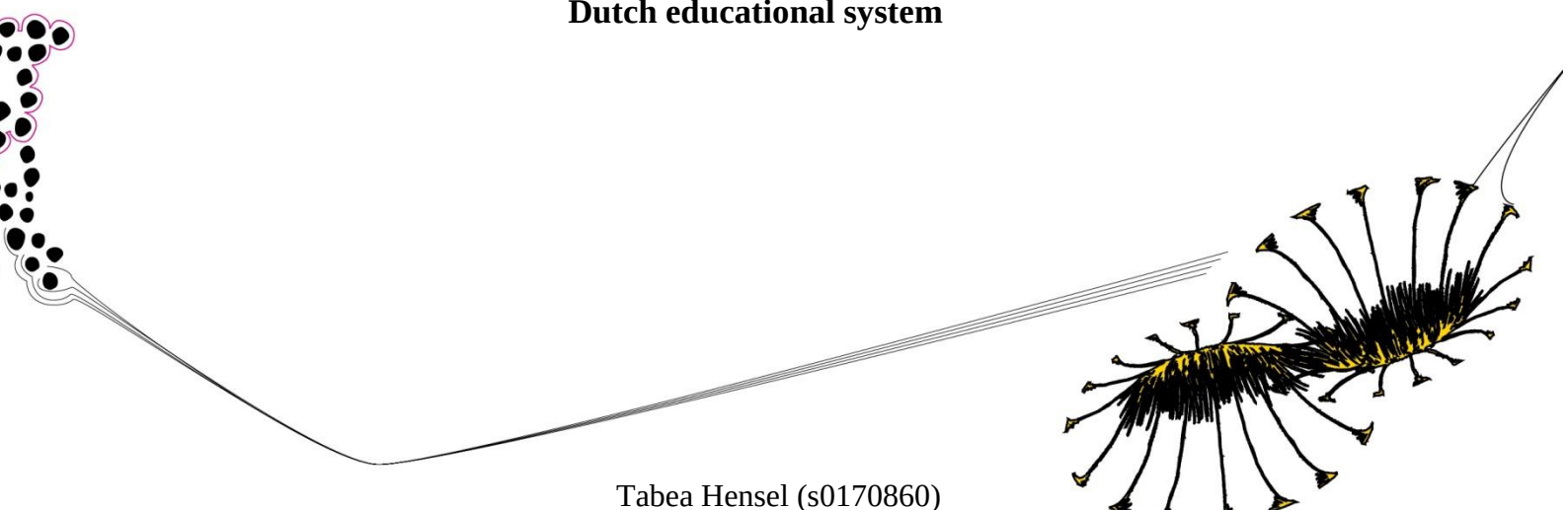


**Validation of the Flesch-Kincaid Grade Level within the
Dutch educational system**



Tabea Hensel (s0170860)

Faculty of Behavioural Sciences

University of Twente, the Netherlands

First supervisor: prof. dr. Jan Maarten Schraagen

Second supervisor: dr. Teun Lucassen

External supervisor: Wim Muskee (Kennisnet)

January 31, 2014

Abstract

Although the Flesch-Kincaid Grade Level is frequently used by researchers in order to estimate the readability of written material, it is barely validated. Therefore, this study investigated whether the Flesch-Kincaid Grade Level is valid for Dutch teaching material. Experiments, in which texts from three types of schools (vmbo, havo and vwo) were used, were carried out with havo students. Results suggest that the formula is not an accurate predictor for readability, but can help to select teaching material more easily by classifying the text as easy or difficult to understand. According to this study, a Flesch-Kincaid Grade Level score below 8 can be used as a cut-off point for easy texts for havo students. Furthermore, it can be proposed that by applying the Flesch-Kincaid Grade Level, text selection can be enhanced for a target group below the age of 15.

Contents

Introduction	5
A brief introduction to readability formulas.....	6
Reading comprehension	9
Weaknesses and shortcomings of readability formulas	11
Justification of this research	12
Method	16
Participants	16
Design.....	17
Variables.....	17
Independent variable	17
Dependent variables	19
<i>Score of comprehension questions</i>	19
<i>Likert scale estimation..</i>	19
Materials.....	19
Justification of text choice.....	20
Calculating the readability scores	21
Task	21
Procedure.....	21
Data analysis	22
Results	23
Flesch-Kincaid Grade Level scores of teaching material.....	23
Binary logistic regression.....	26
Discussion & conclusion	28
Limitations & future research	31
Recommendations for Kennisnet	32
References	33
APPENDIX A	37
APPENDIX B	42
Texts for the easy condition	42
<i>First grade</i>	42
<i>Second grade</i>	43
<i>Third grade</i>	47
<i>Fourth grade</i>	49

APPENDIX C	52
Texts for the difficult condition	52
<i>First grade</i>	52
<i>Second grade</i>	54
<i>Third grade</i>	57
<i>Fourth grade</i>	61

Introduction

A good readability of a text should always be of concern to every author. It is important that people understand written material such as insurances, contracts, package inserts or instruction manuals. However, several studies have identified shortcomings of the readability of written text (Bruce, Rubin & Starr, 1981; Masson & Waldron, 1994; Wegner & Girasek, 2003; Lucassen, Dijkstra & Schraagen, 2012). A study by Wegner and Girasek (2003) revealed that insufficient readability can even end with fatal consequences such as the death of a number of children in America in 1998. The wrong installation of child-safety seats in cars was the cause for the deadly accidents. Wegner and Girasek (2003) assumed that the underlying cause for the wrong installation was the poor comprehension of the installation instruction, which consequently led to the infants' deaths. Another example for poor readability is the mismatch of patient-education materials and the readability skills of patients (Meade & Smith, 1991). Patient-education materials often seem to be too complex and complicated for the intended audience (Pichert & Elam, 1984), which can also result in undesirable outcomes such as intoxication or death. For this reason, health professionals have argued that readability needs to be enhanced for these types of materials.

In order to detect which texts were too difficult, they suggested applying readability formulas. Eventually, the reading material should match the patients' level of education and readability skills, in order for patients to understand the matter they are confronted with (Leichter, Nieman, Moore, Collins & Rhodes, 1981). A non-lethal, but still disconcerting example of poor readability is the finding of Lucassen, Dijkstra and Schraagen (2012) that approximately half of the English Wikipedia articles are too difficult to comprehend for a large amount of people. This condition must certainly be perceived as undesirable for an open encyclopedia that aims to reach a broad variety of users.

In the following sections, an introduction to readability and readability formulas is given. The focus lies on the Flesch-Kincaid Grade Level readability formula, used for the Dutch language and for educational purposes. Furthermore, ways to measure reading comprehension are presented. Thereafter, some weaknesses and shortcomings of readability formulas are discussed. This finally leads to the purpose of this study, namely the validation of the Flesch-Kincaid Grade Level for Dutch teaching material. Subsequently, methodology and data analysis are given in detail. At the end, the results and limitations of this study are discussed.

A brief introduction to readability formulas

First of all, in order to understand what readability formulas indicate and their purpose entails, readability needs to be defined. Readability, not to be confused with legibility, can be defined as a characteristic of a written text, which implies an adequate understanding of the message that was intended by the writer. Legibility, in contrast, relates to the layout and typeface of written text (DuBay, 2004). According to Heydari (2012), the following definition by Dale & Chall (1949) of readability is most comprehensive: “The sum total (including all the interactions) of all those elements within a given piece of printed material that affect the success a group of readers has with it. The success is the extent to which they understand it, read it at an optimal speed and find it interesting.” (Dale & Chall, 1949, p. 23).

Readability formulas are mathematical formulas that are composed of different variables concerning amongst others word and sentence length and constant weights (Van Oosten, Tanghe & Hoste, 2010). These formulas are supposed to predict the required reading ability one should have in order to understand the content of the text (Redish, 2000). Using readability formulas offers several advantages. Firstly, it is truly easy to apply them to the respective piece of text by automated computer programs. Secondly, texts can be chosen according to their level of difficulty without reading the text beforehand. By this means, the process of selecting suitable texts for the target audience will be simplified.

In the 1920s the first attempts were made to predict the readability of written text. This was done by estimating vocabulary difficulty and counting sentence length (DuBay, 2007). These properties seemed to be a suitable measure for the estimation of reading difficulty, since these properties are still often used in the current readability formulas in order to predict readability. During the last century, about 200 readability formulas were invented. Among them are the Dale-Chall Readability Formula (Dale & Chall, 1948), the Fry Graph (Fry, 1969), the SMOG formula (McLaughlin, 1969), and readability tests developed by Rudolf Flesch, such as the Flesch Reading Ease Test (Flesch, 1948) and the Flesch-Kincaid Grade Level (Kincaid, Fishbure, Rogers & Chissom, 1975). All of these formulas were initially developed for the English language, but are also used for other languages with certain adjustments, as for example the Flesch-Douma formula for the Dutch language. Although this formula is adjusted in order to give a better indication of comprehension difficulty for the Dutch language, the Flesch-Douma formula has not been validated yet (Kraf, Lentz & Pander Maat, 2011).

Amongst all formulas, the Flesch Reading Ease and the Flesch-Kincaid Grade Level have gained most popularity. Microsoft even integrated the Flesch Reading Ease formula in

MS Word to check a document's readability (Marnell, 2008). Both formulas by Flesch focus on the same characteristics of the text. These are the average sentence length (calculated by dividing the total of words by the total of sentences) and average word length (calculated by dividing the total of syllables by the total of words). The Flesch-Kincaid Readability formula is a recalculation of the Flesch Reading Ease formula (Kincaid, Aagard, O'Hara & Cottrell, 1981). What distinguishes the two formulas is the readability score that is calculated. Whereas the score of the Flesch-Reading-Ease Formula lies between 1 and 100, defining 100 as very easy to understand and 0 as very difficult, the score of the Flesch-Kincaid indicates a grade level, or to be more precisely the years of education one should have in order to comprehend the text. It is assumed that increasing length of both, words and sentences, makes it harder to read and to understand a text.

The formulas of the Flesch Reading Ease and Flesch-Kincaid Grade Level formula are as follows:

Flesch Reading Ease:

$$206.835 - 1.015 (\text{total words/total sentences}) - 84.6 (\text{total syllables/total words})$$

Flesch-Kincaid Grade Level:

$$0.39 (\text{total words/total sentences}) + 11.8 (\text{total syllables/total words}) - 15.59$$

The score of the Flesch-Kincaid Grade Level reflects the grade level of the American educational system, which indicates the number of years of education the reader has had since age 6. For example, if the score of a text is 8, the text should be comprehensible for an average American student that has had 8 years of education. Hence, an adolescent between the ages of 13 and 14.

In the table below (table 1), an interpretation of the Flesch Reading Ease scores is given (Du Bay, 2004). This table was published in *The Art of Readable Writing* by Flesch in 1949. Besides the Reading Ease score and an interpretation, an estimated reading grade was added.

Table 1

Interpretation of the Reading Ease scores with estimated reading grade by Flesch (1949)

Reading Ease Score	Interpretation	Estimated Reading Grade
0 to 30	Very difficult	College graduate
30 to 40	Difficult	13 th to 16 th grade
50 to 60	Fairly difficult	10 th to 12 th grade
60 to 70	Standard	8 th and 9 th grade

70 to 80	Fairly easy	7 th grade
80 to 90	Easy	6 th grade
90 to 100	Very easy	5 th grade

In contrast to most countries in Europe, in the United States, there is no division in educational levels according to the skills of the child. In the Netherlands, there are several types of schools for children with different abilities. It lies outside the scope of this study to explain the entire Dutch educational system in detail, which is why a limited description of the three main types of school is given.

After primary school, about the age of 12, Dutch students attend secondary school. Broadly speaking, there are three types of secondary schools. These are firstly the preparatory middle-level vocational education (voorbereidend middelbaar beroepsonderwijs, in short vmbo), higher general continued education (hoger algemeen voortgezet onderwijs, in short havo) and thirdly the preparatory scientific education (voorbereidend wetenschappelijk onderwijs, in short vwo). Education at vmbo lasts four years. Students who have attended havo usually finish school by the age of 17, after five years of education since primary school. Vwo has six grades. Usually, students finish this type of school by the age of 18 (Figure 1).

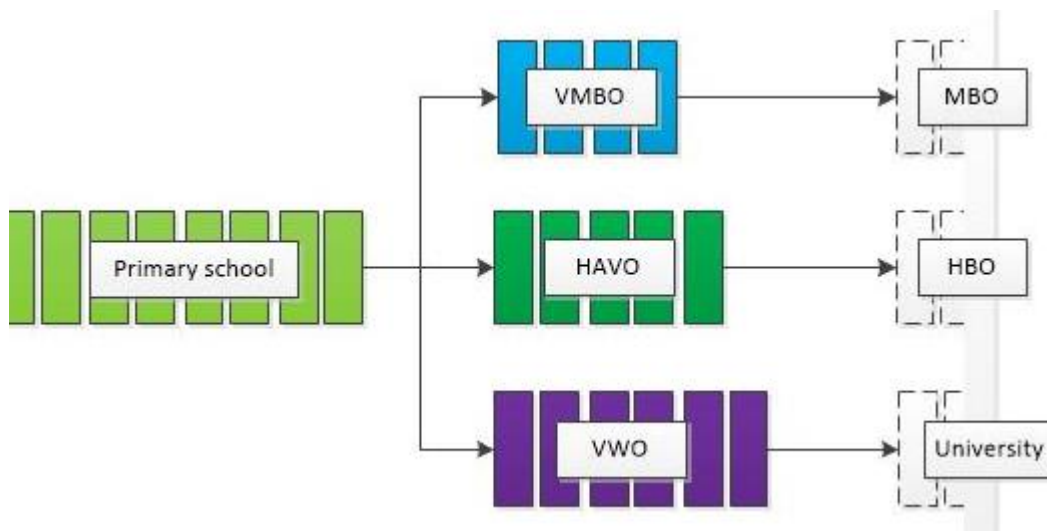


Figure 1: An illustration of the Dutch educational system

Within these three types of schools, the level of difficulty rises from vmbo (the lowest level) to vwo (the highest level), with havo being positioned between the two types of school.

Besides the increasing difficulty level, vmbo emphasizes practical skills, whereas havo and vwo provide more theoretical knowledge.

With regard to the Flesch-Kincaid Grade Level, a translation from grade level to age does not appear applicable according to the Dutch educational system: Since students of different types of schools follow an educational system complementing different skills, reading comprehension skills are assumed not to be comparably well developed. For example, a text with a Flesch-Kincaid Grade Level score of 8 would presumably not be equally readable for a second grader of vmbo, compared to a second grader of vwo. Grades in America and the Netherlands are numbered differently in secondary school. In the table below (table 2), the grade levels according to the American and Dutch educational system and the respective age is presented.

Table 2

American and Dutch grade levels and respective age

American Grade Level	Dutch Grade Level	Age
7	1	12-13
8	2	13-14
9	3	14-15
10	4	15-16
11	5	16-17
12	6	17-18

Reading comprehension

Readability formulas give an indication of the difficulty of the text. In other words, the extent to which the reader understands what he or she is reading. However, what is reading comprehension and which factors affect reading comprehension? Snow (2002) defines reading comprehension as “the process of simultaneously extracting and constructing meaning through interaction and involvement with written language.” (p.11). Heydari (2012) classifies factors that influence reading comprehension in two categories, namely reader variables and text variables. Reader variables refer to factors that are internal to the reader, such as the background knowledge, motivation, skills or abilities. Text variables, on the other hand, are concerned with the textual factors as for instance genre, typographical features, text content and readability.

Comprehension of written text by itself is very complex. DuBay (2007) gives two reasons for this. First, there are many definitions for comprehension and second, testing reading comprehension is challenging. McLaughlin (1974) defines four ways of measuring reading comprehension. The first one is the *activity* test, which is basically only suitable for instructional material such as instruction manuals for, for instance constructing a cupboard. In an activity test, participants are asked to carry out a task according to instructions they received. Controlling comprehension is simple, for example when a cupboard is built up correctly, it might consequently be concluded that the instructions were clearly written and well understood.

A second type of comprehension measurement is a *replacement* test, also called Cloze (readability) procedure (Bormuth, 1967). In this procedure, every fifth word in a passage is deleted and in lieu thereof, a blank space is inserted. It is the participants' task to fill in the blank space with words that were deleted beforehand and they assume would fit in the specific context. McLaughlin (1974) states that, although comprehension is measured, it is a different kind of comprehension that is intended to be tested by the Cloze procedure since the comprehension of a mutilated text is tested.

Paraphrasing a passage of text is the third kind of measuring comprehension according to McLaughlin (1974). Participants are asked to summarize a text passage in order to test to which degree the reader understood the written material. However, this method comes along with several weaknesses such as the scoring of the paraphrases and the fact that summing up a passage does not necessarily mean that the subject truly comprehends what he or she read.

The fourth type of measuring comprehension is a *question* test. While doing this, questions about a text passage are asked. McLaughlin raises many questions expressing his doubts with regard to this method. His concerns were the following: "To what extent are the answers implicit in the questions? Can a subject answer them anyhow, without having understood the text? Are the questions more difficult than the text? To what extent do the questions test reasoning and memory, rather than comprehension?" (McLaughlin, 1974, p. 369).

The question test can be designed in three ways. Open-ended questions can be asked which means that the subject usually has to write down the correct answer in a blank field. It is up to the participant to use his or her own words or to literally copy the answer from the text. Closed-ended questions often appear in the form of multiple-choice options. Several answers are presented and the subject has to choose the correct answer. One weakness of this

form of testing is that in case the answer is incorrect, the question rises whether this is due to a lack of comprehension or a lack of memory (DuBay, 2007). Another form of closed-ended questions is posing true/false questions. In this case, statements are given that need to be marked as either true or false.

Weaknesses and shortcomings of readability formulas

Although readability formulas ought to give an indication of the difficulty of the text, there are some shortcomings of the formulas. Many studies have discussed several pitfalls of readability formulas in general (Bruce, Rubin & Starr 1981; Oakland & Lane, 2004; Marnell, 2008). Bruce, Rubin and Starr (1981) mention three weaknesses of these formulas. First, what is known about reading and its process seems to be ignored by these formulas. The authors emphasize that it needs more than sentence length and word difficulty to predict the difficulty of a text. What also plays an important role in determining text difficulty is for example the rhetorical structure, complexity of ideas and background knowledge. Reader-specific factors such as interest in the subject or motivation are also neglected and not integrated in the readability formulas. Second, the formulas lack a statistical basis and validation. Even though validity studies have been carried out for many readability formulas, the focus lay on earlier formulas which were validated by using the McCall-Crabbs Standard Test Lessons in Reading. However, these lessons were initially developed for the purpose of practicing exercises in reading, not to determine comprehension. Third, the inappropriate use of these formulas is criticized. Although readability formulas were initially developed to determine which books are suitable for children, there is no evidence that a readability formula can correctly predict the interaction of a reader with a certain book. Besides determining the readability of a certain piece of text, readability formulas also serve the purpose of modifying texts. A general (mis)conception regarding the use of readability formulas is that shorter words and sentences are easier to understand. However, this can often lead to increased difficulty of the text rather than enhancing reading ease. In addition to the weaknesses that Bruce, Rubin and Starr (1981) sum up, Marnell (2008) addresses further shortcomings of the Flesch Reading Ease formula. In his criticism, he refers to textual attributes more in detail. Due to calculating a readability score by the number of words and syllables, a change of the word order of a grammatically correct sentence into a nonsense sentence still has the same readability score but might not make sense anymore. Another point Marnell stresses is that the Flesch readability formula does not take the change of meaning by punctuation into account.

Using a full stop rather than a question mark can change the meaning of a text as well as using a hyphen. The same is true for typographical cueing, which helps to accentuate certain parts of the text in order to make the text passage more comprehensible. Using bold text or italic script can heavily influence the meaning of written text. However, the Flesch Reading Ease fails to detect that too. Marnell (2008) concludes that it takes more to define or cause readability than sentence length or syllable count. Rather than focusing on surface features of written text, the actual reader and the content of what is read should be considered more carefully when measuring readability.

Although there are many studies that criticize the utility and validity of readability formulas, they are still in common use to estimate readability (Rameezdeen & Rajapakse, 2007; Rajagoplanan, Khanna, Leiter, Stott, Showalter, Dicker & Lawrence, 2011; Lucassen, Dijkstra & Schraagen, 2012). The Flesch Reading Ease score is also deep-rooted in the state law of Florida. The reading ease score of insurances in Florida is demanded to be at least 45 in order to make sure that a large part of the population can understand the documents (Florida Laws: FL Statutes - Title XXXVII Insurance Section 627.4145). What needs to be kept in mind is that this requirement is based on a readability formula that was invented in the 1950's and has not known many validation studies since then. Therefore, it is time to validate the Flesch-Kincaid Grade Level formula. Since it is nearly impossible to validate the Flesch-Kincaid Grade Level for all kinds of populations, academic background, ages and kinds of text, in order to make a first step, this study will focus on havo students in the Netherlands.

Justification of this research

Readability formulas were used as means to estimate readability in many contexts as was described in the previous sections. The educational sector, however has not received such research attention so far. As with every other written document, sufficient readability is also of great importance for teaching material. Since the Flesch-Kincaid Grade Level can easily be applied on written text by computed programs, it could be very useful for the selection of teaching material. One Dutch organization that is engaged with educational affairs and would benefit from the addition of readability scores (if the formula is valid) is Kennisnet. Kennisnet is a public educational organization that is specialized in the interaction between Dutch primary and secondary school and vocational institutions and ICT. For Kennisnet, the question *where* ICT can be deployed most effectively is crucial. There are several educational services and applications that were developed in collaboration with the partners of Kennisnet.

Some of their services are merely offered to teachers as can be seen with Wikiwijs, an online platform in which teachers can share and search digital course material with their colleagues. Other offers for teachers are services that support the professionalization of the teacher. Leraar24 is an online platform, created for educators in order to support them in their daily functioning as a teacher. On this platform, teachers get the opportunity to share experiences with their colleagues. Besides the services for teachers, Kennisnet also offers special services for students. One of them is Davindi, a safe search engine for students of primary schools. This program helps students to search for the information they need for presentations and work pieces safely. Kennisnet also created applications for tablets. The Kenny App is especially developed for toddlers. When using this app, toddlers can become acquainted with skills such as writing and logical and spatial thinking. Furthermore, two supportive services are provided by Kennisnet. Kennisnet entitled Edurep and Kennisnet Federatie. Kennisnet Federatie is an association of educational institutions and suppliers of teaching material.

Edurep is a service that is offered by the foundation Kennisnet and is focused on educational material. It is a digital collection of metadata of materials that are used for teaching purposes. The material is digitally made available and labeled with metadata by for example publishers or educational institutions. Examples of metadata are the features of teaching material such as author, title, level of education of key terms. Institutions and publishers who possess collections of teaching material can choose to join Edurep and hence make the metadata of their material accessible for other parties, for example for teachers who are searching for new material.

Edurep aims to collect as much educational metadata as possible to enhance the process of searching for teaching material. Besides the name of the author or the level of education, metadata such as the readability score of the material are desirable. By means of this score, finding the material that is convenient regarding learning level will be facilitated.

Although the Flesch formulas enjoy popularity and are broadly used in order to determine how readable and understandable written text is, the formulas' validation studies were often carried out decades ago. The second point of criticism of Bruce et al. (1981), the lack of validation in combination with the aim of expanding the metadata of Edurep by readability scores, led to this study. Since the Dutch educational system knows more diversity than the separation due to age, a grade level has little explanatory power when it comes to a readability score. Nevertheless, to begin with, the Flesch-Kincaid Grade Level readability formula is aimed to be validated on havo students. In a study on the use of readability formulas within patient-education material, Pichert and Elam (1984) point out that it is

important to use a readability formula that is validated on the same population as the one that is intended to read the written material. Hence, it is important to validate the Flesch-Kincaid Grade Level on a student population.

Therefore the main research question of this study is as follows:

RQ1: Is the Flesch-Kincaid Grade Level a valid predictor for text comprehension of Dutch teaching material and havo students?

In order to answer this question, other research questions need to be answered. The assumption was made that teaching material differs between types of schools and grades. To check whether this is also reflected in the Flesch-Kincaid Grade Level scores, the scores were compared with each other. This led to the following research questions and hypotheses:

RQ2: Can differences in Flesch-Kincaid Grade Level scores be established in texts from vmbo, havo and vwo?

H1: Flesch-Kincaid Grade Level scores rise with type of school (from vmbo to vwo), because vwo texts are expected to be more difficult than vmbo and havo texts.

H2: Flesch-Kincaid Grade Level scores for vmbo texts are lower than for havo texts.

H3: Flesch-Kincaid Grade Level scores for havo texts are lower than for vwo texts.

RQ3: Can differences in Flesch-Kincaid Grade Level readability scores be established in texts from all four grades?

H4: With each grade, the Flesch-Kincaid Grade Level readability scores will increase.

The other research questions focus on the actual performance of the experiment and its results.

RQ4: Are the differences in types of schools reflected in the comprehension scores?

H5: Havo students perform better on vmbo texts than on havo and vwo texts.

H6: Havo students perform worse on vwo texts than on havo and vmbo texts.

RQ5: Is there a difference in comprehension scores between texts classified as easy and difficult according to the Flesch-Kincaid Grade Level score?

- H7:** Less incorrect answers will be given in the easy condition, compared to the difficult condition.
- RQ6:** What is the likelihood of giving an incorrect answer to the comprehension questions for each grade?

The validation of the Flesch-Kincaid Grade Level in this study took place in several stages. First of all, the assumption was made that teaching material differs in difficulty not only concerning the age of students, but also within types of schools since different types of schools (vmbo, havo and vwo) exist. Therefore, it was chosen to use authentic teaching material that is currently used at schools in the Netherlands. Teaching material is specifically created for students in different grades and different schools. Schoolbooks for the subject Dutch for four grades of vmbo, havo and vwo were collected. All sections “reading” (“lezen”) were digitized and subsequently, the Flesch-Kincaid Grade Level was applied. By doing this, it was checked whether this assumed difference in teaching material difficulty across types of schools was also reflected in the scores of the formula.

Secondly, after these texts were marked with their respective readability scores, the texts with the lowest (easy condition) and the highest scores (difficult condition) of each grade and of each type of school were chosen as materials for the experiment. Two forms (parallelklassen) of each havo grade were used as experimental group (two 1st grades, two 2nd grades and so on). Each participant received three texts, each text representing a different type of school (vmbo, havo and vwo). The texts only differed according to the type of school, not in grades. Each 1st grader received one 1st grade vmbo text, one 1st grade havo text and one 1st grade vwo text and so on. There were two groups in this experiment. One group received texts from its own grade with the lowest readability score of all collected texts. The other group received texts with the highest readability score. Additionally, both groups received questions about each text that were intended to assess comprehension of the texts. Alongside each text, at least four and at the most seven questions were posed. Keeping McLaughlin’s criticism on the question test in mind, a pretest was carried out. Six college students read all texts, answered the respective questions and gave feedback on the texts. Questions that could be answered by general knowledge or common sense were excluded. Also, questions that implicitly contained the answer were ruled out. To rule out a memory effect, the participants were allowed to use the texts and questions simultaneously.

Method

Participants

A total of 211 high school students at havo level took part in this experiment. They were all recruited at a single school in Hengelo, Overijssel, the Netherlands. Havo students were chosen for this study because they resembled the “average” student more than vmbo or vwo students would. All students participated voluntarily as a part of their Dutch lessons and without any reward. One participant was excluded from the analysis due to the missing demographical information and missing Likert scale rating. The average age was 14.15 ($SD = 1.381$). 120 of the high school students were male (57.1%) and 90 were female (42.9%). 92.9% was Dutch. 7.1% had another nationality. It was assumed that all participants were proficient in reading Dutch since for most participants Dutch is their mother tongue. In the tables below (table 3 and 4), descriptive statistics of the participants of all tested grades are presented. In the following, the grades that received the texts with the lowest readability scores are referred to as *easy condition*. Grades that received texts with the highest readability scores are referred to as *difficult condition*.

Table 3

Descriptive statistics of participants from easy condition

	1st grade	2nd grade	3rd grade	4th grade
	$N = 30$	$N = 26$	$N = 24$	$N = 24$
Age	$M = 12.63$	$M = 13.77$	$M = 14.79$	$M = 15.83$
	($SD = 0.490$)	($SD = 0.587$)	($SD = 0.588$)	($SD = 0.917$)
Gender	Male = 60%	Male = 65.4%	Male = 70.8 %	Male = 66.7%
	Female = 40 %	Female = 34.6%	Female = 29.2 %	Female = 33.3%

Table 4

Descriptive statistics of participants from difficult condition

	1st grade	2nd grade	3rd grade	4th grade
	$N = 29$	$N = 29$	$N = 27$	$N = 21$

Age	$M = 12.76$ ($SD = 0.511$)	$M = 13.48$ ($SD = 0.574$)	$M = 14.67$ ($SD = 0.480$)	$M = 16.29$ ($SD = 0.717$)
Gender	Male = 69% Female = 31%	Male = 31% Female = 69%	Male = 48.1% Female = 51.9%	Male = 47.6% Female = 52.4%

Design

The design of this study was a between-subjects design with two conditions. In both conditions, the participants received three texts that were all used for the respective grade they were in but differed in type of school (vmbo, havo, vwo). All first texts in the experiments were vmbo texts, all second texts stem from havo textbooks and all third texts are taken from vwo books. The difference between the conditions was the readability score of each text. Whereas one group received texts with low readability scores (easy condition), the other group received texts with high readability scores (difficult condition).

Variables

Independent variable

Flesch-Kincaid Grade Level score. The independent variable was the readability score of the Flesch-Kincaid Grade Level. The score was calculated automatically by website that was created for this study after the respective text had been pasted into the input module of the website. In order to choose which texts were to be used for the actual experiment, the readability formula was applied to all texts from the reading sections texts of each grade (1-4) and type of school (vmbo, havo and vwo). The mean scores of all reading passages are presented in the table below (table 5).

Table 5

Means of Flesch-Kincaid Grade Level scores of all reading passages

Grade	Vmbo	Havo	Vwo
1st grade	$N = 22$ $M = 8.330$	$N = 21$ $M = 8.288$	$N = 19$ $M = 8.644$

2nd grade	<i>N</i> = 26 <i>M</i> = 8.572	<i>N</i> = 28 <i>M</i> = 9.210	<i>N</i> = 20 <i>M</i> = 9.796
3rd grade	<i>N</i> = 22 <i>M</i> = 8.861	<i>N</i> = 18 <i>M</i> = 10.239	<i>N</i> = 21 <i>M</i> = 10.046
4th grade	<i>N</i> = 22 <i>M</i> = 9.848	<i>N</i> = 14 <i>M</i> = 9.921	<i>N</i> = 14 <i>M</i> = 10.069

The texts with the lowest and highest readability score within each grade and level were chosen, resulting in 24 texts. In the table (table 6) below, the readability scores of the texts that were used for the experiments are presented. The smallest difference between the readability scores of the two conditions was 4.639 for texts that stem from the vwo 1 textbook. The biggest difference in readability scores could be detected for vmbo 3 texts. A difference of 10.358 grade levels was found between the lowest and highest readability score.

Table 6

Flesch-Kincaid Grade Level scores of both conditions for four grades of vmbo, havo and vwo

Grade	Easy condition	Difficult condition
Vmbo 1	5.097	11.866
Vmbo 2	4.358	12.956
Vmbo 3	4.219	14.577
Vmbo 4	5.881	12.832
Havo 1	4.523	12.200
Havo 2	5.621	13.722
Havo 3	7.398	13.217
Havo 4	7.122	14.247
Vwo 1	6.801	11.44
Vwo 2	7.213	12.861
Vwo 3	5.353	12.572
Vwo 4	7.709	13.413

Dependent variables

Score of comprehension questions. Participants had to answer between four to seven open questions on each of the three texts they received. Although McLaughlin (1974) has expressed his concern about the question test, this measurement of comprehension seemed most suitable for this study. Because multiple choice questions offer the chance of guessing the right answer, in this study, open-ended questions were used in order to test comprehension. Another rationale for this choice was that the high school students are already familiar with this way of questioning since this resembles the questions that are used in the text books and exams.

After the questions were answered, the answers were coded. A point system with zero points for an incorrect answer and one point for a correct answer was used. There were 111 questions for 24 texts (see Appendix A). Out of these 111 questions, 29 were adopted from the reading sections of the schoolbooks. The rest of the questions was generated by the researcher. For the data analysis, nine questions had to be split up in two questions because questions were too ambiguous or contained two answers.

Likert scale estimation. For a subjective estimation of text difficulty, the participants were also asked to rate how difficult they thought each of the three texts was on a 7-point Likert scale (1 = *very easy* and 7 = *very difficult*).

Materials

The texts that were presented to the participants stem from books that are currently in use at schools in Hengelo, Overijssel, the Netherlands. They all come from the same publisher (Noordhoff & Wolters) but differ in editions. A total of twelve books of the subject Dutch was collected for four grades and three different types of schools. Since the books stem from the same publisher, the structure of all books was basically the same (except for havo 5 and vwo 4, which had a slightly different structure). All books were divided in eleven sections. These were:

- Reading (Lezen)
- Speaking, watching, listening (Spreken/kijken/luisteren)
- Writing (Schrijven)
- Task (Taak)
- Study skills (Studievaardigheid)
- Language and vocabulary (Taal en woordenschat)

- Grammar (Grammatica)
- Spelling (Spelling)
- Fiction (Fictie)
- Examination (Test)
- Project (Project)

Since the reading section was also provided with comprehension and grammatical questions and deliberately selected for reading purposes, this section was chosen for the experiments. All reading sections of the twelve books were scanned by an optical character recognition (OCR) scanner. A total of 247 texts was extracted. Subsequently, the texts were copied to Microsoft's MS Word and adjusted, if necessary, in order to apply the Flesch-Kincaid Grade Level formula to them. Besides adjusting misspelling, the captions needed to be provided with a full stop. This was done in order to avoid that the number of words in a caption is added to the following sentence, which results in an incorrect readability score. For an example, see the text below.

Ster in de stad

Bombaytour voor 6-13 jaar

In Tropenmuseum Junior kan iedereen een ster worden! Ga mee op ontdekkingsreis naar Bombay, de grootste stad van India. Duik in het leven van deze miljoenenstad en maak kennis met z'n inwoners. Ervaar hoe ze hun dromen proberen waar te maken. Het meisje Gauri bijvoorbeeld: met maar één been, danstalent en doorzettingsvermogen, straalt ze als een ster in eigen stad. Ook jij kan een poging wagen. Begin onderop, grijp je kans en wie weet dans je wel mee in een Indiase videoclip. Niet zomaar een videoclip. Het is Thoemka Laga, gezongen door de beroemde Bollywoodzangeres Asha Bhosle. Na afloop kun je de clip bestellen waarin je zelf meedanst.

Van de straat tot superster. Kom naar Tropenmuseum Junior en maak het in Bombay. Ook leuk voor een verjaardagspartijtje.

Kijk voor tips op www.sterindestead.nl.

Without adding full stops the Flesch Kincaid score would be 7.62. However, when adding full stops, the readability score of this text is 7.12.

Justification of text choice

Out of the 247 texts, 24 were chosen for the experiment. In order to validate the readability formula, the texts with the highest and lowest readability score of each grade and type of school were selected for the experiments (see Appendix B). Only one exception was

made for the lowest vwo text for the 2nd grade: Rather than the text with the lowest score, the text with second lowest score was selected because an illustration was necessary in order to fully understand the former.

Calculating the readability scores

The calculation of the readability scores was not carried out manually, but by a computer program, Python 2.7. Additionally, python-hyphenator 0.5.1, Natural Language Toolkit 2.0_beta9 and myspell-nl-2.10g needed to be added to the program. To enhance the usability of calculating the readability score, a website was built where the respective texts could be pasted. This website can be visited on <http://bam.student.utwente.nl/readability.php>. Besides the readability scores (both Flesch-Kincaid Grade Level and Flesch Reading Ease), character count, syllable count, word count and sentence count was given.

Task

Participants were instructed to read three texts and answer open questions concerning these texts. Besides answering the questions, participants were also asked to indicate how difficult the texts were to them on a 7-point Likert scale.

Procedure

In the beginning of each session, all participants were told that the present study was carried out in order to investigate which teaching material was suitable for them. They were not told about readability formulas and the actual purpose of this experiment. All participants received two booklets. In the first booklet, there were three texts they had to read. In the second booklet, a short introduction and open questions for each of the three texts with blank space for the answers were presented. Additionally, participants were also asked to rate the perceived difficulty of the texts on a 7-point Likert scale. After the booklets were distributed, the experimenter repeated the instructions on the procedure verbally and gave time to ask questions. Since this experiment was not set up to test memory, participants were allowed to look for the correct answers by going back and forth in the text. During the experiment, participants were allowed to use both booklets simultaneously. Also, no time restriction for

answering the questions was given. Sessions lasted at most 45 minutes (one teaching lesson). Most sessions, however, lasted 30 minutes on average.

Data analysis

For the data analysis IBM SPSS Statistics 21 was used. ANOVA's were carried out to test whether Flesch-Kincaid Grade Level scores differed between grades and types of schools. For the comparison of comprehension scores between the three types of schools, a Kruskal-Wallis test was conducted. In order to test whether differences in comprehension scores were present, Mann-Whitney tests were carried out. To estimate the likelihood of giving an incorrect answer to the comprehension questions a binary logistic regression was performed.

Results

Flesch-Kincaid Grade Level scores of teaching material

The second and third research question referred to the Flesch-Kincaid Grade Level readability scores of all texts that were collected for this study. A main effect was found for type of school with $F(2, 246) = 3.040$, $p = 0.050$, indicating that there are differences between the means. The first hypothesis was that the Flesch-Kincaid Grade Level scores rise from vmbo to vwo. This was confirmed by a Bonferroni post-hoc test (95% CI [-1.463, -0.006], $p = 0.047$), which showed that there is a significant difference between the mean of vmbo texts ($M = 8.888$, $SD = 2.117$) and the mean of vwo texts ($M = 9.623$, $SD = 1.630$). It was hypothesized that readability scores of texts from vmbo are lower than from havo. Although there is a difference in means between vmbo texts ($M = 8.888$, $SD = 2.117$) and havo texts ($M = 9.325$, $SD = 1.975$), this difference was not significant as was shown by a Bonferroni post-hoc test (95% CI [-1.147, 0.274], $p = 0.420$). The third hypothesis was that Flesch-Kincaid Grade Level scores from havo texts are lower than from vwo texts. A Bonferroni post-hoc test showed that the difference between the mean of havo texts ($M = 9.325$, $SD = 1.975$) and the mean of vwo texts ($M = 9.623$, $SD = 1.630$) were not significant either (95% CI [-1.048, 0.452], $p = 1.000$). In the table below, sample sizes, means and standard deviations of the readability scores of texts for the three types of schools are presented (table 7).

Table 7

Sample sizes, means and standard deviations of the Flesch-Kincaid Grade Level scores for all three types of schools

	<i>N</i>	<i>M</i>	<i>SD</i>
Vmbo	92	8.888	2.117
Havo	81	9.325	1.975
Vwo	74	9.623	1.630

An ANOVA was carried out to detect differences between grades, without making a distinction between different types of schools. It was hypothesized that Flesch-Kincaid Grade Level scores will increase with an increase in grades. For the four grades, a main effect was found ($F(3, 246) = 7.428$, $p = 0.000$). In the table below, sample sizes, means and standard deviations of the readability scores of texts for each grade is presented (table 8). Post-hoc tests revealed where the differences stemmed from. Readability scores in the first grade ($M =$

8.412, $SD = 1.652$) were significantly lower than in the third grade ($M = 9.679$, $SD = 2.080$) as was shown by a Bonferonni post-hoc test (95% CI [-2.168, -0.365], $p = 0.001$). The mean of the scores of the first grade was also significantly lower than from the fourth grade ($M = 9.930$, $SD = 1.921$), revealed by a Bonferroni post-hoc test (95% CI [-2.468, -0.568], $p = 0.000$). This confirms the fourth hypothesis.

Table 8

Sample sizes, means and standard deviations of the Flesch-Kincaid Grade Level scores for each grade

	<i>N</i>	<i>M</i>	<i>SD</i>
1st grade	62	8.412	1.652
2nd grade	74	9.144	1.854
3rd grade	61	9.679	2.080
4th grade	50	9.930	1.921

Difference in answers on comprehension questions between vmbo, havo and vwo texts

The fourth research question referred to the difference between the number of incorrect answers between vmbo, havo and vwo texts. Hypothesis five was that participants will perform better on the comprehension questions for vmbo texts than on havo and vwo texts. The sixth hypothesis was that the participants will perform worse on comprehension questions regarding vwo texts than on havo and vmbo texts. A Kruskal-Wallis test was conducted in order to find out whether there are differences in the number of incorrect answers between the three types of schools. The number of incorrect answers differed significantly between the three types of schools ($H(2) = 60.483$, $p = 0.000$). In order to find out where the differences stem from, Mann-Whitney tests were carried out. To avoid the inflation of the Type I error, the critical value of 0.05 for significance was divided by the number of tests that were carried out. Two tests were carried out in order to answer the fourth research question, which is why a critical value of 0.025 was set.

Comparing the number of incorrect answers of vmbo and havo with each other, revealed a significant difference ($U = 472678.500$, $z = -5.307$, $p = 0.000$) with mean rank for havo being lower (Mean Rank = 975.19) than for vmbo (Mean Rank = 1064.97), which means that more incorrect answers were given to comprehension questions on havo texts. The same was done to detect differences between responses of havo and vwo. Also here, a significant difference was detected ($U = 585739.000$, $z = -2.487$, $p = 0.013$). Fewer incorrect

answers were given to questions that referred to havo texts. Hence, both hypotheses are confirmed. In the figure below (figure 2), the percentage of incorrect and correct answers for all texts of vmbo, havo and vwo are illustrated.

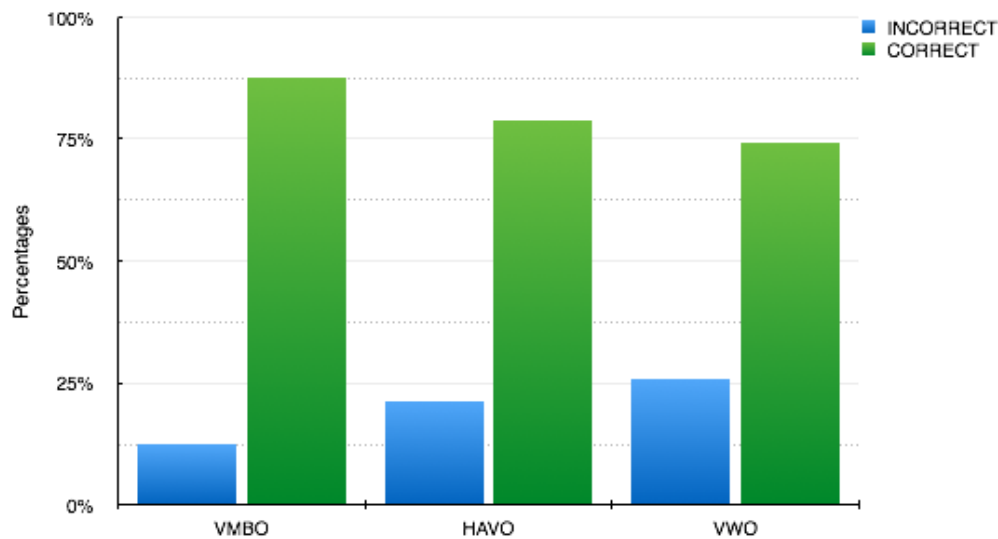


Figure 2: Percentages of correct and incorrect answers to text comprehension questions for each type of school

Difference of given answers between easy and difficult condition

The fourth research question referred to the difference of the given answers between the easy and difficult condition. It was hypothesized that fewer questions will be answered wrong in the easy condition, in contrast to the difficult condition. A Mann-Whitney test was carried out for each grade individually. In the first grade, there was a difference in the proportion of correct and incorrect answers ($U = 85864.00$, $z = -6.635$, $p = 0.000$). In the easy condition, 11.6% of all given answers were incorrect. In the difficult condition, 29.3% of the answers were incorrect. Also a difference between given answers was detected in the third grade ($U = 57474.00$, $z = -8.595$, $p = 0.000$). Of all given answers, 5.2% were incorrect in the easy condition, whereas 31.3% of the answers in the difficult condition were answered incorrectly. In the second grade, the number of incorrect answers was higher in the easy condition (27.7%), compared to the incorrect answers in the difficult condition (27.1%). No significant difference was detected for the second grade ($U = 78678.500$, $z = -0.196$, $p = 0.845$). In the fourth grade, 9.8% of all answers were incorrect in the easy condition. In the difficult condition, 6.5% of the given answers were incorrect. This difference was not significant either ($U = 54600.00$, $z = -1.547$, $p = 0.122$).

Binary logistic regression

In order to answer the sixth research question binary logistic regression was carried out. The Flesch-Kincaid Grade Level score was used as covariate. Different from linear regression where a prediction of a score on some outcome measure is made, logistic regression expresses the likelihood of an event to fall in either of the two categories of the outcome variable (Maroof, 2012). In the SPSS output, this is expressed as odds ratio and can be found in the $\text{Exp}(B)$ column. If the odds ratio is larger than 1, that means that with a one-unit increase of the independent variable, the chance of a certain outcome increases as well. If the odds ratio is smaller than 1, the chance of a certain event to occur will decrease. In case the odds ratio equals 1, the independent variable has no effect on the prediction and can be ignored.

Difference between grades

A binary logistic regression was carried out for the four grades separately in order to take the age, or rather the grade of the participants into account.

In the first grade (12-13 years), the Flesch-Kincaid Grade Level can be significantly associated with giving an incorrect answer (Wald $\chi^2(1) = 37.810$, $p = 0.000$, $\text{Exp}(B) = 1.197$). With every one unit increase of the Flesch-Kincaid Grade Level score, the odds of giving an incorrect answer are 1.197 (19.7%) times greater than the odds of giving a correct answer.

In the third grade (14-15 years), giving an incorrect answer is 1.267(26.7%) times more likely with every one unit increase in the Flesch-Kincaid Grade Level score (Wald $\chi^2(1) = 61.294$, $p = 0.000$, $\text{Exp}(B) = 1.267$).

For the second grade, the Flesch-Kincaid Grade Level score was not a significant predictor for giving an incorrect answer. For the fourth grade (15-16 years), the Flesch-Kincaid Grade Level is a marginally significant predictor for giving an incorrect answer (Wald $\chi^2(1) = 61.294$, $p = 0.056$, $\text{Exp}(B) = 0.920$). Here, the $\text{Exp}(B)$ value is lower than one, which means that the likelihood of giving an incorrect answer decreases by the factor 0.92 with every one-unit increase in the Flesch-Kincaid Grade Level.

Other findings

Likert scale estimation

For each grade separately, a t-test was performed to check for differences in means of the Likert scores between the two conditions. For the first grade, a significant difference was

found. The mean of the easy condition significantly differed from the mean of the difficult condition ($t(881) = -7.762, p = 0.000$). Likert scores were lower for the easy texts ($M = 2.345, SD = 1.285$), compared to the difficult texts ($M = 3.087, SD = 1.541$). A significant difference between the means of the two conditions was also found for the second grade ($t(782) = -2.611, p = 0.009$). The mean Likert score for the easy condition was $M = 2.45 (SD = 1.431)$, whereas the mean of the difficult condition was higher ($M = 2.721, SD = 1.435$). For the third grade, means of the two conditions also differed significantly from each other ($t(790) = -17.378, p = 0.000$). The mean of the easy condition was significantly lower ($M = 1.556, SD = 0.763$), compared to the difficult condition ($M = 3.00, SD = 1.298$). Also in the fourth grade, a significant difference between the means of the Likert scores of the two conditions was found ($t(628) = -5.409, p = 0.000$). The mean Likert score was higher in the difficult condition ($M = 2.643, SD = 1.525$) than in the easy condition ($M = 2.048, SD = 1.188$).

Discussion & conclusion

The present study examined the validity of the Flesch-Kincaid Grade Level readability formula. It was tested whether the Flesch-Kincaid Grade Level is a valid predictor for text difficulty of Dutch teaching material. In order to answer this main research question, other research questions need to be answered first. The second and third research question referred to the teaching material that was collected for the execution of the experiment of this study. It was tested whether the assumption that teaching material differs in difficulty between the different types of schools and grades was also reflected in the Flesch-Kincaid Grade Level scores. A rise in Flesch-Kincaid Grade Level scores was detected in the teaching material from vmbo to vwo. In other words, texts from vmbo are easier to understand than texts from havo and vwo according to the readability score. Also a rise in difficulty of the teaching material from the first to the fourth grade was reflected in the Flesch-Kincaid Grade Level scores. Flesch-Kincaid Grade Level scores of the collected teaching materials from the first grade were significantly lower than from the material from the third and fourth grade. According to the Flesch-Kincaid Grade Level scores, the texts with a lower score are expected to be easier to understand than texts with higher scores. Although the Flesch-Kincaid Grade Level score did increase with increasing grade level, the Flesch-Kincaid Grade Level score did not fully comply with the Dutch grade levels. For the first grade and second grade, the readability score was one grade higher than the actual grade the students are in. The mean readability score of the texts from the third class, in fact, reflects the equivalent American grade. In the fourth class, the mean Flesch-Kincaid Grade Level score is slightly lower than the Dutch grade level. However, the differences between the actual grade level and the Flesch-Kincaid Grade Level Grade Level score are quite small. Pikulski (2002) pointed out that it is quite common that readability scores are erroneously thought of being a precise score. Keeping that in mind, the slight difference between actual grade level and Flesch-Kincaid Grade Level can be regarded as practically negligible. When it comes to the estimation of readability, the score of the Flesch-Kincaid Grade Level should be handled as a guideline rather than be regarded as an infallible tool for accurate prediction (Pikulski, 2002).

The fourth research question referred to whether the assumed difference in difficulty of texts between types of schools is also reflected in the comprehension scores derived from the experiments. When comparing the percentages of wrong and correct answers, fewer questions about the vmbo texts were answered incorrectly in comparison with the questions referring to havo and vwo texts. Also, fewer incorrect answers were given for havo texts,

compared to the answers on vwo texts. This is in line with the assumptions that were made according to the types of schools and the rising difficulty from vmbo to vwo.

The fifth research question related to the difference of comprehension scores between the easy and difficult condition. It was hypothesized that fewer answers would be incorrect in the easy condition. This was true for the first and third grade. Significantly more wrong answers were given in the difficult condition. However, this was not the case for the second and fourth grade. In fact, in the fourth grade strikingly few incorrect answers were given compared to the other grades. The percentage of wrong answers was even slightly lower in the difficult condition. One explanation for this might be that from the age of 15-16 students' reading comprehension skills are developed better and comprehension becomes a more stable skill. Support for this assumption is provided by the literacy researcher Jeanne Chall. Chall (1983) defined six stages of reading development, which start with stage 0 when the child is approximately six months old. The last stage is stage 5 and takes place from age 18. Stages three and four are of interest for this study since they cover the age of participants that took place in the experiments. At stage three, which ranges from 9 to 14 years of age, children's listening comprehension is more developed than reading comprehension. However, during the fourth stage, from age 15 to 17, a shift of comprehension skills takes place. Reading comprehension has improved a lot at that stage and is now better developed than listening comprehension. Adolescents are now able to better understand materials of difficult content and readability. Due to the increase of reading comprehension skills, it might be that the readability score of a piece of text becomes less important from the age of 15.

The sixth research question dealt with the likelihood of giving an incorrect answer to the comprehension questions. In the first and third grade, results indicated that the more difficult the text is according to the Flesch-Kincaid Grade Level score, the more likely it is to give an incorrect answer. In other words, when the text is more difficult, comprehension is worse. This is consistent with the underlying principle of the Flesch-Kincaid Grade Level readability formula indicating increasing text difficulty with increasing Flesch-Kincaid Grade Level scores. Keeping this principle in mind, the outcome of the fourth grade might be against one's expectation. A change in likelihood can be detected in the fourth grade, where students were *less* likely to give an incorrect answer with increasing Flesch-Kincaid Grade Level score rather than more likely as it was detected for the first and third grade. As was already addressed when answering the fifth research question, this might be due to the fact that reading comprehension becomes more stable. Furthermore, other factors such as familiarity of the topic, motivation or text length might be influencing comprehension less than in earlier

stages of reading development. Another explanation might be that the comprehension questions were too easy for the fourth grade. Most questions were created by the researcher and have not been validated, so it might also be that the questions for the fourth grade were a lot easier than for the other three grades.

It is noticeable that no difference in percentages of incorrect answers between the two conditions was detected in the second grade. Percentages of incorrect answers were nearly the same in both conditions. A plausible explanation for this might be that the posed questions for this grade were too difficult or formulated imprecisely. Since the performance on the comprehension tests of two different forms were compared with each other, it might also be that students in the easy condition were less skilled in reading comprehension than were the students in the difficult condition. Even if this was the case, it is noticeable that the Likert scores for both conditions were quite low, which implies that both of the texts were estimated as being easy to understand.

After discussing the results of this study, it cannot be concluded that the Flesch-Kincaid Grade Level is a valid predictor for Dutch teaching material. McLaughlin (1974) put it in a nutshell by saying that the only thing a prediction formula should do is to predict correctly. Although results from this study suggest that the Flesch-Kincaid Grade Level cannot be used as an accurate prediction tool for text comprehension for the Dutch language, it does not mean that the Flesch-Kincaid Grade Level is useless. This is for several reasons. First of all, the formula succeeded in reflecting the assumptions of increasing difficulty with increasing grade and within the three types of schools in the Dutch educational system. As a matter of fact, it can be concluded that this readability formula is able to provide a rough estimation of the difficulty level of Dutch teaching material. Secondly, the results from this study suggest that texts with a lower Flesch-Kincaid Grade Level score (lower than 8) offer a greater likelihood of being understood, compared to texts with a score above 11. Consequently, the Flesch-Kincaid Grade Level can be used as an indication for text difficulty when it comes to the selection of texts.

Furthermore, the findings of this study reveal that in the fourth grade notably little incorrect answers were given when compared to the other three grades. By the age of 15 a shift in reading comprehension takes place which is why readability scores seem to have lost their explanatory power when the reader has reached a certain level of reading comprehension.

Limitations & future research

As was stated earlier, measuring reading comprehension is challenging. Because of that and reasons that are addressed in this section, results of this study should be treated with caution. In the present study, the comprehension test that has been generated for this study has not been validated before using it. Some of the questions that were used for the test were originally taken from teaching books. The rest of the questions have been created from scratch. It was desirable for this validation study to use a reading comprehension test that is already validated. This, however, was impossible due to the fact that it was a priority in this study to use authentic teaching material. Reason for that was the underlying assumption that teaching material differs between grade and type of school and the assumption that the publishers of teaching material are proficient in the design of these books. Nevertheless, even teaching material can be rated as being too easy or too difficult according to teachers and students as was mentioned by several teachers in preparation of this study. So, even the categorization of teaching material in different grades and types of schools cannot be taken for granted and should be evaluated carefully. Limiting the used texts to the reading sections for this study was mainly due to practical implications of this study. A broader analysis of all texts of the textbooks for the subject Dutch and also other subjects appears to be desirable.

Another limitation of this study is the lack of identifying the actual reading level of the participants. In the Flesch Reading Ease validation study of Wolf (2013), reading level of the participants was identified beforehand and compared with the scores of comprehension tests afterwards. In this study, this was not done. One reason for that was that the Flesch-Kincaid Grade Level scores reflect the necessary reading ability of an average reader in grade levels. Therefore, it seemed superfluous to identify reading ability of each participant.

For the experiments of this study, texts with low and high readability score were used. Rationale for that was the need for dispersion of the data. Klare (1976) found that it is easier to find significant differences in comprehension when the range of readability scores is wide. For future research, it would be interesting to also test texts with less extreme readability scores as for instance the same grade level scores as the actual grade level of the tested participants. Also comparing texts with the same score with each other would be interesting.

In this study, only havo students were tested. In order to generalize the results to the other types of schools, it is recommended to perform this experiment with vmbo and vwo students in future studies. In doing so, it can be investigated whether there are different cut-off points that mark the transition from easy to difficult texts. Furthermore, the design of a follow-up study should be a within-subjects design. Participants should get as well the easy as

the difficult texts. Comparing the performance on the comprehension questions from the easy and difficult condition with each other might have more explanatory power than comparing the performance of two forms with each other.

Recommendations for Kennisnet

The findings of this study can have implications for the selection of Dutch teaching material for havo students with Edurep. Although the Flesch-Kincaid Grade Level is not as accurate as it would be desirable, it can be used to make a rough selection of texts. This is less time consuming than an elaborate estimation of text difficulty by a human being. For the first and third grade, choosing texts with a Flesch-Kincaid Grade Level score that is lower than a certain score might increase the likelihood of understanding the text. According to this experiment this would be a Flesch-Kincaid Grade Level score lower than 8. Furthermore, the focus of the application of the Flesch-Kincaid Grade Level should lie on teaching material for students below the age of 15. From this age, the readability of a text might not be as important for text comprehension as in earlier years.

References

- Bormuth, J.R. (1967). Cloze Readability Procedure. *CSEIP Occasional Report No.1*, University of California. Los Angeles.
- Bruce, B., Rubin, A. & Starr, K.A. (1981). Why readability formulas fail. *IEEE Transactions on Professional Communication*. 24, 50-52.
- Chall, J. S. (1983). *Stages of reading development*. New York: McGraw-Hill.
- DuBay, W. H. (2004). *The principles of readability*. Costa Mesa, CA: Impact Information.
- DuBay, W. (2007). *Smart language: Readers, readability, and the grading of text*. Costa Mesa, CA: Impact Information.
- Dale, E. & Chall, J.S. (1948). A formula for Predicting Readability. *Educational Research Bulletin*. 27 (1). 11-28.
- Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, 32 (3), 221-233.
- Florida Laws: FL Statutes - Title XXXVII Insurance Section 627.4145 (n.d.). Retrieved September 30th, 2013 from <http://law.onecle.com/florida/insurance/627.4145.html>.
- Fry, E.B. (1968). A readability formula that saves time. *Journal of reading*, 11, 513-516.
- Heydari, P. (2012). The Validity of Some Readability Formulas. *Mediterranean Journal of Social Sciences*, 3 (2), 423-435.
- Heydari, P. & Riazi, A.M. (2012). Readability of texts: Human evaluation versus computer index. *Mediterranean Journal of Social Sciences*, 3 (1), 177-190.
- Kincaid, J. P., R. P. Fishburne, R. L. Rogers, & B. S. Chissom. (1975). Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading

- Ease Formula) for Navy enlisted personnel. *CNTECHTRA Research Branch Report*. 8-75.
- Kincaid, J. P., Aagard, J. A., O'Hara, J. W., & Cottrell, L. K. (1981). Computer readability editing system. *IEEE Transactions on Professional Communication*, 24 (1), 38-42.
- Klare, G. R. (1976). A Second Look at the Validity of Readability Formulas. *Journal of Literacy Research*, 8 (2), 129-152.
- Kraf, R., Lentz, L., & Pander Maat, H. (2011). Drie Nederlandse instrumenten voor het automatisch voorspellen van begrijpelijkheid. *Tijdschrift voor Taalbeheersing*, 33 (3), 249-265.
- Leichter, S.B., Nieman, J.A., Moore, R.W., Collins, P. & Rhodes, A. (1981). Readability of self-care instructional pamphlets for diabetic patients. *Diabetes Care*, 4 (6), 627-630.
- Lucassen, T., Dijkstra, R. L., & Schraagen, J. M. (2012). Readability of Wikipedia. *First Monday*, 17 (9).
- Marnell, G. (2008). Measuring readability, part 1: the spirit is willing but the Flesch is weak. *Southern Communicator*, 14 (1), 1-16.
- Maroof, D.A. (2012). *Methods in Neuropsychology: Common Procedures Made Comprehensible*. Springer, New York.
- Masson, M. E. J., & Waldron, M. A. (1994). Comprehension of legal contracts by non-experts: Effectiveness of plain language redrafting. *Applied Cognitive Psychology*, 8, 67 - 85.
- McLaughlin, G.H. (1969). SMOG grading – a new readability formula. *Journal of Reading*, 22, 639-646.
- McLaughlin, G.H. (1974). Temptations of the Flesch. *Instructional Science*, 2 (4), 367-383.

- Meade, C. D., & Smith, C. F. (1991). Readability formulas: cautions and criteria. *Patient education and counseling*, 17 (2), 153-158.
- van Oosten, P., Tanghe, D., & Hoste, V. (2010). Towards an improved methodology for automated readability prediction. *Proceedings of the seventh International Conference on Language Resource Evaluation (LREC'10)*. 775-782.
- Pichert, J. W., & Elam, P. (1985). Readability formulas may mislead you. *Patient Education and Counseling*, 7 (2), 181-191.
- Pikulski, J.J. (2002). *Readability*. Retrieved from <http://www.eduplace.com/state/author/pikulski.pdf>
- Rajagopalan, M. S., Khanna, V. K., Leiter, Y., Stott, M., Showalter, T. N., Dicker, A. P., & Lawrence, Y. R. (2011). Patient-oriented cancer information on the internet: a comparison of Wikipedia and a professionally maintained database. *Journal of Oncology Practice*, 7 (5), 319-323.
- Rameezdeen, R., & Rajapakse, C. (2007). Contract interpretation: the impact of readability. *Construction Management and Economics*, 25 (7), 729-737.
- Redish, J. (2000). Readability Formulas Have Even More Limitations Than Klare Discusses. *ACM Journal of Computer Documentation*, 24 (3), 132-137.
- Richards, J.C., Platt, J. & Platt, H. (1992). *Longman dictionary of language teaching and applied linguistics*. London: Longman.
- Snow, C. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Rand Corporation.
- Wegner, M. & D. Girasek. (2003). How readable are child safety seat installation instructions? *Pediatrics*, 111 (3), 588-591.

Wolf, R. (2013). The Flesch Readability Formula: Still alive or still life? *Bachelor Thesis*, University of Twente, the Netherlands.

APPENDIX A

Posed questions to each text

Questions: * = originally taken from schoolbooks

** = needed to be split up for analysis because two questions were asked in one question or answer could be interpreted in more than one way.

Grade 1

	Low readability scores	High readability scores
VMBO	1) Waarover gaat aflevering 9 van het programma “Naar de brugklas”? * 2) Wie geeft tips over het maken van nieuwe vrienden? 3) Hoe vaak komt het volgens de tekst voor dat je in groep 8 vrienden in de eerste klas kent? 4) In welke klas zit Gülsah? 5) In welke klas zit Halima?	1) Wie heeft de scholenactie bedacht? * 2) Welke opdracht kregen scholieren in deze actie?* 3) Waarom kregen ze deze opdracht?* 4) Waarom heeft de jury “vriespect” gekozen?* 5) Wat betekent het woord “vriespect”?
HAVO	1) Hoeveel jeugdteams zijn er? 2) Hoeveel leden zijn onder de 18 jaar? 3) Waarom is de auteur van de tekst blij dat er een net zit tussen de teams? 4) Wat vindt de auteur leuk aan teamsport?	1) Wat is er ontwikkeld tegen stress bij huisdieren? 2) Op grond waarvan komt de schrijver tot de conclusie dat er veel behoefte is aan gespecialiseerde huisdierproducten? 3) Waar komen de meeste nieuwe huisdierproducten vandaan? 4) Waardoor kan verklaard worden dat huisdieren als modeaccessoire gezien worden?
VWO	1) Waarom zijn de mensen begonnen met het vieren van carnaval?* 2) Waarom wordt carnaval het meest gevierd in de provincies Brabant en Limburg? 3) Wat betekent “carnaval” letterlijk?	1) Wat was de aanleiding voor Peter Benenson voor het schrijven van het artikel dat in het Engelse dagblad The Observer verscheen? 2) Wat betekent het woord “amnestie”? 3) Wat is het verschil tussen

- | | |
|--|--|
| <p>4) Hoeveel dagen wordt er tegenwoordig carnaval gevierd?</p> <p>5) Noem de twee verschillende redenen op die in de tekst genoemd worden waarom mensen zich met carnaval verkleden. **</p> | <p>gewetensgevangenen en politieke gevangenen? **</p> <p>4) Wat is de houding van Amnesty International ten opzichte van de doodstraf?</p> <p>5) Noem twee manieren om Amnesty International te helpen. **</p> |
|--|--|

Grade 2

Low readability scores

High readability scores

VMBO

- | | |
|---|--|
| <p>1) Wat is een andere naam voor heel veel overgewicht?*</p> <p>2) Hoeveel procent van de Amerikaanse bevolking heeft overgewicht?</p> <p>3) Hoe ontwikkelt overgewicht zich bij de jeugd in Nederland?</p> <p>4) Hoeveel procent van de Nederlandse kinderen heeft overgewicht?</p> <p>5) Wat vindt Zandbol mooier? Dun of dik?</p> | <p>1) Waarvoor ben je met een doorlopende reisverzekering verzekerd?</p> <p>2) Wanneer ontvang je niet direct een polisnummer bij het online afsluiten van een reisverzekering?</p> <p>3) Wanneer is een doorlopende reisverzekering voordeliger dan een gewone reisverzekering?</p> <p>4) Wanneer kan je niet direct de schade online melden?</p> |
|---|--|

HAVO

- | | |
|--|---|
| <p>1) Waarom had Godwin “waarschijnlijk wel” zijn pas afgegeven?*</p> <p>2) Wat gebeurde er met de pas van de kennis van Sculli?</p> <p>3) Waarom kon de broer van de zestienjarige jongen lange tijd geen rekening openen?</p> <p>4) Wat soort jongen was de jongen die Cynthia om haar pas vroeg?</p> <p>5) Wat vindt Cynthia van de actie met de blauwe plastic pasjes?</p> | <p>1) In welk opzicht verschilt het nieuwe spel van eerdere spellen?</p> <p>2) Wanneer vindt de transformatie plaats?</p> <p>3) Hoeveel karakters heeft het nieuwe spel Pokémon?</p> <p>4) Hoe wordt bepaald naar welke Pokémon je transformeert?</p> |
|--|---|

VWO

- | | |
|---|--|
| <p>1) Waarom worden Janine Jansen en haar vriend de “Posh ‘n’ Becks van de viool”</p> | <p>1) Wat heeft Christiaan Huygens drie eeuwen geleden al voorspeld?</p> |
|---|--|

- | | |
|---|---|
| <ul style="list-style-type: none"> 2) genoemd?* 3) Welk instrument speelt Janine Jansen? 4) Waarom vond Paul Popma de ophef over de foto op de voorkant van de tweede cd van Janine niet erg?* 5) Waarom is Janine multimediaal?* | <ul style="list-style-type: none"> 2) Waarom is het lastig planeten die op de aarde lijken te vinden? *, ** 3) Wat is het probleem met kleine planeten i.v.m. metingen? 4) Wat zou een ruimtetelescoop met een spiegel van dertig centimeter kunnen vastleggen? 5) Waarom zijn de beelden die de ruimtesondes maken niet alleen waardevol voor het ontdekken van andere planeten? |
|---|---|

Grade 3

Low readability scores

High readability scores

VMBO

- | | |
|--|--|
| <ul style="list-style-type: none"> 1) Wat is de reden dat Jesse aan het bezuinigen is?* 2) Welke bezuinigingstip heeft Jesse?* 3) Waar kan je nog meer tips vinden?* 4) Wat kan je doen als je zelf een tip hebt?* | <ul style="list-style-type: none"> 1) Waardoor lopen veel leerlingen metaalbewerking risico tijdens praktijklessen?* 2) Waarop is de uitspraak van vraag 1 gebaseerd? 3) In hoeveel procent van de praktijklokalen waren de werkplekken veilig? 4) Wat is volgens de Arbeidsinspectie vooral het probleem? |
|--|--|

HAVO

- | | |
|--|---|
| <ul style="list-style-type: none"> 1) Lezen tegenwoordig meer of minder mensen een boek in vergelijking met 30 jaar geleden? 2) Hoeveel minuten lezen mensen dertig jaar geleden minimaal per week? 3) Wat is de “nieuwe vorm van lezen”? 4) Wat is de verklaring voor het verschil in leesgedrag? | <ul style="list-style-type: none"> 1) Welke positieve gevolgen verwacht de auteur van accijnsverhoging? *, ** 2) Welk indirect gevolg van een verbod op happy hours verwacht de auteur?*, ** 3) Welke twee effecten zijn er volgens de auteur verbonden aan hogere leeftijdsgrenzen voor alcoholverkoop?* 4) Welke mogelijke verklaring geeft de auteur voor het feit dat Nederland geen serieus alcoholmatigingsbeleid kan voeren?*, ** 5) Wat wordt bedoeld met “zet echt geen |
|--|---|

- 6) zoden aan de dijk” in: “Die stijgingen veranderen ook nooit zo veel, net als nu: de aangekondigde verhoging van het bieraccijns met twee cent (van negen naar elf cent per flesje) zet echt geen zoden aan de dijk.”

VWO

- | | |
|--|---|
| <ol style="list-style-type: none"> 1) Wat heeft aan Renates verandering bijgedragen? 2) Hoe beschrijft Renate zichzelf? 3) Besteedt Renate veel tijd aan haar uiterlijk? 4) Waardoor kan Renate chagrijnig worden? | <ol style="list-style-type: none"> 1) Waarom werkten Amerika en de Sovjet-Unie eerst met proefdieren?* 2) Om welke twee redenen beloofde John F. Kennedy dat er voor het einde van het decennium een Amerikaan op de maan zou lopen?*, ** 3) Waardoor kwam de staatsfinanciering van de Russische ruimtevaart op een lager niveau te liggen? ** 4) Hoeveel toeristen hebben al een bezoek gebracht aan de ISS? 5) Waarnaar wordt onderzoek gedaan in ruimtestations? |
|--|---|

Grade 4**Low readability scores****High readability scores****VMBO**

- | | |
|--|--|
| <ol style="list-style-type: none"> 1) Wat is de nationaliteit van Samira, Souad en Mina? 2) Wat is de mening van Pieter over getrouwde vrouwen die willen werken?* 3) Wat is de mening van Samira over werken als je getrouwd bent?* 4) Hoe wil Samira later bereiken dat zij kan blijven werken als zij getrouwd is?* 5) Wat is de mening van Pieter over werkende moeders?* 6) Waarom is Mina zo gedreven om haar diploma's te halen?* | <ol style="list-style-type: none"> 1) Wat leek Jacqueline leuk aan het beroep van popjournaliste?* 2) Waarom kon Jacqueline geen Journalistiek studeren? 3) Waarom kon Jacqueline niet als ziekenverzorgende in verpleeghuizen werken? 4) Wat is de mening van Jacqueline over haar huidige baan?* 5) Wat is de doelgroep van de zorginstelling waar zij werkt? |
|--|--|

HAVO

- | | |
|---|--|
| <ul style="list-style-type: none">1) Voor welke leeftijd is de Bombaytour geschikt?2) Waar vindt de Bombaytour plaats?3) Wat is de grootste stad van India?4) Wat maakt het meisje Gauri zo bijzonder? | <ul style="list-style-type: none">1) Wat is het criterium dat Marc Abrahams hanteert voor het kiezen van de genomineerde?2) Waarom kozen Bart Knols en Ruurd de Jong ervoor kaas te gebruiken voor hun experimenten?3) Waarom krijgt een specht geen hoofdpijn als hij zijn snavel in een boomstam jast?4) Wat houdt de “puberverjager” in?5) Voor welke uitvinding kregen Francis Fesmier en collega’s de Ig Nobelprijs?6) Wat was de taak van Miss Sweetie Poo? |
|---|--|

VWO

- | | |
|--|--|
| <ul style="list-style-type: none">1) Op welke manieren zoeken ouders een voornaam voor hun kind?2) Waarom wil Rolf zijn naam veranderen?3) Heeft de rechtbank de naamsverandering toegestaan?4) Was Rolf tevreden met de uitspraak? | <ul style="list-style-type: none">1) Wie plegen minder delicten binnen zes jaar na ontslag uit de gevangenis? Langgestrafte gedetineerden of ex-tbs’ers?2) Hoe kan verklaard worden dat men het gevoel heeft dat tbs’ers op verlof een grote bedreiging zijn voor de maatschappelijke veiligheid?3) Wat wordt hier bedoeld met “een kat in het nauw maakt immers rare sprongen”?4) Wat draagt bij aan de effectiviteit van de behandeling van een tbs’er?5) Worden tbs’ers gestraft voor hun delict? |
|--|--|
-

APPENDIX B

Texts for the easy condition

First grade

Tekst 1:

Nieuwe vrienden

In aflevering 9 van het programma ‘Naar de brugklas’ praat Gülsah over het maken van nieuwe vrienden. In groep 8 heb je soms vrienden die je uit de eerste klas kent. Het is dan ook raar als je misschien volgend jaar niet meer bij elkaar op school zit. Hoe maak je nieuwe vrienden? En moet je dat de eerste dag al doen of kan het ook halverwege het schooljaar? Op de middelbare school ontmoeten Gülsah en de groep 8-ers brugger Halima. Zij heeft heel veel nieuwe vrienden gemaakt en zij geeft tips hoe je dat moet doen.

Tekst 2:

Mijn Sport

“Ik zit al vier jaar op volleybal. Ik ben ermee begonnen toen ik in groep 6 zat. Eerst speelde ik bij de mini’s; dat zijn kinderen tot 11 jaar. Maar nu speel ik bij de jeugd: meisjes C. Als je een C-speler bent, betekent dat dat je 12 of 13 jaar bent. Ik ben nu 13. Onze club heeft meer dan 350 leden. Ongeveer de helft daarvan is jeugd, dat wil zeggen: onder de 18 jaar. Bij de senioren zijn er tien damesteams en tien herenteams. Bij de jeugd zijn er vijf teams bij de meisjes en drie bij de jongens. En dan zijn er nog heel veel miniteams. Volleybal is de leukste sport die er is. Je moet goed kijken en snel bewegen, want je moet altijd op de plek staan waar de bal komt. En je moet veel springen. Bijvoorbeeld bij het net als je wilt aanvallen of als je de bal juist wilt tegenhouden. Of bij de sprongservice, want die hebben wij ook al geleerd. Ik ben blij dat er bij volleybal een net zit tussen de teams. Je kunt dan niet geduwd of geschopt worden als je aan het spelen bent, zoals bij voetbal of handbal. Verder vind ik het heel leuk dat je altijd met een team samen speelt. Je ziet elkaar heel vaak: twee keer bij de training en natuurlijk bij een wedstrijd. En sommige teamgenootjes zitten ook nog bij mij op school, dus zo word je echte vriendinnen.

Tekst 3:

Vier dagen feest

Het plaatsje Schaijk is uitgestorven. Je ziet geen mens op straat. De mensen bereiden zich voor op het carnaval van komend weekend. Iedereen verkleedt zich, want met carnaval ben je iemand anders dan de rest van het jaar. Met carnaval verdwijnen voor even de verschillen tussen rijk en arm. Vier dagen lang is iedereen gelijk. Carnaval is een heel oud feest. Het heeft te maken met het rooms-katholieke geloof. Vroeger gingen de katholieken veertig dagen voor Pasen vasten. Dat betekende dat je minder snoep at en geen vlees. Zo komt het feest ook aan zijn naam: Carne-Vale betekent zoiets als 'vlees vaarwel'. Voordat de tijd van vasten begon, gingen de mensen nog één keer uit hun bol. Nog één keer heel veel eten, snoepen en feesten dus. Die vastenavond noemen we nu carnaval. Alleen wordt er inmiddels niet één avond feestgevierd, maar vier dagen. Niet alleen in Nederland wordt carnaval gevierd. Ook in andere landen waar katholieken wonen, vieren de mensen dit feest. Bijvoorbeeld in de beroemde Braziliaanse stad Rio de Janeiro, en ook in Duitsland, België en Amerika. In Nederland wordt carnaval het meest gevierd in Brabant en in Limburg. Daar wonen de meeste katholieken. Met carnaval zijn alle cafés open. Gedurende het feest is iedereen gelijk, of je nu arm bent of rijk. Daarom verkleedt iedereen zich, dan zie je geen verschillen meer tussen de mensen. In Brabant dragen veel mensen boerenkleding.

Second grade

Tekst 1:

Obesitas

Als je veel te veel weegt en dus overgewicht hebt, dan heet dat obesitas. Dit probleem komt steeds vaker voor, ook in Nederland. In de Verenigde Staten komt obesitas voor bij 25%-30% van de bevolking. Het percentage van de mensen met obesitas in Nederland is nu ongeveer 10%.

Kinderen

Hoewel obesitas bij kinderen nog maar voorkomt bij minder dan 1%, heeft al 10%-16% van de kinderen in Nederland overgewicht. Bij de hele jeugd van 3-21 jaar is een sterke stijging te zien van overgewicht en obesitas de afgelopen jaren. Deze stijging van obesitas heeft grote invloed op de volksgezondheid.

Poll/stelling: Is dun mooier dan dik?

22-08-05 door La_Lina: Ik vind van niet. Want stevigere mensen kunnen toch heel mooi zijn? Kijk maar naar Beyonce. Niet dat zij heel stevig is.

22-08-05 door AngelOfLove: Ik vind het ook niet echt mooi als iemand dun is. Nou ja, ik ben het gewoon met jou eens.

24-08-05 door Yummy: Ik vind het best mooi als iemand dun is, zolang we maar niet over mager gaan spreken, een beetje vet om de armen en benen enzo, mag van mij wel.

24-08-05 door Zandbol: Ik vind dun echt lelijk. Vooral omdat ik het zelf ook ben! Dikker is mooier.

25-08-05 door La_Lina: Dus jij vindt jezelf niet mooi, Zandbol?

25-08-05 door Yummy: Ik weet niet hoe je eruit ziet, maar je moet jezelf niet lelijk vinden hoor!!! Echt niet, want iedereen is mooi.

26-08-05 door La Lina: Dat vind ik ook!! Iedereen op z'n eigen manier. mooi gezegd trouwens, Yummy!

31-08-05 door @dje: Het is allebei mooi: je hebt mensen die er dun mooi uitzien en mensen die er juist mooier uitzien als ze wat dikker zijn. Zolang ze maar niet té dik of té dun zijn, vind ik het best.

Tekst 2:

Iedereen weet toch dat het niet klopt

Godwin Arhin (13) van het Augustinus College in Zuidoost heeft zijn jaszak vol met blauwe plastic pasjes gepropt. De pasjes werden gisteren op scholen in onder meer Zuidoost uitgedeeld. Met de tekst 'Wees slim, laat je niet vastpinnen' werden scholieren gewaarschuwd hun bankpas niet af te geven in ruil voor geld. Daardoor is

Arhin nu op de hoogte. ‘Als iemand me een dag geleden voor honderd euro om mijn pas had gevraagd, had ik het waarschijnlijk wel gedaan. De meeste kinderen doen dingen voor geld. Nu zou ik het niet meer doen. Ik weet nu dat je dan zelf ook crimineel bent.’

Sculli (16) vertelt dat een kennis van hem een jaar geleden zijn pas wel heeft gegeven. ‘Er is toen heel veel geld op zijn rekening gestort. Hij heeft later aangifte gedaan en gezegd dat zijn pas gestolen was. Ik weet niet hoe het is afgelopen.’

Ook bij de Open Scholengemeenschap Bijlmer kent iedereen de verhalen. Een zestienjarige jongen vertelt dat zijn broer, destijds vijftien, een paar jaar geleden zijn pas had uitgeleend. ‘Een vriend van hem had het ook gedaan. Op diens rekening werd heel veel geld gestort. Die vriend kreeg tien procent van dat bedrag en kocht daar mooie kleren en schoenen van. Mijn broer wilde dat ook, hij wilde ook merkkleding. Zijn vriend bracht hem daarom in contact met die jongens. Maar mijn broer liep tegen de lamp. De bank belde naar mijn ouders omdat er heel veel geld op zijn rekening was gestort. Dat vonden ze ongewoon.’ Zijn ouders hebben het er thuis nog steeds over, vertelt de jongen. ‘Mijn broer heeft het moeilijk gehad. Hij studeert nu, maar hij kon al die tijd geen rekening meer openen. Als je bij zo’n criminele actie betrokken bent geweest, dan wil geen bank je namelijk hebben. Hij had zelfs

problemen met zijn studiefinanciering. Pas na vijf jaar wordt alles kwijtgescholden. Nu gaat het weer goed met hem. Hij heeft er spijt van dat hij zichzelf zo in de problemen heeft gebracht.’ De achttienjarige havoscholiere Cynthia van Scholengemeenschap Reigersbos is het al enkele keren gevraagd. De eerste keer was het: ‘Wil je geld verdienen? Geef je pinpas maar, je krijgt 150 euro. Ik vroeg wat ze dan deden. Ze zeiden dat ze geld op mijn rekening zouden storten en dat er binnen een dag weer af zouden halen. Meer niet. Maar iedereen weet toch dat het niet klopt.’ Een andere keer kreeg Cynthia zelfs vijfduizend euro aangeboden. Ze zou haar pas dan wel een week kwijt zijn. ‘Ik heb bedankt. Ik verdien het liever wit. Je pincode is voor jezelf.’ Ze kende de jongen die haar om haar pas vroeg, trouwens goed. ‘Het is een goede vriend van mij. Hij studeert nu op de Hogeschool van Amsterdam. Het is een heel normale jongen, met een overhemd en een spijkerbroek. Dit soort dingen verbaast me niet. Ik woon in Zuidoost. Veel mensen zijn hier met foute dingen bezig.’ Op het schoolplein en in de omgeving liggen tal van blauwe waarschuwingspasjes op de grond. Dat de politie nu is komen waarschuwen, vindt Cynthia overbodig. ‘Iedereen op school kent de verhalen van de pinpassen. We weten heus wel dat mensen bezig zijn om je pasje te lenen. Iedereen heeft het erover. Daarom zijn al die waarschuwingspasjes echt niet meer nodig.’

Tekst 3:

Hoe doet ze het? Met haar stijl, een rode jurk en talent!

Aan beide kanten van de zaal staat op de gang een lange, dubbele rij jonge mensen te wachten. Als het belsignaal gaat, sprint het publiek de zaal in. De mensen duwen en trekken voor een staplaats vooraan. Een concert van Marco Borsato of van Anouk? Nee, vanavond speelt in de Weense Musikverein de Nederlandse violiste Janine Jansen. De zaal is uitverkocht! Janine Jansen (28) weet een breed publiek enthousiast te maken voor klassieke muziek. Ze staat symbool voor een nieuwe generatie musici. Hoe doet deze ‘popster’ het?

Ze speelt goed. Jansen steekt met kop en schouders boven de andere violisten in Nederland uit, zei dirigent Jaap van Zweden ooit over haar. ‘Muziekster Vanessa Mae sprak ook een breder publiek aan, maar zij was minder gewild bij muziekcritici. Jansen weet met haar spel ook de gevestigde orde te overtuigen.’ Als enige Nederlandse soliste heeft Jansen een contract bij de grote platenmaatschappij Decca. In 2004 en 2005 won ze de Edison publieksprijs Klassiek. In 2005 opende ze de Night of the Proms in Londen. Jansen voelt zich niet te goed voor lichte muziek. Vroeger was dat toch meer een taboe, zegt ze zelf. Maar dat is nu niet meer zo. Jansen speelt ook met Trijntje Oosterhuis en Karin Bloemen. Met Trijntje nam ze een single op voor prinses Amalia. Jansen: ‘Wat mij betreft kan alles, als je er zelf maar achter staat. Muziek is muziek en emoties zijn emoties.’ Een single met Marco Borsato, zou dat dan geen geheide kaskraker zijn? ‘Jazeker,’ zegt ze lachend. Ach, wie weet. Als-ie me zou vragen. Daarnaast weet ze zich te verkopen. Jansen en haar vriend Rachlin, ook violist, zijn de Posh ‘n’ Becks van de viool, schreef de Britse krant The Daily Telegraph ooit, in een verwijzing naar het duo Victoria en David Beckham. Wat ze delen, is gevoel voor marketing. Zo droeg Jansen op haar eerste cd-hoes dezelfde knalrode jurk van couturier Mart Visser, die ze ook had gedragen tijdens het Bevrijdingsconcert. Dat concert was uitgezonden op tv. Bovendien is ze mooi en durft ze daar voor uit te komen. Voor haar tweede cd, een opname van Vivaldi’s Vier jaargetijden, liet Jansen zich in jurken fotograferen die zo bloot waren, dat de pers er niet over uitgepraat raakte. ‘De halve wereld viel erover,’ zegt Popma van Universal Music, ‘maar zoiets geeft wél extra publiciteit.’ En dat komt goed uit, want het publiek voor klassieke muziek wordt kleiner. Vroeger was er geen marketing nodig als er een nieuwe klassieke plaat uitkwam; die verkocht zichzelf. Daarbij is ze ook multimediaal. Jansen was met haar Vivaldi-album de meest gedownloade klassieke artieste op itunes. Ook zijn op www.janinejansen.com fragmenten van haar nieuwe cd te beluisteren. Het is mogelijk een e-

card te sturen naar een vriend, met foto's en geluid. En er is een gastenboek, waar fans berichtjes achterlaten. Martin schreef: 'Jouw vioolspel raakt me echt! Ik word er week van en raak ontroerd, terwijl ik eigenlijk meer een liefhebber van pop, rock en metal ben.' En Dante: 'If I weren't married, I would propose to you. Lovely music played very well by a beautiful woman.' Paul Popma van Universal: 'Zulke musici zijn mensen van deze tijd. Ze begrijpen de media van nu.' Jansen kijkt dus goed om zich heen. Hoewel ze muziek speelt die al eeuwen oud is, op een bijna even oud instrument, weet Janine Jansen goed wat er in de huidige tijd mogelijk is. Ze gebruikt deze mogelijkheden om een breed publiek enthousiast te maken voor 'haar' muziek.

Third grade

Tekst 1:

Tips & trucs

Jesse (16): 'Ik ben aan het bezuinigen omdat ik een tijdje te veel geld heb uitgegeven. Maar ja, wat doe je dan? Ik kon veel besparen door mijn abonnement op een tijdschrift op te zeggen, maar dat wilde ik niet. Vertelde iemand me dat ik dit tijdschrift gratis in de bibliotheek kon lezen!' Bezuinigen, besparen en lekker veel geld hebben. Iedereen houdt zich daar wel mee bezig. Maar hoe pak je dat nou aan? Wat moet je wel doen en wat niet? Tips en trucs zijn dan van harte welkom. En ook waar je niet in moet stinken!

Tipmachine

Heb je een gat in je hand? Kost je mobiel je teveel geld? Onze tipmachine levert je tips die precies bij jouw financiële probleem passen. Probeer het uit!

Op = op spel

Heb je zelf een wereldtip? Daarmee maak je kans op een leuke prijs. Stuur je tip in, dan ontvang je een lotnummer waarmee je direct kunt checken of je gewonnen hebt. Kijk snel.

Tekst 2:

Het boek verliest terrein

AMSTERDAM. Volgens het Sociaal en Cultureel Planbureau las dertig jaar geleden 49 procent van de Nederlanders wekelijks minimaal een kwartier in een boek. Dit jaar is dat 31 procent. Toch wordt er meer dan ooit gelezen in Nederland. Ra, ra, hoe kan dat? Internet en sms hebben een nieuwe vorm van lezen geïntroduceerd,' stelt directeur Aad Meinderts van de Stichting Lezen. De gestage neergaande lijn van lezen in boeken wordt ook veroorzaakt door een wijziging in het vrijetijdspatroon. Er valt steeds meer te kiezen.'

Tekst 3:

Een nieuwe look

Renate is totaal veranderd! Natuurlijk heeft haar nieuwe kapsel daar een groot aandeel in. Net als de trendy make-up die Thijs adviseerde. Nu maar hopen dat ze dit voortaan elke dag doet voordat ze naar school gaat. Van 'girl-next-door' naar prachtmeid!

Wat vind jij van jezelf?

Ik vind mezelf onopvallend en sportief. Eigenlijk houd ik er niet van om in het middelpunt van de belangstelling te staan. Maar vandaag vond ik het wel leuk om tips en adviezen te krijgen. Ik ga ongeveer een keer in de twee maanden naar de kapper, maar verder doe ik weinig aan mijn uiterlijk.

Gebruik je vaak make-up?

's Ochtends moet ik altijd stressen om de bus te halen. Make-up gebruik ik alleen als ik een feestje heb. Een beetje mascara en lippenstift, meer niet. De lippenstift van LOLA is fijn, want die blijft lekker lang zitten.

Hoe belangrijk is je uiterlijk voor jou?

Als mijn haar niet goed zit, kan ik daar wel even chagrijnig van worden. Op die momenten doe ik een elastiekje in mijn haar en kijk ik er de rest van de dag niet meer naar om.

De nieuwe kleuren lippenstift van LOLA zijn geschikt voor elke gelegenheid. Er zijn minder opvallende kleuren die ervoor zorgen datje er toch verzorgd uitziet en er zijn knalkleuren voor als je indruk wilt maken op die leuke jongen. LOLA lipstick blijft lang zitten en verzorgt je lippen. En de prijs? Die is ook prettig. Dus ren naar de winkel en scoor zo'n LOLA lipstick!

Fourth grade

Tekst 1:

‘Mijn vrouw mag best werken’

Het denken en doen van jongeren is behoudend, bleek uit de Emancipatiemonitor. Allochtone meisjes zijn het progressiefst.

nooit zeker, van tevoren,’ zegt Mina. ‘Na het huwelijk kunnen mannen erg veranderen. Eigenlijk moet je eerst samenwonen. Dan leer je elkaar goed kennen.’ Zij doet dat zeker, tenminste als haar ouders het goed vinden. ‘Bij Marokkanen is het niet zo gebruikelijk.’ Hoe gaan ze dat regelen als er kinderen komen? Dat is een heikel punt. ‘De vrouw zorgt voor de kinderen want die hebben dat, eh, meer in zich,’ zegt Pieter. ‘Dat is hun instinct.’ Deze keer wordt zijn uitspraak niet ontvangen met hoongelach, de meisjes denken na. ‘Het is waar dat een kind meer verbonden is met de moeder,’ zegt Samira aarzelend. ‘Zeker het eerste jaar. Bovendien, kinderen krijgen is het mooiste wat er is. En ze zijn snel groot, dat zou ik niet willen missen.’ ‘Maar voor een man is het toch ook goed om mee te maken hoe zijn kinderen opgroe

Tekst 3:

Kevin hockeyt niet

De voornaam is een belangrijk bezit. Niet voor niets zijn aanstaande ouders vaak maanden bezig om er één (of meer) te verzinnen. Ze ploegen door voornaamboeken heen, spellen de geboorteberichten in de krant en vuren voortdurend ideeën op elkaar af ('Wat vind je van Mark?' 'Nee! Er zat een vreselijke Mark bij mij in de klas.'). Ondanks die zorgvuldige voorbereiding gaat het toch wel eens fout. Zo wil de Utrechtse student Rolf graag Rolf Dudock Maurits heten. Daarom smeekte hij de rechtbank van Utrecht begin vorige maand zijn naam te veranderen. Zijn vrienden van het studentencorps zouden wegens zijn schamele voornaam op hem neerkijken, klaagde Rolf. Bovendien, pleitte de student vol vuur, zijn ouders hadden bij zijn geboorte al overwogen om de komst van hun telg met drie voornamen te bekronen. Ondanks de bezweringen van Rolf zag de rechtbank de noodzaak van

APPENDIX C

Texts for the difficult condition

First grade

Tekst 1:

Wat staat er tegenover geweld? Vriespect!

De Nederlandse taal is vanaf vandaag een nieuw woord rijker: vriespect. Dit woord is namelijk als winnaar uit de bus gekomen in de landelijke scholenactie van SOS Kinderdorpen, waarbij scholieren gevraagd werden een tegenh

rustgevende pillen voor het geval dat de huisdieren bij al die schoonheidsbehandelingen in de stress raken. En voor hondeneigenaren met een gevoelige neus is er nu ook een eau de toilette, onder de naam Fruits and Passions. Dit is, zoals 58 procent van de nieuwe huisdierproducten, afkomstig uit de VS. Volgens het onderzoeksbureau houden veel mensen evenveel van hun huisdieren als van hun kinderen. Ze willen hun lievelingen daarom alleen de beste producten geven. Daarnaast worden huisdieren steeds meer als modeaccessoires gezien. Dat komt volgens de trendwatchers van Mintel door sterren als Paris Hilton en Geri Halliwell, die vaak met een hondje in hun tas lopen.

proces. Dat proces moet niet al te lang op zich laten wachten. Sommige politieke gevangenen zitten maanden of jaren vast voordat hun zaak voor de rechter komt.

Mensenrechten

Amnesty International vindt dat iedereen het recht heeft op leven en op bescherming tegen marteling. Verder vindt Amnesty dat iedereen het recht heeft om te denken en te zeggen wat hij wil. Ook mag volgens Amnesty niemand worden gediscrimineerd. Amnesty doet onderzoek naar ernstige schendingen van deze mensenrechten en

Direct online reisschade melden

Als je schade hebt, kun je dit online melden en aan de bank mailen, behalve als het om schade gaat waarbij je het ingevulde formulier moet printen en samen met je bewijsstukken moet opsturen (bijvoorbeeld bij waterschade).

Voorwaarden

Geldig voor jongeren tussen de 12 en 18 jaar,

M

Wereld moet men dan met het verstand beseffen: zoo vele Zonnen, zoo vele Aardklooten,' schreef hij in zijn in 1698 postuum gepubliceerde werk 'Cosmotheoros'. Het zou tot 1995 duren voordat de Zwitserse astronomen Michel Mayor en Didier Queloz van het Geneefse Observatorium voor het eerst echt een planeet bij een verre ster, een exoplaneet, waarnamen. Inmiddels zijn er rond de 210 exoplaneten gevonden, maar nooit leken die op onze planeet 'aarde'. Bijna steeds hadden ze het formaat van de reuzenplaneet Jupiter of nog groter en cirk

Third grade

Tekst 1:

Leerlingen lopen veel risico in praktijklokalen

DEN HAAG (ANP) —

Zo'n elfduizend leerlingen metaalbewerking lopen dagelijks risico tijdens praktijk lessen doordat ze met onveilige machines werken. Dat constateert de Arbeidsinspectie na een onderzoek bij 261 scholen. In 90 procent van de praktijklokalen werden onveilige situaties ontdekt. Het ministerie van Sociale Zaken en Werkgelegenheid heeft de resultaten van het onderzoek woensdag bekendgemaakt. Volgens de Arbeidsinspectie is het 'zorgelijk gesteld' met de veiligheid in de lokalen. De leerlingen hebben in de lessen te maken met onveilige apparaten zoals freesmachines en draaibanken. Het gaat volgens de Arbeidsinspectie vooral om het ontbreken van beschermingskappen. Veel van de onderzochte scholen namen overigens maatregelen nadat de inspectie langs was geweest.

Tekst 2:

De overheid steunt Koning Alcohol

Tot de landen waarin het meeste bier ter wereld wordt gedronken behoren Ierland, Engeland en Nederland. Vooral het alcoholmisbruik onder jongeren is in deze landen zeer hoog. Dit probleem is echter niet het enige wat de drie naties gemeenschappelijk hebben, blijktens door de Universiteit van Tilburg gecoördineerde enquêtes onder de bevolking: ze kennen ook de hoogste percentages slachtoffers van openbaar geweld in Europa. De maatschappelijke schade die wordt toegebracht door overmatig alcoholgebruik is hoog. Alcoholmisbruik is niet alleen de belangrijkste oorzaak van zinloos geweld, huiselijk geweld en seksueel geweld. Het is ook de primaire oorzaak van verkeersongelukken en een hoofdoorzaak van uiteenlopende ziektes waaronder, naast levercirrose (een chronische leverziekte) en dementie, ook kanker en hartinfarcten. Hersenen van jonge mensen lopen bovendien ernstige en onomkeerbare schade op ten gevolge van alcoholmisbruik. Het is dus hoog tijd dat de overheid met serieuze maatregelen het overmatig gebruik van alcohol tegengaat. Zij kan dat het beste doen met accijnsverhoging, een verbod op happy hours en hogere leeftijdsgrenzen voor alcoholverkoop. Accijnsverhoging leidt bijvoorbeeld onmiddellijk tot minder alcoholconsumptie onder

jongeren en vervolgens tot minder geweld, inclusief moorden. Dat hebben de hogere accijnzen in Zweden wel bewezen. Deze statistische verbanden zijn sterk. Toen Finland bij de toetreding van buurland Estland tot de Europese Unie zich gedwongen voelde om de accijnzen te verlagen, ging zowel het alcoholgebruik onder jongeren als het aantal geweldplegingen meteen omhoog. Een verbod op happy hours werkt ook. Zo'n verbod heeft bijvoorbeeld in Engeland geleid tot een directe daling van het aantal letselopnames in ziekenhuizen in de grote steden. Vooral jongeren gingen zich nu minder vaak te buiten aan alcoholmisbruik en dus waren er minder vechtpartijen. Hogere leeftijdsgrenzen voor alcoholverkoop, zoals 18 of 21 jaar, zijn eveneens bewezen effectief. Dat zagen we in de Verenigde Staten. De hogere leeftijdsgrenzen zijn daar vooral ingevoerd onder invloed van de Organisatie Moeders tegen Dronken Rijden. Zij hebben echter ook op het aantal openbare geweldplegingen een gunstige invloed gehad. Maar accijnsverhogingen hebben bij ons toch nog nooit invloed gehad op het alcoholgebruik? Dat klopt. Die stijgingen veranderden ook nooit zo veel, net als nu: de aangekondigde verhoging van het bieraccijns met twee cent (van negen naar elf cent per flesje) zet echt geen zoden aan de dijk. Vanwege inflatie en hogere inkomens is bier ten opzichte van vroeger zelfs goedkoper geworden. De kolossale externe kosten van alcoholmisbruik — de kosten van verpleging en criminaliteitsbestrijding waarvoor de gemeenschap opdraait — zouden in de prijzen moeten worden doorberekend. Deze operatie zou leiden tot accijnzen dubbeltjes omhooggaan. Door de huidige lage accijnzen te handhaven, subsidieert de overheid in feite de alcoholindustrie. Prijzen we onszelf de markt niet uit met hogere accijnzen? Nee, dat is niet het geval. De accijnzen zijn in Nederland niet hoger dan elders in Europa, hoewel dit vaak wordt beweerd. Er kan best een forse schep bovenop zonder dat hoeft te worden gevreesd dat Nederlandse jongeren massaal naar het buitenland afreizen om zich daar te laten vollopen. Ook tegen hogere leeftijdsgrenzen kan geen redelijk mens bezwaar maken. Een leeftijdsgrens van 18 of 21 jaar is veel gemakkelijker te handhaven dan een van 16 jaar. Op zijn minst zal het gemakkelijker worden om jongeren beneden de 16 uit de drankwinkels en kroegen te weren. Bovendien is er voor zo'n maatregel brede steun onder de bevolking. Blijkens landelijke enquêtes is 75% van de Nederlanders er voorstander van. Hoe is het toch te verklaren dat Nederland met zijn gigantische alcoholgerelateerde geweldsproblemen politiek niet in staat is een serieus alcoholmatigingsbeleid te voeren? Zou het kunnen zijn dat de bierlobby in ons land te machtig is? Zouden de aandeelhouders van Koning Alcohol de belangen van onze gezinnen behartigen in plaats van de politieke partijen? Ik hoop vurig van niet. Want in dat geval heeft

de Nederlandse overheid met een serieus alcoholmatigingsbeleid het probleem van alcoholmisbruik nog binnen de regeerperiode van dit kabinet opgelost.

Tekst 3:

Naar de maan

Verhalen en fantasieën over hoe de mens de aarde verlaat, zijn al heel oud. Ongeveer twee eeuwen na Christus publiceerde de Griek Lucianus van Sarnosata het verhaal *Verae Hisroriae* waarin hij beschrijft hoe een schip in een zware storm door de wind lucht in wordt geblazen en op de maan terechtkomt. In de negentiende eeuw publiceerde Jules Verne zijn boek *De la Terre à la Lune* (De reis naar de maan), waarin hij nauwkeurig omschrijft hoe een ruimtevaartuig wordt afgeschoten door een kanon en naar de maan reist. Frappant zijn de vele overeenkomsten tussen zijn boek en de Apollovluchten naar de maan die meer dan honderd jaar later plaatsvonden. Ook in de verhalen van de schrijver H.G. Wells, onder andere *War of the Worlds* en *The First Men on the Moon* spelen reizen door de ruimte een grote rol. De Sovjet-Unie ontwikkelde de eerste kunstmaan, Spoetnik T. In hetzelfde jaar bracht men het eerste levende wezen in de ruimte, het hondje Laika met de Spoetnik 2. Even later volgden de Amerikanen met de Explorer 1. De Amerikanen werkten veelal met aapjes. De reden hiervoor was dat beide naties met die proefdieren veel ervaring hadden. Hun fysieke reacties op het ruimteleven zouden daardoor snel herkenbaar zijn. Nadat de Amerikanen hun eerste aap de ruimte inschoten, hadden de Russen ook de volgende primeur: Joeri Gagarin werd in een baan om de aarde gebracht, de eerste mens in de ruimte. Als reactie hierop en om de achterstand in de ruimtewedloop in te halen beloofde de Amerikaanse president John F. Kennedy dat er voor het einde van het decennium een Amerikaan op de maan zou lopen. Door de NASA werd toen het voorbereidende Gemini gestart, waarmee ervaring werd opgedaan met bemande ruimtevluchten en zelfs met ruimtewandelingen. Het Project Apollo resulteerde acht jaar later uiteindelijk in wat Kennedy had beloofd: Neil Armstrong zette als eerste mens voet op de maan. Hij sprak de befaamde woorden: ‘That’s one small step for a man, but a giant leap for mankind’. (Dit is een kleine stap voor een mens, maar een grote sprong voor de mensheid). Ondertussen zat de Russische ruimtevaart ook niet stil. Meer op de achtergrond waren de Russen bedreven geraakt in het in een baan om de aarde brengen van kunstmanen. Daarnaast begonnen zij met het sturen van sondes naar diverse bestemmingen in ons zonnestelsel en bouwden zij ruimtestations op. De voorkeur van de Sovjets lag bij Venus, waar de Amerikanen gingen koersen op Mars. Na de val van de Sovjet-Unie en het einde van de

Koude Oorlog, kwam de staatsfinanciering van de Russische ruimtevaart op een veel lager niveau te liggen. Ook de Amerikaanse ruimtevaart kreeg ernstige klappen te verwerken: na het verongelukken van de Space Shuttle Challenger lag de bemande ruimtevaart in Amerika bijna drie jaar lang stil. NASA en Rosaviakosmos, de ruimtevaartorganisaties van de VS en van Rusland, besloten daarom tot een verregaande vorm van samenwerking. De Russen zouden hun langlopende project, het ruimtestation Mir, opgeven en samen met de Amerikanen en de Europese ruimtevaartorganisatie ESA werken aan het International Space Station ISS. Door de ramp met de Space Shuttle Columbia, inmiddels vijf jaar geleden, kwam de bevoorrading en de afbouw van de ISS in gevaar. Om te voorkomen dat er in de toekomst nog grotere problemen of financiële rampen ontstaan, moet de ISS absoluut afgebouwd worden. Nieuwe modules moeten met de Space Shuttle omhoog gebracht worden. De laatste tijd zijn er spanningen tussen NASA en Rosaviakosmos door het meenemen van ruimtetoeristen naar de ISS. Er zijn al vier toeristen op bezoek geweest in de ISS tegen betaling van 20 miljoen dollar. Tegelijkertijd doen de ruimtevaartorganisaties onafhankelijk van elkaar nog steeds onderzoek naar het leven in de ruimte. Hoewel het moeilijk te voorspellen is wat Rusland, Amerika en andere landen precies in petto hebben, is de verwachting dat de ruimtevaart in de toekomst vooral onbemand zal zijn. Op ruimtestations wordt veel onderzoek gedaan naar het leven in de ruimte en dat zal ook zeker zo blijven, maar hiervoor zijn mensen niet meer nodig. Satellieten en geavanceerde computerapparatuur leveren voldoende informatie. De mensen die wel de ruimte ingeschoten worden, zijn voornamelijk rijke toeristen zijn die wel eens wat anders willen dan de zoveelste zonvakantie op de Canarische eilanden. Al met al kan gezegd worden dat het reizen in de ruimte de mensheid al eeuwenlang bezighoudt. En zo lang een reisje naar de maan tot de verbeelding spreekt, zullen wetenschappers nieuwe technieken blijven ontwikkelen om de mysteries van het heelal te ontrafelen.

Fourth grade*Tekst 1:*

Jacqueline schreef op 27 september 2006 om 10:46:

Als klein meisje wilde ik ballerina worden, daarna stewardess en eenmaal puberend wilde ik popjournaliste worden. Lekker optredens van bands recenseren en zo. Maar ik heb voortijdig de school verlaten, dus toelating tot de school voor Journalistiek kon ik wel vergeten. Uiteindelijk is het na vele omzwervingen de ouderenzorg geworden. Ik heb eerst als ziekenverzorgende in verpleeghuizen gewerkt, maar moest hier helaas vanwege rugproblemen mee stoppen. Nu heb ik alweer acht jaar een leuke kantoorbaan bij een thuiszorginstelling, het coördineren van aanvullende diensten aan huis (kapper, pedicure, maaltijden, etc.) bij nog zelfstandig wonende ouderen.

*Tekst 2:***Stinkkaas, piepende krijtjes en papieren vliegtuigjes**

Voor de zestiende keer zijn vannacht de Ig Nobelprijzen uitgereikt, de alternatieve Nobelprijzen voor onderzoek dat je eerst aan het lachen maakt en daarna aan het denken zet. De Ig Nobelprijzen zijn een ludieke variant op de eerbiedwaardige Nobelprijzen, die deze week eveneens worden toebedeeld. Het brein achter de Ig Nobelprijzen is Marc Abrahams, hoofdredacteur van het cultijdschrift *Annois of Improbable Research* en evenals vorige jaren vannacht weer ceremoniemeester tijdens de uitreiking. Jaarlijks vist Abrahams tien merkwaardige studies uit de zee van wetenschappelijke publicaties die in zijn ogen een Ig Nobelprijs waard zijn. Zijn criterium? 'First it makes you laugh. Then it makes you think.' Ook Nederland was dit jaar van de partij. Vannacht stond de Wageningse malariaonderzoeker Bart Knols op het podium van het Sanders Theater in Harvard (VS) om daar onder luid gejoel van het te samen gekomen publiek de Ig Nobelprijs voor biologie in ontvangst te nemen. Het begon als een haast studentikoze associatie van twee Wageningse onderzoekers. Bij een nogal macaber experiment waarbij een proefpersoon naakt plaatsnam onder een klamboe waarin muggen werden losgelaten hadden de twee ontdekt dat Afrikaanse malariamuggen niet zomaar in het wilde weg steken. Ze staken bij voorkeur in het gebied rond de enkels en de voeten. Malariaonderzoekers Bart Knols en Ruurd de Jong associeerden vervolgens vrijelijk

van ‘voetengeur’ naar ‘tenenkaas’ naar ‘stinkkaas’, en lieten zich daarna door de plaatselijke kaasboer informeren over de meest onwelriekende kaasjes uit het assortiment. De Limburgse stinkkaas die ze uiteindelijk meenamen, bleek een uitstekend muggenlokmiddel. In 1996 publiceerden ze hun bevindingen in het medische vakblad *The Lancet*, onder de titel ‘On Human Odour, Malaria Mosquitoes and Limburger Cheese’. De Amerikaanse vogelonderzoeker Ivan Schwab kreeg de trofee omdat hij uitpuzzelde waarom een specht eigenlijk geen hoofdpijn krijgt bij zijn dagelijkse bezigheden. Hogesnelheidsfilmpjes tonen aan dat een gemiddelde specht met een vaartje van 25 kilometer per uur zijn snavel in een boomstam jast. Dat is zoiets als in volle vaart tegen een muur aan rennen. De hersenen van het dier blijken gevat te zijn in een soort schokdempers, zodat de ergste klappen worden opgevangen. Ook Howard Stapleton van het Amerikaanse bedrijfje Compound Security mag zich sinds vandaag eveneens de trotse eigenaar van een Ig Nobelprijs noemen. Stapleton ontwikkelde de ‘puberverjager’, een elektronisch apparaatje dat een irritant hoge pieptoon voortbrengt die alleen door jonge oren te horen is. Zo’n ‘teenage repellent’ verjaagt de buurtjeugd als die toevallig jouw portiek heeft uitgekozen om een avondje te hangen. Het geluid is overigens ook als ringtone te downloaden op een mobiele telefoon. Handig voor diezelfde pubers, die zo in de klas gebeld kunnen worden zonder dat de leraar het doorheeft. Het spaghettionderzoek waarmee twee Fransen vorig jaar het tijdschrift *Physical Review Letters* haalden, was vannacht eveneens goed voor een Ig Nobelprijs. De Fransen onderzochten hoe het komt dat gedroogde spaghetti nooit precies in twee stukken breekt, maar altijd in drie of meer. In de categorie ‘etenswaren’ viel ook het bekroonde Spaanse onderzoek naar de snelheid van ultrasoon geluid in cheddarkaas, gemeten bij verschillende temperaturen. En dan waren er nog de twee Australische onderzoekers die uitzochten hoeveel foto’s je van een groep mensen moet maken om er zeker van te zijn dat niemand zijn ogen per ongeluk dicht heeft, Lynn Halper die uitzocht waarom mensen zo slecht tegen het geluid van een nagel of een krijtje over een schoolbord kunnen, en last but not least: de Ig Nobelprijs voor geneeskunde ging dit jaar naar Francis Fesmier en collega’s, die een inventieve manier bedachten om de hik te verhelpen: een vinger in de anus. Geen flauwekul, maar, net als de negen andere bekroonde onderzoeken, gepubliceerd in een heus wetenschappelijk tijdschrift, *Annals of Emergency Medicine* in dit geval, onder de titel ‘Termination of intractable hiccups with digital rectal massage’. Tijdens een carnavaleske ceremonie in een uitverkocht Sanders Theater — waar de winnaars overigens op eigen kosten naar toe moeten reizen — werden de prijzen uitgereikt door échte Nobelprijswinnaars. Dit jaar waren dat onder meer Frank Wilczek (natuurkunde, 2004), William Lipscomb (scheikunde, 1976) en Ron Glauber

(natuurkunde, 2005). Glaubers Nobelprijs weerhield hem er evenwel niet van zijn vaste taak tijdens de ceremonie ter hand te nemen: het podium ontdoen van de papieren vliegtuigjes die het publiek traditie getrouw naar voren werpt. De winnaars - negen van de tien waren aanwezig - kregen elk een minuut de tijd om een speech te houden. Aangezien wetenschappers niet direct bekend staan om hun bondigheid, werden ze aan de tijd gehouden door Miss Sweetie Poo, een opdringerig achtjarig meisje dat, al roepend: 'Please stop, I'm bored' tergend langzaam dichterbij kwam. Al met al was de uitreiking van de vreemdste prijs die de wetenschap kent, ook dit jaar weer een hilarische gebeurtenis.

Tekst 3:

Zonder tbs-verlof wordt Nederland minder veilig

De tbs-klinieken in Nederland hebben de laatste jaren flink aan de weg getimmerd. Risico-inschattingen worden gedaan met wetenschappelijk onderbouwde instrumenten waarvan is aangetoond dat zij betrouwbare voorspellingen opleveren. Ook is de aanvraagprocedure voor verlof flink aangescherpt. De resultaten van die verbeterde aanpak zijn duidelijk zichtbaar. Het aantal patiënten dat in de eerste tien jaar na de tbs opnieuw een ernstig delict pleegt, daalde in vijf jaar van 49 naar 34 procent. Terwijl 25 procent van de langgestrafte gedetineerden binnen zes jaar na ontslag uit de gevangenis terugvalt met een zeer ernstig delict, geldt dat maar voor 9 procent van de ex tbs'ers. Van de ruim 300.000 verlofbewegingen in de periode 2000-2005 leidden er 488 (dat is nog geen 0,2 procent) tot een onttrekking. Daarbij kwam het 22 keer tot een (niet altijd ernstig) delict. De kans dat een verlofbeweging uit een kliniek eindigt met een recidive is daarmee 0,007 procent. 's Morgens in een auto stappen is vele malen gevaarlijker voor een individuele burger dan een tbs'er op verlof tegenkomen. Toch krijg je als mediaconsument niet bepaald de indruk dat het goed gaat met de tbs. Als je afgaat op de kranten, lijkt het alsof tbs'ers op verlof een van de grootste bedreigingen zijn voor onze maatschappelijke veiligheid. Inmiddels is de politieke en publicitaire obsessie met de tbs zo ver uit de hand gelopen, dat iedere ontvluchting voorpaginanieuws is. Dat is om meerdere redenen bijzonder tragisch. Allereerst omdat volledig buiten proporties geraakte onveiligheidsgevoelens daarmee alleen maar verder escaleren. De beleving dat er van alles mis is met de tbs, volgt immers niet uit de feitelijke prestaties van de tbs klinieken, maar uit de keuze van journalisten en politici om van de tbs voorpaginanieuws en campagnevoer te maken. Dat deze keuze weinig rationeel is, blijkt al uit

het opmerkelijke feit dat de recidive van reguliere ex-gedetineerden volledig buiten beeld blijft in de media en de Tweede Kamer. Ook maakt de klopjacht op ongeoorloofd afwezige tbs'ers hun delictgevaar eerder groter dan kleiner: een kat in het nauw maakt immers rare sprongen. Een repressiever klimaat, met veel minder vanzelfsprekend contact met de buitenwereld, zal trouwens de neiging tot ontvluchting eerder voeden dan afremmen. Repressie is een prima manier om desperado's te kweken. Tragisch is bovendien dat ieder incident de politiek tot weer nieuwe maatregelen drijft die de uitvoering van de tbs zoals zij bedoeld is, verder hinderen. Tornen aan verlofmogelijkheden betekent het verlies van een cruciaal behandelinstrument: pas in interactie met de buitenwereld wordt de behandeling effectief. Ook is aangetoond dat de recidivekans groeit naarmate de overgang van binnen naar buiten abrupter is doordat er weinig of geen kans is geweest te oefenen in verlof. Tornen aan verlofmogelijkheden betekent verder het nodeloos binnenhouden van de meerderheid van patiënten die helemaal geen gevaar vormen. Tbs'ers zijn patiënten in een psychiatrisch ziekenhuis, die vanwege hun psychische stoornis niet gestraft worden voor hun delict, maar behandeld worden om herhaling te voorkomen. Toch bevinden zij zich tegenwoordig in de weinig benijdenswaardige positie dat hun wel en wee eerder door politieke dan door behandelinhoudelijke overwegingen wordt bepaald. Natuurlijk is het verder terugdringen van ongeoorloofde afwezigheid belangrijk en natuurlijk moet al het mogelijke worden gedaan om de kans op recidive te verkleinen. Dat zijn precies de doelen die werkers in de tbs dagelijks nastreven. Maar wie eist dat de uitvoering van de maatregel geen enkel incident met zich mee brengt, vraagt iets onmogelijks. Het is niet voor niets dat Kamerleden die zich opwinden over een ontvluchting, het antwoord schuldig blijven als het over alternatieven gaat. Als de politiek niet met de onvermijdelijke risico's van de tbs kan leven, moet zij niet blijven steken in halfslachtige maatregelen die van de tbs stiekem een onrechtmatige levenslange gevangenisstraf maken. Een dergelijk misbruik van de psychiatrie past niet een democratische rechtstaat. Schaf de tbs dan gewoon af. Stuur mensen die nu de tbs ingaan, dan maar naar de gevangenis. Helaas zullen ze dan wel na het uitzitten van hun straf tweemaal zo vaak ernstig recidiveren als na tbs. Nu al is het een pijnlijke ironie dat alle korte termijn maatregelen die in naam van de veiligheid door het parlement worden afgedwongen, de kansen op een succesvolle behandeling verkleinen en op de langere duur waarschijnlijk tot meer recidiveslachtoffers zullen leiden. Als we dat tij niet keren, wordt Nederland gevaarlijker in plaats van veiliger. Tbs en tbs-verlof moeten met rust gelaten worden.