Educational Assessment, Evaluation and Accountability
https://doi.org/10.1007/s11092-025-09477-6

A theoretical model of stakes and their impacts

Suzanne Rice[1] · **Therese Hopfenbeck**[2] · **Christopher Corbel**[3]

Received: 10 October 2023 / Accepted: 21 December 2025

Abstract
High-stakes testing (HST) of entire student cohorts, in which students undertake compulsory tests with an uncertain outcome but a potential benefit, is increasingly used as a policy tool by countries seeking to improve achievement. Long established in the United States, high-stakes testing programs have become prominent features of the accountability landscape in many countries. Researchers have documented the impacts of high stakes testing on schools, teachers, and students (e.g. Polesel et al., Journal of Education Policy, 29(5), 640–657. 2014); (Segool et al., Psychology in the Schools, 50(5), 489–499. 2013); (Baker et al., Education Evaluation and Policy Analysis Archives, 21, 1–71. 2013). Often such research assumes that those involved in testing will value tests and their associated stakes equally and will respond in similar ways. However, different stakeholder groups – students, parents, teachers, school leaders and policymakers – will have different stakes in the outcomes of tests, and within each of these groups, the value placed on the stakes is likely to vary among individuals. How these stakes differ between stakeholders, the role that the value placed on stakes plays, and the effects of these aspects on responses to testing, often remain unexamined. In this paper we propose a theoretical framework of stakeholders and stakes drawing on expectancy-value theory (Wigfield & Eccles, Contemporary Educational Psychology, 25, 68–81. 2000). We theorise that the impact of testing will be influenced by stakeholder motivation and expectations in relation to stakes. We argue that this framework provides a starting point for the development of a more nuanced, multifaceted exploration of the impacts of testing, one that can encompass different stakes and differences in the value stakeholders place on them.

Keywords Test consequences · High-stakes tests · Test-based accountability · Large-scale assessments

Extended author information available on the last page of the article

Published online: 20 February 2026

1 Introduction - what is at stake in testing?

Influential concepts in educational assessment may have generally accepted meanings, yet remain contested within the field of testing itself, depending on the underlying assumptions on which they are based (Elwood et al., 2017; Pastore & Andrade, 2019). Testing of a high-, medium- or low-stakes nature fits this category. While the subject of continuing debate, frequently to do with aspects of *testing* (e.g. Elwood et al., 2017), one of the central concepts, the notion of *stake*, frequently remains unexamined. The purpose of this discussion paper is to examine critically the notion of stakes in testing, and propose a theoretical model of stakes, the value that might be placed on them by stakeholders in the testing regime (principals, teachers, schools, parents and policymakers) and how values and stakes are likely to interact to create the flow-on effects that have been documented in testing regimes. To our knowledge, there is no current broad theory that allows these relationships to be examined and understood. This in turn limits the capacity for more nuanced policy and practice to be developed; such policy and practice would minimise the negative effects of testing regimes, while enhancing the positive effects.

Because of its potential impact on outcomes for students, schools and teachers, high-stakes testing has been the focus of much research. High stakes testing has been defined in various ways. Johnson et al. (2008) argue that a test or testing program is high stakes when it has consequences for student grade promotion or graduation, teacher accountability, the reputation of schools or their funding. According to Au (2008), testing programs are high stakes when performance in such tests leads to some form of sanction or reward for students, their teachers, school administrators, schools, local school regions or systems. In addition to these direct rewards or sanctions attached to test performance, some authors have argued that where school results are made public and there is capacity for students to change schools, the tests become high stakes for schools as the action of parents transferring their of children to schools performing well in the test effectively punishes those schools perceived to have performed poorly (Ball, 2008; Howe et al., 2001; Lobashcher, 2011; Lingard, 2010). Implicit in all these definitions of high stakes testing is the notion of the stake: what is at risk or of value to those involved.

A stake has been defined as "A share or interest in a business, situation or system", or "A sum of money or something else of value gambled on the outcome of a risky game or venture" (Oxford online). Cambridge online dictionary notes that, "If you have a stake in something it is important to you because you have a personal interest or involvement in it". A common connection to a financial element is captured by two other definitions in Cambridge: "The amount of money that you risk on the result of something such as a game or competition." "In an activity or competition, the stakes are the reward for the person who wins or succeeds in it" (Cambridge online). A *stakeholder* is thus "A person with an interest or concern in something, especially a business" or "A person, such as an employer, customer or citizen who is involved with an organisation, society etc. and therefore has responsibilities towards it and *an interest in its success*" (Oxford online, our emphasis).

The essence of these dictionary definitions of *stake* and *stakeholder* is the sense that something of value is put at risk in a process with an uncertain outcome but a

potential benefit for one or more persons involved. That which is put *at stake* has an effect on the person(s) doing the staking, who clearly has an interest or involvement in the outcome. But the nature of the stake or stakes, and the range of stakeholder or stakeholders, is not usually specified beyond the stated or implied interests of gamblers, investors and/or parents.

To summarise, there is often a generic examination of tests that are characterised as high stakes without any serious examination of what the stakes are, to whom those stakes are likely to matter, and to what degree. This is important because the reaction of various stakeholders to a given test is likely to depend on the stakes associated with the test, and more importantly, the value that each stakeholder attaches to a given stake. Further, it matters in both policy and practical terms, because many of the stakes associated with testing (class promotion, publication of school results) are the result of policy decisions at the system or school level, and it is here that there is capacity to influence the impact testing has on students, teachers and schools.

Expectancy-value theory (Wigfield & Eccles, 2000) would suggest that the way in which key actors or stakeholders respond to testing regimes, and in particular, the motivation to do well or for school staff and parents to encourage students to do well, will be affected by the value a student, teacher or school leader places the stakes attached to test performance and their expectation that they are able to perform well. Expectancy-value theory is a well-established theory that has been extensively explored in relation to learning and motivation (see Rosenzweig et al. (2019) for a summary), and career pathways and choices (e.g. Rice et al., 2024). Much of the discussion on testing, particularly at a policy level, assumes that stakes are equivalent for all actors: that if schools place a lot of importance on a test, then students will as well, and that negative reactions such as stress will be the same across all groups. While there has been some use of expectancy-value theory to consider student motivation and effort during tests (e.g. Dowley & Rice, 2022; Penk & Richter, 2016) there appears to be little exploring the relationships with stakes.

This discussion paper provides a more systematic examination of the impacts of the stakes associated with testing by examining in detail what is at stake, for whom, and with what likely effects, in testing programs, with a particular focus on high-stakes testing. It begins with an overview of the current research surrounding the impact of high-stakes testing. This overview is used to create a typology of impacts on high-stakes testing on stakeholders.

However, not all impacts are stakes and not all stakes are the same for all stakeholders. Consequences among stakeholders vary in relation to the degree of achievement in the test, so not all stakes are valued equally by all stakeholders. A framework is presented that captures the potential stakes, stakeholders and the consequences of poor test performance for those most affected, those in schools.

Expectancy-value theory is then used as a basis for examining ways in which different stakes may be valued by different stakeholders and the relationship between values, performance and stakeholder responses. Various stakes are likely to be valued differently by stakeholders. The reputation of the school is likely to be valued highly be the principal, for example, but potentially less so by the students, particularly if they are very young. There are also likely to be different values within groups of stakeholders as well, with some teachers, for example, valuing the school reputation

more than others. If a test is used to determine entry to a course, those students wanting to enter the course are likely to place a strong positive value on doing well in the test, in contrast to those not seeking course entry.

The paper concludes by considering the implications of the discussion for future research, policy and practice. It argues for a more nuanced approach to examining the impact of testing programs on schools and students, one that takes into account the values of differing stakes to differing stakeholders.

2 Background and method - the impact of high stakes testing

This perspective article draws on the collective expertise and diverse professional experiences of the author team, comprising academics with backgrounds in teacher education, psychometrics, test development, and national and international assessment systems. Our perspective is informed by the authors' direct involvement in both high-stakes and low-stakes assessment research across multiple international contexts, including large-scale implementation studies, policy evaluations, and classroom-based investigations spanning Australia, Norway, the United Kingdom, and beyond.

Our approach combined narrative review methods with expert consultation. Initial literature searches focused on identifying key review articles and empirical studies examining the relationship between high-stakes and low-stakes assessment practices, with particular attention to formative assessment implementation in different policy contexts. We conducted iterative searches following citation trails from foundational texts and recent publications, prioritising work that addressed high and low stakes tests. Article selection was guided by the team's collective expertise in assessment theory, psychometric methods, and educational policy implementation, allowing us to identify and synthesise literature most relevant to understanding how different assessment purposes and stakes might support or challenge teacher-led formative assessment in secondary school contexts. This selective, expertise-driven approach enabled us to critically examine current scholarship while maintaining focus on practical implications for classroom practice and teacher professional learning.

High stakes testing has generated concern among some educators and researchers because of its perceived negative impacts on students and schools, and there is a very significant body of literature that explores these impacts (e.g. Abrams, 2004; Ashadi & Rice, 2016; Au, 2008; Boguslawski et al., 2021; Cunningham & Sanzo, 2002; Polesel et al., 2014; Madhaus et al., 2009; David, 2011; Gregory & Clark, 2003; Thompson & Cook, 2014; Yui, 2020). However, it should be noted that positive impacts have also been found, including increased student learning and motivation (Klapp, 2015; Phelps. 2012), and student excitement in meeting the challenges associated with testing (Dowley and Rice, 2022). A summary of both the positive and negative impacts noted in the literature is outlined in Table 1 below. We then narrow our focus to those impacts particularly affecting school leaders, teachers and students.

In our review of the research literature, we identified six categories of stakeholders who have been identified as having an interest in, and potentially being impacted by, high stakes testing: students, teachers, school leaders, policymakers, parents

Table 1 Documented impacts of high stakes testing on students, teachers and school leaders

Potential stakeholder/ impact	Students	Teachers	School leaders
Curriculum	Narrower curriculum with less access to non-tested subjects and topics (Polesel et al., 2014) Learning to the test (Elwood et al., 2017)	Narrower curriculum with less access to non-tested subjects and topics (Berliner, 2009, 2011; Boguslawski et al., 2021; Polesel et al., 2014; Thompson & Harbaugh, 2013)	Narrower curriculum with less access to non-tested subjects and topics (Berliner, 2009, 2011; Boguslawski et al., 2021; Polesel et al., 2014)
Teaching and learning approaches	Surface learning approaches (Yerdelen-Damar & Elby, 2016). Test-focused learning strategies (Damankesh & Babaii, 2015)	Use of teacher-centred pedagogies (Avalos et al., 2020) Focus on drilling/test practice/ test skills (Polesel et al., 2014; Tóth & Csapó, 2022) Focus on lower-order thinking (Thompson & Harbaugh, 2013; Zohar & Alboher Agmon, 2018).	
Achievement	Increased achievement in tested areas (Phelps, 2012)		
Motivation, engagement and morale	Reduced intrinsic motivation (Demir & Keles, 2021) Reduced engagement, potentially leading to dropout (Menken, 2016) Increased morale where performance is better than expected (Klapp, 2018) Reduced motivation for low-achieving students (Klapp, 2015) Excitement (Dowley & Rice, 2022).	Reduced professional satisfaction/morale (Polesel et al., 2014; Smith & Holloway, 2020) Reduced motivation (Demir & Keles, 2021).	
Professional opportunities		Less access to professional learning for teachers in untested subjects or year levels (Ashadi & Rice, 2016).	
Professionalism		Reduced sense of professionalism (Smith & Holloway, 2020, Garver, 2019) Staff stratification within a school (Ashadi & Rice, 2016).	Reduced staff collaboration
Stress and anxiety	Stress through pressure to perform (Banks & Smyth, 2015; Heissel et al., 2021; Rice et al., 2016; Segool et al., 2013, 2014	Increased anxiety/sense of pressure & reduced morale (Garver, 2019; Tóth & Csapo, 2022)	
Personal resources	Less free time due to tutoring or additional classes (Kwon et al., 2017)		

Table 1 (continued)

Potential stakeholder/ impact	Students	Teachers	School leaders
Equity	Justification of current educational inequities (Au, 2011) Focus on borderline students to the detriment of high and low achievers (Welsh et al. 2019)	Reconceptualisation of the "good teacher" (Thompson & Cook 2014).	

and the broader community. Teachers were the group most frequently considered and research often focused on those who taught students preparing for the test (e.g. Learned et al., 2020), but included their colleagues as well (e.g. Ashadi & Rice, 2016; Boguslaski et al., 2021). Students were the second strongest focus, perhaps reflecting the call of Polesel et al. (2014) for a greater focus on their well-being. School leaders were a third category and included principals but also those in the school with responsibility for making decisions about curricula, pedagogy and resources. Policymakers were the fourth group identified, those with responsibility for setting the frameworks for education within a school system or district. Parents were the fifth identified stakeholder group, however, while parents clearly have an interest in testing regimes and their outcomes, there appears to be scant research on the impact of high stakes testing regimes on parents. Finally, the literature identifies the broader community as one potential stakeholder in high stakes testing, but very little research focuses on either what they want from testing regime, or any impact (see Rice et al., 2012a, b, for an example).

Our review of the literature identified a range of impacts of high-stakes testing regimes on stakeholders, although the research overwhelmingly focuses on the first 3 groups: students, teachers, and school leaders. We classified these impacts into the areas of curriculum, pedagogy, achievement, motivation, engagement and morale, funding, professional opportunities, sense of professionalism, stress and anxiety, personal resources, and equity. These impacts are summarised in Table 1 below.

Negative impacts on students noted in the table include being offered a narrow curriculum (Polesel et al., 2014) and learning opportunities that are tailored to test content, use surface learning approaches and employ test-focused teaching (Elwood et al., 2017; Yerdelen-Damar & Elby, 2016). Negative impacts on student motivation and engagement have been identified (Demir & Keles, 2021; Menken, 2016), particularly for low-performing students (Klapp, 2015), together with increased student stress (Banks & Smyth, 2015; Heissel et al., 2021; Rice et al., 2016; Segool et al., 2013, 2014). Positive impacts may include increased achievement (Phelps, 2012), and enhanced motivation for students who achieve more highly than they anticipate and for those who enjoy testing (Dowley & Rice, 2022; Klapp, 2018).

Documented negative impacts on teachers of high-stakes testing regimes include a narrower curriculum with less emphasis on non-tested subjects (Berliner, 2009, 2011; Boguslawski et al., 2021; Polesel et al., 2014; Thompson & Harbaugh, 2013). Ashadi and Rice (2016) found that teachers not teaching tested subjects or year levels have reported a status stratification among staff, with those teaching tested subjects in a privileged position, and teachers not teaching tested subjects or year levels given

less access to professional learning than their peers. Teachers also report using narrower, more teacher-centred pedagogies, with a focus on drilling and surface learning (Avalos et al., 2020; Polesel et al., 2014; Thompson & Harbaugh, 2013; Tóth & Csapó, 2022 Zohar & Alboher Agmon, 2018).) and that testing regimes have reduced their professional satisfaction or morale (Polesel et al., 2014; Smith & Holloway, 2020). Other studies have reported that high-stakes testing regimes have reduced motivation and the sense of professionalism among teachers (Demir & Keles, 2021), with some teachers reporting increased anxiety (Garver, 2019; Tóth & Csapo, 2022). The negative impacts of testing on school leaders have been less examined, but include the offering of a narrow curriculum, and reduced staff collegiality (Berliner, 2009, 2011; Boguslawski et al., 2021; Polesel et al., 2014).

The impacts outlined in Table 1 have been the focus for much of the research on high-stakes testing. However, the impact of any test will depend on the stakes attached to performance on the test, and the value placed on the test by each of the stakeholders. We now move to propose a theoretical model of how stakes, values and expectancies may be related to the impact of high-stakes tests on the key stakeholders.

3 A hypothesised model of the relationships between stakes, motivations, expectancies and testing program impact

Tests in and of themselves are a routine part of the educational experience of school students. However, in high-stakes testing regimes it is the decisions of policymakers (including politicians) and school staff that determine most of the stakes associated with any given test or tests for each of the key stakeholders, ranging from merit pay eligibility for teachers, through to grade promotion or course access for students.

Where a stake is assigned to a test result, expectancy-value theory (Wigfield & Eccles, 2000) would suggest that individual stakeholders will place a value on that stake according to its importance to them. The theory would predict that a student who wishes to gain access to a higher course stream such as advanced maths will place a higher value on doing well in the maths test that determines stream placement than a student who has no desire to undertake advanced mathematics; principals who are concerned about falling enrolments in their school are likely to place a higher value on their students doing well in state tests for which results are published, if they believe poor results will lead to further drops in enrolments, than principals whose schools are oversubscribed.

Expectancy-value theory would predict that value alone is not enough to predict motivation to take actions to improve performance here. Instead, values placed on the stake will interact with expectancy beliefs. Where a person places a high value on doing well, and that this value is accompanied by a belief in a capacity to do well or to improve performance (expectancy), motivation and effort will be high. Applied to high-stakes testing programs, we would argue that where a stakeholder values the stake, and believes that a specific action (such as test practice or increased curriculum focus on tested areas) will lead to either a high-level performance or an improvement in performance, then motivation to implement these actions will be strong.

In contrast, where the value assigned to the stake is high, but there is little expectation that an action will lead to high-level performance or an improvement in performance, the motivation to implement that action is likely to be low.

in summary, low value placed on a stake by a stakeholder is likely to be associated with a lack of action to achieve highly or improve performance on the test, even if there is a belief that the action will be effective. High value placed on the stake is also likely to be associated with a lack of action to achieve highly or improve performance on the test, if the stakeholder believes the action will probably be ineffective. Action will be taken by stakeholders to perform well or lift performance where there is the combination of high value placed on the stake, and high expectancy that a given action will be effective.

It is these actions that are frequently the documented impacts of high-stakes testing. By actions we mean specific changes to individual or group behaviour, practice and/or policy at the local level that influence students' and teachers' experiences of education. Examples of such actions include:

- Shaping the curriculum to focus on tested subjects or materials (Avalos et al., 2020; Berliner, 2009, 2011; Thompson & Harbaugh, 2013).
- Resorting to "safer" pedagogies such as drilling and whole class instruction (Thompson & Harbaugh, 2013; Zohar & Alboher Agmon, 2018).
- Conducting practice tests (Polesel et al., 2014).
- Assigning teachers who are better qualified, more experienced and/or seen as better teachers to examined year levels (Ashadi & Rice, 2016).
- Cheating (Amrein-Beardsley, Berliner & Rideau, 2010).
- Excluding poor-performing students (Hofflinger & von Hippel, 2018).
- Skewing school resources to students taking the tests (Yiu, 2020).

Figure 1 below summarises our hypothesised model of impact from high stakes tests. We explore this figure in the text below it.

We propose that within testing regimes, tests may have stakes attached to them, ranging from access to courses at a student level to school reputation at a system level. Such stakes are determined by decisions at a policy or school level, for example, a secondary school setting a specific maths test mark as the requirement to enter the advanced maths stream in senior secondary. In this example, entry to the course would be the stake. We then propose that such a stake will be attached a value that is likely to vary among stakeholders. For the student who dreams of undertaking an engineering course at university for which advanced maths is a prerequisite, the value assigned to the test is high. In contrast, for the student who is keen to enter a university course focused on literature, with no maths requirement, the value assigned is likely to be low.

Simultaneously, stakeholders will hold beliefs (expectancies) about whether they are able to improve their performance and/or reach the requisite level of the performance on the test. Such actions may be legitimate (studying more) or conversely may affect the integrity of test results (such as where schools encourage low-achieving students to be absent from school on testing days to improve the school's results) (Save Our Schools, n.d.) Where such expectancies are low (the student, for example,

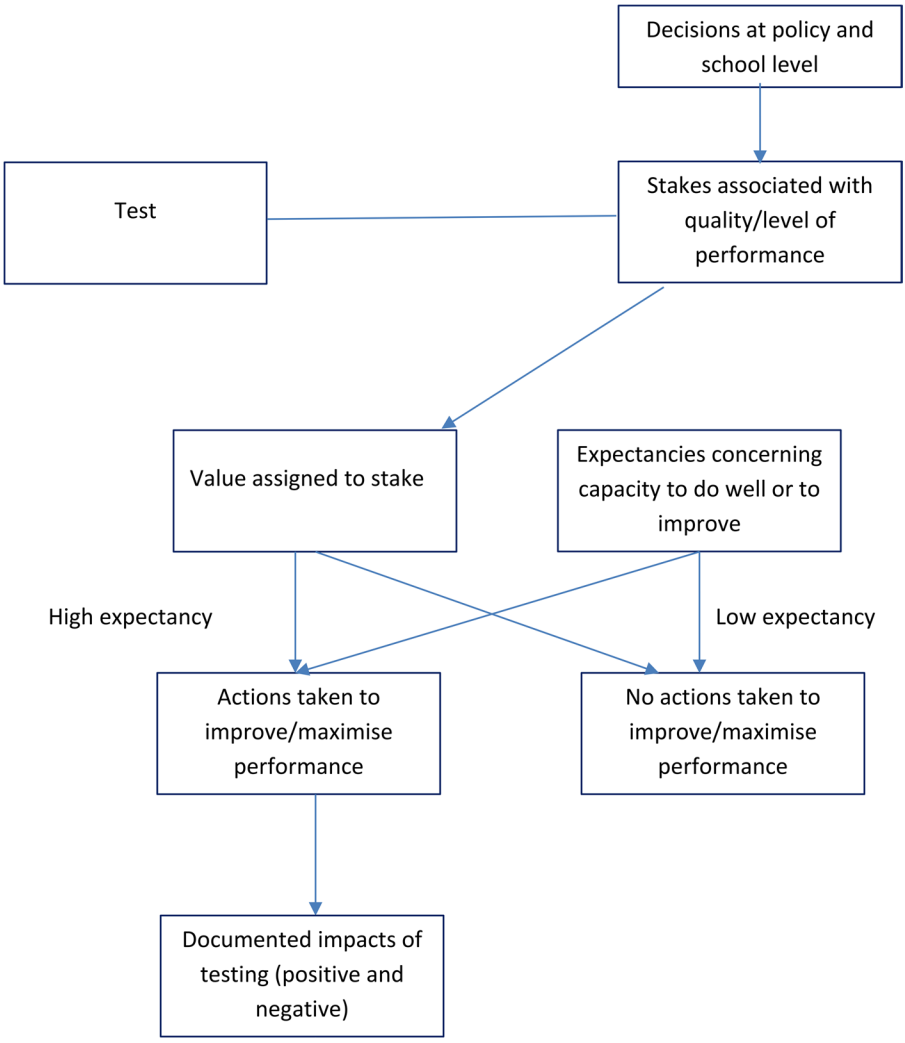

Fig. 1 The proposed relationships between stakes, values, expectancies and test impacts

believes they are poor at maths and that nothing they do will alter this) they are unlikely to change behaviour in the lead up to the test, even if they place a high value on doing well. Similarly, a stakeholder who places a low value on doing well in a test (such as our hypothetical literature student) is also unlikely to alter behaviour in preparing for the test, or to feel much anxiety in relation to it. Actions taken to maximise performance are unlikely to be taken. However, where the value placed on doing well is high *and* the expectancy that changed behaviour will lead to the requisite performance both exist, behaviour is likely to be changed and actions are taken to maximise performance. The student seeking entry to the engineering course may put in many hours of additional study, seek extra advice from their maths teacher, watch mathematics videos online, undertake practice tests, and even seek outside tutoring.

We argue that these actions drive the impacts of testing regimes that have been documented in the literature as summarised in Table 1.

4 Stakes, stakeholders and the consequences of high stakes testing

The model outlined in Fig. 1 predicts that the way in which stakeholders respond will respond to a stake will be a function of the value placed on the stake by the stakeholder, and the expectancy that a given action will lead to high individual or group performance, or an improvement in performance. What, then, are the types of stakes that have been associated with performance in testing programs, and which stakeholders are likely to place a high value on these stakes?

We identified from the literature a range of stakes associated with tests, and we classified these as relating to opportunity, position, financial rewards or identity. In addition, there can be sanctions associated with poor performance, a form of negative stake, which motivates effort in order for the stakeholder to avoid the sanction. Table 2 below summarises the literature on the stakes that have been associated with test performance. Note that in some cases, stakes are based on perceptions of the stakeholders (for example, the reputation of the school or principal) rather than on empirical evidence linking test results to the stakes. We also include some stakes as being possible and hypothesise that they are likely to exist in some contexts (for example, a teacher's promotion or lack thereof depending on test results) where there is as yet no empirical investigation of the matter.

For students, stakes can include course access, particularly in the final years of secondary school, access to qualifications, self-esteem as a student and respect among peers, or financial or other rewards for performance from parents. In some education systems, grade promotion may also depend on test performance. For teachers, stakes may include access to professional learning or to teaching higher status student streams (such as classes for the gifted and talented), access to promotion or tenure, the provision of merit pay (particularly in the United States), and professional self-esteem and reputation.

For school leaders, stakes are likely to include the reputation of the school ("a good school because it has good results") with flow-on effects to student enrolments and staff numbers. A school's reputation as either an academically high-achieving school or a poor-performing school may influence the attraction and retention of staff, and may also have an impact on parental support for the school. In some education systems, merit pay for school leaders attached to school test results may be a powerful stake.

5 The effects of stakes on stakeholder performance

We argue that, according to the pathways described in Fig. 1, stakeholder responses are likely to vary depending on the value placed by any stakeholder on the stake, and their expectation that they can achieve to the desired level on a high-stakes test, or that a particular action (such as increased study time or test practice) will increase

Table 2 Real, perceived and possible stakes for students, teachers and school leaders associated with performance in high-stakes tests

Stakeholder Stake	Students	Teachers	School leaders & the school
Opportunity	Access to specific courses or streams at school or tertiary level	Access to professional learning opportunities (Ashadi & Rice, 2016)	Maintained or increased enrolments (Polesel et al., 2014)
	Access to qualifications (hurdle requirements)		Maintained or increased staff numbers
			Ability to attract and/or retain quality staff through reputation (Lewis & Hardy, 2015; Polesel et al., 2014)
Position	Grade promotion	Access to teaching high status classes	School reputation (Lewis & Hardy, 2015; Polesel, Rice & Dulfer)
		Promotion to leadership roles	Parental support (Polesel et al., 2014)
		Access to tenure/ongoing statues (Lavigne, 2014)	Maintenance of teaching expertise as teachers stay
Financial rewards	Financial rewards from parents	Merit pay (Pham et al., 2021; La Londe, 2017)	Merit pay (Lavigne, 2014)
			Personal performance bonuses
Identity	Self-esteem/ self-concept as a student	Self-esteem/ self-concept as a teacher	Staff morale across a school
	Status among student peers	Professional status with other staff	Staff cooperation

their performance on the test. We provide an outline of what we believe the likely effects of stakes on stakeholder responses will be depending on the stakeholder value (V) and the expectancy (E) in Table 3 below.

Where a stakeholder such as a student places a high value on doing well and believes that particular actions will maximise performance, there will be a strong motivation to undertake such actions. Where a teacher is under pressure to raise test results (thus placing high value on their students doing well) and believes that test practice will maximise the performance of their students they are highly likely to

Table 3 Predicted effects of stakes on stakeholder responses

Value placed on stake (V)	Expectations of capacity to do well or to improve through a given action(E)	Likely stakeholder response	Examples
High	High	Take action to maximise performance	School leaders: Reduce time spent on non-tested subjects Teacher: Teach to the test Student: Private study Parent: Provide tutoring
High	Low	Anxiety/emotional distress Apathy Avoidance/reduced engagement Rationalisation to reduce the importance of the test System gaming	School leaders: Reclassifying students to prevent low achievers sitting the test Teacher: Providing answers or hints to students Student: Disengagement Parent: Underplaying importance of test to child
Low	High	No action/low motivation	School leaders: Continue current practice Teacher: Continue current practice Student: Continue current practice Parent: Continue current practice
Low	Low	No action/low motivation	School leaders: Continue current practice Teacher: Continue current practice Student: Continue current practice Parent: Continue current practice

implement such a practice. Similarly, a principal who is aware that their school will face funding consequences for poor student test performance is likely to place a high value on students doing well. If this principal believes that focusing school effort and time on tested subjects (at the expense of non-tested subjects such as art or sport) will improve student performance, their actions in the school will likely reflect this belief.

Where a stakeholder places a high value on doing well in a test, but has low expectations for doing well, there will most likely be a negative emotional response. Such responses might include anxiety or distress, apathy, avoidance (such as skipping school on the day of the test) or a rationalisation to reduce the importance of the test and thus alleviate any emotional distress ("It's not that important anyway"). If the stakeholder is a student, motivation and effort in the testing situation is likely to be low.

If a stakeholder places a low value on a stake, the level of motivation to apply effort or change behaviour is also likely to be low, regardless of whether there is a high expectation of doing well or a low expectation. It is also possible that a stakeholder who has high expectations of doing well experiences positive emotions in relation to the test, whether or not they place a high or low value on doing well. Where the value placed on the stake by a particular stakeholder is low, we predict that the maintenance of current practice is likely to be the outcome.

As noted, the stakeholder responses outlined above constitute many (although not all) of the flow-on effects of high stakes testing documented in the research literature. It should also be noted that while we have focused in our discussion on the implications of the theoretical model for schools and students, the theory is equally applicable to other situations in which testing is used. Universities, for example, have stakes they attach to performance in tests (for example, a certain mark may be a prerequisite for entering the honours stream or gaining a scholarship) and test performance is likely to influence job opportunities for graduates. In response, university students will place different values on the attached stakes and have expectancies about their capacity to improve performance and/or reach the required standard and, where values and expectancies and high, will take actions to maximise their performance.

6 Towards a more nuanced approach to understanding the impacts of high stakes testing

The nature of the stakes, how they are implemented and monitored, and who decides on these stakes, are often rendered both unitary and invisible. The multi-faceted nature of stakes and the responses of stakeholders to them are only just beginning to be explored in the research, although they have important implications for testing practice and for validity. Our paper proposes a theoretical framework that will support the examination of the relationships between stakes, stakeholders, value, effort, and potential consequences or flow-on effects. This framework draws on motivational concepts from psychology through expectancy-value theory (Wigfield & Eccles, 2000), and applies them to a broad network of actors in the policy landscape. While the focus has been on high-stakes testing programs because of the extensive documentation on their effects, the theoretical framework is relevant for understanding the

impact of medium- and low-stakes testing, and indeed will allow an exploration of whether the classification of certain tests as low- or medium-stakes by researchers or policymakers is consistent with the perceptions of the relevant stakeholders.

The nature of the stakes, and how they are perceived and valued by stakeholders, are important for several reasons. First, the stakes and the values placed on them by stakeholders are likely to be the central influence on the reactions, emotional responses and motivations of stakeholders. Most importantly, they are likely to shape the actions taken by stakeholders that constitute or contribute to many of the documented negative impacts or flow-on effects of high-stakes regimes. Second, they are important because policymakers, principals, teachers and sometimes parents influence or determine the stakes attached to the quality of performance in tests. Policymakers decide whether test results will be used to judge schools or evaluate teacher performance; school leaders and teachers determine whether tests results are used to gain entry to subjects or streams, parents may offer rewards or sanctions for test performance. Decisions about whether results will be published, sanctions or rewards associated with performance (at system, school, class or student level), conversations and framing around levels of performance will produce different flow-on effects. Foregrounding stakes and developing our knowledge of their relationship with responses to testing may allow the shaping of policy and practice to minimise the negative flow-on effects of testing and maximise the positive ones.

They are also important because the nature of stakes and their relationship with student motivation and effort have implications for validity. Testing programs are premised on the notion that stakeholders (school leaders, teachers, students) will respond to stakes by, first, aligning their efforts to the tested curriculum (Supovitz, 2009), and second, increasing motivation and effort (Munter & Haines, 2019). There is an underlying assumption that the stakeholders will place a high value on stakes and respond accordingly, and that all groups will respond in similar ways. This assumption is particularly important for the validity of the use of the outcomes from tests by policymakers to set improvement agendas and monitor student progress, but also for the appropriateness of using test results for school or teacher accountability.

Those in favour of high-stakes testing programs often position data from the tests as more "objective" than internal school tests and teacher judgements (Phelps, 2006). For example, the Scottish National Standardised Assessments website states that "The SNSA and the MCNG were introduced as part of the National Improvement Framework for Scottish Education to provide teachers with *objective*, comparable information about progress, which will help improve outcomes for all children and young people." (Scottish Government, n.d., our emphasis). Eklöf (2010) notes that embedded in high-stakes testing programs is a presumption of optimal performance, i.e. that students will do their best. However, if students perceive a state or national test to be low stakes or do not value any stakes attached, they are unlikely to expend effort to do well. If this is the case, national and state test results may be a poor reflection of their actual learning. Student results on such tests may be a less accurate representation of their knowledge than internal school tests for which there are more immediate consequences that matter to the students, such as access to courses or streams, or reports to parents.

There is a growing body of research that examines test-taking motivation (TTM), and the likely implications for the validity of test results. There is considerable evi-

dence now that students with high motivation outperform equivalent students with low motivation (Baumert & Demmrich, 2001; Cole et al., 2008; Eklöf et al., 2014; Thelk et al., 2009; Wise & DeMars, 2005). Slim, Must, Täht, & Pedaste (2020) found motivational indicators for high stakes tests added 15% predictive power to their structural equation model of test performance, suggesting non-trivial effects. Le (2020) raised questions around the use of testing programs for accountability and policy purposes given their meta-analysis demonstrated that monetary incentives raise student test scores in mathematics, suggesting lower-than-optimal effort in some test situations.

However, as yet there has only been a small amount of research that considers the degree of motivation and effort students expend in testing programs and the relationship with stakes. Hopfenbeck and Kjærnsli (2016) examined Norwegian students' motivation for PISA (a low-stakes test as there are no direct consequences for students) and found students reported relatively high levels of motivation, although they did not compare student motivation levels with motivation in internal school tests. Dowley and Rice (2022) compared secondary students' motivation for internal school tests and those for the Australian testing program, NAPLAN, which has relatively low stakes for students, but high stakes for schools. They found students were significantly more likely to report high levels of motivation and effort for internal school tests than for NAPLAN.

An examination of stakes, values and expectancies is also likely to shed explanatory light on some of the discrepancies in findings on the impacts of high stakes testing. For example, while some researchers have found that high-stakes tests generate high levels of test anxiety among students (e.g. Banks & Smyth, 2015; Howell, 2017; Segool et al., 2013) other research has found that little evidence of negative impacts on student emotional wellbeing (e.g. Dowley & Rice, 2022; Rogers et al., 2016; Whitney & Candelaria, 2017). These discrepancies may well be underpinned by differences in the stakes or perceived stakes associated with the tests, and the value students, teachers and school leaders place on doing well. A closer examination of stakes, values and expectancies is also likely to provide explanatory power where the stakes generated by different policy actors are different for the same test. For example, Hopfenbeck (2016) found that the national tests in Norway were introduced with low stakes, but certain Norwegian municipalities used them to hold schools accountable for their results, so the purpose and the stakes of the tests had shifted for some schools but not others (Hopfenbeck, 2016). Our framework predicts that this is likely to have an effect on the flow-on effects of testing.

Finally, foregrounding the value placed on stakes by individual stakeholders (this student or teacher) and groups of stakeholders (teachers generally) has the capacity to bring greater rigour to the research on the impacts of testing. At present, classifying tests as high-, medium- or low-stakes appears to be as much a personal judgement of the researcher or policymaker as anything, with stakeholders and their values often ill-defined. There appears to be little other than perception to determine whether a test or testing program is high-, medium- or low-stakes. Researching values attached to stakes for individual and group stakeholders will allow for a stronger empirical basis on which judgements about stakes can be made, and more accurate research conducted.

Much important research has been done examining the impact of testing on students, teachers and school leaders. We have outlined a theoretical framework for examining the relationships between tests, stakes and stakeholder responses, and believe the time is ripe for greater nuance in the research on the uses of test results, the associated stakes, how these stakes are valued and the consequent actions and emotional responses of stakeholders. Further research is needed to examine the following questions in greater detail:

- What types of stakes are associated with various tests and testing programs for the broad range of stakeholders involved in testing, and how might they be classified? How well do the stakeholders understand the stakes associated with tests? What are their perceptions about how test results are used and the potential impact on them? How accurate are these perceptions? How does this affect their responses to testing?
- What value is placed on the stakes by stakeholders, and how does this vary both across and within stakeholder groups? Do different groups of stakeholders value stakes differently?
- What expectations do students hold about their test performance and how are these related to their motivation in the test? How are student expectations and values related to other responses to testing (e.g. student anxiety)?
- What beliefs do teachers, principals, policymakers and other stakeholders such as parents hold in regard to their capacity to influence student performance on tests? How is this related to the actions they take and the documented positive and negative flow-on effects of testing?
- How are stakes determined and by whom? How can decisions about stakes be better informed by more nuanced information about likely impacts?

Research to identify stakes from the perspectives of different groups of stakeholders needs to be undertaken and used as the basis of a more rigorous classification of stakes - we have outlined those stakes identified in the literature in Table 2, but as noted, there is a paucity of literature on stakes, and there are doubtless other stakes that are yet to be identified. Both qualitative and quantitative research studies are needed to identify what stakeholders perceive to be the stakes associated with test performance, the value that they place on such stakes, their expectancies to perform well, stakeholder responses to testing, and the relationships between these variables. There is a rich field of work yet to be carried out.

Findings from research that explores the questions we outline will provide a better foundation for determining the validity of student outcomes from testing programs, and thus exploring the consequential validity of the ways in which test results can and should be used. It will also enable the development of better testing practices that support high-quality teaching and strong student learning outcomes, while reducing some of the documented negative impacts of high-stakes testing.

Author contributions Author 1 led conceptualization of the work, theoretical work and the writing of the paper. They contributed to literature searching and revision. Author 2 contributed to conceptualization of the work, theoretical work and the writing of the paper. They contributed to literature searching and revi-

sion. Author 1 contributed to conceptualization of the work, theoretical work and the writing of the paper. They contributed to literature searching and revision.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions.

Data Availability This paper is theoretical and there are no data sources associated with it.

Declarations

Competing interests The authors have no competing interests, financial or non-financial, to declare. There was no funding associated with this research.

References

Abrams, L. M. (2004). Teachers' views on high-stakes testing: Implications for the classroom. *Policy Brief: Education Policy Studies Laboratory. Arizona State University College of Education*.

Amrein-Beadsley, A., Berliner, D., & Rideau, S. (2010). Cheating in the first, second and third degree: Educators' responses to high-stakes testing. *Education Policy Analysis Archives*, *18*(14).

Ashadi, A., & Rice, S. (2016). High stakes testing and teacher access to professional opportunities: Lessons from Indonesia. *Journal of Education Policy*, *31*(6), 727–741. https://doi.org/10.1080/02680939.2016.1193901

Au, W. (2008). Between education and the economy: High-stakes testing and the contradictory location of the new middle class. *Journal of Education Policy*, *23*(5), 501–513.

Au, W. (2011). Teaching under the new taylorism: High-stakes testing and the standardization of the 21st century curriculum. *Journal of Curriculum Studies*, *43*(1), 25–45. https://doi.org/10.1080/00220272.2010.521261

Avalos, M. A., Perez, X., & Thorrington, V. (2020). Comparing secondary English teachers' ideal and actual writing practices for diverse learners: Constrained professionalism in figured worlds of high-stakes testing. *Reading & Writing Quarterly, 36*(3), 225–242.

Baker, B. D., Oluwole, J., & Green, P. C. (2013). The legal consequences of mandating high stakes decisions based on low quality information: Teacher evaluation in the race-to-the-top era. *Education Evaluation and Policy Analysis Archives, 21*, 1–71.

Ball, S. J. (2008). *The education debate*. The Policy Press.

Banks, J., & Smyth, E. (2015). Your whole life depends on it': Academic stress and high-stakes testing in Ireland. *Journal of Youth Studies*, *18*(5), 598–616. https://doi.org/10.1080/13676261.2014.992317

Baumert, J. & Demmrich, A. (2001). Test motivation in the assessment of student skills: The effects of incentives on motivation and performance. *European Journal of the Psychology of Education, 16*, 441–462. https://doi.org/10.1007/BF03173192

Berliner, D. (2009). MLCB (Much curriculum left behind): A U.S. Calamity in the making. *Educational Forum*, *v73*, n4.

Berliner, D. (2011). Rational responses to high stakes testing: The case of curriculum narrowing and the harm that follows. *Cambridge Journal of Education*, *41*(3), 287–302. https://doi.org/10.1080/0305764x.2011.607151

Boguslawski, M. K., Lohrmann, D. K., Sherwood-Laughlin, C., Eckes, S., Chomistek, A. K., & Applegate, T. (2021). Elementary school personnel and cultural factors affecting health education implementation in the high-stakes testing era. *Journal of School Health, 91*(10), 846–856.

Cole, J. S., Bergin, D. A., & Whittaker, T. A. (2008). Predicting student achievement for low stakes tests with effort and task value. *Contemporary Educational Psychology*, *33*(4), 609–624. https://doi.org/10.1016/j.cedpsych.2007.10.002

Cunningham, W., & Sanzo, T. (2002). Is high-stakes testing harming lower socio-economic schools? National Association of Secondary School Principals NASSP Bulletin, pp. 62–75

Damankesh, M., & Babaii, E. (2015). The washback effect of Iranian high school final examinations on students' test-taking and test-preparation strategies. *Studies in Educational Evaluation, 45*, 62–69.

David, J. (2011). What students need to learn: High-stakes testing narrows the curriculum. *Educational Leadership*, *68*(6), 78–80.

Demir, C., & Keles, O. (2021). The impact of high-stakes testing on the teaching and learning processes of mathematics. *Journal of Pedagogical Research*, *5*(2), 119–137. https://doi.org/10.33902/JPR.2021269677

Dowley, M. & Rice, S. (2022). Comparing student motivations for and emotional responses to national standardised tests and internal school tests: The devil in the detail. *Australian Journal of Education.* https://doi.org/10.1177/00049441211061889

Eklöf, H. (2010). Skill and will: Test-taking motivation and assessment quality. *Assessment in Education: Principles, Policy & Practice, 17*(4), 345–356. https://doi.org/10.1080/0969594X.2010.516569

Eklöf, H., Pavešič, B. J., & Grønmo, L. S. (2014). A cross-national comparison of reported effort and mathematics performance in TIMSS Advanced. *Applied Measurement in Education*, 27(1), 31–45. https://doi.org/10.1080/08957347.2013.853070

Elwood, J., Hopfenbeck, T., & Baird, J. A. (2017). Predictability in high-stakes examinations: Students' perspectives on a perennial assessment dilemma. *Research Papers in Education*, *32*(1), 1–17. https://doi.org/10.1080/02671522.2015.1086015

Garver, R. (2019). *Evaluative relationships: Teacher accountability and professional culture.* Department of Educational Leadership Scholarship and Creative Works. 16. https://digitalcommons.montclair.edu/educ-leadership-facpubs/16

Gregory, K., & Clarke, M. (2003). High stakes assessment in England and Singapore. *Theory into Practice*, *42*(1), 42–50.

Heissel, J. A., Adam, E. K., Doleac, J. L., Figlio, D. N., & Meer, J. (2021). Testing, stress and performance: How students respond to high-stakes testing. *Education Finance and Policy, 16*(2), 183–208.

Hofflinger, A., & von Hippel, P. (2018). The response to high-stakes testing in Chile: Did schools improve learning or merely inflate test scores? SSRN. https://doi.org/10.2139/ssrn.2906552

Hopfenbeck, T. N., & Kjaernsli, M. (2016). Student's test motivation in PISA: The case of Norway. *The Curriculum Journal, 27*(3), 406–422.

Howe, K., Eisenhart, M., & Betebenner, D. (2001). School choice crucible: A case study of Boulder Valley. *Phi Delta Kappan*, *83*(2), 137–146.

Howell, A. (2017). Because then you could never ever get a job!': Children's constructions of NAPLAN as high-stakes. *Journal of Education Policy*, *32*(5), 564–587. https://doi.org/10.1080/02680939.2017.1305451

Johnson, D. D., Johnson, B., Farenga, S., & Ness, D. (2008). *Stop high stakes testing: An appeal to America's conscience.* Rowman & Littlefield.

Klapp, A. (2015). Does grading affect educational attainment? A longitudinal study. *Assessment in Education: Principles, Policy and Practice, 22*(3), 302–323.

Klapp, A. (2018). Does academic and social self-concept and motivation explain the effect of grading on students' achievement? *European Journal of Psychology of Education*, *33*, 355–376. https://doi.org/10.1007/s10212-017-0331-3

Kwon, S. K., Lee, M., & Shin, D. (2017). Educational assessment in the Republic of Korea: Lights and shadows of high-stake exam-based education system. *Assessment in Education: Principles Policy & Practice*, *24*(1), 60–77. https://doi.org/10.1080/0969594X.2015.1074540

La Londe, P. G. (2017). A failed marriage between standardization and incentivism: divergent perspectives on the aims of performance-based compensation in Shanghai, China. *Education Policy Analysis Archives*, *25*(88). https://doi.org/10.14507/epaa.25.2891

Lavigne, A. L. (2014). Exploring the intended and unintended consequences of high-stakes teacher evaluation on schools, teachers, and students. *Teachers College Record, 116*(010308), 29.

Le,V-N. (2020). Do student-level incentives increase student achievement? A review of the effect of monetary incentives on test performance. *Teachers College Record, 122*, 3. https://doi.org/10.1177/016146812012200304

Learned, J. E., Dacus, L. C., Morgan, M., Schiller, K. S., & Gorgun, G. (2020). "The tail wagging the dog": high-stakes testing as a mediating context in secondary literacy-related. *Teachers College Record, 122*(11), 1–47. https://doi.org/10.1177/016146812012201115

Lewis, S., & Hardy, I. (2015). Funding, reputation and targets: The discursive logics of high-stakes testing. *Cambridge Journal of Education*, *45*(2), 245–264. https://doi.org/10.1080/0305764x.2014.936826

Lingard, B. (2010). Policy borrowing, policy learning: Testing times in Australian schooling. *Critical Studies in Education*, *51*(2), 129–147. https://doi.org/10.1080/17508481003731026

Lobascher, S. (2011). What are the potential impacts of high-stakes testing on literacy education in Australia?. *Australian Journal of Language & Literacy*, *34*(2), 9–19.

Madhaus, G., Russell, M., & Higgins, J. (2009). *The paradoxes of high-stakes testing: How they affect students, their teachers, principals, schools and society*. Information Age Publishing.

Menken, K. (2016). High-stakes tests as de facto language education policies. In E. Shohamy, I. G. Or, & S. May (Eds.), *Language Testing and Assessment* (pp. 1–12). Springer International Publishing.

Munter, C., & Haines, C. (2019). Students get what flows downward: district leaders' rationalizations of the standardized testing of children. *Educational Forum, 83*(2), 160–180. https://doi.org/10.1080/00131725.2019.1567891

Pastore, S., & Andrade, H. L. (2019). Teacher assessment literacy: A three-dimensional model. *Teaching and Teacher Education, 84*, 128–138. https://doi.org/10.1016/j.tate.2019.05.003

Penk, C., & Richter, D. (2016). Change in test-taking motivation and its relationship to test performance in low-stakes assessments. *Educational Assessment Evaluation and Accountability*, *29*, 55–79.

Pham, L., Nguyen, T., & Springer, M. (2021). Teacher merit pay: A meta-analysis. *American Educational Research Journal*, *58*(3), 527–566.

Phelps, R. P. (2006). Characteristics of an effective student testing system. *Educational Horizons, 85*(1), 19–29.

Phelps, R. P. (2012). The effect of testing on student achievement, (1910–2010). *International Journal of Testing, 12*, 21–43.

Polesel, J., Rice, S., & Dulfer, N. (2014). The impact of high-stakes testing on curriculum and pedagogy: A teacher perspective from Australia. *Journal of Education Policy*, *29*(5), 640–657. https://doi.org/10.1080/02680939.2013.865082

Rice, S., Care, E., & Griffin, P. (2012a). Compulsory literacy and numeracy exit standards for senior secondary students: The right direction for Australia? *Australian Journal of Education, 56*, 40–55. https://doi.org/10.1177/000494411205600104

Rice, S., Dulfer, N., Polesel, J., & O'Hanlon, C. (2016a). NAPLAN and student wellbeing: Teacher perceptions of the impact of NAPLAN on students. In B. Lingard, G. Thompson, & S. Sellar (Eds.), *National testing in schools: An Australian assessment*. Routledge.

Rice, S., Griffin, P., Care, E., & McPherson, J. (2012b). Compulsory hurdle literacy and numeracy requirements for senior secondary students: What do the stakeholders think? *Research in Post-Compulsory Education*, *17*(2), 143–161.

Rice, S., Watt, H. M. G., Richardson, P. W., & Crebbin, S. (2024). Interest in science and mathematics teaching among non-science/mathematics initial teacher education candidates. *Educational Review*. https://doi.org/10.1080/00131911.2024.2379418

Rogers, S. L., Barblett, L., & Robinson, K. (2016). Investigating the impact of NAPLAN on student, parent and teacher emotional distress in independent schools. *Australian Educational Researcher*, *43*, 327–343. https://doi.org/10.1007/s13384-016-0203-x

Rosenweig, E., Wigfield, A., & Eccles. (2019). Expectancy-value theory and its relevance for student motivation and learning. In K. A. Renninger, & S. E. Hidi, (Eds.) *The Cambridge handbook of motivation and learning* (pp. 617–644).

Save our schools (n.d.) Encouraging students to stay away on test day is a rort to increase school test scores – SOS Australia.

Segool, N. K., Carlson, J. S., Goforth, A. N., von der Embse, N., & Barterian, J. A. (2013). Heightened test anxiety among young children: Elementary school students' anxious responses to high-stakes testing. *Psychology in the Schools, 50*(5), 489–499. https://doi.org/10.1002/pits.21689

Segool, N. K., von der Embse, N. P., Mata, A. D., & Gallant, J. (2014). Cognitive behavioral model of test anxiety in a high-stakes context: An exploratory study. *School Mental Health, 6*(1), 50–61. https://doi.org/10.1007/s12310-013-9111-7

Silm, G., Must, O., That, K., & Pedaste, M. (2020). Does test-taking motivation predict test results in a high-stakes testing context? *Educational Research and Evaluation*, *26*(7–8), 387–413.
Smith, W. C., & Holloway, J. (2020). School testing culture and teacher satisfaction. *Educational Assessment, Evaluation and Accountability, 32*, 461–479. https://doi.org/10.1007/s11092-020-09342-8
Supovitz, J. (2009). Can high stakes testing leverage educational improvement? Prospects from the last decade of testing and accountability reform. *The Journal of Educational Change*, *10*(2–3), 211–227. https://doi.org/10.1007/s10833-009-9105-2
Thelk, A. D., Sundre, D. L., Horst, S. J., & Finney, S. J. (2009). Motivation matters: Using the Student Opinion Scale to make valid inferences about student performance. *Journal of General Education*, *58*(3), 129–151. https://doi.org/10.2307/27798135
Thompson, G., & Cook, I. (2014). Manipulating the data: Teaching and NAPLAN in the control society. *Discourse: Studies in the Cultural Politics of Education*, *35*(1), 129–142. https://doi.org/10.1080/01596306.2012.739472
Thompson, G., & Harbaugh, A. G. (2013). A preliminary analysis of teacher perceptions of the effects of NAPLAN on pedagogy and curriculum. *Australian Educational Researcher*, *40*(3), 299–314. https://doi.org/10.1007/s13384-013-0093-0
Tóth, E., & Csapó, B. (2022). Teachers' beliefs about assessment and accountability. Educational Assessment. *Evaluation & accountability*, *34*, 459–481. https://doi.org/10.1007/s11092-022-09396-w
Welsh, R., Graham, J., & Williams, S. (2019). Acing the test: An examination of teachers' perceptions of and responses to the threat of state takeover. *Educational Assessment, Evaluation and Accountability, 31*, 315–347.
Whitney, C., & Candelaria, C. (2017). The effects of No Child Left Behind on children's socioemotional outcomes. *AERA Open*. https://doi.org/10.1177/2332858417726324
Wigfield, A., & Eccles, J. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, *25*, 68–81.
Wise, S. L., & DeMars, C. E. (2005). Low examinee effort in low-stakes assessment: Problems and potential solutions. *Educational Assessment*, *10*(1), 1–17. https://doi.org/10.1207/s15326977ea1001_1
Yerdelen-Damar, S., & Elby, A. (2016). Sophisticated epistemologies of physics versus high-stakes tests: How do elite high school students respond to competing influences about how to learn physics? *Physical Review Physics Education Research*, *12*(1), 010118.
Yui, L. (2020). Educational injustice in a high-stakes testing context: A mixed methods study on rural migrant children's academic experiences in Shanghai public schools. *Comparative Education Review*, *64*(3), 498–524.
Zohar, A., & Alboher Agmon, V. (2018). Raising test scores vs. teaching higher order thinking (HOT): Senior science teachers' views on how several concurrent policies affect classroom practices. *Research in Science & Technological Education*, *36*(2), 243–260.

Authors and Affiliations

Suzanne Rice[1] · Therese Hopfenbeck[2] · Christopher Corbel[3]

✉ Suzanne Rice
s.rice@unimelb.edu.au

[1] Youth Research Collective, Faculty of Education, University of Melbourne, Melbourne, Australia

[2] Assessment and Evaluation Research Centre, Faculty of Education, University of Melbourne, Melbourne, Australia

[3] Language and Literacy Research Hub, Faculty of Education, University of Melbourne, Melbourne, Australia