



# Test-based accountability and perceived pressure in an autonomous education system: does school performance affect teacher experience?

Natalie Browes<sup>1</sup> 

Received: 5 June 2020 / Accepted: 9 June 2021

© The Author(s), under exclusive licence to Springer Nature B.V. 2021

## Abstract

Across the globe, education quality has become synonymous with student performance. The shift towards test-based accountability (TBA) has changed what is required of schools and what it means to be a ‘good teacher’. Different tools may trigger a performance orientation within schools, from administrative (such as the Inspectorate) to market (schools competing for students). It is logical to assume that TBA policies will be interpreted and enacted differently in schools at different ends of the performance spectrum, and this, in turn will affect the expectations on teachers and the pressures they feel. Based on interviews with teachers ( $n=15$ ), principals ( $n=4$ ) and the school board ( $n=1$ ), this study compares the experiences of teachers in two ‘high’ and two ‘low’ performing primary schools under the same management in one Dutch city. Findings reveal that the schools respond differently to TBA, and are facing different performance pressures, yet in all four, test data was found to significantly shape educational practices. It was further found that teachers experience pressure in different ways; however, it cannot be said that those in high-performing schools experience *less pressure* compared to those in low-performing schools, or vice versa. Rather, teachers’ experience of pressure is more closely connected to their schools’ logics of action: the practices the schools adopted in response to accountability measures and their relative market position.

**Keywords** Test-based accountability · Policy enactment · School performance · Teacher pressure · Logics of action

---

✉ Natalie Browes  
[browesn@gmail.com](mailto:browesn@gmail.com)

<sup>1</sup> Department of Sociology, Universitat Autònoma de Barcelona, 08193 Cerdanyola, Barcelona, Spain

# 1 Introduction

With schools worldwide being held accountable through their student test results, education has become a matter of performativity. Teachers are expected to align their teaching to core learning standards, demonstrate data-oriented working, and show improvements in their students' test scores. Holding actors accountable for education quality in this way is known as 'test-based accountability' or TBA. Research, particularly in Anglo-Saxon contexts, has shown such results-oriented policies to have exerted considerable pressures on teachers: not only changing what is required of them, but changing what it means to be a good teacher and the very identity of the profession (Ball, 2003; Day, 2002; Valli & Buese, 2007).

A large number of studies have detailed the impact of (high-stakes) TBA policies on teachers; (Berryhill et al., 2009; Nathaniel et al., 2016; Saha & Dworkin, 2009; Verger & Parcerisa, 2017; Wilkins, 2011, to name a few) resulting in notorious maxims such as '...the terrors of performativity' (Ball, 2003) and '(life in) the pressure cooker' (Agrey, 2004; Perryman et al., 2011). A far smaller number have directly compared the experiences of teachers in schools that occupy different positions within the same accountability system (DeBray 2000; Diamond & Spillane, 2004; Mintrop, 2007). This study adds to this scant literature, examining teacher experience of TBA and performative pressure in schools under the same management, yet that face different circumstances. This is done in the context of the Netherlands: a system where school (board) autonomy is high, parents have substantial choice in a diverse market of government-funded schools, and the inspectorate has a longstanding quality assurance role. The paper analyses two 'high' and two 'low' performing primary schools under the same board in one city. Dutch primary schools are particularly useful sites to research the impacts of TBA given the significant role of compulsory standardised testing throughout. These tests are used by the government to measure student and school performance and often, though more informally, used by boards and schools to gauge teacher performance. Drawing on the analysis of interviews with teachers, principals, and the school board as well as key school documents, this research seeks to uncover cultural and contextual elements of the schools, to understand their interpretation and enactment of TBA, conceived through the concept 'logics of action,' and, by extension, the performance pressures facing their teachers. The study highlights similarities and differences between teachers' experiences, seeking to answer why these might exist and to what extent they can be understood through school performance level. It is shaped by two main research questions:

- I. What accountability pressures do high- and low-performing schools face, and what logics of action are adopted?
- II. Do teachers in high- and low-performing schools experience pressure differently: in what ways do they feel under pressure to perform, and what do they perceive to be the main sources of this pressure?

This research adds to our knowledge of the enactment and impact of TBA policies in various ways. Firstly, the field is dominated by studies conducted in Anglo-Saxon, ‘high-stakes’ contexts; that is, contexts where sanctions or rewards are attached to test performance, often at the school *and* teacher level. By researching enactment in an accountability system that does not implement such measures, the paper adds to the limited research (see for example Thiel & Bellmann, 2017) that questions implicit assumptions that lower stakes mean no or limited undesired impacts. Secondly, previous studies have tended to focus on particular accountability measures and responses at the *school level*; the state imposed administrative sanctions facing probation schools for example (Mintrop, 2004), or the market pressures facing schools in competitive systems (Van Zanten, 2009). By *also* focusing on the individual experiences of teachers and how they perceive pressure, the study investigates whether fewer (external) performative demands on schools does in fact result in reduced pressure felt by teachers. Finally, by directly comparing teachers’ experiences of TBA in different school settings, the research attempts to highlight assumptions on which these policies are based, providing a deeper understanding of policy enactment and bringing into question ‘what works, where and for whom?’ The research forms part of a wider, comparative project called ReformEd ([reformproject.eu](http://reformproject.eu)), which investigates the evolution and enactment of TBA in various countries around the world.

## 2 TBA, performativity and enactment theory

Comparable test results are central to the TBA model. Tests are designed to measure the attainment of centralised standards and used to gauge student and school performance. Attaching these test results to sanctions or rewards (managed at the municipal, state, or national level) is often referred to *administrative accountability* (see Verger et al., 2019). The school inspectorate or equivalent authority may also play a central role here; administering quality labels and interventions or even closures should a school not meet performance targets. These interventions, or the threat of them, place particular pressure on those schools on the borderline of these targets.

Working hand in hand with these administrative measures described above, the market has also been leveraged as an accountability mechanism. *Market accountability* usually involving the publication of a school’s test results and the Inspectorate’s quality label which are meant to guide parents when it comes to school choice, promotes competition between schools (Verger et al., 2019). Underperforming schools may struggle to attract and/or keep students and may even have to close. A sufficiently performing school that does not adequately diversify and innovate may also struggle to attract students (OECD, 2010, 2013). In these competitive environments, schools’ reputations and the pressure of maintaining them become particularly important (West et al., 2011).

The publication and availability of results also serves to inform the parents of current students about the schools’ performance and their child’s progress. This establishes parents as (another) forum by which schools are held to account (see Bovens, 2007). West et al. (2011) refer to this as *participative accountability*. This

is particularly significant when tests are high stakes for students; holding a certification, streaming or selection function (see Verger et al., 2019). In this regard, schools are also accountable to their students (West et al., 2011).

These policy tools are designed to ensure that teachers are focused on results and the job of results improvement; whether their work is formally assessed through student test scores or not. Studies investigating the enactment of these policies often employ the term ‘performativity’ (Ball, 2003) to denote generally the changes to teachers’ practices and identities, imposed and encouraged, through performance-focused, test-based school accountability policies. The term, used throughout this paper, does therefore not only refer to the process of testing itself, but to the accoutrements of TBA. Many enactment-based studies have highlighted the burden that performativity has placed on teachers:

The spiralling demands of government initiatives, incessant record-keeping, education plans, targeting and inspections, have left teachers reeling (Mathison & Freeman, 2006, p43)

Given the centrality of standardised test results within accountability systems, school performance level is a particularly important factor to consider when examining the pressures of performativity and the enactment of TBA. Enactment theory has underlined the complex dynamics involved in ‘doing policy;’ understanding school actors as *policy shapers*, pivotal in their interpretation and translation roles (Ball et al., 2011). It has also underscored time and again the utmost importance of context (see for example, Braun et al., 2011): to understand enactment, we must understand something about performance context (school competition, reputation, student population...), as well as school (performance) culture (Ball & Maroy, 2009). These are fundamental factors which both shape and reflect policy translation. The extent to which a school’s image is tied into academic performance is key here. Scholars’ work on schools’ ‘logics of action’, provides an excellent perspective from which to better understand this notion of performance culture.

Ball and Maroy (2009) define logics of action as consistencies in practices and decisions regarding various aspects of a school’s external and internal functioning, including student enrolment, curricular content, organisation of classes and performance. Linking to earlier work (Bernstein, 1996; Hargreaves, 1995), these practices or ‘functions’ are thought to exist within two domains: schools must ‘deal with an instrumental function, or task achievement, but also with an expressive function, or maintaining good social relationships’ (Hargreaves, 1995, p25). According to Hargreaves, these two domains are at the core of school culture.

Schools with a predominately *instrumental order* may appear more exclusive, with a high level of student homogeneity and an emphasis on academic performance and results maximisation. Ability differentiation and programmes that target high-performing students are common and hierarchy and authority are likely to play a strong role in shaping the relationships between students and teachers as well as teachers and management. Conversely, schools where the *expressive order* is dominant can be recognised through the importance attributed to inclusivity and equity, often concentrating on low-performing students and emphasising student socio-emotional development. Interpersonal relationships play a central role in the

organisation of the school, including more open relationships between teachers and management and greater collaboration and teacher involvement in decision-making (Ball & Maroy, 2009).

These functions can be understood as a response to regulatory processes (accountability structures for example), as well as to more local processes (the nature of and the school's position within the education market) (Maroy & van Zanten, 2009). They serve to promote a particular 'image,' to attract and retain a certain profile of student. We would therefore expect to find a relationship between a school's logics of action, student population, performance, and market position. An understanding of these factors is fundamental to understand the impact of accountability measures on schools, and by extension, the expectations on teachers and the pressures they experience.

## 2.1 State of the art

A number of studies have looked at the enactment and impact of TBA, mediated by school performance within a system (Diamond & Spillane, 2004; Mathison & Freeman, 2006; Mintrop, 2004), the nature of the accountability system: 'high' or 'low' stakes (Nathaniel et al., 2016; Thiel & Bellmann, 2017), and evolution along the TBA timeline (Holloway & Brass, 2018). Comparing high- and low-performing schools in two 'high-stakes' American states, Diamond and Spillane (2004) reported particular pressures and undesirable responses in underperforming probation schools. These schools were generally attended by high percentages of poor and ethnic minority students. The studies found that school responses to accountability sanctions focused on superficially complying with policy demands (Diamond & Spillane, 2004) resulting in narrowed curricula and educational triage, whereas high-performing schools implemented more comprehensive and meaningful changes that were 'closer to the intentions of policy makers' p.1157). Significantly, Thiel and Bellmann (2017) revealed that undesirable side effects are not only present in schools operating in high-stakes contexts, but also in 'low' and 'no- stakes' contexts. Indeed, even in contexts with no formal administrative sanctions or free school choice, the fact that test scores have become inherently tied to perceptions of school and teacher quality, exacerbated by the publication of these scores, may result in considerable reputational pressures.

At the teacher level, empirical studies such as those conducted in the English and US contexts have found (high-stakes) accountability to negatively affect the curricular and pedagogical freedom of teachers; whether they are explicitly required to follow increasingly detailed prescribed curricula (Dobbins, 2009; Webb et al., 2004) or manipulated into narrowing curricula due to the stakes attached to tests (Berryhill et al., 2009; Mathison & Freeman, 2006). In a particularly comprehensive paper, Valli and Buese (2007) reported that teachers in one US, mid-Atlantic state experienced a loss of control over the curriculum in terms of its pace, content and organisation. Mathison and Freeman (conducting research in the state of New York) found that, as a result of such performative pressures, teacher stress was prevalent across schools, regardless of their performance level.

Within these neoliberal systems of accountability and surveillance, core aspects of the teaching profession appear to be changing. *Regulated autonomy* for example (Dale, 1982), describes the existence of teachers' autonomy within a limited scope, constrained by increasing external control. Perryman et al (2011) present the notion of *earned autonomy*: awarded and dependant on (continued) good results. Such change is not only evident in teachers' work environments, but also in their expectations and perceptions of the profession. Guided by Ball's work (2003, 2016), Holloway & Brass' findings (2018) reflect the tightening grip that TBA is having on the teaching profession. Comparing two separate studies, US teachers interviewed during the 'second wave' of TBA (Obama's 'Race to the Top') expressed markedly different views from those interviewed almost a decade earlier amidst Bush's 'No Child Left Behind' Act. Whereas teachers experiencing TBA in its early stages saw it as an external intrusion with clearly negative impacts on their autonomy, practices, and professionalism, 10 years later these policy tools had apparently become woven into the fabric of teaching. Not only were test data, standards, and performance indicators valued, but TBA mechanisms had become 'the very modes by which they (teachers) knew themselves and their quality' (Holloway & Brass, 2018 p 362). In such environments, with quality reduced to measurable output and value placed on recording work over the work itself (see Ball, 2016; Day, 2002), it has been argued that a professional teacher is more akin to a technician, expected to 'do things right' rather than 'do the right thing' (Darling-Hammond, 1990 p31). How these issues play out in a system such as the Dutch one, characterised by its high levels of decentralisation and school and teacher 'freedoms,' needs to be explored further.

### 3 Accountability in Dutch schools

In Dutch primary schools, centralised core learning standards are measured primarily through the final or 'end-test' (eindtoets) which takes place in group 8, the final year of primary school (akin to grade 6). Schools must also be able to show they are following the development of students in relation to these standards through a student tracking system 'leerling volg systeem' or 'LVS.' This system is test-based, with tests usually taking place twice each year and compulsory as of group 3 (when students are 6–7 years old), although many schools already start them in groups 1 or 2. There are a number of companies that provide the end-test and LVS tests, and schools can choose which to opt for.

Beyond these learning objectives, there is no core curriculum that schools must follow; however, spotting an opportunity, a number of private companies have developed subject curricula, known as 'methods' in line with learning standards and test content. These 'methods' are highly-structured and detailed, and often employ end-of-unit testing. As part of the tracking culture, various student administrative systems, also developed privately, are utilised by schools. With these, teachers are encouraged to record and analyse student results, develop 'class' or 'group plans' (the results-based ability grouping of students and subsequent curricular plans), and keep (daily) notes on students' work, progress, and behaviour, as well as a record of all parental interactions. Using these teaching methods and administrative systems is

not compulsory, although the inspectorate states that schools must be able to show they are following the development of their students in a systematic way (Inspectie van het Onderwijs, 2020).

Through the standardised tests, Dutch primary schools are held accountable in a number of ways. If a school performs below average in the end-test for three consecutive years,<sup>1</sup> it is labelled as ‘very weak’ by the inspectorate and undergoes an intensive inspection and monitoring process, whereby a selection of other standardised test results and education processes will also be examined. Schools must produce improvement plans and are encouraged to seek external support from one of a number of providers to help get performance ‘back on track’ (Inspectie van het Onderwijs, 2020). If a school consistently underperforms, showing little or no sign of improvement, the Minister of Education has the right to withdraw funding, effectively closing the school (Waslander, 2010). However, policy experts—interviewed at an earlier stage of the research—could only recall this happening on ‘one or two’ occasions (Browes & Altinyelken, 2019). Teachers’ salaries and bonuses are not connected to test scores and neither are teachers *formally* assessed through scores<sup>2</sup>; however, administrative requirements are powerful in shaping the priorities of schoolboards and principals, and by extension, shaping teachers’ practices.

Market accountability mechanisms work alongside the administrative. In the Netherlands, national- and municipal-level websites provide easy-to-read, school-level data (see [vensters.nl](http://vensters.nl) or [allecijfers.nl/basisscholen](http://allecijfers.nl/basisscholen)). This includes some qualitative data (such as school plans), but is mainly quantitative (test scores and secondary school advice). Performance data is shown against national averages and can also be compared directly with other schools. Given that funding is attached to the student, based on parental education level, a poor-performing school may not only suffer from a reduction in student numbers but also a reduction in money. To encourage market dynamics, satisfactory-performing schools may now also request a quality inspection. Here, the inspectorate assesses various aspects of a school’s organisation and functioning, with the view of awarding a label of ‘good’ or even ‘excellent’ (see Inspectie van het Onderwijs, 2020).

TBA in the Netherlands therefore differs in important ways from TBA in high-stakes systems, such as those found across the USA. According to Thiel and Bellmann (2017), given the lack of formal (and utilised) sanctions and rewards attached to test performance, the Netherlands could be classed as a ‘low-stakes’ system. Yet, the looming presence of the Inspectorate and the publication of performance indicators provide a constant reminder to school actors of the need to play the performance game, and teacher stress and burnout are significant (Inspectorate of Education, 2019). The very fact that results are published in a market where parents exercise school choice, makes a ‘low-stakes’ label problematic. Beyond this, given their role in secondary school streaming, standardised tests are high stakes for students, particularly in the upper years (groups 6, 7 and 8). Dutch secondary education varies

<sup>1</sup> Compared to schools with a similar SES.

<sup>2</sup> Indeed, there is no centralised framework for teacher evaluation, rather this is left up to the schools and boards to decide.



considerably, ranging from more practically focused schools ‘voorbereidend middelbaar beroepsonderwijs’ or ‘VMBO’ (leading to vocational training upon graduation), to ‘hoger algemeen voortgezet onderwijs’ or ‘HAVO’ (leading to university of applied science), to ‘voorbereidend wetenschappelijk onderwijs’ or ‘VWO’ (theoretical education leading to research university) (see Nusche et al. 2014). Movement between streams is challenging, making performance at primary school a core determinant of access to higher education and future career path. This creates an environment in which participative accountability may also play a significant role (see for example CNV Onderwijs & EenVandaag, 2018; Duo Onderwijsonderzoek & De Monitor, 2016). In summary, the current definitions and binary ‘high-stakes’, ‘low-stakes’ categorisations of accountability systems are limiting. For the reasons outlined above, it is perhaps most accurate to describe Dutch accountability as a ‘middle-stakes’ system.

Therefore, despite important differences regarding how accountability operates, Dutch teachers are regulated in many of the same ways as teachers in high-stakes contexts. An international comparison of teacher professionalism structures (Voisin & Dumay, 2020) indicates that the Dutch system can be best understood as a ‘market model:’ market and standards-based regulation, characterised by, amongst other things, relatively low levels of professional autonomy (p.7). This model is also found in England and the USA (amongst others). Given the high rates of school autonomy that define the Dutch system, this may seem counterintuitive. Indeed, considering the numerous intermediary stakeholders involved in the system and the numerous ways in which schools are driven to improve, it is important to examine these mechanisms and their effects on teachers more carefully.

## 4 Methods

The research adopts an exploratory case study approach (Yin, 2009). It compares the experiences of teachers across four primary schools, differing in their performance levels. As such, it can best be understood as a multiple case, embedded approach (Yin, 2009). Purposive sampling (Patton, 2015) was used to select schools, based on the criteria of school management, school ‘type,’ school location, and school performance level.

Schools under the same board were selected to keep external management conditions as equal as possible. All of the schools are public (‘openbaar’), meaning they are government-established<sup>3</sup> as well as government-funded, and cannot discriminate or tailor education based on religious or philosophical grounds. In total, the board manages 26 schools, 21 of which are general primary schools. The schools are located in a small but densely populated city in the province of North Holland. It was necessary to select an urban area with a competitive education market, given the study’s interest in the role of the market in accountability and performance pressure.

<sup>3</sup> However, one of the schools was created through a merger of a public and a general independent school.



Big cities were avoided for fear of research fatigue, indicated during an initial exploration of the field. The researcher created an excel sheet of all schools under the board, ‘ranking’ them based (only) on performance data, and contacting two of the top and two of the bottom ranked schools for participation in the study. Given the focus of the research, ‘performance’ was determined in a very narrow, quantitative way, based on (1) the schools’ average scores in the end-test and (2) the percentage of students going to VWO (the most academic level of secondary education). Beyond these figures, the term does not reflect school or teacher quality and says nothing about student, parent or inspectorate satisfaction. It is noteworthy that while both of the ‘low performing’ schools serve highly-mixed and non-ethnic Dutch populations, the ‘high performing’ schools are attended by almost exclusively middle class, ethnic Dutch populations. This high correspondence between socio-economic composition and school performance is not coincidental (Ball & Maroy, 2009; Diamond & Spillane, 2004).

In each school, interviews were sought with the principal and four teachers, including one group 8 teacher (the ‘end-test year’), another teacher from the ‘bovenbouw’ (groups 6–8, where tests are more high stakes for schools and students), and two group teachers from other tested years. Teacher recruitment for the research was challenging for various reasons, resulting in an incomplete sample and unavoidable instances of selection bias. In all schools, the researcher first made contact with the principal to explain the study and request participation. Principals were concerned about adding to the high workloads of their teachers and would not enforce participation, rather agreed to inform teachers should any show interest. Akin to voluntary response sampling, or self-selection bias (Collier & Mahoney, 1996), it is worth noting that these participants may be particularly politically engaged, or have particularly strong opinions about the policies under investigation. In case there were still gaps in the sample, ‘missing’ teachers were contacted by the researcher. To encourage participation and show appreciation, all interviewees were presented with a book voucher upon completion of their interview. In one school, neither of the two group 8 teachers agreed to participate, bringing the number of teacher interviews to three. However, this included a teacher who had recently taught group 8. In another school, the vice principal stood in for the principal who was unavailable for an interview. This brought the interview total to 20; school board ( $n=1$ ), (vice) principal ( $n=4$ ), teacher ( $n=15$ ). For a full list of participants, see Table 2 in the appendix. All schools and participants have been given pseudonyms to maintain their anonymity.

Interviews were conducted one-to-one, lasted between 32 and 57 min and were semi-structured. The interviewer followed a script but remained flexible to pursue relevant issues that were raised. Interview scripts differed slightly between actors, yet contained the same themes and many overlapping questions so that data could be triangulated. Themes included teachers’ work routines and working environments, testing practices and opinions on TBA, and experience of performance pressures and autonomy. Interviews were transcribed verbatim and then analysed using ‘Atlas.ti’. To gain an insight into schools’ logics of action, the most recently available school guide was downloaded and also coded. Schools are required by law to produce this document annually. It provides information to parents of prospective

and current students about the mission, vision, working methods, education plans and performance of the school. The comprehensive nature of these documents and their function to inform but also promote the school, make them an excellent data source. Content analysis, using a directed approach (Hsieh & Shannon, 2005), was conducted on both the interview scripts and the school documents. For this purpose, a code list was developed iteratively, guided by theory and the research questions. Examples of codes include school context\_perceived reputation, policy interpretation\_opinion on TBA, policy translation\_test preparation, policy translation\_student differentiation, perceived pressure\_inspectorate, perceived pressure\_parents. The interview and coding scripts were developed with fellow ReformEd researchers as comparative research tools, but allowed for the addition of case-specific issues.

## 5 Findings

Findings are presented according to the study's research questions. The first section focuses on the school level: examining school contextual and cultural factors to try and determine the importance of performance and the degree of performance pressure in each school. The next section shifts to the teacher level, to see how this impacts teachers' experiences of performance pressure, and if and how this experience varies between schools.

### 5.1 Performance context and logics of actions: a study of four schools

Table 1 (below) presents the contexts of each school, including their population, performance, and market position as perceived by the school's (vice) principal. This is followed by a more in-depth qualitative analysis of the schools, focusing on; accountability pressures, school values, and key school practices such as student recruitment and the organisation of learning.

#### 5.1.1 High-performing schools

Barlaeusschool and Basisschool De Witt are amongst the top performing schools under the board. Located in affluent parts of the city, no students receive extra 'dis-advantage' funding and percentages of NOAT students are very low. Parents at De Witt are described as highly educated, wealthy, working professionals, while Barlaeus' population is reportedly somewhat lower middle class, and increasingly consists of blue-collar workers. As will be presented below, the schools' organisational values and their organisation and content of education reveal that an 'instrumental order' dominates in the schools.

Given that both schools comfortably exceed quality requirements, pressures concerning potential administrative sanctions are low. As the result of sufficient performance, it had been some time since Barlaeus received a school inspection (status: 'awaiting current judgement'). Conversely, at the time of research, De Witt had just

**Table 1** School characteristics

	Performance level	No. students (2019/20)	% 'extra funding' students (2018/19)*	% 'NOAT**' students (2019/20)	Av. end-test scores (2017, 18, 19)	% attending VWO (2017, 18, 19)	Inspection status	Enrolment and perceived competition
Barlaeusschool	'High'	432	0%	1%	(IEP)*** 88.6, 85, 86.9	47, 49, 50	'Without current judgement'	Recent decline in enrolment. A moderate-strong amount of competition perceived (particularly with a recently-opened nearby school)
Basisschool De Witt	'High'	603	0%	3%	(CITO)**** 542.1, 541, 539,6	68, 55, 49	Requested quality inspection (2019). 'Good'	The school has a long waiting list. A sense of competition is not felt
Erasmusschool	'Low'	170	35.5%	88%	(CITO)**** 525.2, 531, 528.6	12, 4, 14	Basic inspection (2016). 'Good'	Sometimes struggles to achieve desirable student numbers. Not perceived to be in competition with other schools
Van Schurman-school	'Low'	234	22.5%	58%	(CITO)**** 522, 528, (IEP)*** 87.7	4, 7, 19	Quality inspection due to underperformance (2018). 'Very weak'	The school has a waiting list. A moderate amount of competition is felt

Schools are listed in no particular order

\*2018/19 is the last year this data is available

\*\* 'NOAT' stands for 'non-native speakers in Dutch education'

\*\*\*IEP scores range from 58 to 97. 25% of schools score < 78 points, 25% score > 84

\*\*\*\*CITO scores range from 513 to 548. 25% of schools score < 532, 25% score > 537

completed an elected quality inspection. Following a short but intense inspection, the school was awarded the label of ‘good.’

Despite the absence of any administrative interventions, pressure to maintain results was evident in principal interviews. This pressure was inherently connected to the schools upholding their academic reputations, and thus continuing to attract somewhat privileged and high-performing students. Both principals mentioned the high involvement and high expectations of parents. Bas, the principal of De Witt, further stressed that given his school’s privileged population, the Inspectorate expects more than ‘just sufficient’ results. Despite this, he reported to feel confident in the positive and distinctive reputation of De Witt, and spoke proudly of its long waiting list. Beyond its academic success, Bas considered the school’s warm and collaborative culture absolutely central to this reputation.

The most different thing was that my policy is ‘you work together.’ Open the doors, speak with each other! Also with the parents we open the doors. Before I started, the parents bring their kids and must wait outside in the morning, and that was the first change – ‘come inside!’

(Bas, principal, Basisschool De Witt)

In contrast, maintaining sufficient student numbers was a considerable concern for Barlaeus’ vice principal, Yvonne, an issue that had only arisen in recent years. This appears to be the result of a change in market position: on the one hand a newly opened school in close proximity had increased competition for students, while on the other, a reportedly changing neighbourhood composition had led to a perceived mismatch between the school’s logics of action and the demands of local families.

“...now we want to do more of the nice things. More creativity, more music, every day from 12.00 until 14.00, so all the groups are working on nice things to do. It had to be more attractive, not simply lessons out of the book.”

(Yvonne, vice Principal, Barlaeusschool)

Certainly, for both of these high-performing schools, offering a broad, well-rounded educational experience where regular time is dedicated to gym, music and art, was of high importance. However, the core learning areas and academic achievement, it was stressed, remained a priority, and a performance orientation was evident at the schools. At Barlaeus, this manifested itself through differentiated working, a particularly strong focus on high-ability students, and an emphasis on results-oriented working (for example, participating in data-literacy projects with the explicit aim of increasing student outcomes). Teachers at the school are constructed first and foremost as individual professionals, encouraged to take the initiative to develop themselves and to work towards their own strengths in order to achieve desirable results. De Witt also promotes a strong emphasis on student achievement and a differentiated or ‘customized’ approach to student education, yet one that also appears tailored to struggling students as well as to high achievers.

Teaching ‘methods’ were used at both schools, however, whereas curricula at Barlaeus appears to be (exclusively) prescribed, De Witt is now moving away from the use of purchased methods for non-core subjects. On this matter, De Witt’s

principal emphasised the importance of curricular autonomy to foster (teacher) innovation and ‘avoid dependency,’ yet, given that this has not been extended to the assessed learning areas, at least for now, performance seemingly takes priority over innovation. Both schools also implement LVS tests from group 2 onwards. These results are recorded in biannual student reports and discussed with parents. The data is discussed individually with management as well as in team meetings, and teachers’ class plans, which are based largely on results, are regularly reviewed. In both schools, the importance of results ‘justification’ was also emphasised: a teacher experiencing a results ‘dip’ was admissible, as long as there was an acceptable explanation (i.e. one not linked to the role of the teacher). ‘Unexplained’ drops in results would be met with concern.

We have school results, which we can see from each other, and we are transparent about it; you can see how your group is doing compared to the other groups of the same age. And, it’s not a problem when there are different scores, but it can be a problem when you can’t explain it.

(Bas, principal, Basisschool De Witt)

Similarly, according to Barlaeus’ vice principal, concern about ‘unexplainable’ results would be followed up with teacher meetings and a member of management observing classes to establish the cause of the performance decline.

### 5.1.2 Low-performing schools

Erasmusschool and Van Schurmanschool are located in majority ‘immigrant’ and mixed neighbourhoods respectively, dynamics that are reflected in the schools’ populations. Income levels in these neighbourhoods are below the city average and significant numbers of students at both schools receive extra funding based on the criterion that parental educational levels are low. At Erasmus in particular, many new students reportedly speak little or no Dutch, and in response, the school has forged itself a reputation as somewhat of a language specialist. Van Schurmanschool is attached to a school for special education (sharing a principal), with an aim of eventually integrating these students. As will be seen further, the dominance of an expressive order is evident in the schools’ oft-discussed values, and organisational and educational dynamics.

The average performance of both schools has been amongst the lowest under the board. Their principals reported to feel the pressure of the inspectorate’s eye upon them, and considered their ‘quality status’ precarious given their student populations. Yet, administrative interventions differ considerably between the two schools. Dipping below the average mark for its end-test results for three consecutive years, Van Schurman was labelled ‘very weak’ by the inspectorate, and was undergoing a period of extensive inspection at the time of research. Classed as an ‘underperforming school’, measures included bringing in external support, monitoring from the inspectorate, and the utilisation of a school improvement plan. Performing at or above average, Erasmusschool had received no such risk-based visit. The principal recounted its last (basic) inspection as a positive process and highlighted the favourable report it had received.

Despite relatively low performance, neither principal perceived their school to be under market pressure. While Erasmus' student numbers were reported to fluctuate slightly from year to year—sometimes making target enrolment numbers difficult to reach—this was not attributed to school competition but to a combination of family relocation (the result of housing developments elsewhere in the city), and a number of its students moving into special education. Conversely, Van Schurmanschool was reported to have a waiting list, with the principal attributing its popularity to its reputation as a safe and pleasant place with good teachers. This reputation was reportedly built through word-of-mouth amongst the Dutch-Moroccan community in particular. However, Annemarie, the school's principal, also acknowledged that many 'white Dutch parents' living nearby choose to send their children elsewhere, something she believes may have been exasperated by the school's classification as 'very weak'.

Unlike the more instrumental orientations of the high-performing schools, at Erasmus and Van Schurman, educational needs are emphasised over educational outcomes. Throughout the school guides and principal interviews, the institutions emphasise first and foremost inclusivity, equality, respect and tolerance.

And what we do is we really look - what does the child need? Every child. We always ask ourselves, 'are our children benefiting from this?' when we want to do something new.

(Monique, principal, Erasmusschool)

As in the high-performing schools, differentiated learning is integral to the structure of education, yet at Erasmus and Van Schurman school, it is more designed to incorporate academically struggling students than to push high achievers. Differentiated learning and support for lower achieving students is particularly significant at Van Schurman, which is working towards dividing learning into a 'theoretical' and an 'entrepreneurial' stream from group 6.

Unlike the broad educational offers at De Witt and Barlaeus, curricula at the two low-performing schools appears to be considerably narrower. Non-core subjects, while desirable, were seen as luxuries for which the schools had insufficient time. This feeling was most prevalent at Van Schurman. Annemarie emphasised her expectations that staff follow set guidelines and structured learning in order to 'stay on track', with the school recently switching to new methods in several areas. Teachers, she believed, still had autonomy within these guidelines:

"If you go to teach in a free school, or in a Montessori school, then you teach in a different way than you teach in this school. But if you go here to school, we all write down the programme on the board, and we all write down what the children learn. We all work with an instruction table. But you are free how you do it."

(Annemarie, principal, Van Schurmanschool)

In comparison, while these methods were equally important at Erasmus, teachers appeared to have a much greater role in their choosing. Monique, Erasmus'

principal, reported an expectation that her teachers take the initiative to explore new ideas, whilst making decisions as part of a team.

I expect them to take their own responsibility... I give them space for that and you can lift them up too and I expect them to commit themselves to the way we work - so we speak to each other together and together we make the rules, so I expect them to keep the rules in mind.

(Monique, principal, Erasmusschool)

As in the high-performing schools, LVS tests at Erasmus and Van Schurman are taken twice yearly from group 2, and results are central in discussions with parents, management, and amongst teachers. In contrast to the high performers however, testing is primarily valued as a useful way to identify student learning problems. While teachers in both schools are encouraged to share ideas and tips, and discuss reasons for well-performing and underperforming groups, it was stressed that these meetings are based on support rather than judgement. Since the school was deemed 'underperforming,' Van Schurman's principal explained that data analysis and keeping records of teaching and learning, are, more than ever, stressed as vital parts of her teachers' work. Classroom observations, she noted, are also much more common place.

The principals in the low-performing schools do, however, adopt different perspectives when it comes to test data. Monique, principal of Erasmus, expressed a real interest in the data generated from tests, understanding it as a useful way to gauge the level and progress of a group and to assess the school's performance against others. In contrast, at Van Schurman, comparative test data was generally perceived negatively. This data was the reason behind the schools' 'very weak' quality label, an assessment that was seen by the principal as essentially unfair, and as having resulted in undesirable effects on student learning; namely, an overly narrow curriculum.

...It has changed because the inspectorate has such a high norm (...) But it is not always good for children – they have to play, they have to draw, they have to make music. But I can't say to the parents 'your child can't count, but he can really dance lovely.

(Annemarie, principal, Van Schurmanschool)

## **5.2 Teachers' experiences of performativity and pressure**

There are clear commonalities between teachers' experiences of TBA across the four schools. Reflecting the numerous documented experiences of teachers in the 'high stakes' Anglo-Saxon contexts presented earlier, interviews revealed that, regardless of their school's performance level and logics of action, teachers felt the weight and frustrations of a performative, outcomes-based agenda. This included, to various extents, a sense of restriction in the classroom: being contained by, and carried along with, a rigid schedule and prescribed curricula shaped by test content (see also Valli & Buese, 2007). The vast majority of participants reported to be conscious



of having to ‘stay within the lines’ when it came to the core learning areas: sticking closely to the subject methods used in the school. All teachers reported to have more autonomy outside of these core areas. Furthermore, regardless of school performance level, almost all respondents reported to engage in test preparation activities. The nature of these activities varied, although mostly teachers reported to take time to ‘familiarise’ their students with the structure of test questions. Whereas in the high-performing schools, test preparation activities were reported to take place only in the few weeks before the tests, teachers in the low-performing schools were more likely to integrate them as a routine part of their teaching throughout the year. Similar findings were recorded by Diamond and Spillane (2004) in the US context.

Drawing further parallels with high-stakes contexts (see for example, Mathison & Freeman, 2006), heavy workloads and a sense of insufficient time for tasks was also experienced by teachers across the four schools. Participants expressed a desire to have *more* time: the time to explore new topics and revisit challenging ones, and the time and flexibility to follow intrigue and to teach impromptu. Teachers also expressed a wish to dedicate more time to non-core subjects and for more creative, exploratory, and student-centred learning. The majority of participants wanted to spend *less time* recording and planning their work, particularly in relation to the time spent on class plans and updating student administrative systems. Several teachers across the schools reported to regularly work on their evenings and weekends in order to manage their tasks.

In sum, regardless of school context and performance level, all teachers interviewed recognised themselves to be working in performative environments and were exposed to some degree of pressure around (test) performance (Diamond & Spillane, 2004; Mathison & Freeman, 2006). Yet, perceptions of pressure, its source, and the extent to which it affected participants, varied. Some teachers found a performative work environment quite natural. These individuals considered testing to be a useful indicator of learning, and valued data and documentation for their ability to inform teaching.

We put everything [the test data] in our system and we have an intern begeleider, and she looks with us... ‘What did you do in your lessons?’ ‘What went wrong or what went well?’ ‘Why are the kids growing or ‘what can we do to make them stronger next time?’ So you look at the system and the system gives the answers.

(Eva, group 8, Barlaeusschool)

In these cases, testing was seen as a clear way for teachers to know their goals, keep them focused, and gauge their own as well as their students’ progress (see also Holloway & Brass, 2018). As a result of these views, they were also more likely to report feeling pressure that their students perform well in tests. Another group of teachers were more sceptical, seeing tests as having a specific, yet limited function. These individuals spoke of their confidence in themselves as professionals, and did not connect their professionalism to their students’ test scores. As such, most also claimed not to internalise performance pressure, although a small number of these teachers confessed they did, despite themselves, instinctively measure their teaching abilities through their students’ results.

Data analysis did not reveal clear links between school performance level and the value that teachers attribute to testing. Rather, teachers spoke about the impact of their experience: those who had been in the job a decade or more were less likely to admit to attaching importance to results and were more likely to feel frustration with the growing centrality of testing and the role of data and learning standards within their work. These teachers reported to be witnessing a changing profession—a slow but steady tightening of the reigns.

...in these developments in the last 10 years, we've become more and more slaves of the method – 'we need to do lesson 5 and then you need to go to this topic and then this...' there's very little room to do something for fun, to do something because it is just *interesting* but that doesn't add up to any measurable results.

(Sara, group 6, Barlaeusschool).

Despite some adopting a more critical view, significantly, no teacher rejected testing outright. All respondents found value in test data. The majority of teachers claimed to appreciate the comparative function of standardised tests, not only accepting, but embracing the role of benchmarking. The ability to compare their class with a national average was seen as a particularly useful gauge of progress and for some, a reassurance of their teaching. Furthermore, while it was acknowledged that a focus on the core standards detracted from other things, the learning goals measured by the tests were widely accepted. No teacher claimed to intrinsically value high test scores, yet all wanted their students to succeed in the system by achieving the results they knew them to be capable of.

### 5.2.1 Sources of teacher pressure

Unlike the individual variations seen in terms of the perceived importance of testing and the degree to which teachers internalised performance pressure, the impact of pressure and its perceived sources were found to be closely linked to school performance. These findings, explored below, further indicate that TBA pressure is not contained to high-stakes systems nor to low-performing school contexts.

**Administrative accountability** Teachers in both of the low-performing schools perceived their school principal to be under performance pressure. This, they believed, came from an external source—the school inspectorate. Yet, the impact of this on teachers varied considerably between the two schools. Despite recognising the precariousness of their schools' situation in the face of administrative accountability measures, teachers at Erasmus reported their experience of performance pressure to be relatively low. The school was performing sufficiently in standardised tests and well overall, and teachers felt confident in their work and encouraged and supported by management. Teachers at Erasmus were also generally satisfied with their levels of professional freedom, and reported to have critical discussions as a team about the value of prescribed methods and the role of testing.

Because some schools, they just do all the lessons, always follow all the lessons. And we already choose – ‘no, just a few lessons are most important.’  
(Mark, group 7, Erasmusschool)

In contrast, managerial pressures at the ‘under-performing’ Van Schurmanschool had clearly filtered down to its teachers. Respondents from the school reported the greatest results-focus and performance pressure out of the four cases; ‘it’s [student results] the most important thing I think’ (Saskia, group 8, Van Schurmanschool). All four teachers interviewed, and Saskia in particular given her responsibility as a group 8 teacher, found the ongoing inspection process to be a highly stressful affair. These teachers expressed greater feelings of insufficient autonomy than at the other schools and a sense of loss of control over teaching and tasks. Several felt that, in the face of administrative interventions, their school was pursuing a particularly narrow educational approach, and lacked a clear and collective vision, resulting in confused priorities and at times, in ill-considered, ill-implemented working methods:

...Instead of, thinking about it properly, they [the school management] implement all these things – so ‘we try this, oh it doesn’t work,’ ‘we try this, nope it doesn’t work’ (...) For me, that’s not the right way to work, but okay, I try it, I don’t try it fully, I just try it a little bit.  
(Jonathan, group 5, Van Schurmanschool)

This is comparable to the ‘superficial’ responses recorded by Diamond and Spillane (2004) in pressured probation schools in Chicago. Jonathan connected these pressures to his school’s population. He believed his workload could be reduced and his overall professional satisfaction increased in a different school.

[I’m] 75% happy. 25% thinking that there’s more, that I’m working too hard for the money I get, the hours I make (...) That could increase at a different school. And that’s probably because right now we have the inspection, and we have a lot of children who need extra attention, need the extra care, and that takes up a lot of time.  
(Jonathan, group 5, Van Schurmanschool)

More so than elsewhere, teachers at Van Schurman expressed a sense of unfairness about the accountability system: that the format of the standardised tests disadvantaged their non-native speaking students and the school was unfairly judged as ‘very weak’ before the inspector even arrived. Results were, as such, primarily attributed with instrumental value—perceived as a route back to greater autonomy and a more rounded school experience.

Conversely, the inspectorate was not once mentioned as a source of results-based pressure at the high-performing schools. Nonetheless, teachers at De Witt did speak of the stress brought about by their recent inspection, despite its ‘reward-based’ focus (i.e. the reputational promotion of the school). This stress did not

spring from test performance, but from a different aspect of the accountability agenda—documentation:

The other week, we had the inspection, so we needed to know – ‘is everything how it’s supposed to be?’ ‘Is everything on paper?’ ‘Does it look right?’  
(Ilse, group 8, Basisschool De Witt)

Indeed, this quality inspection was reported as having a significant effect on teachers’ working practices significantly, with extra demands from management to ensure thorough accounting of their work should proof be required. Yet, whilst this form of administrative pressure did result in extra work, it did not appear to impact teaching and teacher autonomy to the extent that it did in Van Schurmanschool. Rather, teachers at De Witt felt confident and clear in their school’s educational approach and reported to have an important degree of freedom in their teaching, particularly in the non-core subjects.

We have a lot of autonomy. ‘Bas’ is not watching over my shoulder; ‘oh you have to do it now, this is what is important, you should do this or that...’ No, not at all. No, no. On the contrary, for geography and history we don’t use working books anymore. Lesson books, but no working books. So we work in projects. (...) and I am one of the founders, one of the ones who had the idea.  
(Thijs, group 8, Basisschool De Witt)

**Market accountability** Teachers at the two high-performing schools perceived there to be a general, underlying pressure to maintain their schools’ strong academic reputations. They reported management to place a greater value on performance and test scores, and several felt an expectation to show a continuous trajectory of improvement. This, teachers found unrealistic, likening it to private sector expectations of year-on-year profit growth:

We are like a commercial company where you are expected to have a credit growth every year, and we’re also expected to increase the results every year... but we’re talking about humans here, it’s not like a pile of money.  
(Sara, group 6, Barlaeusschool)

Sometimes I do have the idea that I’m working for a company, working for a car factory. The quality has to be impeccable, and we don’t need any deviations, it has to be a linear process (of improvement), but it isn’t, we’re working with kids.

(Finn, group 5, Basisschool De Witt)

This reputational pressure is innately tied into the schools’ academic image and predominately instrumental logics. Basisschool De Witt was not found to experience market competition, yet maintaining its reputation as one of the best schools in the city was a pressure in itself:

In *this* class, a lot of people don't go to HAVO. A lot of children need extra instruction and they go to VMBO-T. And that is average in Holland - it's okay. But for this area here, we are 540 on Cito end-test, 60% goes to VWO... Those are our goals.

(Thijs, group 8, Basisschool De Witt)

Reputation was perhaps even more pressing to Barlaeusschool, given its struggle to attract students. This market pressure was clearly felt by the vice principal and did not go unnoticed by teachers: it was the only school where a number of enrolment strategies were highlighted. One of these strategies, as noted, was a reorientation of school practices to spend more time on non-core learning areas. According to teachers, this approach had not (yet) resulted in a reduction of performance pressure. Rather, it seems that letting go of an outcome-based *raison d'être* was proving difficult for the school. Test scores, quite possibly, have become integral to how Barlaeus views its own success, an image that is difficult to relinquish.

So you have the head, the hands and the heart and we always say we try to focus on everything, but in the end, it's a lot of head.

(Marianne, group 3, Barlaeusschool)

This in turn, appeared to affect Barlaeus' teachers' experience of autonomy, who reported to be limited by a rigid curriculum and strict working methods, particularly when it came to the core learning areas. Respondents expressed a desire to break free from the methods and to have more fun and engagement with their students. 'I would try and have more meaningful moments' (Marianne, group 3, Barlaeusschool). Unlike at Erasmus and De Witt, teachers did not mention their involvement in the selection of these methods.

**Participative accountability** In the Dutch system, standardised tests not only have stakes for schools but for students too, functioning as a key determinant of their future career path. This adds to the burden on teachers. Pressure that was described as 'self-imposed' was reported as a significant pressure source in all four schools. Not only did some teachers measure their own abilities through test results, but almost all felt a keen responsibility not to let their students down by allowing a sub-optimal performance. Even greater than this however, was the pressure teachers felt from parents. Reportedly, parental expectations were not always realistic:

Most of the parents say... 'I want HAVO, because HAVO is the best.' Yes, HAVO is the best for *some* children, but not for your child...

(Mark, group 7, Erasmusschool)

Parental pressure was present in all schools, but was reported to be a greater concern in the high-performing Barlaeus and De Witt. Interviews at Barlaeusschool in particular revealed the considerable stress felt during the period when students' secondary school advice is given. It was not uncommon that parents did not see

eye-to-eye with the teacher on this: 'I got into *so* many fights with really abusive parents' (Sara, group 6, Barlaeusschool).

Teachers at Basisschool De Witt did not feel this parental pressure as directly (or at least seemed to internalise it less), yet expressed significant concern about its effect on students. They understood it to be a result of the school's population—high-earning, well-educated parents with high expectations and considerable involvement in their child's education (see also Mathison & Freeman, 2006). Interviews frequently revealed that many parents at the school sign-up their children for (extra-curricular) test training. This was not mentioned in any of the other schools. There were also accounts of a small number of parents year buying copies of an upcoming test that had been 'leaked' one year. As well as worrying that this undermined the validity and use of the tests, teachers were significantly concerned about the stress this placed their students under. This concern was one of the key reasons why teachers in high-performing schools reported to only discuss and practice test questions in the few weeks before the test.

## 6 Discussion

The aim of this study has been to better understand the impact of test-based accountability policies by examining the experiences of teachers in differently performing primary schools in the Netherlands. More specifically, it has inquired into the value that schools and teachers attach to academic performance (and standardised testing in particular), the extent to which teachers experience performance pressure, and the impacts of this on their work. Comparing two 'low' and two 'high' performing schools, the research addresses the question: In such performative environments, (how) does school performance level affect teacher experience?

Findings show that schools operating within the same TBA structures under the same management are confronted with quite different challenges. It is clear that accountability mechanisms do not have an equal impact, rather, the four schools were found to be vulnerable to different (combinations of) pressures. The lower performing schools are, by policy design, under greater external administrative pressures than their high-performing counterparts. Echoing the experiences of probation schools in the USA (Diamond & Spillane, 2004; Mintrop, 2004), Van Schurmanschool was found to be experiencing the most immediate challenges and direct pressures due to the school-level interventions attached to underperformance, and in response was instigating various, and reportedly superficial, educational changes. Market pressure was not found to be related to school performance; the 'underperforming' Van Schurmanschool and the 'high performing' De Witt had waiting lists, whilst Barlaeusschool, also high performing, was under considerable pressure to attract new students. This has more to do with the schools' logics of action. Whereas the other schools were currently occupying a particular position within the education market (Maroy & Van Zanten, 2009), Barlaeus' position had become jeopardised. The school was attempting to adjust its educational offer as a result.

Indeed, we see considerable interschool variances regarding the ways in which the schools manoeuvre themselves in response to accountability pressures, and the significance placed on performance. ‘Instrumental’ and ‘expressive’ are, of course, ideal types, yet clear consistencies or ‘logics’ could be found, unique to each school (Ball & Maroy, 2009). Van Schurman and Erasmusschool were consistent in their emphasis on student needs, equality and diversity. This was also apparent in their organisation of education and approach to testing. In comparison, Barlaeus and Basisschool De Witt were much more explicit in the promotion of their results and perceived good performance as central to their reputation. There were further differences regarding the expected role of teachers and relational structures in the schools. Collaboration and an ‘open-door’ approach were repeatedly emphasised as central to Basisschool De Witt’s culture, yet not mentioned at Barlaeus or Van Schurman. Similarly, meaningful teacher involvement in *core* curricular decisions was only reported at Erasmus. As will be discussed shortly, the importance and impact of these logics of action on teachers appears significant, apparently explaining the unexpectedly low or high levels of pressure they feel, regardless of school performance.

Despite differing external pressures and internal logics, the considerable role of standardised testing in shaping education is evident across the schools. All four conduct standardised tests from an early age, follow prescribed methods, engage in test preparation activities, practice ability differentiation on the basis of results, shape staff meetings, school reports and teacher monitoring and evaluation around results, and keep careful records of student learning. In this way, test data used as a representation of learning and signifier of quality has become entrenched in Dutch schools as it has in high-stakes neoliberal systems (Ball, 2016; Holloway & Brass, 2018). Essentially, regardless of the ‘autonomy’ label so often afforded to Dutch schools, performance pressure, stemming from a variety of sources, restricts this freedom.

Inevitably, these pressures have, in turn, impacted teachers’ work and their own experiences of performativity and professionalism, and reflect what is best understood as a market model approach to professional regulation (Voisin & Dumay, 2020). When considered alongside prominent studies conducted in other ‘market model’ contexts, (Day & Smethem, 2009; Perryman et al., 2011; Valli & Buese, 2007), they indicate a commonality to teacher experience that stands apart from the accountability structures in place. Indeed, in the Netherlands no formal administrative or economic stakes are attached to ‘teacher performance’, yet various other mechanisms have closely tied this performance to reputation, professional identity, and students’ futures. In this way, TBA policies have changed the very identity of teaching; both from outside and inside the profession (Ball, 2016).

While some elements of TBA have become naturalised or taken for granted, frustration and dissatisfaction amongst teachers remained apparent. Prescribed and test-heavy working methods may be considered the first line offender here, yet they do not work alone. Time and scheduling pressures, exacerbated by the burden of accountability-related admin means that even teachers who described



themselves as autonomous, confessed to feeling restricted by their workload. Interschool variances were evident in teachers' experiences of their work and of performance pressure. School performance and logics of action play an important role here. Although TBA interventions in the Netherlands are designed to have a primarily informative and supportive role—contrasting with the higher-stakes sanctioning role described by Mintrop (2003) in the USA—their impact on teachers is nonetheless significant. Despite Van Schurman's primarily expressive orientation, teachers perceived there to be a substantial, even overwhelming, focus on test results. These teachers reported considerable performance pressure and restricted autonomy. In contrast, despite also finding itself at the low end of the performance spectrum, teachers at Erasmus spoke of their work satisfaction and claimed to experience relatively low results-based pressure. Furthermore, rather than finding the system unfair, the principal welcomed the comparison of her school with others. Two key differences exist here. Firstly, Erasmus school has managed to achieve consistent satisfactory results, meaning it is not subjected in the same way to external administrative pressures as Van Schurman. Secondly, while the two schools share many of the same priorities and 'expressive' orders, unlike Van Schurman, Erasmus seems to also adopt a collaborative decision-making approach, with teachers reporting a meaningful involvement in key curricular decisions. This is significant: despite being one of the schools most focused on core learning, Erasmus' teachers expressed the greatest sense of control over their work. Indeed, even in 'disadvantaged' schools where educational content may be narrowed in an effort to meet core quality standards, teachers may still be able to enjoy a meaningful sense of curricular and pedagogical freedom.

A more expressive approach to educational decision-making was also reported at De Witt. Teachers' involvement in (non-core) curricular choice appears to provide them with a deeper understanding and greater sense of ownership over what they are teaching, and with a greater chance of working with methods they feel are suited to their students. As such, policy is experienced more as something they are *doing*, rather than as something being *done to* them (see Ball et al., 2011). Despite recognising the presence of external pressures, teachers at Erasmus and De Witt believed their schools to have clear and collective missions that were not susceptible to changeable government policies and societal expectations.

Despite the 'autonomy' of the Dutch school system therefore, and despite what findings from studies focussing on *administrative* pressure may suggest, this research indicates that in a system so shaped by accountability, it is inaccurate to assume that well-performing schools and their teachers will experience no or low performative pressure. It is true that administrative sanctions or 'actions' should not trouble such schools, but pressures come from elsewhere. In the more instrumentally oriented Basisschool De Witt and Barlaeusschool, where image and market position are so dependent on performance, teachers reported to feel greater pressure from management and parents: the pressure to 'stay on top.' In De Witt, where the school and its board applied for a quality label, the importance of reputation also resulted in the stress and increased workload associated with inspection. Similarly, it is important to emphasise

that low-performing schools and their teachers will not necessarily experience high degrees of performance pressure. Due to adequate performance, a primarily ‘needs over outcomes’ focus, and an inclusive management approach, teachers at Erasmus reported to experience some of the lowest levels of pressure across the four schools.

## 7 Conclusions

This study has examined the enactment and impact of TBA at high- and low-performing primary schools in what can be described as a ‘middle-stakes’ system. It has shown that, despite the differences that exist between the Dutch and those ‘higher-stakes’ Anglo-Saxon systems explored in the paper, notable similarities in policy impacts are found. An interplay of various accountability mechanisms create pressure on different parts of the education system which is significant in steering schools’ and, inevitably, teachers’ practices and experiences. As a result, perceived performance pressures within the school may be considerable, even when administrative accountability pressures are not. Conversely, policy tools designed to exert pressure on schools may not always achieve their intended impact.

Despite adopting many of the same test-based practices, schools place different significance on performance. The nature and sources of performative pressures vary between high- and low-performing schools, yet—given ‘sufficient’ performance—it cannot be said that one context is *less* pressured than the other. While teachers in the more ‘disadvantaged’ schools felt their work to be, in many ways, more challenging and more conditioned by TBA, findings suggest that a more privileged student population is no guarantee of escape. Rather, schools and their teachers are susceptible to performative pressures in various ways and for various reasons. Overall, teachers expressed higher levels of professional satisfaction and autonomy in the schools with a clear and collective mission, reporting to be less vulnerable to changing government demands and external pressures. Further research, exploring this relationship and its direction is warranted here.

While this study does not purport to make broad generalizations, it does reveal the various pressures—shared and diverging—felt by teachers in schools that face different challenges within the context of performativity. In doing so, it adds to the growing body of research that draws attention to the assumptions and oversights on which these, now globally spread, accountability policies are based.

## Appendix

**Table 2** List of participants

Organisation	Name	Role
School board Barlaeusschool	Thomas	Director
	Yvonne	Vice principal
	Eva	Teacher: group 8
	Daan	Teacher: group 7/8
	Sara	Teacher: group 6
	Marianne	Teacher: group 3
Basisschool De Witt	Bas	Principal
	Thijs	Teacher: group 8
	Ilse	Teacher: group 8/behaviour coach
	Anna	Teacher: group 7
	Finn	Teacher: group 5
	Monique	Principal
Erasmusschool	Mark	Teacher: group 7
	Felix	Teacher: group 6
	Els	Teacher: group 2
	Annemarie	Principal
Van Schurmanschool	Saskia	Teacher: group 8
	Thea	Teacher: group 7/8 (SBO)
	Marieke	Teacher: group 6
	Jonathan	Teacher: group 5

**Funding** This work was supported by the H2020 European Research Council [StG-2015–680172].

**Data availability** Anonymized data will be available for researchers after request via the REFORMED website. These data will be embargoed until the end of the project (July 2022).

**Code availability** Atlas.ti coded. Coding protocol can be made available upon request.

**Declarations**

**Conflict of interest** None.

## References

- Agrey, L. (2004). The pressure cooker in education: Standardized assessment and high stakes. *Canadian Social Studies*, 38(3), 1–12.
- Ball, S. (2003). The teacher's soul and the terrors of performativity. *Journal of Education Policy*, 18(2), 215–228.

- Ball, S. J. (2016). Neoliberal education? Confronting the slouching beast. *Policy Futures in Education*, 14(8), 1046–1059.
- Ball, S., & Maroy, C. (2009). School's logics of action as mediation and compromise between internal dynamics and external constraints and pressures. *Compare: A Journal of Comparative and International Education*, 39(1), 99–112.
- Ball, S., Maguire, M., Braun, A., & Hoskins, K. (2011). Policy actors: Doing policy work in schools. *Discourse: Studies in the Cultural Politics of Education*, 32, 625–639.
- Bernstein, B. (1996). *Pedagogy symbolic control and identity theory, research, critique*. Taylor and Francis.
- Berryhill, J., Linney, J. A., & Fromewick, J. (2009). The effects of education accountability on teachers: Are policies too stress provoking for their own good? *International Journal of Education Policy and Leadership*, 4(5), 1–14.
- Bovens, M. (2007). Analysing and assessing accountability: A conceptual framework 1. *European Law Journal*, 13(4), 447–468.
- Braun, A., Ball, S. J., Maguire, M., & Hoskins, K. (2011). Taking context seriously: Towards explaining policy enactments in the secondary school. *Discourse: Studies in the Cultural Politics of Education*, 32(4), 585–596.
- Browes, N., & K Altinyelken, H. (2019). The instrumentation of test-based accountability in the autonomous Dutch system. *Journal of Education Policy*, 1–22.
- CNV Onderwijs & EenVandaag. (2018). *Onderzoek: 'Leraren onder druk om schooladvies.'*
- Collier, D., & Mahoney, J. (1996). Insights and pitfalls: Selection bias in qualitative research. *World Politics*, 49(1), 56–91.
- Dale, R. (1982). Education and the capitalist state: contributions and contradictions. In M. Apple (Ed.), *Cultural and economic reproduction in education*. Routledge and Kegan Paul.
- Darling-Hammond, L. (1990). Teacher professionalism: Why and how. In A. Lieberman (Ed.), *Schools as collaborative cultures: Creating the future now*. Falmer Press.
- Day, C. (2002). School reform and transitions in teacher professionalism and identity. *International Journal of Educational Research*, 37(8), 677–692.
- Day, C., & Smethem, L. (2009). The effects of reform: Have teachers really lost their sense of professionalism? *Journal of Educational Change*, 10(2), 141–157.
- Diamond, J. B., & Spillane, J. P. (2004). High-stakes accountability in urban elementary schools: Challenging or reproducing inequality? *Teachers College Record*, 106(6), 1145–1176.
- Dobbins, K. (2009). Teacher creativity within the current education system: a case study of the perceptions of primary teachers. *Education 3–13*, 37, 95–104.
- DUO Onderwijsonderzoek & De Monitor (KRO-NCRV). (2016). *Rapportage: Werkdruk leerkrachten in het basisonderwijs*.
- Hargreaves, D. (1995). School culture, school effectiveness and school improvement, school effectiveness and school improvement: An International Journal of Research. *Policy and Practice*, 6(1), 23–46.
- Holloway, J., & Brass, J. (2018). Making accountable teachers: The terrors and pleasures of performativity. *Journal of Education Policy*, 33(3), 361–382.
- Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277–1288.
- Inspectie van het Onderwijs. (2020). *Onderzoekskader 2017 voor het toezicht op de voorschoolse educatie en het primair onderwijs*.
- Inspectorate of Education. (2019). *State of education, 2020, summary report*. Inspectie van het Onderwijs.
- Maroy, C., & Van Zanten, A. (2009). Regulation and competition among schools in six European localities. *Sociologie Du Travail*, 51, e67–e79.
- Mathison, S., & Freeman, M. (2006). Teacher stress and high stakes testing. In R. G. Lambert & C. J. McCarthy (Eds.), *Understanding teacher stress in an age of accountability* (pp. 43–63). IAP.
- Mintrop, H. (2003). The Limits of Sanctions in Low-Performing Schools. *Education policy analysis archives*, 11, 3.
- Mintrop, H. (2004). *Schools on probation: How accountability works (and doesn't work)*. Teachers College Press.
- Mintrop, H. (2007). School improvement under test-driven accountability: A comparison of high- and low-performing middle schools in California CSE Report 717 Heinrich Mintrop National Center for Research on Evaluation, Standards, and Student Testing (CRESST)/Universit, 1522(310).

- Nathaniel, P., Pendergast, L. L., Segool, N., Saeki, E., & Ryan, S. (2016). The influence of test-based accountability policies on school climate and teacher stress across four states. *Teaching and Teacher Education*, 59, 492–502.
- Nusche et al. (2014). “School education in the Netherlands”. In *OECD Reviews of Evaluation and Assessment in Education: Netherlands 2014*. OECD Publishing.
- OECD (2010). *PISA 2009 results: What makes schools successful? Policies and Practices (Volume IV)*. PISA, OECD Publishing.
- OECD. (2013). *PISA 2012 results: What makes schools successful? Policies and Practices (Volume IV)*. PISA, OECD Publishing.
- Patton, M. Q. (2015). *Qualitative research and evaluation methods: Integrating theory and practice* (4th ed.). Sage.
- Perryman, J., Ball, S., Maguire, M., & Braun, A. (2011). Life in the pressure cooker – School league tables and English and Mathematics teachers’ responses to accountability in a results-driven era. *British Journal of Educational Studies*, 59(2), 179–195.
- Saha, L. J., & Dworkin, A. G. (2009). Teachers and teaching in an era of heightened school accountability: a forward look. In *International handbook of research on teachers and teaching* (pp. 1177–1185). Springer.
- Thiel, C., & Bellmann, J. (2017). Rethinking side effects of accountability in education: Insights from a multiple methods study in four German school systems. *Education Policy Analysis Archives*, 25(93).
- Valli, L., & Buese, D. (2007). The changing roles of teachers in an era of high-stakes accountability. *American Educational Research Journal*, 44(3), 519–558.
- Van Zanten, A. (2009). Competitive arenas and schools’ logics of action: a European comparison. *Compare: A Journal of Comparative and International Education*, 39(1), 85–98.
- Verger, A., & Parcerisa, L. (2017). A difficult relationship: Accountability policies and teachers. International evidence and key premises for future research. In M. Akiba & G. LeTendre (Eds.), *International handbook of teacher quality and policy* (pp. 241–254). Routledge.
- Verger, A., Parcerisa, L., & Fontdevila, C. (2019). The growth and spread of large-scale assessments and test-based accountabilities: A political sociology of global education reforms. *Educational Review*, 71(1), 5–30.
- Voisin, A., & Dumay, X. (2020). How do educational systems regulate the teaching profession and teachers’ work? A typological approach to institutional foundations and models of regulation. *Teaching and Teacher Education*, 96, 103144.
- Waslander, S. (2010). Government, school autonomy, and legitimacy: Why the Dutch government is adopting an unprecedented level of interference with independent schools. *Journal of School Choice*, 4(4), 398–417.
- Webb, R., Vulliamy, G., Hämäläinen, S., Sarja, A., Kimonen, E., & Nevalainen, R. (2004). A comparative analysis of primary teacher professionalism in England and Finland. *Comparative Education*, 40(1), 83–107.
- West, A., Mattei, P., & Roberts, J. (2011). Accountability and sanctions in English schools. *British Journal of Educational Studies*, 59(1), 41–62.
- Wilkins, C. (2011). Professionalism and the post-performative teacher: New teachers reflect on autonomy and accountability in the English school system. *Professional Development in Education*, 37(3), 389–409.
- Yin, R. K. (2009). *Case study research: Design and methods* (4th ed.). e-book, Sage (accessed 20 December 2010).

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.