



Research Article

© 2022 Abdulrahman Albeladi.

This is an open access article licensed under the Creative Commons Attribution-NonCommercial 4.0 International License (<https://creativecommons.org/licenses/by-nc/4.0/>)

Received: 14 March 2022 / Accepted: 11 April 2022 / Published: 5 May 2022

Assessing Students' Learning Outcomes in the Saudi Education System from the Perspective of Various Stakeholders

Abdulrahman Albeladi

Islamic University of Madinah,
Abo Bakr Al Siddiq,
Al Jamiah, Medina 42351,
Saudi Arabia

DOI: <https://doi.org/10.36941/jesr-2022-0087>

Abstract

Assessing students' learning outcomes is a significant part of the education system. This study investigated the perspective of a new policy that focuses on assessing students' learning outcomes. The method of implementation and the impact of the new policy were discussed. As this study was exploratory in nature, a qualitative research approach was applied, which allowed for the compilation of rich data through document analysis and semi-structured interviews. This study recruited 13 participants, including five supervisors, six head teachers, and two deputies. Results revealed the following: (1) a wide gap between the policy and actual practice; (2) limited involvement of the school community; (3) supervisors' examinations suffered from poor quality; (4) positive and negative impacts related to teaching and learning outcomes; and (5) several obstacles that needed to be resolved. Finally, this paper provides numerous recommendations to develop the process of assessing students' learning outcomes.

Keywords: student evaluation, school supervision, evaluation methods, leadership role, policy in practice

1. Introduction

The Kingdom of Saudi Arabia (KSA) is one of the G20 countries and has the highest population growth rate in the Middle Eastern region (Al-Hanawi et al., 2019). Of its population of more than 35 million, 31% are under 19 years and may be receiving primary education (General Authority for Statistics, 2020). One of the main objectives of Saudi Vision 2030 is the improvement of basic learning outcomes. Of late, the education sector in KSA has received substantial funding from the government: for example, in 2020, the government allocated SAR 193 billion to the education sector, accounting for 18.9% of the total national budget (Ministry of Finance [MOF], 2019). Yet, according to the Organisation for Economic Co-operation and Development (OECD, 2020), the Saudi education sector still fares poorly with regard to global assessment yardsticks such as Trends in International Mathematics and Science Study (TIMSS), Progress in International Reading Literacy Study (PIRLS), and the Programme for International Student Assessment (PISA).

Evidently, the huge investment in education has not been reflected in academic performance. For example, in the PISA 2018 assessment, which assessed proficiency in Mathematics, Science, and Reading of 15-year-old students, KSA was ranked 79, far below the average rank of OECD countries.

Further, in Mathematics and Science, Saudi students performed poorly in comparison to those from other participating countries from the Middle East and North Africa (MENA) region. These lower academic achievements indicate that developing student learning outcomes is of critical importance if KSA aspires to meet its economic targets outlined in Vision 2030 (OECD, 2020). The national strategy (2016–2020) has also stressed on improving students' learning outcomes and methods of student evaluation by taking advantage of the PISA, TIMMS, and PIRLS results (National Strategy for Education Development in Saudi Arabia [NSEDSA], 2015).

1.1 *The process of assessing students' outcomes:*

For developing the quality of students' learning outcomes, assessments, and supervision of schools, the MOE has suspended the traditional method of direct supervision of schools and adopted a new policy that seeks to employ Key Performance Indicators (KPIs). The MOE applied indicators of learning outcomes (MOE, 2017), which are the main focus of this research. Generally, in KSA, there are four types of students' learning outcomes assessments:

First, there exist teachers' assessments for each subject during and at the end of each academic term, which are mostly tests across all school levels (primary, intermediate, and secondary). However, these tests have low validity and do not offer an accurate picture of students' achievement: they only test a small set of cognitive skills and focus on memorising of information (EEBERSA, 2018) instead. Consequently, a new assessment body consisting of supervisors has been introduced.

Second, supervisors (inspectors) and head teachers test students on each subject to measure their learning outcomes. Further details are provided in the next section.

Third, supervisors conduct stage assessments of all pupils in grades 4,7, and 10, for Mathematics, Science, and Arabic.

Fourth, there exist national and international tests. National tests conducted by the Education and Training Evaluation Commission (ETEC), such as the General Aptitude Test (GAT) and the Scholastic Achievement Admission Test (SAAT) for grades 11 and 12, are required for university admissions and not all students need to take them. International tests such as TIMSS for grades 4 and 8, PIRLS for grade 4, and PISA for grade 9 also exist. However, this study will just focus on the second and third types of assessment as they are related to the new policy under deliberation.

1.2 *National Framework for Public Education Curricula Standards:*

Apart from the new direction taken by the MOE, another government agency, the ETEC, introduced the National Framework for Public Education Curricula Standards in 2018 (NFPECS, 2018). This new national framework provided a theoretical and procedural reference for building, implementing, and evaluating educational curricula standards. The present study will use this framework to evaluate the practice in the field and shed light on issues related to the implementation of the aforementioned policy. The framework specifies eight principles of assessment: (1) planning for assessment; (2) diversity of assessment methods; (3) validity and consistency of assessment methods; (4) assessment as a continuous process; (5) learner as a partner in the assessment processes; (6) use of performance measurement reports as benchmarks; (7) utilising assessment results for developing the education system; and (8) emphasising the learner's participation in the formative assessment processes and employing assessment strategies to raise their motivation, self-confidence, and the level of achievement. The NFPECS (2018) also stressed the use of assessment results to develop the educational system in the KSA by analysing national and international test results. Further, it suggested using these results to identify the strengths and weaknesses of the educational system and develop enhancement plans to improve the quality of educational outcomes.

1.3 Research questions

This study aimed to answer the following questions:

1. What is the role of supervisors and school leaders in the process of measuring students' learning outcomes?
2. How was the new policy implemented?
3. What was the impact of the new policy?
4. Are there ways to improve the process of implementing the new policy?

2. Literature Review

The Saudi education system is centralised in terms of decision-making, with limited space for schools to take significant decisions (Al Sadaawi, 2010; OECD, 2020). This centralisation can appear in different areas, including curricula, students' evaluation, supervision, management, funding, recruitment of teachers and head teachers, among others, and restrict school communities from taking decisions in different segments of the education process (Algarni & Male, 2014; Albeladi, 2016).

2.1 Assessment practices in KSA:

In KSA, teachers often lack the training to conduct proper assessments. Therefore, improving student assessment practices is vital. The current mode of assessment prioritises memorisation skills and passing exams instead of urging students to develop a deep understanding of various subjects (Al Alhareth & Al Dighrir, 2014; Alotabi, 2014). Moreover, a decade ago, Al Sadaawi (2010) had confirmed that student learning outcomes are not measured by a national assessment system in KSA. Currently, there is a huge emphasis on predominantly summative assessments without a clear understanding of the purpose of the assessment, and a national assessment framework is lacking (OECD, 2020). Despite ETEC's introduction of the NFPECS in 2018, this national evaluation framework has not been implemented in schools, as this study will clarify.

Students' assessments should boost academic performance and be part of the learning process and not an end in itself, which is the case in KSA (Al Alhareth & Al Dighrir, 2014; OECD, 2020). There is an increasing body of evidence that suggests that incorporating assessments into pedagogy can be a powerful strategy to improve both students' involvement and learning outcomes (Wiliam, 2011). Evidence also suggests that improving teachers' effectiveness, teachers' assessment skills, and using a well-designed system of evaluation can improve students' learning outcomes (Akiri, 2013; Christoforidou & Kyriakides, 2021).

There are several studies related to the implementation of the new policy. On the positive front, the new policy contributed to improve the professionalism of supervisors and head teachers' work in practically human relationships, planning, curriculum, and teachers' professional development (Alashqar, 2016; A. Alzahrani, 2019). On the negative front, the DOE, in a bid to protect its reputation, exerts considerable pressure on supervisors when schools under their supervision experience decreases in educational achievement (Alashqar, 2016). Consequently, supervisors often focus on quantitative rather than qualitative outcomes (Alashqar, 2016; Almasoudi, 2016). Supervisors also complain about a lack of incentives (Almasoudi, 2016; Alkhiliwy & Alyahya, 2017; Alqahtani, 2016) and a lack of financial resources and support, which increases their transportation costs (Almasoudi, 2016).

Furthermore, supervisors and head teachers receive an overwhelming number of instructive circulars from the DOE and MOE, which distract them from important work (Alqahtani, 2016). Moreover, Alshehri (2017) pointed out the lack of training opportunities for school deputies. There was an insufficient number of supervisors and assessment specialists in the DOE. This situation led to an overload of work for supervisors, who had a huge number of schools and teachers under their supervision and were responsible for tasks that did not belong to their supervisory roles (Almasoudi,

2016; Alqahtani, 2016; Nawawi, 2016).

In summary, these studies highlighted important aspects such as the state of the new policy, obstacles to its implementation, and the need to improve supervisors' effectiveness. They raised several important issues such as needs pertaining to professional development, increasing the number of supervisors charged with implementing the new policy, and establishing incentives. However, little is known about how the new policy was implemented, the process of implementation, and how the supervisors, head teachers, and deputies at the school level perceived the new policy. This study primarily investigates these considerations.

3. Methodology

This study used a qualitative approach: semi-structured interviews served as the primary data collection tool and documents related to the new policy were analysed. The interviews, ranging between 40 and 60 min each, were digitally recorded and saved in a secure computer. Additionally, it was useful to refer to and analyse the documents published by the MOE and ETEC related to the new policy and students' learning outcomes assessments in the KSA. The interviews were transcribed and analysed using the MAXQDA software. To maintain trustworthiness, data were triangulated from multiple sources. For member checking, participants were asked to verify final transcripts to ensure their intended meaning was presented accurately. Thematic analysis was used as a technique, and the aforementioned process resulted in the creation of 361 'initial codes', five main themes, and 20 sub-themes.

3.1 Selection of participants

The total number of interviewees in this research was 13. To achieve triangulation, the participants included five teachers' supervisors, six head teachers, and two deputies at the school level. The participants were selected based on the following purposive sampling criteria: (1) having more than 15 years of experience in education in primary, secondary, and intermediate schools in Madinah; and (2) being involved in the implementation of the new policy.

3.2 Ethical considerations

Prior to the study, ethical approval from the institution committee was obtained, all participants were informed about the nature and the aim of this research as well as the interview process, and all participants confirmed their willingness to engage in the study. Moreover, 'written informed consent' was obtained and confidentiality was assured. Due to the sensitivity of the data and the nature of this research, the identities of the participants were anonymised, and they were coded as follows: for example, teachers' supervisor (1), HT (3).

4. Results and Discussion

This section will present and discuss the findings of this study.

RQ 1: *What is the role of supervisors and school leaders in the process of measuring students' learning outcomes?*

4.1 A new role for supervisors and head teachers

The new policy was initially adopted in 2013 and has developed over time. It created a new role for supervisors and head teachers to measure students' learning outcomes. Data on students' achievements was used to calculate the value added by individual teachers and schools. The DOE led the implementation of the policy and controlled the process of measuring students' learning

outcomes through a central committee by setting objectives, providing training, controlling data collection and analysis, and ranking schools. Almasoudi (2016) and Alshehri (2017) confirmed this centralised characteristic of the system, which limited the autonomy of school leaders.

Each supervisor and head teacher are required to visit teachers under their supervision two times during each academic term, and they required to create, print, and conduct the test for students. Subsequently, they correct the answer sheets to calculate the general mean and value added for each teacher. These tests do not affect the students' annual evaluation or promotion, nor do they affect appraisals based on teachers' performance. With the proliferation of online education during the COVID-19 pandemic, the number of tests and (online) visits have increased, and supervisors are conducting more than two tests for each class during the academic term.

The head teachers interviewed for this study indicated that they had much knowledge about the schools and the teachers under their supervision. Several head teachers had little belief in the policy, their role in it, and its impact on student learning. HT (3) stated, 'I do not know the purpose and the objective of this policy; the supervisor visited our schools and carried out the tests, which have no genuine impact on students' learning.' HT (3) added:

'From my point of view, supervisors are not qualified to measure students' learning outcomes and cannot evaluate teachers' efforts by visiting schools twice each term; I know my teachers better than they do. Also, I believe supervisors' efforts must focus on following them if they are on the right track and improving their teaching skills only.'

Moreover, HT (1) said, 'I believe in the result of the ETEC tests (GAT and SAAT); these tests have more credibility because of their quality, they come from a neutral body, and they rank the schools in each district fairly'. These findings are consistent with those of OECD (2020), which indicated that school staff did not trust supervisors' evaluations due to their focus on a limited range of teaching activities. Further, most supervisors were former teachers with limited experience in assessing school staff and their efforts. This finding was confirmed by Nawawi (2016), who pointed out the lack of evaluation specialists in the DOE. Moreover, findings from global research have suggested that school-based evaluation is more reliable (OECD, 2020), which was not supported by the new policy.

This new role for supervisors and head teachers demonstrates the MOE's lack of trust in teachers' assessments in the classroom and other national assessments such as the GAT and SAAT, which are controlled by the ETEC. Furthermore, it reflects the centralisation of the education system. As McFadden and Williams (2020) have suggested in their review, there is a significant benefit to improving teachers' preparation and training by focusing on evaluation and evaluative thinking, which can have positive impacts on teachers' professional practices and students' learning outcomes.

RQ 2: *How was the new policy implemented?*

4.2 *Gap between the policy and practice*

4.2.1 *One strategy to measure schools' outcomes:*

Interestingly, all participants confirmed that there was only one strategy for measuring school performance, namely, tests. This contradicted the policy, which emphasised both quantitative and qualitative measurement strategies. Likewise, the NFPECS (2018) framework highlighted the importance of diversity of assessment methods to evaluate students' learning outcomes. These tests, designed by supervisors and head teachers to evaluate students' achievement, are also used to calculate the value added by each teacher, which is then linked to teachers' annual evaluations and for ranking schools based on performance. That said, linking teachers' quality with students' achievement is a controversial model as these tests are not designed to assess teachers (Goe, 2007). The new policy adopted a model of measuring teachers' quality via value-added models. Such a model also makes it difficult to identify other factors that affect students' learning outcomes, such as students' socioeconomic and cultural backgrounds, the influence of peers, curricula, the learning

environment, and so on. Indeed, several scholars have expressed their reservations about the use of value-added models to evaluate teachers' performance (Braun, 2005). The limitations of value-added models are also evidenced in this study, and it is posited that this will lead teachers to focus more on exams rather than on teaching and students' learning.

Supervisor (4) also highlighted the centralised nature of education wherein the MOE distributes textbooks for each module, defines students' learning outcomes, and sets targets for learning outcomes to be achieved in each subject. These quantitative strategies have a long history in Saudi education and are mostly used because they are clear, simple, and easy to implement on a large scale. This finding was confirmed by Al Alhareth and Al Dighrir (2014), Almasoudi (2016), and Alashqar (2016).

4.2.2 Lack of enhancement plans at the school level

The most surprising aspect of the findings is the lack of enhancement plans in schools, which are essential for improving students' learning outcomes. All head teachers complained about not receiving a comprehensive assessment report. Head teachers design enhancement plans based on data from supervisor reports. However, neither these reports nor any feedback on students' learning outcomes were submitted to the schools. This report and feedback, when provided, can help schools devise an enhancement plan for classes or students. Head teachers only received the general mean for their school. HT (1) indicated that their school had an enhancement plan to improve students' learning outcomes but said, 'Teachers' supervisors are not focusing on a real measure of the students' learning outcomes; rather, they are just looking to complete the documents sent by the MOE. Also, they did not use the data to help our students by showing us who were performing well and who needed more help and attention.'

This complaint was consistent with other findings by prior studies, which confirmed the prioritisation of grading rather than truly assessing students' learning within schools (Alotabi, 2014; Bramwell-Lalor & Rainford, 2015). This is also against the NFPECS (2018) framework which confirmed the 'use of assessment results in order to develop the education system'.

In contrast, supervisor (3) mentioned that supervisors were required to enter the value added in the NOOR system (a system for registration service) and that schools could extract the data from the system. This finding shows the tensions and miscommunication between schools and supervisors as they blame each other for shortcomings. Moreover, it appeared that supervisors were more intent on meeting the MOE requirements than on improving schools through their work. This can perhaps be attributed to centralisation and the desire of supervisors to please their superiors.

Feedback on assessment can help with learning and clarifying grades and improving students' outcomes has been linked with constructive feedback (Li & De Luca, 2014). Similarly, the NFPECS (2018) also notes that students' participation in the formative assessment processes and the use of assessment strategies can lead to increases in students' motivation, self-confidence, and level of achievement. Students must participate actively in the assessment process (Sadler, 1989) to develop the ability to analyse their own learning and become self-directed learners (OECD, 2013). In contrast, neither the policy nor the practices of supervisors involved students and their parents in the assessment process, and this needs to be revised in the future. Despite evidence suggesting that community participation can improve students' learning outcomes (Bin Salamah, 2001), community engagement is not prevalent in the Saudi education system (Aljughaiman & Grigorenko, 2013).

4.2.3 Lack of high-quality tests

One unexpected finding was that the tests designed by teachers' supervisors were of poor quality, and most of the head teachers confirmed this. A possible explanation for this might be the fact that supervisors are far removed from the classroom educational process. HT (2) stated, 'Some teachers' supervisors worked for a long time in administration roles and may have had less engagement with

the subject. Therefore, they designed unsuitable tests for the students. Teachers complained about the tests not having credibility.' HT (3) further added, 'Teachers' supervisors are not qualified to design tests and evaluate teachers' efforts through these poor tests.' HT (1), meanwhile, pointed out that tests designed by supervisors were far below the quality standards of international tests such as TIMSS, PIRLS, and PISA. Although the NFPECS (2018) guidelines had specified the use of high-quality assessment methods by supervisors, evidence suggests that this has not been the case. This finding can perhaps be explained by Almasoudi's (2016) study, which found a lack of professional preparation and training for supervisors.

Furthermore, HT (5) stated, 'These tests provided by teachers' supervisors focus on memorising information, not using a higher order of thinking.' These findings were confirmed by several studies (Alotabi, 2014), which found that assessments in schools were highly content-focused and did not concentrate on complex skills such as critical thinking and problem-solving. The majority of test questions contained one correct answer option, which demotivated deep learning (OECD, 2020). Therefore, multidimensional measures for assessing non-cognitive skills are required, and the government is pushing for initiatives to improve non-cognitive skills (EEBERSA, 2018). While the MOE has undertaken efforts recently to alter curricula to newly introduce such skills, assessments are yet to experience a related shift.

In general, the new policy supports the notion of 'teaching to the test' rather than helping students gain mastery over subjects. Al Sadaawi (2010) had similar findings. Further, analysis of policy documents revealed that the policy concentrated on 'assessment of learning' through grading of students and schools rather than 'assessment for learning', despite extant literature highlighting the benefits of the latter (Bramwell-Lalor & Rainford, 2015). Assessment of learning has been linked to a behaviourist perspective, which puts emphasis on outcomes rather than the actual learning process (Shepard, 2000), while assessment for learning largely falls within the paradigm of constructivism (Bramwell-Lalor & Rainford, 2015). Although the education minister has stated that 'tests drive and influence the student in the classroom and on the level of the learning outcomes' (Minister video, 2019), this has not been reflected in practice.

4.2.4 Non-standardised tests

While it is acceptable for internal assessments to be non-standardised, standardisation is vital for external assessments that compare school performance to rank schools. The lack of standardisation in external assessments can lead to unfair evaluations of schools and teachers. Supervisor (5) confirmed this, saying, 'The tests are not of the same standard: some supervisors make them easy, and others make them difficult. This difference can be resolved by using a template or specific criteria for each module.' Similarly, Al Sadaawi (2010) found that educational outcomes in KSA had not been monitored objectively over time using reliable indicators.

To sum up, the investigation of the 'gap between policy and practice' was one of the main contributions and most evident finding to emerge from the analysis of policy documents and participants' interviews about actual practice. There are several explanations for this gap, including but not limited to: (1) the lack of educational assessment specialists among supervisors and head teachers (Almasoudi, 2016; Nawawi, 2016); (2) the inadequacy of preparation and training programmes provided by the MOE and its agent, the DOE (Almasoudi, 2016); and (3) practitioners' perception of the policy as being inadequate and burdensome (Almasoudi, 2016; Alqahtani, 2016; Nawawi, 2016).

RQ 3: *What was the impact of the new policy?*

4.2.5 Positive and negative impacts of the policy

The new policy has had a positive impact on the role of supervisors in improving teachers' teaching skills. As supervisor (4) confirmed, 'When a teacher's value added is low, I give the teacher further

attention by arranging additional visits to observe classroom teaching, and by offering more training courses and arranging professional visits with expert teachers.' HT (3), meanwhile, said, 'About half of our teachers increase their efforts to achieve better results and raise our school rank.' Supervisor (3) also stated that the policy helped prepare students for international tests. In a similar vein, deputy (1) said, 'I notice there is an increased interest in students' results in international tests after implementing the new policy.'

4.2.6 Limited impact on students' learning outcomes

While there are some positive impacts of the new policy, the participants of this study have mentioned some negative impacts, too. There is a general agreement among participants that the policy has had a slight impact on students' learning outcomes [supervisor (4), HT (3), (5), and (6)]. This is consistent with the findings of Alshamry (2018). Further, the TIMSS (2019) report confirmed the same, with KSA ranking 53rd among 58 participating countries. Saudi students in the fourth grade attained average scores (398), which were slightly higher than their 2015 scores (383) and lower than their 2011 (410) scores. Similarly, D (2) said, 'The tests implemented by supervisors or head teachers have no effect.' When asked why, the participants replied, 'Students know it will not affect their academic performance and will not count in their final result, so they do not take it seriously.' This finding is consistent with that of OECD (2020), which found that students and teachers prioritise marks earned and believe unmarked tests to be worthless.

4.2.7 Impact on school management

Research on the new policy has not concentrated on its impact on school management. This study highlights this important issue raised by head teachers, who confirmed that the new policy had a negative impact on their ability to manage schools. All the head teachers confirmed that supervisors abused their power and visited schools without prior notice or arrangement. This caused specific problems in schools, such as: (1) changes in school schedules based on supervisors' requirements; (2) interference with teachers' weekly lesson plans or teachers' scheduled tests [D (2)]; (3) lack of readiness among students to take tests; and (4) neglect of lessons that were disrupted by these unannounced and sudden tests [HT (3)]. Despite these problems, an analysis of policy documents revealed that the policy had specific instructions for unannounced supervisor visits to schools. Additionally, this also highlights the head teachers' relative lack of power in comparison to the power held by supervisors.

RQ 4: *Are there ways to improve the process of implementing the new policy?*

4.2.8 Obstacles to the successful implementation of the new policy

Oversized classes: Previous studies have not dealt with obstacles to the implementation of the new policy at the classroom level. The present study found that oversized classes hindered the participants' ability to follow the procedures prescribed by the new policy. For instance, D (2) said, 'Every school's circumstance is different, and we faced difficulties in measuring students' learning outcomes due to the huge number of students, which reach 40 to 50 in each class.' There is evidence that overcrowded classes can have adverse effects on the implementation of new assessment practices (Al-Abdulkareem, 2004) and may prevent teachers from keeping track of students' progress and providing appropriate feedback (Alotabi, 2014).

4.2.9 Lack of training courses

The MOE does not provide training courses, and each DOE must provide these courses during the academic year. Supervisor (5) criticised this approach and demanded training courses from the MOE

on account of their superior quality. Many participants reported a lack of preparation for implementing the new policy. HT (5), for instance, said, 'There was a lack of preparation before launching the new policy, we need more training before implementing the new policy.' This finding was consistent with those of previous studies (Almasoudi, 2016; Alqahtani, 2016).

Further, HT (1) mentioned that most of the training courses related to the measurement of students' learning outcomes targeted head teachers and supervisors, not the teachers, and each head teacher was required to pass on their knowledge to the schools' teachers. This is also consistent with the findings of Alshehri's (2017) study. In some cases, there was no training provided to schools: HT (6) reported, 'There was no training course provided to our school for the measurement of students' learning outcomes.' This is surprising, considering that previous research has highlighted the importance of training educators, including supervisors (K. Alzahrani, 2019; Alotabi, 2014; OECD, 2020). HT (4) clarifies this: 'There is no training course for teachers to improve their skills for measuring students' learning outcomes.' These findings coincide with the findings by the OECD (2020) that the MOE delegated the task of training to its agents, the DOEs, and that the training centres did not focus on improving teachers' assessment skills. Moreover, this lack of training can explain the overall weakness of assessment literacy among teachers (Al-Abdulkareem, 2004; OECD, 2020).

5. Recommendations to Improve the Process of Measuring Students' Outcomes

5.1 Freedom to act for teachers:

Several participants called for greater freedom and flexibility for teachers in teaching subjects. Supervisor (5) stated that teachers were obligated to cover all topics in the curriculum, even when some students fell behind and required more attention. This lack of flexibility understandably affects learning outcomes, and in some cases, also leads to poor school performance.

5.2 Incentive system for hard workers

There is a need for an incentive system that rewards good performance by supervisors, head teachers, teachers, students, and the school as a whole. When asked about incentives for outstanding performance, all participants confirmed that there was no incentive system in place to reward those who worked harder, and they only received a certificate of appreciation in some cases. Earlier studies had reported similar findings (Alkhiliwy & Alyahya, 2017; Almasoudi, 2016; Alqahtani, 2016). The absence of an incentive system can limit practitioners' motivation to implement the policy.

5.3 Religious motivations

Regarding the implementation of the new policy, head teachers' religious beliefs can motivate them to work harder and show more concern about students' learning outcomes. HT (5) explained, 'You know, our job is trustworthiness, and God (Allah) will bless my money, children and my life when I do all I can for the students. So, I work hard, as (Allah) God observes my work.' In a similar vein, supervisor (2) said, 'We receive enough incentives from God—Allah.'

This cultural concept is a phenomenon in Muslim communities, which encourages people seeking blessings through honest work (Husin and Norhasniah, 2012). In the Saudi context, some conservative people express this belief more than others who feel that working hard is part of professionalism and is required to meet job expectations. In their study, Husin and Norhasniah (2012) found that Islam stresses on trustworthiness and fairness in the execution of tasks, and work is considered equivalent to the worship of Allah: the Al-Quran Surah Al-Bayyinah, verse 7, states, 'Those who believe and do good, they are the best of creatures.' This is a significant finding which can explain some aspects of head teachers' motivations and can be used to improve their effectiveness in

implementing the new policy or other initiatives.

5.4 Building trust

Building trust can be part of the solution to improve schools' performance. As it stands, head teachers and supervisors test students due to stakeholders' lack of trust in teachers' assessments (OECD, 2020). Supervisor (1) said that it would take several years to build trust, but there is a need to trust head teachers and teachers to make a difference. Moreover, HT (3) and both D (1) and D (2) said that teachers were unparalleled in their ability to measure students' learning outcomes, as they knew their students' skills, abilities, differences, and limitations. However, at present, a head teacher's power is limited, and this lack of trust can stifle their creativity and initiative.

To illustrate with an example, HT (5) said:

'We have a group of low-achieving students in key subjects such as Maths and Reading. The school's team and I changed the weekly lesson plan and launched a new initiative. We cut 10 minutes from some lessons such as 'Religion: Fiqh', PE, and Art, to focus on this group of students. After working for a few weeks, a supervisor visited our school and we informed him of our initiative. He told us this was not right, and we had to stick with the lesson plan provided by the MOE. Then, he forced us to cancel this initiative.'

This story suggests that the rigidity of the centralised nature of the education system can limit creative approaches to some problems. It further suggests that adherence to prescribed rules/procedures is considered more important than helping students improve.

6. Implications and Conclusion

One limitation of this study is that the researcher could not gather the supervisors' test results due to DOE restrictions and the lack of a culture of sharing data publicly and for research (Almasoudi, 2016). Such data could have helped identify progress, if any, in students' learning outcomes. Another limitation is that the perspectives of teachers and students who are not involved in policy implementation were not included. This can be considered in future research. Further research using quantitative methods can be undertaken on a large scale to test the generalisability of the themes identified in this study, such as the gap between policy and practice, assessment strategies, test quality, and the role of school communities, among other pertinent variables. Additionally, further research could investigate the long-term consequences of different kinds of assessment—international, national, or local—based upon students' learning outcomes.

Measuring students' learning outcomes is indispensable for the improvement of the education system in terms of efficacy and performance. Policymakers need to shift their focus from controlling the assessment process to delegating the process to neutral bodies such as the ETEC. It is recommended that efforts of supervisors to implement the new policy should: (1) shift from direct assessment of students to be more supportive of the ETEC's efforts, and (2) be supported by the MOE through the use of data sharing and comprehensive training programmes. There are some reasons for these suggestions: the policy (1) contradicted the ETEC examination plan and evaluation framework (NFPECS, 2018); (2) had a limited positive impact on students; (3) had a negative impact on supervisors, head teachers, and teachers; and (4) resulted in the use of assessments of questionable quality.

The study makes important contributions to the fields of students' assessment, policy in practice, and implementation processes. The main objective was to evaluate the process of implementing the new policy and the study illustrated how the participants in the field perceived it. Critical evaluation of the literature and qualitative data analysis guided the development of the main themes presented in the current study. Although this study is based on a small sample of participants

(13), its findings are significant as it enhances our understanding of the challenges that may arise while implementing a new policy in the education system.

For international readers, the findings also suggest that good design, adequate training for practitioners, community engagement, enhancement plans, and data sharing can play significant roles in the successful implementation of a policy. Engagement is key to the successful implementation of a policy. Leaders in the field of education need to reconsider ways to engage all stakeholders at different levels of policy implementation. It was evident in this study that a lack of engagement led to undesired behaviours and outcomes. Therefore, to increase the probability of successful implementation of any new policy: (1) there is a need for engagement at all levels, and a bottom-up approach is likely to be better suited; (2) there is a need to provide specialised training and support for both supervisors and school staff in order to implement any new policy. In conclusion, the lack of teamwork and a collaborative culture among supervisors from the DOE and school staff can lead to negative impacts. Hence, reducing centralisation and transferring power to schools can enable schools to become agents of change.

References

- Akiri, A. A. (2013). Effects of teachers' effectiveness on students' academic performance in public secondary schools; Delta State—Nigeria. *Journal of Educational and Social Research*, 3(3), 105. Retrieved from doi:10.5901/jesr.2013.v3n3p105.
- Al Alhareth, Y., & Al Dighrir, I. (2014). The assessment process of pupils' learning in Saudi education system: A literature review. *American Journal of Educational Research*, 2(10), 883–891. Retrieved from doi:10.12691/education-2-10-6.
- Albeladi, A. (2016). The perceptions of stakeholders in the field of education regarding new decentralisation efforts in Saudi Arabian schools (Doctoral dissertation, University of Southampton).
- Al Sadaawi, A. S. (2010). Saudi national assessment of educational progress (SNAEP). *International Journal of Education Policy and Leadership*, 5(11). Retrieved from doi:10.22230/ijep.2010v5n11a202.
- Al-Abdulkareem, S. A. (2004). Investigating science teachers' beliefs about science and science teaching: Struggles in implementing science education reform in Saudi Arabia [Doctoral dissertation]. West Virginia University. Retrieved from https://search.proquest.com/openview/955c232f651da0d9732b6bbf3191b021/1?pq-origsite=gscholar&cbl=18750&diss=y&casa_token=VCaypNKNNGnoAAAAA:Pvj5_PvpP8hbXPhiu5wYwYpO7GpYK5ySzP6iH3C_eJiqaGn2-wMP2lJbzRnmwv6MAKZjmHilbWA.
- Alashqar, H. (2016). The role of the system of indicators of supervisory performance leadership in increased the effectiveness of educational supervisors in Alahsah region—Saudi Arabia [Master's Thesis]. Imam Mohammed bin Saud University. Retrieved from <https://drive.google.com/file/d/1CdHQJ-mMLkgsREK2taKgRR0kPik-Khi/view?usp=sharing>. Accessed February 17, 2021
- Algarni, F., & Male, T. (2014). Leadership in Saudi Arabian public schools: Time for devolution? *International studies in educational administration (Commonwealth Council for Educational administration and management (CCEAM))*, 42(3). Retrieved from https://www.academia.edu/download/40180533/2014_alga_rni_male_ISEA.pdf.
- Al-Hanawi, M. K., Khan, S. A., & Al-Borie, H. M. (2019). Healthcare human resource development in Saudi Arabia: Emerging challenges and opportunities—a critical review. *Public Health Reviews*, 40, 1. Retrieved from doi:10.1186/s40985-019-0112-4.
- Aljughaiman, A. M., & Grigorenko, E. L. (2013). Growing up under pressure: The cultural and religious context of the Saudi system of gifted education. *Journal for the Education of the Gifted*, 36(3), 307–322. Retrieved from doi:10.1177/0162353213493153.
- Alkhiliw, L., & Alyahya, D. (2017). The degree of the system of indicators of supervisory performance leadership contributes to achieving administrative transparency in Education Office in Riyadh. *Journal of Scientific Research in Education*, 18(8), 613–642. Retrieved from <https://bit.ly/3bHXydf>.
- Almasoudi, M. (2016). Difficulties facing educational supervisors in applying the supervisory performance leadership system in the education administration at subya governorate [Master's Thesis]. Umm Al Qura University. Retrieved from <https://drive.google.com/file/d/18O-FHLesNpaqV4Ymd5qmw6W4XV4URIOJ/view?usp=sharing>.
- Alotabi, K. A. (2014). Student assessment strategies in Saudi Arabia: A case study of pre- and post-classroom practices. *Literacy Information and Computer Education Journal (LICEJ)*, 3(1), 1786–1763. Retrieved from doi:10.20533/licej.2040.2589.2014.0234.

- Alqahtani, M. (2016). The obstacles of implementing the System of Indicators of Supervisory Performance Leadership in the DOE in Riyadh-Saudi Arabia [Master's Thesis]. Imam Mohammed bin Saud University. Retrieved from <https://drive.google.com/file/d/1lMbtBwU8dlQrs38qhvghk5upsWUS8VP/view?usp=sharin.g>.
- Alshamry, N. (2018). The role of the supervisory system in improving the level of school performance in the secondary stage in the city of Riyadh from the viewpoint of educational supervisors. *Journal of Scientific Research in Education*, 19(15, autumn), 129–184. Retrieved from https://jsre.journals.ekb.eg/article_33781.html.
- Alshehri, N. (2017). Reality of administrative empowerment of private high school leadership in light of system indicators of leadership of school supervisory performance at Riyadh city [Master's Thesis]. Arab East College. Retrieved from https://drive.google.com/file/d/1YN_jQaaKWGMz4vqRMHenniPOms5bkO4g/view?usp=sharing.
- Alzaharani, A. (2019). The role of headteachers in Albaha region in implementing the System of Indicators of Supervisory Performance Leadership. *Scientific Journal of Faculty of Education, Assiut University*, 35(6–2). Retrieved from http://www.aun.edu.eg/faculty_education/journal/papers.php?P_ID=883.
- Alzaharani, K. (2019). An attitude of educational female supervisors towards using the system of indicators of supervisory performance leadership for achieving balance in some areas of their work. *Journal of Scientific Research in Education*, 20(2), 1–49. Retrieved from <https://bit.ly/3v9yWBJ>.
- Bin Salamah, M. A. M. (2001). An investigation of the relationship between Saudi teachers' curriculum perspectives and their preference of curriculum [Dissertation]. West Virginia University. Retrieved from <https://www.proquest.com/openview/3023da2b2ef9eaf3537dofbbab74e2d/1?pq-origsite=gscholar&cbl=18750&diss=y>.
- Bramwell-Lalor, S., & Rainford, M. (2015). Advanced level biology teachers' attitudes towards assessment and their engagement in assessment for learning. *European Journal of Science and Mathematics Education*, 4(3), 380–396. Retrieved from doi:10.30935/scimath/9478.
- Braun, H. I. (2005). Using student progress to evaluate teachers: A primer on value-added models. Princeton, NJ: Educational Testing Service. Retrieved from <https://eric.ed.gov/?id=ED529977>
- Christoforidou, M., & Kyriakides, L. (2021). Developing teacher assessment skills: The impact of the dynamic approach to teacher professional development. *Studies in Educational Evaluation*, 70, 101051. Retrieved from doi:10.1016/j.stueduc.2021.101051.
- EEBERSA. (2018). Enabling evidence-based education reform in Saudi Arabia. Retrieved from <https://drive.google.com/file/d/1klCLnEdvcMKPkb2viKvZzTKTcgaDSkb4/view?usp=sharing>.
- General Authority for Statistics. (2020). Population by age groups, and gender mid-year 2020. Retrieved from <https://www.stats.gov.sa/en/43>.
- Goe, L. (2007). The link between teacher quality and student outcomes: A research synthesis. National Comprehensive Center for Teacher Quality. Retrieved from <https://eric.ed.gov/?id=ED521219>.
- Husin, W. N. W., & Norhasniah, W. (2012). Work ethics from the Islamic perspective in Malaysia. *European Journal of Social Sciences*, 29(1), 51–60. Retrieved from doi:10.5339/rels.2015.work.16.
- Li, J., & De Luca, R. (2014). Review of assessment feedback. *Studies in Higher Education*, 39(2). Retrieved from doi:10.1080/03075079.2012.709494.
- McFadden, A., & Williams, K. E. (2020). Teachers as evaluators: results from a systematic literature review. *Studies in Educational Evaluation*, 64, 100830. Retrieved from doi:10.1016/j.stueduc.2019.100830.
- MOE. (2017). Indicators of learning outcomes, system of supervisory and school performance leadership (5th version). Retrieved from <https://drive.google.com/file/d/13ijEJ8e7OzjBemdH5nP-3NGJsgBKtuwj/view?usp=sharing>. MOE Publication.
- MOF. (2019). Saudi Arabia announces 2020 budget: Delivering efficiently and powering private sector growth. Retrieved from https://www.mof.gov.sa/en/MediaCenter/news/Pages/News_09122019.aspx.
- Minister video. (2019). MOE third meeting devoted to discussing results of national exams and learning outcomes. Retrieved from https://twitter.com/iEN_tv/status/1202891605050634240?s=20.
- Nawawi, E. (2016). The reality of the application of the supervisory performance leadership system competencies and its relationship to professional educators for supervisors from perspective management of education offices in Makkah area [Master's Thesis]. Umm Al Qura University. Retrieved from <https://drive.google.com/file/d/1ZPOG2-9cEKdyaLz36Ho8vT1jW4Ugye/view?usp=sharing>.
- NFPECS. (2018). National framework for public education curricula standards. Retrieved from <https://bit.ly/3opjIG5>.
- NSEDSA. (2015). National strategy for education development in Saudi Arabia. Retrieved from <https://www.moe.gov.sa/ar/aboutus/aboutministry/Pages/MinistryStrategy.aspx>.

- OECD. (2013). Student assessment: Putting the learner at the centre. In OECD reviews of evaluation and assessment in education (pp. 139–269). OECD Publishing. Retrieved from doi:10.1787/9789264190658-7-en.
- OECD. (2020). Education in Saudi Arabia. In Reviews of national policies for education. OECD Publishing. Retrieved from doi:10.1787/76df15a2-en.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), 119–144. Retrieved from doi:10.1007/BF00117714.
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4–14. Retrieved from doi:10.3102/0013189X029007004.
- TIMSS. (2019). Report: First Insight into mathematics and science achievements of Saudi Arabia students in Grades 4 and 8 in an international context. Education and training evaluation commission. Retrieved from <https://www.etc.gov.sa/ar/Researchers/Research-Studies/Documents/TIMSS%202019-%20Saudi%20Arabia%20national%20report.pdf>.
- William, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, 37(1), 3–14. doi:10.1016/j.stueduc.2011.03.001.