Validating a tangible test method for basic language skills in primary education

Joyce Krijgsman

ID 486299

March, 2011

Supervisors:

Willem Fontijn

Hanneke van Mier

serious toys®

Maastricht University

Overview

Abstract

Testing of children in primary education is a necessary, but often a problematic thing. In the Netherlands, every school is required by law to follow the progress of their pupils. However, regular tests on paper or computer often cause a lot of a hassle for the school, the teachers and the children, especially in the lowest grades of primary education. When making a test on paper, typically all children are tested at the same time, for which they need to sit behind a desk and have to work in silence for quite some time. This can be a stressful occasion for such young children who are not used to work in such circumstances. Modern technology makes it possible to let the children make the test on a computer as well. However, interacting with a computer is often done via a mouse or a keyboard, which is unsuitable for young children, because their fine motor skills are still developing (Hourcade et al., 2004). In a search for a better solution, a new test was developed, which is called the TikToets. The TikToets is a Dutch language test for children in grade 2 of primary education, which is done on a TagTiles board. In the present research the main question we asked ourselves was: Does this test on TagTiles constitute a suitable way of testing of children in grade 2? Thirty-one primary school children participated in this study and the usability of the TikToets and the intrinsic motivation of the children for making the test were investigated. Overall we found that the TikToets is very usable for children in grade 2 and the children were highly motivated to make the test. They were very willing to try their best to succeed in the task and also enjoyed making the TikToets because they had the experience of playing a game. These findings are important because we can assume that if the children are highly intrinsically motivated to make the TikToets, they will not find it very stressful to make the TikToets. Therefore we can conclude that a test on TagTiles offers a good alternative for testing children in grade 2 of the elementary school.

Introduction

In this report we describe the validation of a new assessment tool for TagTiles within the company Serious Toys. Serious Toys is a spin out of Royal Philips Electronics and became a company on its own in October 2008. It develops tangible, electronic toys, games and learning aids that offer a combination of entertainment and personalized learning.

TagTiles is the first product Serious Toys has brought to the market. TagTiles is an A3 size electronic board to be used with tangible learning materials that can be used by children alone or together, to play games addressing language, math and cognitive skills (see figure 1). In cooperation with leading educational publishers, Serious Toys has already developed a large variety of games of different levels for children in primary school education and more games are constantly being developed.

Figure 1. TagTiles.

The new tool for TagTiles we have investigated in this study was made in cooperation with KlasseWerk. Within the educational sector, KlasseWerk is for instance known from the learning method 'Onderbouwd' for the youngest children in primary education. Their new tool for TagTiles provides a Dutch language test for this age group. This test is called TikToets. TikToets is linked to the method 'Onderbouwd', but is not part of it. Therefore schools that do not work with 'Onderbouwd' can still use the TikToets.

Tests like the one from KlasseWerk are important since every school is required by law to follow the progress of their pupils. Since 1993 the Dutch Ministry of Education provides core objectives for all primary schools in The Netherlands. These core objectives offer the schools targets to focus on with the development of their pupils. To assist schools in translating these core objectives into education, 'Stichting Leerplan Ontwikkeling' has developed teaching guidelines and sub objectives. Every school is free to choose which tests it wants to use to monitor whether their pupils meet the core- and sub objectives ("Vraag en antwoord: Wat zijn kerndoelen voor het basisonderwijs?", 2010).

Typically the pupils will make such a test on paper. In most cases the whole class will make this test at the same time. Therefore all children need to sit behind a desk and have to work in silence for

quite some time. Many teachers in the lowest grades of primary education complain about this way of testing. They report that making these tests can be quite stressful for such young children, because they are not used working in such circumstances yet. This can have a negative influence on the scores of the test for some children, while others remain unaffected. (Queens, 2010). Therefore the test results are less reliable.

Modern technology makes it possible to let the children make the test on a computer. By using a computer some of the aforementioned problems can be resolved. However, interacting with a computer is often done via a mouse or a keyboard. This is unsuitable for young children, because their fine motor skills are still developing. In a study by Hourcade, Bederson, Druin and Guimbretiere (2004) the performance of both young children and adults using mice was compared. Thirteen four year old children, thirteen five year old children and thirteen adults in the age of 19 till 22 participated in the study. They were asked to do a simple point-and-click task, which consists of moving a cursor from a home area towards a target and then click on the target. The paths taken by the different age groups are shown in figure 2.

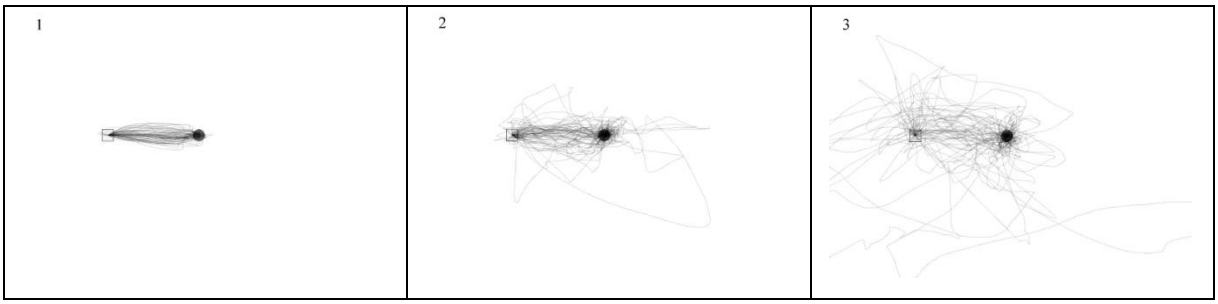


Figure 2. (1) Paths taken by all adult participants. (2) Paths taken by all five year old participants. (3) Paths taken by all four year old participants.

The paths taken by four and five year olds show a completely different picture than those of the adults. The paths of the children are all over the screen and look jagged, while the paths of the adults look much smoother. In addition, there is a halo of activity around the target that is not present in the paths taken by adults. Further analysis of the data showed that children were a lot slower than adults and more importantly, they were significantly less accurate and had significantly less control over the mouse than adults did. These results are important, because as Hourcade et al. (2004) conclude, this 'provides evidence of the need to take into account children's pointing abilities when designing software for them'. As a consequence, the performance of a child on a computer based test is typically determined by a large part by the computer skills of the child, which is undesirable.

Another drawback of the computer is the cognitive load of interpreting the 2D interface. The severity of this cognitive load varies in an undetermined way over children, polluting the test results (Verhaeg, Resing, Jacobs & Fontijn, 2009).

Taking this in consideration, the need for a better solution is apparent. The TikToets was developed to create a more suitable test for young children. A test on TagTiles has several major advantages over both the traditional written test as well as the regular computer based test, not only for the pupils but most certainly for the school and the teachers as well. First of all, the TikToets is designed in such a way that the child can work fully independently. TagTiles will guide the child through the test. While one child is making the test, the rest of the class can continue with their normal program. This makes reorganisation of the classroom unnecessary and the test can be part of the normal classroom routine. Another advantage is that TagTiles automatically records all the scores of the pupils. This means that the teacher does not have to correct the tests of the pupils by hand. When the teacher connects TagTiles with a computer, all the results can be seen on the computer screen. Later on we will explain how this works.

Such an overview of the scores after making the test provides the teacher with a good insight in which topics the children still find difficult. This aids the teacher in selecting suitable tasks that offer more practice on those specific topics. After a few months, the teacher might want to test the child again. A major advantage of the TikToets is that it never offers the exact same test twice. For every single task there are two different questions and TagTiles selects randomly one of them. This makes the test unpredictable, which reduces the chance on training effects. In this way the teacher can let the child make the test as a whole or in parts, multiple times within a school year.

Other major benefits that a test on TagTiles has centre on the children. They can make the TikToets completely by themselves and therefore they can work at their own speed. This contributes to a decrease of stress for having to make the test. Another advantage is the way the children have to indicate their answers to the questions, since they only need to tap the pictures on the board with one of the cubes which are provided with TagTiles. This is a lot easier than writing with a pencil or clicking with a mouse.

So far we described the many benefits a test for TagTiles would have. Table 1 on the next page gives a summary of these benefits.

Advantages of the TikToets

• Child works independently, individual testing
• No impact on regular classroom routine
• Keeps track of the scores and progression
• Provides a clear overview of the scores
• Easy variation of the test items, enabling multiple retest without training effect
• Test can be easily segmented
• No computer skills required (neither specific motor skills, nor specific cognitive skills) hence more reliable then a computer based test
• Less stressful hence more reliable results than a written test

Table 1. Summary of the advantages of the TikToets.

The aim of the first TikToets is that it provides a test for pupils to assess whether they meet the sub objectives for language in the second grade of primary education in The Netherlands. The TikToets is done by putting the right foils on the TagTiles board (see figure 3). The test consists of eight different foils: one log in foil and seven foils with assignments. By using the blue log in foil, the child can log in with his / her own picture in order to begin the test. Lights and sounds that come from the TagTiles board guide the child through the TikToets, through which the child can make the test fully independently. By making the TikToets, the child becomes tested at seven different linguistic categories, namely, Letter Knowledge, Rhymes, Sound Groups, Auditory Synthesis, Critical Listening, Passive Vocabulary and Visual Discrimination. For each new category, the child puts a new foil with assignments on the TagTiles board. TagTiles then selects randomly the questions that are going to be asked. Indicating the answers to the questions is done by tapping the answer with one of the cubes which are provided with TagTiles, so using a pencil or a mouse is not needed. Because every child logs in his / her own picture, the teacher gets a good overview of the performance of each child. When the teacher puts the SD card, on which TagTiles saves the results, in a PC or laptop, an overview of all results appears on the screen. The program provides four different ways of presenting the scores (figure 4). First there is a table that gives an overview of the scores for all children, then there is a bar chart that presents the scores of each child for the whole test, separated per test category and there

Figure 3. Foils of the TikToets.

is a bar chart that presents the scores within each test category for all children. Furthermore, an overview of the personal progression of a child can be shown for each category, which turns out to be very helpful in monitoring the meeting of core- and sub objectives. These four manners of score presentation should provide the school with a comprehensive overview of the development of their pupils.

Figure 4. Examples of different set ups of scores overviews.

At this stage of our research, it is of prime importance to find out whether young children can handle this way of testing. Therefore, the main question we ask ourselves is: Does this test on TagTiles constitute a suitable way of testing of children in grade 2?

Pilot study

We started our research with a pilot study with five first grade and five second grade children. They were asked to make the TikToets to determine whether there were any bugs in the program. We also wanted to see the children's reaction on the test. While observing the children working with the TikToets, we filled in a questionnaire on usability to get answers on questions like 'Does the child know how to start the program?' and 'Does the child understand the meaning of all pictures?'. The TikToets is developed for children of grade 2 and as expected the test itself was too difficult for first grade children. However, almost all first grade children knew how to complete the required actions and therefore we can deduce that almost all second grade children will be able to work in this manner. Based on the answers on the questionnaire and additional notes that were made during the test, the TikToets was improved for a second testing round.

Method

Participants

We recruited the participants via two elementary schools in ‘s-Hertogenbosch and surroundings. Thirty-one Dutch children, twelve boys and nineteen girls, participated in our study. They were all 4-6 years old.

Materials

For this study we used one TagTiles board with coloured cubes plus a speaker for sound. The TikToets consist of one log in foil and six foils with assignments, because the seventh foil, Visual Discrimination, was not ready yet by the time of testing. A separate foil was used to let the children fill in a questionnaire on intrinsic motivation on TagTiles.

Measures

The measures that were applied in the test came from observations on the basis of a questionnaire of usability filled in by the experimenter and questionnaire of intrinsic motivation filled in by the participants.

Questionnaire on usability

While observing how the children took the TikToets, the experimenter filled in the usablitity questionnaire. This questionnaire consisted of twelve questions that served as a guideline for the observations and was specifically developed for this study. It contained questions like ‘Begrijpt het kind hoe het zijn/haar antwoord moet bevestigen?’ (Does the child understand how to confirm his/her answer?) and ‘Begrijpt het kind de overgang naar de volgende folie?’ (Does the child understand the transition to the next foil?). These questions were answered with yes or no. Also the time in minutes the child spend making the test was measured and filled in on this form. The complete questionnaire can be found in appendix A.

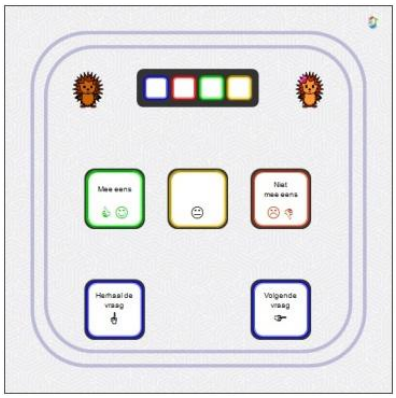
Questionnaire on intrinsic motivation

This survey has already been used for previous studies on TagTiles. For example the section Developmental and Educational Psychology of Leiden University used this questionnaire for their research on games for TagTiles. It is based on the Intrinsic Motivation Inventory (IMI) ("Intrinsic Motivation Inventory", 2010). Although the overall questionnaire is called the Intrinsic Motivation Inventory, it includes only one subscale that actually assesses intrinsic motivation. The original questionnaire includes seven subscales that were developed for assessing participants' subjective experience related to target activities in laboratory experiments. McAuley, Duncan, and Tammen (1989) did a study to examine the validity of the IMI and found strong support for its validity.

For research on TagTiles three of the IMI subscales were considered to be usable and relevant (Philips, 2006). IMI has been tested for separate use of subscales and it was found that this should not influence validity or reliability of the questionnaire. The following three subscales of the IMI were included in the questionnaire:

1. Interest and Enjoyment: this subscale is considered to be a self-report measure of intrinsic motivation.
2. Perceived Competence: this subscale is used to predict maintained behaviour change, effective performance, and internalization of ambient values (Williams, Freedman & Deci, 1998).
3. Effort and Importance: effort is a separate variable that is relevant to some motivation questions, and is therefore used in addition to the other two subscales.

The original IMI is in English and was developed for adults. This required translation of the items into Dutch that would be understandable for children. For example, the word 'competent' was translated as 'het goed doen' (being good at). The translation of the items was done according to a scheme (see appendix B), to make sure that the items were translated consistently (Philips, 2006).
The original questionnaire uses a seven point scale that ranges from 'not at all true' (1) to 'somewhat true' (3) to 'very true' (7). To make the scale more suitable for young children, the seven point scale was turned into a three point scale consisting of 'smileys' (☺ = agree, 😐 = do not agree, do not disagree, ☹ = disagree).
To make the questionnaire more relevant for the purpose of assessing aspects of TagTiles, some items were added to the subscales, for example 'Van tevoren leek het me geen leuk spel.' ('Beforehand, this game seemed not very enjoyable.'). In total the adjusted questionnaire included sixteen items. The complete questionnaire can be found in appendix C.



The questionnaire was filled out on TagTiles, so the answers could be digitally stored (Figure 5). Thereby, filling out the questionnaire on TagTiles is for children a relaxed way of working.

Figure 5. Foil of the questionnaire.

Procedure

In advance the parents of the children had received a letter with information about the test and had signed for agreement of their child's participation. The participants were picked up in their classroom one by one and taken to a separate room where the test took place. The child was asked to take a seat behind the TagTiles board and it was explained to the child that it was here to try a brand new 'werkje' (educational task). Sometimes we also talked about 'game' as the Dutch words 'werkje' and 'spelletje' are often used for the same in a conversation with young children. Also in the motivation questionnaire the activity was called a 'game'.

When the child was ready to begin, the login foil was put on the TagTiles board which automatically started the TikToets. This also meant that the child had seen how one should put a foil on the TagTiles board. Next the TikToets asked the child to tap on his or her log in picture. The experimenter told the child which picture he or she should pick and asked if he or she agreed with this picture. This was to assess whether there were any pictures that had a negative association for the children. After the child was logged in, the TikToets started. While the child was working on the test, the experimenter filled in the usability questionnaire and made pictures. After the child had finished the TikToets and had filled in the questionnaire of motivation, it was brought back to the classroom and the next participant was picked up.

Analytic strategy

The following statistical analysis was used. We started with recoding the negative items of the motivation questionnaire, so that they counted the same way as the rest of the items. This was done for the items 3, 5, 6, 8, 10, 11 and 15 (see appendix C). The usability questionnaire did not contain

reversed items. Since three IMI subscales were included in the motivation questionnaire, principal axis factoring with Oblimin rotation was conducted to examine whether these subscales were actually present. Then we performed a reliability analysis on that same questionnaire in order to find out how strong the individual items of each subscale would correlate to each other. The next step was to compute the frequencies of the answers to the usability questionnaire and the motivation questionnaire. After that we compared these scores for boys and girls, the different schools and the different age groups by using a one-way ANOVA. Finally we determined the correlation between the time the children spent making the TikToets and the results of the two questionnaires and their subscales.

Figure 6. Children taking the TikToets.

Results

Reliability and factor analysis

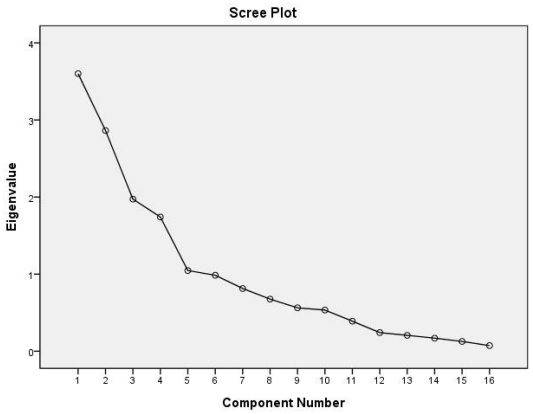


Figure 7. Scree plot of the Eigenvalues.

First the reliability of the questionnaire of intrinsic motivation was checked. The test in SPSS showed that the reliability of the subscales was not very high. For the three subscales we found respectively α = .449, α = .567 and α = .342. A factor analysis showed the questionnaire might even contain five subscales instead of three, though the fifth factor has an Eigenvalue that is only just over 1 (see figure 7).

Frequencies

Figure 8 gives an overview of to what extent the questions of the usability questionnaire were answered positively. Overall we can conclude that the outcome of this questionnaire is really positive. We can see that all children (100%) know how to use the cube to 'tap' their answers (item 3), understand the meaning of all the pictures (item 8) and understand how to continue to the next foil (item 9). Also very important is that most children (93,5%) know without further explanation how to start the TikToets (item 2) and how they should confirm their answers (83,9%) (item 7). However, fewer children (48,4%) know how they can ask for repetition of the question (item 6) and 9,7% of the children held the suggested sequence (item 10).

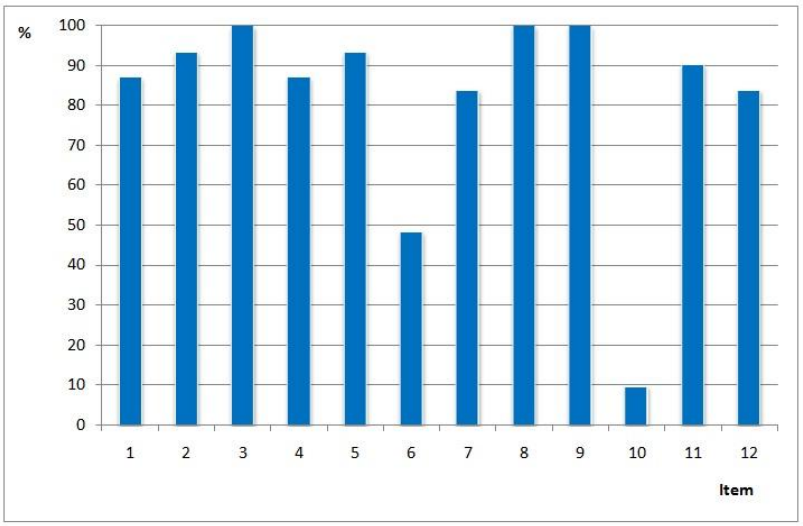


Figure 8. Frequencies of the positive answers of questionnaire of usability.

Figure 9 gives an overview of the answers given by the children to the questionnaire of intrinsic motivation. Because of the recoding of the negative items, the bar chart of figure 9 us all positive answers indicated in blue, the negative answers are indicated in red and yellow is neither positive nor negative. Overall, we can see that the children are quite positive about working with the TikToets.

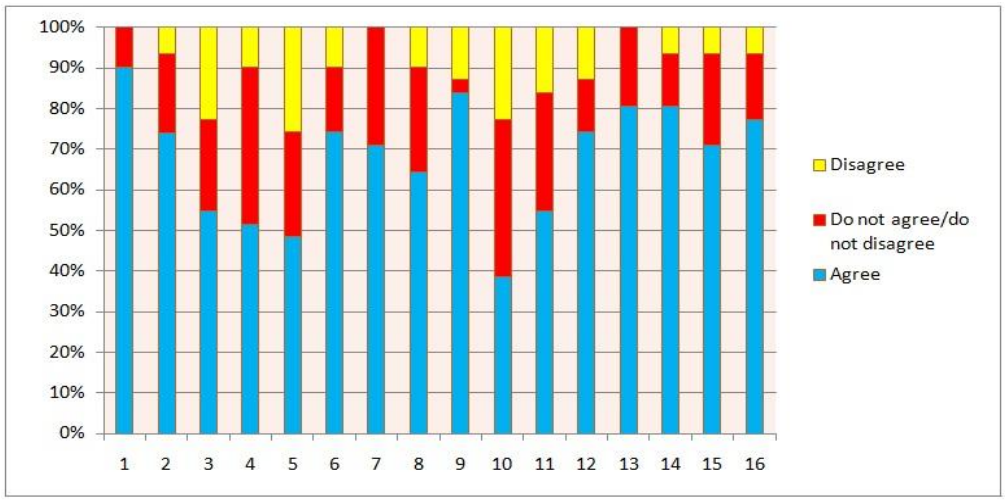


Figure 9. Frequencies of the answers of questionnaire of intrinsic motivation.

Within the category interest and enjoyment (item 1-9) 90,3% of the children answer they really enjoyed playing the game (item 1) and 83,9% would like to play the game more often (item 9). However, when asked if the children find the game boring (item 3) only 54,8% of them disagreed. 22,6% agreed in that and the other 22,6% had no opinion. Furthermore, when was asked if 'once you're in this game, you automatically want to continue' (Als je eenmaal bezig bent met dit spel, wil je er vanzelf mee doorgaan) only 51,6% agreed, 9,7% disagreed and the other 38,7% had no opinion. Another item that stands out is the question if the children could stay focused during the game (item 5). Only 48,4% agreed, 25,8% had no opinion and 25,8% disagreed.

Within the category perceived competence (item 10-13) we noticed that when was asked 'I am not so good at playing this game' (Ik ben niet zo goed in dit spel) (item 10) 38,7% of the children answered they 'do not know' or had no opinion, 38,7% disagreed and 22,6% agreed. Though, when they had to answer 'I am quite good at playing this game' (Ik ben best goed in dit spel) (item 13) 80,6% agreed and the remaining 19,4% did not agree neither disagree. Another remarkable finding was the answers given on 'during the game I did many things wrong' (Tijdens het spel deed ik veel dingen verkeerd) (item 11). Here only 54,8% disagreed, while 29% had no opinion or did not know and 16,1% agreed.

When we look at the third category (item 14-16) we can see that the children were quite willing to put some effort in playing the game. For every item we see that at least over 70% of the children indicated they were willing to try their best to play the game (item 14-15) and they find it important to succeed in the game (item 16).

Analysis of variance

When we compare the answers of the two different questionnaires, looking at the different subscales and the time the children spent making the TikToets, we observe the following. For starters, there is no significant difference between boys and girls on both the usability questionnaire, $F(1, 26) = 1.65$, $p = .21$, and the motivation questionnaire, $F(1, 29) = .14$, $p = .71$. However, we did find significant differences between the two schools for intrinsic motivation, $F(1, 29) = 5.42$, $p = .027$. When we look at the different subscales we can see this difference is probably caused by interest and enjoyment, $F(1, 29) = 4.78$, $p = .037$, and effort and importance, $F(1, 29) = 4.57$, $p = .041$. We also found a small difference in how much time children of the different schools on average spent on making the test, although this was not significant, $F(1, 29) = 3.59$, $p = .068$. Furthermore, we found a clear difference in duration for the different age groups, $F(2, 28) = 11.28$, $p = .000$, however this was only due to one outlier who needed a lot more time to make the test than all the others (see figure 10). This was a child of 4 years old, but her teacher asked to test her as well, because next year she will skip grade 2 and move on to grade 3. In a second analysis we removed this outlier to see whether this would have any consequences for other results apart from duration. This was not the case, so we stuck to our first analysis.

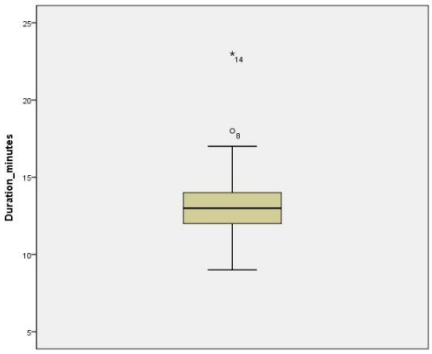


Figure 10. Boxplot of the time in minutes children spend on the TikToets.

Correlations

Finally we wanted to determine whether there was any correlation between how motivated the children were to play the game and how applicable the game and the board were for them. We found there was no correlation between the time the child spend on making the test and how motivated they were or how well they could handle the test on TagTiles. Also between the total score of the motivation questionnaire and the usability questionnaire we found no significant

correlation, $r(26) = .325$, $p > .05$. However, when we looked at the different subscales of the motivation questionnaire compared to usability, we did find a strong positive correlation, $r(26) = .535$, $p < .01$, between the usability of the test and the perceived competence of the children.

Discussion and conclusions

When we look at the outcome of our research we can conclude that a test on TagTiles offers a good alternative for testing children in grade 2 of the elementary school. For one, the TikToets proved to be very easy to use for the children. Especially the items 2, 3, 6, 7 and 9 in the questionnaire of usability were major contributors to this. On most of these items the TikToets scored very high. Only item 6 forms an exception on this. However, it should be noted that for all participants in this study it was the first time they ever worked with TagTiles. In practice, when a school decides to purchase a TagTiles board for the use of testing their pupils, it is quite likely that it will be used for other games as well. Then the children will be more used to working with TagTiles, which will make it even easier to work with it in general. Also item 10 scored very low, though this does not give any problems for the usability at all. The TikToets only gives a suggestion in which order of categories the children can make the test. The children are always free to choose the order they like best.

Looking at the intrinsic motivation of the children gives a good insight in how willing the children are to make the test. Although the adjusted questionnaire of intrinsic motivation has been earlier, in this study it was not tested as very reliable. We assume that if we would have had more participants, the reliability would have been a lot higher. Also the children in this study may have been too young to work with this particular questionnaire, as it was developed for children 7-10 years of age. Some words like ‘childish’ (kinderachtig) and ‘performances’ (prestaties) were quite hard to understand for children in this age. Also items that were formulated in the form of a denial caused some children difficulties. Factor analysis showed that especially item 5 ‘I totally could not focus on the game’ (Ik kon mijn aandacht helemaal niet bij het spel houden) did not correlate at all with any other item in the questionnaire. This was probably because the children just did not understand the meaning of the sentence.

Despite some items of the questionnaire might have been difficult to understand, overall the children rated doing the TikToets as very positive. Although it is meant as a test, the children have the experience as if they are playing a game. The game might not be very exciting, as indicated in question 3, the children do enjoy playing it. Because of the fact it is a test, the children will not get any feedback on their work. Still they rate their perceived competence as quite high. As mentioned before the children are also willing to try their best to succeed in the game. These findings are important because we can assume that if the children are highly intrinsically motivated to make the TikToets, they will not find it very stressful to make the TikToets.

Overall we can conclude that the TikToets offers a good alternative for language tests for young children that are normally made on paper or computer. As mentioned before, it has many benefits for the school, the teacher and the children. It could save a considerable amount of time, because it is not necessary to reorganize the classroom and the teacher immediately gets an overview of the results, so he or she does not have to spend time grading the work of the children. Other nice benefits are the advantages it has for the children. Our research showed that children know well how to work with TagTiles and the board guides them through the test easily. They only need to tap their answer, so they do not have to work with a pencil or a mouse. Also, they can work at their own speed and get the feeling they are playing a game. This makes that making a test becomes fun to do instead of a stressful occasion.

Future work

Though we now know that TagTiles can be used to test children in grade 2 of Dutch elementary schools. There is still a lot of research that can be done. KlasseWerk has designed the TikToets based on the knowledge and experience of its educational scientist. It is advised to do some more research on the validity of this Dutch language test, for example by comparing the results of the children on the TikToets with their results on a regular language test. However, in that case we have to be certain of the validity of the other test as well. Another important topic is that pupils are not only to be tested on their language skills. Also their knowledge of mathematics needs to be tested. By publishing a mathematics test together with the language test, KlasseWerk and Serious Toys can provide elementary schools in The Netherlands a complete and useful test for children in grade 2.

References

Hourcade, J. P., Bederson, B. B., Druin, A., & Guimbretiere, F. (2004). Differences in pointing task performance between preschool children and adults using mice, *ACM Transactions on Computer-Human Interaction, 11*(4), 357-386.

Intrinsic Motivation Inventory. (2010). Retrieved November 15, 2010, from http://www.psych.rochester.edu/SDT/measures/IMI_description.php

McAuley, E., Duncan, T., & Tammen, V. V. (1987). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. *Research Quarterly for Exercise and Sport, 60,* 48-58.

Philips Research Europe. (2006, September). *TagTiles, an educational game on a tangible interaction console* (Technical Note PR-TN 2006/00308). Eindhoven: Verhaegh, J.

Queens, T. (2010, Februari 18). Kleuters mogen geen kleuters meer zijn. *Dé Weekkrant.nl.* Retrieved December 28, 2010, from http://www.deweekkrant.nl/artikel/2010/februari/18/_kleuters_mogen_geen_kleuters_meer_zijn_3_foto_s/

Verhaeg, J., Resing, W. C. M., Jacobs, A. P. A., & Fontijn, W. F. J. (2009). Playing with blocks or with the computer? Solving complex visual-spatial reasoning tasks: Comparing children's performance on tangible and virtual puzzles. *Educational & Child Psychology, 26*(3), 18-29.

Vraag en antwoord: Wat zijn kerndoelen voor het basisonderwijs? (2010). Retrieved December 15, 2010, from http://www.rijksoverheid.nl/documenten-en-publicaties/vragen-en-antwoorden/wat-zijn-kerndoelen-voor-het-basisonderwijs.html

Williams, G. C., Freedman, Z. R., & Deci, E. L. (1998). Supporting autonomy to motivate glucose control in patients with diabetes. *Diabetes Care, 21,* 1644-1651.

Appendix A Questionnaire of usability

1. Vindt het kind zijn/haar inlogplaatje leuk? ja / nee

2. Kan het kind zelf met de taak starten? ja / nee

3. Begrijpt het kind het blokje bedoeld is om mee te 'tikken'? ja / nee

4. Begrijpt het kind de gestelde vragen? ja / nee

5. Kan het kind de test maken zonder dat er vragen herhaald worden? ja / nee

6. Begrijpt het kind hoe het om hulp/herhaling kan vragen? ja / nee

7. Begrijpt het kind hoe het zijn/haar antwoord moet bevestigen? ja / nee

8. Zijn alle plaatjes duidelijk voor het kind? ja / nee

9. Begrijpt het kind de overgang naar de volgende folie? ja / nee

10. Houdt het kind zich aan de gesuggereerde volgorde? ja / nee

11. Kan het kind goed geconcentreerd bij het spel blijven? ja / nee

12. Verliep de toets zonder dat er door mij ingegrepen moest worden? ja / nee

Hoe lang heeft het kind over de test gedaan? ……minuten.

Appendix B Translation scheme

Original IMI	**Translation into Dutch**
enjoyed	ik had plezier
this activity	dit spel
…I was thinking about how much I enjoyed it.	….merkte ik hoe leuk ik het vond.
I think I am pretty good	Ik denk dat ik best goed ben
I think I did pretty well	Ik denk dat ik best goed gespeeld heb
After working at this activity for a while, I felt pretty competent.	Terwijl ik het spel speelde, had ik het gevoel dat ik het best goed deed.

Appendix C Questionnaire of motivation

This questionnaire is based on the Intrinsic Motivation Inventory (IMI).

Interest and enjoyment

1. Ik had veel plezier in het spelen van dit spel

2. Het was leuk om dit spel te doen

3. Ik vond het spel saai

4. Als je eenmaal bezig bent met dit spel, wil je er vanzelf mee doorgaan

5. Ik kon mijn aandacht helemaal niet bij het spel houden

6. Ik vond het spel heel kinderachtig om mee te spelen

7. Ik vond het spel best leuk

8. Van tevoren leek het me geen leuk spel

9. Ik zou dit spel wel vaker willen spelen

Perceived competence

10. Ik ben niet zo goed in dit spel

11. Tijdens het spel deed ik veel dingen verkeerd

12. Ik ben tevreden met mijn prestaties in het spel

13. Ik ben best goed in dit spel

Effort and importance

14. Ik heb erg mijn best gedaan tijdens het spelen

15. Ik heb niet zo erg mijn best gedaan om goed te spelen

16. Het was belangrijk voor me om het spel goed te doen