

Negative impacts of high-stakes testing

Michaela Minarechová

Abstract: High-stakes testing is not a new phenomenon in education. It has become part of the education system in many countries. These tests affect the school systems, teachers, students, politicians and parents, whether that is in a positive or negative sense. High-stakes testing is associated with concepts such as a school's accountability, funding and parental choice of school. The study aims to explain high-stakes testing, how it is created and developed in selected countries and look at the negative impacts of tests on various actors within this relationship.

Keywords: High-stakes testing, negative impacts, students, teaching, learning, teachers, schools

Introduction

High-stakes testing usually involves national or state-wide standardized achievement tests (Marchant, 2004). At present, post-communist countries (for example Slovakia, the Czech Republic, Romania) are dealing with the introduction of high-stakes testing in education. These countries use national and international tests; for example, in Slovakia a national test called Monitor 9 is used in ninth grade, in the Czech Republic the first high-stakes testing is just beginning and will test pupils at level 5 and 9 this grade and in Romania the National Test was introduced in the school year 2008/2009 at the end of Grade 8 and also at the end of Grade 7 (Niculae & Doncu, 2007). In these countries (Slovakia, the Czech Republic and Romania) TIMSS, PIRLS and PISA international tests are also used. These post-communist countries only pay attention to the positive issues relating to the introduction of tests and we therefore focus on the negative consequences of high-stakes testing in this article.

High – Stakes Testing

High-stakes testing consists of a series of tests in which the results eval-

uate schools, teachers and pupils. These tests “are increasingly seen as a means of raising academic standards, holding educators and students accountable for meeting those standards, and boosting public confidence in schools” (Heubert & Hauser, 1999). The consequences of these tests affect all those involved in the education process. Every state in America uses a high-stakes test to comply with the No Child Left Behind (NCLB) policy. For example 25 states in America offer financial rewards to successful schools or schools that have shown improved test results. Similarly, in 25 states the government has the power to close, reconstitute, or take control of failed institutions (Amrein & Berliner, 2002). We shall explore the impacts of high-stakes testing on various factors associated with this theme. First, we will look at the mechanism of high-stakes testing and its application in selected countries.

Creation and expansion of high-stakes testing in England

High-stakes testing has been used for over twenty years in England. West (2010) explores the creation and development of high-stakes testing in England. High-stakes testing has been used since the introduction of the national curriculum and related assessment in 1988 during the reforms (Education Reform Act, an Act to amend the law relating to education, 29th July 1988). One of the key roles of testing was accountability. We might say that testing has become a tool of accountability for schools. The test results provide the means by which schools are responsible for the education they provide (West, 2010). Since 1980, education policy in England has been influenced by both the neoliberal and neoconservative approaches of the ‘New Right’ (Chitty, 1989, in West, 2010). The two key facets of the education reforms included, on one hand, the introduction of choice and diversity and, on the other hand, the national curriculum and testing programme. Greater importance was placed on parental “choice” in the Education Act of 1980 (An Act to amend the law relating to education, 3rd April 1980). The consequence of the reforms (Education Reform Act) was an increased diversity in the education on offer, a change in school funding and disclosure of test results. Since 1992, the results of these tests have been presented in “school performance tables” and “league tables”. Diversity has been encouraged through the creation of 15 City Technology Colleges and the specialist schools programme. By 1996 this programme had covered 181 technology colleges and language colleges (Woods et al., 1998). The technology colleges were funded by the government and private sponsors (Walford, 1997, in West, 2010) and state-maintained schools may opt out of local government control to become so-called “grant-maintained” schools. These schools are able to obtain finan-

cial grants from government agencies for their students (Funding Agency for Schools) (Whitty & Edwards, 1998, in Petrová, 2011). Changes resulting from the laws of 1980 and 1988 created a quasi market environment in the area of education. According to West (2010), on the supply side, there is a “battle” between the schools over customers (the students) and on the demand side there are the parents themselves. The parents, who were initially considered only to be “passive” recipients of education, found themselves in the role of “partners and customers” (DES, 1991; DCSF, 2008, in Wilkins, 2012). They found themselves in the role of consumers. Using information obtained from school inspections, parents can assess and compare schools using the data on performance which is provided and controlled by the government. They may also use other complementary/supporting forms of evaluation of schools. For example school brochures and websites, parent and teacher exchanges and school visits (Wilkins, 2012). The policy of introducing a quasi-market environment led to schools devoting more time to preparing students for testing rather than improving the quality of education itself, because high-stakes testing has become the main external indicator of school quality (Petrová, 2011).

The basic principles of this quasi-market environment were maintained by the Labour government in 1997. But the following year (i.e. in 1998) “the School Standards and Framework Act abolished grant-maintained status and from September 1999 schools were designed as one of three new types of school-community (formerly country), voluntary (aided or controlled) and foundation schools” (West & Pennell, 2002).

As mentioned above, the national curriculum and programme of evaluation was introduced with the 1998 reform. The national curriculum is determined on the basis of Key Stages. The results of tests in mathematics, English and science (at age 11) are published as “school performance tables” and “league tables” (West, 2010). In England, in addition to these tests there are two public examinations: the *General Certificate of Secondary Education* (GCSEs) and the *General Certificate of Education Advanced*, (GCEA – A Levels). The first set of these exams (GCSEs) is taken during the final year of compulsory secondary school education. At this point, students can either leave education and obtain a job or go on to further studies. These exams represent the end of Key Stage 4 – for students aged around 16 years. The second set of public exams is the General Certificate of Education Advanced, A Levels, which students usually take at the end of upper secondary education. As West (2010) notes, the tests undertaken at the age of 11 years and the public examinations sat between the ages of 16 and 18 years can be considered to be national high-stakes testing. These tests determine or help determine the futures of schools, pupils and teachers. A survey in 2005

(MORI, in West, 2010) found that 52% of parents agreed that a school's position in the performance tables would affect their decision in choosing a school for their child/children and 27% of parents reported that they were influenced by other factors such as the location of schools or the resources on offer.

Creation and expansion of high-stakes testing in USA

As in England, high-stakes testing has an established tradition in the United States. Since the 1980s high-stakes testing in United States can be referred to as being so-called "full-value". The introduction of high-stakes testing was preceded by decades of earlier attempts to improve education in the United States going back to 1960. Duncan and Stevens (2011) report on the origin and development of high-stakes testing in the United States. American educators were particularly interested in the question of improving the education system after the launch of Sputnik in 1957. This led to the adoption of an education law on elementary and secondary schools (Elementary and Secondary Education Act, ESEA) in 1965, and the subsequent development and introduction of minimum qualification tests. In this case, almost no impact was found to be had on the teachers and schools involved. However, the tests have been criticized "for being relatively easy to pass since they were concerned with minimums to be learned: the achievement floor and not the achievement ceiling" (Nichols & Berliner, 2007). Title I of this Act was one of the most important pieces of legislation in US history. The law was adopted mainly in relation to federal aid for schools with high concentrations of poor children and it remained the largest source of federal support for public schools. This education law on elementary and secondary schools has had an unquantifiable impact on the expansion of standardized testing in American schools (Sacks, 1999, in Duncan & Stevens, 2011). In 1969 the NAEP project (National Assessment of Educational Progress) was launched, which represented a major step toward national assessment (Grant, 2004). The NAEP is a mandatory survey designed to measure what students know and can do. Student assessments have been conducted periodically in reading, mathematics, science, writing, social studies, civics, US history, geography, citizenship, literature, music, career development, art, and computer competence (Johnson, 1992). The first national assessment was implemented in 1969 and students were assessed on science. Following the introduction of NAEP, "federal government support of national or large-scale testing grew" (Grant, 2004).

The limited academic approach of the next twenty years was published in the report of the Presidential Office in 1983 and was entitled *A Nation*

at Risk. This report pushed the topic of the quality of education onto the national political agenda, since the data apparently suggested that American education was not of the quality found in other countries. It implied that foreign competition had overtaken US economic superiority, since schools in other countries were better, produced better and more educated workers than did the American ones. The results of the standardized tests were an indication of the true status of national education (Duncan & Stevens, 2011). This situation resulted in a “test boom” and state governors tested all students at the district level with the aim of raising and improving the economy (Fiske 2008, in Duncan & Stevens, 2011). In September 1989 a summit (The National Education Summit in Charlottesville, VA) was held which focused on how to improve America’s educational performance. The summit was attended by President George Bush and the nation’s governors and it led to the adoption of six National Education Goals to be reached by the year 2000. The result of this summit was a renewed federal commitment to improving educational achievement and increasing the nation’s commitment to students, teachers, and schools. The issue of standardized tests was dealt with by The National Assessment Government Board created by Congress in 1988. The result of their investigation was the passing of a law entitled No Child Left Behind in 2002 signed by George W. Bush. This Act has four main aspects: flexible and local control, consolidation of parental control, and a focus on the operation of the system (U.S. Department of Education 2007, in Lobascher, 2011). It introduced the legal requirement of annual testing for students. It was the first time in US history that they began testing children in all public schools and the results were used as the unit of school assessment (Matthews, 2006, by Duncan & Stevens, 2011). The tests were referred to as “high” because they had the ability to change lives as a consequence of the implications associated with progress in test scores (Duncan & Stevens 2011).

Creation and expansion of high-stakes testing in Australia

The implementation of high-stakes testing in Australia occurred much later in comparison to the previous examples on creating and developing this mechanism in England and the United States.

The *American Educational Research Association* has developed a scale that can be used to help decide whether tests are defined as high-stakes testing. High-stakes tests have serious consequences for students and teachers. The NAPLAN (*National Assessment Program – Literacy and Numeracy*) has become a high-stakes test as a result of the website *MySchool*. This website was created in 2008 when Australian education ministers agreed

to greater transparency and accountability in schools. *MySchool* provides details on student performance in NAPLAN and basic descriptive and demographic data (Caldwell, 2011).

Lingard (2010, in Lobascher, 2011) referred to the introduction of high-stakes testing in Australia as a global policy convergence. In addition, testing is considered to be a key part of the globalized educational discourse. The educational policy of Australia belatedly followed in the direction of the leaders, such as England and the United States (Lobascher, 2011). The current educational reforms under way in Australia borrow from those adopted in the United States and England. Lingard (2010) believes that this situation is a result of policy loans and the flow of people between Australia and England, and to a lesser extent between Australia and the United States. However neither of these nations has education that could be considered best practice. The *Queensland Studies Authority* (QSA) criticized Australia's move towards high-stakes testing and invited politicians not to repeat the mistakes that have occurred in England; Lobascher (2010) gives some examples:

high-stakes testing produces 'defensive pedagogies'; the effects of the English policy regime are negative: deprofessionalisation of teachers with reductive effects on schools, which means it is difficult for them to achieve their policy goals.

In the United States policy (for example, the No Child Left Behind (2002 Act) has caused a decline in the quality of learning and teaching in US schools and a narrowing of the focus in schools serving disadvantaged students (Lobascher, 2010).

The Impact of High – Stakes Testing upon Students

High-stakes testing has already established a stable base in various countries. It has become a natural part of school life and for students it is part of everyday schooling. High-stakes testing has also become the subject of investigation amongst many researchers (West 2010; Redell 2010; Amrein & Berliner, 2002) who focus on the consequences of high-stakes testing in terms of accountability, teaching and learning. Teachers and students have learned to find their way around this area since the implementation of high-stakes testing in schools through the hidden curriculum. The school curriculum is generally understood to be an explicit, intentional, planned and formal training course with specific objectives. However, besides the formal school curriculum students also come into contact with the so-called unwrit-

ten curriculum, which is characterized by its informality and “unplanned” nature as opposed to the school curriculum. This unwritten process of education is known as the hidden curriculum. The hidden curriculum consists of values and inter-group relations, which enable students to integrate into society (Kentli, 2009). Many studies have been carried out (Dreeben, 1968; Lynch, 1989; Margolis, 2001, Giroux, 2001) which aim to examine the theory of the hidden curriculum found in Kentli’s investigation (2009). Teachers often teach subconsciously through the hidden curriculum. The hidden curriculum manifests itself in subtle features of the teaching: how the different activities, social interactions and relationships are structured. However, “supplying” the hidden curriculum is not a one-way process. It is a process that affects the things that are shaped and reconstructed through interactions (Renold, 2001, in Booher-Jennings, 2007). The relationship between gender and the hidden curriculum appears to be significant. Schools maintain and transmit “gender codes” through a formal structure as well as through informal practices (Arnot 1982, 2002; Dillabough, 2003, in Booher-Jennings, 2007). It is important to realize that the hidden curriculum does not exist *a priori*, but is created in the daily interaction occurring between students, their teachers and classmates (Booher-Jennings, 2007). Teachers respond to the success/failure of students in high-stakes testing on the basis of gender. Booher-Jennings (2007) found that teachers symbolically separate boys and girls who have failed the test. When a girl was unsuccessful in passing the test, teacher’s reaction was negative. They claimed that girls in general “do their best” and so encouraged them to resit. On the other hand, boys who failed were identified as “bad” or as not having tried to pass the test. Thus the experience of success and failure was very different for girls and for boys. Girls adopted the teacher’s criterion that they should “do their best” and applied it to their own behaviour. Girls who did not pass the test faced a double burden. In front of their classmates and parents, they portrayed themselves as studious girls who wanted to pass the test and be the best. In addition, the girls worked separately from the boys who were unsuccessful in the test. The boys were the target of criticism by the girls claiming that they were not committed enough to pass the test. The boys adopted a version of the successful ideology: they felt that they must work harder and not “goof off” (Booher-Jennings, 2007). The boys also received support for the legitimacy of their initiative. But three unsuccessful boys regarded this requirement as unfair. They would rather take the female version of a successful ideology and work within their powers and abilities. Also, in her research Booher-Jennings (2007) shows that girls and boys react differently to failing/passing the test. Girls who passed the test said that their parents were proud of them. Parents of girls who failed the

test preferred the principle of “doing one’s best”. The boys were afraid that if they were unsuccessful in the test their parents would be angry at them. The girls were afraid that their parents would be disappointed. In schools the idea that the staff and the parents know what is the best for children is widespread: girls need more confidence to be able to pass the exams and boys more self-discipline (Booher-Jennings, 2007).

High-stakes testing is associated with various consequences which may include “retention of students in grade before tests”, and suspension, expulsion or reclassification of students prior to testing (Amrein & Berliner, 2002). Bowie (2002, in Marchant, 2004) described the case of students being retained in grade. In this case, over 20,000 elementary and middle school students in Baltimore had to repeat the grade. Marchant (2004) added that this affects the lives of all the students and that they “feel the effects of high-stakes testing”.

Some students are simply expelled due to poor marks and in the interest of school ratings. Students are expelled from school before or during high-stakes testing. In this way schools artificially raise or rather influence their position in the performance tables. An interview with one principal showed that students who tend to score low are expelled from school before testing because they are not “ready to test” (Kelleher, 1999, in Amrein & Berliner, 2002). Other research also found that less successful students were released from school on the day of testing or they were sent on a trip so as to be excused from the test (Haladyna et al., 1991, in Amrein & Berliner, 2002).

In some countries, students with limited language skills (*Limited English Proficient, LEP*) are exempted from high-stakes testing. Language-minority students represent a threat to schools because their “performances” threaten the future of the school. Teachers were therefore familiar with various measures so that LEP students as well as language-minority students were exempted from testing for example (Downs, 2000; Haney, 2000, in Amrein & Berliner, 2002). Students with low test results who are Hispanic but speak fluent English are exempted from high-stakes testing simply on the basis of their Hispanic surnames. In Louisiana, parents requested the *Office for Civil Rights* to investigate why almost half of the students from poor and minority districts had failed the state test, even after taking it for a second time (Sadker & Zittleman, 2004). Another example is provided by the state of Georgia, where two out of three students from low-income families failed the mathematics, English and reading tests, while all the students from wealthy families were successful. Moreover, these students even exceeded standards. These differences are also visible within the middle class. Almost half of the students in Ohio from families with incomes below \$20,000 did not pass tests. By contrast, 80% of students from families with

incomes over \$30,000 did (Amrein & Berliner, 2002; Heubert, 2002/2003; Neill, 2003, in Sadker & Zittleman, 2004).

Greater attention is given to so-called “borderline” students. This group consists of students who are on the border of passing/failing the test. These students have an enormous impact on the school or district score: whether they fail or pass the tests. Researchers who conducted the recent *National Science Foundation* study interviewed teachers who openly admitted that they focused more on the borderline students than the students who would certainly have failed the test. Borderline students were put together in one group and the teachers gave them extra lessons to prepare for the tests (Madaus et al., 1992, in Amrein & Berliner, 2002). In another study these students were labelled as “bubble” students. Teachers said they would not spend time with children who would never pass the test. Thus the school day for “bubble students” was altered so that they could continue practising for the test until tests were taken. By contrast, the weakest and the cleverest students had a traditional school day (McNeil, 2000).

Disadvantaged students may appear in other statistics such as those on race or on material disparities. Gifted and talented students are also negatively affected by high-stakes testing, but this manifests itself in other ways than it does in the minority and poorer students. Gifted and talented students are often bored in class or feel disadvantaged. Many students feel frustration and anger because of the slow pace of teaching, the disproportionate amount of time spent on test preparation and the constant repetition of basic concepts. Some students simply stopped paying attention during lessons or were engaged in various non-academic activities such as independently reading their own texts, resting or chatting with their friends as they waited for teachers who were focusing their attention on the “slower” classmates (Moon et. al., 2007).

The interests of students are in conflict with those of the school, so attention is primarily focused on schools that adopt strategies that are in the interests of the overall evaluation of the school. This means that the primary goals of schools focus on their own interests at the expense of the primary needs of the students and their well-being (Ball et. al., 2010, in Kaščák & Pupala, 2011). Another example of the impact of high-stakes testing on students comes from the Canadian province of Alberta, where high-stakes testing was introduced into education in the 1980s. This was supervised by the MACOSA commission (*Minister's Advisory Committee on Student Achievement*). The commission issued recommendations relating to high-stakes testing, where there is no mention of any positive benefits for the students who are involved in testing (Graham & Neu, 2004, in Kaščák & Pupala, 2011). It is obvious that in some cases, schools and principals are

willing to commit unfair practices, to manipulate students and put them in disadvantaged groups simply to ensure a higher place in the performance tables. However, this manipulation ultimately affects the students.

We mentioned above the negative impact of high-stakes testing upon students taking the test. We have described the ways in which the various intervening parties impact upon the students taking the test. In this subsection we will direct our attention to the pupils, to what happens during testing and their experiences. The greatest impact of high-stakes testing is felt by the individual pupils as students. Reddell (2010) describes experiences of testing in her research, where she noted a wide range of impacts, on students that is, manifest in various forms such as stress and tension in students, teachers' preferences for "better" students, undermining of student self-esteem and in some cases even student fear of failure and the associated consequences. The students suffer in various psychological and physical ways, for example, anxiety, stress, exhaustion resulting from lack of food and water, an increase in blood pressure and the rate of respiration, elevated body temperature, gastrointestinal problems, headaches, difficulty sleeping, and muscle spasms (Blazer, 2011). During the test, students are exposed to an enormous amount of stress caused by public comparisons between students in the classroom or school. Some schools produce test result comparisons in visual format comparing individual students whose test scores fell. In this way, students can see the evaluations of all their classmates, which may be very unpleasant for pupils who have not achieved the required score. In addition it is only the test result that will affect the further development of the students in education (Reddell, 2010). Students cope with the pressures placed on them in different ways from adults. If a child constantly works under pressure, we cannot expect there to be no impact on their psychological or emotional well-being. Stress in children manifests itself in a variety of behaviour issues such as frequent, unjustified and unpredictable "explosions" or problematic behaviour at school or at home. Other possible symptoms of stress are jumpiness, nervousness and poor concentration, which can affect the student's performance in school. The symptoms of stress also include a lack of appetite and the student being frequently ill.

The negative impacts of high-stakes testing upon teaching and learning

1. High-stakes testing interferes with teaching and learning. Under high-stakes testing, the way students are taught is changing along with the methods used and the way in which teachers approach instruction. Creative interdisciplinary activities and project-based investigations are being

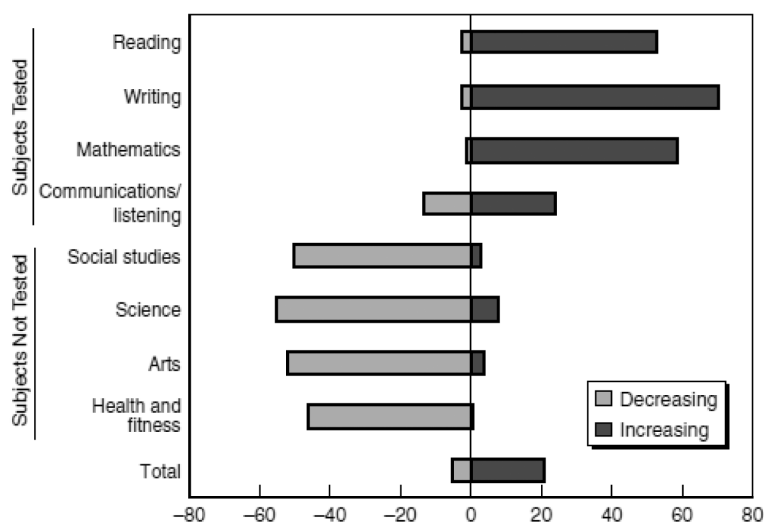


Figure 1 Focus of teachers upon subjects
 Stecher et al., 2000a, p. 21, In: Stecher, 2002

left out. Teachers prefer more “traditional lecture and recitation” strategies to “innovative instructional strategies such as cooperative learning and creative projects” (Blazer, 2011). The teachers put considerable emphasis on students achieving the maximum test score because the results of high-stakes testing are linked to national test results, for example, through “teaching to tests”, providing training, and “coaching” students in how to answer questions (Harlen 2007, in West, 2010). High-stakes testing also leads to an increased emphasis on tested content (Westchester Institute for Human Services Research, 2003). Most teaching time is devoted to preparing for the testing or doing the testing. This is called “teaching to tests”. Simply put, certain subjects are given preference over others in the classroom or increased time is allocated to subjects which are part of the high-stakes testing at the expense of subjects not included in the tests. If the teacher spends a lot of school time preparing students for tests, the quality of teaching decreases. Other subjects (which are not included in the testing, for example, art, physical education, social studies and science) are simply marginalized. Since high-stakes testing demands that students be prepared for tests, it “promote[s] ways of teaching that are often boring and neglectful of problems and issues concerned with race, class, gender, and sexuality” (Grant, 2004). One of the first studies on the effects of high-stakes testing (carried out in two Arizona schools in 1980) showed how subjects were reallocated. The

study showed that reduced emphasis was placed on important subjects. The study also revealed that teachers neglected subjects such as science, social studies and writing, i.e. the subjects not included in the test (Smith et al., 1991, in Stecher, 2002). Figure 1 shows the changed emphasis on subjects in the State of Washington after the introduction of high-stakes testing.

A similar example is given by the national survey of teachers, where nearly 80% of the teachers confirmed that they had paid greater attention to teaching subjects that are part of the test. In addition 50% of teachers testified that the subjects that are not part of the test (art, physical education, foreign language, industrial/occupational training) are allocated less time (Pedulla et al., 2003, in Madaus & Russell, 2010). One teacher said that at their school, teachers were “forbidden” to teach social science before March (Hoffman et al., 2001, in Sadker & Zittleman, 2004). The time devoted to test subjects is increasing at the expense of non-test subjects; this is evident from the above examples. Moreover these “timing changes” affect the child’s breaks during the school day. Since it was necessary to increase the time devoted to testing subject knowledge, the solution was found in restricting breaks. In response to this step, the *Rescuing Recess campaign* was launched, starting on 13 March 2006, by the *National Parent Teacher Association* (PTA) and *Cartoon Network*. This is a campaign that promotes the importance of recess for children and focuses on work that will help sustain and revive the idea in schools across the country (National PTA, 2006).

The negative impacts of high-stakes testing upon schools

As mentioned in the introduction, the influence parents have over their children’s education is manifested through personal choice over decisions about which school is best for their child/children. However, it can happen the other way round and the choice falls on the side of the school, especially if there are more candidates than places available and where schools publish admission criteria to select which students are to be offered places (West et. al., 2004, 2006, in West, 2010). The criteria for selecting students for entry into school divide them up according to those who are likely to be successful in tests and those who are likely to achieve a high score. The league tables also affect private schools. There were concerns that schools prevent pupils from participating in testing in some cases, especially if they are certain that students will not achieve good results. The chairman of the Associations of Independent Schools said that this was exactly what was going on in independent schools and he referred to league tables as an expression of tyranny, which has a toxic effect on the schools, forcing teach-

ing to the test and a narrow curriculum (Times Educational Supplement, 2008b in West, 2010). Stecher (2002) refers to two basic effects (positive and negative) of high-stakes testing upon students, administrators and politicians. We will focus only on the negative effects since that is the aim of our study (see Table 1).

Table 1 The negative effects of high-stakes testing

	Negative effect of high-stakes testing
Impact on Students	<ul style="list-style-type: none"> – Frustration of students and a subsequent decrease in effort – Makes students more competitive – Causes students to devalue grades and school assessments
Impact on Teachers	<ul style="list-style-type: none"> – Encourages teachers to focus more on specific test content than on curriculum standards – Leads teachers to engage in inappropriate test preparation – Devalues teachers' sense of professional worth – Encourages teachers to cheat when preparing or administering tests
Impact on Administrators	<ul style="list-style-type: none"> – Leads administrators to enact policies to increase test scores but not necessarily increase learning – Causes administrators to reallocate resources to tested subjects at the expense of other subjects – Leads administrators to waste resources on test preparation – Distracts administrators from other school needs and problems
Impact on Policymakers	<ul style="list-style-type: none"> – Provides misleading information that leads policymakers to suboptimum decisions – Fosters a "blame the victims" spirit among policymakers – Encourages a simplistic view of education and its goals

Source: Stecher, 2002

Negative impacts of high-stakes testing upon teachers

Besides these negative practices, high-stakes testing affects teachers in other ways. Teachers are also exposed to pressures which are also not without consequences. This is reflected in various, almost unacceptable, practices such as cheating. Many teachers said they had helped students do their coursework in order to achieve better results (Times Educational Supplement 2008a, in West, 2010). A new regulator of qualifications and tests, *Ofqual* (*Office of Qualifications and Examinations Regulation*) was created in England amid fears that students were cheating in coursework (help from

teachers, parents or copying from the Internet). All this is in aid of one thing only: to improve the test and exam performance of pupils. High-stakes testing also affects teachers' instructional and assessment practises. For example, McMillan (1999) surveyed 570 secondary teachers and 152 elementary teachers in Virginia. He found that more than 80% had indicated that the Standards of Learning (SOL) test had affected teachers' instruction. Teachers often leave the teaching profession because of what is going on under the name "accountability" and "tougher standards" (Kohn, 2000, in Reddell, 2010). Teachers are exposed to more stress by high-stakes testing. There are claims that the introduction of high-stakes testing has worsened the problems associated with keeping teachers. Many teachers handed in their resignations to schools after the introduction of high-stakes testing (Kohn, 2000, Jones et al. 1999, Haney, 2001, in Amrein & Berliner, 2002). Jones et al. (1999) conducted a survey among teachers in North Carolina, which found that roughly 75% of teachers had left school during the summer. They left because the State had designated the school as a low-performing school and they wanted to avoid the consequences. Teachers leave education because of the pressure associated with high-stakes testing. In short, teachers have been leaving state schools to go to private schools since the introduction of high-stakes tests because private schools are exempt from the policy of high-stakes testing (Amrein & Berliner, 2002).

Teachers, like students, find themselves under great pressure and stress due to high-stakes testing. They often face pressure from higher places and live with the knowledge of the known effects and disadvantages that result from possible failure in high-stakes testing. As an example, the national study includes interviews with teachers about high-stakes testing. The report states that seven out of ten teachers feel stressed as a consequence of high-stakes testing and two out of three teachers believe that preparation for the testing takes up time that would otherwise be available for teaching important topics (Quality Counts 2001, in Sadker & Zittleman, 2004).

High-stakes testing and its influence

We have already described the impact of high-stakes testing upon students, teachers and even politicians, but we have not discussed the tests themselves. Just as teachers or students can be affected, so can the process of testing itself or the evaluation tests.

As Reddell (2010) noted those involved in correcting and evaluating tests may not mark tests correctly and fairly. Most of the tests consist of parts that can only be marked by human hand, for example essays. Other parts of the test can be electronically marked, but there is no guarantee that the

results are evaluated correctly. In other words, machines may also succumb to the same influences. Questions with multiple choice answers are easy to correct: the answer is either right or wrong. However, computers can also be used to evaluate the essays. The criterion is the name of the categorising software used to determine a student's skill in writing essays. Jehlen (2007) replaced every occurrence of "the" with "chimpanzee" in a student's work. The programme corrected the essay and assessed it to be "cogent" and "well-articulated".

However, the tests themselves may be wrong. One example is the story of Martin Swaden and his daughter who did not pass a maths test. Swaden wanted to see his daughter's test to find out where she had made mistakes and what still needed working on in her order for her to succeed in the next test. Together with a state official, Swaden discovered that 6 out of 68 responses were scored incorrectly not only in his daughter's case but for all the students in Minnesota as well (Draper, 2002, in Sadker & Zittleman, 2004).

Impact of high-stakes testing on the state purse

The question of money and finance also concerns high-stakes testing. Testing is carried out, logically, through tests. However, these tests must be produced by someone, delivered to each school and marked, which is naturally not free. This is where different businesses and companies operating in this area enter into the equation. But when we imagine how many schools and students annually participate in the testing, we find that the cost of preparing and administering the tests is not negligible. On the contrary, this trade is a "money-spinner". It is estimated that in the United States, following the adoption of the *No Child Left Behind* policy in 2002, the industry that produces high-stakes testing for state education departments, increased its annual revenue to \$1.1 trillion (Vu, 2008, in Graber, 2011). In an interview about the integration of high-stakes testing into teaching, teachers said that these tests were a way in which the state demonstrated its efforts to improve educational standards through no small funds (Barksdale-Ladd & Thomas, 2000). Money invested in high-stakes testing (on the preparation, administration and evaluation) could be used in other ways, such as supplying tools to schools, purchasing technical equipment to be used in lessons and so on. The money could be used to enrich the curriculum and the professional development of teachers in relation to assessment (Stiggins, 2002, in Afflerbach, 2005). The reality is quite different. The government annually invests exorbitant amounts in high-stakes testing. Eighty per cent of the school budget was allocated to standard education in 1967. By the end of 1990, the proportion of resources devoted to standard teaching had been

reduced to about 50 per cent (Rothstein & Miles, 1996, in Baines & Stanley, 2004). Some schools (particularly urban and rural ones) that have insufficient financial resources have to pay for the tests out of the money originally allocated for the recruitment of teachers, to repair a leaky roof or for buying new books (Baines & Stanley, 2004). As mentioned above, the funds needed to implement high-stakes testing are enormous. The annual cost associated with implementing high-stakes testing is comparable to the gross national product of a small country: ranging from \$20 billion to \$50 billion (Centre for Education Policy, 2003, in Baines & Stanley, 2004). It is evident that high-stakes testing digs deep into the purses and the financial budget of states. But the question remains as to whether the state is investing in the right things or whether high-stakes testing is as beneficial to the state/government as is expected.

Conclusion

As West (2010) pointed out both Conservative and Labour governments use the test results to place schools in England in the performance tables, which form the central principle of the quasi-market environment. However, the situation in other countries in the United Kingdom is different. For example national testing was eliminated in Wales in 2004 and in Northern Ireland it was replaced in 2010. National testing of pupils does not occur in Scotland, where there is a completely different education system. But a Scottish Survey of Achievement has been in use since 2005. This survey provides information about the national level achieved in order to assess the education system (West, 2010).

The above examples indicate that high-stakes testing comes in different forms in different countries. While some countries have yet to introduce high-stakes testing, other countries have replaced or abolished these tests. In this article, we have pointed out the negative side of high-stakes testing, which countries introducing high-stakes testing into their education systems should take into account. Although the situation associated with high-stakes testing differs according to country, the negative impact is evident almost everywhere.

References

- Afflerbach, P. (2005). National reading conference policy brief, high stakes testing and reading assessment. *Journal of Literacy Research*, 37 (2), 151-162.
- Amrein, A.L. & Berliner, D.C. (2002). *An analysis of some unintended and negative consequences of high-stakes testing*. Retrieved June, 20, 2012, from <http://nepc.colorado.edu/files/EPSTL-0211-125-EPRU.pdf>

- Baines, L. A. & Stanley, G. K. (2004). High – stakes hustle: Public schools and the new billion dollar accountability. *The Educational Forum*, 69 (1), 8-15.
- Barksdale-Ladd, M. A. & Thomas, K., F. (2000). What's at stake in high-stakes testing: teachers and parents speak out. *Journal of Teacher Education*, 51(5), 384-397.
- Blazer, Ch. (2011). *Unintended consequences of high-stakes testing*. Research Services 1008.
- Booher-Jennings J. (2008). Learning to label: socialisation, gender, and the hidden curriculum of high-stakes testing. *British Journal of Sociology of Education*, 29(2), 149-160.
- Caldwell, B., J. (2011). *Educational reform and change in Australia*. Retrieved June, 20, 2012, from HKIED website <http://www.ied.edu.hk/apclc/roundtable2011/paper/Brian%20J.Caldwell.pdf>
- Duncan, B. A. & Stevens, A. (2011). High – stakes standardized testing: Help or hindrance to public education. *National Social Science Journal*, 36 (2), 35-43.
- Educational Act. (1980). Retrieved June, 20, 2012, from Education in England website http://www.legislation.gov.uk/ukpga/1980/20/pdfs/ukpga_19800020_en.pdf
- Education Reform Act. (1988). Retrieved June, 20, 2012, from Education in England website http://www.legislation.gov.uk/ukpga/1988/40/pdfs/ukpga_19880040_en.pdf
- French, D. (2003). A New vision of authentic assessment to overcome the flaws in high stakes testing. *Middle School Journal*, 35(1), 1-15.
- Graber, D. (2011). *The Problem with high stakes testing*. Retrieved June, 20, 2012, from Manchester College website <http://users.manchester.edu/Student/asgraber/ProfWeb/GraberDrewS230HStkspaper.pdf>
- Grant, C. A. (2004). Oppression, privilege and high – stakes testing. *Multicultural Perspectives*, 6 (1), 3-11.
- Heubert, P. & Hauser, R. M. (1999). *High stakes: Testing for tracking, promotion and graduation*. Washington, DC: National Academy of Sciences.
- Jehlen, A. (2007). *Testing: How the sausage is made*. Retrieved June, 20, 2012, from The National Education Association website: <http://www.nea.org/home/11366.htm>
- Johnson, E.G. (1992). The Design of the national assessment of educational progress. *Journal of Educational Measurement*, 29 (2), 95-110.
- Jones, M. G., Jones, B. D., Hardin, B., Chapman, L., Yarbrough, T. & Davis, M. (1999). The impact of high-stakes testing on teachers and students in North Carolina. *Phi Delta Kappan*, 81 (3), 199-203.
- Kaščák, O. & Pupala, B. (2011). Nový režim „kvality“. In O. Kaščák & B. Pupala (Eds.). *Školy v prúde reforiem* (pp. 137-194). Bratislava: Renesans.
- Kentli, F. D. (2009). Comparison of hidden curriculum theories. *European Journal of Educational Studies*, 1(2). Retrieved June, 20, 2012, from http://www.ozelacademy.com/EJES_v1n2_Kentli.pdf

- Lobascher, S. (2011). What are the potential impacts of high – stakes testing on literacy education in Australia? *Australian Journal of Language & Literacy. Literacy Learning: the Middle Years*, 19(2), 9.
- Madaus, G., Russell, M. (2010). Paradoxes of high – stakes testing. *Journal of Education*, 190 (1/2), 21-30.
- Marchant, G. (2004). What is at stake with high stakes testing? A Discussion of issues and research. *Ohio Journal of Science*, 104 (2), 2-7.
- McMillan, J., Myran, S. & Workman, D. (1999). *The impact of mandated statewide testing on teachers' classroom assessment and instructional practices*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Quebec, Canada, p. 10.
- McNeil, L. (2000). *Contradictions of school reform*. New York, NY: Routledge
- Moon, T., R; Brighton, C., M.; Jarvis, J., M. & Hall, C. J (2007). *State standardized testing programs: Their effects on teachers and students*. Retrieved June, 20, 2012, from National Research Center on the Gifted and Talented <http://www.gifted.uconn.edu/nrcgt.html>
- National PTA. (2006). *Recess is at risk, new campaign comes to the rescue*. Retrieved June, 20, 2012, from <http://www.peacefulplaygrounds.com/pdf/right-to-recess/national-pta-recess-at-risk.pdf>
- Niculae, M., & Doncu, R. (2007). *Romanian Educational System: European Conntext*. Analele Universității din Bucuresti. Retrieved July, 10, 2012, from http://anale.fizica.unibuc.ro/arhiva/2007/An_Fiz_2007_5.pdf
- Nichols, S., L. & Berliner, D., C. (2007). *Collateral damage: How high – stakes testing corrupts America's schools*. Cambridge, MA: Harvard Education Press.
- Petrová, Z. (2011). Dopady vzdelávacej reformy v Anglicku a vo Walese. In O. Kaščák & B. Pupala (Eds.). *Školy v prúde reforiem* (pp. 84-106). Bratislava: Renesans
- Reddell, S. (2010). *High stakes testing: Our children at risk*. Retrieved June. 20, 2012, from http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_ERICExtSearch_SearchValue_0=ED513338&ERICExtSearch_SearchType_0=no&accno=ED513338
- Sadker, D. & Zittleman, K. (2004). *Test Anxiety: Are students failing tests or are tests failing students?* Retrieved June, 20, 2012, from http://www.sadker.org/PDF/test_anxiety.pdf
- Stecher, B. M. (2002). Consequences of large-scale, high stakes testing on school and classroom practice. L. S. Hamilton, B. M. Stecher & S. P. Klein (Eds.). *Making sense of test-based accountability in education* (pp. 79-100). Santa Monica: RAND Corporation.
- West, A. & Pennell, H. (2002). *How new is new labour? The quasi-market and english schools 1997 to 2001*. Retrieved June, 20, 2012, from <http://eprints.lse.ac.uk/214/>
- West, A. (2010). High stakes testing, accountability, incentives and consequences in English schools. *Policy & Politics*, 38(1), 23-39.

Westchester Institute for Human Services Research. (2003). *High – stakes testing. The Balanced View*, 7(1). Retrieved June, 20, 2012, from <http://www.sharingsuccess.org/code/bv/testing.pdf>

Wilkins, A.(2012). School choice and the commodification of education: a visual approach to school brochures and websites. *Critical Social Policy*,32 (1), 69-86.

Woods, P., A., Bagley, C. & Glatter, R. (1998). *School choice and competition: Markets in the public interest?* London: Routledge.

Author:

Michaela Minarechová, M.Ed., Ph.D. candidate
Trnava University of Trnava
Faculty of Education
Department of Preschool and Primary Education
Priemyselná 4
Trnava
918 43
Slovakia
Email: michaela.minarechova@gmail.com