

Chapter 16

Testing in Mathematics Education in the Netherlands



Floor Scheltens, Judith Hollenberg-Vos, Ger Limpens and Ruud Stolwijk

Abstract Mathematics testing in the Netherlands focusses on informing schools, teachers, and students about student performance for both formative and summative purposes. The tests are used to monitor whether educational objectives have been achieved and whether content-specific standards have been mastered by the students. In our chapter, we describe the content and objectives of the different national primary and secondary standardised tests. The focus is on the primary function of these tests, but their secondary function where tests are used for accountability is also discussed. In general, the tests are classified into four types: tests to adjust instruction; tests to evaluate proficiency and make decisions about students; tests to evaluate proficiency and make decisions about classes and schools; and tests to evaluate proficiency and make decisions about the quality of the educational system. We show that in practice these types often blend together, as test results are aggregated into class- and school-based indicators at the student level for school evaluation and accountability.

Judith Hollenberg was previously employed at Cito.

F. Scheltens (✉) · G. Limpens · R. Stolwijk
Cito, Arnhem, The Netherlands
e-mail: floor.scheltens@cito.nl

G. Limpens
e-mail: ger.limpens@cito.nl

R. Stolwijk
e-mail: ruud.stolwijk@cito.nl

J. Hollenberg-Vos
Daniel de Brouwerschool, Wilp, The Netherlands
e-mail: j.hollenberg@danieldebrouwerschool.nl

© The Author(s) 2020
M. Van den Heuvel-Panhuizen (ed.), *National Reflections
on the Netherlands Didactics of Mathematics*, ICME-13 Monographs,
https://doi.org/10.1007/978-3-030-33824-4_16

16.1 Introduction

Dutch primary and secondary schools use a variety of tests, each with its own function. In this chapter, we will focus on the most important mathematics tests, describe the primary function of each of these tests, and explain how the tests are used for accountability. In the concluding discussion section, we will identify the difficulties associated with testing mathematics in the Netherlands.

Testing in the Netherlands focusses on monitoring whether or not educational objectives have been achieved and whether the students have mastered content-specific standards. These standards play an important role in allowing Dutch schools to operate relatively autonomously and design their own programs. The standards for different points in a student's educational career are formulated by SLO (2008), the Netherlands Institute for Curriculum Development, under supervision of the Ministry of Education. Schools need to use these standards as guidelines for setting up the content of their educational programme. So, *what* schools have to teach is determined, but schools can choose *how* they work towards the objectives. A sample of mathematics objectives to be achieved at the end of primary education is shown in Fig. 16.1.

In addition to the relatively broad educational objectives that have been in place for quite some time, more detailed content standards have recently been introduced for basic competencies. As the minimum proficiency level of basic skills in the Dutch language and mathematics¹ in secondary education, and particularly in teacher training programs, were considered too low, the Dutch government introduced content standards: the so-called 'Referentieniveaus' (Reference standards) for Dutch language and arithmetic (Expertgroep Doorlopende Leerlijnen voor Taal en Rekenen/wiskunde, 2007). These standards are described for the main transition points in the Dutch educational system: end of primary education, end of secondary education, and end of vocational education. For each transition point, a foundation level (1F, 2F, and 3F) and an ambition level (1S, 2S, and 3S) are specified. All students in a particular school type or track should be able to master the foundation level, while a substantial percentage of students should also be able to master the more challenging ambition level.

Finally, there is a set of standards in the guiding material and test specifications for the construction of national tests and examinations. In so-called syllabi, more detailed descriptions are given of the objectives. These syllabi are specified by the

¹In the Netherlands, a distinction is made between mathematics ('wiskunde') and arithmetic ('rekenen'). In primary school, the term arithmetic is usually used, although this subject also covers other mathematical domains than numbers and operations. In the chapter, we generally use the term 'mathematics', and 'arithmetic' is used to refer specifically to the domain of numbers and operations.

- Students learn to count and do mathematics using estimations
- Students learn to do addition, subtraction, multiplication and division using a flexible strategy
- Students learn to do addition, subtraction, multiplication and division using an algorithm
- Students learn to use a calculator with insight

Fig. 16.1 Sample of objectives of primary mathematics

College voor Toetsen en Examens (CvTE)² for all subjects in secondary education and for mathematics and the Dutch language in primary education. They contain examples of potential examination problems to indicate both the difficulty level and the content of the national examinations and tests. Part of the mathematics syllabus for the pre-university level of secondary education in the domain of algebra is shown in Fig. 16.2.

The objectives, syllabi, and content standards together form the base for testing mathematics in the Netherlands. With this framework in mind, we will now describe the different tests used in primary and secondary education.

Subdomain B1 Algebra

The candidate is able to calculate with numbers and variables using arithmetical and algebraic calculations and understands the use of brackets.

- Readily available knowledge

The candidate knows

- the concepts of absolute and relative

The candidate is able to

- make calculations with and without variables using different arithmetical rules, including power and roots

- Productive abilities

The candidate is able to

- use arithmetic rules to reduce or verify algebraic expressions
- calculate with ratios, percentages, and fractions including one or more variables
- calculate with quantities, composite quantities, and metrics and convert units

Fig. 16.2 Part of the mathematics syllabus for the pre-university level of secondary education for the domain of algebra

²Board of Tests and Examinations.

16.2 Testing Mathematics in the Netherlands

16.2.1 Dutch Education System

Figure 16.3 shows the main elements of the Dutch education system. Primary education includes eight years, starting with two kindergarten years. Children can go to school at the age of four. From the age of five, school is mandatory.

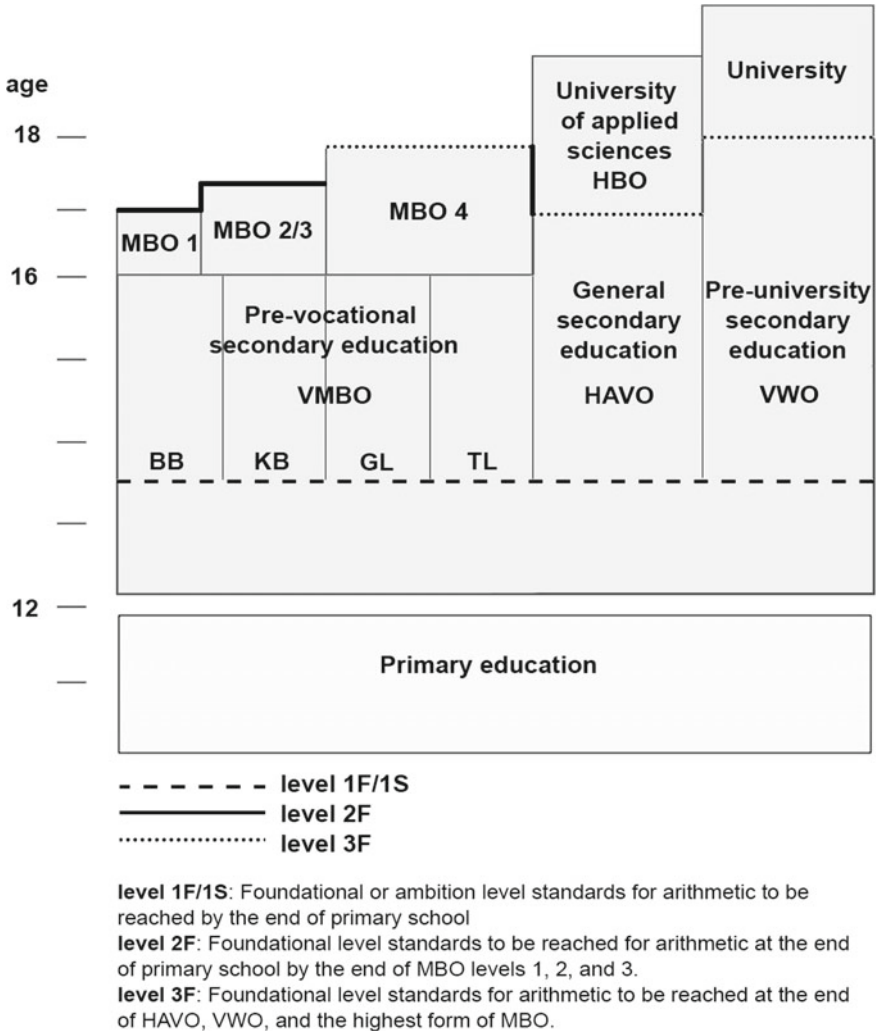


Fig. 16.3 The Dutch school system

Students finish primary education around age twelve and enter secondary education. Secondary education is tracked into three school types:

- VMBO: Pre-vocational secondary education, duration 4 years, subdivided in different levels
- HAVO: General secondary education, duration 5 years
- VWO: Pre-university secondary education, duration 6 years.

Hereafter, students can go to different levels of further education:

- MBO: Intermediate vocational education, duration 1–4 years, subdivided in different levels
- HBO: Higher professional education (also called ‘universities of applied sciences’)
- University.

At the end of each school level students have to reach particular achievement standards for mathematics/arithmetic (Fig. 16.3).

16.2.2 Primary Education

The main objective in primary education, meant for students aged 4–12 is that students (1) gain, gradually and in meaningful contexts, familiarity with numbers, measures, shapes, structures, and their appropriate relationships and calculations; (2) learn to use the language of mathematics; and (3) are able to deal with various sources of content, including daily life, other courses, and pure mathematics (OCW, 2015).³

At the end of primary education, teachers advise students on their secondary education track. To confirm this advice, schools are obliged to administer a test in Dutch language and mathematics. Schools can choose between a number of different tests. When the teacher recommends a lower educational track than that indicated by the test, the teacher’s advice can be reconsidered. The end of primary school test also measures whether students have mastered the foundational (1F) or ambition (1S) level standards for mathematics (and Dutch language).

In addition to the test’s primary function of indicating a secondary education track or verifying the teacher’s recommendation, the aggregated test results for all students can also be used to diagnose areas of improvement for the school (Béguin & Ehren, 2010). For example, they can be used to determine which subjects require more attention and to determine whether measures for improvement have been effective. Most schools use the End Primary School Test, developed by Cito, the Netherlands national institute for educational measurement. The CvTE is mandated by the government of the Netherlands to ensure the quality and proper administration of these national tests and examinations.

³Ministerie van Onderwijs, Cultuur en Wetenschappen (Dutch Ministry of Education).

To monitor the development of primary students in a more formative way, a large number of schools uses a monitoring and evaluation system. One commonly-used system is LOVS⁴ developed by Cito. This system contains tests for different subject domains and sub-domains (e.g., Dutch vocabulary and spelling, and mathematics) for Grade 1–6, with assessments twice a year. There is also a system for pre-schoolers (4- and 5-year-old children) for Dutch language and mathematics. The monitoring system for primary school mathematics is a mixture of mostly open-ended items covering different domains. Each assessment results in an ability score.

Because all mathematics tests in the monitoring system are correlated to each other, teachers can compare test results to those of a previously administered test to monitor student growth. The tests are standardised across the country, enabling teachers to compare individual or class test results and growth with the national average. In addition to indicating a student's overall mathematics ability, the tests also provide information for further analysis. For example, the teacher can analyse whether a student scores very poorly or very high in specific areas. Is the result for the sub-domain Numbers and Operations relatively low and for the sub-domain measurement high, then this could indicate that numbers and operations require additional attention.

The Cito Entrance Test for Grades 4–5 with an assessment once a year is an alternative to the student monitoring system. This test uses a multiple-choice format. It provides a complete overview of the student's skills in mathematics as well as in different sub-domains of Dutch language. In Grade 5, the Cito Entrance Test also provides information to indicate the appropriate secondary education track. All the aforementioned tests are also suitable for students with special educational needs.

The Cito LOVS does not assess mathematical fluency (quickly and correctly solving problems). Therefore, schools use several other tests to monitor this aspect of mathematics.

Along with the national standardised tests from Cito and other test providers, schools use other tests for mathematics such as the tests included in textbooks, various exercises, and (digital) test systems.

Appendix A shows some examples of the type of items which are incorporated in the Cito End Primary School Test and the Cito LOVS tests.

16.2.3 Secondary Education

Mathematics is taught in different ways in the different secondary education tracks. In the first few years of VMBO, the lower tracks of secondary education, the focus is on acquiring insight and skills in the sub-domains of numbers and operations, shapes and figures, quantities and measures, patterns, relations, and functions. Because of the vocational focus of this secondary education track, it is important to provide contexts in which mathematics can be applied: contexts related to everyday life,

⁴Leerling- en Onderwijsvolgsysteem (Student and Education Monitoring System).

other subjects, further education, the workplace, and mathematics itself. In the later years of VMBO, mathematics is only a compulsory subject in the technical sectors; for other students it is an optional subject.

In the higher secondary education tracks, covering HAVO and VWO, of which the highest grades are subdivided into the profiles Nature & Technology, Nature & Health, Economy & Society, and Culture & Society, mathematics is a compulsory subject.⁵ There are different mathematics courses targeted at different profiles:

- Mathematics A, targeted at the Society profiles but also permissible for students in the Nature & Health profile; the focus is more on using mathematical methods and on applications of mathematics.
- Mathematics B, targeted at the Nature profiles and compulsory for students in the Nature & Technology profile; the focus is more on the abstract nature of mathematics.
- Mathematics C, exclusively for students in pre-university education in the Culture & Society profile; the course has some overlap with Mathematics A.
- Mathematics D, a supplementary mathematics course in the specialised or optional component of their profile, for students already taking Mathematics B. Schools are not required to offer a Mathematics D course.

Secondary education ends with a final examination in each subject. For most subjects, the final examination comprises a school examination and a national examination; some subjects, such as physical education, only have a school examination. The school examination is prepared by the individual school and is administered in the final school year or years. Tests can be written, oral, and practical. The national final examination is the same for all schools of a certain type and takes place at the same time in all schools. The student's final mark in a subject is the average of the marks in the school and national examinations. In the Appendices C, D, and E, examples of examination items for the various mathematics courses are shown. These items illustrate the significant differences among the mathematics courses.

The national final examinations in the Netherlands are developed by Cito under the supervision of the CvTE. In the lower secondary educational track (VMBO), there are three national final mathematics examinations, differing in level. These examinations exist in both a paper-based and a computer-based version. In the higher secondary educational tracks (HAVO and VWO), there are, as mentioned before, national final mathematics examinations for Mathematics A, B, and C for each school level. Mathematics D has only a school examination. All the examinations are exclusively paper-based.

For secondary education, there are also monitoring and evaluation systems available for mathematics. An example of such a system is the Cito Monitoring System Secondary Education. This system contains four tests which can be administered over the first three years of secondary education. Students can be evaluated on a vertical equated scale (Béguin & Ehren, 2010). Schools have to monitor student progress in a standardised way, but can choose (or develop) their own system of tests. In addition to

⁵This is not the case for students in HAVO with a Culture & Society profile.

these standardised tests, secondary schools use—similar to primary schools—other tests as well, such as those prepared by the teacher.

16.3 Function of Tests

The previous section of this chapter describes different types of mathematics test. In this section, we will describe the different functions of tests, followed by an outline of functions for the most commonly used tests. Tests can have four different functions: to evaluate and adjust instruction, to evaluate proficiency and make decisions about students, to evaluate proficiency and make decisions about classes and schools, and to evaluate proficiency and make decisions about the quality of the educational system.

16.3.1 Tests to Evaluate and Adjust Instruction

Tests, especially formative tests, ensure that instruction can be adjusted to the students. Tests are designed to provide information not only about the general level of the students but also about student development. Ideally, teachers can use test results to diagnose the specific help or instruction that students need. Examples of tests for evaluating and adjusting instruction are textbook tests and student monitoring systems. The goal of a textbook test is to assess whether students have mastered specific content. When a student answers (almost) all questions correctly, the teacher knows he or she can go on in the textbook. The goal of monitoring systems is to indicate students' current ability levels and growth. These systems contain questions at different levels and in all categories. Teachers can use them to identify specific students who need more instruction or practice and which sub-domains need more attention. In primary education, the student monitoring systems do not aim to classify students. In practice, here the tests are used to identify students who need extra attention or extra challenges. Both in primary and secondary school the tests of the student monitoring systems are also used to choose a secondary education track.

16.3.2 Tests to Evaluate Proficiency and Make Decisions About Students

Tests can also be used to evaluate students' proficiency and make decisions about students. Naturally, these two functions are related. In order to make decisions about a student, the teacher has to figure out whether the student meets the requirements for his or her grade level. This indicates a direction for student's future education.

There are four types of tests for evaluating proficiency and making decisions about students, specifically:

- Tests for selecting students. An example is the examination a student has to pass in order to be admitted to further education, such as succeeding in the national examination for HAVO or VWO, with special requirements regarding the subjects that have been chosen, as a condition for acceptance to higher education.
- Tests for classifying students. Examples are the end of primary school tests. The results of the tests indicate what type of secondary education is best suited for a student.
- Tests for placement. An example is placement in special education. The results of the student monitoring systems are one indicator used to place a student in special education. For special education placements, these results must show that a student's growth is below the growth one might expect for a student at a particular age.
- Tests for certification. The best-known certification test in the Netherlands is the national examination at the end of secondary education.

16.3.3 Tests to Evaluate Proficiency and Make Decisions About Classes and Schools

Tests to evaluate proficiency of students can also be used to evaluate classes and schools. Class growth is central in making decisions about classes. When making these decisions several questions come up. What is the relationship between an increase in ability of a class and the past scores of this class? How is the increase in ability of a class compared to the national increase? But it is also possible to compare the current improvement with previous increases in ability within one school population. How does the improvement of this year's Grade 2 class compare to that of last year's Grade 2? The Cito LOVS incorporates these analyses. Appendix B illustrates and explains a trend analysis at school level.

16.3.4 Tests to Evaluate Proficiency and Make Decisions About the Quality of Education

Schools, school organisations, and also the education inspectorate can evaluate the quality of education. National and international assessments are used to evaluate the quality of education.

An example of a national assessment carried out by Cito is PPON.⁶ This assessment is used to evaluate primary school education in detail every five years.

⁶Periodieke Peiling van het Onderwijsniveau (Periodic Assessment of the Education Level).

Information from this study is used by content experts and decision makers (Béguin & Ehren, 2010). The last PPON for mathematics, carried out in 2011, evaluated 22 different mathematical sub-domains (Scheltens, Vermeulen, & Van Weerden, 2013). In 2014, the responsibility for PPON-like national assessment shifted from the Ministry of Education to the Inspectorate of Education. This change of responsibility will lead to some differences in approach, but the necessity of a national assessment is beyond dispute.

Examples of international assessments are PISA and TIMSS. PISA (the Programme for International Student Assessment) takes place every three years and compares the knowledge and abilities of 15-year-olds in reading, mathematics, and science (Kordes, Bolsinova, Limpens, & Stolwijk, 2013).

The Netherlands also participates in TIMSS (Trends in International Mathematics and Science Study). TIMSS takes place every three years in Grade 4 and 8 and assesses mathematical and science skills. Like in PISA, Dutch students score on average significantly higher than the international average (Meelissen et al., 2012). Table 16.1 summarises the different functions of the most commonly used mathematics tests.

16.4 Use of Tests for Accountability

In the Netherlands, test scores are important for educational accountability. In addition to test evaluations, schools are evaluated by school inspectors who visit the schools. As the Inspectorate of Education is required by law to assess the educational quality that schools offer (including whether the school offers a safe learning environment to students), tests and annual reports are assumed to measure the quality of the school's educational process. The inspectorate uses test scores to identify low-quality schools. Schools that have declining test scores or low test scores over a period of three years are considered to be failing or at risk of failing (Béguin & Ehren, 2010).

Based on the summative or formative function of the tests, it can be assumed that they are valid for measuring the proficiency of an individual student. However, this is not necessarily the case for the aggregated results that are used to indicate educational quality at the school level. Two aspects are important. First, aggregated results as an indicator can misrepresent educational quality if parts of the curriculum are not represented in the tests at hand. For example, the student monitoring tests for primary education do not contain rather open problems in which the student is asked to combine different (mathematical) skills to reach a solution. Nevertheless, a relatively low score in the student monitoring test can still be validly interpreted as a potential lack of quality. Second, one can argue that a test that is a valid measurement of individual students must have different characteristics and content than a test that measures schools (Béguin & Ehren, 2010).

Table 16.1 Outline of functions of the most commonly used mathematics tests

Test	Function						
	Adjusting instruction	Evaluating proficiency and making decisions about students			Evaluating and making decisions on classes and schools		Evaluating and making decisions on quality of education
		Selection	Classification	Placement	Certification		
Testbook test	★						
Student monitoring system mathematics, primary education	★		★			★	
Student monitoring system mathematical fluency, primary education	★					★	
Student monitoring system mathematics, secondary education	★			★		★	
Grade entrance test	★		★			★	★
End of primary school test							
National final examination						★	★
PPON + successors							★
PISA + TIMSS							★

16.4.1 Primary Education

Until recently, the Inspectorate of Education used interim results on the student monitoring system and an end of primary school test as indicators to evaluate the proficiency of primary schools. A new framework for accountability has been available since 2016. This framework focusses on how schools use their test results. The inspectorate no longer sets standards for the interim results of the student monitoring system, but standards are still used for the end of primary school tests (OCW, 2016).

16.4.2 Secondary Education

Since 2016, the Inspectorate of Education has used indicators to judge the quality of a secondary school. First, the inspectorate compares the level of third-year secondary students (Grade 9) to the secondary school track advice that was given at the end of primary school. Next, the inspectorate looks at the percentage of students that pass the first year of secondary school without delay and the percentage of students that pass the last part of secondary school without delay. Finally, the results of the national examination are taken into account. These indicators are compared to a standard established by the inspectorate. The combination of the values achieved on these indicators form a score for the school as a whole. Each of the components contributes to this score and overall it is a balanced system (OCW, 2015). The basic idea of this system of judgement is that schools might do better at one component but worse at another and that this compensates. So if, for example, a school challenges students to achieve a higher level of education than advised, the average scores of these students on the national final examinations can potentially be lower than the scores of students who follow the advised level of secondary school. This will affect the school's indicator for results on the national examinations. Also, it is possible that these students might even need an extra year to finish their secondary education.

16.5 Discussion

16.5.1 Content-Related Issues

16.5.1.1 Testing with or Without Context

Dutch mathematical education has a strong tradition of Realistic Mathematics Education (RME). Mathematics has to be learned in meaningful situations. In the last ten years, a group of experts in mathematics education has advocated for more attention to learning algorithms, teaching fixed procedures for every operation, and teaching mathematics in less meaningful situations. This group has written their

own primary education textbooks. As schools have autonomy, they are free to use an RME-based textbook or a mechanistic algorithm-based textbook (or something between the two). This also has consequences for the tests. Today's assessments contain context problems as well as bare number problems. Nevertheless, schools may vary in the attention they pay to bare number problems and context problems. Therefore, it is possible that there are differences in the extent to which the assessments measure what is actually taught in the school.

Another point about RME is that, in problems that relate to real situations, students are faced with more complex situations in which different mathematical competences have to be combined. In tests, however, different competences are tested in isolation. This is partly because tests have to determine whether there are any gaps in mathematical skills. In order to determine this, it is necessary that each question focusses on one particular competence. This is because in more complex computational problems, the outcome is less clear and analysis is more difficult for teachers, making the results less reliable.

16.5.1.2 Should Mathematics be a Compulsory Subject?

In the Dutch educational system, in the lower grades of secondary education all students at each level must do mathematics, but this does not continue through the end of secondary education. In the pre-university secondary school track (VWO), all students are required to do mathematics. For the other levels, mathematics is not obligatory. So, the system requires that pre-university students know about mathematical relations and be able to do some mathematical thinking at a certain level, but for the majority of secondary students, mathematics is an elective. One could ask oneself what this means for society as a whole: will this lead to a social gap (or an increase in an existing gap) between university-educated citizens and others?

16.5.2 Use of Test Scores

Almost all tests, whether monitoring tests, diagnostic tests, or examinations, provide information about student progress towards content standards. In all these cases, mathematical ability is expressed as a value, for example, an ability score. To ascertain whether a student has obtained a content standard, these standards are connected to an ability score. This is a convenient and effective way to access whether a student has attained a particular standard. A disadvantage of this procedure is that mathematical ability is squeezed into one value. If a student scores strongly in one domain, this may compensate for a weakness in another domain. Therefore, students may pass a certain content standard according to the test without mastering the specific goals of all the reference standards because they exceed the standards in some other domains. A passing score, therefore, should always be considered into the light of a domain analysis. If a student scores equally (well) in all domains, it may be concluded with

reasonable certainty that he or she has mastered the skills described in the reference standards. If the student scores relatively poorly in one or several domains, it is advisable to review the points from the ambition level reference standards in order to establish whether there are gaps to work on with the student.

16.5.3 Use of Tests

16.5.3.1 Autonomy Versus Control

Schools in the Netherlands have the freedom to organise their own teaching programme. As a consequence, they have to account for their choices, for example to the school inspectors. This accountability policy places pressure on schools; they are busy fulfilling all necessary requirements. As a result, autonomy is not what schools experience. By focussing on controlling what schools do, and therefore on collecting test data, there is the risk that the tests partly prescribe the content of the teaching programme. Schools feel that they are judged by the results of the tests, and so they will try to achieve the highest scores. For some schools, this means that the tests determine what they emphasise in their teaching. In these cases, the school does not autonomously decide what they offer their students, but, to put it bluntly, the teaching programme is dictated by the tests.

16.5.3.2 Resistance Against Testing

As mentioned above, schools experience a lot of pressure from testing. Since primary education assessment occurs twice a year for about six subjects, it takes two weeks a year to administer these tests. In addition to that, the use of the student monitoring system is often seen as 'testing for the school inspectors or the school board' rather than a monitoring system for students. It is very counterproductive to use these tests for accountability.

Another type of resistance is against tests for pre-schoolers. As most Dutch children enter primary school at the age of 4 (attending school is required from age 5 on), there is a monitoring system for these young students, too. For mathematics, these tests measure some elementary knowledge of numbers, such as the ability to count, adding small collections of objects, and knowledge of mathematics-related terms such as long(er), short(er), first, and last. Like all monitoring tests, these tests are also ability tests. As a result, the test contains questions that the average student can do well, but it also includes questions designed for lower- and higher-than-average-level students. Since the test includes questions for the more proficient students, some of these assignments demand more than strictly required for the average-level goals. Especially when dealing with young children, this calls for a lot of discussions with teachers. On the one hand, they want to give their students an experience of success, thus, they want their students to pass as many questions as possible. On the other

hand, teachers often talk about the importance of play for children aged 4 or 5. Some teachers find that, in order to cover all the topics that are in the tests, they cannot let the students play as much as they think is necessary for their students' development. The message that students are also allowed to make mistakes in the monitoring tests is a difficult one and has been insufficiently communicated to schools.

16.5.3.3 Teaching-to-the-Test

The main goal of the (monitoring) tests is to monitor the development of students in order to adjust instruction to their potential and needs. The use of test results to assess the quality of schools is of minor importance. In actual practice, however, it seems that the main purpose of the monitoring tests is for external parties to assess school quality. The result of this is that schools, against all advice, adapt their teaching to the tests and have their students practise for the test. The consequence of this is that the expectations of the Inspectorate of Education rise further, since the average assessment test scores increase. The fact that the average assessment score increases does not mean that mathematics proficiency has automatically increased. In this case, the higher assessment scores are the result of more frequent passes of certain parts of the test, not an indicator that teaching methods in mathematics have improved overall. To fairly determine student ability, it is necessary to update the tests frequently. However, this is (very) expensive. In the future, adaptive testing, in which questions in each test are different for different students, may be one solution. Teaching aimed at specific test problems would then be less feasible for schools. Furthermore, schools should be encouraged to keep in mind the real purpose of the tests.

Teaching-to-the-test is a phenomenon that occurs not just with monitoring tests, but also, for example, with end of school tests and examinations. In these cases, however, it is less 'helpful' for schools because these types of test are updated annually. Therefore, teaching to specific assignments is not possible, and, in fact, it is never advisable.

16.5.3.4 Misuse of Tests

An end of primary school test is administered to ascertain whether a student has successfully completed the curriculum in order to decide whether he or she is ready for a certain type of secondary education. Monitoring tests serve a different purpose. They are, as mentioned, primarily meant to steer teaching efficiently towards students' abilities. However, as ministerial policy on secondary school advice changes, there is a danger that monitoring tests could be given a different, more serious function than their original purpose. In 2015, both the time set for administering the end of primary school test and the aims of this test changed. Before 2015, it was meant to be an objective test, indicating a direction for a student along with teacher advice. The test was administered in February, before students had to choose a secondary

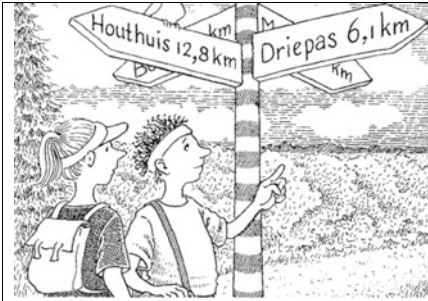
school. A large number of secondary schools required a minimum score on the end of primary school test for admission. This made this test a very important one for students and their parents. To avoid misuse of the test, the government decided to move the time for this test to after students have registered for secondary school. Now, teacher advice is the primary factor in secondary school choice, and students can change their choice only when they get a higher than expected score on the end of primary school test. The result of this is that some secondary schools now require minimum scores on the students' monitoring test progress results—again, a misuse, in this case of the monitoring tests instead of the end of school tests. Schools should use the monitoring tests only as a means of diagnosis and not as a selection tool. A positive development is that the results of the monitoring tests are no longer part of the inspectorate's evaluation framework. As a result, the emphasis in schools moves to the primary goal: namely, identifying students' capabilities and challenges.

16.5.3.5 One Test, Different Functions

On a related point, attention should be given to the use of a test for more than one purpose. Different types of tests each have their own goal and contribute to the quality of Dutch mathematics instruction in their own way. One is aimed at informing teachers and schools about student ability, while another test provides information about the school as a whole, and again other tests aim to determine the national level. As can be seen in Table 16.1, many tests are used for more than one purpose. In order to facilitate accurate assessments, each test should have its own goal(s) and, moreover, that goal (or those goals) should be clear for all parties involved. Only in this way can tests be used for the intended purpose, that is, as a means of improving education. Ultimately, all tests serve this purpose. Whether a test is meant to tune teaching to the needs of students or to determine the quality of teaching, all tests should finally contribute to the best possible education to prepare students for their future as much as possible.

Appendix A

Sample of items from the Cito End Primary School Test and the Cito Student Monitoring System Primary School (LOVS).

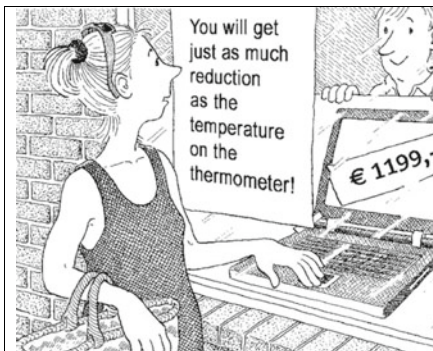


Jean and Peter are walking towards the village of Driepas.

How many metres do they have to walk from this sign to Driepas?

- A 61 metres
- B 610 metres
- C 6100 metres
- D 61000 metres

Cito End Primary School Test (2007)

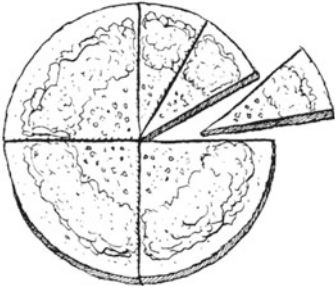


The thermometer indicates 24.9°C. So Wanda will get 24.9% reduction on this comp

How much reduction will Wanda get, roughly speaking?

- A € 50.-
- B € 250.-
- C € 300.-
- D € 400.-

Cito End Primary School Test (2007)




Mother takes a piece of pie.
Which fraction of the pie is this?

A $\frac{1}{12}$ **C** $\frac{1}{5}$

B $\frac{1}{6}$ **D** $\frac{1}{3}$

Cito Student Monitoring System, Grade 4 (2009)



Romy wins first prize. She gives her four best friends €100 each.
How many Euros are left?

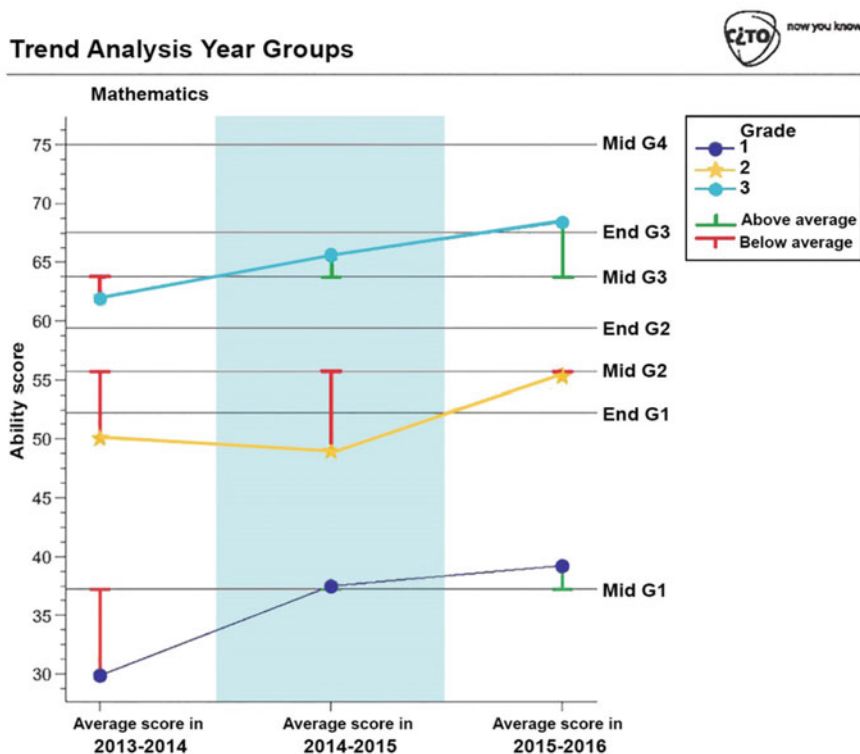
€ _____

Cito Student Monitoring System, Grade 4 (2009)

Appendix B

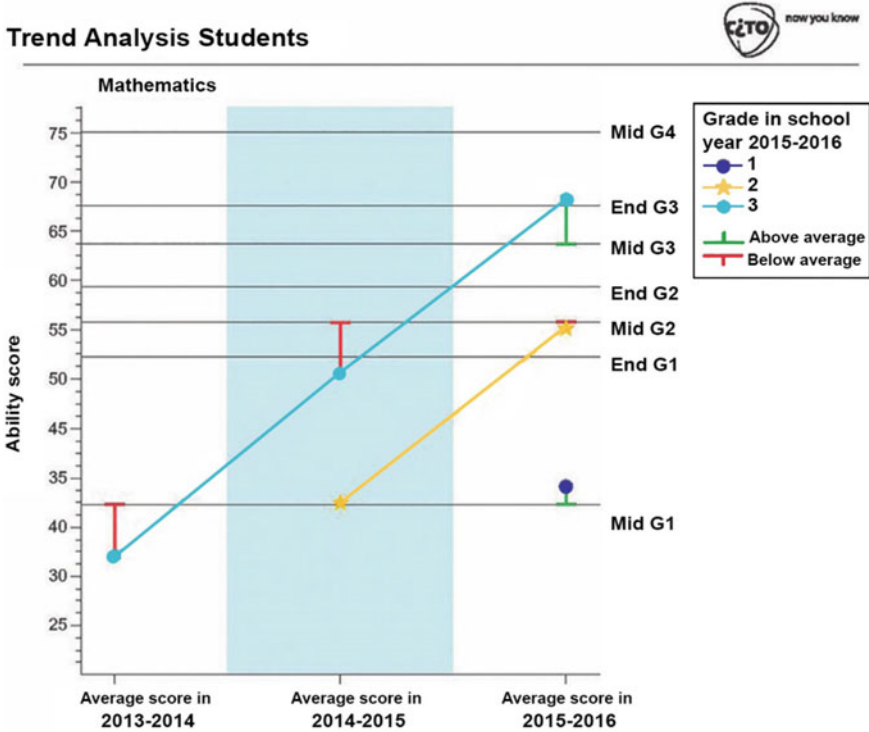
School level analysis based on the scores in the Cito Student Monitoring System Primary School (LOVS).

The Cito LOVS contains a digital module providing a school analysis. This module produces different reports with an overview of the results on school level. Two different types of trend analysis can be made: trend analysis of year groups and of students.



The trend analysis of year groups answers the question: How are the results of this year’s Grade 3 (or 1, 2, etc.) compared to the results of Grade 3 (or 1, 2, etc.) in previous years? In this analysis comparisons are made of different groups of students. Answering this question can be effective for monitoring the effect of a change in teaching approach, like an increase in attention for mathematical fluency. If the results of this year’s Grade 3 are better than those of previous years, the school can confirm that the change has been effective. When the results of Grade 3 were above average this year and below average last year (horizontal lines indicate the average score), the school has an indication that mathematics education needs more

attention. The school can then search for an explanation for this downturn. Has there been a change in student population? Or is a change of teacher a possible (part of the) explanation?



The trend analysis of the students' scores answers the question: How are the results of the students in Grade 3 compared to the results of these students when they were in Grade 2 and in Grade 1? This analysis follows the same group of students. If a group starts in Grade 1 and their scores are above average, one can also expect above average scores in Grades 2 to 4, assuming an average growth. The graph in Trend Analysis Students, shows a group of students scoring below average in Grades 1 and 2 and above average in Grade 3. This shows the growth of this is above average. School can use this information to identify successful factors in their education. What are causes of this growth? Is it due to more attention and time to mathematics in Grade 3? Or is it the effect of a teacher-training programme? Identifying the cause of successes helps schools identify the strengths in their educational approach. The trend analyses identify weak and strong points in education, providing ideas to (further) improve school quality.

Appendix C

Example from national examination Mathematics A (Pre-university Track of Secondary Education, 2014).

The population of Uganda

In 2012, Wali published a study into the population size of the African country of Uganda. According to Wali, this size can be described by a model of the form:

$$U_w = \frac{a}{1 + b \cdot g^t}$$

Here, U_w is the number of inhabitants of Uganda and t is the time in years with $t = 0$ in 1980. Wali used the values $a = 295,267,612$, $b = 22.78367259$ and $g = 0.965$. In the table, you can see that for the years 1980–2010 his model produced values for U_w that matched surprisingly well with the actual values.

Years	Actual population	Calculated population	Years	Actual population	Calculated population
1980	12,414,719	12,414,719	1996	21,248,718	21,266,298
1981	12,725,252	12,845,405	1997	21,861,011	21,980,197
1982	13,078,930	13,290,330	1998	22,502,140	22,716,074
1983	13,470,393	13,749,915	1999	23,227,669	23,474,471
1984	13,919,514	14,224,592	2000	23,955,822	24,255,934
1985	14,391,743	14,714,799	2001	24,690,002	25,061,014
1986	14,910,724	15,220,984	2002	25,469,579	25,890,262
1987	15,520,093	15,743,605	2003	26,321,962	26,744,234
1988	16,176,418	16,283,127	2004	27,233,661	27,623,485
1989	16,832,384	16,840,024	2005	28,199,390	28,528,571
1990	17,455,758	17,414,779	2006	29,206,503	29,460,048
1991	18,082,137	18,007,881	2007	30,262,610	30,418,471
1992	18,729,453	18,619,830	2008	31,367,972	31,404,390
1993	19,424,376	19,251,129	2009	32,369,558	32,418,352
1994	20,127,590	19,902,293	2010	33,398,682	33,460,902
1995	20,689,516	20,573,841			

Some people were impressed by the degree of agreement between the two series of numbers. “Nowhere does the model deviate more than 2% from reality”, one of them said.

Question 1

Using a calculation, demonstrate that this statement is incorrect by giving a year in which the deviation exceeds 2%.

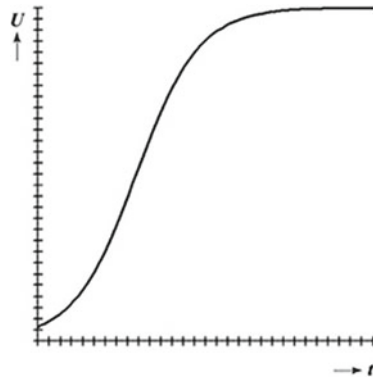
It is not practical when the constants of a model have many digits in front of or after the decimal point. In the sequel of this problem, we therefore work with the following model:

$$U = \frac{300}{1 + 22.8 \cdot 0.965^t}$$

Here, U is the number of inhabitants of Uganda in millions and t is the time in years with $t = 0$ in 1980.

In the figure, you can see that the model predicts a limit value for the population size of Uganda. The horizontal axis runs from 1980 to 2280.

Figure



Question 2

Explain, without substituting numbers into the formula, which limit value goes with this model.

For the derivative of U one has:

$$\frac{dU}{dt} \approx \frac{244 \cdot 0.965^t}{(1 + 22.8 \cdot 0.965^t)^2}$$

Question 3

Demonstrate this.

Question 4

With the aid of the derivative, investigate in which year the population of Uganda increases fastest according to the model.

The population of Uganda – marking scheme

1	maximum score 3	
	<ul style="list-style-type: none"> • A calculation of a percentage greater than 2, for example for 1983: $\frac{13749915-13470393}{13470393} \cdot 100\% \approx 2.1\%$ 	2
	<ul style="list-style-type: none"> • In 1983 the model deviates by more than 2%, hence the statement is incorrect 	1
2	maximum score 3	
	<ul style="list-style-type: none"> • For large t, 0.965^t approaches (arbitrarily close to) 0 	1
	<ul style="list-style-type: none"> • Then the denominator approaches 1 	1
	<ul style="list-style-type: none"> • Then U approaches to 300 million 	1
3	maximum score 4	
	<ul style="list-style-type: none"> • $[0.965^t]' = 0.965^t \cdot \ln(0.965)$ 	1
	<ul style="list-style-type: none"> • $\frac{dU}{dt} = \frac{(1+22.8 \cdot 0.965^t) \cdot 0 - 300 \cdot 22.8 \cdot 0.965^t \cdot \ln(0.965)}{(1+22.8 \cdot 0.965^t)^2}$ 	2
	<ul style="list-style-type: none"> • $\frac{dU}{dt} = \frac{-300 \cdot 22.8 \cdot 0.965^t \cdot \ln(0.965)}{(1+22.8 \cdot 0.965^t)^2} \approx \frac{244 \cdot 0.965^t}{(1+22.8 \cdot 0.965^t)^2}$ 	1
	or	
	<ul style="list-style-type: none"> • $U = 300(1 + 22.8 \cdot 0.965^t)^{-1}$ 	1
	<ul style="list-style-type: none"> • $[0.965^t]' = 0.965^t \cdot \ln(0.965)$ 	1
	<ul style="list-style-type: none"> • $\frac{dU}{dt} = -300(1 + 22.8 \cdot 0.965^t)^{-2} \cdot 22.8 \cdot 0.965^t \cdot \ln(0.965)$ 	1
	<ul style="list-style-type: none"> • $\frac{dU}{dt} = \frac{-300 \cdot 22.8 \cdot 0.965^t \cdot \ln(0.965)}{(1+22.8 \cdot 0.965^t)^2} \approx \frac{244 \cdot 0.965^t}{(1+22.8 \cdot 0.965^t)^2}$ 	1
	4	maximum score 4
<ul style="list-style-type: none"> • The maximum of the derivative needs to be determined 		1
<ul style="list-style-type: none"> • Describing how with the GC can be determined for which t this derivative is maximal 		1
<ul style="list-style-type: none"> • $t \approx 87.8$ 		1
<ul style="list-style-type: none"> • The answer: in 2067 (or 2068) 		1

Appendix D

Example from national examination Mathematics B (Higher Secondary Education, 2014).

Two Functions

The functions f and g are given by $f(x) = (x + 2)\sqrt{x + 2}$ and $g(x) = x(x + 2)$

The graphs of f and g intersect in the points A and B .

1 Determine the x -coordinates of A and B exactly.

Point C lies on the graph of f . The tangent line to the graph of f at C has slope 6.

2 Determine the x -coordinate of C exactly.

Two functions – marking scheme

1	maximum score 4	
	• From $x(x + 2) = (x + 2)\sqrt{x + 2}$ it follows that $x = -2$ or $x = \sqrt{x + 2}$	1
	• $1x = \sqrt{x + 2}$ gives $x^2 = x + 2$ (with $x \geq 0$)	1
	• Describing how the equation $x^2 = x + 2$ (with $x \geq 0$) can be solved exactly	1
	• (The x -coordinates of A and B are) $x = -2$ and $x = 2$	1
Remark: If $x = -1$ is named as a solution of the equation, award a score of no more than 3 points.		
2	maximum score 5	
	• $f(x) = (x + 2)^{1\frac{1}{2}}$	1
	• $f'(x) = 1\frac{1}{2}(x + 2)^{\frac{1}{2}}$ (or a comparable form)	1
	• For the x -coordinate of C one has $1\frac{1}{2}(x + 2)^{\frac{1}{2}} = 6$	1
	• From this it follows that $(x + 2)^{\frac{1}{2}} = 4$ (i.e., $\sqrt{x + 2} = 4$)	1
• This gives $x + 2 = 16$ hence $x = 14$	1	

Example from national examination Mathematics B (VWO National Examination, 2014).

Fractional Trigonometric Function

For every a with $a \neq 0$ the function f_a is given by: $f_a(x) = \frac{\sin(ax)}{1 - 2 \cos(ax)}$

1 Determine for which values of a the line with equation $x = \pi$ is a vertical asymptote of the graph of f_a

2 Prove that the graph of f_2 is symmetric about the point $(\frac{1}{2}\pi, 0)$

Fractional trigonometric function—marking scheme

1	maximum score 4	
	• One must have: $1 - 2 \cos(a\pi) = 0$, hence $\cos(a\pi) = \frac{1}{2}$	1
	• This gives $a\pi = \frac{1}{3}\pi + k \cdot 2\pi$ or $a\pi = -\frac{1}{3}\pi + k \cdot 2\pi$ (for integer k)	1
	• Hence $a = \frac{1}{3} + k \cdot 2$ or $a = -\frac{1}{3} + k \cdot 2$ (for integer k)	1
	• For these values of a one has $\sin(a\pi) \neq 0$ (hence, for these values of a the line with equation $x = \pi$ is a vertical asymptote of the graph of f_a)	1
Remark: If only the solutions $\frac{1}{3}$ and $-\frac{1}{3}$ are found, award a maximum of 2 score points for this question.		
2	maximum score 5	
	• One needs to prove that $f_2(\frac{1}{2}\pi - p) = -f_2(\frac{1}{2}\pi + p)$ (for every p)	2

(continued)

(continued)

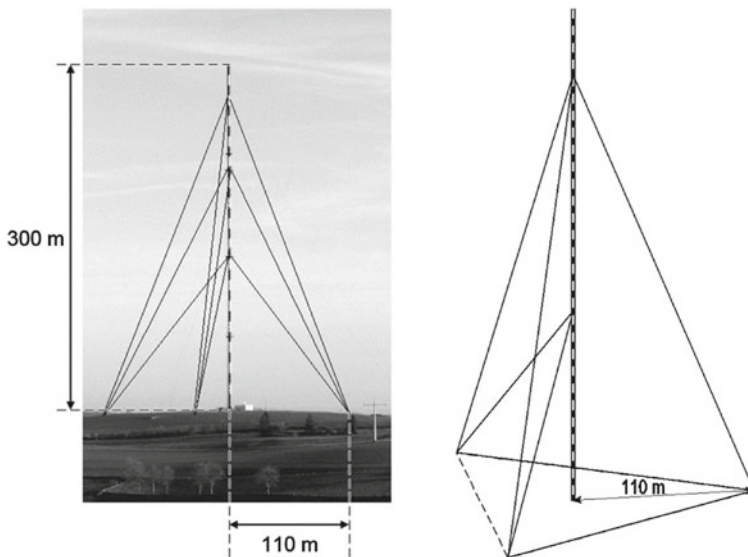
• $f_2(\frac{1}{2}\pi - p) = \frac{\sin(\pi-2p)}{1-2\cos(\pi-2p)}$ and $f_2(\frac{1}{2}\pi + p) = \frac{\sin(\pi+2p)}{1-2\cos(\pi+2p)}$	1
• $(\sin(\pi - 2p) = \sin(2p)$ and $\sin(\pi + 2p) = -\sin(2p)$, hence) $\sin(\pi - 2p) = -\sin(\pi + 2p)$	1
• $(\cos(\pi - 2p) = -\cos(2p)$ and $\cos(\pi + 2p) = -\cos(2p)$, hence) $\cos(\pi - 2p) = \cos(\pi + 2p)$ (hence $f_2(\frac{1}{2}\pi - p) = -f_2(\frac{1}{2}\pi + p)$ for every p)	1
Remark: If a value is substituted for p , do not award any score points for this question.	

Appendix E

Example from national examination mathematics (Lower Secondary Education, 2014).

Radio Mast

The radio mast of Radio Luxembourg is located in Hosingen, Luxembourg. The radio mast is 300 m tall.



The radio mast is held up from three sides by three guy-wires from each side. See the picture.

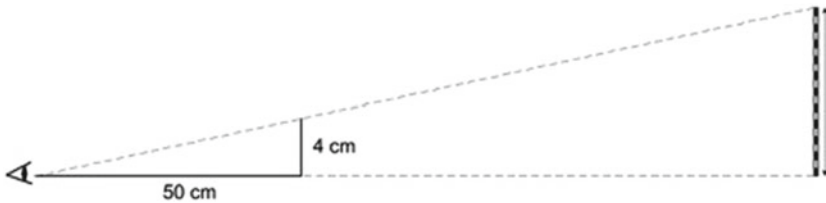
The three guy-wires from one side are all anchored to the ground at one and the same point. The distance from the foot of the radio mast to this point is 110 m. In the

drawing on the right you see the radio mast together with the highest and the lowest guy-wires.

- 1 The highest guy-wire is attached to the radio mast at a height of 270 metres.
→ Determine how many metres the length of the highest guy-wire is. Write down your calculation. Round your answer to the nearest whole number.
- 2 The lowest guy-wire is attached to the radio mast at a height of 120 metres.
→ Determine how many degrees the angle between the lowest guy-wire and the ground is. Write down your calculation.
- 3 The guy-wires are anchored to the ground at equal distances from the radio mast. On the worksheet a map of the three anchor points is shown.
→ Indicate the place of the radio mast with a dot on the map. Show how you have obtained your answer.
- 4 The radio mast can be seen from miles around. Agatha wants to know the distance to the radio mast. She stretches her arm and indicates the size of the radio mast with her fingers.
→ Determine the distance from her eye to her fingers is 50 cm. The height she indicates with her fingers is 4 cm. You see a sketch of the situation.



The distance from her eye to her fingers is 50 cm. The height she indicates with her fingers is 4 cm. You see a sketch of the situation.



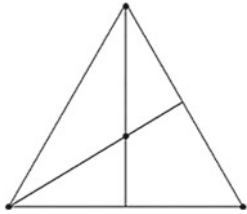
→ Determine the distance from Agatha to the radio mast in whole metres. Write down your calculation.

Radio mast—marking scheme

1	maximum score 3	
	• The length of the highest wire is equal to $\sqrt{270^2 + 110^2} = 291.5 \dots$ (m)	2
	• This is equal to 292 (m) when rounded to the nearest whole number	1
2	maximum score 3	
	• $\tan \text{ angle} = \frac{120}{110}$	2
	• The answer: 47° (or more accurately)	1
3	maximum score 3	
	• For example, two medians drawn correctly	2

(continued)

(continued)

	<ul style="list-style-type: none"> The place of the radio mast has been indicated in the correct position. 	1						
4	<p>maximum score 2</p> <ul style="list-style-type: none"> A ratio table such as <table border="1" data-bbox="404 522 780 583"> <tr> <td>horizontal distance</td> <td>50</td> <td>...</td> </tr> <tr> <td>height</td> <td>4</td> <td>300</td> </tr> </table>	horizontal distance	50	...	height	4	300	1
horizontal distance	50	...						
height	4	300						
	<ul style="list-style-type: none"> The distance is equal to 3750 (m). 	1						

References

- Béguin, A., & Ehren, M. (2010). Aspects of accountability and assessment in the Netherlands. *Zeitschrift für Erziehungswissenschaft*, 13, 25–36.
- Expertgroep Doorlopende leerlijnen voor taal en rekenen/wiskunde. (2007). *Eerste deeladvies* [First partial advice]. Enschede, the Netherlands: SLO.
- Kordes, J., Bolsinova, M., Limpens, G., & Stolwijk, R. (2013). *Resultaten PISA-2012 in vogelvlucht. Praktische kennis en vaardigheden van 15-jarigen. Nederlandse uitkomsten van het Programme for International Student Assessment (PISA) op het gebied van wiskunde, natuurwetenschappen en leesvaardigheid in het jaar 2012* [Results PISA-2012 in a nutshell. Practical knowledge and skills of 15-year-olds. Dutch outcomes of the Program for International Student Assessment (PISA) in the field of mathematics, natural sciences and reading skills in the year 2012]. Arnhem, the Netherlands: Cito.
- Meelissen, M. R. M., Netten, A., Drent, M., Punter, R. A., Droop, M., & Verhoeven, L. (2012). *PIRLS-en TIMSS-2011: Trends in leerprestaties in lezen, rekenen en natuuronderwijs* [PIRLS and TIMSS-2011: Trends in learning achievements in reading, mathematics and science education]. Nijmegen/Enschede, the Netherlands: Radboud Universiteit/Universiteit Twente.
- OCW. (2015). *Onderwijsresultatenmodel* [Education results model]. Den Haag, the Netherlands: OCW.
- OCW. (2016). *Onderzoekskader 2017 voor het toezicht op voorschoolse educatie en het primair onderwijs* [Research framework 2017 for the supervision of pre-school education and primary education]. Den Haag, the Netherlands: OCW.
- Scheltens, F., Hemker, B., & Vermeulen, J. (2013). *Balans van het reken-wiskundeonderwijs aan het einde van de basisschool 5* [Balance of mathematics education at the end of primary school 5]. Arnhem, the Netherlands: Cito.
- SLO. (2008). *TULE—Rekenen-wiskunde. Inhouden en activiteiten bij de kerndoelen* [TULE—mathematics. Contents and activities]. Enschede, the Netherlands: SLO.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

