University of Reading

An Investigation into Experiences and Perceptions of School Performance Evaluation (SPE) in Secondary Schools in Saudi Arabia

Thesis submitted for the degree of Doctor of Philosophy

Institute of Education

Azzah Hadhil Alsubaie

Declaration

Declaration of Original Authorship:

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Azzah Alsubaie

Dedication

To my mother, Norah Alsubaie: the strongest and most beautiful woman I have ever known in my life; the one who taught me that education transforms a woman's life and that women change the world for the better.

Acknowledgements

Abstract

School performance evaluation (SPE) and the use of performance indicators is one of the most common educational reforms sweeping the globe as a consequence of neoliberalism and increased managerialism in education (Ball, 2012a). Like many countries, Saudi Arabia has implemented SPE to improve educational outcomes. The *School Performance Indicators System* (SPIS) is the most recent programme, but not the first (Al Hakamy, 2008). It was preceded by many other SPE programmes, some of which operated concurrently. Although SPE has generated widespread global debate and is known to cause increased workload and emotional strain on teachers and head teachers (Perryman, Ball, Maguire, & Braun, 2011), there is a dearth of empirical research into teachers and head teachers experiences of SPE in Saudi Arabia. This study is aimed at contributing to closing that gap. Giving a voice to an underrepresented group in research by focusing on female teachers and school head teachers, specifically, the aim of this study is to explore how they make sense of and experience SPIS monitoring and inspection visits, as well as their views of SPIS key performance indicators (KPIs) and accountability within the system. In addition, the study explores the effects of SPE on their stress levels, workload and morale.

The data collection and analysis are embedded in an interpretivist paradigm, using explanatory sequential multi-site case study mixed methods. This consisted of administering a questionnaire to 64 female head teachers and 109 female secondary school teachers in Jeddah. These data were then enriched by conducting interviews with three female head teachers and nine female teachers from three schools to gather more in-depth views of SPIS. The conceptual framework for this study centres upon three main concepts, which are discussed in depth: SPE, performativity and accountability.

The results of this study point to the importance of teachers and head teachers participating in the design and implementation of new programmes, aimed at education reform. The study reveals the impact of centralisation on the reduced efficiency of SPE implementation, which casts doubt over the efficacy of such education reforms.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Introduction

Many countries across the world, including the Kingdom of Saudi Arabia (KSA), have established the aim of reforming their education systems, with the goal of promoting their overall economy and society. Financial returns from international students enrolling at universities and from scientific research, funded by companies and governments to generate innovation and create products that will yield high financial returns – as in the field of medicine, computer programming or smart devices – have made education an important factor in the growth of national economies worldwide. This is why so many countries are keen to implement education reform (Au & Hollar, 2016; Bowe, Ball, & Gold, 2017), according to the World Bank (2008) and OECD (2013). As part of the package of education reform, one of the recommendations is for school performance evaluation (SPE), which involves the deployment of performance indicators to evaluate school performance (Ball, 2012a). This type of assessment is based on data that are collected and documented by head teachers and the teachers themselves, including students' exam results and parents' opinions. According to Ehren and Swanborn (2012), once collected, they are used as evidence to help judge how well a teacher or school is performing.

Although there are high expectations of applying SPE and its results, even its supporters do not deny its negative influence on teachers and schools (McVeigh, 2016). These influences have become the subject of considerable discussion in the research community, dedicated to school improvement and teachers' development, identity, wellbeing and professionalism. In Saudi Arabia, even though several SPE programmes have been implemented, such as the School Performance Indicators System (SPIS), where school performance is evaluated according to specific indicators (see Appendix I), the research to date has failed to ascertain its influence, particularly from the perspective of teachers and head teachers. Therefore, in this study, head teachers' perceptions and experience of SPIS have been investigated, as well as the influence of SPIS on:

1- School monitoring
2- Head teachers' stress levels, workload and morale
3- Teachers' stress levels, workload and morale

4- School improvement and accountability for this improvement (as defined by head teachers, teachers and SPIS).

These aspects are highlighted as the research focus for several reasons. The first point for consideration is school monitoring, since most SPIS processes take place during inspection visits to schools. Therefore, teachers' and head teachers' experiences of these visits are explored. The second and third points refer to stress, workload and morale, which are important issues that are likely to influence head teachers and teachers. These can be influenced by details such as the date of the inspection visit, how well staff are prepared for the visit, their understanding of its purpose or the focus of the inspection, and its results.

Empirical evidence on stress and workload resulting from SPE has been published in a considerable number of Western studies; for example, (Perryman, Ball, Maguire & Braun, 2011; Perryman, Maguire, Braun, & Ball, 2018), as well as in studies from the Arab world (see Jaradin, 2004, in the Jordanian context). Therefore, building on that evidence base, the current research has sought to establish whether teachers and head teachers experience issues of stress, workload and declining morale as a result of SPIS.

Finally, the relationship between accountability and school improvement is important for several reasons: the goal of SPE is to improve schools, and its results are then used to judge school performance, including the performance of head teachers and teachers. Consequently, head teachers' and teachers' voices are crucial in this case.

The current chapter establishes the study setting and gives an overview of the thesis, outlining the important points of this study. It identifies the research problem, questions and aims, and presents the main study objectives. Additionally, it clarifies the author's personal and professional interest in the topic, together with the study's significance and outcomes. Likewise, the conceptual framework of this research is described, and the basic structure of this thesis and its remaining chapters is outlined.

1.2 Identifying the Research Problem

Saudi Arabia is one of the most important countries in the Middle East and Islamic world, due to its geographical location, economy and religious status (see section 2.3). For example, its physical location, with a long border that is shared with the majority of Arab countries in the Middle East, gives KSA great political importance. The nation is also a member of G20: a

group of the world's biggest economies. In addition, it is acknowledged as the leader of the Islamic world, which is made up of around two billion Muslims distributed across the globe. This is because the Saudi cities of Mecca and Medina host the two most important mosques in the world: the Haram Mosque and Al-Masjid an-Nabawi.

Saudi Arabia has made considerable efforts to preserve and improve its position as a nation. It has therefore invested a huge budget in various service sectors, most importantly education (Al Sulaimani, 2010). In 2018, according to the Ministry of Economy and Planning (2015), KSA's education budget amounted to 33% of the total state budget of SR 364 billion. In addition, it became a member of the World Trade Organisation (WTO) in 2005, and that same year, participated in the International Mathematics and Science Study (TIMSS) and the Programme of International Student Assessment (PISA) (Tayan, 2017). PISA, designed and administered by the OECD (Rutkowski & Rutkowski, 2016), consists of a triennial international survey, aimed at evaluating education systems around the world by testing the skills and knowledge of 15-year olds. To date, students representing more than 70 economies have participated in this assessment since the year 2000 (Carnoy, Khavenson, & Ivanova, 2015). Breakspear (2012) claims that PISA has had a definite impact on national educational reform and policymaking. It is designed to monitor outcomes in relation to time, as well as shedding light on the factors that could account for differences in performance within and between different systems. Stremfel (2014) explains that PISA results, are:

> presented in comparative achievement scales, [aim to] provide an insight into how one educational system performs in comparison to other systems and also how one educational system contributes to the achievement of common goals of a particular group of participating countries. (p.29)

Following every PISA survey, there has been considerable debate over the test in many countries. For example, Germany's PISA results for the years 2000 and 2003 shocked the nation, in what is now known as the 'PISA shock' in the education landscape. Because of the unexpectedly poor results, compared to those achieved elsewhere in the world, Germany was obliged to reform its education policy (Breakspear, 2012). Meanwhile, in 2012, the PISA results demonstrated serious underperformance amongst Slovenian students in terms of literacy, emphasising a critical need for improvement (Breakspear, 2012).

In the Middle East, specifically in the United Arab Emirates (UAE), PISA scores were found to have improved in mathematics (by 13 points), reading literacy (by 10 points) and science (by 11 points) from 2009 to 2012 (PISA, 2009). However, in certain other countries, such as

Russia, the PISA results were lower than those obtained from other surveys, whereby Russia's TIMSS scores were higher than its PISA results. This has led to extensive debate surrounding the value of PISA as a trusted assessment tool (Carnoy, Khavenson, & Ivanova, 2015).

Nevertheless, despite high expenditure and substantial efforts, the Saudi education system has ranked very low in these international tests: its last result for TIMMS 2015 was 383 (TIMMS, 2015), which was almost half the score achieved by Singapore in the same test (see Figure 1.1). It was embarrassing for Saudi Arabia, which subsequently called for education reform, especially to narrow the gap between economic goals and educational outcomes. According to Maroun, Samman and Moujaes (2008), it consequently became an important topic of debate about education in Saudi Arabia, motivating the nation to reform its education system. These discussions have continued and focused on issues surrounding the quality of Saudi education, such as the lack of qualified teachers in rural areas; the unavailability of proper training for new head teachers, and the question of whether there is adequate educational investment in the population and its skills to enable the Saudi economy to meet future challenges and compete globally, should oil reserves run dry (Al Hakamy, 2008). Due to various issues, including a lack of effectiveness, school performance has especially suffered, which has attracted significant criticism over recent years.

It is against this background that Saudi Arabia has embarked on a programme of reform for its education system, seeking to find solutions to its problems. As a result, the government has launched Tatweer: a set of education reforms that reflect the requirements of neoliberalism in education (Wiseman, Astiz, & Baker, 2013). SPE is especially concentrated in a number of performance programmes, such as the Education Excellence Award, the Education and Training Evaluation Commission (ETEC) and SPIS (Ministry of Education, 2019). All these programmes apply indicators to evaluate school performance (ETEC, 2019; Ministry of Education, 2019).

Unsurprisingly, SPE and its influence on head teachers and teachers is widely discussed in many contexts across the world (Ball, 2003; Galton & MacBeath, 2008; Bailey & Colley, 2015). The above-mentioned studies refer to several effects of SPE based on indicators, such as increased stress and workload among teachers, but also a positive influence on school improvement (see section 3.4).

In Saudi Arabia, SPE has received considerable attention (for example, see Al Hakamy, 2008; Alrwqee, 2012), but its influence on head teachers and teachers has not been given the same importance. It is important to discuss the influence of performativity here, because the evaluation of school performance is new, with many points that need clarification to avoid negative effects, as well as numerous areas in need of improvement.

Teachers and head teachers are partners in the education process and the greatest effort in education always lies in their tasks. In fact, education reform is widely discussed in terms of the teacher's role in its implementation and the way that it is approached by teachers; for instance, whether they contribute to the reform, resist its application, or ignore it altogether (Fullan, 2007; Taylor, Yates, Meyer, & Kinsella, 2011; Jiang, Sporte, & Luppescu, 2015). Although teachers' participation in educational decision-making engenders their trust in the system and reduces their resistance to it (Hallinger & Lee, 2011), they are rarely involved in decisions over education reform (Jiang et al., 2015). However, this highlights the importance of teachers' commitment to the system, if they are to see anything positive in it (Moses, Berry, Saab, & Admiraal, 2017). In addition, if they are to develop a positive opinion of the system, teachers must be provided with clear information (Moye, Henkin, & Egley, 2005). This confirms the necessity to research teachers' opinions and experiences of the process of education reform. The benefits gained will relate to an understanding of their opinions and the impact of this process on them, so that their acceptance of reform can be addressed.

Therefore, teachers' opinions and experiences should be considered, as they have an influence on the application and effectiveness of government efforts to implement managerial approaches to education reform, which are rooted in neoliberalism (see subsection 2.2.2), including performance evaluation (see section 3.2). In particular, it is important to understand how SPE applies to teachers within a centralised education system, as is the case in Saudi Arabia. The negative impact of performance evaluation can have a significant impact on teaching performance, due to pressure and additional work; thereby indicating the importance of studying these effects, so that they can be addressed or avoided.

Figure 1.1 Saudi Arabia's ranking in TIMSS 2015

1.3. The Researcher's Personal and Professional Interest in the Topic

I was born and raised in KSA, and hold a Master's degree in Educational Administration from Umm Al-Qura University. I worked as a secondary school teacher for seven years, but I am also a writer with a column in the Saudi newspaper, *Al-Watan*. Most of my articles published in this newspaper deal with education reform, and I am well known in academic and education circles for my interest in issues of education reform and my criticism of the Saudi education system. My writing in the field of education has enabled me to monitor many issues in the above-mentioned context, especially regarding the implementation of new programmes, such as SPIS (see subsection 2.3.6.2) and its influence on schools and teachers. I have also been able to highlight many problems that could be addressed to help improve the system. Additionally, my work as a teacher has reinforced my contributions to *Al-Watan,* in terms of being able to expose the way in which female staff and other stakeholders suffer. This includes their lack of voice in decision-making, whether in education or in everyday life, and their lack of presence in leadership positions in the Ministry. Aside from this, I have already published a paper on the

topic of Saudi women in academic positions (see Alsubaie & Jones, 2017). All these aspects of my professional background and experience have influenced my choice of research subject for this doctorate, which is associated with education reform, as well as giving teachers the opportunity to express their views and discuss their experiences. This topic specifically concerns the influence of performance evaluation on teachers and head teachers, using SPIS as an example and exploring the perceptions of these education professionals. This proved to be a major challenge for me, but I did everything within my power to gain as much knowledge and experience as I could. I therefore believe that the findings will make a difference to my professional and personal life. Additionally, the Saudi education system is prone to change, which means that my results will be given careful consideration, especially during this period of implementing Saudi Vision 2030, wherein education reform is one of the most important pillars of the nation's development (Vision 2030, 2018)

1.4. Research Aims, Objectives and Questions

1.4.1 Study Aims

This study aims to explore perceptions and experiences of education reform among head teachers and teachers in secondary schools in Jeddah, using SPIS as a specific example. It also aims to fill several knowledge gaps in the field of performance management and its tools of evaluation. Specifically, this study investigates the perceptions of head teachers and teachers in secondary schools in Jeddah, as well as exploring the impact of SPIS processes, namely the use and awareness of key performance indicators (KPIs), SPIS, and the grading of outcomes relating to various aspects of school life, such as school monitoring.

1.4.2 Study Objectives

1- To explore the level of awareness of KPIs and their use in SPIS evaluation, as well as the ability of head teachers and teachers to read performance tables and demonstrate how KPIs indicate key strengths and weaknesses in SPIS evaluation.

2- To explore the frequency of SPIS monitoring of teachers, head teachers' awareness of when this monitoring will take place, and the various monitoring techniques used in SPIS.

3- To examine the extent to which SPIS evaluation can lead to head teachers' and teachers' perceptions and experience of heavier workload in schools.

1.4.3 Research Questions

The main research question in this study is worded as follows:

What are head teachers' and teachers' experience and perceptions of the influence of SPIS on school performance?

This question raises three sub-questions (RSQ):

RSQ1. What are head teachers' and teachers' experiences and perceptions of the influence of SPIS on school monitoring?

RSQ 2. What are head teachers' and teachers' experiences and perceptions of the influence of SPIS on their stress levels, workload and morale?

RSQ 3. How do head teachers and teachers describe and understand their accountability under SPIS in relation to school improvement?

1.5 Overview of the Theoretical and Conceptual Framework

Since the growth of neoliberalism in the 1980s (see subsection 2.2.2), most developed countries have looked for the most effective approach to education reform (McDonald, Pini, & Bartlett, 2019). For example, Ball (2017) states that during the 1990s, many schools and universities in the UK began to phase out traditional forms of educational governance; adopting practices from the private sector and corporate management instead – allegedly for political reasons. According to Tolofari (2005), the widespread economic problems encountered around the world, such as in the UK and USA, caused concern at this time amongst many governments, particularly in OECD countries (for example, the UK and USA), over receiving value for money (Curristine, Lonti, & Joumard, 2007). These countries, together with China, reported that their systems had become more consistent with the adoption of neoliberalism. According to Harvey (2007), this represents a turning point in economic history, whereby the leaders of these countries decided to re-design their systems according to neoliberalism.

To illustrate the above, Chairman Deng Xiaoping led the Chinese nation towards neoliberalism to recover the country's economy (Harvey, 2005). Meanwhile, in the USA, President Ronald

Reagan steered US government policy towards neoliberalism (Harvey, 2005; Robertson, 2008), and British Prime Minister, Margret Thatcher reduced tax to allow citizens greater freedom and personal control (George, 1999). This led to the rolling back of state provision, and the privatisation and marketisation of education in a reconfiguration of state powers. What is more, technological development and the exponential growth in the Internet, with the excessive use of communication tools, ease of information exchange, emergence of competition in the field of computer programming, and protection of information, could not be accommodated within existing forms of traditional education. However, centralised and bureaucratic procedures helped de-escalate the constraints on financial and technological development, which bore upon the learning process (Fusarelli & Johnson, 2004). All these factors drove governments to reform their education systems, while many experts for example, Ferlie, Hartley and Martin (2003) suggested new public management (NPM) (see subsection 2.2.3) as a solution to such challenges.

The most important changes brought about by NPM (see subsection 2.2.3) have included the local authority of powers over budget, appointments, personnel and planning; the sharing of powers amongst all stakeholders (head teacher, teachers, parents, governors and the neighbouring community); competition for pupils; and public accountability. (Ferlie, Hartley, & Martin, 2003)

Tolofari (2008) discusses these points in detail, claiming that roles and relationships in education have come to resemble those of the private sector. Thus, staff participation in school management is encouraged and school heads are delegated more authority than previously. The pattern of accountability under NPM means that all stakeholders, including teachers and parents, are made more accountable for the provision of education. Patterns of governance therefore offer schools greater autonomy, with NPM giving schools more control over their budgets, strategies and mechanisms. Moreover, the educational and other values that underpin schooling have been transformed by NPM, thereby impacting on the value of education and adding value to society. However, in order for NPM to be effective in meeting the requirements of education reform, it is proposed that performance management should be implemented in education (Tolofari, 2008). Moreover, the system of evaluation for education needs to be adapted, so that school performance can be evaluated (Wadongo, 2014).

The changes outlined above fall within the marketing of education, so that it can be pushed towards becoming a tributary of the economy. This means that education draws out students

with high-level skills to serve the economy in all fields, whether in industry, construction or the social and political sciences. However, this is impossible to achieve without improving performance (Proudfoot, 2018), meaning that performance management needs to be implemented in education (see subsection 2.2.4). Furthermore, to ensure continued improvement in education performance, teachers and head teachers must be held accountable. This notion would appear to be the cornerstone of education reform. However, neoliberal, NPM and other performance management policies have led politicians to impose their authority on education by determining the manner of its management, objectives and output priorities, in relation to what can be measured. In fact, performance is always linked to measurement (Wadongo, 2014), giving rise to performance indicators. These point to the level of success in school performance, often associated with student achievement. As such, performance indicators usually consist of national exam results, with teaching performance being evaluated according to students' performance in these tests (Rothstein, 2010), since students with high exam scores are seen as a reflection of teaching quality (Vinh, Chetty, Coppel, & Wangikar, 2011).

This evaluation of school performance has elicited intense debate amongst researchers in terms of its capacity to measure competence and ability, as well as its influence on teachers. This is especially pertinent where the results of performance evaluation are linked with teachers' salaries, promotion, incentives and school budgets, as is the case in the UK (Ball,2017), USA (Dee, Wyckoff, & Force, 2013) and KSA (Ministry of Education, 2019). In fact, evaluation has an influence, whether negative or positive, deliberate or unintentional (The World Bank, 2008). Researchers have explored this influence when applying indicators to evaluate school performance, which is referred to as performativity (see section 3.3.3; see also Ball, 2003; Perryman, 2006). In addition, it is discussed in the context of accountability, because governments use the results of this type of evaluation to ensure accountability in education, especially amongst teachers (Barzanò, 2009).

Nevertheless, the effects of the above have been found to vary. For example, Ball (2003) claims that performance-based accountability and evaluation influence the independence of the teaching profession, robbing teachers of their professional identity. In contrast, Yia and Kimb (2019) consider evaluation to have a positive influence, especially on the performance of school leaders, but do not ignore its negative influence on teaching performance. Additionally, Barzanò (2009) discusses this influence in terms of accountability based on performance

evaluation, including the actions involved in the latter. These range from the dissemination of school performance results, to the stress experienced by teachers. This stress arises from their scepticism that these results will be understood correctly by the general public and the sense of threat to their professionalism. In addition, researchers such as Ball (2003), Perryman (2006) and Mayer, Mitchell, Santoro and White (2011) have studied the influence of SPE on teachers' stress levels and workload, as well as on school improvement, in a debate that is addressed in detail in subsections 3.3.4 and 3.3.5.

Correspondingly, the literature reviewed here deals with three strong themes relating to the current research questions: the evaluation of school performance (see section 3.2), as this is the mission of SPIS in evaluating schools in Saudi Arabia; teachers' accountability and its influence on overall school improvement, and teachers' stress, workload and morale (see subsections 3.3.4 and 3.3.5). Stress and workload have been found to increase following SPE; ensuring continued improvement and holding teachers accountable for their performance in schools (The World Bank, 2003). Besides, this performativity (see section 3.3), which involves the use of performance indicators to evaluate school performance in numerous programmes across the world (for example, Ofsted in the UK and SPIS in Saudi Arabia) can have a profound impact on teachers' stress, workload and morale. These concepts are examined here to try and understand and analyse the perceptions of head teachers and teachers, regarding the influence of SPIS on school performance, according to its effect on school monitoring; teachers' stress levels; workload; morale, and school improvement. However, as may be seen from Figure 1.2, there are relationships between neoliberalism, NPM and performance management, whereby neoliberalist theory has given rise to NPM. This approach uses performance management and its evaluation to make judgements about performance, applying specific indicators and then holding the actors accountable. It points to the potential influence of performance management on schools, school head teachers, and teachers in this study, with respect to the development of school performance, and the stress levels, workload and morale of head teachers and teachers.

.
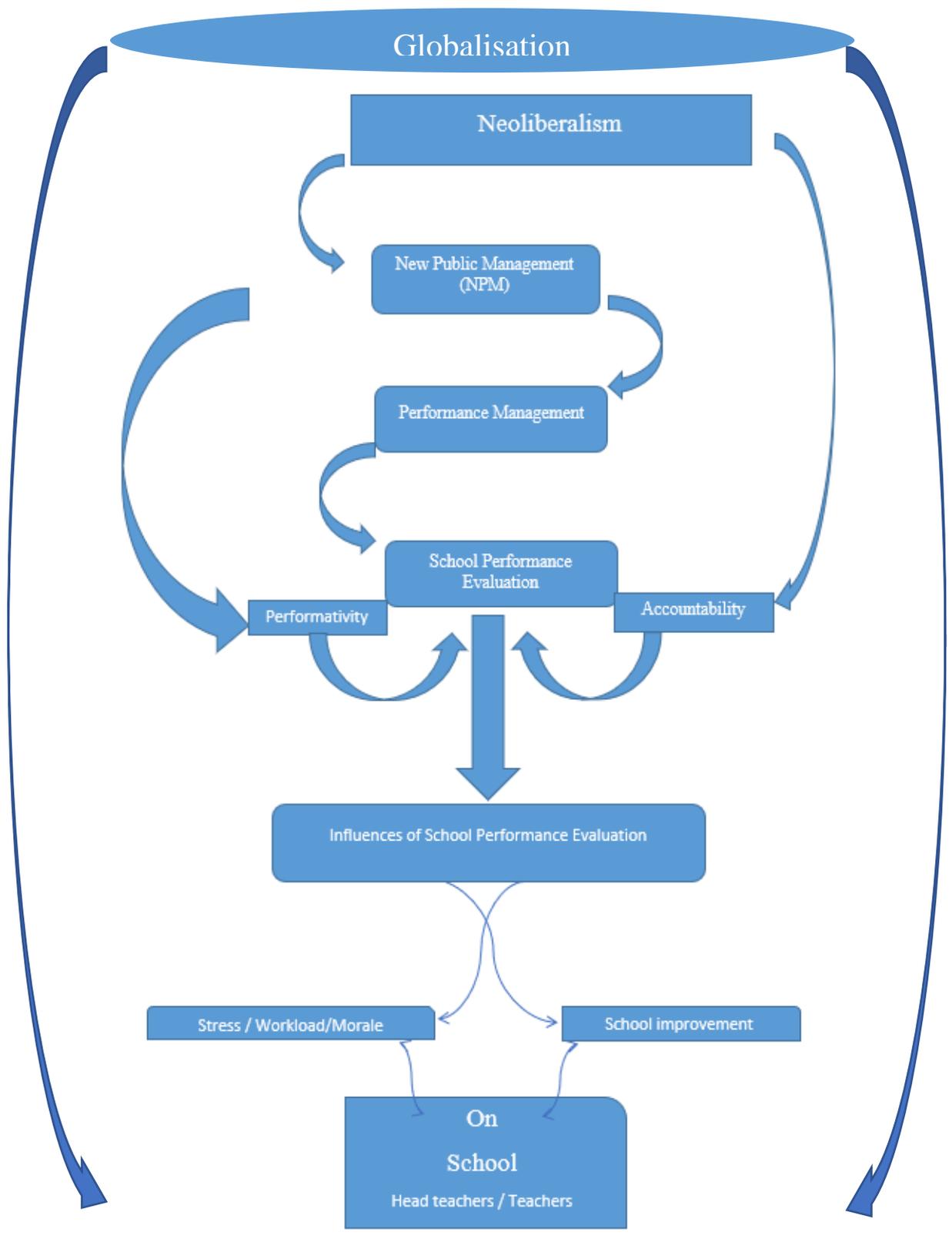
Figure 1.2 The relationship between the three main concepts of SPE, accountability and performativity

1.6 Overview of the Methodology

An explanatory sequential multi-site case study mixed methods approach was adopted in this study. As the researcher, I selected both quantitative and qualitative methods, which is appropriate for explanatory sequential mixed methods. This was initiated with two Google Drive surveys of all head teachers and teachers in secondary schools in Jeddah, whereupon a link to the surveys was distributed to the administrators of WhatsApp groups by the directors of Education Offices in Jeddah, the second largest city in Saudi Arabia in economic terms. An invitation to participate in the questionnaire was consequently sent to the head teachers and a link to the questionnaire was sent to teachers via these school WhatsApp groups. As a result, 64 head teachers and 109 teachers participated voluntarily in this study. The second research phase began with the selection of three girls' secondary schools in Jeddah, located in different parts of the city: Central, North and South. These schools were of different sizes: large, medium and small (see section 4.7). There followed meetings with the schools' head teachers and three teachers from each school in face-to-face, semi-structured interviews. The research methodology will be discussed in detail in Chapter 4.

1.7 Significance and Outcomes of the Study

The most significant contribution of this study is the female 'voice' of a group of people who are largely absent from the literature on education leadership in KSA. An important aim was to begin bridging the gap in the extant literature on what is known about the effects and impact of SPIS evaluation on schools, teachers and head teachers in Saudi Arabia. From a practical perspective, it was anticipated that the results of this study would suggest recommendations for improving SPIS in Saudi education. More specifically, this research was intended to deepen understanding of the perceived effectiveness of SPIS on school performance; while also helping SPIS evaluators to improve the process and refine the performance indicators, in light of the teachers' and head teachers' views of the strengths and weaknesses of SPIS evaluation. This was believed to be useful, as the Saudi economy was facing serious issues in 2016 when the study commenced, after a significant drop in oil prices. As a result, the Saudi government was forced to develop a new economic plan, referred to as its Vision 2030 (2018). This plan was aimed at finding alternative economic resources, other than oil (Vision 2030, 2018). One of the resources identified was investment in people. Thus, it was recommended that the Saudi Ministry of Education develop its national educational criteria and performance indicators.

It is hoped that this study will inform and facilitate the selection of criteria and indicators for an appropriate evaluation system to enhance Saudi education.

- ➢ This study demonstrates the relationship between performance evaluation and increased pressure on head teachers and teachers; pointing to the relationship between accountability and its negative influence on head teachers and teachers, as well as its importance in pushing these professionals to work towards reforming their performance.

- ➢ This study introduces a new context into the literature on performativity. To my knowledge as the author, in the Saudi context and in the literature related to Saudi education, this study is the first to express a research interest in this area.

- ➢ This study also highlights the importance of having a balanced accountability system, which can push a teacher to perform complex tasks without exerting any additional pressure or demanding extra work. It also reveals the relationship between the quality of the procedures in the evaluation process, and the enormous increase in the burden and pressure on teachers as they undertake their duties in school.

- ➢ This study points to the importance of redefining the teaching profession, so that it accounts for the influences of neoliberalism and the requirements of economics of knowledge.

1.8 Structure of the Thesis

Chapter 2 'contextualises the study', providing an overview of its international context. It concentrates on globalisation, neoliberalism and performance management, and provides detailed information about the study context of Saudi Arabia, its education reform and system of education evaluation, giving an overview of its geography, history, social context and economy.

Chapter 3, the 'Literature Review', describes and analyses the relevant literature in greater detail, including work on SPE, accountability and performativity. Additionally, empirical research studies on three conceptual frameworks (SPE, performativity and accountability) are presented.

Chapter 4, the 'Research Design and Methodology', presents a detailed view of the study's research paradigm and the justification for its adoption in this study. It gives a profile of the

participants in the sample, as well as describing the methods of data collection and analysis, defending the study's validity, and outlining the ethical considerations.

Chapter 5 presents the 'Results and Findings' in an analysis of the quantitative and qualitative findings.

Chapter 6, the 'Discussion', interprets the results in relation to the Literature Review and associated theories, which are discussed in this study in light of the research questions.

Finally, Chapter 7 summarises the key findings of this study for each research question, in a 'Conclusion' to the thesis. This is followed by a number of recommendations, as well as the limitations and implications of the research. The Chapter concludes with some suggestions for future studies.

Chapter 2: Contextualising the Study

2.1 Introduction

This chapter sheds light on significant aspects of the international education context in general and the Saudi context in particular. It begins with an overview of globalisation as a driver of worldwide education reform, explaining its roots in neoliberalism. Correspondingly, the current chapter examines NPM and its relationship to education policy and performance management. It then proceeds to describe the specific context of this study, namely KSA and its geography, demographics, society, culture, language, religion, economics, and education system, including the pathways that it has taken towards education reform and globalisation.

2.2 The International Educational Context

2.2.1 Globalisation

The effects of globalisation on education policy have become a controversial issue worldwide, especially in countries such as Saudi Arabia, where the Islamic religion governs all aspects of life, including education (Oyaid, 2009). There are approximately 200 definitions of globalisation in the academic literature (Sheffield, Korotayev, & Grinin, 2013), but no single uniform meaning (Dreher, Gaston, & Martens, 2008; Lingard & Rizvi, 2010). Some scholars have defined it according to its impact. For instance, Sheffield et al. (2013) state: "we see globalisation as the expansion of social systems and the increase and complexity of common social bonds among societies" (p.22). In contrast, many other scholars refer to globalisation in terms of its approaches. For example, Angus (2007) views it as an external phenomenon, which has given rise to mechanisms such as neoliberalism in administrative competition and market arrangements. However, even these two definitions fail to provide a comprehensive description of globalisation, because it is greater than its impact or approach. Going some way towards bridging this gap, Ritzer (2010) states that globalisation is "a transplanetary process or set of processes involving increasing liquidity and the growing multi-directional flows of people, objects, places and information as well as the structures they encounter and create that are barriers to, or expedite, those flows" (p.2). This is the working definition adopted for this thesis.

According to Liu (2015), the whole world is undoubtedly in an era of globalisation. Coleman and Jones (2004) also highlight its inevitability. It is generally agreed that globalisation has led

to greater ease of movement, and the consequent emergence of free trade between countries, which has not only enabled the transportation of goods, but also of cultures (Baldwin, 2006). Generally, this flows outwards from the West to the rest of the world (Ritzer & Dean, 2015). One outcome relevant to the current thesis topic is the emergence of the Global Education Industry (GEI) (Verger, Lubienski, & Steiner-Khamsi, 2016). Verger et al. (2016) state that "school improvement services, online education, or tutoring complementary education /'shadow', e-learning and marketing, advisory services to governments and schools, [and] test preparation services" (p.4) are the components of the education industry. An interesting example is the Omega School Foundation, established by James Tooley in 2008 and financed by Pearson's Affordable Learning Fund to institute high quality schools at low cost across the world (Omega, 2019). Other providers sell international test systems, such as PISA and TIMSS, according to Sahlberg (2016). This has been made possible by global standards for determining the best education, and the standardisation of teaching and learning in schools (Sleeter & Carmona, 2017). In Saudi Arabia, the Ministry of Education has transferred ideas from Ofsted in the UK and adapted these to the latest comprehensive evaluation system (Alrwqee, 2012). Meanwhile, Mukherjee (2015) attributes such a flow of ideas to the power of the West and its ability to influence the culture, politics and economies of the rest of the world (Baldwin, 2006; Dreher et al., 2008; Lingard & Rizvi, 2010).

As a result, serious debate has surrounded the topic of globalisation (Liu, 2015). There are those who believe that globalisation leads to conflict between cultures and social groups, while others see it as the potential fusion of cultures to form new ones (Pieterse, 2019). Saudi commentator, Al-Ghadhami (2013) considers globalisation to be a platform for a system that represents the triumph of Western capitalism over other cultures. In contrast, Mukherjee (2015) believes that globalisation has done a great deal for mankind, such as helping to tackle poverty and gender inequality, as well as promoting cultural exchange and stimulating debate on human rights. Certainly, globalisation has encouraged many countries to introduce new policies and systems as part of a package of education reforms that are sweeping the globe. Mundy, Green, Lingard and Verger (2016) claim that this is partly inspired by the view that education is a driver of economic growth and wellbeing. Thus, education reform may address broader economic and social problems within a country. Although this study discusses education reform in Saudi Arabia, its focus is on its influence on head teachers and teachers. Nevertheless, it provides insights into the effects of globalisation, whether positive or negative, and especially on the most important partners in the education process, namely head teachers and teachers.

It is widely acknowledged that globalisation has far-reaching effects and needs to be understood from different perspectives; for example, in terms of social interactions, whereby people from different countries around the world can communicate with each other easily via the Internet and social media websites. This can lead to a flow of ideas and greater understanding between people from different cultures and as a result, greater acceptance of different views, especially among young people from around the world. It can ease the transmission of ideas and beliefs and therefore shape culture worldwide; giving rise to the creation of united communities from different regions, who converge to address global issues and matters such as women's rights. Additionally, globalisation has an influence on political decision-making and policy. The rapprochement between nations in other parts of the world has caused many Saudis to call for political change, which will give them the same opportunities for political and democratic participation as the developed world, as well as a word-class standard of education. Globalisation has therefore contributed to change in the relationship between the people and the state. As a consequence of this shift, the public's demands have increased, based on the expectation that the government will improve their situation. One very recent example of the power of the will of the people in Saudi Arabia was the government's decision to permit women to drive; making Saudi Arabia the last country in the world to do so.

Turning now to globalisation and its role in education reform, the following subsection will discuss neoliberalism as a theoretical concept that drives policy initiatives.

2.2.2 Neoliberalism

Harvey (2007) refers to neoliberalism as "a theory of political economic practices that proposes that human well-being can best be advanced by liberating individual entrepreneurial freedoms and skills within an institutional framework characterised by strong private property rights, free markets, and free trade" (p.22). This definition places an emphasis on individual skills, which can contribute to wellbeing and the good of the economy. Harvey's inclusion of 'skills' in this definition therefore highlights the role of education in the development of human competences, which equip people with the necessary abilities to conduct business and become entrepreneurs, in the assumption that they will be successful. From this definition, we find a strong link between politics, the economy and education. According to Morel, Palier and Palme (2009), it

is the economic challenges faced by European countries that have led them to invest in education. This establishes a clear link between education and neoliberalism.

Based on these principles, critical theorists such as Ball (2003) argue that neoliberalism serves to generate profit for the private sector. He is particularly critical of the way in which neoliberalism has led to new managerial practices, such as NPM or managerialism, and in turn, performance management practices in education. Ball (2003) argues that the use of indicators to evaluate school performance has resulted in a culture of performativity (see section 3.3) and has become a new mode of state regulation, making it possible to govern in an 'advanced liberal' way. The terrors of performativity for teachers have been of particular concern to Ball (2003) and Perryman (2006), amongst others.

2.2.3 New Public Management (NPM)

NPM is generally considered to be a global phenomenon. It has therefore generated a substantial body of literature (Tolofari, 2005). Much of this literature is dedicated to the definition of NPM and its similarity to, or differentiation from managerialism. Therefore, simple definitions are not easy to form (Pollitt & Dan, 2011) and there is an apparent lack of general agreement on the definition of these two terms, rendering any discussion of them difficult. For instance, it is unclear whether they refer to one management topic or two. According to Randle and Brady (1997), managerialism and NPM are synonymous. In contrast, Boyne (1996) claims that NPM is merely an element of managerialism. Although Tolofari (2005) supports this claim, he also argues that managerialism is a feature of marketisation, which characterises NPM. The critical questions that need to be addressed consist of whether NPM comes from managerialism or managerialism comes from NPM and if there is a good reason for differentiating between them. There is in fact strong evidence that the term, 'managerialism' is sometimes used when discussing NPM, such as in Deem, Hillyard, Reed and Reed (2007), and Lynch, Grummell and Devine (2012). As a result, the terms, 'NPM' and 'managerialism' appear to have been used interchangeably by some commentators (Tolofari, 2005).

Similarly, there is disagreement over the specific meaning of NPM (Kalimullah, Alam, & Nour, 2012). However, for the purpose of this study, NPM is defined as the process of transferring private sector management systems to government and public sectors.

2.2.3.1 Application of New Public Management (NPM) in Education

From the late 1970s to the 1990s, governments across the world applied NPM to various industry sectors, including education. This practice has continued through to the present day (Lynch et al., 2012). Focusing on education, Lynch et al. (2012) state that during the 1990s, numerous schools and universities in the UK began to phase out traditional forms of educational governance, adopting private and corporate management practices instead. In fact, there were political reasons for this, as the new Conservatives – in power in the UK between 1978 and the early 1990s – were the first to implement NPM in education (Lynch et al., 2012). Under NPM, "schools would be given autonomy, reducing government intervention and allowing them to manage their own financial affairs" (Exley & Ball, 2013). Even when Tony Blair, leader of Britain's Labour Party, was subsequently elected Prime Minister, NPM was retained in schools and universities across the UK (Taylor et al., 2011; Exley & Ball, 2013). These similar practices of the New Conservatives and New Labour were a consequence of neoliberal principles of freedom of choice, competition and market choice, which led to the appearance of independent schools, academies, etc.

Meanwhile, on the other side of the world, Saudi Arabia began to implement many features of NPM, such as privatisation (Fattore, Dubois, & Lapenta, 2012), as well as establishing companies to manage initiatives to improve education, supervised by independent authorities like the Tatweer Company (Tayan, 2017). There are a number of reasons for this. According to Tolofari (2005), some countries around the world, such as the UK and USA, faced economic problems during the 1980s. Therefore, the governments of OECD member countries (in particular, the UK, USA and Canada) were concerned about achieving value for money in state-run sectors such as education (Curristine et al., 2007). Added to this, technological developments could not be accommodated within traditional forms of education. Constraints on financial and technological development, which affected learning, were in part de-escalated through being decentralised from bureaucratic procedures (Fusarelli & Johnson, 2004). All these factors drove governments to reform their education systems. As Hartley (1997) notes, the use of managerialism was viewed as a solution to the accompanying challenges.

The most important changes brought about by NPM involve the "local authority of powers over budget, appointments, personnel and planning; the sharing of powers amongst all stakeholders (head teachers, teachers, parents, governors and the neighbouring community); competition for pupils; and public accountability". Ferlie et al. (2003) explore these points in

detail, claiming that in many parts of the world, roles and relationships in education now resemble those of the private sector, where staff participation in school management is encouraged, and heads of schools are delegated a greater degree of authority. Under NPM, all stakeholders, including teachers and parents, become more accountable for the provision of education, and patterns of governance grant schools a higher level of autonomy. Thus, NPM allows education institutes more control over school budgets, strategies and mechanisms. Meanwhile, NPM has altered the educational and other values that underpin schooling. Although many principles of NPM have been adopted in Saudi Arabia, the country has traditionally had a centralised education system (see subsection 2.3.6).

2.2.3.2 Debates Surrounding New Public Management (NPM)

There is considerable debate surrounding the effects of NPM. Early commentators, such as Hartley (1997), claimed that NPM had been a successful approach to solving problems related to administration, but this belief is no longer accepted in the management literature (Overeem & Tholen, 2011). Dibben, Higgins, Dibben and Roper (2004) also note that NPM has failed to address efficiency and accountability, and has not achieved leaner government or stronger governance. Bessant, Robinson and Ormerod (2015) highlight how the application of NPM has resulted in significant changes to the public sector ethos and management practices; for example, the development of new management practices, marketisation, the contracting out of core services to private companies and non-profit organisations, and the creation of executive agencies to take responsibility for implementing NPM. Dunn and Miller (2007) highlight that this has caused some commentators to reject NPM, such as Kalimullah et al. (2012), whose criticisms are based on the view that NPM contravenes the rules of democracy, because it does not allow the government to control public services. Therefore, such researchers argue that NPM can lead to an unfair distribution of services, even though NPM incorporates the means of overseeing its own operation. Critics have also warned of diminished political accountability, due to changes in the public sector. However, although this may be true initially, reduced accountability of this nature will only persist in reforms for a limited time, after which political accountability should increase.

As observed by Hood and Peters (2004), elements of NPM are still developing, with some evidence of success in countries like India and Japan, where PISA results have actually improved (OECD, 2013). Consequently, it is claimed that these countries have found it to be a

fitting solution to improve their public sector systems (Tolofari, 2005). Thus, it could be stated that education reform stemming from NPM needs time to develop and demonstrate its impact. There is no doubt that the success of any approach will depend on how well it is implemented. Some experts in the field of NPM, such as Hughes and Teicher (2004), have warned of the difficulties involved in applying NPM effectively in the developing world. In fact, developing countries that have attempted to implement NPM have experienced uneven effects (Pollitt & Bouckaert, 2004). Nevertheless, although attitudes to NPM vary across the literature, there is evidence that it is still popular on a global scale.

2.2.4 Performance Management

The term, 'performance management' has evolved due to expanding research in the field since the 1980s. As a result, there has been a broad range of interpretations of the term in the literature (Tam, 2008). The traditional view of 'performance' refers to individual or organisational performance (Mackie, 2008), as well as the performance of leaders and employees, or individual achievement under specific circumstances. Some scholars argue that 'performance management' is a term that is used extensively but loosely (Guzzo & Dickson, 1996). Conversely, others consider it to be a multidimensional term, with a range of meanings, which makes it rather difficult to formulate a precise definition (Javadi, 2013). Although not everyone agrees with this argument, Tangen (2005) claims that the definition is clear, but it depends on the angle from which the performance is considered. Similarly, Dickinson, Watters, Graham, Montgomery and Collins (2009) suggest that the term can be defined from various perspectives. The first of these positions considers performance in accomplishing tasks, evaluating all the steps towards goal achievement and the correction of any mistakes. The second looks at conveying this performance to stakeholders, so that accountability and governance can be established (Mackie, 2008). From this perspective, "performance is not only a concept, but also an agenda" (Van Dooren, Bouckaert, & Halligan, 2010), presented with all components of the task performed to deliver quality (Dubnick, 2005).

Correspondingly, these distinct functions of performance are clear in many of the definitions provided in the literature. For example, Lebas (1995) suggests that performance management offers a means of successfully achieving future targets and objectives. This is supported by Dickinson et al. (2009), who state that it may be defined as the efforts made to achieve an organisation's objectives in an efficient and responsible manner. However, both these

definitions ignore certain important elements of performance management, such as a focus on performance evaluation, and the improvement in individual and team performance within an organisation. Moreover, recent studies have contributed diverse definitions of performance management; for example, according to Nielsen (2013), it refers to a continuous process, where the top management first identifies and then sets organisational goals and targets, for which subordinate managers are accountable. Another definition of performance management states that performance improvement is as important as its development, leading to the achievement of individual, team or organisational effectiveness (O'Reilly, 2009). Meanwhile, Fletcher (2004) draws attention to the fact that performance management is not 'a package solution'; instead, it is something that needs to be developed specifically and individually for the organisation concerned. The crucial goal of performance management in education is to improve the performance of a school and its head teachers and teachers.

In addition to the above, there are two terms that relate to performance management. The first of these is 'performance measurement' and the second is 'performance indicator'. According to Wadongo (2014), performance management is applied in the public sector to achieve appraisal results. The development of managerial practices ensures that the organisation meets its aims and objectives and continues to satisfy its stakeholders, including parents, government representatives, and school leaders in the case of schools (Franco-Santos, Rivera, & Bourne, 2014). Therefore, performance is a crucial component of performance management, whereby performance indicators must be applied to enable performance measurement. In an education context, these indicators are designed to provide the system with information about education performance (Nuttall, 2017). For the purpose of this research, performance evaluation is defined as the specification of indicators, with which authorities can measure the activities and actions that take place in schools.

2.2.4.1 Application of Performance Management in Education

After the Second World War, education was reformed for economic reasons over much of the world (Whalley, 2011). The period 1960-1980 was then largely characterised by the relative autonomy of the teaching profession and the management systems that controlled it. Subsequently, according to Whalley (2011), the 1980s represented a golden age of teacher autonomy in the UK. During this period, teachers remained predominantly self-accountable through internal reflection and peer review. At the same time, they worked in organisations

that were established and controlled by head teachers and governing bodies. Most of the evidence relating to the performance of pupils, teachers and schools was collected informally, while views on highly valorised 'professional' behaviour centred on ethical commitment, bureaucratic administrative skill, and expert judgement (Whalley, 2011).

From 1987-1994, however, education in many countries such as Australia, New Zealand and the UK faced significant shifts in terms of the appraisal process and development of new strategies for changing these systems as a whole (Whalley, 2011). Consequently, traditional inspection regimes were abandoned in these countries and many new mechanisms appeared. This is because the implications of NPM, as in the case of the UK, required important reforms to education – although this has not been the only factor driving education reform in the UK. Worldwide, economic problems have also increased over time, such as the growth of competition (Huggins & Williams, 2011), especially in countries like China and Taiwan, which are characterised by cheap labour. Many products have also lost value, like coal, which has been replaced by oil. Moreover, China and Taiwan have experienced high unemployment, and in future, many jobs will involve knowledge of economics and the service sector (Kaplinsky & Morris, 2016).

Another concern that arises in any discussion of NPM in education is the reason for its emergence. According to Naidoo and Jamieson (2005), education has become a product to be bought and sold; a view that is supported by Barr and Christie (2015), who highlight the existence of NPM in many areas related to education in the UK. First, education can have economic benefits, such as improving a recipient's potential for employment, or generating income from enrolment for the providers, thereby allowing them to improve facilities or staffing to benefit students. Second, education can create value for the state by producing a skilled workforce. Given these potential assets, standards should be in place to measure the performance of all stakeholders involved in education provision, particularly following a move towards standardisation that will affect all aspects of life (Barr & Christie, 2015). More specifically, from the years 1987-1994, the standardisation of education was considered to be a matter of priority in the UK and USA, as a means of ensuring the success of their education reforms, thereby easing the pressure on politicians. Tolofari (2005) adds that one of the most important contributions made by liberal governments in the UK is performativity, which may be defined as a culture of quantitative performance measurement, based on input and output (Brown, 2015) (this is explained further in section 3.3).

In 2003, just a few years after the above-mentioned period, the UK government established a new measurement for professional teaching standards, as well as performance management arrangements for teachers and head teachers (Walker, Jeffes, Hart, Lord, & Kinder, 2011). Meanwhile, in KSA, the Ministry of Education launched a comprehensive schools evaluation project in 2004, adapted from Ofsted's performance appraisal criteria in the UK (Alrwqee, 2012).

Furthermore, in the UK, head teachers and governors showed themselves to be very aware of the revised performance management regulations introduced in September 2007 (Walker et al., 2011). The new measurement of professional standards clearly identified the teaching skills, knowledge and understanding required by schoolteachers in the UK context (Evans, 2011). However, Miller, Ochs and Mulvaney (2008) warned that nearly 11,000 teachers would consequently be obliged to abandon their teaching careers, because they would still be unable to meet the standards, even a year after their implementation. This indicates the importance of discussing the impact of applying standards on teachers and head teachers, which is what is addressed in the current study in the Saudi context, following the introduction of SPIS. Saudi education established its education license in 2019, relying on specific performance indicators to measure teaching performance (Ministry of Education, 2019). These indicators included any academic qualifications attained and the outcomes of a performance test (Ministry of Education, 2019).

Interestingly, the methods implemented to establish performance management in school systems worldwide share a number of important features. One is the promotion of school autonomy, which can be observed in countries such as the USA, UK and Australia (Apple, 2004; Clark, 2009). It is rooted in the notion that autonomy enables schools to control their own performance in meeting education policy goals, while at the same time remaining subject to government control. This is because they are required to apply the performance indicators that are specified by stakeholders, including the government (Apple, 2004). It therefore renders the whole idea of autonomy in schools questionable. The second impact of performance management is highly visible, taking the form of school league tables. In the UK, these are published annually by the Department for Children, Schools and Families (DCSF) (Leckie & Goldstein, 2009). According to Herbert and Thomas (1998), league tables have been available in the UK since 1991, whereby students' results are increasingly viewed as a reflection of teaching performance and ultimately, of school performance. It is consequently argued that this

has led to a shift in accountability from teacher professionalism, characterised by accountability of teachers to themselves, their colleagues and their students (self-regulation), to accountability to agencies such as the Qualifications and Curriculum Authority (QCA) and Ofsted (Perryman et al., 2011). However, the pressures accompanying the implementation of performance management have had an effect on teachers' wellbeing and teaching performance, which will be discussed further in subsection 3.3.3. Conversely, as noted previously, Hill and Andrews (2005) claim that the implementation of performance management has had a strongly positive impact on education, which will be discussed in more detail below.

2.2.4.2 Apparatus for Evaluating School Performance

School Monitoring/Inspection Visits

The most prominent features of performance measurement in schools consists of monitoring and inspection visits (see subsection 2.2.4). This may be observed from the education charters of several countries, such as the UK's Education Act of 1988 (Jones & Tymms, 2014), New Zealand's Education Act of 1989 (Sakura, 2007), and Saudi Arabia's Seventh Development Plan 2000-2004 (Ministry of Economy and Planning, 2015). All the above share the common purpose of improving the quality of school performance (Ahmed, 2019).

There is no doubt that the evaluation of school performance can be useful in identifying the strengths to be supported and the weaknesses to be addressed, thereby enhancing school performance and achieving the government's educational goals – which are often linked to economic neoliberalism (Tomlinson, 2005; Altrichter & Kemethofer, 2015). Governments across the world have consequently begun implementing such systems to monitor education throughout all its phases. Perryman (2006) calls this 'panoptic performativity' (p.25). The procedures for these controls vary from one country to another; for example, in the UK, Ofsted is responsible for evaluating school performance (Altrichter & Kemethofer, 2015), while in Germany, this responsibility falls upon Konferenz der Kultusminister (KMK), a standing Ministry committee for education and cultural affairs (Dedering & Müller, 2011). The difference between these bodies rests in the fact that Ofsted is part of a national system of school inspection, while KMK is not (Dedering & Müller, 2011). In contrast, Saudi Arabia's SPIS is a national system of school inspection, but it is not the only one that evaluates school performance; the administrative supervisor in school administration departments also requires assistants to evaluate school performance. For this purpose, the director selects teachers at

random. Moreover, the educational supervisor inspects every teacher in a school to assess his or her performance (Ministry of Education, 2019). In contrast, when Ofsted was first formed in the UK, its work was mainly organisational, with the responsibility for evaluating school performance being transferred from the Local Education Authorities (LEAs) (Ouston, Earley, & Fidler, 2017). Meanwhile, in Saudi Arabia, school supervision of all kinds is still in operation (see subsection 2.3.6.2).

Ofsted has changed its methods many times. For example, at one time, it would evaluate schools every four years (Moreton, 2015), but more recently, each school has been visited according to its grading in the previous inspection. For instance, schools that have previously achieved a high grade are visited every three years, but if they require improvement, they are visited every two years (Jones & Tymms, 2014; Roberts & Abreu, 2016). Conversely, according to the Organizational Guide for Assessing School Performance, SPIS evaluates schools every year (SPIS, 2017).

In the UK, regarding issues surrounding the advance notice of inspection, schools are generally informed at midday, the day before the inspection (Roberts & Abreu, 2016). However, in Saudi Arabia, no notice of this kind is provided for in the Organizational Guide for Assessing School Performance (SPIS, 2019) or on the Ministry of Education website, which means that each education office in KSA establishes its own system of notifying schools of the dates of monitoring and inspection visits. Besides, in terms of informing head teachers and teachers of the evaluation system, Ofsted encourages schools to conduct self-evaluation after an external inspection, meeting head teachers and teachers to discuss the results (Ouston et al., 2017). Consequently, it has been suggested that teachers become less wary of the system and are less anxious about the process (Lowes, 2016).

Performance Indicators

Performance indicators represent contemporary methods of judging the quality of education (Drummond, 2011). They have been the topic of extensive debate, particularly with regard to their definition, their potential to enable conventional judgment, the effects of their implementation, and the difficulties of applying them in schools (Evans, 2011). Admittedly, the definition and establishment of performance indicators as a concept has taken highest priority. Between the 1980s and 1990s, a performance indicator was defined as "an item of information collected at regular intervals to track the performance of the system" (Fitz-Gibbon

& Taylor, 1990). In reality, performance indicators offer a means of measuring achievement in schools, so that a picture of their performance can be obtained (Rowe, 2004). However, according to Parmenter (2015), the benefit of implementing performance indicators is limited, if an organisation fails to redefine them according to its evolving goals and expectations, with a view to enhancing its organisational outcomes.

Nevertheless, the success of an organisation does not depend on the quality of the performance indicators applied, but rather on the success of its implementation procedures (O'Reilly, 2009). O'Reilly (2009) identifies common errors of implementation; citing the example of a mismatch between organisational values, a lack of clarity, inconsistency, and a failure to prioritise. This may stem from a problem highlighted by Mayston (1985), namely that performance indicators stem from welfare theory and information economics, while school performance indicators originate in education theory. To be more precise, performance indicators flow from implementation (input) and use (process and output), to the results (outcomes) of performance measurement systems (Johnsen, 2005). For example, efficiency can be measured by looking at the relationship between input (teachers) and output volumes, as well as the amount of time spent per pupil, which relates to cost-efficiency (Sutherland, Price, Joumard, & Nicq, 2007). Thus, school systems should use available inputs in the best possible way to be considered efficient and avoid unnecessary expenditure of public funds (Afonso & Aubyn, 2006).

Meanwhile, effectiveness is defined as the relationship between outcomes and input. In order to measure this efficiency, academic results from schools are considered to be an indicator of school efficiency (Drummond, 2011). This has encouraged teachers to focus on students' test performance (Perryman et al., 2011), which raises questions over the role of education in furnishing students with the values and skills that exams do not measure. According to Ball (2003), these are currently relegated to secondary importance. Moreover, Goldstein and Thomas (1996) add that there are difficulties involved in the application of performance indicators. First, instead of trying to establish factors to explain differences between schools, performance indicators mainly refer to school rankings. Second, some studies have shown the usefulness of performance indicators to be limited, with unreliable judgements being formed about institutions; for example, some experts remain unconvinced that performance indicators enable accurate judgments to be made. Still, many education systems have established monitoring programmes based on performance indicators (Willms, 2003).

School Performance Leagues

Since 1992, school performance league tables have been published in many countries across the world, such as the UK and USA (Leckie & Goldstein, 2009). These have been the result of increasing demands for education reform, a culture of accountability, and an increase in the number of ways in which education can be monitored (Perryman et al., 2011). For example, this growth in accountability culture has led to the UK establishing school league tables, so that parents can choose good schools for their children. The schools that are most frequently selected by parents receive more funding than those that are less sought after. This was confirmed by government charter (Duggett, 1998): "your choice of school directly affects the school's budget - every extra pupil means extra money for the school" (p.121). However, according to researchers like Leckie and Goldstein (2009) and Perryman et al. (2011), the heavy emphasis on school rankings in league tables gives the government greater control over schools and more leverage to impose its polices.

In addition to autonomy, however, there is also the issue of accountability (Parameshwaran & Thomson, 2015). Various stakeholders hold teachers accountable for their performance, and schools apply performance indicators, such as those issued by the government. Conversely, the validity of performance league tables has been widely contested, because they are based on the results of public examinations; for example, GCSEs in the UK (Ball, 2017)

 According to Leckie and Goldstein (2009), this has led the UK government to include 'added value' as an indicator of student performance, taking into account previous achievements. As a consequence, teachers are subjected to closer and closer monitoring (Apple, 2004).

Thus far, this chapter has concentrated on contextualising the study on a broader global scale, highlighting the key factor of globalisation in the neoliberal approach to education reform, of which NPM is an example. In the next section, the specific context of this study, namely KSA, will be discussed.

2.3 Saudi Arabia

2.3.1 History

According to Abisaab (2015), "writing on any aspect of Saudi Arabia's history is an arduous task", due to conflicting historical records and unsubstantiated evidence (p.2). The First Saudi

State was founded in 1744 through an important agreement known as the Al-Dir'iya Agreement between Muhamad Ibn Saud (founder of the First Saudi State) and Muhammad Ibn Abdul Wahhab, a leading Imam from the Sunni Muslim branch of Islam on the Arabian Peninsula. Some sources suggest that this led to close co-operation between politics and religion, which played a role in building the Saudi State. A closer look at the history of KSA will reveal that earlier kings and princes of the territory actually resisted any religious extremists who attempted to impose their beliefs or government on neighbouring nations. Nevertheless, despite this balance being struck between politics and religion, and efforts to establish the First Saudi State in the middle of the 18th century, it was conquered by the Ottomans. Later, between 1891 and 1918, the second Saudi State was almost brought to an end when Prince Ibn Al Rasheed, ruler of the north of the Arabian Peninsula, led an army to occupy and destroy Riyadh, which was the capital of Saudi Arabia at the time. In 1902, the third Saudi State underwent various phases and then in 1932, the modern Kingdom of Saudi Arabia (KSA) was established (Vassiliev, 2013). Today, it is one of the few countries in the world to be governed by an absolute monarchy, and following a Royal Decree in 1992, the Saudi King must comply with Shariah law and the Qur'an.

2.3.2 Geography and Demographics

KSA comprises 90% of the Arabian Peninsula (Bowen, 2014) (see Figure 2.1), making it the second largest country in the Arab world after Algeria. It shares borders with Iraq, Kuwait, Jordan, Qatar, the UAE, Oman, the Red Sea and the Gulf Sea (see Figure 2.1). Saudi Arabia has two capital cities: its religious capital of Mecca, and Riyadh, the seat of government.

In a survey conducted in 2017, the population of Saudi Arabia was estimated at 32,552,336, out of which 57.48% were male. Moreover, the native Saudi population has been estimated at 20,408,362, a mere 62.69% of the actual population (General Authority for Statistics, 2018).

Figure 2.1 Map of Saudi Arabia (General Commission for Surveys, 2019)

2.3.3 Culture

Segments of the population of Saudi Arabia live in cities and others live in rural areas. There are also nomads who live in migrations established by the government to encourage them to leave their desert way of life. Aside from this, there are Saudis from large tribes and Saudi immigrants of other origins. All this has produced great diversity and tolerance in Saudi culture, (Baki, 2004).

Islam is the official religion of Saudi Arabia. For millions of Muslims around the world, Saudi Arabia has great religious significance, because the Holy Cities of Mecca and Medina are

situated there. These are the main places of pilgrimage in Islam. Moreover, when Muslims all over the world pray, they face the direction of Mecca.

2.3.4 Saudi Women

Saudi Arabia is a deeply conservative country with strict rules for female behaviour, including dress code and communication with the opposite sex. Traditionally, marriage to someone outside one's tribe has not been permitted, but following changes made by Crown Prince Mohammed bin Salman since 2015, the Saudi courts no longer accept objections over women marrying outside their tribe; allowing Saudi women to marry anyone they want nowadays. Moreover, Saudi women were prohibited from driving for many years, but in 2018, Crown Prince Mohammed bin Salman reformed the laws governing this, as well as various other restrictions on women in Saudi society. These changes have had a profound impact on Saudi culture and freed women from out-dated customs. However, many women are still unable to exercise these rights, because of their family's control over them. This power wielded over women by their families is difficult to counteract through law, but needs cultural, social and educational change to raise awareness of the right of women to be treated as free and autonomous individuals.

In terms of professional life, women have traditionally only been permitted to become teachers, doctors or nurses, with no other domains open to them in the KSA. However, after successfully fighting for their rights and constantly demanding respect for their aims and ambitions, women in Saudi Arabia have finally been granted greater equality. Saudi Crown Prince Mohammed bin Salman recently pledged to empower Saudi women; this includes allowing women to take on leadership positions. For example, he has appointed Saudi Arabia's first female ambassador and has promised to appoint a female Minister in the near future.

In Saudi education, however, Saudi women still work separately from men (van Geel, 2016). Nevertheless, despite Saudi Arabia's conservative traditions, teaching is a profession that has been widely accessible to women in the Kingdom. The number of teachers affiliated with the Ministry of Education is estimated at 504,819, which includes 270,584 female teachers. However, the status of female teachers and head teachers in Saudi Arabia is still strongly influenced by traditional and cultural beliefs. This means that in spite of the number of women working for the Ministry of Education, their chances of gaining a leadership position is very limited. In fact, the percentage of women who occupy leadership positions under the Ministry

is just 1.5% and there are only four women who hold the position of deputy minister: for educational programmes, private university education, and scholarships for private education (Ministry of Education, 2019). This low proportion of women in leadership positions can be explained by legislation that prohibited women from becoming leaders, prior to policy reforms in 2017. As a consequence, all leadership positions were held by men until very recently. Additionally, women were previously not permitted to nominate themselves for any leadership positions in education, except for school management or the supervision of female teachers.

Thus, the upper echelons of the educational sphere are still male dominated, with the exception of just one woman, who is in charge of female recruitment in the government's Department of Education. This inevitably affects how well women's voices are heard by decision-makers, which means that they are merely followers, not participants, in Saudi Arabia's education sector. However, with its Vision 2030, the Saudi government has opened up the door to women teachers being nominated for leadership positions. This is evidenced by the measures adopted in the Kingdom, including two Royal Decrees, permitting women to drive, cancelling the need for male guardianship, and thereby permitting freedom of movement for women. These steps have increased women's chances of running for leadership positions and have persuaded senior leaders of their ability to be present and fulfill their responsibilities (Alsubaie & Jones, 2017).

Although a very limited number of women are in leadership positions at present, this could increase with time, as the opportunities facilitated by Vision 2030 gain traction among Saudi women.

2.3.5 The Economy

The oil boom of the 1970s completely transformed KSA in terms of its environment, lifestyle, education and economy. The revenue from oil was used for development; enabling a modern and affluent nation to be constructed, thereby lifting its population out of poverty. People who had once lived in tents or small traditional dwellings now lived in high-rise apartments and large houses. In addition, education expanded from a few schools, to thousands of schools and numerous universities. Finally, as a result of the oil boom, Saudi Arabia developed an oil industry that included oil-refining and other related activities such as plastics manufacture.

In more recent years, however, oil prices have declined, leading the Saudi government to make the important decision to steer away from reliance on oil as its only source of income. This

decision underpins a large-scale development plan known as Saudi Vision 2030, which sits on three axes: a vital society, a booming economy and an ambitious nation (Vision 2030, 2018). The first of these axes refers to building a vibrant society, whose members live according to Islamic principles, but with moderation. The second axis relates to a thriving economy, with the Saudi government dedicated to providing opportunities for all by building an education system that is linked to the needs of the labour market, while also developing opportunities for entrepreneurs, whether on a small or large scale. Finally, the third pillar of the Vision involves the public sector, where the Saudi government is seeking to enhance efficiency, transparency and accountability; thereby promoting a culture of performance to maximise the impact of human resources and efforts.

2.3.6 The Education System

The Saudi Ministry of Education was established in 1953 (Al-Sonbol, 2008) as an extension and development of the Directorate of Knowledge. Initially it was entrusted with planning and supervising the general education of boys across three educational stages (primary, intermediate and secondary), with King Fahd as the first Minister of Education. In 1960, the General Presidency for Girls' Education was established (with a budget of SAR 4,400,000). Women's education in State institutions in Saudi Arabia was launched in 1960, three years behind men's education, as male religious factions refused to allow women to study until the State promised that this would take place under religious supervision. The goal of the basic education of women in Saudi Arabia was originally based on the belief that the role of a woman was solely to be a wife and mother (Hamdan, 2005).

Nowadays, the Saudi Ministry of Education is responsible for providing education free of charge to all children (girls and boys) in Saudi Arabia, with education being compulsory between the ages of six and 18. The Ministry of Education is responsible for five levels of public education, including nurseries and kindergartens (age 3-5 years), primary schools (age 6-12 years), intermediate (middle) school (age 13-15 years), and secondary school (age 15-18 years).

Beyond compulsory education, many students continue their education to graduate and postgraduate level. Since 2003, this has included women. The religious authorities' oversight of women's education was established to reassure society that education for women would be bound by the religious controls that shape women's lives in a conservative manner.

In the case of schools, students are provided with textbooks, and the Ministry designs the national curriculum, provides budgets for all public-sector schools, and supervises over 34,000 schools, more than 650,000 teachers, and in excess of 50,000 administrative staff across 42 Districts (Ministry of Education, 2019). Furthermore, the Saudi government invests large sums of money in initiatives to improve education. For example, in 2010, this amounted to SAR 90,000,620,000, increasing to SAR 124,319,484,000 in 2016 (Ministry of Education, 2019). Such investment was made to try and increase education uptake nationwide.

Saudi Arabia's student population currently numbers 3,043,875 female and 2,937,844 male students (Ministry of Education, 2019). This level of investment in education has led to enormous changes in the Saudi education system over the years, with numerous programmes being implemented to improve it, especially in the area of evaluation, for which more than eight different programmes have been implemented (see section 2.3.6.2). The Ministry of Civil Service is responsible for recruiting teachers and assigning them the work that they are required to perform. According to the Ministry of Civil Service (2019), the teacher must satisfy his or her job requirements, and does not have the right to object to any task that is assigned, as long as the job description is included in his or her contract with the Ministry of Education.

2.3.6.1 The Ministry of Education's Features and Approach

The Saudi education system is governed by the Ministry of Education under guidelines issued by the Supreme Council for Educational Policy in 2016 Ministry of (Education, 2019). The Ministry is led by one Minister and a deputy, with three Ministerial assistants (see Figure 2.2), in a rather centralised approach (Al Essa, 2009; Abisaab, 2015). However, a new organisational structure was recently announced in early May 2019 (Ministry of Education, 2019).

The Saudi Ministry of Education is made up of 13 agencies, 61 public administrations, and 14 offices and secretariats, all of which are linked to the Minister of Education (see Figure 2.2). The magnitude of the effect of neoliberalism should also be noted in the presence of a performance agency, which oversees the SPIS to evaluate the performance of schools, academic offices, and the Agency for Communication and International Cooperation. This also represents an effect of globalisation; evident in the establishment of offices to supervise and evaluate performance, and to set standards.
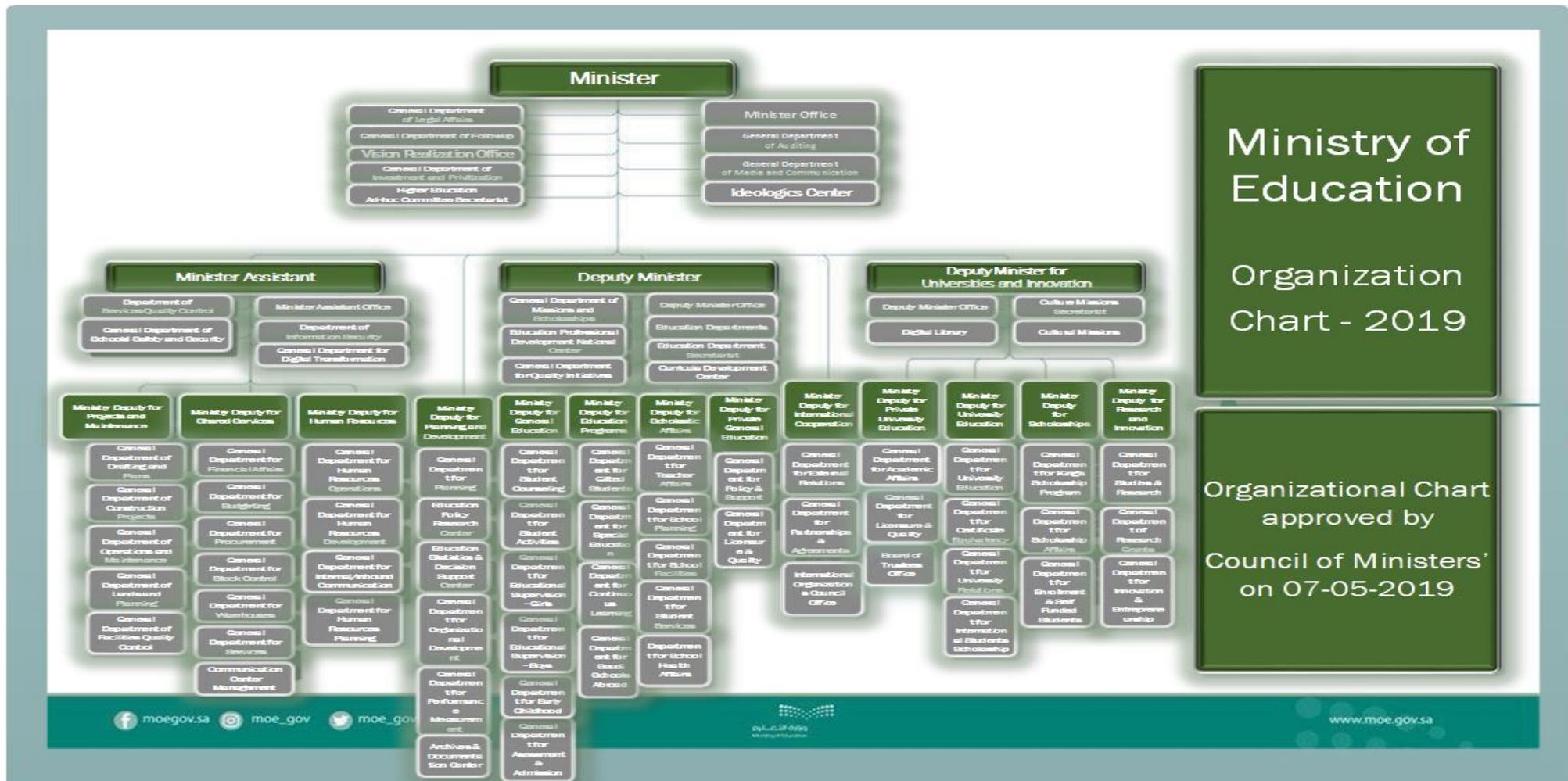
Figure 2.2 Organisational structure of the education system in Saudi Arabia (Ministry of Education, 2019)
https://www.moe.gov.sa/en/TheMinistry/AboutMinistry/Pages/MinistryOStructure.aspx

Under the current system of organisation, the Saudi Ministry of Education controls the national education system. This means that there are no opportunities for schools to manage themselves. Moreover, head teachers are not responsible for designing the curriculum, or for recruiting or dismissing teachers. Consequently, there is a great deal of distance between policymakers and schools, and teachers and head teachers. This can have a significant impact on performance (see subsection 3.4.4). In fact, there are managers in every education department (27 departments in total; see section 2.3), appointed to supervise communication between schools and the Ministry of Education (2019). As illustrated in Figure 2.3, these are situated in key locations in Saudi Arabia.

Figure 2.3 Map of Saudi education departments. These are indicated in red, while cities with offices are indicated in blue, and small towns are indicated in green (Jeddah, 2019)

In recent years, the Saudi Ministry of Education has begun giving education departments permission to make certain decisions such as controlling teachers' mobility between education offices (Al-Sonbol, 2008). However, this is unlikely to have an effect on education reform, as it is centralised (Al Essa, 2009).

2.3.6.2 Qualifications and Development of Saudi teachers

There are no gender-related differences in the Saudi system of teacher recruitment; both sexes undergo the same procedure. However, male teachers are not permitted to teach in girls' schools, whereas female teachers are allowed to teach boys until the third year of elementary education.

There are two conditions that a person must satisfy in order to be accepted as a candidate for a teaching position. First, he or she must hold a university qualification, either obtained from a college of education, or a public college that specialises in teacher training. In addition, an educational diploma is required, focusing on the teacher-training element. The second condition involves passing a test for the teaching licence, which involves a supervised written examination in two parts. The first part tests general knowledge and includes criteria relating to linguistic and mathematical ability. Meanwhile, the second part covers the basics of various teaching specialisations. The results of this test remain valid for a period of five years. Moreover, there is a requirement for teachers in government schools to be Saudi nationals, although this is not a condition imposed on private schools.

As far as teacher training is concerned, one of the most prominent problems encountered in the Saudi education system stems from poor teacher preparation. In consequence, continuing professional development has been identified as a priority by the Saudi Ministry of Education; most recently in a report dating from 2017 (Ministry of Education, 2019). There are two types of training received by Saudi teachers. The first is provided by colleges of education to bachelor's degree holders, often in the final academic year at a public or private university, or educational studies institute. Because this training takes place prior to entry into the teaching profession, it is often referred to as pre-service or practical education. The other type of training received by Saudi teachers is in-service training, and the providing authority is the Ministry of Education's General Administration of Training and Scholarships.

Additionally, the Saudi Ministry of Education sends teachers to Western universities to obtain higher degrees for example, in education, management, evaluation, and modern education technologies. It also provides opportunities for teachers to visit schools in America, Britain and Australia through a programme called Khebrat (Ministry of Education, 2019). In fact, a distinguished teacher can attain a supervisory position in the Ministry of Education and also serve as a school principal after only two years of working as a teacher, provided that good grades are obtained for teaching performance while working as a teacher. Still, even though male and female teachers have equal opportunities in recruitment and in terms of continuing professional development support from the Ministry of Education, female teachers are still far from equal in terms of them holding senior positions of power and decision making in Saudi Arabia's education system. One key explanation is that traditional attitudes towards the women prevail, which limit the roles women can perform. For instance, a woman is expected

to fulfill the role of a dutiful wife, who respect her husband's decisions, rather than making her own decisions. Women still hold a subjugated position in many families and communities. Nevertheless, this does not stop the Ministry of Education from encouraging female teachers to pursue leadership opportunities, albeit this is a relatively recent development (see 2.3.4).

2.3.6.3 School Performance Evaluation (SPE) in the Kingdom of Saudi Arabia (KSA)

The flow of information mentioned by Ritzer (2010) (see subsection 2.2.1), when defining globalisation, is evident in the Saudi education system. Wiseman et al. (2013) confirms that the enhancement of Saudi education has depended on neoliberal principles since the Higher Committee of Education Policy of 1997; introducing many weaknesses into the Saudi education system (Ministry of Economy and Planning, 2015). The most notable of these is the inadequate performance of teachers and head teachers. Thus, it is recommended that the government build a programme of evaluation to facilitate the identification of strengths and weaknesses in education. This recommendation has been taken seriously by the Saudi government (Alrwqee, 2012), and in recent years, the Saudi Ministry of Education has invited many experts, such as Andreas Schleicher (head of the OECD) and Ranjit Malhi, spokesperson at the First International Conference of Total Quality Management (TQM) in Public Education in 2011. As a result, the government has established numerous programmes to improve educational standards. According to Al Hakamy (2008), these consist of:

1- National assessment tests;
2- A project to ensure quality in academic achievement;
3- Basic skills tests for educators;
4- Standardised achievement tests;
5- Comprehensive evaluation of schools;
6- An educational accreditation system for private schools;
7- Comprehensive evaluation, and
8- An award for educational excellence.

Educational supervision is provided by more than 8000 technical supervisors who oversee administrative matters in schools (Ministry of Education, 2019). Their tasks vary and include evaluating school performance, the provision of teacher training, and co-ordination between schools and the Ministry (Alzahrani & Alghamdi, 2016). These supervisors use teacher

performance assessment, in which specific indicators are applied (Ministry of Education, 2019).

The Saudi Ministry of Education has sought to build a profile of education supervision and school performance: designing and applying school and supervisory performance indicators (see Figure 2.4), following approaches adopted by successful institutions around the world. First, the Saudi Minister of Education's Resolution No. 37350005 was issued and released on 12th May 2014. It is referred to as the 'School Performance Indicators System' (SPIS) (SPIS, 2017). Correspondingly, the organisational guide for evaluating the performance of SPIS itself includes the following elements: leadership indicators of teaching performance, the performance of supervisors and education offices, and the performance of head teachers in terms of educational supervision and school performance. These indicators are assessed by the body that oversees school supervisory performance in Saudi Arabia. The support, control and supervision of educational accountability in an organisation poses questions about its commitment to explaining specific tasks and learning outcomes. This depends on smart professionalism and questioning, especially regarding tasks, programmes and output.

SPIS, which involves the use of performance indicators, began evaluating schools along with educational supervision, where Education Supervision was not cancelled but continued to perform its duties. There are some differences and similarities between the two programmes, which are illustrated in the following Figure: indicating that although SPIS and Education Supervision differ in their mission and occupy different offices, the teams undertaking this evaluation all work towards education supervision.
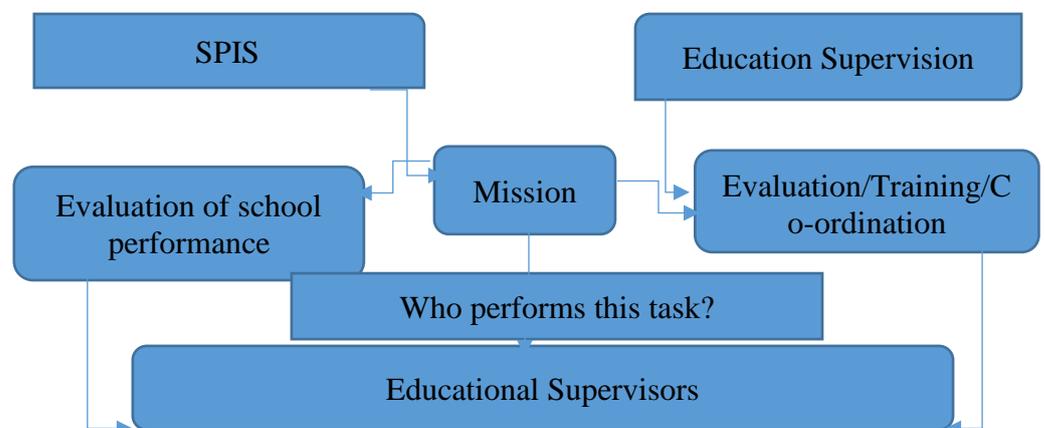
Figure 2.4 Differences and similarities between SPIS and Education Supervision

SPIS will now take on greater importance, since the Cabinet has issued a decision to link teachers' salaries with performance.

2.3.7 Conclusion

This chapter consists of an attempt to contextualise the study within the global sphere by highlighting globalisation as a key influence on the neoliberal approach (such as NPM) to education reform around the world. In particular, it has provided some background on Saudi Arabia as a country and an overview of its population and education system. Also outlined in this chapter is the prevailing influence of neoliberalism on school evaluation.

Chapter 3: Literature Review

3.1 Introduction

Education reform has become a cornerstone in the contest for economic growth and global competition, with many countries around the world reforming their educational policies to achieve their economic goals (Evans, 2011). The most important efforts in these reforms have focused on improving school performance, in order to achieve quantifiable outcomes and produce academically successful and skilled students, who can play a role in building a strong and competitive economy (Walker et al., 2011). Therefore, the first stage in improving school performance is to evaluate that performance, determine points of weakness and strength, and provide the system with appropriate feedback. This is so that improvements can be made and deficiencies addressed, which would include suggesting methods of helping schools to improve their outcomes (Evans, 2011). One evaluation technique involves measuring school performance through the use of indicators. This approach is widely employed, especially among OECD countries like the UK (Whalley, 2011) and Germany (Dedering & Müller, 2011).

Within the field of education research, however, the use of indicators to evaluate school performance across the world has been a topic of much debate, leading to considerable doubts about its real effects (Evans, 2011). Scholars whose work has been of particular influence, gaining significant global attention from education experts, include Perryman (2006) and Ball (2012a), who believe that evaluating the performance of a school on the basis of indicators leads to some aspects of classroom teaching practice being ignored, if these cannot be measured using indicators. An example of this would be the efforts made by teachers to make their teaching successful and effective. This distances the teacher from the actual practice of teaching as a basic task. In addition, certain tasks that are commissioned from teachers, such as documentation, lead to extra work. Although there is abundant evidence of this in the UK and many other OECD countries, there is a dearth of research on these systems, their effects, their successes, and any overlapping characteristics that they might have in Saudi Arabia. The goal of this thesis is consequently to explore the perceptions and experiences of head teachers and teachers with regard to the latest SPE system, known as SPIS.

In light of the above, this chapter is divided into six sections, which introduce, outline and summarise the study's conceptual framework and the empirical literature on which it is based. This is followed by a description of the emergence of SPE procedures, before presenting the

basic concept of school performance management. Secondly, the concept of performativity is examined. Thirdly, the concept of accountability is looked at in detail, including its definition, a model of accountability, and some of the corresponding empirical literature. The chapter is then summarised in a concluding section.

3.2 School Performance Evaluation (SPE)

In the past, schools used traditional means of evaluation, such as formative and summative assessment (Dangerfield, 2012), but since the rise of neoliberalism in the 1980s, many developed countries have looked at more effective ways of enhancing school performance (Evans, 2011). For instance, over the past three decades, the UK has mainly used two different approaches to evaluation. The first of these is performance appraisal, which was originally implemented during the 1980s, becoming obligatory for all teachers in 1991 (McDonald et al., 2019). Its main goal is to provide information that will enable managers to improve employees' performance in an optimal manner (DeNisi & Pritchard, 2006). Leithwood, Jantzi, Silins and Dart (1991) define performance appraisal as a 'standard operating procedure' (SOP) in all mature organisations. Manasa and Reddy (2009) provide a further explanation, stating that "a performance appraisal is a systematic and periodic process that assesses an individual employee's job performance and productivity in relation to certain pre-established criteria and organisational objectives" (p.1). This includes performance indicators, because appraisal tools depend on specific criteria, as discussed later in this section. In contrast, since 1991, performance appraisal has been framed in terms of accountability and professional development (Evans, 2011), which is now widely accepted in the UK (Johnson & Regan, 2014).

In this vein, SPE is seen to be related to performance management (see subsection 2.2.4). However, there are many critics of performance appraisal, for instance, Ball (2017) argues that it tends to sanitise and exclude real values, because it is dependent on students' results (Gaertner, Wurster, & Pant, 2014) as a true reflection of teachers' efforts (Dedering & Müller, 2011). Researchers like Ball, Maguire, Braun, Perryman and Hoskins (2012) have looked at the effect of raising the standard of GCSE achievement, which means increasing pressure on teachers to ensure that their students obtain good results, thereby adding to their teaching stress and workload. Moreover, Gaertner et al. (2014) claim that linking students' results with teachers' performance has had a negative effect on students' performance. Indeed, the impact of stress and workload at the core of the teaching profession cannot be overlooked. It is also

illogical to consider a students' exam results to be an honest measure of the teacher's efforts, because the student may simply have failed to invest enough study hours of their own. However, Plowright (2007) highlights the positive influence of evaluation based on students' achievement on school performance.

Aside from the above, there are those who do not believe that it is evaluation per se that drives teachers to improve, but rather accountability and rewards (Gustafsson et al., 2015). However, ensuring the quality of school performance is not limited purely to evaluating school performance; it also requires accountability (Penninckx, Vanhoof, De Maeyer, & Van Petegem, 2014). That said, school evaluation has other implications for school improvement and teachers' stress, workload and morale, which will be examined in more detail in subsection 3.3.3.

3.2.1 Types of School Performance Evaluation (SPE)

3.2.1.1 Internal Self-evaluation

According to Nevo (2002), internal self-evaluation "can be performed by a teacher or a group of teachers, by other members of the school's professional personnel, by the principal or other school administrators, or by a special staff member designated by the school to serve as a 'school evaluator'" (p.10). In contrast, external evaluation (inspection) must be implemented by an independent organisation (Nevo, 2002). According Chapman and Sammons (2013), internal self-evaluation serves multiple purposes, including preparation for inspection; the driving of collaborative internal school improvement efforts; motivation for teachers to achieve high standards, while at the same time assisting them in identifying needs; goals analysis; the selection of instructional strategies, and the planning and monitoring of work. It also serves to improve decision-making processes, rendering them more effective.

3.2.1.2 External Evaluation

In order to apply standards with confidence, most countries have evolved governing bodies to ensure conformity with national standards. There are two examples of such organisations: The Office for Standards in Education (Ofsted) in the UK (Nevo, 2002) and the Education and Training Evaluation Commission in Saudi Arabia (ETEC, 2019). These inspections are mainly concerned with whether money is being spent effectively. However, according to Elliott (2012),

the inspectors who conduct them have been criticised by politicians for several reasons: some have been found to be unreliable and untrained, while others have failed to spend sufficient time on their missions.

As a consequence, the definition of inspection has changed. Whereas it was once considered to be a process of ensuring that expenditure takes place in accordance with various standards (Nevo, 2002), Clarke and Ozga (2011) describe it as just "one of a cluster of processes that have accompanied changes towards 'governing at a distance' (others include audit; targets; PM and standards), as a way of managing or improving the performance of public services" (p.2). This definition considers inspection to be an advanced version of the external evaluation that emerged in the UK, once standards were established in the mid-1970s (Townsend, Porter, & Mawdsley, 2002). Many other countries, such as New Zealand, Singapore, the Netherlands (Whitby, 2010), and Saudi Arabia (ETEC, 2019) have since followed suit in this endeavour.

3.3 Performativity

As a result of implementing performance management in education, most education systems evaluate the performance of schools and their teachers through student achievement, using performance indicators (see section 2.2). SPE systems thereby use indicators to ensure that schools achieve their goals. However, researchers such as Perryman et al. (2011) and Ball (2012) have described performativity as a tool of the performance evaluation era. It is argued that it has led to a culture of performativity in education (Ball, 2012a), with a significant influence on education worldwide (Wilkins, Busher, Kakos, Mohamed, & Smith, 2012; see also subsection 2.2.4). However, this concept of performativity needs to be clarified in more detail, and so the formulation of a clear definition takes priority in this case.

3.3.1 Definition of Performativity

According to Lockheed and Hanushek (1988) and Perryman (2006), the concept of performativity was first put forward by Lyotard (1984), who suggested that employees should be judged according to their performance; in other words, by what they have accomplished. Moreover, Avis (2005) cites performativity as a process linked to accountability and the objectives of the institution, which aspires to achieve through the performance of its employees in a culture of blaming teachers. In fact, the latter leads to the use of numbers, technology and the documentation of information about employee performance. Therefore, Ball (2012a)

describes performativity as "a technology, a culture and a mode of regulation that employs judgements, comparisons and displays as means of incentive, control, attrition and change based on rewards and sanctions" (p.216). This definition confirms that judgement in performance management should take place in accordance with input or output (Perryman, 2006). Meanwhile, Wilkins et al. (2012) state that what is important about performativity is its role in measuring the performance of teachers and schools, using quantitative data. Thus, performativity deploys performance indictors to evaluate the performance of teachers and schools. It is generally acknowledged that performativity bears a strong relationship to many aspects of education evaluation, such as indicators and management systems, and teachers' identity and professionalism (Ball, 2016). Hence, this view of performativity will be adopted in the present thesis.

3.3.2 Features of Performativity

Many features that characterise performativity are discussed in this thesis, the most important of which is its strong relationship to politics and governments and their goals for education. This is where governments impose certain indicators that education is expected to meet, primarily for economic gains. Through these indicators, teachers are also held accountable for the extent to which their students achieve high scores in national tests (see subsection 2.2.4). In addition, performativity encourages competition between schools in a country (see subsection 2.2.4.3), as well as competition between countries in education, like what happens through exam results in the PISA and TIMSS (see section 1.2). It is based on numbers, data collection and documentation (see subsection 2.2.4.3). This leads to teachers having to accomplish additional work, other than their teaching work, which has encouraged intense debate about the magnitude of the impact of performance on teachers. The application of performativity is explored in the following subsection.

3.3.3 Performativity and Teachers

Performativity has changed teachers' job requirements. Teaching per se is no longer the only requirement of a teacher, but other factors come into play when judging teachers' performance in schools, such as the results achieved by their students, the school's provision of social services, the wider environment, etc. (see subsection 2.2.4.2). Some experts have stated that this presents teachers with numerous obstacles. Ball (2003) claims that the emphasis on

performance is intended to intimidate the teacher, causing him or her to function in an atmosphere of fear about the job and caution over exercising professional judgment. Perryman et al. (2011) also state that this evaluation has influenced teachers and contributed to an increase in the pressure placed upon them. The impact of performativity on teachers has elicited intense debate, mainly negative, in countries such as the UK and USA.

Illustrative of the level of concern over the pressures experienced by teachers, articles in the British media and on British websites have addressed this issue with great interest, under headlines such as 'No game changer: Ofsted framework proposals won't reduce stress, say unions' (Whittaker, 2019) and 'Ofsted under fire in its own survey of teachers' wellbeing' (*The Guardian*, 2019), thereby highlighting its importance. Furthermore, Galton and MacBeath (2008) emphasise the negative influence of performativity on teachers' morale, gathering the experiences of teachers in primary schools within five years of implementing a number of educational change initiatives. From their analysis of these data, five categories of pressure were identified: the frustration faced by those who try to make inclusive education work in practice; the impact of constantly changing policies for the staff required to implement them; loss of status within the teaching profession; reasons cited for teachers choosing to leave the profession, and the consequences of continuing to work in education and fighting for what one believes in. The above authors concluded that teachers were under intense pressure, compared to their previous working conditions, before performance appraisal initiatives were in place. Consequently, this had a huge impact on their confidence and ability to perform within their own profession. The impact of paperwork was especially noticeable on their personal lives and leisure pursuits, as they attempted to achieve good results. Similarly, this influenced their satisfaction with their performance, and in turn, their professional practice. Nevertheless, in Saudi Arabia the effects of SPE have not received any attention within the professional domain or from the media, but this study will attempt to shed light on the topic by exploring stress, workload and morale amongst teachers and head teachers.

3.3.4 Stress and Workload

According to the UK's National Health Service (NHS, 2019), stress is "the feeling of being under too much mental or emotional pressure" and this definition is adopted in the present study. Ramos and Unda (2016) found stress to be strongly linked with teaching as an occupation, perhaps because of the demands imposed on teachers to manage a workload that

they cannot accomplish (Mulholland, McKinlay, & Sproule, 2013). In the UK, 98% of participants in a survey of 4,500 teachers stated that they were stressed (Lightfoot, 2014). It is possible that this was due to Ofsted, and the fear of being designated as a school that requires special measures (Sugrue & Mertkan, 2017). One illustration of the lack of respect for teachers is that some of the procedures in evaluation programmes interrupt teaching schedules, with the result that teachers are obliged to engage in other activities, instead of focusing on their actual job (Alrwqee, 2012; Alkarni, 2015). Jaradin (2004) and Alkarni (2015) also claim that the administrative procedures in schools can lead to teacher burnout. In addition, the rise in stress levels experienced by teachers is due to their accountability for students achieving the desired results, while Al-Omari and Wuzynani (2013) refer to the link between accountability and increased stress among head teachers.

Stress of this nature can affect teachers' performance and increase the frequency of absenteeism (Li & Sullivan, 2016), given that teachers suffering from stress can also experience "sustained physical and mental health problems" as a result (Naghieh, Montgomery, Bonell, Thompson, & Aber, 2015). Conversely, Mousavi (2007) claims that stress can provide positive momentum for employees, providing that they have the time and place to refresh their minds (Siltaloppi, Kinnunen, & Feldt, 2009).

According to Flores and Derrington (2017), SPE leads to increased workload in schools, because it means implementing new teaching strategies and indicators. Mulholland et al. (2013) assert that excessive workload can be stressful for teachers., and so it was a matter that needed to be discussed with them and with head teachers in this study to investigate their views of workload in relation to SPIS evaluation. Finally, morale was another matter raised with the participants, as defined in the following paragraphs.

3.3.5 Morale

Brion (2015) considers morale to be the feelings that a teacher has about his or her job or role within an organisation, and the extent to which this satisfies his or her goals. Morale is specified here as the self-confidence and satisfaction experienced by teachers and head teachers with regard to SPIS, as well as their sense of being effective within their schools. Thus, Bousquet (2012) refers to workload and stress as the reasons behind reduced morale among teachers, while Reid (2010) claims that a lack of respect for teachers has led to a decline in their morale. However, there is no contradiction in this, because exposing a teacher to exhaustion and an

unmanageable workload, despite all efforts, indicates a lack of respect from the management for the essential role of a teacher. In fact, Goldstein (2015) asserts that low morale can lead to teachers leaving their jobs. Therefore, education systems need to be aware of teachers' morale on an ongoing basis, so that high morale can be maintained (Govindarajan, 2012). This will not only serve to keep teachers in the profession, but will lead to many other positive outcomes, such as co-operation, the sharing of ideas, and peer support (Littleford, 2007).

One factor that can affect teachers' morale consists of students' results (Caprara, Barbaranelli, Steca, & Malone, 2006). This is because SPE uses these as an important item for comparing schools and judging the quality of their performance (Perryman et al., 2011). However, in order for schools to achieve good results, they need to adopt strategies that will avoid placing additional burdens on teachers, because stress has a negative effect on teachers' morale (Govindarajan, 2012). Consequently, this study addresses the topic of teachers' morale and related issues encountered by the participants. The effect on teachers' morale will also depend on their commitment to implementing education reform programmes.

Aside from the above, according to Torabi and Sotoudeh (2010), teachers tend to be more committed to implementing a new system or programme, if they feel that they are considered to be an important part of the school. Evidence of this esteem for teachers may include them receiving feedback on their programmes or an invitation to participate in decision-making. Moreover, Singh (2007), Aydin, Sarier and Uysal (2013), and Saljooghi and Salehi (2016) found that teachers' participation in school activities, like planning, organisation, and the application of new programmes, can improve teaching quality and teachers' commitment. In the next section, the third concept will be presented, namely accountability.

3.4 Accountability

The second core concept of this study is accountability, because SPE requires a foundation of systematic accountability (Anderson, 2005). The information collected in evaluating school performance is then used to judge teachers, head teachers and school performance (O'Neill, 2013). In fact, accountability has not only become a tool of the system, but it actually institutes the system itself (Møller, 2009). In this section, its definition and various forms are consequently examined, as well as its relationship with school improvement, before reviewing the empirical research on accountability.

3.4.1 Definition of Accountability

Researchers have recognised different approaches to conceptualising accountability. According to Pollock and Winton (2016), the scholarship on this topic defines it from many different perspectives, which are not without their similarities and differences. That said, Barzanò's (2009) definition, worded as follows, has been accepted by many researchers, such as Watts (2012):

> Accountability in the sense of a set of formal and informal mechanisms making schools answerable to different constituencies interested in educational results, represents one of the major challenges schools—and headteachers in particular—are dealing with. (p.190)

Barzanò considers accountability in education to be made up of a number of processes, which enable schools to answer questions from different stakeholders, whether formally or informally. This means that accountability does not merely relate to a higher authority; empowered to question teachers and head teachers about their school's performance and results, which is a major challenge. Instead, it drives school staff to make their best efforts to achieve the level of quality that all stakeholders require. Nevertheless, Barzanò tends to overlook individual aspects of school performance, possibly because he looks at it as a whole, with all outcomes worthy of consideration. In contrast, McCallum (2018) believes accountability to be a system of making judgements to categorise schools, teachers, and head teachers, based on school output from all school activity, wherein results are evaluated by Ofsted. However, students' results are widely used as indicators by many experts when defining accountability (Richardson, 2015), which will be addressed in specific detail below.

3.4.2 Model of Accountability

A number of researchers, such as Poole (2011) and Yia and Kimb (2019), have classified accountability into two broad types, external and internal. Poole (2011) refers to internal accountability as 'school accountability', describing it as "a process by which agents exert pressure to ensure that schools meet their goals" (p.3). In this operation, all school staff work together to determine school standards and collect all the necessary information for improvement, while also using peer pressure as a means of realising their ambitions. In addition, Poole (2011) identifies external accountability as the external authority that sets the standards and objectives to be met in schools. In the present case, this authority is the government, which can cut its support if a school fails to achieve its goals. In contrast, Kim, Harris and Pham (2018)

claim that external accountability is determined by the OECD's PISA assessments, which force countries and governments to evaluate school performance for economic reasons, using global standards. However, this is not a huge difference, but rather a difference in the number of representatives deployed to implement external accountability, as opposed to the actions or objectives of these representatives.

In addition, Møller (2009) outlines various forms of accountability, including "political, public, managerial, professional and personal". This means that schools are politically accountable in their output to all segments of society and this perspective should include responsibility for goals, standards and needs in education in general, and schools and school leadership in particular. The above author refers to teachers' personal feelings about their duties and students. However, this indicates that Møller (2009) does not define stakeholders according to their position within or outside schools. Moreover, Kwok (2011) defines accountability as follows: "[A]accountability relates to external clients, stakeholders and supporters of the school" (p.16). To conclude, these researchers consider accountability models according to the actors whom they perceive to be responsible for education, with the authority to dictate to schools what they should achieve.

3.4.3 The Influence of Accountability

The wave of neoliberalism and performance management has encouraged many countries around the world to employ accountability in attempts to reform their education systems and improve school performance (Poole, 2011; Bessant et al., 2015; Yia & Kimb, 2019), thereby ensuring consistent progress in education (Larsen, 2009). However, multiple researchers have tried to assess whether this adaptation has had any effect on the education systems concerned or the improvement of school performance (Møller, 2009). According to Bessant et al. (2015), the introduction of accountability leads head teachers to apply standards, so that the requirements of the accountability system can be met, resulting in enhanced school performance. Thiel and Bellmann (2017) corroborate this view, finding that accountability can help schools, in that they are provided with feedback, which can be acted upon to address weaknesses and enhance the quality of their performance. However, Kwok (2011) argues that schools need to ensure compliance with accountability, in order for any benefit to be derived from it:

As the leader of the school, the principal is held responsible for the performance of the school and for the improvement of educational quality. In development planning, long term participation in implementing accountability can also enable the school and stakeholders to close the performance gap between the planned and achieved targets.

In contrast, Møller (2009) highlights the potential of accountability to reform school performance and education in general, pointing to the many different actions involved in classroom accountability – which are too numerous to measure – as well as the role of accountability in evaluating students' results, with teachers held accountable for them. In addition, the above author refers to poor school informants, who fail to support teachers; posing the question of how teachers can be held responsible for that. He states that "effective accountability requires teachers with high skills and adequate knowledge of the accountability system". Hence, he considers that all requirements should be met in accountability systems, if they are to be successful. However, this does not provide strong grounds for rejecting accountability, as it is still a factor that can motivate schools to meet standards.

3.4.4 Accountability and Decentralisation

The most important impact of neoliberalism (see subsection 2.2.2) is decentralisation, which calls for reducing the government's dominance over organisations (Lynch et al., 2012), including those that fall within the education sector. It has led to the implementation of neoliberalism in various types of school, such as independent schools, with head teachers being granted more authority over the running of these institutions. However, one World Bank (2003) report asserts that decentralisation should be subject to accountability, so that school improvement can progress. This interest in decentralisation has emerged as a result of various opinions that assert its beneficial role, particularly in education. These benefits range from the monitoring of finances, reduced corruption, and the satisfaction of local aims. Besides, knowledge transfer facilitates the achievement of local objectives in school communities (Gertler, Patrinos, & Codina, 2007). Nevertheless, according to Carr-Hill, Rolleston, Schendel, and Waddington (2018), the benefits of decentralisation require a rich environment and educated stakeholders, because these are the parties who make decisions. They will consequently need to possess adequate knowledge and abilities, so that they can perform this task effectively in the education context. In contrast, there are some authors who state that decentralisation can lead to favouritism, especially in employment, as well as the use of power for personal gain, drawing upon favours from personal connections in the community (Galiani, Gertler, & Schargrodsky, 2008).

Notwithstanding the above, decentralisation has been found to be successful around the world, such as in the UK and USA (Carr-Hill et al., 2018). Although the Saudi education system has been influenced by neoliberalism (see Chapter 2), it is still highly centralised (Al-Sonbol, 2008) and this has had a negative effect on the system (Al Essa, 2009). In short, the Saudi Ministry of Education adopts a top-down approach (Alzaidi, 2008), as discussed by numerous researchers, particularly in terms of its efficiency and effectiveness in delivering the required evaluation programmes to school teaching staff and head teachers. For instance, Camburn and Han (2015) investigated the impact of implementing a top-down approach, referring to decision-makers changing direction, while teachers encounter difficulties and are left wondering how to apply the previous decisions issued from the top. It clearly indicates the reality faced by teachers when the system marginalises them at the planning stage of change, despite the fact that teachers are crucial to the success of education reform (Camburn & Han, 2015).

However, it is not just the marginalisation and neglect of teachers and their views that result from a top-down approach to planning education reform; it can also create an environment that is hostile to and opposes the regime (Ravitch, 2016). Such negative effects of this approach are evident in Saudi education, whereby Mansour, Heba, Alshamrani and Aldahmash (2014) point out that schools in Saudi Arabia fail to reflect the expectations of planners. Therefore, according to Ravitch (2016), successful reform should work upward from the bottom, although the above author seems to be unaware that this is already being implemented. Education reform has consequently become an arena of experimentation, giving rise to even more errors (Priestley, Miller, Barrett, & Wallace, 2011). In fact, both top-down and bottom-up approaches can have a negative impact on education reform, according to Fullan (2007), as the top-down approach is a threat to teachers' commitment to change, and the bottom-top approach is far too centred on individual experience. Consequently, it is rarely successful or sustainable in its outcomes. However, according to Brezicha, Bergmark and Mitra (2015), co-ordinating the two approaches can be effective for reform. The next section will therefore examine empirical research studies based on three conceptual frameworks.

3.5 Theories of Performance Management
This study is guided by two main theoretical models, both relating to performance management theory: goal setting and expectancy.

1- Goal setting

Goal-setting theory focuses on the relationship between improving organisational or individual performance and the setting of goals (Latham & Locke, 2007). This theory was established by American psychologist, Edwin Locke in 1968 and contributed to by Professor of Organisational Effectiveness, Gary Latham in 1970 (Miner, 2015). It claims that goals can affect performance in many ways; for example, through the existence and clarity of the goal and the procedures for accessing it, as well as the difficulty involved in achieving the goal and its value (Latham & Locke, 2007). This theory stresses the importance of having specific goals to accomplish, which must be very clear to the teacher (Moeller, Theiler, & Wu, 2012). Teachers' ignorance of the goal and its application in a specific system may lead them to view the goal in different ways, which might not be productive or fit in with the school's objectives. Therefore, it is essential that the goal is clear to the teacher. However, according to this theory, goal setting leads to commitment (Lunenburg, 2011). This is because, in the absence of goals, there will be no commitment from the teacher. Conversely, when a goal is clear, a consensual desire to achieve it may be found amongst teachers.

The most important component of goal setting involves teachers setting their own goals, whereby they choose to pursue and achieve them based on an internal stimulus, stemming from their sense of the goals' importance (Lunenburg, 2011). This indicates the necessity for teachers to participate in decision-making, and their ability to set programme objectives, especially with regard to education reform such as SPIS. It also demonstrates that due to teachers being excluded from participating in decision-making and their lack of voice in choosing educational goals, reform programmes in centralised systems tend to fail. Locke (1996) claims that when individuals set their own goals, they are likely to want to invest greater effort in achieving them, rather than being set goals that are too easy by someone else. However, goals should still be attainable (Lunenburg, 2011). Moreover, high or seemingly elusive goals will encourage the teacher to make more effort to achieve them and become a challenge, unlike easy goals that can be achieved by anyone without much effort. According to Lunenburg (2011), easy goals have little impact on the reality of a school or other institution in terms of developing performance, which means that setting such goals may be a waste of time and effort.

However, in this current study, an attempt is made to determine the influence of SPIS when implemented in schools, if indicators are applied to ensure that the goals and plans of decision-makers have been achieved (see subsection 3.3.3). SPIS indicators are targeted towards

improving school performance. What makes this theory important in the current study is that it reveals the reasons for school performance improving or deteriorating, and it can also help ascertain the efficiency or effectiveness of the indicators. Most importantly, it promotes performance amongst teachers and can explain the complexities that they encounter when endeavouring to understand the indicators.

2- Expectancy Theory

Expectancy theory is the result of work by US Business Management Professor, Dr. Victor Vroom during the 1960s, when he found a gap between industrial psychology research and employee motivation in the workplace (Lloyd & Mertens, 2018). Vroom identified a relationship between employees' expectations, and their effort, performance and rewards (Mullins, 2007). He observed that an employee's expectation of a certain amount of effort being required to achieve a set goal, caused him or her to exert that effort. Moreover, the expectation of reward also appeared to affect the level of effort made. Therefore, an individual may believe that a certain degree of effort will lead to a specific level of performance, against which a reward will be earned (Lloyd & Mertens, 2018).

Additionally, expectancy theory claims that an employee's confidence that he or she will receive a fair reward for effort will be a motivation to do the job (Isaac, Zerbe, & Pitt, 2001). The theory also states that the reward must be equal to the performance (Parijat & Bagga, 2014), in the sense that the employee needs to believe that he or she has received a fair reward for the effort and performance delivered. The question that arises here concerns the relationship between this theory and the present study, which investigates the influence of evaluation on school performance. According to the employees' knowledge, the results of their evaluation will affect their expectations of their own efforts (Lunenburg, 2011). As a result, the evaluation of school performance gives teachers a view of their own performance and how much or what they have achieved. Therefore, it is possible to imagine the negative impact on teachers' performance if this evaluation is not performed correctly. For example, teachers can suffer, and their confidence may be negatively affected.

All questions answered in this study relate to the evaluation of school performance through SPIS, based on the performance of head teachers and teachers. Therefore, according to this theory, the evaluation process should be carried out correctly (Pulakos, 2004). This study also

investigates the influence of performance evaluation and its role in increasing pressure on head teachers and teachers by demanding more effort on their part.

3.6 Empirical Research Studies on Three Conceptual Frameworks

The various systems of performance management and school evaluation found worldwide have become an increasingly popular focus of education research (Dunn & Miller, 2007; Pollitt & Dan, 2011; Kalimullah et al., 2012; Jaksic & Jaksic, 2013).

As clarified earlier, the use of indicators and standards for evaluating school performance and making judgements about teachers is referred to as 'performativity', and this is frequently and extensively debated by scholars. However, doubts have been raised about the real effects of performance management criteria (Evans, 2011). For example, some critics argue that performance management has had a detrimental effect on teachers' professionalism and professional identity, resulting in unhealthy and unnecessary levels of stress and workload. The negative influence of performativity has therefore attracted enormous attention from education experts around the world, including Perryman (2006) and Ball (2012a). In contrast, scholars like Whalley (2011) argue that performance management can result in greater accountability among teachers and head teachers. Hence, there are some researchers who believe neoliberalism, NPM/managerialism and performativity to be positive in their impact on education in general and school performance in particular (Ehren & Visscher, 2008; Dedering & Müller, 2011). It should also be considered that while there is now an established body of research into performance management and the effects of SPE in the UK and many other parts of the world, research in this area is still in its infancy in Saudi Arabia. This is despite the fact that Saudi Arabia has already witnessed two decades of educational reform, resulting in fundamental changes in its education system and the strategies adopted to improve it, including SPE systems. However, what will be scrutinised here are peer-reviewed empirical studies on the three concepts, beginning with SPE.

3.6.1 Empirical Research Studies on School Performance Evaluation (SPE)

In the following paragraphs, peer-reviewed empirical studies on SPE are examined; in particular, studies on programmes that resemble SPIS, such as the UK's Ofsted. According to Scheerens (2014), SPE is an important aspect of research on school effectiveness. However, much of the research to date has addressed aspects and areas of school performance that can be

categorised into three main types, whereupon similar programmes to the SPIS are examined, including Ofsted and Germany's KMK (Watts, 2012). In this subsection, the primary measures implemented by Ofsted are described. Here, a course of action to be undertaken by a school is identified to meet the necessary standards, according to the indicators in place and drawing upon observed changes implemented by Ofsted since 2005. In this study, the context of the investigation is specifically secondary education, including the perspectives of head teachers.

Watts (2012) explains that in 2005, Ofsted facilitated inspection procedures and made a number of changes; for example, with regard to the notice issued to schools in advance of inspections, which was initially reduced from several weeks to just a few days. However, this notice period has since been further reduced to half a day, whereby schools are now informed of a pending inspection at midday, the previous day (Roberts & Abreu, 2016). However, the outcome of this is an overwhelming concern with school rankings in the league tables, which has caused many stakeholders to criticise schools and the education system as a whole (Baroutsis, 2016).

Conversely, some researchers have looked at the apparatus that is used to evaluate school performance, including performance indicators, as illustrated and discussed by Rowe (2004), (Evans, 2011) and Dangerfield (2012), amongst others (see subsection 2.2.4). In addition, Leckie and Goldstein (2009) address the use of student achievement in national exams as an indicator to compare performance between institutions, stressing that it is not a fair means of differentiating between schools; a criticism supported by Ehren and Honingh (2011). There is no doubt that relying on just one index to detect differences in school performance cannot give a true picture of school or teaching performance. In fact, it can lead to erroneous judgments being made about certain schools.

Studies on the topic of school performance league tables include Leckie and Goldstein (2009) and Perryman et al. (2011), amongst others (see subsection 2.2.4.3). Besides, some similar studies of interest to this research have been conducted in the UK and more recently, in the Netherlands (Ehren & Visscher, 2008). Here, a number of studies are reviewed and classified geographically. Thus, four studies conducted in Western countries are presented, with a special focus on Germany (Dedering & Müller, 2011), and the UK (Janssens & van Amelsvoort, 2008; McVeigh, 2016; Perryman et al., 2018), and one study on SPE in various contexts, including the Netherlands, Belgium and Germany. Meanwhile, to a lesser extent, studies from Saudi Arabia and other Gulf countries are examined.

The first study (Dedering & Müller, 2011) was conducted in a Western context, and it "focuses on the question of to what extent the purpose of school inspections as laid down in the programme, namely, to promote quality improvement in schools, has been accomplished". The above study describes Germany's experience of SPE, which differs from the SPIS addressed in this study, but is similar in some respects; for example, it focuses on the performance of teachers and schools, and the use of indicators to judge their performance. Hence, it relates to this current study, with a view to providing some detail on Saudi Arabia's SPIS evaluation of school performance. In addition, both studies explore head teachers' perceptions of how well school inspections function and influence practice. The above study sampled teachers in an enquiry into the extent to which inspections promote quality and achievement in schools, similar to this current study. It involved a quantitative approach, using a questionnaire. Some experts consider this approach to be "better… leading to results that are more believable" (Lichtman, 2012, p.44). However, Dedering and Müller's (2011) results cannot be generalised, because the study was conducted on a small sample (468 completed questionnaires). In contrast, a mixed-methods approach was adopted in this current study, comprising both quantitative and qualitative research methods. In addition, Dedering and Müller (2011) describe certain features of SPE in the context of the German education system, which is concentrated in individual school systems, as opposed to the more centralised Saudi system. In Germany, according to Dedering and Müller (2011), there is no national, unified Federal school evaluation, but each Federal State has its own distinct system.

Nevertheless, there are other important points involving SPE in Germany; according to Dedering and Müller (2011), all 16 of Germany's Federal States have the same school inspection objectives, and the school inspection itself provides detailed information on the quality of individual schools. This is subsequently used by the schools themselves to target quality improvement, as well as by administrative bodies to formulate measures of support and governance. In addition, school inspection is implemented as a systematic, evaluative assessment of working conditions, methods, and outcomes of individual schools, thus complementing the former State school supervision. Moreover, the inspections are based on standardised criteria for evaluating good instruction and good schools, according to the expectations determined by the administrative bodies. These criteria are laid down in the framework for school quality across all the Federal States.

Further to the above, no individual feedback is given to teachers as part of their monitoring in Germany. Instead, 50-70% of lessons are observed by a team of inspectors, while the head teachers, teachers, students and parents are interviewed (Dedering & Müller, 2011, p.303). These school inspections are always conducted by teams of three or four experts and schools are informed in advance of the date of the school visit. In all the Federal States, internal and external data concerning school locations are therefore collected and processed by school inspectors, prior to their visit (ibid., p.304).

Despite differences in the systems, Dedering and Müller (2011) findings are crucial for this study. First, the degree of authenticity and comprehensibility of the inspection reports was regarded as positive, with 89.9% of the interviewees considering them to be appropriate in their scope, aptly concrete (87.8%), and comprehensive (81%). In addition, most of the participants who were categorised as school head teachers considered the inspection report to be relevant for school development and believed that it would help the school management with further administration processes. Although these findings are credible, according to Newby (2014), they would have been more in-depth, if the researcher had used mixed methods. For example, interviews would have enabled a deeper understanding of the results.

The second study, with strong links to the present research, was conducted by McVeigh (2016), who reviewed the Ofsted criteria. This corresponds to the present study, whereby the effectiveness of the SPIS and its criteria are addressed. The above study produced interesting findings in response to the 23 Ofsted inspections reviewed. In addition, most of the head teachers studied did not feel that Ofsted's criteria undermined their professionalism or autonomy, while most of the teachers investigated appeared to consider that Ofsted's criteria, influenced by government policies, were necessary. On being interviewed for the above study, the head teachers and teachers agreed that the focus on teaching performance being shifted towards the bigger picture, evaluated over time, and taking into account pupils' progress and attainment, was a positive development, because Ofsted had linked its teaching criteria to teaching standards established in 2012, but had failed to justify them. McVeigh's (2016) main aims were therefore "to review the development of the criteria devised by Ofsted by which inspectors judge the quality of teaching in mainstream schools and to gain primary head teachers' and teachers' views on the criteria and their enactment" (p.2). This shares some common ground with the present research, which discusses the influence of the SPIS process.

In terms of the generalisability of the research, Bryman (2016) claims that the results of interviews cannot be generalised, which also applies to McVeigh's (2015) study. However, his research poses the following questions: How have Ofsted's criteria for evaluating the quality of teaching changed since 1993 (when Ofsted inspections first began)? What have been the key policy drivers and other influences on teaching criteria? How do the 10 selected primary head teachers view the Ofsted criteria and use them to influence classroom practice in their schools? Finally, how do primary school teachers from the head teachers' schools view the Ofsted criteria and use them to influence their own classroom practice? From these questions, it may be observed that there are several differences and similarities between McVeigh's (2015) study and the present research. The first difference is that McVeigh (2015) focussed on Ofsted's history and development, while this study is interested solely in its influence on the SPIS process and not its history. Moreover, McVeigh (2015) aimed to identify the influence of teaching criteria, whereas this study discusses the influence of SPIS in secondary schools, based on the perceptions of teachers and head teachers. Furthermore, McVeigh's sample includes primary school heads, teachers and inspectors.

In particular, McVeigh (2015) used semi-structured interviews with key actors: Her Majesty's Inspectors (HMI). More specifically, he used semi-structured thematic interviews with 10 primary school head teachers and 19 teachers. In addition, historical policy documents were studied, thereby including important features of Ofsted. Ofsted inspections were established in 1993 for secondary schools and the following year for primary and special education schools. To prepare for these inspections, schools receive an outline of the framework and a handbook. The framework for inspection has been revised many times, with versions in 1993, 1994, 1995, 2000, 2003, 2005, 2009, January 2012, September 2012, 2013 and 2014, and a further change in September 2015. The handbook covers performance and procedures in all areas of school life. Moreover, it makes explicit statements about what to look for in lessons. These handbooks have been made available to all schools for use by head teachers. However, McVeigh (2016) reports that the majority of head teachers do not feel that performativity (Ofsted criteria and indicators) undermines their professionalism or autonomy. Meanwhile, the annual reports provide information for schools and teachers, which help define what Ofsted considers to be best practice. Finally, Ofsted has produced guidance on what inspectors look for when they undertake subject survey visits (McVeigh, 2015, p.77).

The third study was conducted by Perryman et al. (2018) to answer the following question: "To what extent do inspection regimes, particularly the Office for Standards in Education (Ofsted), influence the work of a school, and how might that influence be conceptualised?" (p.45). The aim of the above study was to investigate how schools endorse policy. To accomplish this, the study had two main objectives – one being theoretical – to advance a theory of policy enactment. The other was experiential, to engage in a critical exploration of the differences in policy implementation in 'similar' contexts (Ball et al., 2012). The research question in Perryman et al.'s (2018) study more or less resembles one of the questions addressed in the present study, with a few differences. For example, the current study focusses on specific aspects of school life, such as school improvement and monitoring, while Perryman et al. (2018) were interested in Ofsted's influence on leadership and management, and whether there was an element of resistance. For this, the above researchers used a qualitative approach, collecting data from four co-educational, non-denominational and non-selective secondary schools. In contrast, the present study uses a mixed-methods approach.

Unfortunately, Perryman et al.'s (2018) study cannot be generalised, because of the small sample size, including just four schools and 95 interviews with "head teachers, senior management, teachers, union representatives and support and advisory staff" (Perryman et al., 2018, p.150). However, most of the participants spoke positively about Ofsted, because it had helped them improve their schools. In addition, it was clear that the schools were not given any notice of the time or date of their evaluations. Instead, the head teachers were committed to continual Ofsted readiness in their schools; a perpetual state of inspection anxiety that aimed for good or outstanding practice each day of every week for the whole academic year. In this way, the leaders' agency used the inspection tool to exert pressure on head teachers, so as to raise standards in schools.

The fourth study was conducted by Janssens and van Amelsvoort (2008), who explored the conditions and use of self-evaluation by schools. It also looked at responses relating to the effects of school inspectorates in countries with similar school self-evaluation (SSE) developments (for example, the Netherlands, Belgium, Denmark, the UK and Germany). It therefore explored the influence of SSE on eight schools in seven European countries. However, the above study differs from the current study in terms of its aims and questions, as the present research focuses on the influence of SPIS on school performance, with SPIS being an external evaluation process. Moreover, this current study involves collecting data via both quantitative

and qualitative methods, in contrast to Janssens and van Amelsvoort (2008), who only used interviews with 17 inspectors. Another difference is that the current study considers the perceptions of teachers and head teachers from all secondary schools in Jeddah, Saudi Arabia, with three teachers and head teachers from three schools also being interviewed. In contrast, the sample size in Janssens and van Amelsvoort's (2008) study is limited.

In all the countries investigated by Janssens et al. (2008), it was found that the legislative position underpinning SSE was aimed at improving school performance. However, the legislation in some parts of the world, such as Belgium, Northern Ireland and Scotland, also suggests SSE reporting, with external evaluation then being conducted on SSE (in England, Lower Saxony and the Netherlands). However, England, the Netherlands, Northern Ireland and Scotland appear to have adequate standards, which contribute to good self-evaluation and school improvement. To support SSE, the country must have frameworks or guidelines aimed at school improvement.

However, these four studies (i.e. Janssens & van Amelsvoort, 2008; Dedering & Müller, 2011; McVeigh, 2016; Perryman et al., 2018) are similar in terms of the context in which they were conducted (namely, the West). However, they are different in several respects, such as the sample type and size. In addition, they all focus on school improvement as an impact of evaluation, except for Perryman et al. (2018), who did not determine the influence of evaluation.

As far as the Arabic context is concerned, some interesting studies have been conducted on school evaluation (Esan & Hamid, 2013), with six being somewhat similar to the current research, as discussed in the following paragraphs. Two studies on evaluation in general are addressed first, followed by four studies on specific programmes of school evaluation.

First, Esan and Hamid's (2013) study is concerned with the reality of school evaluation in the UAE, Qatar and Sultanate of Oman. It consequently compares assessment systems across these three countries; identifying their need for revision and making recommendations to help develop them. In contrast, the current study focuses on SPIS, a Saudi SPE programme. Meanwhile, Esan and Hamid (2013) adopted a comparative descriptive approach to analysing documents relating to the evaluation institutions in the above-mentioned countries, whereas this current study does not use any documentation, because it is focused on the perceptions of head teachers and teachers with regard to various aspects of the influence of SPIS. Moreover, Esan and Hamid's (2013) study posed the following questions: "What is the philosophy of school

institutional evaluation and what are its tools and procedures?", "What are the updates facing the institutional and school calendar specifically?", "What is the reality of institutional and school assessment in the UAE, Qatar and Oman?", and "What are the suggested recommendations for developing the system?" (p.338).

Esan and Hamid's (2013) research helps elucidate school evaluation in some Gulf countries. However, there are other important features of school evaluation in this zone (for example, in the UAE and Oman). Multiple and conflicting objectives and priorities point to a gap between assumed and real institutional values, with a discrepancy between the values of senior leadership and the values of school staff. Furthermore, there were found to be insufficient resources for solving problems in schools.

To elaborate on the above, Esan and Hamid (2013) refer to the system and some of the school evaluation policies in use in the UAE. The UAE began implementing its academic accreditation programme in 2008, with British expertise being engaged to establish a bespoke SPE unit. All education leaders and school head teachers were trained in implementing the evaluation programme and methods of supervision. School accreditation in the UAE now involves academic recognition by the UAE's Ministry of Education that a school has achieved a certain standard of quality. Similarly, Oman has established a programme for enhancing school performance, which includes a sub-programme for evaluating school performance and developing the schools concerned. Here, an internal and external evaluation system is applied, with Oman using the Ofsted programme to develop a school performance assessment plan. In contrast, Qatar has established an Education Assessment Board with four offices: The Student Evaluation Office; the Evaluation Office, which awards teaching licenses; the Office of Performance Evaluation of the Practitioner, and the Office of Information and Data Collection. Meanwhile, Qatari Academic Accreditation means that the quality of education in a school is recognised by the Qatari Ministry of Education.

Another study, conducted by Moussa (2012), set out to explore the foundations and methods of evaluating the quality of modern schools and their approaches, as well as reviewing global best practices and models. In addition, it aimed to uncover the reality of private schools in KSA and the prevailing practices for assessing the quality of performance and the obstacles hindering improvement in the quality of output. The above study relied on analytical descriptive and comparative approaches, in stark contrast to this current study, which explores staff perceptions and experiences of education reform in secondary schools in Jeddah, KSA (using SPIS as a

specific example). In addition, this current research involves a case study, using mixed methods and setting out broad terms for assessing the effects of the SPIS process and grading outcomes on various aspects of school life.

However, although this current study differs from Moussa's (2012) work, it draws upon the research questions of the latter to gain a deeper understanding of SPE in Saudi Arabia. Moussa's (2012) study addressed the following question: "How can the quality assessment practices of private schools in Saudi Arabia be developed in light of foreign experiences?" This then gave rise to the following sub-questions: "What are the intellectual and organizational bases related to evaluating the quality of the performance of private schools?", "What are the comparisons and lessons learned from international experiences in evaluating the performance of private schools?", "What is the reality of the performance of private schools in Saudi Arabia?", and "What are the most important recommendations and proposals for the development of quality assessment practices for private education in the KSA?" (p.342).

According to Moussa (2012, p.340), the reality of private sector schools in KSA may be described as follows: there is a significant lack of indicators on which to classify them. In addition, teachers and head teachers rarely possess the appropriate educational qualifications for working in private schools, because of the lack of clear specifications and standards for the personnel employed in these schools. Moreover, there is no correlation between the results of evaluation and the realisation of the Ministry's vision. However, there are five Saudi studies, which have concentrated on a specific programme, namely comprehensive evaluation; this being the SPE programme implemented between 2001 and 2011. These studies comprise Al Dossary (2006), Alballawi (2009), Al Sheikh (2010) and Alrwqee (2012).

The above-mentioned Saudi studies mainly explored the topic of comprehensive evaluation and the extent to which it affected school performance. Al Dossary (2006) addressed the question of the extent of comprehensive school evaluation's effectiveness, from the perspective of a supervisor, head teachers and teachers in Mecca, KSA. Later, Alballawi (2009) questioned the degree of effectiveness of a comprehensive assessment programme in diagnosing the reality of a school, from the point of view of school head teachers in Tabuk education in KSA. Additionally, Al Sheikh (2010) posed the same question, but conducted his study in a different region; investigating the role of comprehensive assessment in improving the performance of education administration in the Asir region of Saudi Arabia. Finally, Alrwqee (2012) investigated the reality of the comprehensive assessment system applied in Saudi state schools,

in light of an appropriate criterion for evaluating a comprehensive evaluation system within total quality management (TQM). Besides, the above author explored the most prominent problems facing this evaluation system from the point of view of the study sample.

All the above authors used a descriptive approach, in addition to quantitative methods, such as questionnaires. However, their samples and sample sizes varied; for example, Al Dossary (2006) included 115 supervisors and a sample comprising 30% of head teachers and teachers from Mecca's primary schools (156 school head teachers; 766 teachers), out of a total study population of 1037 individuals. Meanwhile, school assistants were added to Al Dossary's sample, while Alballawi (2009) included school head teachers. These studies were concentrated in specific parts of KSA.

In contrast, Alrwqee (2012) conducted a more general study, with samples drawn from all over KSA. The findings from most of these studies generally support comprehensive evaluation. Al Sheikh (2010) also describes comprehensive evaluation as producing high scores for general factors and medium results for more specific factors. This may be because most of these studies, such as Al Dossary (2006) and Al Sheikh (2010), were supervised by individuals working within the programme, which means that the results could be subjective, as it is not always easy for individuals to criticise their own work.

Overall, it would seem that Alrwqee's (2012) results are more objective than those obtained in the other studies. For example, Alrwqee (2012) presents both negative and positive results; revealing that the comprehensive evaluation system had weaknesses in several areas; referring to a lack of quality in the evaluation performed, as a specific reason for the increase in teachers' stress and workload, especially in matters of organisation and co-ordination. According to Alrwqee (2012), this resulted in dissatisfaction among school head teachers, concerning the evaluation system.

The studies described above share similarities with the present research, which looks at the effectiveness of evaluation on school performance, but there are also significant differences. First, the present study is mainly concerned with external evaluation, while previous studies have concentrated on internal evaluation. In addition, this study uses mixed quantitative and qualitative methods, whereas quantitative methods were used in all the previous research reviewed. Moreover, this study investigates the evaluation of performance management using SPIS in Saudi Arabia.

Overall, the authors of these earlier studies refer to the importance of self-evaluation. They also consider its format, which should be similar to that of external evaluation, offering significant benefits to schools. In addition, previous researchers have argued that teachers need to become more aware of self-evaluation and understand its importance. Finally, it is posited that self-evaluation criteria and indicators require further development to render them effective. In addition, most of the study participants found these evaluation programmes necessary, believing that they would help enhance the performance of their schools. Finally, it was found that these programmes helped increase accountability among teachers and head teachers in Saudi education (Alrwqee, 2012).

3.6.2 Empirical Research Studies on Performativity

Several studies have been conducted on the topic of performativity; these include Penninckx, Vanhoof, De Maeyer and Van Petegem (2016) and Ehren, Perryman and Shackleton (2015) . These researchers followed the research stream initiated by Plowright (2007), and the four studies that come closest to this current research project are discussed below. This subsection provides a descriptive overview of researchers' efforts to characterise performativity. According to Mayer, Mitchell, Santoro and White (2011), a great deal of research on teacher training has focused on performativity. Wilkins (2011) claims that over the past decade, several attempts have been made to describe its effects. Correspondingly, this subsection discusses the influence of performativity on some of the general empirical research, as well as in the four studies that share a common interest with this current research.

The first of these studies was conducted by Penninckx et al. (2016), using a quantitative approach to discover the implications of performativity from every angle. Thus, quantitative data were gathered using an online survey, which was sent to 2202 teachers in primary and secondary schools that had recently been evaluated by the Flemish Inspectorate. All the schools in Penninckx et al.'s (2016) study had been inspected at least once every 10 years. The researchers were interested in teachers' perceptions in primary and secondary schools, with respect to "the impact of the inspection judgement in terms of these effects, the schools' policy-making capacities, and the inspection quality" (p.336). The above findings were influential in shaping this current study; for example, due to their interest in the influence of evaluation. Besides, Penninckx et al. (2016) used an online survey, which resembles the one implemented in the present study.

In addition, the above researchers were interested in teachers' opinions, as is the case in this Saudi study, but the latter also includes head teachers' perceptions of secondary schools. Other differences between the two studies refer to the geographical context, with the previous study being conducted in Belgium, while the present study was undertaken in Saudi schools, where there is a very different culture. Furthermore, mixed methods were used in this current study, while the previous research was purely quantitative. Penninckx et al.'s (2016) findings suggest a significant increase in stress and anxiety in schools with strong policy-making capacity, but this increase is less evident where the quality of the inspection is enhanced. Thus, the researchers recommend encouraging inspection systems to improve teachers' understanding of the system. In short, findings demonstrating that teachers who perceive their school to have a high level of agency in determining its own policy are more likely to be stressed and anxious.

Similarly, Ehren and Visscher (2008) discuss the effects of school inspections on school improvement in Dutch schools. The study used an exploratory approach to test six hypotheses, choosing case studies on 10 Dutch primary schools from the years 2002-2005. The researchers began by administering questionnaires to 567 Dutch primary schools undergoing inspection, with just 190 schools completing the survey. The authors subsequently selected the "ten per cent of schools with the highest innovation capacity and the ten per cent of schools with the lowest innovation capacity" (Ehren & Visscher, 2008, p.215). Furthermore, head teachers, co-ordinators, pupil care staff and teachers were interviewed before and after inspection, while a questionnaire was administered to the head teachers after the inspection and observations were conducted during the inspection visits. Hypotheses were subsequently discussed for school improvement and inspection, in terms of whether the inspection should include elements to support school improvement and whether the school had a high level of innovation capacity. In addition, the study considered whether the performers/bodies working towards improving school performance were being sufficiently creative.

Ehren and Visscher (2008) found that the greater the number of school improvement initiatives to complement inspection, the greater the positive influence on schools. In addition, if a school had the opportunity to discover its strengths and weaknesses, it was more likely to accept feedback. The above study found that schools made remarkable progress after the inspection, but there were two areas where progress was lacking, namely in the schools' capacity for innovation, and in the school environment. Additionally, school improvement processes were

unaffected by the low scores awarded by inspectors, as well as their feedback, suggestions for improvement, and the agreements made.

In light of the above, some differences become evident between Ehren and Visscher (2008) and the current study. For instance, the former is based on the policy theory underpinning the Dutch Educational School Supervision Act, while the present research is based on SPIS in Saudi Arabia, which is an independent organisation. Therefore, although both studies involve on external evaluation, qualitative methods were used in the Dutch study and mixed methods were used in the current study, in response to two research questions. However, one of the sub-questions of the Saudi study is similar to one posed by Ehren and Visscher (2008), although the Saudi study differs in that it investigates the influence of SPIS evaluation.

Likewise, Plowright (2007) investigated how teaching staff and school managers felt about their school's self-evaluation procedures, which were implemented in preparation for an Ofsted inspection. However, there are two main differences between the present research and Plowright's study. First, this current study is not merely concerned with the participants' views on external evaluation or the relationship with school performance improvement, but also investigates the participants' views of the effectiveness of this evaluation. Meanwhile, Plowright used a case study method, concentrated in a single school, whereas the present study sampled different schools to collect data. Moreover, Plowright combined a questionnaire survey – involving all teaching staff – and semi-structured interviews with individual members of the senior leadership team, as well as group interviews with a cross-section of heads of department for the three national core curriculum subjects: Science, Mathematics and English. Although this approach resembles that of the present research, the case study differs in many ways. For example, Plowright (2007) classified the data collection methods according to the type of participant, whereas the same methods were adopted for all the participants in the present study.

It should also be noted that Plowright's main weakness was limiting the data collection to a single school, as it meant that just one environment was explored. Nevertheless, the study findings are interesting and show that the head teachers believed self-evaluation and preparation to be helpful for managing an Ofsted visit, while the teachers considered preparation to be a positive approach to addressing weaknesses. However, according to Savin-Baden and Major (2013), effective and representative sampling leads to robust research. Therefore, since the sample studied in Plowright's research was limited to a single setting, this may have affected the validity of the findings.

Some researchers claim that SSE, where performativity is applied, generates extensive discussion on the nature of professionalism (Dedering & Müller, 2011). Dedering and Müller (2011) describe the impact of performativity evaluation in Germany, having undertaken a survey study with 468 school head teachers from 2005-2008. Although there are similarities between Dedering and Müller's (2011) study and the current research in terms of aims and goals, they differ in other respects. For instance, this Saudi study involved mixed methods, while Dedering and Müller (2011) adopted a quantitative approach. Moreover, this study focussed on Saudi schools in Jeddah, while Dedering and Müller (2011) examined German schools.

Ultimately, Dedering and Müller (2011) found that the degree of authenticity and comprehensibility of inspection reports was regarded as positive, with 89.9% of the interviewees finding the inspection report to be appropriate in its scope, aptly concrete (87.8%), and comprehensive (81%). In addition, most of the participants categorised as head teachers considered the inspection report to be relevant to school development, in the belief that it would help the school's management with further administrative processes. Although these findings are credible; according to Newby (2014), they could have been explored in more depth if the researcher had used mixed methods. Aside from this, disturbing effects were rather scarce after the inspection.

The four studies described in this section were conducted in a Western context, but they differ in the way that the data were collected, because Dedering and Müller (2011) and Penninckx et al. (2016) used quantitative methods, while Plowright (2007) and Ehren and Visscher (2008) used mixed methods. Additionally, all these studies were concerned with external evaluation, except for Plowright (2007), where the main interest lay in internal evaluation. However, there are several reasons why the findings from all the above could promote understanding in this study: mixed methods were used, and internal evaluation was important for preparing external evaluation. Thus, it could be helpful to understand how internal evaluation works.

Nevertheless, in Saudi Arabia, there appeared to be a dearth of studies on the effects of this type of evaluation, while studies conducted in the West tend not to be sufficiently comprehensive to fill the gap identified in the literature, because the Saudi education system is centralised, unlike the Western context.

3.6.3 Empirical Research on Accountability

Accountability is a subject that has attracted substantial attention in the literature, especially over the past 20 years, with important debates on accountability taking place around the world (Kwok, 2011). This may be due to the fact that it is one form of apparatus for education reform in the wake of neoliberalism (see section 2.2). However, the main focus of this attention has been the influence of accountability on teachers (for example, Berryhill, Linney, Fromewick, 2009; Buchanan, 2015; Ingersoll, Merrill, & May, 2016). In terms of teachers' identities, Buchanan (2015) concurs with claims made by Sloan (2006) and Pease-Alvarez, Samway and Cifka-Herrera (2010) that accountability has shaped the identity of teachers by its demands. In fact, the areas for which teachers expect to be held accountable shape their identity, irrespective of what they experience in their profession, or their existing identity.

Teachers' stress and accountability were discussed by Berryhill et al. (2009) in a study that examined the perceptions of primary school teachers in the USA, concerning State policy on accountability, especially the impact of policy on functional participation. One hundred teachers from nine primary schools participated in the above study, with Berryhill et al. (2009) using mixed quantitative and qualitative methods, whereby 100 teachers completed 1000 questionnaires and nine teachers were interviewed. One of the most important results generated by the above study was that the pressure resulting from accountability led to conflict and low self-efficacy amongst teachers. The participating teachers proposed several suggestions to avoid this: the most important being the enhancement of the school environment and appointment of an assistant for each teacher. However, the suggestions made by these teachers merely seemed to be an attempt to create an environment that would help them meet the accountability requirements, rather than an effort to make accountability more flexible.

Although Berryhill et al.'s (2009) research is similar to the current study, in that both quantitative and qualitative methods were used, they differ in many aspects. Berryhill et al. (2009) conducted their research in the USA, where the system is decentralised, in contrast to this current study, which refers to the centralised Saudi education system. The present study also examines the views of teachers and head teachers in terms of a specific programme of evaluation, while the US study examined accountability and its impact. Nevertheless, in accountability and teachers' efforts to help their students achieve in national tests, Rockoff and Turner (2008) found a link between accountability and enhanced student achievement in US

schools. This may be due to the fact that one of the most important indicators of school performance is student achievement in national tests, which are motivating to lead teachers.

In this subsection, the three studies appear to share some common interests with the current research. These studies were conducted in the UK, Saudi Arabia, and elsewhere in the Middle East. The first claims that the UK education system has the highest level of accountability in the world (Barzanò, 2009), based on the views of head teachers involved in implementing accountability and on document analysis, triangulated by interviews with five head teachers and two policy-makers. The participants were asked about their perceptions of the accountability system in UK education, specifically in England. The study findings proved to be interesting; suggesting that the head teachers had positive opinions of the potential for performance standards to provide them with robust feedback, thereby helping them improve their school's performance. However, they also criticised the way in which the results were presented, especially as these were made available to non-specialists who could misjudge teachers, due to their lack of knowledge of the standards. The above study found that policymakers faced two issues in the system concerned, namely their duty to ensure that teachers did their best to bring about school improvement, and "their commitment to defend teachers from the intrusiveness of formal accountability" (p.193). Barzanò's (2009) work shares some common ground with the current study, in that the head teachers' opinions on performance evaluation were of interest, representing accountability in the context of English education. Meanwhile in this current study, the focus is on SPIS in Saudi education.

Additionally, Esan and Hamid (2013) conducted a study to determine the application of management in the percentage of accountability within government schools in the Sultanate of Oman, and the requirements for indicating differences in these requirements based on gender variables and years of career experience. The sample was selected in statistical form, comprising employees from government departments in Oman. To collect the data, a questionnaire was developed to include three axes: a culture of administrative accountability and transparency, accountability and administrative instruments, and the terms for the management's accountability team. The Statistical Package for Social Sciences (SPSS) was applied to process the data statistically, generating the results in percentages. For example, the study sample estimated that the requirements for applying administrative accountability in the government of the Sultanate of Oman was high. However, there were some differences in what these statistics indicated, due to the variables for accountability, the management team, and the

benefit of having achieved higher degrees (PhD/Master's). There were also differences in statistical indications, based on years of experience in general functional requirements, the management accountability team axes, and the figures for most experience.

In addition, Alguhidan (2009) sought to achieve the following objectives: identifying the degree of application of educational accountability, in terms of discipline and professional performance, ethics, and personal behaviour. These objectives were specified for determining social relations in public sector girls' schools in Mecca, from the perspective of the study community. Moreover, it sought to identify the requirements for activating educational accountability from the point of view of directors of public sector schools. Furthermore, it attempted to identify the obstacles facing the implementation of educational accountability in public sector girls' schools in Mecca, as well as from the perspective of the directors of these schools. Finally, it sought to determine whether there were any statistically significant differences between the three populations sampled in the study, in order to determine the degree of application, and the existence or absence of statistically significant differences across the study community, with regard to the obstacles faced as a result of applying accountability.

Alguhidan's (2009) study was based on a descriptive approach, with a questionnaire used as the data collection tool. A random sample was selected from across the previous stages. An important result of this study was the participants' perception that educational accountability is generally applied in public sector girls' schools, whereby the study community strongly agreed with the requirements for activating educational accountability in these schools. Moreover, the participants agreed that there were constraints on the application of accountability. In contrast, SPIS evaluates school performance by testing implementation against indicators. According to Perryman (2006) and Ball (2012), this is performativity. The current study considered the perceptions of head teachers and teachers regarding the influence of SPIS on school performance, especially in terms of monitoring, teachers' wellbeing, and school improvement.

However, the four studies described here differ in some respects. For example, although Berryhill et al. (2009) used mixed methods, similar to Barzanò (2009), the samples recruited were different; Berryhill et al. (2009) concentrated on head teachers, while Barzanò (2009) investigated both head teachers and teachers. Meanwhile, in studies conducted in the Arab world by Alguhidan (2009) and Esan and Hamid (2013), using a quantitative approach, the sample in the second study involved various employees from different government departments in Oman, whereas in Alguhidan's (2009) study, teachers and head teachers were surveyed.

Consequently, this renders the present study worthwhile, as it took place in the context of Saudi education and was concerned with SPIS, a specific calendar programme. It also gives a voice to teachers and school head teachers to clarify their opinion of SPIS application, using mixed methods to increase its validity and reliability when collecting data from different sources.

3.7 Research Questions Emanating from the Literature Review

This literature review explored the forces driving the global growth of SPE systems; revealing that SPE constitutes a key recommendation made by international organisations, such as the World Bank, since education is viewed as a vehicle for driving the growth of national economies (Regmi, 2017). In consequence, there has been much debate among scholars, concerning the benefits and problems associated with SPE. It transpires from this review, however, that although a large body of evidence exists to provide understanding of many aspects of SPE in Western contexts, there is comparably less empirical research from Saudi Arabia. The little research that has been undertaken in Arabic contexts tends to focus on the extent to which evaluation affects school performance (for example, as explored by Al Dossary, 2006; Alballawi, 2009; Al Sheikh, 2010; Alrwqee, 2012; see also subsection 3.5.1). Other Arab studies (for example, Alguhidan, 2009; Esan & Hamid, 2013; see also section 3.5) have mainly focused on the relationship between evaluation and centralisation.

In contrast, research in Western contexts has explored more novel aspects of SPE, such as the factors that ensure its effective implementation, and the efficacy of indicators used to evaluate school performance. Examples of these efforts include Rowe (2004), Evans (2011) and Dangerfield (2012) (see subsection 3.5.1). Thus, the review of the empirical literature revealed a research gap in Saudi Arabia, which this present study could begin to close. Hence, the first research question was formulated to explore teachers' and head teachers' perceptions and experiences of SPIS school monitoring and whether it was viewed by them as improving school performance.

The review also identified Western studies that explored the influence of SPE on teachers and head teachers. Notable examples included De Maeyer, Van Petege and Plowright (2007) and Perryman, Ball and Maguire (2011). Coupled with this, related theories on performativity and accountability in research, which formed part of the conceptual framework of this study, shaped the direction of the second research question, which concerns the experiences and perceptions

of teachers and head teachers, regarding the influence of SPIS on their stress levels, workload and morale.

Alongside the review of theoretical debates and empirical research that highlighted the role of accountability in SPE, numerous voices had called for educational reform using performance evaluation in Saudi Arabia (see1.2). Consequently, this current review of the literature, coupled with the researcher's local knowledge of SPE policy and practice, generated research questions relating to teachers' and head teachers' understanding of accountability under SPIS, specifically in relation to school improvement. Accountability was also an imperative focus of this research, since there have been successive SPE programmes and operational overlaps between some of those programmes over the past decade, with little evidence of improvement in educational outcomes in Saudi Arabia. Furthermore, the direction that the research questions took in this study was underlined by the realisation that to the best of the researcher's knowledge, these aspects of SPE have not been the subject of empirical research in Saudi Arabia.

In sum, the process of formulating the research questions was an iterative process, which involved moving back and forth between the key theories discussed in the literature review that laid the foundation for the conceptual framework for this study, as well as exploring prior empirical research in Western and Saudi contexts, and drawing upon the researcher's knowledge of policy and practice in Saudi Arabia.

3.8 Summary

This literature review has explored the influence of SPE on school life, according to three main concepts: the evaluation of school performance, accountability, and performativity. Five conclusions may be drawn from the arguments surrounding these concepts. First, there is a strong link between SPE and governments' economic goals worldwide, especially for achieving high-level outcomes that will be competitive in the labour market. In addition, this evaluation leads to increased accountability amongst teachers in their professional lives. School evaluation involves two types of evaluation: external (inspection) and internal. Both types of inspection should work in harmony and consistently, in order to achieve the goals agreed upon between schools, the government, stakeholders, parents, social scientists and economists. External evaluation can help the government set its goals and reform education according to national and societal needs, while internal self-evaluation can help teachers carry out evaluation and identify

weaknesses and strengths themselves, so that weaknesses can be addressed and strengths, supported.

The positive and negative aspects of SPE have been mentioned in multiple studies, but there is a consensus that evaluation procedures should be clear and easy to implement in schools. Evaluation reports should also include recommendations that are easy to adopt to correct weaknesses. Nevertheless, there is a scarcity of literature that includes teachers in discussion of the effects of performance evaluation on their stress and workload. Moreover, in studies that consider the views of head teachers and inspectors in Saudi Arabia, the SPIS programme has not been discussed, while other evaluation programmes have only been considered in quantitative studies. This makes the present study unique in that it uses both quantitative and qualitative methods. As such, the reviews in this chapter will be taken into account when analysing the research results in Chapter 5.

Chapter 4: Research Design and Methodology

4.1 Introduction

This chapter introduces the methodology applied in the present study, with the aim of determining the extent and ways in which the evaluation of school performance via SPIS influences teachers and head teachers in Jeddah, Saudi Arabia. Since the 1990s, performance and its evaluation have attracted increasing attention from scholars and various stakeholders (Pollitt & Bouckaert, 2004; Dunn & Miller, 2007; Kalimullah et al., 2012; Jaksic & Jaksic, 2013). According to Zakaria, Yaacob, Noordin, Sawal and Zakaria (2011), performance management can even be adopted to improve the performance of governments. As discussed in Chapter 2 of this thesis, it is evidenced in the strong movement towards NPM (Dunn & Miller, 2007), which has emerged from neoliberalism worldwide (Betzel,2013). In education, performance management has been used to evaluate schools' total performance (Willms, 2003), with the goal of achieving school effectiveness and efficiency (Evans, 2011).

However, the application of performance evaluation in schools has helped give rise to the debate over its influence on teachers and school performance, in terms of the wellbeing of education professionals and school improvement. Globally, several researchers have investigated this influence (for example, Murray, 2012; Bailey & Colley, 2015; McVeigh, 2016). Therefore, in common with other countries around the world, KSA aims to improve its education system (see section 1.2), with the Saudi Ministry of Education establishing a programme to evaluate school performance, namely SPIS (see subsection 2.3.6.2).

In order to meet the aim of this study, the following research question was formulated:

> What are head teachers' and teachers' experience and perceptions of the influence of SPIS on school performance?

This question raises three sub-questions (RSQ):

> RSQ1. What are head teachers' and teachers' experiences and perceptions of the influence of SPIS on school monitoring?

> RSQ 2. What are head teachers' and teachers' experiences and perceptions of the influence of SPIS on their stress levels, workload and morale?

RSQ 3. How do head teachers and teachers describe and understand their accountability under SPIS in relation to school improvement?

This chapter begins by explaining the rationale for the selected methodology and discussing the ontological and epistemological aspects, the research paradigm, and methodological perspectives. It also attempts to justify the choice of a case study approach to address the research question, and to clarify the sampling strategy, data collection methods, techniques of analysis, validity, reliability and ethical issues.

4.2 Underpinning Rationale and Research Paradigm

A person's beliefs, values, language and experiences will affect his or her perspective of the truth and determine the form of knowledge acquired, the way that it is interpreted, and the individual's position in relation to it. According to Morrison (2012), these are the components of the paradigm that form the framework of a person's worldview. For Schwandt, Lincoln and Guba (1990), paradigms are beliefs or points of view that inform the investigation of educational phenomena. In addition, paradigms "are models, perspectives or conceptual frameworks that help us to organise our thoughts, beliefs, views and practices into a logical whole and consequently inform our research design" (Basit, 2010, p.14). Nonetheless, it is clearly important for researchers to identify the corresponding research paradigm, because this will indicate their philosophical position (Newby, 2014). Moreover, Bryman (2016) argues that the paradigm selected by a researcher will determine the way in which data are collected and interpreted. Furthermore, Easterby-Smith, Thorpe, Lowe and Chapman (1994) highlight the negative effect of omitting to establish or define a philosophical paradigm in research. Thus, the researcher must select a paradigm that relates to his or her philosophical perspective, especially in social science research, which has a complex relationship with philosophical theories. In addition, the research paradigm will lead the research methods; however, where no paradigm is specified, these methods will have a negative impact on the research methodology. Consequently, the research ontology and ethical considerations are essential components of a study to ensure the validity of the methods deployed. Therefore, as the researcher, I identified the philosophical position to be adopted in this study and selected the appropriate paradigm to address the research questions.

Conversely, Blanche, Blanche, Durrheim and Painter (2006) state that the research process consists of three major elements: epistemology, or the nature of knowledge; ontology, or the

nature of reality, and methodology. According to Singh (2007), the epistemological and ontological approaches adopted in a study will demonstrate how a person perceives the world.

4.3 Ontology

Primarily, ontology is defined as "the science or study of being" (Blaikie & Priest, 2019, p.92) and it is "concerned with what is real or the nature of reality" (Blaikie & Priest, 2019, p.25). Correspondingly, it refers to the researcher's view of reality and what he or she claims to exist (Guba, 1990). This means that reality is the result of the researcher's experience and education, and what he or she believes to be true. This refers to the subjective ends which adopted by this current research, in which there is no one reality (Bryman, 2016). Moreover, it indicates that the researcher's ontological position is determined by his or her experience, education and knowledge. Therefore, Grix (2002) declares that researchers need to understand the research ontology, in order to be able to clarify their position and ensure that none of these elements affect it negatively. From another perspective, Bryman (2016) states that the ontology applied enables an exploration of individuals' perceptions as a means of exploring reality. Therefore, the ontological position adopted in this current study relies on the perceptions and experiences of its participants. Here, reality is investigated by exploring the perceptions and experiences of head teachers and teachers, with regard to SPIS and its influence on their professional practice. Therefore, I chose a subjective view of reality, because I was interested in understanding the respondents' perceptions of reality. However, there were two possible paradigms that could be adopted (Bryman, 2016): ontological structuralism and ontological constructivism (Grix, 2002). According to Bryman (2016), the difference between these lies in the fact that the former views human behaviour as the result of individuals' values and the rules that they observe in their respective society or communities (for example, schools), while the second asserts that the interaction between people can explain social reality. In fact, these ontologies were bridged in this study, because the participants' school environment managed their professional lives and influenced their behaviour, but the participants interacted with each other and with other stakeholders, meaning that their experiences and perceptions were also affected by these interactions.

4.4 Epistemology

Epistemology is defined as "the theory of knowledge embedded in the theoretical perspective and thereby in the methodology" (Crotty,1998, p.3). In addition, according to Guba (1990) epistemology means "the nature of the relationship between the knower (the inquirer) and the known (or knowable)" (p.18). Additionally, Richardson (2015) asserts that knowledge may be acquired by identifying the relationship between the individual and the environment, while the way in which individuals view their knowledge is referred to as epistemology (Bryman, 2016). It follows that knowledge of reality comes from the art, culture, beliefs and tools used, such as the language of expression and way of life (Kaplan & Maxwell III, 1994). This means that the researcher's epistemological standpoint will be based on his or her beliefs, values or assumptions and the way in which he or she ascertains these to be true.

However, there are two key epistemological positions: positivism and interpretivism (Crotty & Unwin, 1998; Bryman, 2016), and both were considered in this study to justify the research approach. According to Major and Savin-Baden (2010, p.19), under a positivist paradigm, "the researcher can gain knowledge by identifying facts", and these facts are established by the senses. Thus, anything that cannot be ascertained through the senses is not true (Williams, 1996), implying that reality is what we can touch, test or listen to (Gray, 2004). Therefore, positivists use quantitative methods, such as questionnaires (Basit, 2010), and the results are generalisable. Positivism is suitable for research where a phenomenon can be measured (Cohen, Manion, & Morrison, 2013) and so it is not appropriate as the main paradigm in this current study, because it investigates head teachers' and teachers' perceptions and experiences of SPIS, which cannot be measured. Moreover, the sample size is small, which means that the results are not generalisable. Consequently, the most suitable paradigm for this study is interpretivism.

The interpretivist paradigm is "the view that human behaviour needs to be described and explained by individuals in the way it is perceived by them" (Basit, 2010, p.14). In other words, social scientists tend to "grasp the subjective" to find their answers (Bryman, 2016, p.29). For this purpose, qualitative methods such as interviews are implemented under this paradigm (Basit, 2010), because there is an attempt to understand how people interact with phenomena in their social environment, the reasons why they interact, and how they influence and are influenced (Orlikowski & Baroudi, 1991). Interestingly, the strongest contribution of the interpretivist paradigm is the voice that it gives to participants, so that they can present their

experiences and express their perception of the research problem (Orlikowski & Baroudi, 1991).

From the interpretivist perspective, exploring the influence of SPIS from the point of view of head teachers and teachers in relation to their experiences can lead to an understanding and perspective of the influence of SPIS on these professionals and on school performance. This is based on my belief that the creativity, knowledge, diverse backgrounds, and rich experience of head teachers and teachers can add depth to this study, leading to a better comprehension al the impact of SPIS and enabling the research questions to be answered in a more satisfactory manner. Therefore, an interpretivist paradigm was adopted in this study and a mixed methods design was applied for the data collection, which will be discussed in the next section.

4.5 Research Design: A Mixed Methods Approach

One of the most reliable research approaches was adopted in this study, namely an explanatory sequential mixed methods design. Mixed methods are often referred to as the third paradigm in research (Lichtman, 2012). This approach simply consists of bringing together quantitative and qualitative research methods (Lichtman, 2012; Newby, 2014). Thus, quantitative methods such as surveys are used to give a general picture and qualitative methods enable an in-depth understanding to be gained (Newby, 2014). Although a mixed methods approach is potentially beneficial for education researchers (Newby, 2014), it is important that researchers are explicit when they use mixed methods, because the outcome may not be predictable (Bryman, 2016). This means that the method should be used with a rationale of cogent reasoning. Undoubtedly, if the aim of the research requires data collection using a questionnaire, interview or observation, mixed methods will be appropriate (Lichtman, 2012).

Aside from the above, the use of mixed methods can strengthen the results and confirm them in a clear manner (Gorard & Taylor, 2004). In fact, each method has its strengths. Therefore, when using the two methods, the researcher can combine their strengths to obtain good results. By adopting a quantitative approach, the researcher can derive answers in response to issues that will merit further exploration using qualitative methods. This can also help with understanding the results that are obtained using a single method and avoid any confusion or ambiguity in certain areas (Sandelowski, 2003). Therefore, both methods were deployed in this study to obtain deep results, which were then interpreted and confirmed in both quantitative and qualitative ways.

Quantitative research reports objectively on reality in general, while qualitative research unlocks an understanding of a situation and its underlying factors. In the current study, the purpose was to discover opinions of a system, but it also looked at matters related to feelings, which meant using quantitative, followed by qualitative methods to identify what could not be deduced from numbers. The participating teachers felt pressured by and obligated to the system; therefore, it was crucial to discern information in their interview responses, in order to answer the research questions.

However, there are issues that arise when a mixed methods approach is applied, and these will now be considered. According to Lichtman (2012), the main problem with this approach is the possibility of gathering uneven or conflicting evidence. This can happen, because quantitative methods are usually conducted on a large scale to generate a high volume of data. This information can then be used by the researcher as a basis for gathering in-depth data, using qualitative methods. It ensures that the evidence gathered is equivalent to quantitative data, which is collected on a larger scale. In addition, this approach requires researchers to have a high level of knowledge about qualitative and quantitative methods (Onwuegbuzie & Johnson, 2006), although all researchers should have adequate knowledge about the methods that they use and be able to analyse the data appropriately. Moreover, since diverse views were collected here, using both quantitative and qualitative methods, important questions about the selected methods needed to be addressed. Firstly, it had to be established whether one method should take precedence over the other and if so, which. To answer these questions, the researcher required sufficient knowledge of the various techniques used in mixed methods research, in order to ascertain which would be most appropriate to answer the research questions (Subedi, 2017).

Mixed method typologies are categorised into four classifications: triangulation, embedded, explanatory and exploratory (Creswell & Clark, 2007). The differences between these four typologies are dependent on timing, variants, weighting and mix (Cameron, 2009). To clarify this, triangulation refers to the simultaneous use of quantitative and qualitative methods, with the methods weighted equally. Meanwhile, embedded mixed methods comprise either concurrent or sequential timing, where the researcher is free to choose which method to start with. In contrast, explanatory mixed methods are initiated by a sequential start, with quantitative and then qualitative methods. Finally, exploratory mixed methods are sequential, but begin with a qualitative technique. For the sake of clarity, these categories are illustrated in Figure 4.1. A

researcher will select one of these four types of mixed methods research, according to the research questions. For Creswell (2009), the choice of the first phase of a study will depend on the research aim and objectives. Correspondingly, the researcher can decide which phase should take priority and initiate the study. In this present case, an explanatory sequential mixed methods design was adopted, which is explained in more depth in the next subsection.
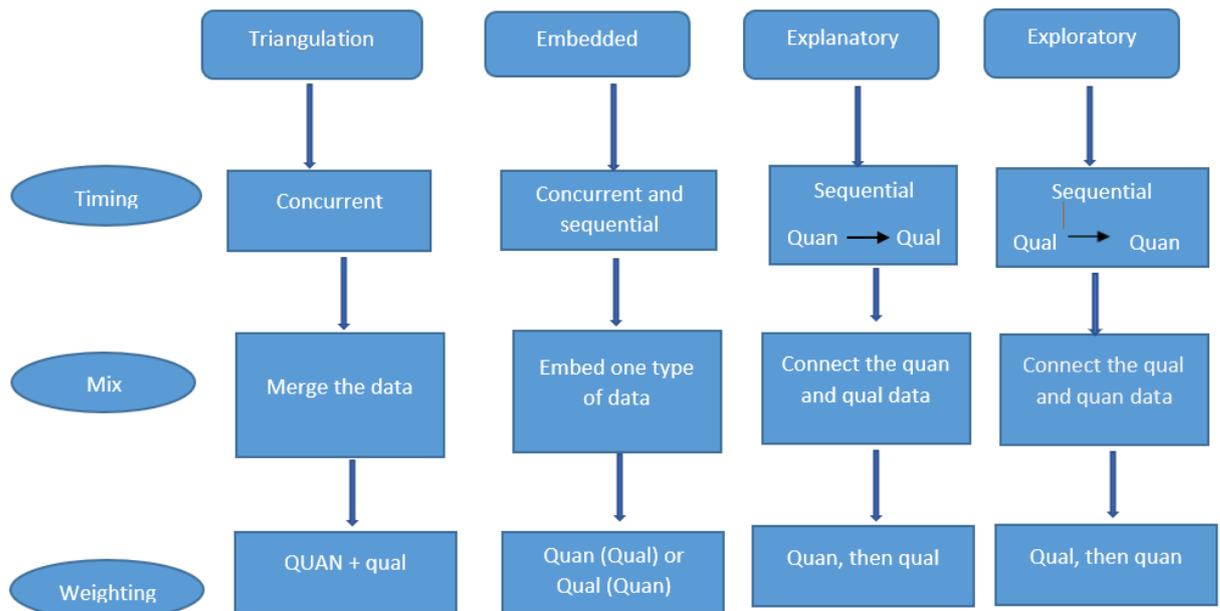
Figure 4.1 Types of mixed methods research design (source: Adapted from Creswell and Clark, 2007, p.70)

4.5.1 An Explanatory Sequential Mixed Methods Design

As indicated above, an explanatory sequential multi-site case study mixed methods design was adopted in this study. According to Creswell, Klassen, Plano Clark and Smith (2011), this type of mixed methods research includes a quantitative phase of data collection, so that points for further identification may be identified. It is then followed by a qualitative phase of data collection (see Figure 4.2).
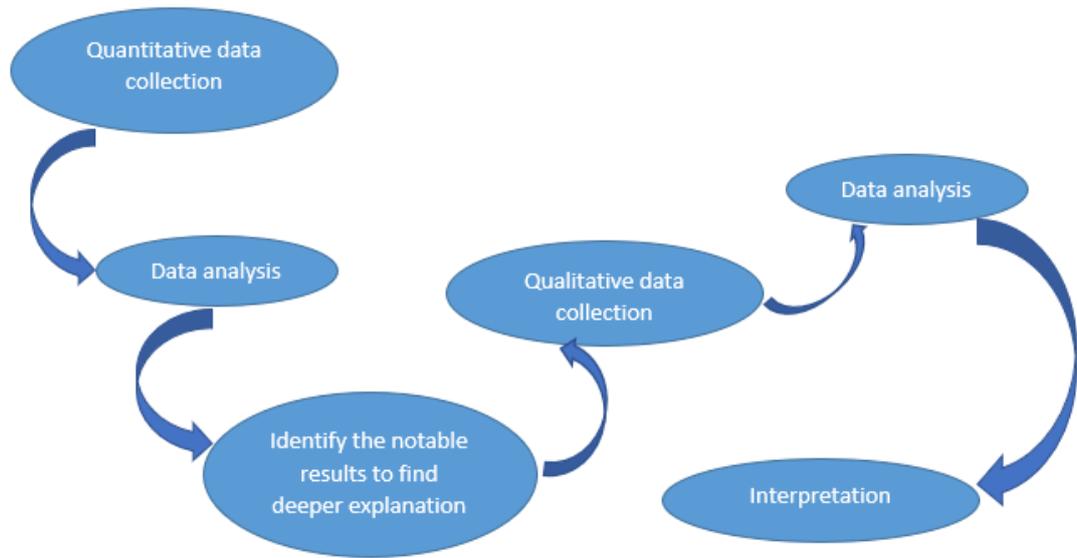
Figure 4.2 An explanatory sequential mixed methods design

Although this type of mixed methods enquiry is popular among researchers, it is not easy to implement (Ivankova, Creswell, & Stick, 2006); numerous challenges may face the researcher in the process of its implementation. The first of these difficulties relates to time, as the researcher should leave an interval between collecting the quantitative and qualitative data, so that the quantitative data can be analysed before the phase of qualitative data collection can begin.

Additionally, the researcher needs to take great care when determining which type of quantitative or qualitative data should be given priority. For this, the researcher must consider the purpose and objectives of the corresponding study, in order to find out which approach will yield general information, and which will yield deep meaning in the data gathered from the participants. As mentioned previously, the data collection in this study began with a quantitative phase to obtain broad information in response to the research questions, conducting Google Drive surveys from the 12th to the 22nd of September 2017. I subsequently closed the surveys and downloaded the participants' responses into an Excel file, before entering them into the SPSS program to obtain percentages for each type of response. However, because of the limited time allotted for the data collection period, beginning on 2nd August and ending on 30th October, I reviewed all the tables of responses and selected any cases that needed further explanation, such as multiple evaluation, increased stress among teachers, and the question of whether this affects their commitment (see Appendix II).

Nevertheless, it should be noted that I faced certain barriers to identifying answers from some important results, such as concerning the restrictions imposed by the Ministry of Civil Service and the Ministry of Education, regarding the unveiling of school performance improvement. These reasons and others are presented in the Discussion Chapter of this thesis (Chapter 6). It was therefore essential to identify emerging questions that would enable me to find out more about the perceptions and experiences of head teachers and teachers, concerning the application of SPIS evaluation in their schools. Figure 4.3, below, illustrates the interpretive paradigm and research design.
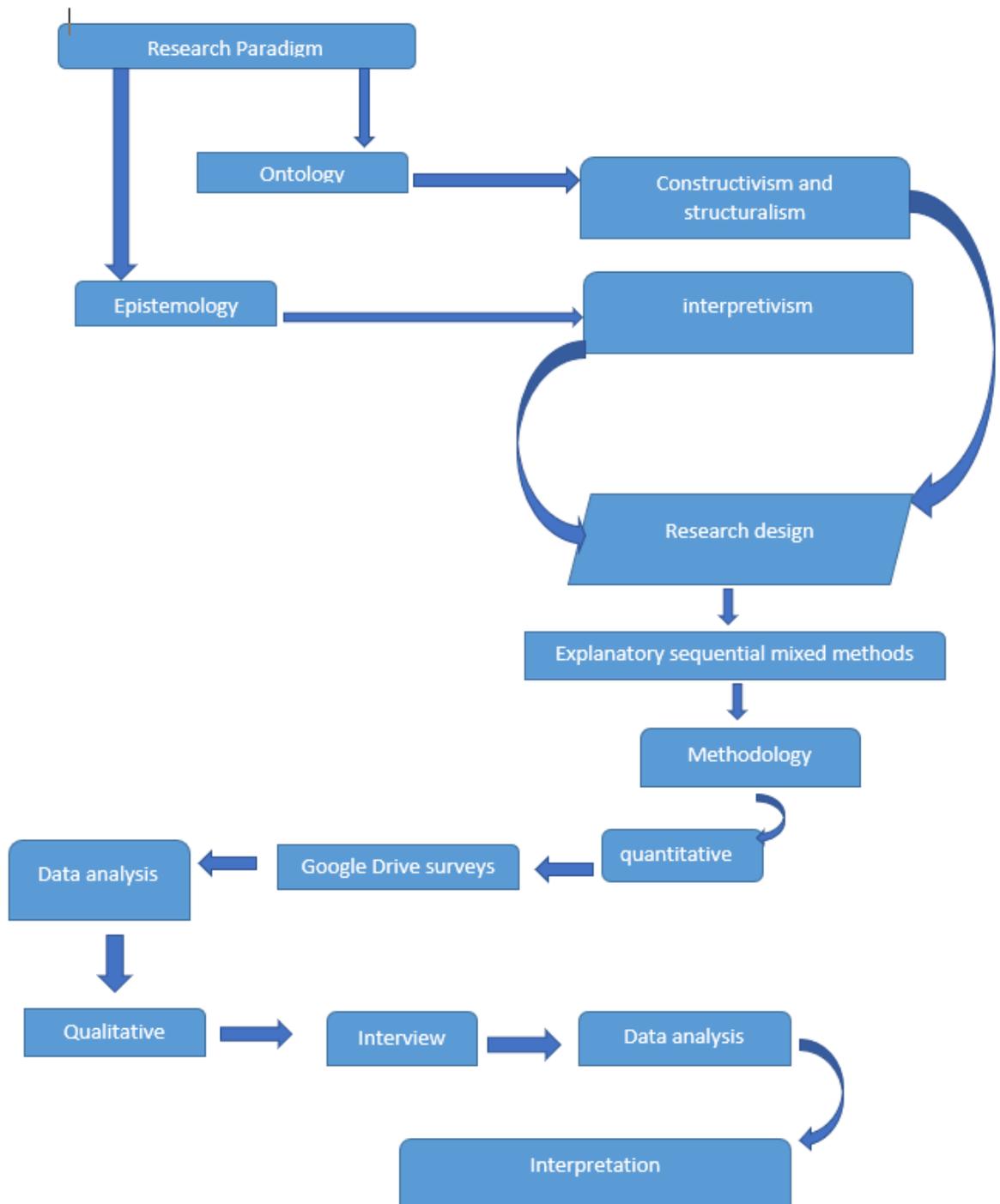
Figure 4.3 Research overview of the interpretive paradigm and research design

4.5.1.1 The Case Study Approach

A small multi-site case study approach was adopted in this study. There were several reasons for this; firstly, according to Basit (2010), a case study is essential for gaining a full understanding of situations or individuals, processes and interactional dynamics in research. Newby (2010) argues that a case study can be used to analyse in detail, individual circumstances

or events that are selected because they are typical, unusual, problematic, or working well. This means that the case study was the most appropriate approach for this research, especially because it would focus on the SPIS process and its influence on head teachers and teachers. In addition, a case study approach can help the researcher to develop theories from one case, as a means of better understanding another (Basit, 2010). Therefore, it was anticipated that this approach would shed light on other evaluation programmes, as a result of understanding SPIS. Furthermore, a case study approach benefits from the use of various methods, such as statistics, questionnaires and interviews (Bryman, 2016). Thus, this approach would enable me to collect data to answer the two questions arising in this study, for which different data collection methods would be needed. Moreover, "case study research can be split into exploratory, descriptive and explanatory methods" (Yin, 2009, pp.5-6), in an attempt to deal with the "'who, what, where and why' research questions" (Cohen et al., 2011). The current study will be performed using an explanatory method. There are clear advantages of using a case study approach, but there are also a number of disadvantages. Cohen et al. (2011) make the following points in this regard:

1- Case study data are strong in reality but difficult to organise.

2- Case studies allow generalisation, either about an instance or from an instance to a class.

3- Case studies recognise the complexity and embeddedness of social truths.

In addition, Baxter and Jack (2008) claim that case studies enable the researcher to gain considerable insights into a case by gathering data from various sources and converging them, so that the case can be illuminated. There are four types of case study, as follows: a single-case (holistic), single-case (embedded), multiple case (holistic), and multiple (embedded) case study design (Yin, 2013, p.50). However, a multi-site case study approach was adopted in this current study, which is an embedded multiple-case study. The two terms (multi-site and multiple case study) are often used interchangeably in research (Audet & d'Amboise, 2001). A multi-site case study refers to the fact that there are multiple cases and each case includes multiple states (Louis, Lawrence, & Keith, 2007). In this study, the researcher collected data from three schools and in each of these schools, one head teacher and three teachers were interviewed. This decision was based on my desire to gain a more profound, varied and detailed understanding of the impact of SPIS evaluation. Moreover, I selected schools from three different areas of the

city of Jeddah, in order to take advantage of the potential differences between them in terms of school environments and experiences.

4.6 Justification of the Selected Methodology

This study is embedded in an interpretivist paradigm, because one of its main objectives was to give head teachers and teachers the opportunity to express their views and narrate their experiences of the SPIS evaluation system. Moreover, researchers who apply an interpretive paradigm do not aim to produce generalisable results (Basit, 2010). This is also true of the current study, although the research paradigm adopted consists of both qualitative and quantitative components (Bryman, 2016). In this paradigm, there are no barriers between the researcher and participants, which facilitates the extraction of information by the researcher from the participants. Moreover, if the researcher has had similar experiences to the participants, it can be very helpful for gaining a deeper understanding and more explanation of the findings. As the researcher, I was fortunate to have accumulated seven years' teaching experience, which meant that I had credibility with education professionals. Moreover, I was able to encourage the participants to give their opinions, talking freely about their experience of SPIS evaluation. The first aim of this investigation was to explore the influence of the SPIS process, which required the use of a questionnaire. Its second aim was to look for in-depth answers to questions concerning the identification of the head teachers' and teachers' perceptions and experiences of the SPIS process, for which a qualitative method was implemented in the form of interviews. Therefore, a mix of different methods was adopted to collect the study data, in order to meet the two main research aims. These mixed methods were selected, because neither a qualitative nor a quantitative approach alone could have provided satisfactory answers to the research questions; separately, they would have been inadequate for a thorough exploration of the attitudes of head teachers and teachers to the SPIS process. Moreover, this research used an explanatory sequential design, because it began with an initial analysis of the data to gather general information from the quantitative data; identifying important points to be discussed in depth in the interviews, where the qualitative data would be collected. This method can be used to obtain notable results, such as those found in this study; for example, the teachers were subjected to more than one type of evaluation system and this raised another question over whether they could differentiate between these types of SPE.

4.7 Sampling

According to Cohen et al. (2013), the sampling technique is one of the most important factors for achieving quality in research. However, the sampling strategy will depend on the type of sampling used by a researcher. There are two types of sample: random probability and non-random (Newby, 2014). The difference between them is that the most reliable representation of an entire population can be achieved using a random technique, whereas non-random sampling relies on the researcher's judgment or an accident. Therefore, it cannot be used to make generalisations about a whole population (Walliman, 2017). However, both random and non-random sampling may be divided into further types. Firstly, there are five types of random non-probability sampling: quota, snowball, modal instant, heterogeneity, and purposive sampling (Etikan, Musa, & Alkassim, 2016; Etikan & Bala, 2017). However, this type of sampling has both advantages and disadvantages; for example, in quota sampling, where the participants are selected according to characteristics that are determined by the researcher, such as age, sex or colour (Acharya, Prakash, Saxena, & Nigam, 2013). One of the main advantages of this is that it affords enough statistical strength to distinguish cluster variations (Yang & Banamah, 2014), but its main disadvantage relates to the fact that there may be an extension of the class or group (Sharma, 2017). For example, women may be divided into working women, non-working women, married women and unmarried women. This is where financial cost, time and effort is demanded from the researcher to inventory the participants. Additionally, according to Etikan, Musa and Alkassim (2016), in purposive sampling, the researcher selects the participants according to their knowledge or experience of the research topic, which then determines the qualities of the participants in relation to their knowledge and experience. Although this type of sampling can be beneficial for a study, if it looks at specific experience with special information (Passmore & Baker, 2005) (for example, students achieving full marks in an important test), it can also reveal the researcher's bias (Sharma, 2017). Nevertheless, if the researcher is keen to avoid this and enriches his or her knowledge though extensive reading, researcher bias can be avoided.

Random probability sampling is likewise further divided into different types: systematic random, stratified random, cluster, multiphase and multistage sampling (Acharya et al., 2013). However, each of these has weaknesses and strengths. For instance, stratified random sampling involves participants being divided into groups according to certain qualities such as age, gender or education. Participants are then randomly selected from each group (Acharya et al.,

2013). Although this helps prevent any bias when selecting participants, it is not useful if the main group cannot be divided into sub-groups or are incorrectly divided. In any case, each type needs to be examined by the researcher, so that he or she can select appropriate subjects for the research questions and objectives. Moreover, the researcher should be able to avoid negative aspects and focus as far as possible on the positive ones, so that a sample is selected that can provide data from which key findings will be drawn within the corresponding field.

Irrespective of the sampling techniques used by the researcher, however, there are four important factors to be considered:

1. Sample size
2. Representativeness
3. Access
4. Sampling strategy (Cohen et al., 2013).

4.7.1 Details of the Research Sample

Before explaining sample size, representativeness and other factors of importance in the present study, there is certain relevant information about the Saudi education system that should be clarified. Saudi education is largely centralised (Al Essa, 2009), which means that all schools, regardless of their location or whether they are public or private, are regulated by the Ministry of Education, which supervises them, designs the curriculum, specifies activities, assigns and manages testing, and conducts evaluation (see Chapter 1, specifically section 1.2). This made it easier to select a representative sample of secondary schools in Jeddah, which would in turn be representative of secondary schools in KSA's big cities, which differ greatly from schools in rural or remote areas in terms of their environment, as well as the strength of law enforcement. In Jeddah, which is the main focus of this study, there are 107 girls' secondary schools and 107 head teachers, while the total number of teachers is 3219 (Ministry of Education, 2019). The schools are distributed across four Districts, represented by the North, South, Central and West Offices, corresponding to the location of the schools. Non-probability sampling was deployed, specifically purposive sampling, because the sample needed to be able to provide data for a specific purpose. In this case, the sample comprised head teachers and teachers from a secondary school in Jeddah, which had been evaluated by SPIS on at least two occasions. According to Bryman (2016), non-probability sampling is the most appropriate method, if the research questions require a specific target group to be sampled. In the

quantitative phase of this study, a large sample of 64 head teachers and 109 teachers was used (see Table 4.1), drawn from girls' secondary schools in Jeddah, where the students were aged 16 to 18 years. According to Kumar (2019), quantitative methods enable the collection of data related to the impact of government policy, which affects large numbers of stakeholders and therefore requires large samples.

In the qualitative phase of this study, the sample included three schools, from which three head teachers and nine teachers were selected (see Table 4.1). The schools varied in size, with one consisting of fewer than 200 students (small school), one accommodating 200-400 students (medium-sized school), and one large school with a capacity of 400-600 students (see Table 4.4). Since this study discusses the influence of the SPIS process on secondary schools in Jeddah, a large sample was appropriate.

Table 4.1: Sample size for the qualitative and quantitative studies

	Qualitative study (3 schools)	Quantitative study (112 schools)
Number of head teachers	3	64
Number of teachers	9	109
Total sample	12	173

Table 4.2: Head teachers' qualifications and experience (survey of head teachers)

Qualification	Diploma	Degree	Master's	-
	1	60	3	-
Practical experience in general	Less than 1 year	1-2 years	3-6 years	7 or more years
	0	0	0	64
Practical experience at current school	Less than 1 year	1-2 years	3-6 years	7 or more years
	0	1	10	53

Table 4.3: Teachers' qualifications and experience (survey of teachers).

Qualification	Diploma	Degree	Master's	-
	12	90	7	-
Practical experience in general	Less than 1 year	1-2 years	3-6 years	7 or more years
	0	0	11	98
Practical experience at current school	Less than 1 year	1-2 years	3-6 years	7 or more years
	8	27	37	37

Table 4.4: Selection of schools

School	Size	Location
School 1	Small with fewer than 200 students	Central Jeddah
School 2	Medium-sized with fewer than 400 students	Southern Jeddah
School 3	Large with fewer than 600 students	Northern Jeddah

Table 4.5: Head teachers' qualifications and names, size and location of school (qualitative interviews with head teachers)

	School 1	School 2	School 3
Names	Sarah	Hind	Leen
School size	Small	Medium	Large
Location	Central Jeddah	Southern	Northern
Qualification	Degree	Degree	Degree

Table 4.6: Teachers' qualifications and names, size and location of schools (qualitative interviews with teachers)

School	School 1			School 2			School 3		
Size	Small			Medium			Large		
Location	Central Jeddah			Southern			Northern		
Names	Noha	Muna	Meachael	Lela	Salma	Souad	Abeer	Marram	Amal
Qualification	Degree	Degree	Degree	Degree	Master's	Degree	Degree	Master's	Degree

The rationale for selecting secondary schools in Jeddah was due to my experience of teaching at a secondary school there for seven years, gaining solid experience and knowledge of the system. Jeddah is also my home city, and it is Saudi Arabia's second largest city. In addition, girls' schools were selected in this study, because I am a female researcher, and the rules in Saudi Arabia do not allow women to enter boys' schools, which would have made it very difficult to collect such data.

All these schools had been evaluated by the SPIS for at least three years and had received their reports. According to the report from the Saudi Ministry of Education, Jeddah's schools had been visited twice for inspection, once in 2015 and again in 2016. Therefore, the respondents were expected to be able to provide rich information.

4.8 Data Collection

4.8.1 The Questionnaire

The questionnaire is one of the most important data collection tools in social research (Sekaran & Bougie, 2016). It is generally used to study people's beliefs, views and perceptions (Hinkle, Wiersma, & Jurs, 2003). Of particular relevance to this study, Gall, Borg and Gall (1996) claim that questionnaires can be used to explore diverse educational phenomena, such as the perceptions of teachers and head teachers concerning proposals for school reform. Moreover, according to Robson (2002), questionnaires can be useful instruments for collecting potentially generalisable data. Therefore, this study will use a questionnaire as the first method of data collection.

In particular, Creswell et al. (2011) and Bryman (2016) explain that a research questionnaire may consist of two main types of question: structured and open-ended. According to Cohen et al. (2013), these types of question render a questionnaire quick to complete and straightforward

to transfer to a computer for analysis. Both types of question were included in the questionnaire for this current study.

There were a considerable number of advantages gained from using a questionnaire, one being the ease with which a large amount of general information could be collected within a short period of time (Robson, 2002; Basit, 2010). However, there are flaws in the questionnaire method; the most important of which being that it can only be completed by people who can read and write, and there is the risk that some participants may leave questions unanswered (Denscombe, 2007). However, in this study, the participants were all head teachers or teachers, who could evidently read and write. I also used Google Drive surveys, which required each participant to answer the department's questions before moving on to others. Therefore, each questionnaire had to be completed in full before it could be returned to me.

The questionnaire implemented in this study consisted of four parts:

1- The introduction to the questionnaire, illustrating the goals and significance of the study, the rights of the participants, and instructions on how to answer the questions;

2- Information related to the participants' names, schools, experience in education and length of employment;

3- Schedule of survey questions divided into four sections. The first of these related to school monitoring and comprised four points: the extent to which head teachers and teachers were monitored by key stakeholders; the head teachers' and teachers' awareness of SPIS monitoring; the head teachers' and teachers' views of awareness of SPIS monitoring techniques for performance evaluation, and the head teachers' and teachers' views of the use and awareness of SPIS KPIs. The second question related to head teachers' and teachers' perceptions of the influence of SPIS on their stress levels, workload and morale, while the final question investigated school improvement.

4- Finally, one question was aimed at determining whether the participants were willing to consent to a follow-up with a qualitative interview (see Appendix II).

In this study, the researcher obtained permission from the Ministry of Education to email this survey to 107 secondary schools in Jeddah, and it was expected that most of these would be returned. The questionnaires were used to collect quantitative data and were administered via links to two Google Drive surveys, one for all teachers in the selected population and one for

93

81 head teachers. These Google Drive surveys required the participants to complete each section of the questionnaire before moving on to the next, which meant that all the questionnaire items would be completed. The questionnaires were then sent to four directors of Education Offices, who distributed them via a WhatsApp group to 107 head teachers. In turn, these head teachers each sent teachers in their schools a link to the questionnaires. I received 109 responses from teachers and 64 from head teachers. It was also important that all data were checked to ensure that there were no obvious flaws (Bryman, 2016), such as unanswered sections, which could potentially affect the results.

4.8.2 The Interviews

According to Mallick and Verma (2005), a questionnaire can provide substantial information, but an interview can provide in-depth data, with many important details that cannot be collected via a questionnaire. Thus, the second method applied in this study was an interview (see Appendix III), which is a qualitative method. The very nature of an interview clearly assumes human interaction, which is essential for knowledge production (Cohen et al., 2013). More specifically, an interview can obtain a description of the interviewees' inner world, with respect to interpreting the meaning of the phenomenon described (Kvale, 2008). Bryman (2016) considers interviews to be the most widely used method in qualitative research.

There are three types of interview: structured, semi-structured and unstructured (McVeigh, 2016). The difference between these types is defined by the questions: for example, in the structured type, there are specific questions. In contrast, unstructured questions are not specific, but simply appear during the course of the conversation. In this study, interviews were used to learn about the experiences and views of head teachers and teachers, with regard to the influence of SPIS on teachers and school performance. These interviews needed to be flexible enough to give the interviewees an opportunity to express their opinion of the SPIS process. For this reason, a semi-structured interview guide was designed, with questions that covered the main themes to be covered during the interviews, rather than specific questions. According to Denscombe (2014), semi-structured interviews are identified as the most appropriate method of gathering research participants' "feelings, emotions and experiences". Carter, Henderson (2005) add that in semi-structured interviews, a degree of flexibility will enable the interviews to be guided by the researcher's interest. However, they also allow interviewees to raise any other issues that they might consider relevant or important. Therefore, this type of interview

was selected for the current study, as a means of discovering the main themes and the interviewees' experiences in response to the research questions. However, there are certain disadvantages to this approach, highlighted by Kajornboon (2005), as some important data may be missed if the interviewer fails to elicit it with appropriate prompts. Another problem can be the interviewer's inexperience or lack of curiosity.

In this study, a series of 12 individual interviews were conducted with head teachers and teachers from girls' secondary schools in Jeddah, all of whom were sought as volunteers in this study. The interviews were guided by a semi-structured list of questions (the interview guide). This guide and a list of themes were drawn up in a way that would allow the interviews to be flexible and conversational (Denscombe, 2014). They typically lasted between 30 and 45 minutes and were carried out at the schools themselves. With the participants' informed consent, they were audio-recorded, with notes also being taken.

4.9 Public Documents

According to Bowen (2009), there are a considerable number of benefits that can be gained from using documents as resources. The three most important of these are as follows:

1- They provide data on the context within which the research participants operate.

2- They contain information that can suggest some questions to be asked and situations that need to be observed as part of the research.

3- They provide supplementary research data.

SPIS evaluation includes important documents such as the Organizational Guide for Performance Evaluation (SPIS, 2017), the Civil Service Instructions for Teachers and Head Teachers (Ministry of Civil Service, 2019), and the Teacher's Charter, which the Ministry of Education issues as part of its mission (Ministry of Education, 2019), under which teachers must act accordingly. All these documents were reviewed, in order to better understand the system, and to be able to build the questionnaire and interview items. These documents also helped interpret the research results, but they were not used as tools for the collection or analysis of data.

4.10 Data Analysis

After the data collection process, the researcher must make sense of the participants' responses and analyse the data, so that the research questions can be answered. In this sense, Creswell (2009) claims that the aim of data analysis is to find answers to research questions and avoid any responses that are not related to any of those questions. This can be divided into meaningful segments for easier interpretation and clarification (Major & Savin-Baden, 2010). According to Onwuegbuzie and Teddlie (2003), there are various ways in which data can be analysed in a mixed methods approach. One of these begins with an analysis of quantitative data, which can take place in clusters, based on exploratory aspects and using descriptive statistics, as in the application of SPSS software. Conversely, qualitative data can be analysed using exploratory thematic analysis. Because this research adopted sequential explanatory mixed methods, the data were first analysed sequentially and then subsequently integrated. After analysing the findings from the quantitative and qualitative methods, I divided them into two groups and compared them (Tashakkori & Teddlie, 2003). The data analysis procedures for the current study are discussed in the next subsection.

4.10.1 Analysing the Survey Data

Since Google Drive surveys were used to collect the data in this current study, the results needed to be downloaded as an Excel file. However, no gaps were found due to missing answers, because Google Drive surveys do not allow participants to move on to the next section before finishing the previous one. I subsequently entered the data manually into the SPSS software. SPSS was run to complete the data coding and computing. After cleaning up the data, the analysis was conducted in the following sequence: provisional through descriptive to inferential (Creswell & Clark, 2007). The items in each section of the questionnaire were given specific codes to differentiate between them. For instance, Section One on monitoring was coded as 'M' and each group of questions in this section had a corresponding number; for example, M1, M2, etc. Therefore, the items were coded as M1A, M1B and so forth.

Next, the Cronbach's alpha was generated for the questionnaire's reliability, with tables being prepared for each group of questions belonging to the same category. Finally, important results were selected to be presented and explained. Concerning the open questions in the questionnaire, these were analysed thematically, in the same way as the qualitative interview data (see subsection 4.8.2). Moreover, the negatively worded items were reversed, so that they

matched the positive scheme of the other scored items. These negative questionnaire items are presented in Table 4.7 below:

Table 4.7: Reversal of negatively worded items

Negative Items	Reversed Items
1- I do not feel pressure due to workload from SPIS evaluation	1. Strongly disagree → Strongly agree 2. Disagree → Agree 3. Undecided → Undecided 4. Agree → Disagree 5. Strongly agree → Strongly disagree

4.10.2 Analysing the Interview Data

According to Huberman and Miles (2002), there are three steps involved in analysing qualitative interview data: data reduction, data display and drawing a conclusion. To follow these steps, the researcher must begin by recording and transcribing the interviews (see Appendix III). The researcher must then send a transcript to each participant for final verification of its content. This is a dependability check, which will give the participants a chance to add any comments to clarify or expand on the opinions that they have expressed in their interviews. To become familiarised with the data and to code it manually, the researcher must read through it carefully to gain a holistic overview of the main themes deliberated on by the interviewees. This comprehensive reading will allow similar statements and ideas to be classified into main themes (Cicourel, 1982).

The coding process and thematic analysis were performed immediately after the data collection (Bryman, 2016). Thematic analysis is the easiest and clearest method of analysing data, as it allows for a deep interpretation of the interview data (Braun & Clarke, 2006). In the present case, colours were used to code the emerging themes; for example, stress was coded as red and workload was coded as green (see Appendix XI). This technique allowed me to find the relationships between themes and to recognise each theme in the interview (Bryman, 2016). Moreover, a randomly selected pseudonym was assigned to each participant, accompanied by their school number to assist in the discussion, when referring back to their comments and to compare the responses given by different participants. This was also done to protect the participants' identities, in conformity with research ethics. I subsequently divided the themes

into sections according to the research questions. Later, all the themes were classified into sub-themes and presented in sections, in accordance with the research questions (see Appendix XII).

4.11 Quality Criteria

The quality of the instruments used by a researcher is the core of a study's reliability, validity (Kumar, 2019), trustworthiness, credibility and transferability (Appleton,1995). Therefore, the following subsection will discuss both reliability, validity, trustworthiness, credibility and transferability in light of this current research.

4.11.1 Reliability

According to Bryman, (2016) research is considered reliable, when the results of a study are repeatable. There are numerous factors that can affect the reliability of a study (Kumar, 2019).

1- Ambiguous words in questions;

2- Physical setting, such as the time spent in an interview;

3- The respondent's mood when asked a question or whilst giving an answer;

4- The interviewer's mood;

5- Nature of the interaction, and

6- The regression effect of an instrument (a statistical measurement used to determine the strength of a relationship between one dependent variable with multiple independent variables).

Therefore, a pilot study was conducted in this research (see section 4.12), with attention to these points; avoiding any words that might be ambiguous and adding explanation where required. Moreover, I conducted several interviews to test the necessary timeframe required for each interviewee.

According to Cohen et al. (2013), if similar results are derived from a repeated study on a similar group of respondents in a comparable context, then the research may be considered reliable. Mertler and Charles (2005) list three steps for achieving reliability:

1- Verifying different sources of qualitative data to ensure that the collected data are consistent.

2- Planning the procedures to obtain the data carefully and thinking about the trustworthiness of the informants.

3- Applying an internal critique (for example, comparing what one respondent says with what is said by other respondents).

However, there are a number of potential weaknesses that can face researchers, such as cultural misunderstandings. For example, in this current case, the questionnaire was adopted from Scanlon (1999). Therefore, it was designed for use in a Western context, namely the UK. As a result, some items were omitted, because they were considered as personal questions, relating to the privacy of the principal or teacher in Arab culture. Moreover, some questions affected the Cronbach's alpha results, which are important for achieving reliability in quantitative data, particularly with regard to the consistency of the questions. Moreover, some of the interview responses were conflicting, especially those referring to when the SPIS evaluation would take place, as discussed in the findings (see Chapter 5) and Discussion chapter (see Chapter 6).

Additionally, before finalising the questionnaire for implementation, I took care to check its validity and reliability and to translate it into Arabic myself. This translation was then proof-read by an expert, who was familiar with academic language. In addition, it was reviewed by a PhD student who is a native Arabic speaker to ensure that the translation met academic standards. Difficulties that emerged in relation to ambiguous language in the questionnaire and to the accuracy of the words selected in the translation, especially with regard to the school league tables or table of school performance, were subsequently dealt with (see Appendix II).

4.11.2 Validity

Kumar (2019) defines validity as "the ability of an instrument to measure what it is designed to measure" (p.178). Moreover, it is defined as "the appropriateness of the interpretations, inferences and actions that we make based on test scores" (Johnson & Christensen, 2019, p.140). For Bryman, validity is concerned with the integrity of the conclusions that are generated from a piece of research. However, validity differs according to the methods used by the researcher. If quantitative methods are used, three different types of validity may be relevant:

1. Face and content validity. This type of validity means that there is a logical link between the research questions and the research objectives, and that the items and questions that are formulated in the study cover the full range of attitudes being measured.

2. Concurrent and predictive validity. This type is used when the researcher tries to develop his/her instruments by comparing them with other assessments.

3. Construct validity. This type is based on statistical procedures (Bryman, 2016; Kumar, 2019), where Yin (2013) mentions three tactics to be used in case studies:

   a) Using various sources of evidence
   b) Establishing a chain of evidence
   c) Reviewing draft case study reports by key informants.

In contrast, the validity of qualitative research is dependent on different criteria (Bryman, 2016; Kumar, 2019), namely:

1- Credibility which parallels internal validity.

2- Transferability which parallels external validity.

To apply these criteria for internal validity, Yin (2013) recommends the matching of patterns, the building of explanations, rival explanations being addressed and the use of logical models. For external validity, he recommends applying theory in single case studies and replication logic in multiple case studies.

4.11.3 Trustworthiness, Credibility and Transferability

As discussed above (subsection 4.11.2), the differences between qualitative and quantitative research raise implications for tests of validity. Cohen et al. (2013) claim that both qualitative and quantitative methods can address internal as well as external validity, although there are limitations when methods are applied in qualitative research. To address this problem, Johnson and Christensen (2019) propose solutions for enhancing the validity of qualitative studies. Their strategies include extended fieldwork; the researcher acting as a detective, and the use of low inference descriptors, triangulation, participant feedback, peer review, external audit, negative case sampling, and pattern matching. However, Corbin and Strauss (2008) reject the use of validity tests in qualitative research; favouring credibility instead, because confidence in the

steps and results of qualitative research not only relates to validity, but is also deep, complex and requires accuracy (Flick, 2009). Other commentators on this topic point out that examining theories, results and methods to ensure their authenticity should depend on the philosophical foundations adopted by the researcher (Carper, 1978). According to Hammersley (2007), the use of reliability tests to demonstrate the reliability of qualitative research can lead to conflicting and ambiguous assumptions, because their philosophical assumptions are different. While this study involved mixed methods, it was embedded within an interpretivist paradigm. Therefore, a more pragmatic view was adopted toward its research philosophy, compared to Hammersly (2008). Nevertheless, as Jasper (1994) and Appleton (1995) claim, one cannot ignore the criticism that is directed at qualitative researchers who rely on approaches related to validity and reliability in quantitative research.

In response to the criticisms and debate surrounding the use of criteria traditionally used in quantitative studies, new alternative terms have subsequently emerged in qualitative research (Denzin & Lincoln, 2018). Notably, according to Bryman (2016), credibility, transferability and transparency may be applied to check qualitative research results, rather than validity. As a result, this research has sought to meet the criteria of credibility, transferability and transparency in qualitative research.

Credibility refers to whether the results of qualitative research are credible or believable, whereas trustworthiness is a term that is tied to credibility; it is often used as an alternative to 'validity' (Lincoln & Guba,1985). In the current study, the pilot study helped to improve the internal validity of this research. Testing out the questions to find out if the respondents understood them and to assess if the questions helped to answer the main research question was vital. In addition, trustworthiness began early in the research process, since it involved building trust with the participants (Guercini, Raich, Müller, & Abfalter, 2014). In this study, the steps taken toward this goal began by encouraging the participants to feel part of the data collection process. It was achieved by developing a collaborative relationship, based on the researcher and participants' joint interest in meeting the research aims (Creswell & Creswell, 2017). Having previously worked as a teacher for several years and subsequently as a writer on education topics, my work would have been familiar to most teachers in Jeddah. Therefore, in the early stages of building trustworthiness and credibility in this study, I referred to my own experience and activities in the field when communicating with the participants. This established an

emotional connection and common ground with them, so that they were encouraged to speak freely, thereby contributing to the credibility and trustworthiness of the data.

Additionally, trustworthiness can be achieved through transparency, namely providing information about the research tools, techniques, and purpose of the research (Goldberg & Allen, 2015; Moon, Brewer, Januchowski-Hartley, Adams, & Blackman, 2016). Steps taken in this regard involved talking to the participants about the goal of contributing to education research in Saudi Arabia through this study; in particular, by enriching it with information about evaluation in education. Further steps in this regard included providing all participants with information letters and consent forms, which explained the purpose of the research and how the data would be used. As part of the ethical requirements of the research, this provided the participants with transparent information about the study. Thus, they all had an opportunity to ask questions or seek clarification, if necessary. As a result, they collaborated willingly and productively, because of their joint interest in supporting Saudi education and the goals of this research. Hence, there was a great deal of enthusiasm amongst them.

Throughout, steps were taken to avoid bias. This can be seen in the questionnaire and the interview questions, since both multiple choice and open questions were included in the questionnaire to give participants a greater voice in the data, while great care was taken to avoid leading the participants towards specific views in the interviews. Other steps involved prompting participants at several points in the interviews to speak freely on their own views and experiences.

Meanwhile, transferability in qualitative research involves judgements on whether it makes sense to transfer the results of a study to another context (Lincoln & Guba, 1985). It is problematic and the results of this study may only be transferable in other Arabic contexts to some degree. Still, steps taken to enable the transferability of this research involved providing transparent information about the research context, methods and procedures, so that others could judge whether the results were transferable to a different context. Other steps taken were related to the transcription and translation of interview data. The interviews were recorded and then transcribed verbatim in Arabic, before being translated into English. As a result, it would be possible to refer to the original versions for clarification, if necessary, in future. Appleton (1995) suggests that an expert colleague or linguist should verify the translation of primary data as part of the collection procedure, in order to achieve credibility and trustworthiness, and to enable transferability. Therefore, a sample of the interviews was checked by an expert

colleague, who is a native Arabic speaker and a PhD student in the field of linguistics. The English translation was also checked and compared for accuracy by the above-mentioned bilingual Arabic/English speaker (see Appendices VII-X).

4.12 Pilot Study

A pilot study is an important step in educational research, and its aim is to achieve validity and reliability in a study (Basit, 2010). Thus, the aim of pilot testing is to clarify the instructions, check for any ambiguity or unclear questions, and measure the time it will take to complete the questionnaire.

There are numerous reasons to support the importance of pilot studies, but they must correspond to the methods used by the researcher. If, for example, a questionnaire is used, it should be checked to determine whether the questions are all easy to understand, with none that might be misunderstood (Basit, 2010; Chen et al., 2013). In this current research, 20 questionnaires were distributed in a pilot study. Once the data were collected, three head teachers and two teachers were interviewed via Skype, a free audio-visual communication platform (Chen et al., 2013). The collected data were then analysed. After this, feedback was obtained, so that the validity and reliability of the research could be achieved by refining the data collection instruments. In this feedback, certain points were raised, in response to which I evaluated and amended the relevant questions (Yin, 2013). As mentioned above, the questionnaire was taken from an existing study (Scanlon, 1999) and then adapted to fit the context of Saudi education. It was also presented to the relevant supervisors to ensure that it had appropriate consistency and was of a suitable length. This questionnaire was subsequently submitted to the Jeddah Education Department, so that they could grant permission for it to be administered to teachers and head teachers in secondary schools in Jeddah. However, the Jeddah Education Department stipulated the condition of removing several questions relating to the anger and negative behaviour or feelings experienced by head teachers. This especially referred to the second question of Section 3 in the questionnaire, which listed seven possible feelings amongst head teachers: 'Irritated', 'Angry', 'Helpless', 'Anxious', 'Depressed', 'Unable to concentrate', and 'Overtired'. I subsequently excluded this question from the head teacher's questionnaire.

All interview questions were also submitted to the Ministry of Education for approval, who subsequently granted permission to start the data collection. This meant that a pilot study could be conducted, using 20 questionnaires: 10 distributed to head teachers and 10 to teachers. The

clarity of the language was verified, with some incomprehensible words being substituted for clearer wording, which could be better understood by the participants. An example of this was the replacement of the option, 'Stressful' with 'Moderately stressful'. The data were then analysed, and reliability was verified; the Cronbach's alpha was tested using SPSS, generating a result of 0.528. Therefore, some questions were deleted to improve this score. These questions are displayed in the following Table. The result then increased to 0.79. According to Berthoud (2000), this value is acceptable.

Table 4.8: Items deleted from the questionnaire

The benefits of the SPIS evaluation outweigh the negative aspects.
Overall, I am satisfied with the clarity of the SPIS evaluation report.
The oral feedback and the written report from PEEC were consistent.
The working environment at school
In your routine when dealing with parents, do they frequently refer to the performance tables?
Is your school's position in the league table very important to you personally?

Following this pilot study, all changes deemed necessary were explained in detail. However, the questionnaires and interview schedules were piloted, so that the pilot study sample could give their feedback on the clarity of the questions and questionnaire items. Secondly, the pilot study gave me the chance to practice my communication skills and interviewing ability. Therefore, the instruments were piloted according to the same approach that was adopted for the data collection. The participants gave their feedback to improve the clarity of the instruments and assist with managing the interview timings.

4.13 Ethical Issues

There are many ethical responsibilities to be considered in social research worldwide. Several documents are instrumental to this process, including guidelines issued by the British Educational Research Association (BERA). And the University of Reading's Code of Ethics (see Appendix V).

According to the above-mentioned documents, educational research must be conducted in an ethical manner. Therefore, it is vital for all researchers to understand what is meant by ethical considerations. Bryman (2016) claims that these revolve around issues such as how the research

participants are treated and the activities in which they should or should not engage. Moreover, the researcher should undertake training in research ethics; for example, through an online course (which I took in 2017). However, according to Robson (2011), data collected via a questionnaire can still have an impact on participants. Therefore, the researcher needs to be careful when formulating the questionnaire items, in order to avoid a negative impact. The following sections details my response, as the researcher, to primary ethical areas.

4.13.1 Informed Consent

According to Cohen et al. (2013) informed consent means that the participants are fully aware of the aims, procedures and measures of a research project and that they understand its potential impact. This study examines mature people, who are head teachers and teachers in secondary schools in Jeddah. I initiated this process by sending a letter containing information about the study and asking the Ministry of Education for permission to conduct the research. In this letter, I also asked for permission to conduct the study and stated that all information gathered would remain confidential. In addition, I ensured that all the participants gave their written consent, confirming that they understood the process, and explaining why their participation was necessary (see subsection 4.13). I translated the questionnaire and interview questions from English into Arabic, and these were then reviewed by a professional translator.

4.13.2 Confidentiality

Any information provided by the participants in this study would remain confidential and only be seen by myself, the researcher, and my supervisors. None of the participants (i.e. the head teachers and teachers) or the school would be identifiable in any published report resulting from this study. Moreover, no information about individuals would be shared with the school, and the collected data would be held in strictest confidence, with no real names used in this study or in any subsequent publications. The records of this study would be kept private, with no identifiers linking the participants or schools to the study being included in any sort of report that might be published. The participants were assigned pseudonyms and numbers and referred to by these in all records. Finally, the research records would be securely stored in a locked filing cabinet and on a password-protected computer, with only myself and my supervisors having access to them. In line with the University's policy on the management of research data,

anonymised data were therefore gathered. The results of this study would be presented at national and international conferences, and in written reports and articles.

As the researcher. I was aware of any potential influence of my position as an employee of the Ministry of Education and a writer for a Saudi newspaper, on the participants' opinions and their desire to speak freely. Therefore, I was keen to affirm my commitment to refraining from publishing any data or evidence referring to the participants' identities. Moreover, none of this information would be used for reasons other than academic research. I emphasised my neutrality and explained my research objectives and desire to deliver the participants' voice. I also stressed that I was keen to benefit from their experience of applying the Ministry's recommendations for improving the SPE programme. Finally, I was keen to remain objective in analysing the data and respecting the participants' views, interpreting them from every angle for the purpose of scientific research. This included monitoring both positive and negative views.

4.14 Limitations and Constraints

This study has a number of limitations, in that it only included government schools in Jeddah. SPIS evaluation was implemented in Jeddah's schools in 2014, 2015 and 2016. As mentioned earlier, Jeddah is the second largest city in Saudi Arabia, located in the western region of the country. It is where I grew up and where I worked for seven years as a teacher, prior to starting my PhD programme. As a Saudi woman, during the data collection, I was not permitted to travel alone between cities, but was required to be accompanied by a male guardian from my family, especially when staying in hotels. This prompted me to choose Jeddah as the source of my sample. Moreover, at the time of the data collection, women were prohibited from entering boys' schools. Therefore, the study sample was limited to female teachers and head teachers in girls' schools.

An explanatory sequential mixed methods research design was applied in this study, even though this approach has been criticised for failing to provide a strong base for scientific generalisation. Therefore, although there is no aim to generalise the study findings, they could form part of a bigger picture and elucidate the perceived impact of SPIS evaluation on secondary schools, thereby promoting an understanding of its impact on all schools in Saudi Arabia. The study's limitations and constraints are discussed in more detail in Chapter 7.

4.15 Summary

This chapter has presented the rationale and an explanation of the methodology adopted for this mixed methods case study, where questionnaires were administered; followed by semi-structured interviews, primarily within a pragmatic paradigm. This was in order to understand from all possible angles, the participants' individual realities and the data collected from them. It also explained the sampling, data collection and analysis strategies adopted, and the process applied to guarantee the quality of the data gathered and the conclusions derived from them. Finally, it set out the key ethical issues underpinning the study, such as informed consent and confidentiality. The findings from the data analysis are presented in the next chapter (Chapter 5).

Chapter 5: Results and Findings

5.1 Introduction

As noted previously (see section 1.1), a key aim of this thesis was to explore head teachers' and teachers' experiences, views and understanding of SPIS school monitoring, and their perceptions of its influence on school performance in Saudi Arabia. In recent years, the issue of school evaluation has grown in importance in Saudi Arabia, in light of concerns that significant spending on education is failing to have the desired effect of improving educational outcomes. Previous work (for example, Ball, 2012) has explored the influence on teachers of school evaluation, using performance indicators. However, prior research has not addressed the topic of school monitoring and evaluation from the perspective of teachers and head teachers in Saudi Arabia. Therefore, little is known about how they experience and perceive the present system. The current study is therefore intended to address this gap and add essential value by giving voice to an under-represented group, namely female teachers and head teachers in Saudi Arabia.

In accordance with the sequential mixed methods design of this study, quantitative data were first collected, using two Google Drive surveys that were administered to teachers and head teachers. The survey results were then analysed in SPSS. In the second stage of the study, qualitative data were collected through semi-structured interviews with head teachers and teachers. This chapter outlines the results of both phases of the study in response to the following research question:

> What are head teachers' and teachers' experience and perceptions of the influence of SPIS on school performance?

This question raises three sub-questions (RSQ):

> RSQ1. What are head teachers' and teachers' experiences and perceptions of the influence of SPIS on school monitoring?

> RSQ2. What are head teachers' and teachers' experiences and perceptions of the influence of SPIS on their stress levels, workload and morale?

> RSQ3. How do head teachers and teachers describe and understand their accountability under SPIS in relation to school improvement?

This chapter begins with a profile of the study participants (see section 5.2), followed by three main sections (see sections 5.3, 5.4 and 5.5), each answering a research question. Every one of these sections is in four parts, as follows: the quantitative data from the teachers' survey; the qualitative findings from the teachers' interviews; the findings from the head teachers' survey, and the qualitative findings from the head teachers' interviews. The main theoretical implications and contributions of this study are discussed in depth in the Discussion chapter (Chapter 6). Supplementary information, such as the interview and survey schedules, can be found in Appendices I-IV.

5.2 Profile of the Schools and Study Participants

Jeddah is the second largest city in Saudi Arabia in terms of its economy, as well as in the number of its public sector schools. According to the latest statistics from the Jeddah Education Office (General Education Administration of Jeddah ,2019), the number of students registered by the Department of Education in Jeddah totalled 462,154 in 2017, studying at 1376 schools in the city. However, this study is specifically dedicated to girls' secondary schools, amounting to 125 establishments, with 3219 teachers and 125 head teachers. Concerning the quantitative data, the questionnaire was distributed electronically to all relevant schools, and all teachers and head teachers (see section 4.8.1). However, just 109 teachers and 64 head teachers participated, drawn from all girls' secondary schools in Jeddah, which came under the Education Offices of the Centre, North, South and East. Most of these head teachers held university degrees, while one had a Diploma, and three had Master's degrees (see Table 5.1). Likewise, most of the teachers also held university degrees: 12 had diplomas and seven had Master's degrees (see Table 5.1). In terms of experience, all the teachers had more than one year's experience. However, without exception, the head teachers had at least seven years' experience, because the Saudi education system requires candidates to accrue many years of professional experience, before attaining the position of head teacher (see Table 5.1).

In terms of qualitative data, I chose three Offices from different parts of Jeddah: the Centre, North and South. The Centre corresponds to the heart of the city, which is largely inhabited by the middle classes. Here, the schools are predominantly old, because the centre is an old part of the city. It is also known as Old Jeddah and its schools are small, because they were built during the early years of Saudi education, when there were fewer students. Therefore, the schools selected from this area were small, with fewer than 200 students in each. Another part of the

city included in this study was North Jeddah, which is a newly developed area. Therefore, most of its schools are modern and large, with over 400 students each. The population in this part of the city tends to be relatively wealthy and have high social status. The third part of the city sampled was the South, which is usually considered to be the hub of most of Jeddah's activity. This includes its schools, because most were built in the 1980s, when Saudi education was still in the process of expanding, and there were fewer students than there are now. All participants who provided qualitative data had high-level academic qualifications, including Master's degrees, and all had previous experience as teachers and head teachers (see Table 5.1).

Table 5.1: The participants' information

Methods	Instrument	Number of Participants	Participants' Qualifications		Practical Experience in General		Practical Experience at Current School	
			Level	**No.**	**Duration**	**No.**	**Duration**	**No.**
Quantitative research	Questionnaire for teachers	109	Diploma	12	Less than 1 year	0	Less than 1 year	8
					1-2 years	0	1-2 years	27
			Degree	90	3-6 years	11	3-6 years	37
			Master's	7	7 or more years	98	7 or more years	37
	Questionnaire for head teachers	64	Diploma	1	Less than 1 year	0	Less than 1 year	0
			Degree	60	1-2 years	0	1-2 years	0
					3-6 years	0	3-6 years	0
			Master's	3	7 or more years	0	7 or more years	0
		School	**Location**	**School Size**	**Interviewees**			
					Name	**Position**		**Qualifications**
Qualitative research	Interviews	School 1	Central Jeddah	Small	Sarah	Head teacher		Bachelor's degree
					Noha	Teacher		Bachelor's degree
					Muna	Teacher		Bachelor's degree
					Meachael	Teacher		Bachelor's degree
		School 2	South Jeddah	Medium	Hind	Head teacher		Bachelor's degree
					Lela	Teacher		Bachelor's degree
					Salma	Teacher		Master's
					Souad	Teacher		Bachelor's degree
		School 3	North Jeddah	Large	Leen	Head teacher		Bachelor's degree
					Abeer	Teacher		Bachelor's degree
					Marram	Teacher		Master's
					Amal	Teacher		Bachelor's degree

5**.3 In-school Monitoring**

5**.3.1 Quantitative Findings: Head Teachers' and Teachers' Experiences of School Monitoring**

To learn about the head teachers' and teachers' experiences of school monitoring and evaluation, it was imperative to begin by assessing their engagement in these processes, particularly their experience of SPIS, a relatively new system. Therefore, the current participants were asked if they had any experience of in-school monitoring by school advisors and SPIS inspectors. The results for the head teachers are set out in Table 5.2.

Table 5**.2: Extent to which head teachers were monitored by key stakeholders**

		Never	1-3 times per year	4 or more times per year
School advisors	n	0	19	45
	%	0	29.7	70.3
SPIS inspectors	n	0	46	18
	%	0	71.9	28.1

Table 5.2 shows that all the participants had been observed by school advisors and SPIS inspectors in the past year. This diversity of evaluation in Saudi education may have developed because the Ministry of Education did not ensure complete elimination of the old system of school supervision – which performed the same tasks as the SPIS – when it adopted the SPIS to evaluate school performance. This was due to weak co-ordination between the Ministry's Departments. It is therefore worth noting that more than a quarter of the participants declared that they were visited four or more times a year by SPIS inspectors, and slightly less than three quarters of the participants reported that they were monitored by school advisors. This raises questions about the nature of the advisors' work and the necessity for them to go into schools so frequently. Overall, it indicated that extensive in-school monitoring took place, with implications for the teachers' and head teachers' workload, morale and wellbeing. It also poses questions over the extent to which these staff had the capacity to engage with the process, and what they thought of its achievement in improving student outcomes. Given the study participants' experience of school monitoring, they were well placed to answer questions on such matters.

Table 5.3 presents important issues related to SPIS monitoring, such as awareness of monitoring techniques, and their appropriateness for SPE. Meanwhile, Table 5.3 presents head teachers' knowledge and experience of SPIS monitoring, specifically their awareness of when SPIS monitoring would take place and the main areas of its focus.

Table 5.3: Head teachers' awareness of SPIS monitoring

		Never	Seldom	Sometimes	Usually	Always
1-Awareness of when SPIS monitoring would take place	n	0	63	1	0	0
	%	0	98.4	1.6	0	0
2-Awareness of what SPIS monitoring would focus on	n	28	10	16	6	4
	%	43.8	15.6	25.0	9.4	6.3

Therefore, it can be seen from the above Table that only one of the 64 head teachers were 'Sometimes' aware of when SPIS monitoring would take place, while the remainder seldom knew. There was also considerable variation in the participants' awareness of the intended focus of SPIS monitoring. The results for these two questions suggest that SPIS had inadequate capacity to identify and determine objectives for schools, and inaccurately informed schools on what or when they would be assessed, thereby negatively affecting their readiness for evaluation. The Organisational Guide includes all the indicators and some of the procedures that the SPIS requires from schools, but this was clearly inadequate for helping schools to understand the system. In addition, the results suggest that the SPIS process could cause misleading evaluation results, as it does not inform schools when monitoring will take place. This could affect schools' ability to prepare for evaluation and hinder them from providing the evaluation team with evidence of their achievements.

Other issues related to the SPIS monitoring process could also affect school evaluation results, including the appropriateness of the monitoring techniques, as shown in Table 5.4.

Table 5.4: Head teachers' views of awareness of the SPIS monitoring techniques used for performance evaluation

		Never	Seldom	Sometimes	Usually	Always
3-Appropriateness of the monitoring techniques	n	6	9	34	10	5
	%	9.4	14.1	53.1	15.6	7.8

The head teachers were asked for their opinions on the appropriateness of the monitoring techniques applied. Most believed that the SPIS techniques were 'Sometimes' appropriate. However, more clarification was required to justify their views on the appropriateness of the monitoring techniques. Therefore, they were asked four questions to ascertain their ability to recall approximate figures from their schools' performance tables, their perceptions of the three most important evaluation tasks expressed in their schools' aims, and the key strengths and weaknesses of their schools' performance. Table 5.5 presents these responses.

Table 5.5: Head teachers' views on the use and awareness of SPIS KPIs

		Never	Seldom	Sometimes	Usually	Always
1-Are you able to recall your school's approximate figures from the performance tables?	n	40	15	3	6	0
	%	62.5	23.4	4.7	9.4	0
2-Are your school's aims expressed in the KPIs?	N	6	14	24	14	6
	%	9.4	21.9	37.5	21.9	9.4
3-Do the KPIs identify the key strengths of your school?	N	3	2	21	12	26
	%	4.7	3.1	32.8	18.8	40.6
4-Do the KPIs help identify the main weaknesses of your school?	N	4	2	32	13	13
	%	6.3	3.1	50.0	20.3	20.3

What is striking in the results displayed in Table 5.5 is the high percentage of head teachers in the study who could not recall their school's figures in the performance league tables. This result suggests that these performance tables, derived from the results of SPIS evaluation, were difficult for the head teachers to recall. A possible explanation for this is that the preparation of these Tables was neither accurate nor clear. Moreover, the head teachers may have had insufficient training in the use of KPIs, or their training may have failed to include guidance on interpreting school monitoring data.

Table 5.5 therefore illustrates that there was some hesitation as to whether the KPI terms expressed the schools' aims. The results indicate that while the SPIS indicators were devised to establish schools' goals, within the centralised Saudi education system, many head teachers were unsure that these indicators expressed the goals of their schools. However, in this case, these results could have been the outcome of head teachers being more instruction-oriented than the policymakers. Therefore, they did not find that the SPIS indicators expressed their schools' aims. This is further explored in the qualitative findings presented later in this chapter.

Another important issue related to the purpose of SPIS monitoring techniques is the indicators' potential to help schools discover the strengths and weaknesses of their performance. Thus, the head teachers in this study were asked about their ability to identify the strengths and weaknesses of their schools. The majority stated that the KPIs took into account and recognised their school's strengths, but more than two thirds of the participants believed that the SPIS KPIs only 'Sometimes' or 'Usually' included their schools' weaknesses. As a possible explanation of these results, the SPIS Organisational Guide provides clear indicators and norms of evidence required in the evaluation to show teachers' achievement. Therefore, based on these documents, it was easy to determine the strengths and weaknesses of each teacher. These results suggest that SPIS monitoring techniques could better help schools identify the strengths rather than the weaknesses of teachers' performance. This survey also investigated teachers' opinions on school monitoring, as explored in the next subsection.

5.3.2 Quantitative Findings: Teachers' Experiences of School Monitoring

The survey asked the teachers how often their teaching had been monitored by head teachers, head teachers' assistants, school advisors, other teaching staff, and SPIS inspectors over the past year.

Table 5.6: Extent of monitoring by key stakeholders, according to classroom teachers

		Never	1–3 times per year	4 or more times per year
By head teachers	n	7	99	3
	%	6.4	90.8	2.8
By head teachers' assistants	n	14	91	4
	%	12.8	83.5	3.7
By other teaching staff	n	35	69	5
	%	32.1	63.3	4.6
By school advisors	n	1	105	3
	%	0.9	96.3	2.8
By SPIS inspectors	n	0	109	0
	%	0	100	0

Overall, Table 5.6 shows school monitoring to be a significant component of SPE. In general, it would seem that the majority of the teachers in this research were monitored through classroom observations, conducted by their head teachers at the rate of one to three times per year. Moreover, it is noticeable that another large segment – more than three quarters of the teachers – were visited by their head teachers' assistants, and slightly more than half were visited by other teaching staff. In addition, most teachers were visited by school advisors at a rate of 1-3 times per year. Nevertheless, surprisingly, a minority of teachers never had their teaching monitored by either their head teachers or their head teachers' assistants, which may be due to misunderstandings among the teachers, or the problem of the data collection taking place at the beginning of the school year. Here, the teachers were asked if they had been visited by head teachers or others in the current year, because it was stated in the schools' procedural guide that head teachers and assistants should visit every teacher at least once a year. In fact, the statistics generated in this study showed that in-class monitoring was part of everyday school life. Consequently, the teachers and head teachers were well versed in this topic and offered important insights into the current system of school evaluation in Saudi Arabia. More remarkably, concerning school monitoring, SPIS inspectors had visited all the teachers who participated in this research.

Additionally, teachers were asked four questions related to different aspects of SPIS monitoring, including their awareness of when SPIS monitoring would take place and what it would primarily investigate. They were then questioned on the appropriateness of the monitoring techniques. Table 5.7 presents the results of this data analysis.

Table 5.7: Teachers' awareness of SPIS monitoring

		Never	Seldom	Sometimes	Usually	Always
Awareness of when SPIS monitoring would take place	n	21	11	46	21	10
	%	19.3	10.1	42.2	19.3	9.1
Awareness of the main focus of the SPIS monitoring	n	12	9	36	31	21
	%	11.0	8.3	33.0	28.4	19.3

Overall, the teachers' responses to the items in Table 5.7 revealed remarkably low awareness of SPIS among them, particularly regarding the planned dates of monitoring and the areas on which the SPIS inspectors would concentrate during their visits A number of explanations could be suggested to justify these results. For instance, the procedures adopted in the Saudi education system merely involve informing head teachers of the regulations, but do not give much attention to training teachers or providing them with details of the new system. This is discussed in depth in the findings from the interviews with the teachers and head teachers. Table 5.8 gives an overview of the teachers' opinions on the appropriateness of SPIS monitoring techniques.

Table 5.8: Teachers' views on the appropriateness of SPIS monitoring techniques

		Never	Seldom	Sometimes	Usually	Always
Appropriateness of SPIS monitoring techniques	n	10	6	59	24	10
	%	9.2	5.5	54.1	22.0	9.2

The question related to Table 5.8 focused on the respondents' opinions about the appropriateness of the monitoring techniques. Overall, the results suggest that most teachers in this study were unconvinced of the appropriateness of these techniques. It is illustrated that only a small proportion of the teachers reported that the techniques were always appropriate, whereas the same percentage of teachers reported that the techniques were never appropriate. However, the majority of the teachers claimed that the monitoring techniques were either 'Seldom', 'Sometimes' or 'Usually' appropriate. As the previous results show, this large proportion of teachers, who were uncertain about the appropriateness of the monitoring techniques, may have lacked information on what the evaluation would examine, or how it would take place, which

was not shared before the evaluations. In addition, SPIS was a new system and had no certain answers, which is likely to have spurred resistance to change, as discussed in detail in the next chapter.

To enable deeper analysis of the appropriateness of the monitoring techniques, the study included four items that were designed to explore the use and awareness of KPIs: these being important SPIS evaluation techniques. The above-mentioned items consisted of the teachers' ability to read performance tables; the fact of whether parents found these tables to be important; the schools' aims, and the ability of KPIs to identify the strengths and weaknesses of school performance.

Table 5.9: Teachers' views on the use and awareness of SPIS KPIs

		Never	Seldom	Sometimes	Usually	Always
1-Are you able to recall your school's approximate figures from the performance tables?	n	30	43	19	17	0
	%	27.5	39.4	17.4	15.6	0
2-Are your school's aims expressed in the terms of the KPIs?	n	40	0	28	38	3
	%	36.7	0	25.7	34.9	2.8
3-Do the KPIs identify the key strengths of your school?	n	22	0	36	24	27
	%	20.2	0	33.0	22.0	24.8
4-Do the KPIs help identify the main weaknesses of your school?	n	20	0	37	33	0
	%	18.3	0	33.9	30.3	0

The data analysis presented in Table 5.9 shows considerable variation in the teachers' responses to these four survey questions. However, the results reveal that most of the teachers did not apply or have any awareness of the SPIS KPIs. Their responses are examined here in turn; whereby it appears that slightly more than half the teachers could 'Seldom' or 'Never' recall figures from the performance tables, while none claimed that they could 'Always' do so. These results make sense, since teachers are not allowed to see performance tables under the SPIS (for example, the report on the operational plan), except for the reports on their students' skills

achievement. However, it should also be noted that the teachers, according to the interview findings, did not receive any details about SPIS, except as provided by their head teachers and the SPIS website, which significantly affected their ability to read the performance tables.

Interestingly, based on the teachers' views of whether the KPI terms expressed their schools' aims, only 2.8% reported that they 'Always' did, while more than half stated that they 'Never' or only 'Sometimes' did. Clearly, the SPIS KPIs were derived from the education policies set out by the Saudi Ministry of Education, while the education system did not impose compulsory school aims. Consequently, most schools established their own goals, which potentially led to diversity between schools. Logically, the SPIS indicators may fail to express this.

Regarding the identification of schools' strengths and weaknesses in the KPIs, approximately the same number of teachers reported that these factors were 'Never' indicated. Similarly, more than a third of the respondents claimed that strengths and weaknesses were 'Sometimes' targeted. However, the results reveal that slightly more than half the teachers believed that the strengths were 'Sometimes' or 'Usually' identified, while no teachers mentioned that the weaknesses were 'Always' identified. It is clear that according to the teachers' perceptions, the SPIS KPIs can identify the strengths better than the weaknesses of school performance. To understand these responses in more depth, interviews were conducted with head teachers and teachers from different schools.

5.3.3 Qualitative Findings: Head Teachers' and Teachers' School Monitoring by the SPIS

School monitoring was found to be an important part of the SPIS evaluation process, and all types of school evaluation in the Saudi education system. Thus, SPIS appeared to be significant and so this research examined teachers' and head teachers' perceptions of it. In particular, three aspects of school SPIS were highlighted: awareness of monitoring, the appropriateness of monitoring techniques, and the participants' experience of SPIS monitoring (see Figure 5.1).
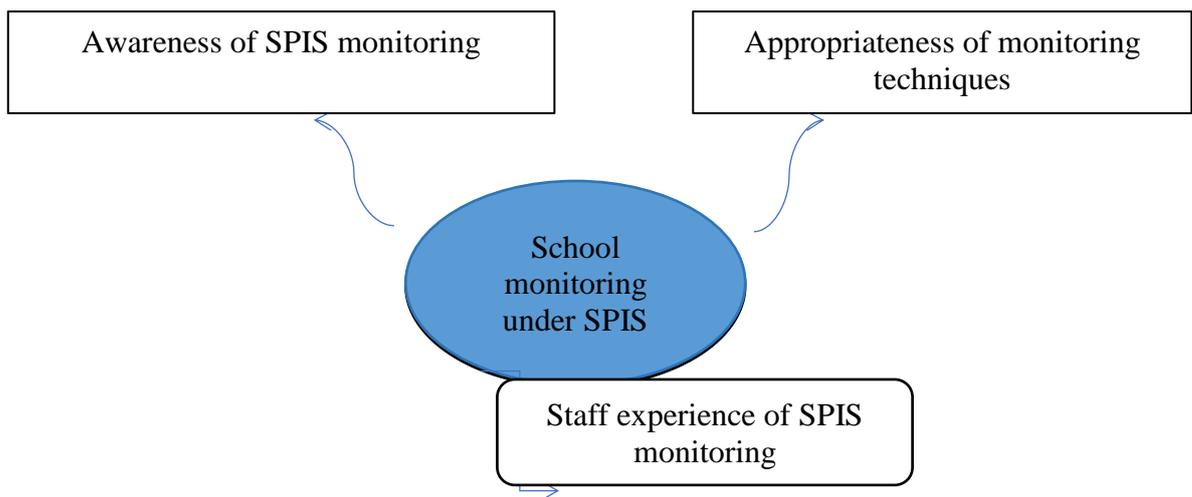
Figure 5.1 Aspects of findings for SPIS school monitoring

School monitoring under SPIS is described in the quantitative findings (see Figure 5.2) for the participants' awareness and views of the appropriateness of the monitoring techniques, which revealed an apparent lack of information, knowledge and training to help them understand the system. These results raise the key question of how the participants obtained knowledge about the system and whether doing so had helped increase their awareness of SPIS monitoring.

Interviews were conducted with three head teachers and nine teachers from schools of three different sizes (small, middle and large) in different areas (North, South and Central Jeddah; see Table 5.1). As a brief reminder, School One was a small school in Central Jeddah, School Two was a medium-sized school in South Jeddah, and School Three was a large school in North Jeddah. The following subsection presents details gathered during these interviews.

5.3.3.1 Awareness of SPIS Monitoring

In this research, awareness of monitoring included three issues related to the participants' awareness (see Figure 5.2).
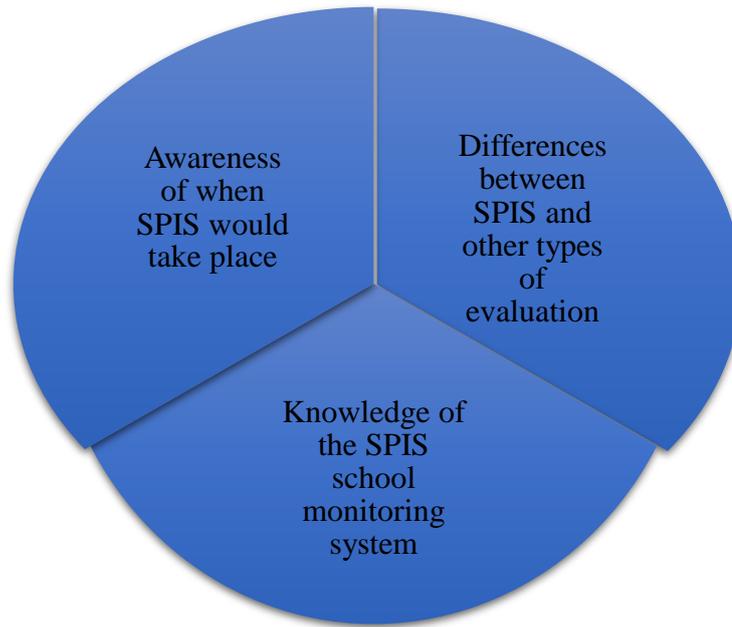
Figure 5.2 *Awareness and knowledge of the SPIS school monitoring system*

Regarding the ways in which the participants received knowledge of the SPIS school monitoring system, most agreed that they did so through school supervisors, who informed head teachers. In turn, these head teachers informed teachers. This important interview finding indicated that the SPIS system considered school supervisors to be a link between the system and the schools' head teachers. Consequently, school supervisors regarded head teachers as a link between them and the teachers being evaluated under SPIS. For example, Sarah, the head teacher of School One, noted:

"I know that through various meetings with my supervisor."

لقد علمت عن ذلك من خلال اجتماعات مع مشرفتي

Meanwhile, Hind, a head teacher of School Two, stated:

"I know about the system from my supervisor, and I have attended some meetings with experts about the system."

لقد علمت عن النظام من خلال مشرفتي وبعدها حضرت اجتماعات مع خبراء حول النظام

Leen, the head teacher of School Three, pointed out:

"I knew about it through courses and from the Organisational Guide."

لقد علمت عن النظام عبر حلقات تدريبية وكذلك الدليل التنظيمي لمنظومة الأداء

121

Leen's answer was no different from that of the other head teachers, because such courses are usually arranged by school supervisors. Additionally, the teachers referred to their head teachers as their main source of knowledge about SPIS. For instance, Salma, from School Two, stated:

"the head teacher told us about the system, its criteria and indicators."

مديرة المدرسة أخبرتنا عن النظام قالت إنه عبارة عن معايير ومؤشرات

Similarly, Muna, a teacher from School One, reported:

"I heard about the system from the head teacher."

سمعت عن النظام من مديرة المدرسة

This SPIS trait, whereupon it is the responsibility of schools to introduce the system, reflects the culture of the education system under the Saudi Ministry of Education, which establishes broad direction through top-down reform. However, it is clear that this system has a considerable number of weaknesses, such as the wide distance between schools and education policymakers, as mentioned in the Discussion Chapter (Chapter 6). Additionally, this approach does not include much direct contact between teachers and policymakers; instead, it relies on third parties to perform this function, usually school supervisors and head teachers. In turn, this undoubtedly bears upon important issues related to the discussion of policies and new systems, particularly information about the application of the system in schools. Such means of obtaining knowledge may be the reason for the lack of awareness of SPIS, indicated in the questionnaire results.

In addition, the interview findings shed light on the teachers' efforts to obtain the requisite information about the system from other sources. For example, Muna from School One claimed:

"I heard about the system from the head teacher. I knew that [the SPIS] would come to evaluate the school, so I browsed the Ministry's website to read the Organisational Guide for SPIS."

"سمعت عن النظام من المديرة وأنهم سيأتون لتقييم آداء المدرسة لذا قمت بزيارة موقع وزارة التعليم والاطلاع على الدليل الإجرائي"

Similarly, Lela, from School Two, reported:

"My head teacher told us about the system, and I have read about it on the Internet and the website."

لقد اخبرتنا المديرة عن النظام ثم قمت أنا بعد ذلك بالقراءة عنه في الانترنت والموقع الخاص به

However, it was clear that the teachers did not find their head teachers to be effective sources of knowledge about the system; instead, they tried to increase their awareness in alternative ways. Nevertheless, although this was a good approach to gaining knowledge, many doubts over the credibility of the information gathered about the system remained. Moreover, the teachers needed to consult experts on the system to answer their questions, which was difficult to do on the website.

Another important finding was that few head teachers and teachers appeared to have referred to the SPIS Organisational Guide as a source of information. This indicates that copies of this Guide had not reached their schools or that the teachers were not sent links to online copies. Undoubtedly, this Guide contains important procedures for teachers and head teachers on the way in which monitoring is to be conducted and on the important indicators that SPIS inspectors apply in their examination of schools. This means that the SPIS has had a negative effect on teachers' ability to deal with monitoring, as it has not helped them access appropriate resources to prepare themselves adequately for its evaluation of school performance.

A comparison between the head teachers' responses revealed another important finding, indicating that there were different ways of providing head teachers with information about the system. One head teacher pointed to courses, while another stated that she obtained knowledge of the system through meetings. In addition, two head teachers referred to their supervisors as sources of information, while another did not mention any of these. This variation may have resulted from the fact that the head teachers fell under different Education Offices in Jeddah, and the Organisational Guide did not outline procedures to clarify the system for teachers or head teachers. Moreover, from the teachers' responses, it seemed clear that most had gained their knowledge of the system from their head teachers. This finding indicates that the Ministry of Education provides just one source of knowledge for the teachers to keep them updated about its monitoring systems and procedures. For example, the teachers saw their head teachers as their primary source of knowledge and their guides at school. Undeniably, this increased the burden on head teachers, while teachers could feel overlooked and undervalued by the Ministry.

5.3.3.2 Awareness of When SPIS Monitoring Will Take Place

The second issue relating to awareness of SPIS monitoring was advance knowledge of when it would take place. This was generally important to the evaluation, as it could negatively affect judgments. Schools need time to prepare for the monitoring day, so that they can demonstrate

what they have done and manage that day at school. However, in response to the questionnaire item asking the head teachers and teachers if they were aware of when school monitoring would take place, most answered 'Seldom' or 'Never' (see Tables 5.3 and 5.7). Therefore, this point was discussed in the interviews with the head teachers, whereby there was a marked contrast between their responses. For example, Sarah, the head teacher of School One, stated that the district office manager had informed her of the date of the evaluation, one week in advance:

> "One week earlier, we had been sent a schedule, which included the organisation of the day for the visiting team and how to prepare our work and arrange a tour for them in the school during school hours."

> "قبل زيارتهم بأسبوع وصلنا منهم جدول يحتوي على التنظيمات الخاصة بالزيارة مثل الموعد وكيف نقوم بتحضير عملنا لهم ليقوموا بمراجعته وكذلك تنظيم جولة للفريق داخل المدرسة أثناء اليوم الدراسي"

In contrast, Hind, the head teacher of School Two, mentioned that the Education Office had

> "informed [her] which semester but did not specify the exact week."

> مكتب التعليم أرسل لنا معلومات الزيارة و الفصل الدراسي الذي ستتم فيه عملية التقييم لكن دون تحديد للأسبوع الذي سوف يزوروننا فيه

The response from the head teacher of School Three was completely different; Leen claimed that she did not have any idea of the day on which the team would come. In fact, there is no specific information in the SPIS Organisational Guide relating to how schools should be informed of the time and date of their evaluations, which may explain the diverse approaches to this process. In addition, some head teachers may have had close relationships with the administration of the Education Offices or the administrative supervisor. Without specific rules for working practice, every Office applies the system in its own way.

The teachers' answers not only differed from one school to another, but also within the same school. For instance, Noha, a teacher at School One, claimed that she had been notified of an evaluation by the school administration, one week in advance:

> "They told us the [SPIS] would come next week to evaluate our work, and the night before, we got a message from the manager saying that there was a committee that would come to us, and we should look decent."

> الإدارة في المدرسة أخبرونا قبل أسبوع أن فريق المنظومة سيأت لتقييم آداء المدرسة أثناء هذا الأسبوع لكن ليلة الزيارة المديرة أرسلت رسالة تخبرنا أنهم سيأتون في اليوم التالي وعلينا أن نبدو في منظر لائق

In contrast, Muna and Meachael mentioned that they only knew about a pending evaluation one day before it took place:

"I knew the day before that they would come. The head teacher sent us a message, so we prepared our files and sent them to the manager's office."

لقد علمت قبل يوم بالزيارة و, المديرة أرسلت تطلب تحضير جميع ملفاتنا والوثائق وإرسالها لقسم الإدارة في المدرسة

"They told the administration, and the administration told us one day beforehand."

هم أخبروا الإدارة والإدارة أخبرونا قبل يوم واحد من الزيارة

This could be due to the fact that some teachers had extra duties at the school, working with the administration or helping the head teachers. Therefore, they knew about the evaluation day, while other teachers did not.

In addition, three teachers from School Three claimed that they had found out about the visit from their head teacher when they arrived at school on the day of the inspection. Salma stated:

"We had not been notified of the time. What happened was that when they came, the head teachers told us and asked us to get our files ready, go to our classes, and wait for their visit."

لم يبلغنا أحد بموعد الزيارة لكن عندما وصل الفريق للمدرسة تلقينا رسالة من المديرة بتجهيز ملفاتنا ‚ الذهاب للفصول وانتظار زيارة الفريق للفصل

Lela added that she knew

"…when they came to the school",

علمنا عندما وصلوا للمدرسة,

while Souad asserted:

"I had no idea. I just knew when they arrived."

لم يكن لدي فكرة من قبل علمت فقط عندما وصلوا للمدرسة

Regarding the conflicting answers given by the teachers and the head teacher from School Three, the head teacher claimed that she had no idea when SPIS evaluation would take place. One teacher's response validated this statement, while the other two teachers' responses diverged. The teacher, Amal stated:

"The head teacher sent us a letter on the same day, saying that she would come with a team to the school to evaluate the work."

المديرة أرسلت لنا في يوم الزيارة تقول أنها قادمة ومعها فريق اليوم لتقييم عملنا

Meanwhile, Marram and Abeer gave completely different answers, with Marram stating:

"I knew the date of the visit, as they told us to prepare our work documents and that they would come to our school within three days, but we did not know when exactly."

125

لقد علمت بموعد الزيارة فلقد أخبرونا أنهم سيأتون خلال ثلاثة أيام لذ يجب أن نحضر كل الوثائق المطلوبة لكن دون تحديد أي يوم بالضبط

Abeer reported,

"The head teacher told me that they would come at this time."

المديرة أخبرتني أنهم سيأتون في هذا التوقيت بالضبط

There was no specific explanation for this difference, but it may have been due to the large size of the school and the fact that the evaluation had taken place a few months previously. Consequently, the staff did not remember what had really happened. These responses indicate that there was no cohesive SPIS process, arrangement or organisation that enabled schools to prepare for inspection team visits. This led to unfairness in the judgments issued by the inspection teams.

5.3.3.3 Differences between SPIS and other Types of Evaluation

Finally, the quantitative data results revealed the multiple evaluation contacts that had performed the same task (see subsections 5.3.1. and 5.3.2). Therefore, the interviewees were asked about their awareness of differences between SPIS and other types of evaluation, including Ministry of Education and comprehensive evaluation. The participants held somewhat varying perceptions of the differences between evaluation systems, as discussed in this subsection.

First, the evaluation conducted by Ministry of Education supervisors was performed by just one person, while SPIS was carried out by a team. Noha, a teacher at School One, reported:

"The evaluation used to be done by a supervisor, while a team came for the evaluation under SPIS. But I did not feel disturbed by them, as they honestly acted very kindly. So, I did not find a serious difference, but of course, there was a difference in focusing on many other things. They had the prestige of being a group and not just one person. The supervisors usually focused on a specific area, while monitoring by SPIS was more comprehensive."

كنا متعودين على تقويم الإشراف التربوي ثم أصبحنا نقوم أيضا بالمنظومة. لم اشعر بالإنزعاج منهم لقد تصرفوا بلطف بالغ لذا لم اجد فرق بينهم وبين الإشراف التربوي لكن كان هناك اختلاف في ما يركزون على تقييمه كما أنهم فريق وليس فقط شخص واحد كما أن المشرف عادة يركز على جانب واحد بينما تقييم المنظومة كان شامل

As pointed out by two teachers from School Three, Amal stated,

"Yes, their work is more general, and they have more prestige and encourage us to feel that we are one team."

"نعم عملهم أكثر عمومية وهم أكثر هيبة من الإشراف ودفعونا للشعور بوجوب العمل كفريق في المدرسة

In addition, Marram responded,

"Yes, the size of the SPIS team was bigger, and they also required several things."

حجم الفريق كبير كما أنهم يطلبون أمور متعددة

Similarly, a teacher from School Two, known as Salma in this study, declared:

"Yes, with the evaluation team, there is more anxiety, because their work is more accurate and comprehensive."

"مع فريق التقويم في المنظومة هناك خوف لأن عملهم دقيق وشامل

It is clear that the size of the team influenced head teachers and teachers in terms of working together and the emergence of some concerns over the accuracy of their work and comprehensiveness. In fact, the comprehensiveness of the assessment is dependent on indicators that require more than just good teaching performance in the classroom. For example, there are indicators relating to student behaviour and communication with the wider community, beyond the school, as explored in the Discussion Chapter (see Chapter 6).

Second, SPIS requires effort from all members of school staff, meaning that they must be adequately prepared to welcome the evaluation team, rather than merely preparing lessons. In contrast, during evaluation by supervisors from the Ministry of Education, the school day ran as normal. For example, Leen, head teacher of School Three, mentioned:

"There is no big difference, but the system team members are more serious."

لايوجد اختلافات كبيرة لكن فريق المنظومة أكثر جدية‘

Finally, evaluation under the SPIS looks at overall school performance, while other types of evaluation tend to focus on just one factor, such as teaching performance. For instance, Abeer, a teacher at School Three, declared that she had noticed an important difference:

"Yes, the Ministry supervisor was coming for the teachers, but the system team was for the school."

الإشراف يأتون لتقييم آداء المعلم لكن المنظومة يأتون لتقييم المدرسة

Overall, these interview findings suggest that the head teachers and teachers were able to distinguish between SPIS evaluation and other school evaluation systems. Although this ability was beneficial for school performance and evaluation in general, the considerable number of

evaluations conducted in Saudi schools could have a negative effect. This is discussed in more detail in Chapter 6. However, the head teachers and teachers were interviewed individually about another issue related to monitoring: the appropriateness of the monitoring techniques, as addressed in the next subsection.

5.3.3.4 Appropriateness of the Monitoring Techniques

SPIS monitoring uses specific methods of evaluating school performance based on KPIs (see section 3.2). This current subsection examines the head teachers' and teachers' perceptions of these techniques; in particular, to determine whether KPIs reflect best practices in school performance and whether the method of system indicators helps identify key factors of success. Moreover, the interviewees were asked if any KPIs in the system evaluation were unachievable, and if so, why. In general, they agreed that the indicators should be suitable for school informants. These questions are addressed in the following paragraphs.

The head teachers' and teachers' perceptions were clearly affected by their school's location and environment. For example, School One in Central Jeddah is in a middle-class part of the city, with low student numbers. Its head teachers and teachers were found to have deeper and more comprehensive perceptions of SPIS indicators, especially the ability of these indicators to measure school performance and identify schools' strengths and weaknesses. In addition, most of the teachers' responses compared SPIS indicators with the indicators in their own schools, making suggestions accordingly. For instance, they recommended that the SPIS draw up a list of indicators to help every school adapt to its environment, also enabling it to achieve this end. Sarah, the head teacher of this school, used an important word in this regard, namely 'realistic':

> "In fact, they were a good team in treating and dealing with the process of evaluation, but their prestige… we were trying to worry too much about the work; it was exhausting. I mean, they must be more realistic."

> "في الحقيقة هم فريق جيد وبارعين في التعامل مع إجراءات التقييم لكن هيبتهم أدت إلى أننا نقلق من تقييمهم لعملنا , لقد كان الأمر مرهق , كان يجب أن يكونوا أكثر واقعية"

She added:

> "Some indicators do not measure reality. Each indicator is supposed to be realistic, not to be evidence of a successful or failed group that does not correspond to the richness of its environment.

> "بعض المؤشرات لاتقيس الواقع وليس مجرد دليل على الفشل والنجاح خاصة في البيئات التي لاتتمكن من تحقيقها"

She also noted:

"I remember some indicators measuring feelings about spiritual things. For instance, there was an indicator discussing the students' prayer performance. It, in my opinion, included the relationship between the worshipper and God, so how could I measure this way or consider it as an indicator? These things are considered to be religious affairs, and I did not know its relation to the indicators or how it was measured. Accordingly, this made me observe the students in something that cannot be measured."

أتذكر بعض المؤشرات التي تقيس مشاعر الأشياء العاطفية. على سبيل المثال ، كان هناك مؤشر يناقش أداء صلاة الطلاب. في رأيي ، العلاقة بين المصلي والله ، فكيف يمكنني قياس هذا الأمر أو اعتبارها مؤشرا؟ تعتبر هذه الأمور شئون دينية ، لا أعرف ما علاقتها بالمؤشرات أو كيف تم قياسها. وفقًا لذلك ، جعلني هذا ألاحظ الطلاب في شيء لا يمكن قياسه.

However, a teacher referred to here as Meachael answered these questions differently:

"To some extent, the indicators must be appropriate to the school environment. When the school is uptown, its indicators are supposed to be different from those in the middle, with a suitable environment and enormous potential."

"إلى حد ما المؤشرات يجب أن تكون مناسبة لبيئة المدرسة فالمدرسة في مستوى مرتفع من المفترض أن تكزن مؤشراتها تختلف عن المدرسة في مستوى متوسط

She also referred to some unrealistic indicators:

"There is some moral success achieved without these indicators. For example, not all strategies fit all students. I have 45 students in my class. How can I apply a strategy that only applies to 25 students?"

هناك بعض الأمور الأخلاقية التي تحققت بدون هذه المؤشرات مثلا ليست كل الاستراتيجيات تناسب أعداد الطلاب مثلا لدي في الفصل 45 طالبة كيف اطبق استراتيجية مناسبة لفصل طلابه فقط 25

In addition, she asserted that some indicators cannot be achieved:

"Yes, there are many, especially those that require a particular environment."

نعم هناك بعض المؤشرات التي من الصعب تحقيقها حيث تتطلب بيئات معينة

Muna, for instance, was unsure whether she could remember the indicators:

"I saw the indicators a long time ago. I'm not sure I remember them, but they were good and made our work appear reasonable."

لقد رأيت المؤشرات منذ وقت طويل فلا أستطيع تذكرها لكنها كانت جيدة لقد جعلت عملنا يبدو معقولاً

However, when asked about the effectiveness of the indicators to help schools achieve their goals, she commented:

"This can be done with the report, if the team members are experts in education."

يمكن القيام بذلك مع التقرير إذا كان أعضاء الفريق خبراء في التعليم

Muna also believed that

"all indicators can be achieved [by schools]."

كل المؤشرات يمكن تحقيقها

However, another teacher, Noha did not agree with her colleagues and considered the SPIS indicators to be weaker than the school indicators:

"I do not think they helped, as we have long-term goals and short-term goals that I feel are stronger than the system indicators."

لا أعتقد أنهم ساعدوا على تطوير آداء المدرسة لأن لدينا أهداف طويلة الأجل وأهداف قصيرة الأجل أشعر أنها أقوى من مؤشرات النظام.

Noha also thought that all the indicators were easy to achieve:

"All the indicators can be applied in the school. It is not a big deal."

كل المؤشرات سهل تطبيقها وهي ليست صعبة أبداً

Similarly, School Three in North Jeddah was wealthy and new, so most of its classrooms had new facilities, but the school was also large and had high student numbers. Therefore, this environment may have affected the head teachers and teachers, encouraging them to give balanced answers. They viewed the indicators positively but were also able to identify their weaknesses. For instance, Leen, the head teacher, commented:

"Yes, they do reflect that very much, but they may overlook some of them, like some minor practices that we cannot document. But… if they let us add what is appropriate for each school and its environment in terms of criteria and indicators, that would be better."

نعم ، إنها تعكس ذلك كثيرًا ، لكنها قد تتجاهل بعضًا منها ، مثل بعض الممارسات الصغيرة التي لا يمكننا توثيقها. لكن ... إذا سمحوا لنا بإضافة ما هو مناسب لكل مدرسة وبيئتها من المعايير والمؤشرات ، فسيكون ذلك أفضل

Similarly, Amal, a teacher at the same school, declared,

"I have to see the indicators in detail and accurately. Then, maybe they can help as the way becomes clear."

أعتقد أنه يجب أن أرى المؤشرات بالتفاصيل , عندما تصبح واضحة أظن يمكن أن تساعدنا يمكنها أن تساعدنا

While Marram pointed out:

"It is possible, but it has interrelated things, and sometimes it does not. I mean, there are unrealistic indicators, while there are other indicators that have begun to show their impact on students, such as behavioural indicators."

إنه أمر ممكن ، لكنه يرتبط ببعض الأشياء ، وأحيانًا لا يحدث ذلك. أعني ، هناك مؤشرات غير واقعية ، بينما هناك مؤشرات أخرى بدأت تظهر آثارها على الطلاب ، مثل المؤشرات السلوكية

Likewise, Abeer stated:

"Yes, but in a good environment only. I think my school environment is bad, and you cannot apply these indicators."

نعم لكن في البيئة الجيدة لكن في مدرستي لاتستطيع تطبيقها إنها بيئة سيئة

Concerning the possibility of achieving the indicators, Amal had an important answer:

"I hope the environment to achieve the indicators will be available. Sometimes, an indicator is good, and I am able to apply it, but the environment is not prepared to help in this. For example, I have 40 students in my class, and some other classes have 50."

أتمنى تكون هناك بيئة مناسبة لتحقيق المؤشرات أحيانا المؤشرات جيدة وأنا قادرة على تطبيقها لكن بيئة المدرسة لاتساعدني في ذلك مثلاً لدي فصل فيه أكثر من أربعين طالبة بينما بعض الفصول فيها خمسين طالبة

Additionally, Marram claimed:

"It is possible, in a way."

إنها ممكنة نوعاً ما

In contrast, however, Abeer had a clear answer and referred to specific indicators when answering the question:

"Indicators of education strategies – as I said, the overcrowded classrooms make it difficult to achieve them."

استراتيجيات التعليم كما قلت الفصول المكتضة تجعل من الصعب تحقيقها

Moreover, the responses from the head teachers and teachers at School Two, which is in a relatively poor environment in South Jeddah, were affected by the location and student numbers. The head teacher, Hind was very optimistic and positive about the system and the monitoring techniques. She chose to discuss the indicator of community partnership, which she clearly found to be a solution to her school's challenges, including meeting students' needs. This indicator had helped her school obtain funding and support from companies and businesses, which had never happened before, as the Ministry of Education's strict regulations prohibited head teachers from seeking financial help from companies for their schools. When Hind was asked whether the KPIs reflected best practices in school performance, she replied:

"The KPIs reflect best practices in school performance, like community partnership. This has supported me and helped me get financial support to help poor students, as well as to benefit from the experiences of mothers to help us train students."

تعكس مؤشرات الأداء الرئيسية أفضل الممارسات في أداء المدارس ، مثل الشراكة المجتمعية. ساعدني هذا في الحصول على الدعم المالي لمساعدة الطلاب الفقراء ، وكذلك الاستفادة من تجارب الأمهات لمساعدتنا في تدريب الطلاب.

She also believed that her school could satisfy the indicators:

"I did not find any indicators difficult to achieve."

لم أجد أي مؤشر صعب التحقيق

Teachers at the school were generally of a similar opinion. For example, Lela stated:

"The indicators can help to identify best practices in the school, because they are clear and organised for our work and cause us to carefully prepare the evidence that we use to deepen our experience."

يمكن أن تساعد المؤشرات في تحديد أفضل الممارسات في المدرسة لأنها واضحة ومنظمة لعملنا وتجعلنا نعد بعناية الأدلة التي نستخدمها لتعميق تجربتنا

In addition, Lela did not find the indicators difficult to achieve, although they needed more time:

"I do not think so, but some of them need more than a year to show their results."

لا أعتقد لكن بعض منها يحتاج لأكثر من عام ليظهر أثر تطبيقها

Based on these responses, there seems to be a consensus amongst the participants, with most referring to the unrealistic nature of some of the indicators. For instance, the indicators relating to education strategies required an upper limit of 30-35 students in the classroom, whereas many schools in Jeddah have approximately 45 students per class. In addition, the teachers agreed on the problems related to the city's school environments. For example, some schools were old and lacked funds, with no facilities for technology use in the classroom, as the indicators required. Consequently, some teachers, such as Amal and Abeer, suggested that these indicators should be adjusted and not applied to every school. However, the head teachers tended to be more positive about the appropriateness of the indicators. For instance, Hind reported that her school received funds from the Ministry to apply the indicator of community partnership. In contrast, Leen found the indicators to be useful, but hoped that her school could become independent and implement its own strategies and methods.

Moreover, there were strong and blatant disagreements over linking teacher performance with student performance. For example, Lela from School Two was very positive about the monitoring techniques, but contended:

"It is good, but there are indicators that I find do not measure accurately, but on the contrary, show the opposite results, such as considering the level of the female students to reflect my performance and efforts, as there are students who are careless and not serious in their studies, despite my multiple efforts and attempts to help them improve."

إنه جيد ، لكن هناك مؤشرات أجدها لا تقيس بدقة ولكن على العكس من ذلك ، تُظهر النتائج المعاكسة ، مثل النظر في مستوى الطالبات لتعكس أدائي وجهودي نظرًا لوجود طلاب غير مهتمين ولا جاد في دراستهم رغم جهودي وتجاربي المتعددة لمساعدتهم على التحسن.

However, this is very common in evaluation in education, especially in a country such as Saudi Arabia, where there is no national testing. The next subsection addresses the final issue related to monitoring: the head teachers' and teachers' perceptions of their experience of SPIS monitoring.

5.3.3.5 Experiences of SPIS Monitoring

SPIS evaluation mainly occurs on a single day of monitoring, when the inspection team visits a school and evaluates its performance. Therefore, most SPIS procedures are performed on the same day. Hence, it was a priority in this study to explore the teachers' and head teachers' perceptions of this process and discuss their experience of SPIS monitoring days.

The head teachers' responses showed a consensus on some organisational procedures. For instance, the teams arrived early at all the schools, required documents, visited classrooms randomly, discussed lessons with students, and spent around 15-20 minutes in any one classroom. However, the teachers' responses varied, as some had never been visited by an evaluation team in their classrooms. In addition, there had not been enough time for all the teachers to show the team their work. These differences may have occurred because the length of the visits depended on school size. Furthermore, a team of three or four could not visit every teacher in a school. The responses gathered illustrate these differences.

In general, the teachers and head teachers reported that the inspection teams were professional and friendly. For example, Sarah, the head teacher of School One, described her experience of SPIS in positive terms and believed that everything went smoothly, as the SPIS inspectors were well prepared for the evaluation; arriving at the school early and treating everyone in a friendly manner. She stated:

> "I knew that when they came, we would spend this day in a normal way. They entered and attended some of the classes, which were randomly selected, and I prepared all the files for them. Then, they read them [very carefully] and reviewed the standards with me… they asked me about what I had achieved. Oh, it was a long day."

> كنت أعلم بمجيئهم ، كنا نقضي هذا اليوم بطريقة طبيعية. دخلوا وحضروا بعض الفصول التي تم اختيارها بشكل عشوائي ، وأعددت جميع الملفات لهم. ثم قرأوها بطريقة دقيقة واستعرضوا المعايير معي ، وسألوني عما حققته. أوه ، لقد كان يوما طويلاً.

The teachers at the school gave similar answers. For example, Noha reported:

"The head teacher told me to get ready, because they would come to visit me in the learning resources room for some time. They entered the room and were impressed by it, and their impression was nice and encouraging. I had a good lesson, and they attended it for a quarter of an hour. They praised me and behaved in a nice way."

المديرة أخبرتني لأكون جاهزة لأنهم قادمون لحضور حصة لي في غرفة مصادر التعلم, دخلوا وأبدوا إعجابهم بغرفة المصادر, انطباعهم كان رائعاً ودرسي كان جيداً لكنهم فقط لربع ساعة ثم غادروا مبدين اعجابهم ولطفهم

Furthermore, Meachael added,

"They arrived and entered the classroom; then, they asked for the work papers from the students to make sure of the strategy, observed the girls' behaviour, and then left."

لقد حضروا ودخلوا الصف عندي ثم سألوا عن أوراق العمل ليتأكدوا من استراتيجية العمل وراقبوا قليلا تصرفات الطالبات وسلوكهن ثم غادروا

Muna, a teacher at the same school, described her visit as follows:

"I did not feel their presence, because they did not get to my classroom, but I only knew that the administration of the school was very busy. There was a lot of noise within the school. The girls did not get their usual full break, as we did not want to disturb the team, and we were keen to calm the students. There was a big mess in the school day."

لم أشعر بحضورهم لأنهم لم يدخلوا فصلي لكني عرفت عندما رأيت الإدارة مشغولين أيضاً كذلك كان هناك الكثير من الضجيج في المدرسة حتى الطالبات لم يحصلن على وقت استراحة كافي مثل العادة لم نرد ازعاج الفريق لذا قمنا بتهدئة الطالبات كانت هناك فوضى كبيرة في المدرسة

Hind, the head teacher of School Two, made more positive comments:

"They came just before eight o'clock. Then, I informed the teachers through a WhatsApp group. I sent one of the employees to inform the teachers in their classrooms, and I asked them to bring the files. My special team, which I formed to help me apply the system, was asked to come and organise the visit. However, most of the required work was here in my office."

لقد جاءوا تقريبًا في الساعة 8 صباحًا. ثم ، أبلغت المعلمين على مجموعة WhatsApp. لقد أرسلت أحد الموظفين لإبلاغ المعلمين في الفصول الدراسية ، وطلبت منهم إحضار الملفات. طُلبت من فريقي الخاص ، الذي شكلته لمساعدتي في تطبيق النظام ، الحضور لتنظيم الزيارة. ومع ذلك ، كان معظم العمل المطلوب هنا في مكتبي.

The teachers' answers perhaps also differed, due to their position in the school. Lela, for instance, replied:

"I was an assistant to the head teacher in clarifying the system and explaining the indicators. We had already prepared our files expecting their visit, so everything went smoothly on this day."

لقد كنت المساعد للمدير في شرح النظام للمعلمات وجميع المؤشرات كنا قد قمنا بإعداد كل الوثائق من أجل الزيارة لذلك جميع الأمور تمت بسلاسة

Meanwhile, Salma noted:

"What happened was that when they came, the head teacher told us and asked us to get our files and go to our classes and wait for them to pass by, but they did not come to my class."

ماحدث في ذلك اليوم أن المديرة قالت لنا أن علينا تجهيز الملفات وإبقاءها معنا في غرف الفصول والانتظار حتى يزورنا فرق التقويم لكنهم لم يحضروا لفصلي

Similarly, Souad stated,

"They came early, and the head teacher asked me to bring all my work, and they visited some classes but not my classroom."

لقد حضروا مبكراً والمديرة طلبت مني احضار عملي ولقد قام الفريق بزيارة بعض الفصول لكن ليس فصلي

At School Three, head teacher Leen viewed SPIS monitoring positively:

"We had prepared and reviewed all the items and made sure to provide the files to the SPIS evaluation team."

لقد أعددنا وراجعنا كل الوثائق وتأكدنا من وجود جميع الملفات لفريق المنظومة

She added:

"They came early in the morning and asked us to bring all the files and allow them to enter the classrooms. This made us feel a bit confused, but it was easy, because we were prepared and had everything ready. It was a good day, and we answered all their questions."

لقد جاءوا في الصباح الباكر وطلبوا منا إحضار جميع الملفات والسماح لهم بدخول الفصول الدراسية. هذا جعلنا نشعر بالارتباك بعض الشيء ، لكنه كان سهلاً لأننا كنا مستعدين وكان كل شيء جاهزًا. كان يومًا جيدًا ، وأجبنا على جميع أسئلتهم.

Among the teachers, Marram claimed:

"I did not notice them when they came. Frankly, I did not see them, as I was not outside my classroom, and they didn't come to my class."

في الحقيقة لم ألاحظهم عندما وصلوا ربما لأني لم أخرج من فصلي كما أنهم لم يدخلوا لفصلي

Amal also mentioned:

"The head teacher sent us a letter saying that the SPIS team had come to the school to evaluate it. And they came, and we brought all our evidence to the head teacher's office…"

المديرة أخبرتنا في رسالة أن فريق المنظومة سوف يحضر لتقييم آداء المدرسة وأن علينا إحضار كل الوثائق لمكتبها .

while Abeer added:

"They entered the classroom for a few minutes and asked for our files. I am not sure if there was enough time for them to evaluate my performance."

لقد دخلوا فصلي لدقائق ثم سألوا عن ملفاتي ولست متأكدة إذا كان الوقت الذي أمضوه في فصلي يسمح لهم بالحكم على آدائي

Thus, some of the teachers questioned whether the inspectors were able to obtain sufficient data within this limited time, as a basis for forming their judgments. This concern may have arisen from the fact that the SPIS monitoring took place on just one day, while other systems, such as school inspection by supervisors, took place in two sessions. This point is elaborated upon in the Discussion Chapter (Chapter 6).

The next section examines the responses to the second research question on the head teachers' and teachers' perceptions of the influence of SPIS on their stress levels, workload and morale.

5.4 Head Teachers' and Teachers' Perceptions of the Influence of SPIS on Their Stress Levels, Workload and Morale

5.4.1 Quantitative Findings: Head Teachers' and Teachers' Workload and Wellbeing

The second research question addressed the head teachers' and teachers' opinions about the influence of SPIS evaluation on their workload, stress and morale. This section discusses the participants' responses to this question.

5.4.1.1 Workload, Stress, Illness and Morale among Head Teachers

The survey included four sections. The results of the first section on 'workload' are presented in Table 5.10.

Table 5.10: Head teachers' views of the influence of SPIS evaluation on workload

		Strongly disagree	Disagree	Undecided	Agree	Strongly agree
1- I do not feel pressure due to workload resulting from SPIS evaluation.	n	7	4	33	12	8
	%	10.9	6.3	51.6	18.8	12.5
2- I work long hours due to SPIS evaluation.	n	17	17	3	20	7
	%	26.6	26.6	4.7	31.3	10.9

Table 5.10 shows that more than half the head teachers avoided disclosing their opinions when asked if they felt pressure due to their workload, while more than one third denied that they felt

any pressure from workload. These results suggest that the head teachers in this study tended not to reveal their feelings about the SPIS system. This could be because they were wary of the strict regulations against disturbing government employees, or afraid that if their answers were revealed, their views might be regarded as complaints against the system, which would be unacceptable to their superiors. Thus, their silence on this point possibly relates to the sensitivity of their positions as head teachers, even though they were informed that their answers would be kept confidential. Surprisingly, more than half of the participants 'Disagreed' or 'Strongly disagreed' that they worked long hours due to SPIS evaluation. This could mean that the system did not directly require more working hours due to the evaluation. It could also indicate that the head teachers' work was usually limited to supervision, while the teachers had other professional duties.

In the second section of the survey, the head teachers were asked about stress relating to SPIS evaluation. Table 5.11 presents these results.

Table 5.11: Head teachers' views on the influence of SPIS evaluation on stress

		Not at all stressful	Mildly stressful	Moderately stressful	Very stressful
1- How stressful do you find being a head teacher in a school evaluated under SPIS?	n	51	10	2	1
	%	79.7	15.6	3.1	1.6

Table 5.11 demonstrates the head teachers' responses to the question on how they found being head teachers of schools that were evaluated by SPIS. The majority stated that it was not stressful at all and so most of their answers were clearly positive. These results may have been due to their practice and feelings, as they wished to present themselves as strong and responsible; successfully bearing the burden of working in the system.

After this question, the head teachers were asked an open-ended question, which sought their opinion on what constituted the most stressful aspects of SPIS evaluation, should any exist. This question elicited 11 non-recurring answers, the most important of which were as follows: sudden visits, sudden changes in the calendar, and a lack of clarification at the beginning of the year, regarding the demands to be met. In addition, this unstable system should be fully explained and not changed every year. From the head teachers' responses, it may be concluded

that stress relating to SPIS was caused by the SPIS regulations not being set out properly or in sufficient detail.

The head teachers were also asked if they had taken any time off due to illness during the six months following an SPIS evaluation. These results are shown in Table 5.12.

Table 5.12: Head teachers' views on time off due to illness after SPIS evaluation

		Never	Seldom	Sometimes	Usually	Always
1- In the six months after SPIS evaluation, have you had any time off due to illness?	n	42	15	5	2	0
	%	65.6	23.4	7.8	3.1	0

The results indicate that head teachers were less likely than teachers to be negatively affected by the evaluation. While more than half the teachers were absent due to illness during the six months following evaluation (see Table 5.12), over half claimed that they never took any time off due to illness. This difference may be the result of head teachers having more experience and greater ability to support evaluation. Additionally, the survey asked the head teachers if their illnesses were linked to evaluation. More than two thirds 'Disagreed', which was unsurprising, given the answers to the previous question.

Finally, this section looks at areas related to the head teachers' morale. The head teachers who contributed to this research were asked to choose one of four values that they believed best described their morale in response to nine statements. The results for this item are presented in Table 5.13.

Table 5.13: Head teachers' views on their morale under SPIS evaluation

		Strongly agree	Agree	Disagree	Strongly disagree
1- Overall, I am satisfied with the administration provided by the SPIS evaluation team.	n	1	3	38	22
	%	1.6	4.7	59.4	34.4
2- The demands placed on me by the SPIS evaluation were reasonable.	n	5	6	12	41
	%	7.8	9.4	18.8	64.1
3- The SPIS evaluation provided appropriate opportunities for the head teacher to work productively with the evaluation team.	n	0	17	15	32
	%	0	26.6	23.4	50.0
4- The oral feedback and written report from the SPIS were consistent.	n	3	4	13	44
	%	4.7	6.3	20.3	68.8
5- The SPIS evaluation identified clear recommendations for improvement.	n	2	5	12	45
	%	3.1	7.8	18.8	70.3
6- I will use the SPIS evaluation's recommendations to move the school/my teaching forward.	n	2	5	14	43
	%	3.1	7.8	21.9	67.2
7- I am satisfied that the views of pupils were explored by the SPIS inspectors.	n	5	0	6	53
	%	7.8	0	9.4	82.8
8- The benefits of SPIS evaluation outweigh the negative aspects.	n	3	4	10	47
	%	4.7	6.3	15.6	73.4
9- Overall, I am satisfied with the way in which the SPIS evaluation was carried out.	n	1	8	11	44
	%	1.6	12.5	17.2	68.8

Overall, from the results, it would appear that most of the head teachers who participated in this research expressed negative attitudes to SPIS evaluation across the nine statements presented in Table 5.13. Looking at the details in relation to satisfaction with the administration undertaken by the SPIS evaluation team, most of the head teachers expressed dissatisfaction, indicating that they 'Disagreed' or 'Strongly disagreed' that it was well performed. This result may be explained by the lack of instructions for the SPIS process, leading to inequality between schools. For example, some schools knew when their evaluation would take place, while others did not know until the actual day of the evaluation.

Similarly, more than three quarters of the sampled head teachers considered the demands of SPIS evaluation to be irrational, and 'Disagreed' or 'Strongly disagreed' with the statement that the demands imposed by it were reasonable. This finding was explored in more detail in the

interviews with these head teachers. Some examples are highlighted here to illuminate the reasons behind the high number of head teachers who perceived SPIS demands as unreasonable.

From their point of view, most of the respondents claimed that SPIS evaluation did not provide appropriate opportunities for them to work productively with the evaluation team. These participants clearly 'Disagreed' or 'Strongly disagreed' with the third statement that SPIS evaluation provided appropriate opportunities for head teachers to work productively with the evaluation team. The Organisational Guide did not cover certain important issues; in particular, the rules shaping and facilitating the relationship between head teachers and the evaluation team. This shortcoming may have led to many problems affecting the head teachers' morale, such as personal bias and the requirement for soft and hard documentation.

In addition, almost all the respondents 'Disagreed' or 'Strongly disagreed' that the oral feedback and written reports from the SPIS evaluation were consistent. A lack of clarity in the evaluation reports was one of the main weaknesses cited in relation to the evaluation, which prevented schools from benefitting from the process, thereby exerting a negative effect on school improvement.

Moreover, most of the head teachers in this study either 'Disagreed' or 'Strongly disagreed' that the SPIS team gave clear recommendations for improvement. Consequently, most of the head teachers seemed to be reluctant to implement the SPIS evaluation's recommendations to move their school and teaching forward. Only seven (around 10%) 'Agreed' or 'Strongly agreed' with the statement that they would use the SPIS evaluation's recommendations to make progress with their school and in their teaching.

Likewise, the head teachers reported dissatisfaction with the inspectors' exploration of pupils' views, the benefits of SPIS evaluation, and the overall process of carrying out the evaluation. A large majority of the participants 'Disagreed' or 'Strongly disagreed' with these items (92.2%, 89% and 86%, respectively). This negative attitude towards SPIS, indicated in the results, could have many explanations, including the head teachers' lack of training or induction into the system. In addition, the SPIS evaluation programme was new and demanded considerable effort; imposing a high level of responsibility, which often spurred resistance to change in schools. The next subsection provides some explanation of these views derived from the interviews with head teachers.

5.4.1.2 Workload, Stress and Illness amongst Teachers

This subsection discusses the teachers' responses to the second question, relating to workload, stress and illness. The first part consists of two statements related to workload. The first was designed to ascertain whether any of the teachers felt pressure from their workload during the SPIS evaluation, while the second enquires about working hours. These findings are presented in Table 5.14.

Table 5.14: Teachers' views on the influence of SPIS evaluation on workload

		Strongly disagree	Disagree	Undecided	Agree	Strongly agree
1- I do not feel pressure due to workload resulting from SPIS.	n	9	20	35	23	22
	%	8.3	18.3	32.1	21.1	20.2
2- I work long hours due to SPIS.	n	4	17	4	47	37
	%	3.7	15.6	3.7	43.1	33.9

Table 5.14 shows the number and percentage of teachers who gave certain responses when asked to describe how strongly they felt about the two statements on 'workload'. Overall, the teachers did not give identical answers about the pressure that they felt due to workload. In Table 5.14, it may be seen that slightly more than one quarter responded negatively ('Strongly disagree' or 'Disagree') to these items. Interestingly, nearly one third (32.1%) selected 'Undecided', while a plurality responded positively ('Agree' or 'Strongly agree'). These results indicate that SPIS evaluation influenced school performance for around half of the participating teachers, resulting in them feeling pressure from their workload. This common response from the teachers is unsurprising, as the evaluation created extra work for them, since they were obliged to prepare documentation and achievement folders. This potentially caused them stress. The results also indicated that the teachers tended to be more ready than the head teachers to share their opinions on this topic (see Table 5.10).

As illustrated in Table 5.14, more than three quarters of the teachers in this study 'Agreed' or 'Strongly agreed' that they worked long hours. A minority of less than one quarter reported that they did not work long hours, while a few were 'Undecided'. These results suggested that according to the teachers' perceptions, SPIS evaluation increased their working hours. Although school hours did not change in Saudi Arabia when SPIS was introduced, the teachers

who perceived that SPIS increased their working hours were likely to have had to complete work at home, as the evaluation required hard and soft copy drafts for its website.

The next part of the questionnaire addressed the feeling of stress during evaluation, with the results presented in Table 5.15.

Table 5.15: Teachers' views of SPIS evaluation's influence on stress

		Not at all stressful	Moderately stressful	Very stressful
1- How stressful do you find being a teacher in a school evaluated by SPIS?	n	8	45	56
	%	7.3	41.3	51.4

The results in Table 5.15 support the findings presented in Table 5.14, regarding the influence of SPIS evaluation on workload. More than half of the teachers in this study believed that it was very stressful to teach in schools evaluated by SPIS. A similar percentage of teachers confirmed that they felt exhausted, due to working in schools subjected to SPIS evaluation. This exhaustion may have resulted from the extra work required by the evaluation process, as mentioned above.

The next question in this survey, which was open-ended, generated 109 responses. The participants were asked if they had experienced stress and were invited to cite which aspects of SPIS evaluation caused them stress, according to their personal opinion. Most of the respondents provided general answers, without specifying any particular procedures, such as 'tension and timing due to heavy burdens' or 'sometimes the teacher is held accountable for things she is unable to apply'. However, some responses referred to specific processes. For example, one teacher stated:

> "Sometimes the visits require examining the paperwork, and this is a hindrance because of the short timeframe. The process of examining the records demands that they must be handwritten, which represents an obstacle, knowing that there are ready-prepared documents like a teacher's guide."

This teacher was clearly referring to the examination of documents. Another teacher mentioned various procedures:

> "We need periodic follow-up, not just tests or attending classes."

In addition, this teacher referred to teachers' needs:

"The system must consider adding what the teachers really need, not just ensuring the papers are filled out."

These answers indicate that the SPIS process contributed to increased stress among the teachers.

Moreover, the teachers were asked if they had taken any time off due to illness in the six months following SPIS evaluation. Table 5.16 presents these results.

Table 5.16: Teachers' views on time off due to illness after SPIS evaluation

		Never	Seldom	Sometimes	Usually	Always
1- In the six months after SPIS evaluation, did you have any time off due to illness?	n	52	0	30	23	4
	%	47.7	0	27.5	21.1	3.7

In brief, it would appear that SPIS evaluation negatively influenced teachers' attendance, confirming the results for 'workload' (see Table 5.16). The teachers stated that SPIS added to their working hours and increased pressure on them; therefore, it seems logical that more than half claimed to have taken days off, due to the effects of evaluation. However, these results are in contrast with those derived from the head teachers' responses, suggesting that the latter were unaffected by SPIS evaluation. There are a considerable number of possible explanations for this difference, for instance, in such cases, the teachers tended to express their suffering more than the head teachers, which may have increased the teachers' feelings of stress, even if they did not inform others about it.

Later, the teachers were asked if they thought that their illnesses were linked to SPIS evaluation, and these results are illustrated in Table 5.17.

Table 5.17: Teachers' views on whether their illnesses were linked to SPIS evaluation

		Major contributing factor	Contributing factor	Minor contributing factor	Not connected
Do you feel that your illness was linked to the evaluation?	n	102	7	0	0
	%	93.6	6.4	0	0

Interestingly, all the participants in this research agreed that their illnesses were linked to SPIS evaluation, but to varying degrees, from major to minor contributing factors. However, most of the teachers considered SPIS evaluation to be a major contributing factor to their illnesses. Thus, evaluation was found to be a negative influence on school performance, and this constituted a major contributing factor of illness among teachers. Although many of the participating teachers stated that they had not taken any time off due to illness in the six months after SPIS evaluation (see Table 5.16), all those who did asserted that it was linked to the evaluation. This result could indicate that most of the teachers attended school despite being sick. Their illnesses could also be psychological or due to feelings of tiredness and stress. Overall, these results imply that SPIS evaluation affected the teachers' wellbeing, according to their perceptions.

Additionally, Table 5.18 shows the participants' responses to certain survey items relating to morale. Here, they were asked to select one of four values, which they believed best described their morale in relation to nine statements.

Table 5.18: Teachers' views on their morale as a result of SPIS evaluation

		Strongly agree	Agree	Disagree	Strongly disagree
1- Overall, I am satisfied with the administration provided by the SPIS evaluation team.	n	5	0	63	41
	%	4.6	0	57.8	37.6
2- The demands placed on me by SPIS evaluation were reasonable.	n	21	13	31	44
	%	19.3	11.9	28.4	40.4
3- The oral feedback and the written report from SPIS were consistent.	n	17	18	35	39
	%	15.6	16.5	32.1	35.8
4- The SPIS evaluation identified clear recommendations for improvement.	n	10	10	44	45
	%	9.2	9.2	40.4	41.3
5- I will use SPIS evaluation's recommendations to move the school/my teaching forward.	n	14	9	41	45
	%	12.8	8.3	37.6	41.3
6- I am satisfied that the views of pupils were explored by the SPIS inspectors.	n	15	18	31	45
	%	13.8	16.5	28.4	41.3
7- The benefits of SPIS evaluation outweigh the negative aspects.	n	21	11	18	59
	%	19.3	10.1	16.5	54.1
8- Overall, I am satisfied with the way in which the SPIS evaluation was carried out.	n	23	19	31	36
	%	21.1	17.4	28.4	33.0

Overall, these results suggest that the teachers' morale was negatively influenced by SPIS evaluation. A deep sense of dissatisfaction with the system was evident from the high rate of negative responses in this regard, which could be explained by examining the results in detail. Surprisingly, however, as shown in Table 5.18, the majority of the teachers who participated in this research either 'Disagreed' or 'Strongly disagreed' with the first statement, namely that they were dissatisfied with the administration provided by the SPIS evaluation team. Only five of the teachers indicated that they 'Strongly agreed' with the statement. This could be due to the fact that SPIS is a new evaluation programme, which needs more work. The teachers were busy and so they may have failed to take this need into account, thereby giving rise to their dissatisfaction and reluctance to change.

Regarding the appropriateness of the demands of SPIS evaluation, a considerable number of the teachers – more than half – 'Disagreed' or 'Strongly disagreed', while one third 'Agreed' or 'Strongly agreed' with the statement that the demands placed on them by SPIS evaluation were reasonable. This high rate of dissatisfaction was explored in the interviews with the teachers, who provided some examples of the irrational and abnormal nature of certain methods, such as asking teachers to use technology that was unavailable in their schools.

In consideration of the oral feedback and written SPIS report provided for the teachers, slightly more than half of the teachers either 'Disagreed' or 'Strongly disagreed' that these were consistent. In contrast, around one third of the participants 'Agreed' or 'Strongly agreed' that they were consistent. It would seem that the reason for these contradictory results was that the team intended to use positive language with the teachers but used official language and set out precise rules in the reports included in the achievement folders. This discrepancy led the teachers to perceive inconsistency between the oral feedback and written reports.

Similarly, most of the teachers 'Disagreed' or 'Strongly disagreed' that the SPIS recommendations for improvement were clear. Unsurprisingly, as a result, almost all the teachers demonstrated reluctance to apply these SPIS recommendations as a means of enhancing their teaching or school (either 'Disagreeing' or 'Strongly disagreeing'). Moreover, over two thirds of the teachers expressed dissatisfaction with the SPIS inspectors' exploration of students' views (either 'Disagreeing' or 'Strongly disagreeing'). Meanwhile, less than a third of the teachers (32 out of 109) viewed the SPIS evaluation positively, either 'Agreeing' or 'Strongly agreeing' that its benefits outweighed its drawbacks. Overall, more than half of the teachers were dissatisfied with the way that SPIS evaluation was carried out (either 'Disagreeing' or 'Strongly disagreeing').

Most of the results were therefore negative, which may have been because the teachers did not receive enough training in the system's procedures and implementation, which was reflected in the results for the last section. The teachers' negative views of the system were likely to have resulted from their estimation of the novelty of SPIS and its focus on assignments that were not considered to be at the core of their traditional work. Moreover, some components of SPIS evaluation are foreign to Arabic culture. For instance, the teachers may have regarded consideration for students' views on their practice to be an insult to them as professionals. The next subsection addresses these issues, as mentioned by three head teachers and nine teachers in the interviews.

5.4.2 The Qualitative Findings: Head Teachers' and Teachers' Perceptions of the Influence of SPIS on Their Stress Levels, Workload and Morale

Issues of stress, workload and morale have recently been raised as problems relating to SPE. Therefore, conducting interviews served to obtain answers to the second sub-question on the head teachers' and teachers' perceptions of the influence of SPIS on their stress levels, workload and morale. Their responses were rich and differed from each other. Some of these responses were affected by the results of evaluation and size of the school. In addition, some head teachers described the impact of SPIS evaluation in positive terms, perhaps because they wished to create a good impression of their ability to manage the issues surrounding evaluation. All these results are discussed in detail in the following subsections, beginning with the topic of stress.

5.4.2.1 Head Teachers' and Teachers' Perceptions of the Influence of SPIS on Stress Levels

The head teachers and teachers were asked about the influence of SPIS on them as individuals. Their answers varied, perhaps because they had different experiences, responsibilities and duties as head teachers and teachers in schools of varying size. Their responses consequently generated interesting findings for this study. For instance, some of the interviewees cited the ill-conceived requirements imposed on them by the SPIS as the reason for their increased stress and tension. In one response, Sarah, the head teacher of School One, described feeling bored on evaluation day and noting it as an important issue, with a major impact on her feelings during the inspection:

> "At this time, there was some tension in the school. When they arrived, I tried to be realistic, but their evaluation was not fair and required ideal manners. For example, during the break, I saw students shouting and playing, which was normal, and it was the only way to get their free time. However, when they heard the noise, they probably considered these things to be abnormal, which decreased the school evaluation score. When they visited a teacher, who chose on that day to teach the lesson without a worksheet or did not use a particular strategy, they saw that as abnormal. In fact, these things happened to us, because they considered that the teachers were intentionally neglecting parts of their professional practice, even though their files were full of evidence of using these strategies… But their way is too ideal, embarrassing, leading to tension and exhaustion. In fact, whatever you do, it will never lead you to perfection."

لقد حاولت أن أكون واقعية لكن تقييمهم كان غير عادل ويتطلب أمور مثالية مثلاً في وقت الفسحة كان الطلاب يحدثون ضوضاء وهذا شئ طبيعي إنه الوقت الوحيد الذي يلعبون فيه أثناء الدوام المدرسي لكنهم اعتبروه غير مقبول في المدرسة كذلك عندما زاروا معلمة لم يكن درسها يتطلب استراتجية تستخدم أوراق العمل لكنهم اعتبروا

147

عدم وجودها في هذا الدرس بالذات نقص رغم أنها أثبتت لهم بالدليل استخدامها لاستراتيجيات وأوراق عمل في دروس أخرى . على كل حال متطلباتهم تبدو مثالية جداً وهذا أصابنا بالاحراج والتوتر والقلق لقد شعرنا أننا مهما فعلنا لن نكون كاملين في نظرهم

Abeer, a teacher from School Three, agreed with Sarah, but referred to time pressure and the number of demands made upon teachers as the reasons for her stress:

"My feelings were negative, because of the accumulation of work and the demand for impossible things, so I could not organise my time."

مشاعري كانت سلبية بسبب تراكم العمل ومتطلباتهم المستحيلة ، لذلك لم يكن لدي القدرة على تنظيم وقتي.

Additionally, the teachers' sense of responsibility for their school's results and position enhanced the results of this evaluation. Noha, from School One, attributed her feelings of pressure to her concerns for her school:

"I was somewhat alarmed. However, it wasn't so bad; we were very normal. But when they told us that they would enter our classrooms, I then felt this kind of stress and tension. I was afraid to affect the results of the school or to make a mistake in the application of strategies, but it passed very quickly."

كنت خائفة نوعا ما. ومع ذلك لم يكن الأمر سيئاً. كنا طبيعيين جدا. لكن عندما أخبرونا أنهم سيدخلون الفصول ، شعرت بهذا النوع من التوتر والقلق. كنت خائفة من التأثير على نتائج المدرسة أو ارتكاب خطأ في تطبيق الاستراتيجيات ، لكنها أي الزيارة مرت بهدوء شديد

A teacher from School Two corroborated Noha's opinion, with Lela noting that she felt a

"little tension because of the fear of the results. The system used it to reduce stress, giving us more time before the visit and specifying the time of the visit."

شعرت بتوتر بسيط بسبب خوفي من النتائج لكن اعطائنا وقت للإعداد ثم اخبارنا بموعد الزيارة خفف من مشاعر التوتر والقلق

However, some teachers did not feel any stress, until they noticed it in everyone around them. For instance, Muna, another teacher at School One, did not feel any pressure. She described the situation as follows:

"I did not feel their presence, because they did not go into my classroom, but I only knew that the administration of the school was very busy. There was a lot of noise inside the school. The girls did not get their usual full break, because we did not want to disturb the team, and we were keen to calm the students. There was a big mess during the school day."

لم أشعر بوجودهم لأنهم لم يحضروا في فصلي ، لكني عرفت فقط أن إدارة المدرسة كانت مشغولة للغاية. كان هناك الكثير من الضوضاء داخل المدرسة. لم تحصل الفتيات على استراحة كاملة كالمعتاد لأننا لم نرغب في إزعاج الفريق ، وكنا حريصين على تهدئة الطلاب. كانت هناك فوضى كبيرة خلال اليوم الدراسي.

Other interviewees believed that the change in the evaluation requirements for teachers, particularly in terms of teaching techniques, had led to their feelings of stress. For instance, Salma described:

> "Fear, anxiety and resentment of some changes to their teaching strategies. The system prompts you to apply modern strategies, which bothers you and forces you to look for a way to improve your performance. I was afraid that we would not get a good ranking, but because of the procedures, not because of our work."

> الخوف والقلق والاستياء من بعض التغييرات في استراتيجيات التدريس. يطالبك النظام بتطبيق استراتيجيات حديثة ، مما يزعجك ويجبرك على البحث عن طريقة لتحسين أدائك. كنت خائفة من أننا لن نحصل على مرتبة جيدة بسبب الإجراءات وليس بسبب عملنا.

Regarding the same changes, Leen, the head teacher of School Three, agreed with Salma about the changes, but considered changes to versions of the SPIS system as the main cause of her stress:

> "I was disappointed that there were rapid changes in versions. Now we are in the fifth version in three years, and that is very tiring and exhausting."

> شعرت بالإحباط بسبب التغييرات السريعة في اصدارات المنظومة لدينا الأن النسخة الخامسة خلال ثلاث سنوات هدا متعب ومرهق

Another teacher justified her negative feelings by referring to the SPIS process. Meachael, a teacher at School Three, believed that the evaluation process had disappointed them:

> "I was shocked that although I was exhausted during the preparation, they did not see the whole lesson, so I almost felt frustrated."

> لقد صدمت فعلاً لم يبقوا حتى نهاية الدرس رغم الإرهاق الكبير الذي أصابني أثناء الإعداد له كنت تقريباً محبطة

However, others described positive feelings about the SPIS evaluation and reports. For example, Hind, the head teacher of School Two, declared:

> "The report was positive. That made me more confident about the methodology of the standards and indicators and their utility in improving performance. I felt happy, because it organised my work and highlighted it."

> التقرير كان إيجابي لقد جعلني أكثر ثقة في منهجية المؤشرات والمعايير وقدرتها على تحسين الأداء لقد كنت سعيدة لقد نظمت وقتي وعملي

In addition, some teachers suggested that early preparation could reduce stress and highlighted that it was the responsibilities imposed upon their schools that had increased their stress. Some teachers emphasised their anxiety and tension, because they had doubts about the fairness of the evaluation. This meant that the teachers' feelings of responsibility for their school and its position had contributed positively to the results of the evaluation.

From another perspective, the interviews focused on workload and more specifically, whether the interviewees felt that the evaluation had increased their responsibilities. This point is discussed in the next subsection.

5.4.2.2 Head Teachers' and Teachers' Perceptions of the Influence of SPIS on Their Workload

When asked in the questionnaire about their workload, 51.6% of the head teachers and 32.1% of the teachers selected 'Undecided'. Consequently, the interviews were used to explore their opinions of this issue in more detail. Most of the respondents reported having more work and duties, due to the problems brought about by the evaluation, but particular features clarified their perceptions. Most prominently, the schools that had already performed self-evaluation before the SPIS evaluation did not consider it to be extra work. For example, Sarah, the head teacher of School One, considered the SPIS evaluation requirements to be part of the usual work undertaken by her school every year:

> "Regardless of the Ministry programmes in the school, it can be included in the policy and guidelines for advancing education and it can be divided up across the teacher, student and achievement files. All the things instructed by the organisation were already applied by us several years ago. Our policy was based on the details that we can now find in the SPIS. Consequently, when the SPIS arrived, we were already applying its standards, except for some differences that only needed to be re-arranged. Let me say that the organisation helped us show our achievements, which we used to do before in a documented way, but without it, we would continue going in the same direction."

> بغض النظر عن برامج الوزارة في المدرسة ، التي يمكن تضمينها في السياسة والتوجيه للنهوض بالتعليم ويمكن تقسيمها بين ملفات المعلم والطالب والإنجاز. تم تطبيق جميع الأشياء التي أصدرتها المنظومة بالفعل من قبلنا قبل عدة سنوات. استندت سياستنا إلى التفاصيل التي يمكننا أن نجدها الآن في المنظومة . نتيجة لذلك ، عندما جاءت المنظومة ، كان لدينا بالفعل معاييرها ، باستثناء بعض الاختلافات التي لا تحتاج إلا للترتيب. اسمحوا لي أن أقول إن المنظومة ساعدتنا في إظهار إنجازاتنا ، والتي اعتدنا القيام بها من قبل بطريقة موثقة. ولكن بدونها ، سنواصل السير في نفس الاتجاه

Teachers in this school were in complete agreement with their head teacher. For example, one teacher, referred to as Noha in this study, claimed:

> "There is no extra work, because the head teacher asked us to do the same things that she used to ask in the past, so we are used to this work. We never find it difficult, and our files are already there."

> لا يوجد عمل إضافي لأن المديرة طلبت منا أن نفعل نفس الأشياء التي اعتدت عليها في الماضي ، نحن معتادون على هذا العمل. لذا لا نجد صعوبة أبدًا ، وملفاتنا موجودة بالفعل.

Muna also mentioned that her school's experience of evaluation made her consider all the work as a routine duty, which she performed every year:

> "There is no extra work, because our school implements an internal evaluation, and we issue a report every year, since our manager has a belief and vision about the importance of this, and I love that in the system; my work is documented and every year, I can go back to my work and document the new work."

<div dir="rtl">

لا يوجد عمل إضافي لأن مدرستنا تطبق تقييمًا داخليًا ، ونصدر تقريراً كل عام لأن مديرتنا لديها فكرة ورؤية حول أهمية التقويم الداخلي . أحب ذلك لقد تم توثيق عملي ، وفي كل عام ، يمكنني العودة إلى عملي وتوثيق العمل الجديد.

</div>

The answers from the participants at this school confirmed the importance of self-evaluation and its ability to facilitate SPIS evaluation, which could help the head teachers and teachers accept the new procedures and SPIS requirements. Additionally, the goals and anticipated results of SPIS could help schools handle any necessary extra work. For instance, Hind, the head teacher of School Two, viewed evaluation as extra work, but also considered its benefits:

> "Yes, there is a lot of work, but it has helped me to understand my work better and organise it."

<div dir="rtl">

هناك الكثير من العمل لكنه ساعدني لفهم عملي أكثر وكيف أستطيع تنظيمه

</div>

This indicates that the teachers were helped by understanding the evaluation and being equipped with the skills and experience to handle any difficulties that they might face during the evaluation process. In addition, it could help increase their acceptance of their workload, especially the work required for the evaluation system.

However, in a large school, where there is already a lot of work, the evaluation could clearly cause resistance and elicit complaints. For example, Leen, the head teacher of School Three (a large school), complained about the work that the evaluation added to her existing duties:

> "There were a lot of papers, files and evidence, and there was some work that did not have clear evidence, or we couldn't find evidence for it… I hoped that they would accept all the work we did, even if we did not find the procedural evidence."

<div dir="rtl">

كان هناك الكثير من الأوراق والملفات والأدلة ، وكانت هناك بعض الأعمال التي لا تحتوي على أدلة واضحة ، أو لم نتمكن من العثور على أدلة لها. وآمل أن يقبلوا كل العمل الذي قمنا به حتى لو لم نجد الأدلة الإجرائية.

</div>

Amal, a teacher at this school, had the same attitude, commenting:

> "Following up each student and the students' portfolios is tiring, and I am unable to do my work, as I spend most of the time doing paperwork rather than teaching, which is my main job."

<div dir="rtl">

إعداد ملفات للطلاب متعب ، وأنا غير قادرة على القيام بعملي حيث أقضي معظم الوقت في القيام بالأعمال الورقية بدلاً من التدريس ، وهو وظيفتي الرئيسية"

</div>

Marram agreed with the head teacher on this matter and described:

> "the challenges of papers that take a long time to prepare or complete. Then, I photocopy them and make several copies."

<div dir="rtl">

تأخذ وقت طويل لأعدادها ثم أقوم بتصويرها وطباعت عدة نسخ منها

</div>

Abeer, a teacher at this school, expressed a similar opinion:

> "A lot of work – I feel I am not coming to school to work as a teacher. I think we need to find assistants for teachers, reduce the number of students, and establish logical criteria."

<div dir="rtl">

الكثير من العمل أشعر أنني لا أذهب للمدرسة للعمل كمعلمة أعتقد يجب أن يكون هناك مساعدة للمعلمة وإنقاص عدد الطلاب في الفصول ومعايير منطقية للتطبيق

</div>

Finally, a number of teachers did not consider some of the work associated with SPIS evaluation, such as the students' portfolios and paperwork, to be their job. They believed that it would distance them from their true job, namely teaching. For example, Souad hoped that the SPIS requirements would be reduced:

> "I think the SPIS evaluation requires great effort, aside from teaching, and they could ease these requirements."

<div dir="rtl">

أعتقد المنظومة تتطلب جهداً كبيراً بعيداً عن التدريس ، أعتقد يمكنهم تخفيف المتطلبات.

</div>

Lela expressed a similar opinion:

> "The huge number of achievement records are very tiring in their preparation, and the student files also need more time to be less tiring."

<div dir="rtl">

إن العدد الهائل من سجلات الإنجاز مرهقة للغاية في إعدادها ، كما أن ملفات الطلاب تحتاج أيضًا إلى وقت لتكون أقل تعبًا

</div>

In addition, Salma referred to another responsibility imposed on teachers during the school evaluation:

> "In this evaluation, I participated in the previous work for the team's attendance, prepared with the director and a group of teachers, so I had extra work in addition to my work as a teacher."

<div dir="rtl">

في هذا التقييم ، شاركت في العمل السابق لحضور الفريق وأعددت مع المدير ومجموعة من المدرسين كل ترتيبات الزيارة، لذلك كان لدي عمل إضافي بالإضافة إلى عملي كمدرس.

</div>

The next subsection describes in detail the head teachers' and teachers' morale, and how they became more confident and satisfied with the SPIS evaluation.

5.4.2.3 Head Teachers' and Teachers' Perceptions of the Influence of SPIS on Morale

Issues of morale, such as confidence, commitment to belonging to an organisation, and interest in participating in the system's development, were discussed with the participants to find out their opinions on how SPIS evaluation influenced their morale. The rich responses differed in their perspectives of teachers' roles in school. These responses clearly reflect the teachers' goals and beliefs in relation to their duties at school. For example, Sarah, from School One, stated:

> "Here, I am like a mother, with transparency… I prefer this description, although many statements have been made about the role of the school principal, I see the school principal as a mother, and she is always the head teacher and responsible for everything. In the end, the mother is always afraid for her children and trying to affect them in positive ways for their own good… Of course, I see myself as very assertive, formal, and professional in my work."

الجميع يعاملني هنا كأم له . صحيح هناك العديد من التوصيف لعمل مديرة المدرسة لكن في اعتقادي المديرة يجب أن تعتبر نفسها أم للجميع والأم هي المسؤولة عن الاطفال وتعليمهم والتأثير عليهم بشكل إيجابي لكني بالطبع في نفس الوقت حازمة ورسمية ومهنية جداً

Meachael, a teacher at the same school, also claimed:

> "I consider myself to be a teacher and a mother."

أعتبر نفسي معلمة وأم أيضاً

In fact, a number of responses described the head teachers' and teachers' role as that of a mother to the students. This could be based on the fact that until just a few years ago, the Saudi Ministry of Education was called the Ministry of Education and Learning. Consequently, many people in the Ministry considered themselves to be parental figures to the students, which undoubtedly affected their work, profession, and maybe even their feelings. Additionally, the responses included references to the duties of a head teacher, which were similar across the Saudi education system, particularly their responsibility for students' behaviour and education.

Meanwhile, most of the participants tended to use supportive language to encourage teamwork and to ensure the impact of SPIS indicators on promoting teamwork in schools. For instance, Hind, the head teacher of School Two, asserted:

> "My role is supportive. I support my team, and I am the leader of the work, but I do not control it."

أنا مساندة وداعمة لفريقي وكذلك القائدة للعمل لكن أتجنب التحكم فيه

The influence of other SPIS indicators was also evident in some of the participants' responses. More specifically, when they referred to their roles in school, they mentioned tasks such as the

educational activities and training required by the SPIS indicators. For example, Leen, the head teacher of School Three, stated:

> "My role is educational and administrative, but I am interested in the behaviour of students, their morale and their preparation for university."

<div dir="rtl">

دوري تعليمي وإداري لكني مهتمة بسلوك التلميذات وأخلاقياتهن وإعدادهن للتعليم الجامعي

</div>

Amal, a teacher at the above school, commented:

> "My role is educational, and I participate in activities."

<div dir="rtl">

دوري تعليمي وأنا كذلك أشارك في الأنشطة

</div>

In addition, Souad claimed:

> "My role comes from my sense of belonging to this school. I am a Maths teacher here, and I do training in national tests."

<div dir="rtl">

دوري يأت من شعوري بالانتماء لهذه المدرسة أنا معلمة رياضيات وأدرب الطالبات على الاختبارات الوطنية

</div>

Additionally, the participants declared that they were committed to SPIS. They all justified it by explaining that they needed SPIS to improve their school's performance, which required commitment to the system; starting with knowledge about how schools could benefit from it. The anticipated results could be strong motivators for commitment to SPIS. For example, Muna declared:

> "What I love about the system is that my work is documented, and every year, I can go back to my work and document the novelty. I also train my students in that, so I think I will commit to it."

Likewise, Hind added:

> "I will be very committed to it, as I am convinced of it to a great extent, since it helped me understand and organise my work more [effectively]."

<div dir="rtl">

ما أحب في المنظومة هو أن عملي موثق ، وفي كل عام ، يمكنني العودة إلى عملي وتوثيق الجديد. أنا أيضاً أقوم بتدريب طلابي على ذلك ، لذلك أعتقد أنني سألتزم به. أضافت هند: "سأكون ملتزمًا جدًا بها لأنني مقتنع به إلى حد كبير حيث ساعدني ذلك على فهم وتنظيم عملي أكثر

</div>

Leen agreed, stating:

> "Yes, because it helped me understand and organise my work better."

<div dir="rtl">

نعم لأنه يساعدني أفهم وأنظم عملي بشكل أفضل

</div>

However, some of the teachers reported that they could not commit to the system without adequate information about it and could not implement the process due to a poor school

environment. It was clear that the potential and scope of different schools and environments were not taken into account during the preparation of the SPIS indicators. For example, Lela emphasised:

"I comply with what I can, but there are other things that I cannot do, such as in the Active Learning card, where I have to apply seven points. I cannot do that, as the lessons are too long, and the session timing is too short: only 45 minutes."

أوافق على ما يمكنني ولكن هناك أشياء أخرى لا يمكنني القيام بها ، كما هو الحال في بطاقة التعلم النشط ، حيث يتعين علي تطبيق سبع نقاط. لا أستطيع أن أفعل ذلك لأن الدروس طويلة جدًا ، ووقت الجلسة قصير جدًا ، 45 دقيقة فقط.

Amal confirmed Lela's statements:

"The school environment is poor, so how do you evaluate us, when the school lacks the equipment? In other words, we are expected to use technology, when our school does not have equipment in the classrooms, such as projectors and smart boards."

البيئة المدرسية سيئة ، فكيف تقيمنا عندما تفتقر المدرسة إلى المعدات؟ بمعنى آخر ، من المتوقع أن نستخدم التكنولوجيا عندما لا يوجد في مدرستنا معدات في الفصول الدراسية ، مثل أجهزة العرض واللوحات الذكية.

However, Marram was the only teacher who refused to commit to the system. She justified her response as follows:

"My commitment! I do not know exactly what that means. For example, documentation is not important to me as I have so much other work, and I do not have time to document [things], especially as they demand that we use the same paper, which is very hard."

التزامي! أنا لا أعرف بالضبط ماذا يعني ذلك. على سبيل المثال ، لا يعد التوثيق مهمًا بالنسبة لي لأن لدي الكثير من الأعمال الأخرى ، وليس لدي وقت لتوثيقها ، خاصة وأنهم يطالبون باستخدام نفس الورقة الخاص بهم ، وهذا أمر صعب للغاية.

Nevertheless, some teachers referred to self-evaluation as a factor that had helped them become more committed to SPIS. Meachael added an important point:

"I wonder what we have benefited from. I would like to have a pre-evaluation definition course to show us how to apply the indicators and how to better distinguish the standards, as the situation is vague, ambiguous, unclear and very superficial."

أتساءل ما الذي استفدنا منه. أرغب في الحصول على دورة تعريف ما قبل التقييم لتوضيح كيفية يتم تطبيق المؤشرات وكيفية التمييز بشكل أفضل بين المعايير حيث أن الوضع غامض وغير واضح وسطحي للغاية.

Noha also mentioned self-evaluation, which her school had already applied:

"I do not know exactly, but I am already committed to these things, and we have functional performance, so we commit to things that we discover within the standards of the system, so it is just [a case of] changing the names. It does not change any of our standards or indicators."

لا أعرف بالضبط ، لكنني ملتزم بالفعل بهذه الأشياء ، ولدينا أداء وظيفي ، لذلك نحن نلتزم بالأشياء التي نكتشفها ضمن معايير النظام. لذلك هو مجرد تغيير الأسماء. لا يغير أي من معاييرنا ومؤشراتنا.

In general, most teachers involved in managerial work relating to SPIS evaluation clearly felt more like they belonged to their schools, were more knowledgeable about the details of SPIS, and were more confident about its application. Consequently, they accepted it and were satisfied with the extra work that it generated, in addition to their teaching. Accordingly, Muna talked about her role in the school:

> "I am a different teacher, and I do managerial work as well. I have managerial as well as teaching skills."

أنا معلمة مختلفة لأني مكلفة بأعمال إدارية لأن لدي مهارات إدارية ومهارات تدريس

This makes it clear that these teachers and head teachers perceived themselves positively and were confident about acting as guides. They may have had these attitudes, because most of them had spent at least seven years at their schools, which could explain their strong sense of affiliation to their establishments.

Concerning suggestions, some of the participants refused to give any, as they did not have enough knowledge about the indicators. However, most were happy to make suggestions for improving the system, which wold increase their affiliation to it and their belief in its ability to improve and help their schools achieve high levels of performance. These suggestions may be divided into three main categories; some involved the system in general, while others were more specific. The first suggestions referred to the importance of creating a list for each school, based on its needs and ability to apply the indicators. For instance, Sarah stated:

> "There should be appropriate standards for each school, depending on the school environment, location and number of students. Using one method of evaluation for every school is not fair, as the same procedure cannot be applied in all schools. There are schools that lack some equipment or facilities, so why should they be evaluated for not using them, if they are not available to them, like the lack of a library or laboratory? There are many schools that have exceeded these standards and need higher standards and bigger challenges to apply each year."

يجب أن تكون هناك معايير مناسبة لكل مدرسة بناءً على بيئة المدرسة وموقعها وعدد الطلاب. إن استخدام طريقة تقييم واحدة لكل مدرسة ليس عادلاً حيث لا يمكن تطبيق نفس الإجراء في جميع المدارس. هناك مدارس تفتقر إلى بعض المعدات أو المرافق ، فلماذا يجب تقييمها لعدم استخدامها إذا لم تكن متاحة لها ، مثل عدم وجود مكتبة أو مختبر؟ كذلك هناك العديد من المدارس التي تجاوزت هذه المعايير وتحتاج إلى معايير أعلى وتحديات أكبر لتطبيق كل عام.

Sarah suggested redesigning the system so that it would be more flexible and convenient for every school environment across the city. Agreeing with Sarah, other teachers referred to the

negative influence of ignoring schools' abilities and needs when dealing with the indicators. In their view, this blanket approach had led to inaccurate judgments about schools.

There was also a widespread belief among the interviewees that some indicators could be improved if changed or fixed. Hind, the head teacher of School Two, reported:

> "I hope that the date of evaluation will be determined. They also measure school performance by recording students' and teachers' tardiness. I mean, they record the percentage of students' and teachers' tardiness. This is not my fault, and it is bad for the performance of the school, as it is something that I cannot change and that I don't have control over. I cannot improve it, because there are no sanctions for late teachers, but salary deductions, while it affects school results."

> آمل أن يتم تحديد موعد التقييم. كما أنهم يقيسون أداء المدارس من خلال تسجيل تأخر الطلاب والمدرسين. أقصد أنها تسجل النسبة المئوية لتأخر الطلاب والمدرسين. هذا ليس خطأي ، وهو أمر سيء بالنسبة لأداء المدرسة لأنه شيء لا يمكنني تغييره ، وليس لدي أي سيطرة عليه. لا يمكنني تحسينه لأنه لا توجد عقوبات على المعلمين المتأخرين ولكن يتم اقتطاع الرواتب ، بينما يؤثر ذلك على نتيجة المدرسة

More than half of the participants regarded advance knowledge of the day of the evaluation as a priority for them, while only one participant argued about it when she considered SPE based on student achievement to be a serious matter. Similarly, other participants referred to specific aspects of evaluation. For example, Leen mentioned the SPIS requirement for authentication:

> "I wish they could accept all the work we do, even if it is not backed up with formal evidence."

> أتمنى قبول كل الأوراق التي نعمل عليها حتى لو لم تتوافق مع النسخ الرسمية التي تستعملها المنظومة

Leen had refused to authenticate her work and instead, wanted her work to be accepted by the SPIS without authentication. Meanwhile, Abeer had important suggestions, possibly relating to what was mentioned by Sarah:

> "I want to remove all that is impossible to apply in the environment of our school. For example, they asked us to use technology, and we do not have any devices to help with this in the school, and there is nothing to add."

> أريد إزالة كل ما هو مستحيل التطبيق في بيئة مدرستنا. على سبيل المثال ، طلبوا منا استخدام التكنولوجيا ، وليس لدينا أي أجهزة مساعدة في المدرسة ، وليس هناك ما أضيفه.

Abeer's use of the term 'impossible' when referring to her ability to implement the SPIS indicators raises important issues relating to teachers' ability and skills to apply SPIS and the efforts made by the Ministry of Education and SPIS to improve them. However, this point will be addressed later in the Discussion Chapter (Chapter 6).

Finally, a few teachers declared that they did not wish to add anything, since they had nothing more to say, or lacked knowledge and information about the SPIS indicators. For example, Marram, from School Three, stated:

> "I do not know what to add, because I need to review the standards of the system. Honestly, I do not care about reviewing them."

<div dir="rtl">

لا أعرف ما الذي يجب إضافته لأنني بحاجة إلى مراجعة معايير النظام. بصراحة ، لا يهمني مراجعتها

</div>

This response clearly indicates that some of the teachers lacked any curiosity about the SPIS indicators. However, Muna, from a small school, claimed that she did not think that there was anything to add, which demonstrates that the teachers were more interested in avoiding co-engagement than giving judgments.

To conclude, although the head teachers had objections to the SPIS, they appeared to be willing to apply the system, perhaps due to the power and role of the Saudi Ministry of Education, which obliges head teachers to adopt it. Some teachers made an important point, asserting that they would commit to SPIS, as their school had applied it before. Therefore, it was easier for them, emphasising the importance of applying this evaluation. This point is addressed later in the Discussion chapter of this thesis (Chapter 6). In addition, some participants claimed that they would apply SPIS based on its results. Overall, however, these findings reveal the importance of training staff and informing them of the system in detail, including all the potential results. The next section outlines the results related to the last research question, concerning the perceptions of the head teachers and teachers with regard to the influence of SPIS on school improvement.

5.5 Head Teachers' and Teachers' Perceptions of their Accountability under SPIS in Relation to School Improvement

5.5.1 The Quantitative Findings

The third research question focuses on the participants' perceptions of their accountability under SPIS, in relation to school improvement. To answer this question, the survey included four statements, designed to explore the participants' opinions concerning school and student improvement. Here, the findings from the head teachers' survey are addressed first.

5.5.1.1 The Head Teachers' Perceptions

This subsection presents the findings from the final section of the questionnaire, with four questions on the head teachers' opinions of the influence of SPIS on school improvement (see Table 5.19).

Table 5.19: Head teachers' views of the influence of SPIS on school improvement

		Improvement	No change	Deterioration	Unable to say
1- Quality of education provided	n	5	6	12	41
	%	7.8	9.4	18.8	64.0
2- Educational standards achieved by pupils	n	36	13	1	14
	%	56.3	20.3	1.6	21.9
3- Pupils' behaviour	n	34	17	1	12
	%	53.1	26.6	1.6	18.7
4- Pupils' attendance	n	35	19	1	9
	%	54.7	29.7	1.6	14.1

The above results show that the head teachers found it difficult to give their opinions on the quality of educational improvement resulting from SPIS. This could be due to laws that strictly prohibit head teachers from expressing their views, and their concern that the data would be shared publicly. In contrast, their reactions to pupils' achievements were positive, indicating 53.1% improvement in their behaviour and 54.7% improvement in their attendance. This inconsistency may have been due to their desire to give the impression that although the system is failing, they have maintained their accountability and continue to perform well. The result indicates that in the opinion of the head teachers, student achievement in schools – as evaluated by the SPIS – is improving, perhaps because the indicators require making an effort to promote high school attendance, achievement and good behaviour, wherein the head teachers recognised that they were accountable for achieving targets.

5.5.1.2 The Teachers' Perceptions

Table 5.20: Teachers' views on the influence of SPIS on school improvement

		Improvement	No change	Deterioration	Unable to say
1- Quality of education provided	n	27	15	36	31
	%	24.8	13.8%	33.0	28.4%
2- Educational standards achieved by pupils	n	32	42	26	9
	%	29.4	38.5	23.9	8.3
3-Pupils' behaviour	n	30	43	19	17
	%	27.5	39.4	17.4	15.6
4-Pupils' attendance	n	32	42	26	9
	%	29.4%	38.5	23.9	8.3

Overall, the results suggest that the teachers had a greater tendency than the head teachers to express their opinions on the ability of the SPIS to improve education quality (see Table 5.19). Different explanations for this may be proposed, including the various positions and requirements of teachers and head teachers. In addition, as the results show, the teachers faced more workload and stress than the head teachers (see Table 5.11), which could have driven them to complain more and seek opportunities to talk about what they had suffered.

Additionally, the results indicate that the teachers were unsure whether SPIS evaluation enhanced education. Their proposed explanations for this are discussed later in the interview findings. However, interestingly, Table 15.20 reveals that a high percentage of teachers believed that there had been no changes in the educational standards achieved by their pupils, or in their attendance and behaviour, following SPIS evaluation, and they felt less accountable for this. Moreover, a moderate proportion of the teachers believed that the educational standards achieved by their students had actually deteriorated, as had their attendance and behaviour. As one possible explanation for this, SPIS requires teaching strategies and methods that correspond to its indicators. Thus, the teachers' workload during the evaluation may have affected their teaching performance. Moreover, the teachers were attempting to apply new teaching techniques and were unsure of how well the results compared with those achieved in their usual practice. Moreover, they may have been less effective, because they were still trying to apply

these teaching methods, but their results had yet to meet their expectations. These results suggest that most of the teachers found SPIS evaluation to be ineffective for improving their pupils' educational standards, behaviour or attendance. This issue was raised in the interviews with the head teachers and teachers, as discussed in the next subsection.

5.5.2 Qualitative Findings: Head Teachers' and Teachers' Perceptions of School Improvement

The participants had different opinions of the influence of SPIS on school improvement, possibly related to their varying experiences during in-school monitoring and evaluation. They also worked at schools of different sizes, which may have affected their views. However, there was a common tendency among the participants who were employed at schools that implemented self-evaluation, which became an important factor of improved school performance. For example, Sarah argued:

> "Let me tell you frankly that our school has a plan for development and a new curriculum for the advancement of education, regardless of the Ministry's programmes in the school. It is probably already included in the policy and guidelines for the advancement of education. Moreover, classifying the teachers, students, achievement files, and everything else that has been introduced by the organisation was already applied by us several years ago. Our policy was based on the details that we can now find in the organisation. Thus, when the organisation appeared, we originally met its standards, except for some differences that just needed to be rearranged."

دعني أخبرك بصراحة أن مدرستنا لديها خطة لتطوير ومناهج لتطوير التعليم ، بغض النظر عن برامج الوزارة في المدرسة. يمكن تضمينها في السياسة والتوجيه لتطوير التعليم ويمكن تقسيمها بين ملفات المعلم والطالب والإنجاز. كل الأشياء التي جاءت من المنظومة تم تطبيقها بالفعل من قبلنا منذ عدة سنوات. استندت سياستنا إلى التفاصيل التي يمكن أن نجدها الآن في المؤسسة. نتيجة لذلك ، عندما جاءت المنظومة ، كان لدينا بالفعل معاييرها ، باستثناء بعض الاختلافات التي لا تحتاج إلا للترتيب.

Although Sarah believed in the important influence of self-evaluation, she admitted,

> "Let me say that the organisation has helped us display our achievements, which we used to do before through documentation, but even without it we were walking in the same direction."

دعيني أقول أن المنظومة ساعدتنا في عرض انجازاتنا بعد أن كانت بدون توثيق معرضة للضياع لكن حتى بدونها سوف نستمر في هذا العمل

Therefore, the SPIS seems to have positively influenced the performance of the above school, but this may be because it had made initial improvements through self-evaluation.

The teachers from this school had similar perceptions but referred to the documentation procedures: an important requirement in all school performance evaluation, although

undoubtedly not the only reason for improved school performance. However, for Noha and Muna, it was enough to improve their school. Noha noted:

> "To be honest with you, I already had a portfolio. I mean, we already had what they demanded from us. We always document our work and we have records of learning resources; we even have records for the laboratory teacher. So, I had done all that was required before the introduction of the system, but perhaps in other schools, they do not have such a director as ours, who is interested in documentation and standards. Perhaps this would contribute to their development."

لأ كون صادقة معك ، لدي بالفعل ملفات موثقة. أعني ، كان لدينا بالفعل ما طلبوه منا. نحن دائمًا نوثق عملنا ، ولدينا سجلات لموارد التعلم ؛ لدينا حتى سجلات للمعلم المختبر. لذا فقد فعلت كل ما هو مطلوب قبل إدخال المنظومة. لكن ربما في المدارس الأخرى ، ليس لديهم مدير مثل مدرستنا ، المهتمة بالتوثيق والمعايير. ربما هذا يساهم في تنميتها.

Muna added:

> "I don't know, but there is nothing in the system that could make a difference, because in our school, we do everything that is required by the SPIS. Our school implements internal evaluation, and we issue a report every year, because our manager believes in and has a vision for its importance."

لا أعرف ، لكن لا يوجد شيء في المنظومة يمكن أن يحدث فرقًا لأننا نقوم بكل ما تتطلبه المنظومة في مدرستنا ، تقوم مدرستنا بإجراء تقييم داخلي ، ونحن نصدر تقريراً كل عام لأن مديرنا لديه فكرة ورؤية حول أهمية ذلك.

Nevertheless, the SPIS was found to have helped improve school performance, as it had contributed to schools achieving educational goals in Saudi Arabia, according to most of the participants. Hind asserted that:

> "This saved our efforts and funds, as we were previously working without awareness and without linking to the education policies. Therefore, all our information became unified and this is the most important thing in the system, as we are benefiting from each other, innovating."

أنقذ هذا جهودنا وأموالنا كما كنا نعمل سابقًا دون وعي ودون ربط بسياسات التعليم. لذلك ، أصبحت جميع معلوماتنا موحدة. وهذا هو أهم شيء في النظام حيث أننا نستفيد من بعضنا البعض الابتكار .

Leen also insisted,

> "Yes, by a large percentage, and they satisfied the Ministry and its policies."

Meanwhile, Lela made the same point:

> "The system has organised our work, documented it and arranged it. Maybe I was doing all these things, but without proof. Now, I am doing the work, organising it and documenting it, so I think it will help us to develop education."

قامت المنظومة بتنظيم عملنا وتوثيقه وترتيبه. ربما كنت أفعل كل هذه الأشياء ولكن دون دليل. الآن ، أقوم بهذا العمل ، وأقوم بتنظيمه وتوثيقه ، لذلك أعتقد أنه سيساعدنا على تطوير التعليم.

These responses clarify that the SPIS can be a roadmap for schools to organise schoolwork and achieve Saudi Arabia's educational goals. However, a few of the teachers refused to answer, as they believed that the SPIS needed many years to realise its influence. For example, Marram declared:

> "I cannot tell you now whether it is useful or not. I think we need many years to see its results in education."

Many of the head teachers and teachers clearly had some positive opinions of the SPIS and found its indicators useful for enhancing education, particularly in helping to organise their work. Moreover, the head teachers and teachers who had longer experience of performance evaluation found the SPIS to be more effective and easier to work with, which appears to have contributed to their confidence in applying the corresponding recommendations. However, some of the respondents said little about school improvement and appeared reluctant to commit themselves. This could have been due to them believing that it would take much longer to realise the impact of SPIS on school improvement. For instance, it could take many years for schools to achieve their target indicators, success and desired improvement. Additionally, one head teacher referred to the system's benefits for explaining Ministry policies that were difficult to understand, because the aim of the system is not to define the education system itself. It could be, however, that the above head teacher was referring to the indicators as the system's requirements.

Regarding the relationship between accountability and school improvement, two important results for this question were derived from the interviews with head teachers and teachers. Firstly, these results relate to the motivation of head teachers and teachers to apply the indicators, and therefore, the development of school performance. This was based on the interviewees' knowledge of the benefits of applying indicators, as well as their role in the development of school performance, or because they acknowledged their responsibility to change and reform education. For example, Sarah stated:

> "In fact, SPIS is great, but I have my own criteria that I am working on, and my goals that I aspire to achieve, in order to improve education and my school's performance. I want people to feel and see in their children that the effort I make with them in the school is very tiring. Education is a great responsibility and I am willing to be the headmaster, so I must exert effort and maintain my sense of purpose in this work."

Conversely, many of the teachers mentioned the weaknesses of their school's environment, which rendered it unsuitable for applying the indicators. This was said to bear upon the

development of their school. They also referred to the unrealistic nature of some of the indicators; suggesting that poor school performance was not solely the responsibility of teachers. For instance, the indicators implied that teachers should not seek excuses for any lack of performance as a prelude to defending themselves over their alleged accountability for any lack of development in their school. In this regard, Souad asserted:

> "I don't think it makes sense to hold us accountable for not applying the indicators that they believe will develop the school, and they did not provide an environment that would help us in their application, such as in the use of technology in the classroom. There is no educational tool to facilitate this."

5.6 Summary

This chapter has presented an analysis of the study findings in three sequential sections, which endeavoured to answer the three research questions. The findings provided significant insight into teachers and head teachers' experiences and perceptions of SPIS, and its influence on school performance. In terms of school monitoring, they raised important points such as the relationship between the lack of awareness of SPIS monitoring and the centralisation of the Saudi education system. Additionally, the findings for the second sub-question are especially important, inspiring debate over the tendency of head teachers to avoid discussing the pressure that they face during SPIS evaluation. The results also showed that teachers have more work than their managers. Likewise, differences were observed between the responses from the teachers and head teachers, regarding the same issue in the third research question, where head teachers were more positive than the teachers about the SPIS system. All these findings will be explored in more depth in the following Discussion Chapter.

Chapter 6: Discussion

6.1 Introduction

This chapter synthesises and discusses the findings of this study (as shown in Chapter 5) with the aim of relating them to the key points that emerged from the Literature Review (see Chapter 3) and to the theoretical ideas outlined in the Methodology (see Chapter 4). Meanwhile, Chapter 5 has already presented the findings from both phases of the study to answer the main research question:

1. *What are head teachers' and teachers' experience and perceptions of the influence of SPIS on school performance?*

In turn, this question raised three sub-questions:

RSQ 1 *What are head teachers' and teachers' experiences and perceptions of the influence of SPIS on school monitoring?*

RSQ 2 *What are head teachers' and teachers' experiences and perceptions of the influence of SPIS on their stress levels, workload and morale?*

RSQ 3 *How do head teachers and teachers describe and understand their accountability under SPIS in relation to school improvement?*

As is clear from the research question and its sub-questions, the purpose of this study was to ascertain the effect of evaluating the performance of head teachers and teachers, using indicators on some aspects of their wellbeing, such as their stress, morale and workload; in addition to the development of school performance, based on their experience and perceptions and using SPIS as an example.

The findings relating to these questions highlight four important issues for the discussion on the reliability and influence of SPIS evaluation in schools in Jeddah. The first issue concerns the challenges facing head teachers and teachers, due to the current method of implementing SPIS as a means of evaluating school performance (see section 6.2.1). Similarly, the second issue raises questions about the reliability of the current system (see section 6.2.2). Next, the third issue concerns performativity and this refers to the wellbeing of head teachers and teachers when subjected to SPIS evaluation and compounded by successive and sometimes concurrent evaluation programmes (see section 6.2.3). Finally, the issue of accountability is addressed, in

relation to school improvement as a result of SPIS evaluation, according to the perceptions of head teachers and teachers (see section 6.2.4).

6.2 What are Head Teachers' and Teachers' Experiences and Perceptions of the Influence of SPIS on School Monitoring?

One of the main purposes of this study was to investigate head teachers' and teachers' experiences and perceptions of the influence of SPIS on school monitoring. In fact, their responses to this research question refer very clearly to the serious challenges facing teachers and head teachers, due to the current method of implementing SPIS school monitoring (see Tables 5.2 and 5.6). These findings stem from two key issues: the first relates to the centralisation of education in Saudi Arabia, meaning that the Saudi Ministry of Education has central control over the management of educational and administrative supervision, as well as of SPIS and other school evaluation systems that run in parallel with it. This system adopts a top-down approach, which gives the top management a monopoly over all decisions on matters of education, without involving school staff (namely head teachers and teachers) in any way (this will be discussed in detail in subsections 6.2.1 and 6.2.2). Illustrative of the importance attached to school monitoring, Saudi Arabia has witnessed a succession of SPE systems, and on occasions multiple monitoring by several bodies that simultaneously assess school performance (see subsection 2.3.6). This raises questions about the nature of what is driving school monitoring both within Saudi Arabia and globally, and what the government's response is to this. Drawing on ideas raised in Chapter 2 of this thesis, the following sections of the current chapter will explore possible answers to these questions. More specifically, consideration will be given to the way that neoliberalism encourages competition between schools and the market forces driving efforts to make education more accountable and efficient. The second key issue concerns the appropriateness of the SPIS monitoring techniques (indicators) and the nature of accountability for meeting the KPIs that are applied to school monitoring and evaluation. Each of these issues is discussed in turn in the following subsections.

6.2.1 Multiple Monitoring Systems

The findings for the first sub-question, which addresses the frequency of evaluation conducted in schools the previous year, reveal how many times head teachers and teachers were monitored

by key stakeholders, because the participants reported experiencing multiple forms of school monitoring (see Tables 5.2., and 5.6). The responses from the head teachers revealed that they were monitored by school supervisors and SPIS inspectors, while the teachers were observed in at least five inspections during 2016. The qualitative findings suggest that the participants were able to distinguish between the parties monitoring the performance of their schools and could identify the differences between them very easily. Therefore, it would appear that despite the similarities between these bodies in the work that they perform, there are clear differences in the way that they operate, although there may be other potential explanations for these discrepancies.

First, as illustrated in subsection 2.2.3.1 education reform in Saudi Arabia has been affected by neoliberalism and NPM in many respects, such as the development of teachers' performance, the establishment of programmes to evaluate the performance of teachers and head teachers, and the identification of the Ministry of Education's requirements, with indicators that need to be satisfied in a school, if it is to achieve a good performance evaluation score.

In general terms, Saudi Arabia, like other countries around the world, has been affected by globalisation, which has opened the doors of economic competition at its most intense. Therefore, the individual school is not only competing with local counterparts, but also globally. International scales of achievement, such as PISA and TIMSS, create the conditions for global competition in educational outcomes and a global platform for countries to triumph or be shamed. This, naturally, puts pressure on entire nations, as they seek to stay ahead of the competition. In fact, the evaluation of school performance has been one of the manifestations of this reform and as the results show, there is more than one system of evaluation in Saudi Araba, namely comprehensive assessment (see subsection 3.5.1) and educational supervision (see subsection 2.3.6.2), put in place by the Ministry to evaluate teaching performance. This indicates the importance of SPE to the Saudi education system.

A possible explanation of this could be that school evaluation has become a topic of interest in the Saudi education system, as explained in the Introduction to this thesis (see Chapter 1), due to the challenges facing the Saudi economy, especially low oil prices, which have forced the government to find other sources of revenue apart from oil. The government has since identified an unavoidable solution, which, according to recommendations from the OECD and other international organisations, involves the application of neoliberalism as a tool for reforming education. Nevertheless, in order to achieve this, the first step is SPE (see subsection 2.2.4).

As discussed in the Literature Review (see section 3.2), neoliberalism can, it is argued, enhance a nation's economy by reforming its education system; thereby providing individuals and society with advanced professional skills and the ability to innovate. Several studies have shown that Saudi education has adopted elements of neoliberalism (see Sakura, 2007; Al Hakamy 2008; Wiseman et al., 2013; Jones & Tymms, 2014). These studies refer to Saudi Arabia's efforts to reform its education system by adopting approaches that place greater emphasis on accountability through school evaluation. The reasoning behind this is that it will help the Saudi economy to develop, so that it no longer relies solely on oil.

In addition, Al Dossary (2006) and Al Sheikh (2011) cite other reasons for implementing SPE. In their view, it has been recognised as a tool for ensuring quality in education. Moreover, both the above researchers refer to the importance of SPE for improving the Saudi education system. Some earlier studies have noted the importance of school performance and its relationship to education reform; for example, Dedering and Müller (2011), who indicate the important role played by school evaluation in improving the quality of school performance. Moreover, McVeigh (2016) highlights the importance of school evaluation for enhancing quality in education. All the above serve to explain why SPE is given so much attention in Saudi education, with the Ministry of Education implementing around 10 evaluation programmes over the past 15 years (see subsection 2.3.6). This finding clarifies that two of these programmes are still active and being implemented in schools: Education Supervision and SPIS (see subsection 2.3.6.1). This enthusiasm for school evaluation has therefore led to multiple monitoring systems, possibly as a result of poor co-ordination. As a result, new school evaluation programmes are introduced, without the old ones being discontinued. Thus, when school inspections take place, it is necessary for the Ministry of Education, inspectors and staff to be clear which evaluation programme is being implemented during each visit.

The second explanation for multiple monitoring could be due to the centralised approach to management in the Saudi education system (Al-Issa, 2009). This leads to top-down management approaches (see subsection 2.3.6), which create distance between schools and the administration and affect the flow of knowledge between parties. This can mean that school performance is evaluated by two different supervisory bodies and SPIS, with the same tasks being performed in different ways. According to Gertler, Patrinos and Rubio (2008), decentralisation aids the transmission of information much more effectively than centralised approaches. The findings that relate to this explanation will certainly be scrutinised more

closely, but there are some issues of immediate importance that need to be addressed at the top level of the Ministry of Education. For instance, it is possible that decision-makers in Saudi education are quite unaware of what is happening in schools, due to the fact that the education system is centralised. Nonetheless it should be apparent that multiple monitoring programmes are being implemented, which places head teachers and teachers under a great deal of pressure. To alleviate this, should this situation continue, then the Ministry of Education needs to minimise confusion and anxiety among head teachers and teachers and for evaluation programmes to be taken seriously. These issues will be discussed in more detail in the next subsection, which deals with awareness of SPIS monitoring.

6.2.2 SPIS Monitoring Awareness

The second key issue arising from the findings for the first sub-question is that teachers and head teachers have remarkably poor knowledge of the SPIS system and its objectives. This result corroborates Alrwqee (2012), who refers to the lack of information given to head teachers and teachers, regarding the comprehensive evaluation applied in schools before SPIS. The participants described the way in which they receive information about the system as hierarchical and top-down. For example, head teachers are provided with information by their supervisors, while teachers refer to head teachers as their source of information. There are certainly many explanations for these results. First, it has been argued that top-down approaches lead to teachers being considered merely as compliant workers, as opposed to professionals who play an active role (Pyhältö, Pietarinen, & Soini, 2014). According to Lowes (2016), this is problematic, because when teachers are involved in making decisions about a programme, they become more aware of the system concerned and place more confidence in it. This finding raises a serious question over the process of education reform in Saudi Arabia, which seems to be devoted to the establishment of new programmes, without instituting procedures and policies that will enable these programmes to succeed. To illustrate this, programmes for evaluating education performance are amongst the most important to be recommended globally in the context of education reform (see subsection 2.2.4), but they need a less centralised regulatory environment in which to achieve success (see subsection 3.4.4). However, this does not seem to be happening in practice within Saudi education, which is predominantly centralised and characterised by a top-down approach

The Saudi teachers in this study did not generally exhibit great confidence in the implementation of SPIS, which may have been due to the inconsistent information that they received about inspection visits, and their limited understanding of the main objectives of this evaluation. They may also have simply viewed it as one of many other evaluation programmes that they are obliged to deal with. The participants also seemed to sense their lack of agency in the process and in turn, appeared to distance themselves from being accountable for applying KPIs. These KPIs are derived from aspects of evaluation that are aimed at improving school performance.

According to the goal-setting theory of performance management systems, there is a strong relationship between whether stakeholders "know the objectives of the system" and whether they "improve their performance" (Koppes, 2014, p.28). This means that the low performance observed in schools in Saudi Arabia may be the result of their limited knowledge of the corresponding evaluation system, rather than their genuine inability to raise standards and improve achievement within their schools. In addition, there were many examples in the findings of teachers' doubts over the current SPE system. The participants questioned the ability of SPIS to make accurate judgments about their competence and performance, precisely because of the short duration of the evaluation visits. The teachers also reported their attempts to search for more information about SPIS on the Internet. Although references to the Ministry of Education programme confirmed the distribution of evidence, none of the teachers or head teachers mentioned the Organisational Guide as a source of information about SPIS. This may indicate that there is no actual guide to SPIS in schools, or there may merely be a lack of encouragement from the Office of Performance to examine and review it.

Additionally, Fullan (2007) argues that under a top-down approach, teachers are less likely to show commitment to an evaluation programme that is implemented by the government and may consequently show resistance to change. However, Fullan's (2007) work was conducted in Western contexts, where there is more freedom than in Middle Eastern cultures. Even though there appeared to be a high level of commitment to SPIS amongst the participants, when they were asked directly in the research interviews whether they would commit to it, they mainly justified their answers with the expectation that it would help improve school performance. According to expectancy theory, employees are most likely to improve their performance and increase their commitment based on their expectations of the results. Thus, if teachers and head

teachers expect SPIS to improve the performance of their school, they will be more committed to it.

6.2.2 Appropriateness of the SPIS Monitoring Indicators and Accountability

The use of performance indicators to evaluate school performance has been a topic of considerable debate across the world (see subsection 3.2), and one of the most important points raised relates to its appropriateness for evaluating school performance (Dangerfield, 2012; Evans, 2011). The findings on the appropriateness of SPIS indicators first revealed disparities between teachers' and head teachers' perceptions of the appropriateness of these indicators, which suggests that they are unsure of its appropriateness. This result has not been previously described or highlighted by any other researchers in the field. Therefore, it is considered as one of the most interesting findings from the current study. For example, half of all the participants considered SPIS indicators to be appropriate 'Sometimes'. Additionally, the findings for this question related to the participants' ability to recall league performance in the SPIS evaluation table, indicating that most of the head teachers and teachers were unable to do so. Clearly, this means that SPIS evaluation does not give any attention to school league performance and does not establish or encourage teachers to refer to it. Additionally, the findings from the qualitative results show that despite the efficiency of the evaluation team, some of the participants mentioned the flawed reasoning behind the indicators. For example, they mentioned issues such as the teacher's inability to apply the indicators, or the difficulties involved in measuring certain aspects, as in the case of emotional issues.

However, the findings of this study are unsurprising in the research field, as regards the performance indicators. Numerous studies have pointed to a lack of confidence in the provisions that are extracted using indicators, as well as the poor suitability of the indicators for application in schools, which is confirmed by Dangerfield (2012) and Evans (2013). In fact, it demonstrates that SPIS indicators lack the most important attributes for effective evaluation, such as clarity and logic (O'Reilly, 2009). These results may be explained by the fact that the Ministry, which is limited through centralisation (see section 6.2) has not yet been able to formulate appropriate or clear indicators, which will enable teachers to apply them professionally, and the SPIS evaluation team to obtain an accurate picture of school performance. Nevertheless, it is perhaps difficult to determine the reasons for these results, because in order to obtain answers, it is

necessary to look at the method of constructing the indicators used in SPIS, but this was not one of the objectives of this current study.

Irrespective of the above, these results highlight the problem of poor efficacy concerning the SPIS indicators; namely, the way in which teachers are held accountable for performance, as measured by such indicators. In particular, neoliberalism and its performance evaluation tools are relevant here (see sections 2.4 and 3.3), in terms of commitment to accountability, because it is through accountability that the government will be sure that its goals are implemented in schools. Most of this requires a high level of performance from teachers to equip students with the skills that will enable them to lead the economy in their country. Thus, accountability can motivate teachers to implement performance indicators (Robinson, 2012; Yia & Kimb, 2019). In turn, it implies that teachers will not be motivated to perform well, if they feel distrust of these indicators and the procedures for their implementation and neither will they accept that they are accountable for poor performance. These factors create a school landscape that is full of conflict, with a rejection of responsibility, because of the lack of conviction amongst all parties involved that they are to blame for any shortcomings or errors. This was confirmed by Taylor and Tyler (2011), who refer to the fact that teachers improve their performance as a result of clear and detailed advice on how best to perform their tasks.

Nonetheless, the SPIS Guide claims that evaluation is not meant to question the competence of school staff, but rather to help them improve their school's performance (SPIS Organisational Guide, 2018). In reality, this concept does not seem obvious; instead, schools are being held accountable for their performance, as is the case with all other evaluation programmes in the UK, Germany, Australia and the USA, where evaluation is used as a means of keeping teachers accountable (Leckie & Goldstein, 2009; Perryman et.al., 2011; Watts, 2012; Goldstein, 2015). Moreover, this is not limited to SPIS evaluation, but applies to all other types of teaching evaluation in Saudi Arabia (although it does not threaten teachers' jobs, because it is carried out on the instructions of the Ministry of Civil Service). Therefore, if a teacher achieves an unsatisfactory result for the first time, he/she will be deprived of the annual allowance, but if this poor performance is observed again the following year, the teacher concerned will be prevented from teaching or performing any other administrative work that is assigned to them. If this poor performance is observed for a third year, the teacher will be dismissed. This illustrates that SPIS and the Ministry of Education do not consider accountability to be a vital factor that can continue to enhance teaching performance, which contradicts a great deal of the

previous literature, which points to the importance of accountability for improving teachers' performance (see Poole, 2011; Robinson, 2012). This may be because accountability helps shape an education system (Møller, 2009) according to the government's goals and the output expected from education, thereby enabling schools and their staff to be categorised (McCallum, 2018). In contrast, Saudi school evaluation does not include any procedures for ensuring accountability, which makes SPIS the real beneficiary. However, this lack may deprive it of the benefits achieved by evaluation systems in other countries, which are linked with assessment and accountability.

6.3 What are Head Teachers' and Teachers' Experiences and Perceptions of the Influence of SPIS on Their Workload, Stress Levels and Morale?

The points that emerged in the findings for this question raise important themes for discussion, such as the workload produced by the evaluation process, and the participants' stress and morale. In this section, all the important findings will be discussed, and possible explanations will be highlighted. The following subsection will begin by looking at the issues relating to workload.

6.3.1. Participants and Workload as an Influence of SPIS Evaluation

The findings for the second sub-question refer to workload; bringing to light three important points for attention: the feelings of stress that result from workload (see subsection 6.3.1.1); the ability of teachers and head teachers to express their opinions freely about the pressure exerted on them by the workload created by SPIS (see subsection 6.3.1.2), and the relationship between self-evaluation and a reduction in the workload imposed by SPIS evaluation (see subsection 6.3.2.2).

6.3.1.1 Increased Workload as a Result of SPIS Evaluation

This study confirms that the workload caused by SPIS evaluation has added to teachers' and head teachers' customary tasks (see Tables 5.10 and 5.14). To the best of my knowledge, no previous study on Saudi education has looked at the relationship between workload and SPE. Most of the studies that have examined a specific evaluation programme, like comprehensive evaluation, in Saudi schools (Al Dossary, 2006; Alballawi, 2009; Al Sheikh, 2010; Al Rwqee, 2012) have done so from the point of view of its effectiveness (see subsection 3.5.1).

Nevertheless, this finding is consistent with Ball (2003), Ball (2012a) and Ball et al. (2012), who considered SPE as the reason for increased workload amongst teachers in schools. The results from the interviews in this present study illustrated that this workload affected day-to-day teaching practice and mainly consisted of paperwork to document achievement.

There are several possible explanations for these results. First, SPE is the result of a wave of neoliberalism. Interestingly, Ball (2012a) refers to this workload as the result of neoliberalism and performance management. The neoliberal approach, which emphasises the creation of competitive market forces, semi-privatisation, accountability and performance evaluation and management seeks to reform education, so that it can make a contribution to the economy; preparing individuals who can drive both the economy and education in a new direction with innovative procedures. One example is the Tatweer school initiative. Nevertheless, performance management has contributed to the introduction of new strategies and mechanisms in education, which require teachers to do more than their traditional teaching work (Whalley, 2011). Therefore, although performance evaluation requires high performance from the teacher, it can also make teachers less keen to exert much effort in meeting the government's expectations of education, as these expectations usually increase their workload.

In terms of the types of workload that teachers normally complain about in relation to SPE, there are several studies that confirm the current research results. Paperwork is the most challenging aspect cited in this regard; for example, Galton and MacBeath (2008) claim that it creates extra work for teachers. Carter (1998), Perryman (2006) and Ball (2012a) confirm this, arguing that paperwork hinders teachers from performing their primary duty of teaching. In fact, paperwork is an important part of performance evaluation, because performance indicators require evidence of what schools have done throughout the year.

The results presented and analysed in subsection 5.3 refer to the lack of available information about SPIS, which may have an impact on perceptions and experiences of SPIS in Saudi schools. Another view that was expressed highlighted how hard it was to apply indicators when there was so little information about them, thereby leading to negative perceptions of the evaluation process. According to Moye, Henkin and Egley (2005), empowering teachers, enabling them to access information that needs to be shared, and initiating discussion can encourage them to believe that all their efforts will help improve education. Moreover, a notable result of this study was that better information provision could help reduce the workload imposed on teachers by school evaluation, which will be discussed in the next subsection.

6.3.1.2 Workload as a Result of SPIS Evaluation and Self-evaluation

The most important finding to emerge from the qualitative data related to workload. It is consequently suggested that self-evaluation can reduce the workload created by SPE and the pressure that it imposes on staff. One of the three schools investigated in this study reported using a self-evaluation procedure. The head teacher and teachers from this school agreed that self-evaluation had the potential to reduce workload. This result reflects those of Chapman and Sammons (2013), who also found that self-evaluation could promote school performance in many ways. The most prominent of these involves preparing for an external evaluation, as this means acquiring the appropriate professional skills to perform the work, with reduced effort required for external evaluation such as SPIS. Moreover, this result corroborates the findings of a substantial body of previous research on self-evaluation such as Janssens and van Amelsvoort (2008), who refer to the benefits of self-evaluation for improving performance. This results from school indicators, which enable head teachers and teachers to prepare for inspection, although these indicators also represent a major hurdle to overcome before external evaluation, even if they do reduce the amount of additional work required. Moreover, the application of performance indicators can be driven by a desire for experimentation in the development of professionalism, ahead of external evaluation

This result may be easily explained in light of self-evaluation as a form of training or practice for teachers; giving them an opportunity to increase their ability and skills for dealing with external evaluation. Through the years, it has become part of their daily work. As a result, they do not consider it to be an extra chore or additional source of pressure that accompanies evaluation. As mentioned in the literature review, Newman at all (2009), and Chapman and Sammons (2013) refer to the importance of self-evaluation and how it can help schools prepare for external evaluation. They also mention its role in motivating teachers to do whatever is required by SPE indicators.

6.3.1.3 Workload and Teachers' Voices

Despite all the procedures that I undertook as the researcher to gain permission and approval from all the relevant bodies, prior to conducting this study, and all the measures implemented to ensure that the participants' identities and data were published anonymously, and that they were satisfied with their responses, a large number of head teachers – more than half – claimed that they were 'Undecided' about the answer to the first interview question, relating to their

feelings of pressure and their workload. This study therefore suggests that teachers and head teachers may find it difficult to express their feelings about the workload that accompanies SPIS evaluation. There are several possible explanations for this surprising result.

The first of these is that the word, 'Undecided' indicates that the interviewees were hesitant about selecting their responses. This could have been because they thought that they needed more experience of the system to be able to decide whether there was any link between the process and their feelings of pressure. Additionally, they may have felt a considerable amount of pressure, but they were unable to identify the source of their stress: whether it resulted from the evaluation or from another factor. In addition, this may be specifically related to Arabic culture, where it is not common for leaders to admit to a lack of self-confidence or that they are experiencing pressure or stress.

Another explanation may be attributed to the instructions provided by the Saudi Ministry of Education, which usually require head teachers and teachers to protect information about their work. However, this policy does not conform to NPM or performance management, which encourages evaluation to identify weaknesses and strengths, so that systems can be improved (see subsections 2.2.3 and 2.2.4). According to Harvey (2007) neoliberalism aims to promote freedom, with employees being allowed to talk openly about their experiences and perceptions. Current Saudi policies therefore contravene normal NPM procedures, which advocate that teachers be considered as stakeholders (Tolofari, 2005). It also raises the question of accountability in relation to evaluation. Head teachers and teachers who are held accountable for school performance are also considered worthy of an opinion, whereby they are encouraged to talk frankly about their experiences of the system and to justify their mistakes, so that they can be supported in avoiding any repetition of them. This also brings the discussion back to the principles of neoliberalism and the extent of its application in the Saudi education system; subsequently reinforcing the process of education reform and identifying whether there is any awareness of the importance of complementarity between these principles, such as freedom of opinion and enhanced performance, so that reform programmes can be more effective. It also calls for a debate on the impact of preventing teachers from speaking out about the increasing prevalence of anxiety and high pressure in their careers.

6.3.2 The Participants' Stress as an Effect of SPIS Evaluation

An initial objective of this project was to identify the participants' perceptions of their stress levels. Their responses to this part of the investigation gave rise to three important themes: the link between accountability and stress, the link between the level of quality in the process and stress, and the relationship between the participants' position and stress. The first of these themes will be discussed in the next subsection.

6.3.2.1 The Link between Accountability and Increased Stress

One interesting qualitative finding was that the reported increase in stress was mainly due to accountability. This finding is consistent with Al-Omari and Wuzynani (2013), who pointed out that the high levels of stress suffered by head teachers in Saudi Arabia and Jordan resulted from their fear of accountability. Despite the common ground between this current finding and that of the above-mentioned study, however, there is one element that should be clarified here, namely that the previous study was conducted on Jordan's performance evaluation system, which places strong emphasis on accountability amongst teaching staff and their leaders. Thus, teachers are fully aware that the results of their performance assessment will impose strong accountability on them, potentially resulting in penalties such as cuts to annual bonuses, dismissal, and investigation or review procedures, which can be very costly to the individuals concerned. In contrast, SPIS does not bear these implications, because SPIS performance evaluation, as indicated in the SPIS Organisational Guide, does not impose real accountability, but rather endeavours to develop the work performed by schools. Nevertheless, it was clear that the participants were unaware of this objective, thereby confirming their lack of knowledge of the aims and processes of SPIS. Consequently, this had led to misconceptions about the programme and a sense of pressure among teachers. Therefore, Penninckx et al. (2016) recommend helping teachers to understand the system of evaluation that is adopted, in order to resolve this problem of stress among teachers.

Additionally, another explanation for increased stress surrounding SPE is accountability. Evaluation is closely linked to accountability, because the information collected by inspectors is used to judge the performance of teachers, head teachers and schools (O'Neill, 2013). Therefore, one of the interviewees justified her increased stress, stating:

> "I then felt this kind of stress and tension. I was afraid to affect the results of the school or to make a mistake in the application of strategies, but it passed very quickly."

It is also appropriate to mention here that more than one interviewee reported feeling reassured when her school's evaluation report proved to be favourable.

This result corroborates the findings of previous work by Sugrue and Mertkan (2017), who refer to increased stress among teachers, if these teachers believe that they will be held accountable for the results of an assessment in which the scores for their school are low. The above-mentioned study was conducted in Britain, where there is a strong system of accountability (Ofsted). Conversely, this current study refers to Saudi education, where accountability does not have the same serious implications as in Britain. Nonetheless, there have been indications from the new Ministry this year that a radical change is about to take place, with the introduction of a system of accountability (see subsection 2.3.6). It points to the importance of the results of this study, because it is clear that a weak system of accountability will cause an increase in pressure. Therefore, as accountability gains momentum in Saudi education, this pressure will also increase. The stress arising from accountability is consequently worth discussing in more detail with further explanation.

6.3.2.2 The Position and Stress Levels of the Participants

Another interesting finding derived from this study was that teachers suffer more frequently than head teachers from stress (see Tables 6.11 and 5.15). This could merely be because few head teachers claimed that they had experienced any stress arising from their work in schools being evaluated by SPIS. There are in fact a number of possible explanations for this result, which will be outlined in this subsection. The first of these is that teachers – as found in this study – tended to receive less information and training than head teachers, and this probably made them more susceptible to pressure than their head teachers. It undoubtedly influenced their understanding of the purpose of evaluation and the implications of its results.

In addition, this result may be attributed to the fact that the work of the head teacher is supervisory, while teachers are required to perform multiple tasks, all in some way related to each other and accompanied by documentation, which imposes intense pressure on them. This explanation is supported by the idea that teaching work can in itself be stressful (Ramos & Unda, 2016).

Another potential explanation is that evaluation can threaten teachers' identity, so that they feel stressed (Evans, 2011). This feeling may have more of an impact on teachers than it does on

other school staff such as head teachers, because performance management (see subsection 2.2.4) and SPE (see section 3.2) change the nature of the teaching profession, so that a teacher not only undertakes teaching duties, but also performs other work such as documentation and services to the community. In this way, students are trained in the skills that will qualify and prepare them to serve the country's economy and promote competitiveness. Consequently, teachers become accountable for their students' exam results. This then drives them to try and improve student achievement in national tests, thereby resulting in more working hours for teachers. All of this has redefined the identity of teachers, leaving them feeling threatened as to their original identity and thereby increasing their sense of pressure.

6.3.2.3 The Quality of the Process and Stress

It is interesting to note that in the qualitative findings relating to stress, it would appear that a lack of quality in the process of evaluation can increase stress levels. This finding is in full agreement with Jaradin (2004), who points out that one of the reasons why head teachers feel stressed consists of the poor quality of supervisors' assessment procedures. Moreover, Al Rwqee (2012) noted that the teachers sampled were disturbed by the poor quality procedures adopted for comprehensive assessment. For instance, there was a lack of co-ordination found between the evaluation team and the school, regarding the date and time of the visit. A possible explanation for this tension was that the teachers felt disrespect for the Ministry, especially given the lack of quality demonstrated, as it pointed to the Ministry's disinterest.

Likewise, this confirms Alkarni (2015), who pointed out that the procedures followed by the Ministry of Education, which demonstrated a lack of respect for head teachers, leads to increased tension between them. It also seems possible that when programme evaluation procedures are of poor quality, they are subject to variation. A lack of understanding therefore leads teachers into problems when they attempt to apply indicators. Hence, teachers inevitably feel under pressure. Additionally, this result may be explained by the fact that a lack of quality sometimes leads to dual action, followed by enormous pressure, especially when the expected results of application are not achieved. Teachers may then have to fix the flaw themselves or re-apply the indicators in the hope of a good result.

However, some interviewees gave examples of successive changes in SPIS policies, where there had been insufficient time for training in each policy. Moreover, this can be explained by the fact that they spend sufficient time practicing the application of each evaluation system but

are caught unawares if there is a change to the system, thereby increasing their stress levels. This result confirms Galton and MacBeath (2008), who refer to the impact of changing policies on stress levels. Additionally, this result in some way supports Penninckx et al. (2016), who claim that stress can be reduced in schools, when the quality of evaluation processes increases.

6.3.3 The Influence of SPIS on Morale

It is interesting to note that the process of SPIS evaluation clearly has an influence on teachers and head teachers, in terms of their perceptions and morale, relating to their level of satisfaction with the system. It also affects collaboration within the school team and their commitment to SPIS. This section discusses all these important themes and the first of these will be addressed in the next subsection.

6.3.3.1 The Participants and Their Satisfaction with the SPIS Process

What is surprising is that the participants' morale was negatively influenced by SPIS evaluation (see Tables 5.13 and 5.18). Most of the head teachers and teachers who were asked in this study about their satisfaction with SPIS processes, including the administration carried out by the SPIS evaluation team and the way in which evaluation was conducted, had a negative impression of the process. However, there are some important explanations that could clarify this result, and these are outlined below.

SPIS uses indicators to evaluate schools. This technique measures their performance in a way that is quite different from the traditional methods applied in the past. Some of these new approaches were familiar to the participants and so they were able to deal with them quite easily. Conversely, those that were not so familiar to them were more difficult for them to handle. Consequently, as the results indicate, this led to a strong sense of dissatisfaction, which confirms the findings of previous studies, such as Ball (2012b)and Bousquet (2012). Moreover, two studies have been conducted in the Saudi context, as confirmed by Al Rwqee (2012) and Alkarni (2015), who highlight the dissatisfaction amongst head teachers and teachers, regarding the Ministry of Education's procedures for evaluation in general. In addition, this result could be a consequence of the hard work that was required to implement this system, such as paperwork and documentation. Moreover, this is what teachers have confirmed previously in related studies. For example, Galton and Macbeth (2008) assert the influence of performativity

on teachers' morale. Bousquet (2012) also supports this explanation, because he considers workload to have a negative influence on teachers' morale.

Aside from the above, numerous procedures appeared in the results, with the teachers explaining their lack of conviction. This shed some light on their reasoning, as well as on their lack of understanding and on the inconsistency of the procedures. Besides, the teachers highlighted the Ministry's lack of interest in delivering information directly to them and providing them with adequate training. These factors are likely to have given them a sense of dissatisfaction with the progress of evaluation and instilled in them the sense that the Ministry does not respect them, thereby contributing even further to their frustration with the system. Reid's (2010) findings are in agreement with this explanation, because he observes the lack of appreciation afforded to teachers, especially in relation to the importance of their role. This was found to have an adverse effect on teachers' morale.

6.3.3.2 Collaboration within the School Team as an Influence of SPIS

It is interesting to note that one of the most significant results in the qualitative part of this study pointed to the school team being encouraged to co-operate in the application of indicators, in order to ensure success in the process. There are a few possibilities that could explain this study finding. First, the wave of neoliberalism and NPM requires greater participation from all stakeholders such as head teachers and teachers (Tolofari, 2005), especially where these actions are aimed at reforming education to serve a nation's economy. This makes for national motivation amongst teachers to work collaboratively, as a means of serving their home country. Similarly, every participant needs to co-operate in a school to achieve its goals, because the indicators require a high level of co-operation and collaboration from staff for their implementation. This will ensure that the work is performed appropriately in school activities, including trips, the completion of projects, observation of student behaviour and follow-up, and meetings with parents.

In addition, an important element of this finding is accountability, which obliges teachers to improve their professional performance, because any failure on their part will have a negative impact on the school evaluation report. This finding may be compared with work by Al Rwqee (2012), who suggests that accountability can help improve school performance, while also promoting co-operation between team members. Therefore, it is considered as one of the most important features of evaluation work, consistent with the literature on improving

administrative work in institutions by introducing accountability (Møller, 2009; Bessant et al., 2015). In fact, under SPE, teachers are observed more closely (Apple, 2004), which causes them to comply with their instructions. In turn, this requires them to work co-operatively. Equally, this finding can be explained from the perspective of expectancy theory, which claims that teachers' performance depends on the goals that they expect to achieve (Mullins, 2007). Therefore, it could be said that if teachers believe that they will achieve goals through co-operation, they are very likely to co-operate.

6.3.3.3 Commitment to SPIS

What was surprising was that although the head teachers' and teachers' levels of satisfaction with the SPIS process were low (see subsection 6.3.3.1), the finding for the extent of their commitment was high, which contradicts the result of a previous study conducted by Torabi and Sotoudeh (2010). It suggests that high commitment is the result of teachers' positive attitudes to the system. However, Torabi and Sotoudeh's (2010) study was conducted in Iran, which differs from Saudi Arabia in terms of its culture and organisational systems. For example, the few organisational systems that exist in Saudi Arabia are not open to investigation. Correspondingly, there are few explanations that can help clarify this finding and provide possible reasons for it.

Nevertheless, the first point that may be made is that labour regulations in Saudi Arabia do not permit employees to refuse to perform any work that they are assigned (Ministry of Civil Service, 2019). Therefore, the head teachers and teachers in this study may have shown their commitment to SPIS, purely because they considered it to be part of their job description, so that they could avoid being threatened with dismissal. Additionally, it may also have related to the updated conditions for awarding the annual bonus, which now demand a high level of performance (Ministry of Education, 2019), placing teachers under great pressure to commit themselves to the education system, with SPIS being an important part of this.

Additionally, there is another potential explanation for this finding, which is very likely to be the commitment of head teachers and teachers to SPIS, because the evaluation system requires more participation from teachers in the work of their school. For example, the SPIS indicators encourage teachers to perform numerous extra-curricular tasks such as morning participation, extracurricular activities, school board membership, and meetings with parents (SPIS, 2018). In fact, teachers tend to become more committed once they become part of a school. This has

been confirmed by a considerable number of studies on teachers' commitment (see Singh, 2007; Aydin et al., 2013; Saljooghi & Salehi, 2016). These researchers refer to the influence of teachers' participation on their commitment at school, and the work that they carry out as a result, which has been found to have a positive impact on school results.

6.4 How do head teachers and teachers describe and understand their accountability under SPIS in relation to school improvement?

The final sub-question of this study relates to school improvement as an influence of the implementation of SPIS performance evaluation, specifically in terms of the quality of education provided; the educational standards achieved by pupils; pupils' behaviour, and pupils' attendance. Here, all findings for this question are important to discuss and explain, but the most important of these were divided into three sections: the low ability of head teachers to determine the influence of SPIS evaluation on quality in education, the teachers' negative perceptions of the influence of evaluation on pupils' behaviour and attendance, and the quality of the organisation of school performance and evaluation. The first of these findings will be discussed in the next subsection.

6.4.1 Head Teachers' Perceptions and the Quality of Education

One unanticipated finding for the third research question in this study, which enquires about the perceptions of head teachers and teachers, was that the head teachers appeared to be unable to define quality in education, specifically at their school. This revealed their capacity to determine whether education quality was influenced by SPIS evaluation. In a review of the relevant literature, no data were found to contradict or confirm this finding. It therefore raises a considerable number of questions over why the head teachers were unable to provide clear answers on this topic, such as whether they believed that revealing the level of quality in the education provided by their school would be tantamount to sharing confidential school information, which should not be shared with the public. Moreover, they may have been reluctant to attract criticism by revealing the quality of their schools. In addition, they may simply have been unable to evaluate education quality. All these possible explanations will be discussed both broadly and impartially here.

First, in the sixth version of the Saudi SPIS Organizational Guide (2018), there are 45 indicators of school performance, with school performance quality being evaluated according to a school's

achievements (SPIS, 2018). Meanwhile, in the UK and USA, for example, the quality of a school will depend on the students' results in national tests (Goldstein & Leckie, 2008; Perryman et al., 2011). However, the approach adopted in SPIS is more useful in the sense that it involves more than merely judging schools based on students' results: the negative effects of which have been widely discussed in previous studies by Ball (2004), and Perryman et al. (2011). Nevertheless, the finding for this item, which deals with the quality of schools following SPIS performance evaluation, demonstrates the low ability of head teachers to determine the quality of their schools. This could be due to the fact that the large number of indicators make evaluation too difficult for school staff. Additionally, there are no school performance league tables in Saudi education, which would otherwise enable the quality of schools to be determined according to their rankings based on achievement.

However, there are other possible explanations, including the fact that head teachers may be afraid to make public any information that could indicate the quality of their school's performance. Therefore, the absence of school performance leagues in the Kingdom makes it easier for them to protect any information that could reveal the actual performance of their school. Interestingly, this explanation corroborates the findings from extensive previous work on the relationship between teachers' fears and stress and published SPE results (for example, Goldstein & Leckie, 2008; Perryman et al., 2011; Baroutsis, 2016). Thus, the selected head teachers were unable to reveal their perceptions of their school's quality in response to questions about quality following SPIS evaluation, due to their fear of criticism.

6.4.2 Teachers' Negative Perceptions of the Influence of SPIS on Pupils' Behaviour

The results obtained from the qualitative data in this part the study, relating to school improvement as an effect of performance evaluation, showed that the teachers had negative views of the improvement to pupils' behaviour as an effect of implementing performance indicators. However, in terms of school and student improvement, regardless of whether it related to results or behaviour, this finding contradicts those of previous studies, which have suggested that SPE can help schools enhance the quality of their performance (Al Dossary, 2006; Alballawi, 2009; Al Sheikh, 2010; Chapman & Sammons, 2013; Gustafsson et al., 2015). In contrast, this finding is consistent with Penninckx et al. (2016), from the perspective that evaluation alone cannot help schools improve their performance; it must be accompanied by

rewards and accountability to achieve goals that include school improvement and improved student behaviour. Tit is a result that can be examined from various angles.

Firstly, the negative opinions expressed by the head teachers could have resulted from teachers' feelings about SPE, as discussed in the Literature Review for this study. More precisely, what has been confirmed by previous studies, concerning the negative impact of SPE requirements on the core of teachers' work has been cited by Ball (1993) and Perryman (2006). This means that teachers tend to view this kind of assessment as a threat to their identity and a disproportionate preoccupation with assigned tasks that are far from their essential teaching duties.

Additionally, this result may be explained by teachers' performance being negatively affected by SPE. Therefore, their students are unlikely to show improvement. This explanation is supported by Yia and Kimb (2019), who refer to the negative influence of SPE on teaching performance, because teachers experience stressful work and multiple sources of pressure, resulting in greater exposure to stress at work. This point is so serious that it needs to be examined further. If the principles of performance management are likely to have a negative effect on teachers' performance, which could then drive them to abandon their profession or perform poorly under pressure, the question arises of whether neoliberalism is really capable of enhancing the economy by improving educational output, or whether it will make things worse, with education becoming a burden on the state, rather than a tributary of the economy.

6.4.3 Improving the Management of School Performance and Evaluation

The most obvious finding to emerge from the data analysis in this study was that SPIS has helped enhance the management of SPE among head teachers' and teachers. For example, teachers reported that they collated evidence of their work and kept this in a more organised fashion than previously, such as in files, in preparation for the inspection. This corroborates what was reported by Clarke and Ozga (2011), who found that SPE can lead to improved school performance. This result is also associated with a UNESCO (2000) report, referring to improved performance as a result of SPE implementation. However, there are a considerable number of explanations that could be proposed for this finding.

For example, one possible explanation is that determining the strengths and weaknesses of school performance can make it easier to enhance and control it, which may be cited as an

important benefit of SPE. It also conforms other research published in this area such as Tomlinson (2005) and Altrichter and Kemethofer (2015). In addition, Thiel and Bellmann (2017) concur that the evaluation report produced by an evaluation team for a school can reduce workload and save time for staff, while they focus on the reform process.

Moreover, it would seem that this result is due to accountability, although there is no strong evidence of the role of accountability in the Saudi education system, especially in reference to SPIS evaluation. Nevertheless, accountability could be the main factor behind improving school performance. This explanation is consistent with numerous studies that have linked it to enhanced performance (Møller, 2009; Kwok, 2011). It may be because head teachers and teachers are keen to apply criteria and indicators, due to their accountability (Robinson, 2012). Hence, this gives rise to specific questions, such as, 'Is it enough to apply the indicators to enhance school performance?' It would suggest that the current performance indicators should be reviewed to see whether they are aligned with organisational objectives and evolving expectations, thereby strengthening organisational results (Parmenter, 2015). Furthermore, the way in which indicators are applied is more likely to contribute to the success of their impact. It would seem that prior to the introduction of evaluation systems, Saudi schools were traditionally managed in a very different way, namely through review by an evaluation team. The benefits of this evaluation were recognised by the participants in this study.

6.4.4 Poor School Environment and Accountability

One of the points explored in relation to the third research question was the relationship of accountability to the development of performance. The participants refused to answer questions about any shortcomings in the performance of their school, but this might have been unrelated to their unwillingness to accept accountability for their work. Instead, they may have observed that the indicators used to measure school performance make no sense in their school.

However, in the responses to the final question, there was variation across the different school environments. It is illogical to hold teachers accountable for failing to apply indicators that are impossible to implement in their schools, such as a learning strategies indicator that requires a minimum number of students in a class, or the use of specific technology, where there is only a whiteboard. There are several possible explanations for this; one being that it is clearly an effect of inefficient centralisation upon performance management and the entire process of education reform, as mentioned by Fullan (2007). Centralisation also reduces the efficiency of

the programmes being implemented, which has been mentioned previously (Alzaidi, 2008). This then requires the application of indicators in advance of any evaluation to ensure that all sides are fully informed, which is not currently achieved in the central system. It has undoubtedly led to a lack of indicators for the most important requirement, which is the identification of priorities for each school (O'Reilly, 2009).

Another possible explanation is that SPIS evaluators have not been eager to construct an accurate system for recording the fruit of their teacher performance evaluation. This has meant that the development of appropriate indicators for underprivileged or affluent school environments has been overlooked. Thus, no consideration has been given to the applicability of the performance indicators, leading to a poor-quality system. However, most studies have demonstrated that neoliberalism is aimed at enhancing education quality and the adoption of a system of accountability to ensure the quality of education performance (see subsection 3.4.3). Nevertheless, the relationship between accountability, performance appraisal, and development is broad and complex, which means that it requires more detailed attention in this study.

6.5 Summary

This chapter has discussed the quantitative and qualitative results of the current study, obtained from a questionnaire and interviews with the research participants. This research has investigated the influence of SPIS evaluation from the perspective of head teachers and teachers. It also compares these findings with those that appear in the relevant literature and juxtaposes them with existing theory. The results reveal a noticeable overlap between the current study and previous studies in this field, especially in terms of the effect of evaluation by applying performance indicators on the participants' stress, workload and morale. In addition, the results of this study align with those of previous research, concerning the influence of this system on the organisation of school performance. However, they also contradict some work carried out on the impact of SPIS on education quality. Here, the participants expressed their negative perceptions of the change in education quality that has been brought about by the current evaluation system. Moreover, this discussion demonstrated a strong relationship between the emergence of these effects of SPE and the organisational approach adopted by the Saudi Ministry of Education (centralisation). Additionally, a strong relationship was found between these influences and accountability, although this is not clearly stated in Saudi SPIS. Nevertheless, despite the fact that accountability had noticeably caused the participants a degree

of stress, it had also helped them to achieve their goals, in that they feared negative results from the evaluation of their schools. Most interestingly, this study generated certain findings that have not been published by other researchers in the field, such as the participants' reluctance to express their views on specific issues.

Chapter 7: Conclusion

7.1 Introduction

As the final chapter in this thesis, the current chapter will conclude and summarise the study (see section 7.2), highlighting the key findings. It will also endeavour to present answers to the main research questions (see section 7.3). Additionally, the study's original contribution to knowledge will be explained (see section 7.4). The limitations of the conceptual framework and research design will also be discussed (see section 7.5). Finally, recommendations for future research will be made (see section 7.6).

7.2 Summary of the Study

As clarified in Chapter 1 of this thesis, Saudi Arabia is a developing country and it has taken many steps to develop all sectors, including education. In particular, SPE has been instituted in an attempt to reform Saudi education (Al Hakamy, 2008). As discussed in Chapters 2 and 3, SPE has expanded worldwide as a key tool of education reform. Attention has also been paid to the influence of SPE on teachers, in terms of increased stress, workload and morale (Perryman et al., 2011; Ball, 2012; McVeigh, 2015). Moreover, in the Saudi context, school evaluation has drawn the attention of numerous researchers (for example, Alballawi, 2009; Alrwqee, 2012), who have discussed Saudi Arabia's comprehensive evaluation programme in terms of its effectiveness, impact and potential for development or research in the quality assessment of schools (see, for example, Moussa, 2012). These researchers have looked at the influence of SPIS – as an approach to the evaluation of school performance – on head teachers and teachers, and some aspects of their wellbeing, such as stress levels and morale. However, the perceptions and experiences of female head teachers and teachers, regarding school improvement in a girls' secondary school in Jeddah has not yet been considered, although a similar topic has been discussed in Western research; for example, by Perryman et al. (2011) in the UK context. Moreover, Penninckx et al. (2016) looked at the Flemish Inspectorate, but purely in terms of stress and anxiety as a result of SPE's impact on teachers. Finally, Ehren and Visscher (2008) investigated the influence of SPE on school improvement, which is similar to the aim of this current study.

In the present study, my knowledge of Saudi education and the current study findings revealed aspects of the influence of SPE on head teachers and teachers. School performance has not been

discussed before in this context, so this study represents an attempt to fill the gap in the existing literature on what is known about the influence of SPE. The effect of evaluating school performance using performance indicators was therefore investigated, with the aim of exploring and gaining an understanding of the perceptions of head teachers and teachers, regarding school reform in secondary schools in Jeddah, Saudi Arabia (using SPIS as a specific example).

The context of SPE is described in detail in Chapter 2 of this thesis. In the same Chapter, the main theories relating to SPE, such as neoliberalism, performance management and the Saudi evaluation system are also highlighted, iterated further in Chapter 3. This study was developed out of a theoretical framework consisting of three conceptual frameworks, namely the evaluation of school performance, accountability and performativity, explored in Chapter 4 from two main perspectives: theoretical and empirical. These three frameworks were important for understanding and explaining the strength of their relationship with SPIS, which is the type of evaluation applied in this study.

Since SPIS is used to evaluate school performance in schools in Jeddah, it was necessary to understand the form and method of its application, as well as the surrounding discussion. It was also important to understand accountability as a concept, because of its strong link with SPE (Anderson, 2005; O'Neill, 2013). There is an interdependency between accountability and performance evaluation, with performance results being used to keep teachers accountable. As this relationship has an impact on teachers (Perryman et al., 2011), accountability is a key concept in this study. The third concept is performativity, which is equally relevant here, because SPIS uses indicators to evaluate school performance as a means of determining achievement amongst schools, head teachers and teachers – referred to here as performativity (Ball, 2012a). It was therefore considered essential in this study to understand the influence of SPIS and what has already appeared in the literature on the topic of performativity and its influence on head teachers and teachers (see section 3.3).

All of the above were taken into consideration when the research questions were formulated. These research questions sought to identify the perceptions and experiences of head teachers and teachers, concerning three important aspects of school performance and school life, namely school monitoring, the influence of SPIS on the stress levels of head teachers and teachers, as well as on their workload and morale, and the influence of SPIS on school improvement. In order to address these questions, an interpretive paradigm was adopted, because the research was focused on the participants' views and opinions. Additionally, to collect the study data, I

used an explanatory sequential multi-site case study mixed methods data collection design, which included quantitative and qualitative approaches. The details and rationale for the selected methodology are explained in Chapter 4, while an analysis of the collected data is presented in Chapter 5, and a discussion of the results in relation to previous studies may be found in Chapter 6. In the next section, the study findings will be summarised.

7.3 Summary of the Findings

The research questions in this study were concerned with demonstrating head teachers' and teachers' experience and perceptions of the influence of SPIS on school performance. As mentioned earlier (see section 1.2), the SPIS programme for evaluating school performance is one of the main components of Saudi Arabia's attempt to reform education. This reform was conceived of for multiple reasons, such as economic and social factors (see section 1.2). Thus, the study findings highlight Saudi Arabia's attempts to reform education, the SPIS programme and its procedures, and the effects of this type of evaluation on head teachers, teachers and school performance.

Moreover, the results clearly demonstrate consistency in many points, with researchers such as Perryman (2006), Ball (2012b) and other contributors in this field, examining the influence of SPE on head teachers and teachers. The next sections will deal with the findings from each sub-question separately, with the main research question worded as follows: What are head teachers' and teachers' perceptions of the influence of SPIS on school performance?

This question was formulated to find out from immediate stakeholders how their work is influenced by school education reform, where the Ministry of Education implements a new programme of reform. These stakeholders were selected, because they are part of the educational context, with experience that would enable them to refer to its weaknesses, strengths and influence. The findings derived through the sub-questions were diverse, but in general, they raised important points; some relating to the system and the way in which it operates, and some concerning its influence on SPIS implementation, such as centralisation and the system's lack of accountability in the Ministry of Education. Moreover, it refers to the negative influence of SPIS implementation on teachers' wellbeing, in terms of increased stress and workload. Additionally, the findings elaborate on the influence of SPIS implementation in terms of school improvement. However, these points will be explained further in the following paragraphs, in a discussion of each sub-question.

RSQ 1: What are head teachers' and teachers' experiences and perceptions of the influence of SPIS on school monitoring?

This first sub-question enquired into mentoring by SPIS, and it sought to explore the perceptions of head teachers and teachers in this regard. The most important issues addressed in this study include the awareness of SPIS monitoring and the techniques applied amongst key stakeholders in their monitoring of teachers and head teachers. The findings for this sub-question will be outlined here, relating to the monitoring of head teachers and teachers by key stakeholders, the appropriateness of the SPIS monitoring techniques applied, and the level of awareness of SPIS monitoring.

7.3.1 Monitoring by Key Stakeholders

The aim of this question was to explore evaluation activities in schools and whether the Ministry of Education considers the necessary procedures when a new programme or system is implemented. This involves an attempt to eliminate any trace of the old system, while encouraging teachers to focus on the new system and its requirements. A number of points emerged from the study results, such as the high level of observation that was found to take place in schools. According to the responses from the head teachers and teachers, various stakeholders evaluated school performance twice in the same year, and the participants were able to easily distinguish between the two types of evaluation. For instance, one type of monitoring was conducted by school stakeholders, head teachers, and the head teachers' assistants, and this was regarded as SSE. However, it is difficult to consider monitoring by school advisors as self-evaluation, because these advisors are from outside the school. Therefore, it was clear that the schools were evaluated differently over the course of the year: once through SPIS and once by school advisors. It illustrates the lack of co-ordination between schools and the Ministry of Education in Saudi Arabia, especially where teachers and policymakers are concerned. Each initiative works toward the same goals and performs the same tasks, using the same people, which results in a waste of time and money. It also increases teachers' workload, and it could be a reason behind the stress and low morale suffered by teachers in this context. From the review of the relevant literature in this study, new programmes of school evaluation in other countries, such as the UK, have involved discontinuing previous evaluation programmes. For example, in the UK, the government established Ofsted and no other programme of school evaluation is in operation (Ouston et al.,

2017). This study reveals that the opposite is true of Saudi education. Therefore, the next subsection will look at the perceptions of head teachers and teachers, concerning the appropriateness of SPIS monitoring techniques.

7.3.2 Awareness of SPIS Monitoring

The most obvious finding to emerge from the data in this case study was the lack of awareness amongst head teachers and teachers of the SPIS monitoring process, especially in terms of the specific time of the evaluation and the main areas of its focus. Moreover, the findings demonstrate the policies followed by SPIS, which consider teachers to be the second level of contact when approaching head teachers and advisers about applying SPIS indicators. Nevertheless, the application of indicators is one of the jobs of a schoolteacher. Moreover, the teachers demonstrated a lack of information and lack of training; they were unclear about how they should respond to inspection and did not understand KPIs, which showed their poor accountability in this matter. Thus, it could become counter-productive, with the indicators being neglected, or a negative staff culture forming in response to having to meet SPIS requirements. Additionally, unclear information about the system could lead to misunderstanding amongst teachers, who may then proceed to implement it in the wrong way. Hence, the next subsection raises another issue related to SPIS monitoring, namely the appropriateness of SPIS monitoring techniques.

7.3.3 The Appropriateness of SPIS Monitoring Techniques

SPIS is used to evaluate school performance and it consists of applying performance indicators. The participants in this current study were asked about the appropriateness of these indicators, but most were not convinced that they were appropriate monitoring techniques. Conversely, another issue emerged in this regard, consisting of the participants' inability to read performance tables. Additionally, the results relating to whether the indicators reflected the school's aims illustrated that the participants were unsure about the representativeness of these indicators in expressing their school's goals or indicating their school's strengths and weaknesses. Nevertheless, the results were interesting, because they showed that while the system was capable of identifying strengths, it was less certain that it could highlight weaknesses. Most of the results in this case refer to a lack of awareness of the focus of monitoring, or when it should take place. This may be related to the lack of information on the

system among head teachers and teachers. Correspondingly, according to Lowes (2016), if teachers do not trust the system, they will not apply it effectively.

RSQ 2: What are head teachers' and teachers' experiences and perceptions of the influence of SPIS on their stress levels, workload and morale?

The perceptions of head teachers and teachers, regarding the influence of SPIS implementation as a means of evaluating their schools, refers to three important issues relating to their wellbeing: workload, stress and morale. These topics have been discussed across the world by many evaluation experts. However, in the Arab context, to the best of my knowledge, there has been no research conducted on the influence of SPE on head teachers and teachers. Instead, there have been studies on evaluation in general, or on pressure and morale from the perspective of head teachers and teachers. Hence, this renders the current study novel in its field; generating results that are important, because they especially target stress, workload and morale. These results will be summarised below, beginning with workload as an influence of SPIS implementation.

7.3.3.1 Workload

The most important results regarding workload indicated that the head teachers did not reveal their thoughts, perceptions or feelings about workload, while the teachers were ready to discuss this issue. However, this result is not confirmed by other research, which makes it crucial to discuss the underlying reasons, especially as the approach to SPE in KSA is rooted in neoliberalism (see subsection 2.2.2). This has been adopted from the Western context, where the principles of free speech encourage head teachers and teachers to join the discussion on education and educational reform. More specifically, as regards the issue of applying a new programme, which must be developed and improved every year to achieve the goal of evaluating school performance, its advantages and shortcomings must be identified by head teachers and teachers, drawing upon their experience of using the programme over a period of years. This provided that they are given the freedom to express their opinions and feel secure enough to do so. Therefore, it could be said that their lack of voice in some way points to their abandonment of the programme, as they do not express their honest perceptions of its improvement or development. However, this is certainly not what the Saudi Ministry of Education wish for.

Nevertheless, this may not be feasible, if participants are reluctant to talk honestly or freely. Additionally, the results relating to workload demonstrate that teachers are obliged to work harder than head teachers, because of this evaluation. Therefore, they are more prone than head teachers to illness and absenteeism. These results confirm the findings from other studies, which also report on the negative relationship between teachers and SPE, which needs to be resolved by the decision-makers.

7.3.3.2 Stress

The results of this section confirm other findings from previous studies, which have revealed increased stress levels due to SPE (Galton & MacBeath, 2008; Ball, 2012b; Bailey & Colley, 2015). For example, this corresponds to Al-Omari and Wuzynani's (2013) findings that accountability, which uses judgment to hold teachers accountable, increases stress levels. In addition, the results presented in this subsection show that head teachers and teachers are more exposed to stress if they work for a school that is evaluated using SPIS. Aside from this, the results of this present study indicate that teachers have greater exposure than head teachers to stress, which is confirmed by Ramos and Unda (2016), who refer to the link between teaching jobs and increased stress. In fact, a considerable number of studies discuss the relationship between teachers' stress and performance evaluation, such as Galton and MacBeath (2008), Bailey and Colley (2015) and Lightfoot (2016), where the findings reveal the weaknesses of the indicators and their application as a source of stress and increased pressure. Additionally, the above authors consider a poor working environment to be the main reason for the struggle experienced by teachers and for increased stress while applying indicators.

7.3.3.3 Morale

Regarding head teachers' and teachers' morale and SPIS evaluation of school performance, a discrepancy was found between the quantitative and qualitative findings. The quantitative results relating to morale were negative, indicating a low opinion of SPIS evaluation, while the qualitative results were generally positive.

The sampled head teachers and teachers were dissatisfied with the process of applying SPIS, including its administration, recommendations and reports. However, in the research interviews, they cited the benefits that they had gained from SPIS implementation, in terms of organising their work and protecting their achievements. However, this result might not be

accurate, as the participants could merely be cautious in their comments about the system, for fear of being identified. Although I did my best as the researcher to make them comfortable and assured them that their identity would not be revealed, the Ministry's rules are strict and presented a block to them speaking freely. Hence, this indicates the importance of discussing the Ministry's rules and encouraging head teachers and teachers to critique the system. The following paragraphs summarise the findings for the second sub-question, relating to the participants' perceptions.

RSQ3: How do head teachers and teachers describe and understand their accountability under SPIS in relation to school improvement?

The important quantitative result corresponding to this question was obtained from the head teachers and teachers with regard to their willingness to reveal their opinions of education quality as an influence of SPIS evaluation; whereupon the head teachers preferred not to share their views, unlike the teachers. Moreover, the head teachers were more positive about the influence of SPIS on improving their students' behaviour and attendance, while the teachers expressed a negative opinion of this aspect of performance. However, the qualitative results demonstrated that the participants were in agreement over the positive influence of SPIS evaluation on many aspects of school performance, such as the organisation and safeguarding of their work and achievements.

In addition, the findings demonstrate the teachers' objection to being asked about improving school evaluation, at a time when school environments are suffering due to a lack of equipment and inability to apply the programme's indicators. The teachers therefore considered this to be inappropriate and illogical. Besides, the results generated by this research question from the interviews with teachers revealed that the SPIS programme did not aim to make teachers accountable for enhancing evaluation; rather, it aimed to help them improve their work, according to the selected indicators established by the Ministry, in order to achieve their educational goals. It was also found that schools in Jeddah suffer because of a lack of appropriate equipment, which would otherwise enable them to apply SPIS indicators.

7.4 Original Contribution

There are six specific contributions to knowledge made by this study, which are as follows:

i. To the researcher's knowledge, the theoretical contribution of this study lies in its discussion of the influence of performativity, as well as the dark side of applying SPE frameworks that are rooted in neoliberalism and performance indicators, such as increased stress amongst teachers and head teachers, or increased workload. These specific points have not been discussed in the Saudi education literature before. This is despite the enthusiasm to implement OECD, World Bank and other recommendations by international institutions, which are interested in reforming education as a means of economic development; for example, by introducing evaluation and strategies for learning. This key aspect, which also supports previous recommendations surrounding the importance of evaluation, has overlooked teachers' satisfaction, psychological wellbeing and self-confidence. Therefore, this study has attempted to shed light on these issues and ignite debate around them, in order to find a solution to the problems that they present; thereby reflecting positively on education reform and avoiding any negative effects on those who play a major role in the process, especially teachers and head teachers.

ii. Although this study is not primarily intended to discuss the empowerment of women, the fact that all the participants were female represents another important contribution of this study, which is to give female head teachers and teachers a voice to express their opinions and illuminate their experience of education reform in Saudi Arabia. This is especially pertinent, since Saudi women do not play a major role in decision-making in education. It makes it even more crucial to enable them to communicate their views on key issues pertaining to the implementation of change and reform in education.

ii. This study contributes to the debate on the theory-practice gap, relating to the influence of evaluation through performance indicators on head teachers and teachers. This study has revealed new issues associated with the SPIS system: its ambiguity, inconsistency, unpredictability and lack of clarity, with implications for its validity, reliability and trustworthiness. Consequently, head teachers and teachers tend to struggle with the system, leading to increased stress and workload.

iii. Accountability in SPIS is unclear to head teachers and teachers, with no consequences being identified for poor performance. This suggests the need for a clear conceptualisation of the relationship between the application of accountability, and the laying down of a foundation for holding schools, head teachers and teachers accountable

for their performance, and for successful SPE implementation. This would encourage schools and their academic staff to meet performance indicators at a high level of quality, resulting in education reform. It could be considered as a further contribution of this present study, with findings that highlight the importance of accountability in facilitating and enhancing education reform; leading to positive change in the direction of neoliberal goals, and consequently helping to develop the economy and wider society.

iv. SPIS measures performance based on specific indicators. It is a system that is rooted in Western neoliberalism, which requires decentralisation. It is therefore incompatible with centralised systems, because centralised systems direct SPIS according to a top-down hierarchy, beginning with the Ministry of Education, and followed by education administrators, head teachers and finally, teachers. Lower down the information trail, teachers are left uninformed about many important aspects of SPIS, their role within it, and their responsibilities and accountability for meeting KPIs. Improved communication, the provision of appropriate training, and the involvement of head teachers and teachers could overcome many of the problems identified in this study. Another alternative would be to move towards a less centralised system, but this alone would offer no guarantee of information reaching the school, head teachers or teachers.

v. In terms of the context, the present study findings could help many Saudi education organisations and entities, such as the Saudi Ministry of Education, to become more effective by encouraging them to pay more attention to teachers. This would ensure that they were partners in the process, listening to their views and benefiting from their experience within the process of education reform. In addition, the results reveal the weaknesses of SPIS. These need to be addressed, in order to improve and develop the system, so that its negative effects on teachers can be reduced or eliminated, while its positive impact is captured.

7.5 Limitations of the Conceptual Framework and Research Design

The limitations of this study are presented below:

1. A lack of capacity to generalise

The first limitation of this study is its non-generalisability. Due to limited time and resources I selected Jeddah as the study context, because it is the second city in KSA, and

because this is where I used to live and work before starting my PhD. Moreover, I chose a secondary school, because I was previously a secondary school teacher, with solid experience of the system. Thus, the findings of the current study cannot be generalised, but they do provide important findings that should inform any future reviews or development of SPE in Saudi Arabia.

2. Scope of the study and sampling

All the participants were female, because I am a female researcher. In Saudi culture, at the time of data collection (2017), a female researcher was not permitted to make face-to-face contact with boys' schools, male head teachers or male teachers. Therefore, only female head teachers and teachers were selected for this study. In addition, the number of head teachers and teachers interviewed was small, which may have limited the potential to gather extensive information, compared to the scope offered by a large number of participants and their perceptions.

3. Limitations in terms of time, cost, population and instruments

The timeframe and resources for this study were limited. Moreover, just two research instruments were used. However, the implementation of more research tools such as school documentation (SPIS reports) could add credibility to the study. Furthermore, this study focused exclusively on teachers and head teachers in Jeddah, while inspectors were not involved.

7.6 Implications for Practice

7.6.1 Implications for Teachers and Head Teachers

I. The role of teachers and head teachers in reforming education first requires them to believe in the process and to demand the right to participate. The most important aspect of their participation is their right to express their views on every application or programme that they are asked to implement, and to avoid any unjustifiable fear. This could be achieved through written communication with the Minister of Education and programme officials, or via professional social media communities of practice, as a platform for expressing their views. They could then find out from the evaluation team the procedures for a school visit, how and when they will learn about plans to visit their school, and how they can contribute to the programme's development. Their opinions

would enhance the process of education reform, as it would relieve the burden on head teachers and teachers; helping them to solve their problems by empowering them to be agents of change. In addition, it would help improve and reform education, thereby achieving the wider community's aspirations towards better educational outcomes and enhanced skills amongst students. Thus, it would support the community both economically and socially.

II. As teachers predominantly turn to online sources of information about SPIS, the implication is that these resources should be easily accessible, high quality, up to date and supplemented with links to help with queries. Therefore, it is recommended in this study to include content on implications for the Ministry of Education.

III. Education reform requires following up the recommendations of global institutions such as UNESCO, which supports the idea that a teacher's work is not limited to teaching. Thus, teachers must be encouraged to accept change in their profession and job requirements, which include documentation and the implementation of new teaching strategies. This new description of the teaching profession should also be provided to new teachers in their initial teacher-training and orientation programmes.

7.6.2 Implications of SPIS

I.   The implementation of multiple evaluation programmes has had a negative influence on teachers and school performance by adding to their teaching workload. Therefore, the Ministry of Education should ensure that their programme of school evaluation is the only one applied in Saudi schools, and that other previous programmes are abandoned if they have the same mission of evaluating school performance, in order to avoid duplication, confusion and excessive burdens on staff.

II. In terms of mentoring, information about the system and the processes for its application should be consistently distributed across all schools, before the monitoring visits take place.

III. The role of SPIS is not clear for teachers. Therefore, the Organizational Guide and the website should include all necessary information through podcasts/videos and PowerPoint slides, which head teachers and teachers can access and apply to achieve success in the system.

IV. The relationship between SPIS and accountability is unclear. Therefore, schools need more information about this and what is expected with regard to their results after evaluation.

V. The performance table has not yet been published and so teachers are unable to read the report, which consequently affects their motivation to improve their school's performance, as they have no awareness of the strengths or weaknesses of their performance.

7.6.3 Policy Recommendations

The following policy recommendations emerged from this study:

I. The findings provide a knowledge base for school evaluation and assessment in Jeddah. There is a need for this to be built on in other urban areas of Saudi Arabia, so that practices can be compared. This will create a national knowledge base from which lessons can be learnt.

II. Teachers and head teachers should be involved in a regular cycle of reviewing school evaluation and assessment practices in Saudi Arabia, in order to respond to the strong interest in this area and the issues that are evident at national level.

III. There is clearly a need for greater consistency in the way that SPIS is implemented across schools; specifically, the information and notice that staff receive prior to an inspection visit, and the way that the inspection results are subsequently communicated. This would ensure the effective application of the results by staff, parents and other stakeholders. Current practices should be reviewed and improved.

IV. There is great need for continuing professional development amongst teachers and head teachers with regard to SPE. Since many teachers search online for information prior to a school inspection visit, authoritative online resources would serve this purpose well. These would then need to be augmented by face-to-face learning and development opportunities for all staff, which should include training for inspectors.

V. Communication channels were impaired, due to the centralised nature of education in Saudi Arabia. By moving towards a more decentralised model, clearer channels of communication might be created, along with autonomy for school staff to address

deficiencies in the current system. This would allow head teachers, for example, to source training for staff.

VI. The results showed that teachers in schools reported a poor classroom environment, especially in technological aspects and were sometimes expected by SPIS inspectors to use technology not provided by the school (see p.155 for example). Therefore, aspects of SPIS relating to technology need to be reviewed and/or funding for technology made a priority.

VII. The results showed an increase in the burden on teachers, which resulted in many working overtime (see 7.3.3.1). There is consequently a need for the Ministry of Education to find ways to support teachers and head teachers, so that their core work is not affected by school evaluation procedures. One solution would be to appoint classroom assistants to help staff prepare documents, thereby relieving the pressure on teaching staff by assisting them with their duties.

VIII. The results also showed an increase in stress and low morale among head teachers and teachers. This underlines the need for free and confidential staff welfare and counselling services, which would help staff withstand work pressure and ensure that school evaluation procedures do not have a negative effect on their performance.

7.7 Recommendations for Future Research

i. The current study has investigated the perceptions of head teachers and teachers from Jeddah, concerning the influence of SPIS. It would therefore be beneficial to compare the results of this current study with those derived for another region or city, in order to identify any similarities or differences.

ii. In this study, a small sample of 64 head teachers and 109 teachers was used. It would consequently be more useful to conduct a similar study with a larger sample size. Additionally, because of the constraints of Saudi culture, as a female researcher, I was only able to collect data from female participants. Therefore, it would be beneficial to include both genders in the sample in future research.

iii. This study did not involve any SPIS inspectors. Hence, it would be insightful to conduct further research that included the perceptions of SPIS inspectors and decision-makers, as well as the views of teachers and head teachers.

iv. The findings from this study demonstrate the importance of accountability in the improvement of the evaluation process. Therefore, it is highly recommended that this variable should be researched further in future work.

7.8 Summary

As the final chapter of this thesis, the current chapter summarises the study (see section 7.2) and its findings (see section 7.3) and discusses its original contribution to knowledge (see section 7.4). The limitations of the study's conceptual framework (see section 7.5) and research design are illustrated, and some of the implications of this study for practice have been discussed. Finally, some suggestions for future research in this area (see section 7.7) have been made. In this study, there has been an attempt to contribute to Saudi education reform; presenting the views and experiences of head teachers and teachers, as they give voice to their own and the community's hopes and desires for change and development in education. Knowledge of their views will undoubtedly lead to a better understanding of many of their attitudes towards and ideas for improvement to the systems that can be applied to education reform. As eloquently articulated by Locke:

> the improvement of understanding is for two ends: first, our own increase of knowledge; secondly, to enable us to deliver that knowledge to others. (Locke, 1998 , p.208)

References

Abisaab, M. (2015). Dörg Matthias Determann: Historiography in Saudi Arabia: Globalization and the State in the Middle East. *The American Historical Review*, *120*(2), 752-753, https://doi.org/10.1093/ahr/120.2.752

Acharya, A. S., Prakash, A., Saxena, P., & Nigam, A. (2013). Sampling: Why and how of it. *Indian Journal of Medical Specialties, 4*(2), 330-333. doi:https://www.researchgate.net/publication/256446902

Adams, R. (2019). Ofsted under fire in its own survey of teachers' wellbeing. *The Guardian,* 21st July.

Afonso, A. & Aubyn, M. S. (2006). Cross-country efficiency of secondary education provision: A semi-parametric analysis with non-discretionary inputs. *Economic Modelling, 23*(3), 476-491. doi: https://www.sciencedirect.com/science/article/pii/S0264999306000149?via%3Dihub

Ahmed, H. (2019). Effectiveness of minimum quality standards for primary education in low and lower middle-income countries. *K4D Helpdesk Report, 548*, 1-12. Retrieved from http://opendocs.ids.ac.uk/opendocs/handle/123456789/14413

Al Dossary, S. (2006). *The Effectiveness of Comprehensive Evaluation for Schools.* (PhD thesis). Uum Alqura University, Makkah.

Al Essa, A. (2009). *Education Reform in Saudi Arabia between Absence of Political Vision, Apprehension of the Religious Culture and Disability of Educational Management* (2nd ed.). Beirut: Dar AlSakee.

Al Hakamy, A. (2008). *Educational Evaluation and Quality Assurance in Education.* Paper presented at the the 14th annual meeting of the Saudi Society for Educational and Psychological Science, 14th May, AlGasem, Saudi Arabia.

Al Sheikh, A. (2010). *The Role of Comprehensive Evaluation in Improving the Performance of Primary School Principals.* (Master's dissertation). King Khalid University, Saudi Arabia.

Al Sulaimani, A. (2010). *The Importance of Teachers in Integrating ICT into Science Teaching in Intermediate Schools in Saudi Arabia: A Mixed Methods Study.* (PhD thesis). RMIT University, Australia.

Al-Omari, A. A. & Wuzynani, M. M. (2013). Factors influencing Jordanian and Saudi Arabian teacher decisions to pursue the principalship: A comparative study. *School Leadership & Management, 33*(5), 473-485.

Al-Ghadhami, A. (2013). *Neoliberalism.* Casablanca: Al Markaz Al Thagafi.

Al-Sonbol, A. (2008). *Education System in the Kingdom of Saudi Arabia* (1st ed.). Riyadh, Saudi Arabia: Dar Al-Khraiji for Publishing and Distribution.

Alballawi, I. (2009). *Degree of Effectiveness of the Comprehensive Evaluation Programme in Diagnosing School Reality from the Perspective of Headteachers in the Educational Area of Tabuk.* (PhD thesis). Mutah University, Jordan. Retrieved from https://search.emarefa.net/ar/login

Alguhidan, S. (2009). *Reality of Educational Accountability in Girls' Public Schools in Mecca, Its Active Requirements and Application Barriers from the Perspective of Headteachers.* (Master's dissertation). Um Alqura University, Makkah. Retrieved from: https://search.mandumah.com/Record/611654

Alkarni, A. (2015). *Re-visioning the Development of Head Teachers in Jeddah: Seeing Head Teachers' Professional Learning through a New Lens.* (PhD thesis). Newcastle University, Newcastle, UK. Retrieved from http://hdl.handle.net/10443/2888

Alrwqee, M. (2012). *Development of a Comprehensive Evaluation System for the Schools of General Education in the Light of Total Quality Management.* (PhD thesis). Al-Imam University, Riyadh. Retrieved from http://ecat.kfnl.gov.sa:88/ipac20/ipac.jsp?session=VL751R0697363.838764&profile=akfnl&uri=full=3100006@!539659@!1&ri=1&aspect=advanced&menu=search&source=172.27.18.30@!kfnl1256&ipp=20&staffonly=&term=Ø§Ù%C2%84Ø±Ù%C2%88Ù%C2%82Ù%C2%8A+Ø%C2%8C+Ù%C2%85Ø·Ù%C2%84Ù%C2%82+Ø¨Ù%C2%86+Ù%C2%85Ù%C2%82Ø¹Ø¯+Ø¨Ù%C2%86+Ù%C2%85Ø·Ù%C2%84Ù%C2%82&index=.AW&uindex=&aspect=advanced&menu=search&ri=1

Alsubaie, A. & Jones, K. (2017). An overview of the current state of women's leadership in higher education in Saudi Arabia and a proposal for future research directions. *Administrative Sciences, 7*(4), 36. doi: https://doi.org/10.3390/admsci7040036

Altrichter, H. & Kemethofer, D. (2015). Does accountability pressure through school inspections promote school improvement? *An International Journal of Research, Policy and Practice, 26*(1), 32-56. Retrieved from https://doi.org/10.1080/09243453.2014.927369

Alzahrani, A. & Alghamdi, M. (2016). Enhancing education supervision in Saudi Arabia – towards Supervision. *Educational Futures, 7*(3), 1-17.

Alzaidi, A. (2008). Secondary school head teachers' job satisfaction in Saudi Arabia: The results of a mixed methods approach. *Annual Review of Education, Communication & Language Sciences*, *5*. https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=Secondary+school+head+teachers%E2%80%99+job+satisfaction+in+Saudi+Arabia%3A+The+results+of+a+mixed+methods+approach.&btnG=

Anderson, J. A. (2005). *Accountability in Education*. Paris: International Institute for Educational Planning

Angus, L. (2007). Globalisation and the reshaping of teacher professional culture: Do we train competent technicians or informed players in the policy process? In: *Handbook of Teacher Education* (pp. 141-156). Springer.

Apple, M. W. (2004). Creating difference: Neo-liberalism, neo-conservatism and the politics of educational reform. *Educational Policy, 18*(1), 12-44. Retrieved from https://doi.org/10.1177/0895904803260022

Appleton, J. V. (1995). Analysing qualitative interview data: addressing issues of validity and reliability. *Journal of Advanced Nursing*, *22*(5), 993-997.

Au, W. & Hollar, J. (2016). Opting out of the education reform industry. *Monthly Review, 67*(10), 29. Retrieved from https://search.proquest.com/docview/1768932327?pq-origsite=gscholar

Audet, J. & d'Amboise, G. (2001). The multi-site study: An innovative research methodology. *The Qualitative Report, 6*(2), 1-18. Retrieved from https://nsuworks.nova.edu/tqr/vol6/iss2/2/

Avis, J. (2005). Beyond performativity: Reflections on activist professionalism and the labour process in further education. *Journal of Education Policy, 20*(2), 209-222.

Aydin, A., Sarier, Y., & Uysal, S. (2013). The Effect of School Principals' Leadership Styles on Teachers' Organizational Commitment and Job Satisfaction. *Educational Sciences: Theory and Practice, 13*(2), 806-811. Retrieved from http://www.edam.com.tr/estp.asp

Bailey, G. & Colley, H. (2015). 'Learner-centred' assessment policies in further education: Putting teachers' time under pressure. *Journal of Vocational Education & Training, 67*(2), 153-168. Retrieved from https://doi.org/10.1080/13636820.2014.983956

Baldwin, R. E. (2006). *Globalisation: The Great Unbundling(s)*. Retrieved from: file:///C:/Users/hb857025/AppData/Local/Packages/Microsoft.MicrosoftEdge_8wekyb3d8bbwe/TempState/Downloads/Baldwin_06-09-20%20(2).pdf

Ball, S. (2003). The teacher's soul and the terrors of performativity. *Journal of Education Policy, 18*(2), 215-228. Retrieved from https://doi.org/10.1080/0268093022000043065

Ball, S. (2012a). Performativity, commodification and commitment: An I-spy guide to the neoliberal university. *British Journal of Educational Studies,* 60(1), 17-28. Retrieved from https://doi.org/10.1080/00071005.2011.650940

Ball, S. (2012b). *Politics and Policy Making in Education: Explorations in Sociology* (Vol. 33). New York: Routledge.

Ball, S. (2016). Neoliberal education? Confronting the slouching beast. *Policy Futures in Education, 14*(8), 1046-1059. Retrieved from https://doi.org/10.1177/1478210316664259

Ball, S. (2017). *The Education Debate* (2nd ed.). Bristol, UK: Policy Press.

Ball, S., Maguire, M., Braun, A., Perryman, J., & Hoskins, K. (2012). Assessment technologies in schools: 'Deliverology'and the 'play of dominations'. *Research Papers in Education, 27*(5), 513-533. Retrieved from https://doi.org/10.1080/02671522.2010.550012

Baki, R. (2004). Gender-segregated education in Saudi Arabia: Its impact on social norms and the Saudi labor market. *Education Policy Analysis Archives, 12*(28), n28. Retrieved from http://epaa.asu.edu

Baroutsis, A. (2016). Media accounts of school performance: Reinforcing dominant practices of accountability. *Journal of Education Policy,* 31(5), 567-582. Retrieved from https://doi.org/10.1080/02680939.2016.1145253

Barr, J. & Christie, A. (2015). Improving the practice of Value for Money Assessment. *CDI Practice Paper, 12*, 1-9. Retrieved from http://opendocs.ids.ac.uk/opendocs/handle/123456789/5977

Barzanò, G. (2009). Reflecting on English educational accountability. *Italian Journal of Sociology of Education, 1*(3), 189-209. Retrieved from https://ijse.padovauniversitypress.it/2009/3/10

Basit, T. N. (2010). *Conducting Research in Educational Contexts.* New York: Bloomsbury Publishing.

Baxter, P., & Jack, S. (2008). Qualitative case study methodology: Study design and implementation for novice researchers. *The Qualitative Report, 13*(4), 544-559.

Berryhill, J., Linney, J. A., & Fromewick, J. (2009). The Effects of Education Accountability on Teachers: Are Policies Too-Stress Provoking for Their Own Good? *International Journal of Education Policy and Leadership*, *4*(5), 1-14.

Berthoud, R. (2000). *A Measure of Changing Health*. Bristol: PolicyPress.

Bessant, S. E., Robinson, Z. P., & Ormerod, R. M. (2015). Neoliberalism, new public management and the sustainable development agenda of higher education: History, contradictions and synergies. *Environmental Education Research, 21*(3), 417-432. *Retrieved* from https://www.tandfonline.com/doi/full/10.1080/13504622.2014.993933

Betzel, A. (2013). *A Discursive Analysis of Education Policy Documentation under the Conservative Liberal Government in England.* (PhD Thesis, pp. 1-250). Institute of Education, University of London.

Blaikie, N., & Priest, J. (2019). *Designing Social Research: The Logic of Anticipation.* Cambridge: John Wiley & Sons.

Blanche, M. T., Blanche, M. J. T., Durrheim, K., & Painter, D. (eds.). (2006). *Research in Practice: Applied Methods for the Social Sciences.* Cape Town: Juta and Company Ltd.

Bousquet, S. (2012). *Teacher Burnout: Causes, Cures and Prevention*. Retrieved from https://eric.ed.gov/?id=ED534527

Bowe, R., Ball, S. J., & Gold, A. (2017). *Reforming Education and Changing Schools: Case Studies in Policy Sociology.* Routledge.

Bowen, W. H. (2014). *The History of Saudi Arabia.* ABC-CLIO.

Boyne, G. A. (1996). Scale, performance and the new public management: An empirical analysis of local authority services. *Journal of Management Studies, 33*(6), 809-826.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101.

Breakspear, S. (2012). The Policy Impact of PISA: An Exploration of the Normative Effects of International Benchmarking in School System Performance. *OECD Education Working Papers, No. 71.* Retrieved from http://www.oecd.org

Brezicha, K., Bergmark, U., & Mitra, D. L. (2015). One size does not fit all: Differentiating leadership to support teachers in school reform. *Educational Administration Quarterly, 51*(1), 96-132. Retrieved from https://doi.org/10.1177/0013161X14521632

Brion, S. A. (2015). *Teacher Morale.* (PhD thesis). Pennsylvania State University, USA. Retrieved from https://etda.libraries.psu.edu/catalog/24978

Brown, R. (2015). The marketisation of higher education: Issues and ironies. *New Vistas, 1*(1), 4-9. Retrieved from https://repository.uwl.ac.uk/id/eprint/3065/1/The%20marketisation%20of%20Higher%20education.pdf

Bryman, A. (2016). *Social Research Methods*. Oxford: Oxford University Press.

Buchanan, R. (2015). Teacher identity and agency in an era of accountability. *Teachers and Teaching, 21*(6), 700-719. Retrieved from https://doi.org/10.1080/13540602.2015.1044329

Camburn, E. M. & Han, S. W. (2015). Infrastructure for teacher reflection and instructional change: An exploratory study. *Journal of Educational Change, 16*(4), 511-533. Retrieved from https://link.springer.com/article/10.1007/s10833-015-9252-6

Cameron, R. (2010). Mixed methods in VET research: Usage and quality. *International Journal of Training Research, 8*(1), 25-39.

Caprara, G. V., Barbaranelli, C., Steca, P., & Malone, P. S. (2006). Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: A study at the school level. *Journal of School Psychology, 44*(6), 473-490. Retrieved from https://www.sciencedirect.com/science/article/pii/S0022440506000847

Carnoy, M., Khavenson, T., & Ivanova, A. (2015). Using TIMSS and PISA results to inform educational policy: A study of Russia and its neighbours. *A Journal of Comparative and International Education, 45*(2), 248-271. Retrieved from https://doi.org/10.1080/03057925.2013.855002

Carper, B. (1978). Fundamental patterns of knowing in nursing. *ANS, 1*(1), 13-23. Reprinted with permission from Aspen Publishers Inc.

Carr-Hill, R., Rolleston, C., Schendel, R., & Waddington, H. (2018). The effectiveness of school-based decision making in improving educational outcomes: A systematic review. *Journal of Development Effectiveness, 10*(1), 61-94. Retrieved from https://doi.org/10.1080/19439342.2018.1440250

Carter, S., & Henderson, L. (2005). Approaches to qualitative data collection in social science. *Handbook of Health Research Methods: Investigation, Measurement and Analysis, 1,* 215-230.

Carter, W. P. (1998). Updated maximum incremental reactivity scale for regulatory applications. *Oecologia,* 1-11.

Chapman, C. & Sammons, P. (2013). *School Self-Evaluation for School Improvement: What Works and Why?* Reading: CfBT Education Trust.

Cicourel, A. V. (1982). Interviews, surveys, and the problem of ecological validity. *The American Sociologist, 17*(1),11-20.

Clark, D. (2009). The performance and competitive effects of school autonomy. *Journal of Political Economy, 117*(4), 745-783. Retrieved from https://www.journals.uchicago.edu/doi/abs/10.1086/605604

Clarke, J. & Ozga, J. (2011). *Governing by Inspection? Comparing School Inspection in Scotland and England.* Paper presented at the Social Policy Association Conference, 4-6 July, University of Lincoln, UK.

Cohen, L., Manion, L., & Morrison, K. (2013). Action research. In: *Research Methods in Education* (pp. 368-385). Routledge.

Coleman, D. & Jones, P. W. (2004). *The United Nations and Education: Multilateralism, Development and Globalisation.* Routledge.

Corbin, J. & Strauss, A. (2008). Strategies for qualitative data analysis. In: Corbin, J. & Strauss, A. (eds.), *Basics of Qualitative Research*: *Techniques and Procedures for Developing Grounded Theory* (3rd ed., pp. 65-86). Thousand Oaks, CA: SAGE Publications, Inc. doi: 10.4135/9781452230153

Creswell, J. W., & Clark, V. L. P. (2007). *Designing and Conducting Mixed Methods Research.* California: Sage Publications.

Creswell, J. W. (2009). *Research Design: Qualitative and Mixed Methods Approaches.* London; Thousand Oaks: Sage Publications.

Creswell, J. W., Klassen, A. C., Plano Clark, V. L., & Smith, K. C. (2011). Best Practices for Mixed Methods Research in the Health Sciences (pp. 541-545). Bethesda, Maryland: National Institutes of Health.

Creswell, J. W. & Creswell, J. D. (2017). *Research Dsign: Qualitative, Quantitative, and Mixed Methods Approaches*. Sage Publications.

Crotty, M. (1998). *The Foundations of Social Science Research*. New South Wales: Allen & Unwin.

Curristine, T., Lonti, Z., & Joumard, I. (2007). Improving Public Sector Efficiency. *OECD Journal on Budgeting*, *7*(1), 1-41. https://www.oecd-ilibrary.org/content/paper/budget-v7-art6-en

Dangerfield, A. S. (2012). School Self-evaluation in Action: A Case Study in the North-West of England. (Doctoral Dissertation). University of Huddersfield.

Dedering, K., & Müller, S. (2011). School improvement through inspections? First empirical insights from Germany. *Journal of Educational Change*, *12*(3), 301-322. https://rd.springer.com/article/10.1007/s10833-010-9151-9

Dee, T., Wyckoff, J., & Force, C. T. (2013). Incentives, Selection, and Teacher Performance. *National Bureau of Ecomomic Research Working Paper Series*. Retrieved from https://doi.org/10.1002/pam.21818

Deem, R., Hillyard, S., Reed, M., & Reed, M. (2007). *Knowledge, Higher Education, and the New Managerialism: The Changing Management of UK Universities.* Oxford University Press.

Denscombe, M. (2003). *The Good Research Guide: For Small-scale Social Research Projects* (2nd ed.). Maidenhead, UK: Open University Press.

DeNisi, A. S. & Pritchard, R. D. (2006). Performance appraisal, performance management and improving individual performance: A motivational framework. *Management and Organization Review*, 2(2), 253-277.

Denzin, N. K., & Lincoln, Y. S. (2018). *The SAGE Handbook of Qualitative Research* (5th ed.). Sage Publications, Inc.

Dibben, P. & Higgins, P. (2004). New Public Management: marketisation, managerialism and consumerism. In: P. Dibben, G. Wood & I. Roper (eds.), *Contesting Public Sector Reforms, Critical Perspectives, National Debates.* Basingstoke: Palgrave Macmillan.   .

Dickinson, I., Watters, D., Graham, I., Montgomery, P., & Collins, J. (2009). Guide to the assessment of competence and performance in practising surgeons. *ANZ Journal of Surgery*, *79*(3), 198-204.

Dreher, A., Gaston, N., & Martens, P. (2008). Measuring globalisation. *Gauging its Consequences.* New York: Springer.

Drummond, A. (2011). *Effects of Performativity on the Professional Practice of Principal Teachers within a Scottish Secondary School.* (PhD thesis). The University of Edinburgh, Edinburgh, UK. Retrieved from https://era.ed.ac.uk/handle/1842/33807

Dubnick, M. (2005). Accountability and the promise of performance: In search of the mechanisms. *Public Performance & Management Review, 28*(3), 376-417. Retrieved from https://www.tandfonline.com/doi/abs/10.1080/15309576.2005.11051839

Duggett, M. (1998). New developments in public administration: Country report; Citizen's Charter: People's charter in the UK. *International Review of Administrative Sciences, 64*(2), 327-330. Retrieved from https://doi.org/10.1177/002085239806400212

Dunn, W. & Miller, D. Y. (2007). A critique of the new public management and the neo-Weberian state: Advancing a critical theory of administrative reform. *Public Organization Review, 7*(4), 345-358. Retrieved from https://link.springer.com/article/10.1007/s11115-007-0042-3

Easterby-Smith, M., Thorpe, R., & Lowe, A. (1994). T*he Philosophy of Research Design. Improving Educational Management through Research and Consultancy.* London: Paul Chapman.

Ehren, M., Perryman, J., & Shackleton, N. (2015). Setting expectations for good education: How Dutch school inspections drive improvement. *School Effectiveness and School Improvement, 26*(2), 296-327. doi:10.1080/09243453.2014.936472

Ehren, M. & Swanborn, M. (2012). Strategic data use of schools in accountability systems. *School Effectiveness and School Improvement, 23*(2), 257-280. Retrieved from https://www.tandfonline.com/doi/full/10.1080/09243453.2011.652127

Ehren, M. & Visscher, A. (2008). The relationships between school inspections, school characteristics and school improvement. *British Journal of Educational Studies, 56*(2), 205-227. Retrieved from https://www.tandfonline.com/doi/full/10.1111/j.1467-8527.2008.00400.x

Ehren, M. C. & Honingh, M. E. (2011). Risk-based school inspections in the Netherlands: A critical reflection on intended effects and causal mechanisms. *Studies in Educational Evaluation, 37*(4),239-248. Retrieved from https://www.sciencedirect.com/science/article/pii/S0191491X12000065

Elliott, A. (2012). Twenty years inspecting English schools–Ofsted 1992–2012. *Research and Information on State Education (Rise) Review,* 1-4. Retrieved from http://risetrust.org.uk/online-publications/45-rise-review-ofsted-1992-2012

Esan, S. & Hamid, I. (2013). The reality of the school institutional evaluation in the UAE and Oman. *Educational and Social Studies - Egypt, 19*(1), 375-333. Retrieved from 469880/Record/com.mandumah.search://http

ETEC. (2019). [Website]. https://www.eec.gov.sa/

Etikan, I. & Bala, K. (2017). Sampling and sampling methods. *Biometrics Biostatistics International Journal, 5*(6), 00149. Retrieved from https://scholar.google.co.uk/scholar?hl=en&as_sdt=0%2C5&scioq=Durand%2C+Francesca+Examining+the+relationship+between+tEACHERS%E2%80%99PERCEPTIONS+OF+SYSTEM+LEADERSHIP%2C+TEACHER+MORALE+AND+TEACHER+ATTENDANCE+IN+THE+CAPITAL+REGION+OF+NEW+YORK+STATE&q=Sampling+and+sampling+methods++Etikan%2C+Illker+Bala%2C+Kabiru&btnG=

Etikan, I., Musa, S., & Alkassim, R. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical Applied Statistics, 5*(1),1-4. Retrieved from https://scholar.google.co.uk/scholar?hl=en&as_sdt=0%2C5&scioq=Durand%2C+Francesca+Examining+the+relationship+between+tEACHERS%E2%80%99PERCEPTIONS+OF+SYSTEM+LEADERSHIP%2C+TEACHER+MORALE+AND+TEACHER+ATTENDANCE+IN+THE+CAPITAL+REGION+OF+NEW+YORK+STATE&q=Comparison+of+convenience+sampling+and+purposive+sampling&btnG=

Evans, L. (2011). The 'shape' of teacher professionalism in England: Professional standards, performance management, professional development and the changes proposed in the 2010 White Paper. *British Educational Research Journal*, *37*(5), 851-870.

Exley, S. & Ball, S. J. (2013). Neo-liberalism and English education. In: Turner, D. & Yolcu, H. (eds.), *Neo-liberal Educational Reforms* (pp. 31-49). Routledge.

Fattore, G., Dubois, H. F., & Lapenta, A. (2012). Measuring new public management and governance in political debate. *Public Administration Review*, *72*(2), 218-227.

Ferlie, E., Hartley, J., & Martin, S. (2003). Changing public service organizations: Current perspectives and future prospects. *British Journal of Management, 14*, S1-S14. Retrieved from https://doi.org/10.1111/j.1467-8551.2003.00389.x

Fitz-Gibbon & Taylor, C. (1990). *Performance Indicators* (Vol. 2, Multilingual Matters). Bristol, UK: Channel View Publications Ltd.

Fletcher, C. (2004). *Appraisal and Feedback: Making Performance Review Work.* CIPD Publishing.

Flick, U. (Ed.) (2009). *The Sage Qualitative Research Kit: Collection*. SAGE Publications Limited.

Flores, M. A., & Derrington, M. L. (2017). School principals' views of teacher evaluation policy: Lessons learned from two empirical studies. *International Journal of Leadership in Education*, 20(4), 416-431.

Franco-Santos, M., Rivera, P., & Bourne, M. (2014). Performance management in UK HEIs: The need for a hybrid approach. *Research and Development Series* (4th ed.). London: Leadership Foundation for Higher Education.

Fullan, M. (2007). *Leading in a Culture of Change*. San Francisco: John Wiley & Sons.

Fusarelli, L. D., & Johnson, B. (2004). Educational governance and the new public management. *Public Administration and Management: An Interactive Journal*, 9(2), 118-127.

Gaertner, H., Wurster, S., & Pant, H. A. (2014). The effect of school inspections on school improvement. *School Effectiveness and School Improvement,* 25(4), 489-508. Retrieved from https://doi.org/10.1080/09243453.2013.811089

Galiani, S., Gertler, P., & Schargrodsky, E. (2008). School decentralization: Helping the good get better, but leaving the poor behind. *Journal of Public Economics*, 92(10-11), 2106-2120.

Gall, M. D., Borg, W. R., & Gall, J. P. (1996). *Educational Research: An Introduction*. Longman Publishing: NewYork.

Galton, M. & MacBeath, J. (2008). *Teachers under Pressure*. New York: Sage.

General Authority for Statistics. (2018). [Website]. Retrieved from https://www.stats.gov.sa/en

General Commission for Surveys (2019). *Map of Saudi Arabia*. Retrieved from https://www.gcs.gov.sa/En/Pages/default.aspx

General Education Administration of Jeddah (2019). (Document in Arabic). Retrieved from https://edu.moe.gov.sa/jeddah/DocumentCentre/Pages/default.aspx?DocId=9e01d8a3-9db3-4d88-8b16-f68ea04a1dd5

George, S. (1999). *A Short History of Neoliberalism.* Paper presented at the Conference on Economic Sovereignty in a Globalising World, 24-26 March, Bangkok, Thailand.

Gertler, P., Patrinos, H., & Codina, M. R. (2007). *Impact Evaluation for School Based Management Reforms.* Washington DC: World Bank.

Goldberg, A. E. & Allen, K. R. (2015). Communicating qualitative research: Some practical guideposts for scholars. *Journal of Marriage and Family*, 77(1), 3-22.

Goldstein, A. (2015). *Teachers' Perceptions of the Influence of Teacher Collaboration on Teacher Morale.* (Doctoral Thesis). Walden University, USA.

Goldstein, H. & Thomas, S. (1996). Using examination results as indicators of school and college performance. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 159(1), 149-163.

Gorard, S. & Taylor, C. (2004). *Combining Methods in Educational and Social Research.* McGraw-Hill Education (UK).

Govindarajan, K. (2012). Teachers' morale. *International Journal of Behavioral Social and Movement Sciences*, 1(2), 57-61.

Gray, M. (2004). *Geodiversity: Valuing and Conserving Abiotic Nature*. England: John Wiley & Sons.

Grix, J. (2002). Introducing students to the generic terminology of social research. *Politics,* 22(3), 175-186.

Guba, E. G. (1990). *The Paradigm Dialog.* Thousand Oaks; San Francisco, CA: SAGE Publications, Inc.

Guercini, S., Raich, M., Müller, J., & Abfalter, D. (2014). Hybrid analysis of textual data. *Management Decision*, 52(4), 737-754. https://doi.org/10.1108/MD-03-2012-0247

Gustafsson, J. E., Ehren, M. C. M., Conyngham, G., McNamara, G., Altrichter, H., & O'Hara, J. (2015). From inspection to quality: Ways in which school inspection influences change in schools. S*tudies in Educational Evaluation,* 47, 47-57.

Guzzo, R. A. & Dickson, M. W. (1996). Teams in organizations: Recent research on performance and effectiveness. *Annual Review of Psychology*, 47(1), 307-338.

Hallinger, P. & Lee, M. (2011). A decade of education reform in Thailand: Broken promise or impossible dream? *Cambridge Journal of Education*, 41(2), 139-158.

Hamdan, A. (2005). Women and education in Saudi Arabia: Challenges and achievements. *International Education Journal*, 6(1), 42-64.

Hammersley, M. (2008). Assessing validity in social research. In: Alasuutari, P., Bickman, L., & Brannen, J. (eds.), *The SAGE Handbook of Social Research Methods* (pp. 42-53). Sage Publications Ltd.

Hartley, D. (1997). The new managerialism in education: A mission impossible? *Cambridge Journal of Education*, 27(1), 47-57.

Harvey, D. (2005). *The New Imperialism.* Oxford: Oxford University Press.

Harvey, D. (2007). Neoliberalism as Creative Destruction. *The ANNALS of the American Academy of Political and Social Science, 610*(1), 21-44. doi:10.1177/0002716206296780

Herbert, D. T. & Thomas, C. J. (1998). School performance, league tables and social geography. *Applied Geography*, 18(3), 199-223.

Hill, H. & Andrews, M. (2005). Reforming budget ritual and budget practice: The case of performance management implementation in Virginia. *International Journal of Public Administration*, 28(3-4), 255-272.

Hinkle, D. E., Wiersma, W., & Jurs, S. G. (2003). A*pplied Statistics for the Behavioral Sciences* (Vol. 663). Houghton Mifflin College Division.

Hood, C. & Peters, G. (2004). The middle aging of new public management: Into the age of paradox?. *Journal of Public Administration Research and Theory*, *14*(3), 267-282.

Huberman, M. & Miles, M. B. (2002). *The Qualitative Researcher's Companion*. Sage Publications Ltd.

Huggins, R. & Williams, N. (2011). Entrepreneurship and regional competitiveness: The role and progression of policy. *Entrepreneurship & Regional Development: An International Journal,* 23(9-10), 907-932.

Hughes, O. & Teicher, J. (2004). *Institutional Requirements for New Public Management in Developing Countries.* Paper presented at the Eighth International Research Symposium on Public Management, 31 March-2 April, Budapest, Hungary.

Ingersoll, R., Merrill, L., & May, H. (2016). Do accountability policies push teachers out? *Educational Leadership*, 73(8), 44-49.

Isaac, R. G., Zerbe, W. J., & Pitt, D. C. (2001). Leadership and motivation: The effective application of expectancy theory. *Journal of Managerial Issues,* 13(2)*,* 212-226. Retrieved from https://www.jstor.org/stable/40604345?seq=1#metadata_info_tab_contents

Ivankova, N. V., Creswell, J. W., & Stick, S. L. (2006). Using mixed-methods sequential explanatory design: From theory to practice. *Field Methods, 18*(1), 3-20. Retrieved from https://journals.sagepub.com/doi/abs/10.1177/1525822x05282260

Jaksic, M. & Jaksic, M. (2013). Performance management and employee satisfaction. *Montenegrin Journal of Economics*, 9(1), 85.

Janssens, F. J. & van Amelsvoort, G. H. (2008). School self-evaluations and school inspections in Europe: An exploratory study. *Studies in Educational Evaluation*, 34(1), 15-23.

Jaradin, N. (2004). *The Performance of the General Secondary School Principals in Jordan in Relation to the Level of Stress and the Use of Their Dealing Skills.* (PhD thesis). Amman Arab University, Jordan.

Jasper, M. A. (1994). Issues in phenomenology for researchers of nursing. *Journal of Advanced Nursing*, 19(2), 309-314.

Javadi, S. (2013). *Performance Management in Higher Education: A Grounded Theory Study.* University of Southampton, UK. Retrieved from https://eprints.soton.ac.uk/378174/

Jiang, J. Y., Sporte, S. E., & Luppescu, S. (2015). Teacher perspectives on evaluation reform: Chicago's REACH students. *Educational Researcher*, 44(2), 105-116.

Johnsen, Å. (2005). What does 25 years of experience tell us about the state of performance measurement in public policy and management? *Public Money and Management*, 25(1), 9-17.

Johnson, D. G. & Regan, P. M. (eds.) (2014). *Transparency and Surveillance as Sociotechnical Accountability: A House of Mirrors.* London: Routledge.

Johnson, R. B. & Christensen, L. (2019). *Educational Research: Quantitative, Qualitative, and Mixed Approaches.* SAGE Publications, Inc.

Jones, K. & Tymms, P. (2014). Ofsted's role in promoting school improvement: The mechanisms of the school inspection system in England. *Oxford Review of Education*, 40(3), 315-330.

Kajornboon, A. B. (2005). Using interviews as research instruments. *E-journal for Research Teachers,* 2(1), 1-9.

Kalimullah, N. A., Alam, K. M. A., & Nour, M. A. (2012). New public management: Emergence and principles. *Bup Journal*, 1(1), 1-22.

Kaplan, R. M. and Maxwell III, J. T. (1994). Text-compression technique using frequency-ordered array of word-number mappers, (June 28), US Patent 5,325,091.

Kaplinsky, R. & Morris, M. (2016). Thinning and thickening: Productive sector policies in the era of global value chains. *The European Journal of Development Research,* 28(4), 625-645. Retrieved from https://link.springer.com/article/10.1057/ejdr.2015.29

Kim, K., Harris, C. J., & Pham, L. (2018). How Character Education Impacts Teachers. *International Journal of Multidisciplinary Perspectives in Higher Education*, 3(1).

Koppes, L. L. (Ed.). (2014). *Historical Perspectives in Industrial and Organizational Psychology.* New York: Psychology Press.

Kvale, S. (2008). *Doing Interviews*. London: Sage.

Kwok, P. K. P. (2011). *Principals' Perception of Quality and Accountability: A Case Study of Lasallian Schools in Hong Kong.* (Doctoral thesis). University of Leicester, UK.

Larsen, E. (2009). *School Accountability and AdministratorIincentives in California.* Public Policy Institute of California.

Latham, G. P. & Locke, E. A. (2007). New developments in and directions for goal-setting research. *European Psychologist,* 12(4), 290-300. Retrieved from https://econtent.hogrefe.com/doi/abs/10.1027/1016-9040.12.4.290

Lebas, M. J. (1995). Performance measurement and performance management. *International Journal of Production Economics,* 41(1-3), 23-35. Retrieved from https://www.sciencedirect.com/science/article/pii/092552739500081X

Leckie, G. & Goldstein, H. (2009). The limitations of using school league tables to inform school choice. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 172*(4), 835-851. Retrieved from https://rss.onlinelibrary.wiley.com/doi/full/10.1111/j.1467-985X.2009.00597.x

Leithwood, K., Jantzi, D., Silins, H., & Dart, B. (1991). Using the appraisal of school leaders as an instrument for school restructuring. *Peabody Journal of Education,* 68(2), 85-109.

Li, D. & Sullivan, W. C. (2016). Impact of views to school landscapes on recovery from stress and mental fatigue. *Landscape and Urban Planning*, 148, 149-158.

Lichtman, M. (2012). *Qualitative Research in Education: A User's Guide.* London, UK: Sage.

Lightfoot, M. D. (2014). *Education Reform for the Knowledge Economy in the Middle East: A Study of Education Policy Making and Enactment in the Kingdom of Bahrain.* (PhD thesis). UCL Institute of Education. Retrieved from https://discovery.ucl.ac.uk/id/eprint/10021657/

Lincoln, Y. S., & Guba, E. G. (1985). Establishing trustworthiness. *Naturalistic inquiry*, 289(331), 289-327.

Lingard, B. & Rizvi, F. (2010). *Globalizing Education Policy*. New York: Routledge.

Littleford, A. R. (2007). *Principal Leadership and Its Perceived Influence on Teacher Morale in Elementary Schools.* (PhD thesis). East Tennessee State University.

Liu, J. (2015). Globalizing indigenous psychology: An East Asian form of hierarchical relationalism with worldwide implications. *Journal for the Theory of Social Behaviour, 45*(1), 82-94. Retrieved from https://onlinelibrary.wiley.com/doi/full/10.1111/jtsb.12058

Lloyd, R. & Mertens, D. (2018). Expecting more out of expectancy theory: History urges inclusion of the social context. *International Management Review,* 14(1), 28-43.

Locke, J. (1996). *Some Thoughts Concerning Education: And, of the Conduct of the Understanding*. USA: Hackett Publishing.

Locke, J. (1998). *An Essay Concerning Human Understanding*. Oxford: Oxford University Press

Lockheed, M. E. & Hanushek, E. J. C. (1988). Improving educational efficiency in developing countries: What do we know? *Compare: A Journal of Comparative and International Education,* 18(1), 21-38.

Louis, C., Lawrence, M., & Keith, M. (2007). *Research Methods in Education* (7th ed.). New York: Routledge.

Lowes, T. D. (2016). *To What Extent Do Ofsted Inspectors' Values Influence the Inspection Process (2005-2012)?: An Examination of Ofsted Inspectors' Perceptions.* (Doctoral dissertation). University of Hull.

Lunenburg, F. C. (2011). Goal-setting theory of motivation. *International Journal of Management, Business, and Administration*, 15(1), 1-6.

Lynch, K., Grummell, B., & Devine, D. (2012). *New Managerialism in Education: Commercialization, Carelessness and Gender.* Springer.

Lyotard, J.-F. (1984). *The Postmodern Condition: A Report on Knowledge* (Vol. 10). University of Minnesota Press.

Mackie, R. (2008). *Organisational Performance Management in a Government Context: A Literature Review*. Scottish Government Social Research.

Major, C. & Savin-Baden, M. (2010). Exploring the relevance of qualitative research synthesis to higher education research and practice. *London Review of Education,* 8(2), 127-140.

Mallick, K. & Verma, G. (2005). *Researching Education: Perspectives and Techniques*. Oxford Routledge.

Manasa, K. V. L. & Reddy, N. (2009). Role of training in improving performance. *The IuP Journal of Soft Skills,* 3(3), 72-80.

Mansour, N., Heba, E. D., Alshamrani, S., & Aldahmash, A. (2014). Rethinking the theory and practice of continuing professional development: Science teachers' perspectives. *Research in Science Education*, 44(6), 949-973.

Maroun, N., Samman, H., Moujaes, C. N., Abouchakra, R., & Ihrke, D. M. (2008). How to succeed at education reform: The case for Saudi Arabia and the broader GCC region. *Abu Dhabi, Ideation Center, Booz & Company, 109*, 113. Retrieved from https://scholar.google.com/scholar?hl=en&as_sdt=0,5&q=How+to+succeed+at+education+reform%3A+The+case+for+Saudi+Arabia+and+the+broader+GCC+region+Maroun,+Rabih+Samman,+Hatem+Moujaes,+Chadi+N+Abouchakra,+Rabih+Insight,+Ideation+Center

Mayer, D., Mitchell, J., Santoro, N., & White, S. (2011). Teacher educators and 'accidental' careers in academe: An Australian perspective. *Journal of Education for Teaching, 37*(3), 247-260. Retrieved from https://www.tandfonline.com/doi/full/10.1080/02607476.2011.588011

Mayston, D. J. (1985). Non-profit performance indicators in the public sector. *Financial Accountability & Management,* 1(1), 51-74. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-0408.1985.tb00244.x

McCallum, A. F. (2018). *Teachers' Constructions of Creativity in Secondary English: Who Gets to Be Creative in Class?* (Doctoral thesis). London Metropolitan University.

McDonald, P., Pini, B., & Bartlett, J. (2019). The emergence of the marketing professional in schools. *British Journal of Sociology of Education*, 33(1), 1-20. DOI: 10.1080/01425692.2012.632864

McVeigh, H. (2016). *An Exploration of the Development of the Ofsted Criteria for Evaluating the Quality of Teaching in Mainstream Schools and How the Criteria Are Perceived and Used by Primary School Headteachers and Teachers.* (Doctoral thesis). University of Roehampton, UK.

Mertler, C. & Charles, C. (2005). *Introduction to Educational Research in Education.* San Francisco: Jossey-Bass.

Miller, P. W., Ochs, K., & Mulvaney, G. (2008). International teacher migration and the Commonwealth Teacher Recruitment Protocol: Assessing its impact and the implementation process in the United Kingdom. *European Education*, 40(3), 89-101.

Miner, J. B. (2015). Goal-setting Theory: Edwin Locke and Gary Latham. In: *Organizational Behavior* 1 (pp. 175-199). Routledge.

Ministry of Civil Service. (2019). [Website]. Retrieved from https://www.mcs.gov.sa/Pages/Default.aspx

Ministry of Economy and Planning. (2015). [Website]. http://www.mep.gov.sa/ar

Ministry of Education. (2017). *Statistical Manual.* Retrieved from https://edu.moe.gov.sa/jeddah/About/Pages/Statistics.aspx

Ministry of Education. (2019). [Website]. https://www.moe.gov.sa/en/pages/default.aspx

Moeller, A. J., Theiler, J. M., & Wu, C. (2012). Goal setting and student achievement: A longitudinal study. *The Modern Language Journal*, 96(2), 153-169.

Møller, J. (2009). School leadership in an age of accountability: Tensions between managerial and professional accountability. *Journal of Educational Change*, *10*(1), 37-46.

Moon, K., Brewer, T. D., Januchowski-Hartley, S. R., Adams, V. M., & Blackman, D. A. (2016). A guideline to improve qualitative social science publishing in ecology and conservation journals. *Ecology and Society*, 21(3).

Morel, N., Palier, B., & Palme, J. (2009). *What Future for Social Investment?* Retrieved from https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=What+future+for+social+investment%3FMorel%2C+Nathalie+Palier%2C+Bruno+Palme%2C+Joakim&btnG=

Moreton, H. J. (2015). *Headteacher Inspectors: Boundaries, Identity and the Potential for System Leadership.* (Doctoral thesis). Sheffield Hallam University, UK.

Morrison, C. (2012). *A Study of Policy Change and Stress amongst Teachers in Scottish Schools* (Doctoral dissertation). University of Strathclyde.

Moses, I., Berry, A., Saab, N., & Admiraal, W. (2017). Who wants to become a teacher? Typology of student-teachers' commitment to teaching. *Journal of Education for Teaching*, 43(4), 444-457.

Mousavi, E. S. (2007). Exploring 'teacher stress' in non-native and native teachers of EFL. *English Language Teacher Education and Development, 10,* 33-41. Retrieved from https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=Exploring+%E2%80%98teacher+stress%E2%80%99+in+non-native+and+native+teachers+of+EFL&btnG=

Moussa, N. (2012). Patterns of Evaluating the Performance of Private Schools (Saudi Arabia). *Altourath, 2*(11), 310-344. Retrieved from 363374/Record/com.mandumah.search

Moye, M. J., Henkin, A. B., & Egley, R. J. (2005). Teacher-principal relationships: Exploring linkages between empowerment and interpersonal trust. *Journal of Educational Administration*, 43(3), 260-277.

Mukherjee, M. (2015). From missionary to new middle-class schooling in the era of global capitalism. In: Sprague, J. (ed.), *Globalization and Transnational Capitalism in Asia and Oceania* (p. 90). Taylor & Francis Group.

Mulholland, R., McKinlay, A., & Sproule, J. (2013). Teacher interrupted: Work stress, strain, and teaching role. *Sage Open*, 3(3), 2158244013500965.

Mullins, L. J. (2007). *Management and Organisational Behaviour*. Edinburgh: Pearson Education.

Mundy, K., Green, A., Lingard, B., & Verger, A. (2016). *Handbook of Global Education Policy*. West Sussex: John Wiley & Sons.

Murray, J. (2012). Performativity cultures and their effects on teacher educators' work. *Research in Teacher Education, 2(2)*, 19-23. Cass School of Education and Communities, University of East London, London.

Naghieh, A., Montgomery, P., Bonell, C. P., Thompson, M., & Aber, J. L. (2015). Organisational interventions for improving wellbeing and reducing work-related stress in teachers. *Cochrane Database of Systematic Reviews,* (4). Retrieved from https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD010306.pub2/abstract

Naidoo, R. & Jamieson, I. (2005). Empowering participants or corroding learning? Towards a research agenda on the impact of student consumerism in higher education. *Journal of Education Policy, 20*(3), 267-281. Retrieved from https://www.tandfonline.com/doi/full/10.1080/02680930500108585

Nevo, D. (2002). Dialogue evaluation: Combining internal and external evaluation. In: *School-based Evaluation: An International Perspective* (pp. 3-16). Emerald Group Publishing Limited.

Newby, P. (2014). *Research Methods for Education.* Routledge.

Newman, D., Karimi, S., & Cavedon, L. (2009, December). External evaluation of topic models. In: Australasian ocument Computing Symposium, 4 December, 2009, Sydney, Australia.

NHS. (2019). Struggling with stress?. *NHSInform*. Retrieved from https://www.nhsinform.scot/healthy-living/mental-wellbeing/stress/struggling-with-stress

Nielsen, P. A. (2013). Performance management, managerial authority, and public service performance. *Journal of Public Administration Research and Theory*, 24(2), 431-458.

Nuttall, D. L. (2017). Choosing indicators: United Kingdom and International Perspectives. In: Riley, K. A. & Nuttall, D. L. (eds.), *Measuring Quality: Education Indicators* (pp. 17-40). London: Routledge.

O'Neill, O. (2013). Intelligent accountability in education. *Oxford Review of Education, 39*(1), 4-16. Retrieved from https://www.tandfonline.com/doi/full/10.1080/03054985.2013.764761

O'Reilly, T. (2009). *What is Web 2.0.* O'Reilly Media, Inc.

OECD. (2013). *OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes.* Retrieved from http://www.oecd.org/education/school/oecdreviewonevaluationandassessmentframeworksforimprovingschooloutcomes.htm

Omega Schools. (2019). [Website]. Retrieved from https://www.omega-schools.com/foundation-overview.php

Onwuegbuzie, A. J. & Johnson, R. B. (2006). The validity issue in mixed research. *Research in the Schools,* 13(1), 48-63.

Ouston, J., Earley, P., & Fidler, B. (2017). *OFSTED Inspections: The Early Experience.* Routledge.

Overeem, P. & Tholen, B. (2011). After managerialism: MacIntyre's lessons for the study of public administration. *Administration Society,* 43(7), 722-748. Retrieved from https://journals.sagepub.com/doi/full/10.1177/0095399711413728

Oyaid, A. (2009). *Education Policy in Saudi Arabia and Its Relation to Secondary School Teachers' ICT Use, Perceptions, and Views of the Future of ICT in Education.* (Doctoral Thesis). University of Exeter, Exeter, UK.

Parameshwaran, M. & Thomson, D. (2015). The impact of accountability reforms on the Key Stage 4 curriculum: How have changes to school and college Performance Tables affected pupil access to qualifications and subjects in secondary schools in England? *London Review of Education,* 13(2), 157-173. Retrieved from https://www.ingentaconnect.com/content/ioep/clre/2015/00000013/00000002/art00013

Parijat, P. & Bagga, S. (2014). Victor Vroom's expectancy theory of motivation–An evaluation. *International Research Journal of Business Management, 7*(9), 1-8.

Parmenter, D. (2015). *Key Performance Indicators: Developing, Implementing, and Using Winning KPIs.* John Wiley & Sons.

Passmore, D. L. & Baker, R. M. (2005). Sampling strategies and power analysis. *Research in Organizations: Foundations and Methods of Inquiry* (pp. 45-56). Berrett-Koehler.

Pease-Alvarez, L., Samway, K. D., & Cifka-Herrera, C. (2010). Working within the system: Teachers of English learners negotiating a literacy instruction mandate. *Language Policy, 9*(4), 313-334. Retrieved from https://link.springer.com/article/10.1007/s10993-010-9180-5

Penninckx, M., Vanhoof, J., De Maeyer, S., & Van Petegem, P. (2016). Effects and side effects of Flemish school inspection. *Educational Management Administration, Leadership,* 44(5), 728-744.

Perryman, J. (2006). Panoptic performativity and school inspection regimes: Disciplinary mechanisms and life under special measures. *Journal of Education Policy,* 21(2), 147-161. Retrieved from https://www.tandfonline.com/doi/full/10.1080/02680930500500138

Perryman, J., Ball, S., Maguire, M., & Braun, A. (2011). Life in the pressure cooker – School league tables and English and mathematics teachers' responses to accountability in a results-driven era. *British Journal of Educational Studies,* 59(2), 179-195. Retrieved from https://www.tandfonline.com/doi/full/10.1080/00071005.2011.578568

Perryman, J., Maguire, M., Braun, A., & Ball, S. (2018). Surveillance, Governmentality and moving the goalposts: The influence of Ofsted on the work of schools in a post-panoptic era. *British Journal of Educational Studies,* 66(2), 145-163. Retrieved from https://www.tandfonline.com/doi/full/10.1080/00071005.2017.1372560

Pieterse, J. N. (2019). *Globalization and Culture: Global Mélange.* Rowman & Littlefield Publishers.

Pietarinen, J., Soini, T., & Pyhältö, K. (2014). Students' emotional and cognitive engagement as the determinants of well-being and achievement in school. *International Journal of Educational Research,* 67, 40-51.

Plowright, D. (2007). Self-evaluation and Ofsted inspection: Developing an integrative model of school improvement. *Educational Management Administration & Leadership,* 35(3), 373-393.

Pollitt, C. & Bouckaert, G. (2004). *Public Management Reform: A Comparative Analysis.* USA: Oxford University Press.

Pollitt, C. & Dan, S. (2011). The impacts of the New Public Management in Europe: A meta-analysis. *Coordinating for Cohesion in the Public Sector of the Future* (pp. 1-66). Retrieved from https://scholar.google.com/scholar?hl=en&as_sdt=0,5&q=The+impacts+of+the+New+Public+Management+in+Europe%3A+A+meta-analysis+Pollitt,+Christopher+Dan,+Sorin

Pollock, K. & Winton, S. (2016). Juggling multiple accountability systems: How three principals manage these tensions in Ontario, Canada. *Educational Assessment, Evaluation Accountability,* 28(4), 323-345. Retrieved from https://link.springer.com/article/10.1007/s11092-015-9224-7

Poole, S. M. (2011). The relationship between external accountability policy and internal accountability: A cross-state analysis of charter and traditional public schools. *Journal of School Choice,* 5(3), 261-280. Retrieved from https://www.tandfonline.com/doi/full/10.1080/15582159.2011.604225

Priestley, M., Miller, K., Barrett, L., & Wallace, C. (2011). Teacher learning communities and educational change in Scotland: The Highland experience. *British Educational Research Journal,* 37(2), 265-284. Retrieved from https://www.tandfonline.com/doi/full/10.1080/01411920903540698

Proudfoot, K. I. (2018). *How Does Neoliberal Performance Management Affect Teachers'* Perceived *Motivations to 'Improve'?* (Doctoral thesis). Lancaster University, UK.

Pulakos, E. D. (2004). *Performance Management: A Roadmap for Developing, Implementing and Evaluating Performance Management Systems.* SHRM Foundation.

Orlikowski, W. J. & Baroudi, J. J. (1991). Studying information technology in organizations: Research approaches and assumptions. *Information Systems Research,* 2(1), 1-28.

Ramos, V. & Unda, X. L. (2016). Work stress and organizational climate in an educational context: A comparison study between teachers and support staff. *Escuela Politécnica Nacional.* Retrieved from https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=Work+stress+and+organizational+climate+in+an+educational+context%3A+A+comparison+study+between+teachers+and+support+staff&btnG=

Randle, K. & Brady, N. (1997). Further education and the new managerialism. *Journal of Further and Higher Education,* 21(2), 229-239. Retrieved from https://www.tandfonline.com/doi/abs/10.1080/0309877970210208

Ravitch, D. (2016). *The Death and Life of the Great American School System: How Testing and Choice are Undermining Education.* Basic Books.

Regmi, K. D. (2017). World Bank in Nepal's education: Three decades of neoliberal reform. *Globalisation, Societies, Education,* 15(2), 188-201. Retrieved from https://www.tandfonline.com/doi/full/10.1080/14767724.2016.1169517

Reid, K. (2010). Teachers plead against cutbacks. *The Record.* Retrieved from from http://www.recordnet.com/apps/pbcs.dll/article? AID=/20100331/A_NEWS/3310319/-1/a_special0258

Richardson, J. (2015). Accountability incentives and academic achievement: Distributional impacts of accountability when standards are set low. *Economics of Education Review,* 44, 1-16. Retrieved from https://www.sciencedirect.com/science/article/pii/S0272775714000983

Ritzer, G. (2010). *Enchanting a Disenchanted World: Continuity and Change in the Cathedrals of Consumption*. Pine Forge Press.

Ritzer, G. & Dean, P. (2015). *Globalization: A Basic Text.* John Wiley & Sons.

Roberts, N. & Abreu, L. (2016). *School Inspections in England: Ofsted* (Briefing Paper 07091). London: House of Commons Library.

Robertson, S. L. (2008). 'Remaking the World' Neoliberalism and the Transformation of Education and Teachers' Labor. In: Compton, M. & Weiner L. (eds.), *The Global Assault on Teaching, Teachers, and Their Unions' Stories for Resistance* (pp. 11-27). New York: Palgrave Macmillan.

Robson, C. (2002). *Real World Research: A Resource for Social Scientists and Practitioner-researchers* (2nd ed.). Oxford, UK: Blackwell Publishers.

Rockoff, J. & Turner, L. J. (2010). Short-run impacts of accountability on school quality. *American Economic Journal: Economic Policy*, 2(4), 119-47.

Rothstein, J. (2010). Teacher quality in educational production: Tracking, decay, and student achievement. *The Quarterly Journal of Economics, 125*(1), 175-214. Retrieved from https://academic.oup.com/qje/article/125/1/175/1880323

Rowe, K. (2004). In good hands? The importance of teacher quality. *The National Newspaper for All Non-government Schools,* (149), 4. Retrieved from https://search.informit.com.au/documentSummary;dn=016042395466475;res=IELHSS

Rutkowski, L., & Rutkowski, D. (2016). A call for a more measured approach to reporting and interpreting PISA results. *Educational Researcher,* 45(4), 252-257.

Sahlberg, P. (2016). The global educational reform movement and its impact on schooling. In: *The Handbook of Global Education Policy* (pp. 128-144). Retrieved from https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=The+global+educational+reform+movement+and+its+impact+on+schooling&btnG=

Sakura, F. (2007). School monitoring and quality assurance in the New Zealand school system. *Educational Research for Policy and Practice*, 6(3), 229-234.

Saljooghi, B. & Salehi, K. (2016). Developing a Teacher Evaluation Model: The Impact of Teachers' Attitude toward the Performance Evaluation System (PES) on Job Satisfaction and Organizational Commitment with the Mediating Role of Teachers' Sense of Efficacy. *Health Sciences,* 5(5), 200-209.

Sandelowski, M. (2003). Tables or tableaux? The challenges of writing and reading mixed methods studies. In: A Tasakkori & C. Teddlie (eds.), *Handbook of Mixed Methods in Social and Behavioral Research* (pp. 321-350). Thousand Oaks, CA: Sage.

Savin-Baden, M. & Major, C. H. (2013). *Qualitative Research: The Essential Guide to Theory and Practice.* Oxford: Routledge.

Scanlon, M. (1999). *The Impact of OFSTED Inspections.* London: National Foundation for Educational Research for the National Union of Teachers.

Scheerens, J. (2014). Evidence based educational policy and practice: The case of applying the educational effectiveness knowledge base. *Journal of Educational, Cultural and Psychological Studies,* 1(9), 83-99.

Sekaran, U. & Bougie, R. (2016). *Research Methods for Business: A Skill Building Approach.* Chichester: John Wiley & Sons.

Sharma, G. (2017). Pros and cons of different sampling techniques. *International Journal of Applied Research,* 3(7), 749-752. Retrieved from https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=Pros+and+cons+of+different+sampling+techniques&btnG=

Sheffield, J., Korotayev, A., & Grinin, L. (2013). *Globalization: Yesterday, Today and Tomorrow*. New York: Routledge.

Siltaloppi, M., Kinnunen, U., & Feldt, T. (2009). Recovery experiences as moderators between psychosocial work characteristics and occupational well-being. *Work Stress,* 23(4), 330-348. Retrieved from https://www.tandfonline.com/doi/full/10.1080/02678370903415572

Singh, K. (2007). *Quantitative Social Research Methods.* Sage.

Sleeter, C. & Carmona, J. F. (2017). *Un-standardizing Curriculum: Multicultural Teaching in the Standards-based Classroom.* Teachers College Press.

Sloan, K. (2006). Teacher identity and agency in school worlds: Beyond the all-good/all-bad discourse on accountability-explicit curriculum policies. *Curriculum Inquiry,* 36(2), 119-152. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-873X.2006.00350.x

SPIS. (2017). *Guidance Manual.* Retrieved from https://departments.moe.gov.sa/EducationAgency/RelatedDepartments/ESUPERVISION/girls/SiteAssets/Pages/مؤشرات-ومنظومة الأداء/__منظومة%20مؤشرات%20قيادة%20الأداء%20الإشرافي%20والمدرسي%20ـ%20الإصدار%20%20_6(1).pdf

Stremfel, U. (2014). Slovenia on its own way towards improving PISA results. *Solsko Polje, 25*(5), 29-51,149-150,169-170. Retrieved from https://search.proquest.com/docview/1664148409?accountid=13460

Subedi, D. (2016). Explanatory sequential mixed method design as the third research community of knowledge claim. *American Journal of Education Research,* 4(7), 570-577.

Sugrue, C. & Mertkan, S. (2017). Professional responsibility, accountability and performativity among teachers: the leavening influence of CPD? *Teachers Teaching,* 23(2), 171-190. Retrieved from https://www.tandfonline.com/doi/full/10.1080/13540602.2016.1203771

Sutherland, D., Price, R., Joumard, I., & Nicq, C. (2007). *Performance Indicators for Public Spending Efficiency in Primary and Secondary Education* (Paper No. 546). Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=967656

Tam, W. H. K. (2008). *Academics' Perspectives of Performance Management in a British University Context.* (Doctoral thesis). University of Leicester, UK.

Tangen, S. (2005). Demystifying productivity and performance. *International Journal of Productivity and Performance Management, 54*(1), 34-46. Retrieved from https://www.emerald.com/insight/content/doi/10.1108/17410400510571437/full/html

Tayan, B. M. (2017). The Saudi Tatweer Education Reforms: Implications of Neoliberal Thought to Saudi Education Policy. *International Education Studies,* 10(5), 61-71. Retrieved from https://eric.ed.gov/?id=EJ1142134

Taylor, E. S., & Tyler, J. H. (2011). *The Effect of Evaluation on Performance: Evidence from Longitudinal Student Achievement Data of Mid-career Teachers* (No. w16877). National Bureau of Economic Research.

Taylor, M., Yates, A., Meyer, L. H., & Kinsella, P. (2011). Teacher professional leadership in support of teacher professional development. *Teaching Teacher Education,* 27(1), 85-94.

Teddlie, C. & Tashakkori, A. (2003). Major issues and controveries inthe use of mixed methods in the social and behvioral sciences. *Handbook of Mixed Methods in Social & Behavioral Research* (pp. 3-50). Thousand Oaks: Sage.Thiel, C. & Bellmann, J. (2017). Rethinking Side Effects of Accountability in Education: Insights from a Multiple Methods Study in Four German School Systems. *Education Policy Analysis Archives, 25*(93). Retrieved from https://eric.ed.gov/?id=EJ1157339

TIMMS. (2015). *TIMSS 2015 International Results in Mathematics.* Retrieved from http://timssandpirls.bc.edu/timss2015/international-results/wp-content/uploads/filebase/full%20pdfs/T15-International-Results-in-Mathematics.pdf

Tolofari, S. (2005). New public management and education. *Policy Futures in Education,* 3(1), 75-89. doi:10.2304/pfie.2005.3.1.11

Tolofari, S. V. (2008). *The Commodification of Higher Education in the Welfare State of Sweden: Exploring the Possibilities.* (Doctoral Thesis). University of Glasgow, UK.

Tomlinson, S. (2005). *Education in a Post Welfare Society.* McGraw-Hill Education (UK).

Torabi, A. & Setodeh, S. (2010). Employees' attitude about the annual performance evaluation of faculties affiliated to Ahvaz University of Medical Sciences in 2007, Iran. *Journal of Health Administration*, 12(38), 17-22.

Townsend, J. G., Porter, G., & Mawdsley, E. (2002). The role of the transnational community of non-government organizations: Governance or poverty reduction? *Journal of International Development,* 14(6), 829-839. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1002/jid.928

UNESCO. (2000). *Final Report.* World Education Forum, Dakar, Senegal, 26-28 April. Retrieved from https://unesdoc.unesco.org/ark:/48223/pf0000121117

Van Dooren, W., Bouckaert, G., & Halligan, J. (2010). *Performance Management in the Public Sector.* Routledge.

Van Geel, A. (2016). Separate or together? Women-only public spaces and participation of Saudi women in the public domain in Saudi Arabia. *Contemporary Islam,* 10(3), 357-378.

Vassiliev, A. (2013). *The History of Saudi Arabia.* London: L Saqi.

Verger, A., Lubienski, C., & Steiner-Khamsi, G. (2016). *World Yearbook of Education 2016: The Global Education Industry.* Routledge.

Vinh, N. X., Chetty, M., Coppel, R., & Wangikar, P. P. (2011). GlobalMIT: Learning globally optimal dynamic bayesian network with the mutual information test criterion. *Bioinformatics,* 27(19), 2765-2766.

Vision 2030. (2015). *National Transformation Program.* Retrieved from http://vision2030.gov.sa/en/ntp

Wadongo, B. I. (2014). *Performance Management and Evaluation in Non-profit Organisations: An Embedded Mixed Methods Approach.* (PhD thesis). University of Bedfordshire, UK. Retrieved from http://uobrep.openrepository.com/uobrep/handle/10547/333444

Walker, M., Jeffes, J., Hart, R., Lord, P., & Kinder, K. (2011). *Making the Links between Teachers' Professional Standards, Induction, Performance Management and Continuing Professional Development RR075.* Manchester: Department for Education.

Walliman, N. (2017). *Research Methods: The Basics.* Routledge: London

Watts, Y. J. (2012). *Ofsted 2005: A New Relationship with Primary Headteachers?* The Open University. Retrieved from http://oro.open.ac.uk/49125/

Whalley, M. A. (2011). *Legitimising Educational Management Identity: Seductive Discourses of Professionalism, Masculinity and Performativity.* (Doctoral thesis). Keele University, UK.

Whitby, K. (2010). School inspection: Recent experiences in high performing education systems. *Reading: CfBT Education Trust.* Retrieved from

https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=School+inspection%3A+Recent+experiences+in+high+performing+education+systems&btnG=

Whittaker, F. (2019). 'No game changer': Ofsted framework proposals won't reduce stress, say unions. *Schools Week.* Retrieved from https://schoolsweek.co.uk/no-game-changer-ofsted-framework-proposals-wont-reduce-stress-say-unions/

Wilkins, C. (2011). Professionalism and the post-performative teacher: New teachers reflect on autonomy and accountability in the English school system. *Professional Development in Education,* 37(3), 389-409. Retrieved from https://www.tandfonline.com/doi/full/10.1080/19415257.2010.514204

Wilkins, C., Busher, H., Kakos, M., Mohamed, C., & Smith, J. (2012). Crossing borders: New teachers co-constructing professional identity in performative times. *Professional Development in Education,* 38(1), 65-77. Retrieved from https://www.tandfonline.com/doi/full/10.1080/19415257.2011.587883

Williams, M., & May, T (1996). *Introduction to the Philosophy of Social Research.* London, UCL Press.

Willms, J. D. (2003). *Monitoring School Performance: A Guide for Educators.* Oxford: Routledge.

Wiseman, A. W., Astiz, M. F., & Baker, D. P. (2013). Globalization and comparative education research: Misconceptions and applications of neo-institutional theory. *Journal of Supranational Policies of Education, 2013*(1).

The World Bank. (2003). The World Bank Annual Report 2003. *World Bank.* Retrieved from http://documents.worldbank.org/curated/en/259381468762619763/The-World-Bank-Annual-Report-2003

The World Bank. (2008). The Road Not Traveled: Education Reform in the Middle East and North Africa. 12/09/2019. *World Bank.* Retrieved from http://siteresources.worldbank.org/INTMENA/Resources/EDU_Flagship_Full_ENG.pdf

Yang, K. & Banamah, A. (2014). Quota sampling as an alternative to probability sampling? An experimental study. *Sociological Research Online, 19*(1), 1-11. Retrieved from https://journals.sagepub.com/doi/full/10.5153/sro.3199

Yia, P. & Kimb, H. J. (2019). Exploring the relationship between external and internal accountability in education: A cross-country analysis with multi-level structural equation modeling *International Journal of Educational Development,* 65. Retrieved from www.elsevier.com/locate/ijedudev

Yin, R. K. (2013). Validity and generalization in future case study evaluations. *Evaluation, 19*(3), 321-332.

Yin, Y. (2009). *Case Study Research: Design and Methods* (4th ed.). California; SAGE Publications.

Zakaria, Z., Yaacob, Z., Noordin, N., Sawal, M. Z. H. M., & Zakaria, Z. (2011). Key performance indicators (KPIs) in the public sector: A study in Malaysia. *Asian Social Science, 7*(7), 102.

Appendices

Appendix I (SPIS Instructional Guide)

Priorities for school performance indicators

1. Academic achievement
2. Student behaviors and values
3. Active learning.
4. -Evaluation (teaching performance - teachers ' functional performance - qualitative performance of the school).
5. School discipline
6. Professional development
7. Evaluation of the construction and implementation of school plans
8. School environment
9. Support services
10. technical visits of subjects supervisors.
11. Technical visits to the school leadership supervisor.
12. Initiatives
13. Student guidance
14. Student activity
15. School operational budget

Instructions for application of the school performance leadership indicators

First: The application of the school performance leadership indicators is assigned to the school leadership section of the education department. Supervision of the application in the education offices is also assigned to one of the school leadership supervisors. The tasks of the evaluation specialist are to be managed by the educational supervision department or the education office in rehabilitation and support.

Second: the school qualitative assessment (the second tool) is applied to all schools at 100% per cent.

Third: The qualitative assessment teams for the school performance are formed as follows:
- Educational Supervisor (Chairman of the team).

- Two members of the <u>schools</u> leaders and their agents (**with the following two conditions**)

1- All the team members (school leaders and agents) should receive the Qualitative Assessment Qualifications Document from the Department of Educational Supervision or the Education Office after passing the qualifying program.

2- The assessment process should be done in a circular rather than reciprocal way, so that the number of teams is equal to the number of educational supervisors.

Fourth: The assessment specialist in the Department of Educational Supervision or the Office of Education, offers three rehabilitation programs, each of which is not less than two days for educational supervisors, school leaders and agents. This is followed by a process of evaluation of the participants through written and practical tests, after that the Department of Educational Supervision will grant a certificate of qualification to those who pass the assessment by 80%, and those who have not yet passed, will take other tests.

Fifth: The educational achievement indicators are distributed among the educational supervisors. The tasks of the educational supervisor are to apply the indicator on the schools and follow up the application until the final statement in the order of the schools is handed over to the supervisor of the school leadership responsible for the implementation process.

Sixth: The Director of Educational Supervision or the Director of the Education office creates a team of (one educational supervisor, members of school leaders and agents) three to five members. Each team shall apply the indicators according to the implementation procedures.

Seventh: In exceptional cases, it is possible to form teams under the chairmanship and membership of school leaders and their agents only without the participation of the educational supervision provided that coordination with the General Directorate of Educational Supervision should be applied before the implementation process, even it is online or electronically.

Eighth: The circular sent to the schools, containing the results of the school's qualitative evaluation (the second tool), should include the following:

1- The circular should include a number and a date.
2- The class order should be classified as five titles classification with the following (distinct performance, good performance, average performance, low performance, inappropriate performance), It is not true otherwise.
3- The circular should reach all schools, and if it does not reach more than one school, it is not considered a circular.
4- The circular should specify the main incentives to be provided to the less performing schools according to the results of the qualitative assessment.
5- The circular should include a specification of the main features of the corrective programs for the schools of the two categories (low performance) and (inappropriate performance).

Ninth: The Department of Educational Supervision communicates with other departments and sections in the Department of Education, then , each department or section has to prepare its assessment card according to the specifications and points specified in this guide

Appendix II

Questionnaire for Headteachers

SECTION A: BACKGROUND INFORMATION

The section below is designed to determine the length of your work experience at the school, position, qualifications and date of last evaluation. (SPIS evaluation)

*I*nstruction: Please, tick (√) the response that corresponds with your background information.

1-What is the highest level of formal education you have completed?

Diploma ☐   Degree ☐   Master's ☐   PhD ☐

Any other? (Please specify) ....

2- How many years have you been working as a headteacher?

Less than 1 year ☐   1 - 2 years ☐   3- 6 years ☐ 7 and more years ☐

3- How many years have you been working as a headteacher in this school?

Less than 1 year ☐   1 - 2 years ☐   3- 6 years ☐ 7 and more years ☐

SECTION B: school monitoring

This section is designed to explore frequency of SPIS monitoring; and, monitored teaching, headteachers' awareness of when the SPIS monitoring will take place, as well as the various monitoring techniques used in SPIS.

School monitoring in this questionnaire focuses on the SPIS evaluation.

1- How often is monitoring conducted by each of the following people:

By school advisor

Never ☐ Once to three times per year ☐ four times and more per year ☐

By SPIS

Never ☐ Once to three times per year ☐ four times and more per year ☐

2- Are you aware of when PISE monitoring will take place

Never ☐ seldom ☐ sometimes ☐ usually ☐ always ☐

3- Are you aware of what the SPIS monitoring will focus on?

Never ☐ seldom ☐ sometimes ☐ usually ☐ always ☐

4- In your opinion are the monitoring techniques appropriate?

Never ☐ seldom ☐ sometimes ☐ usually ☐ always ☐

Please explain your answer ……………………………………………………………………………………

………

2

5- Are you able to recall your school's approximate figures from the performance tables?

Never ☐ seldom ☐ sometimes ☐ usually ☐ always ☐

6- Are your school's aims expressed in KPI terms?

Never ☐ seldom ☐ sometimes ☐ usually ☐ always ☐

7- Do KPIs identify the key strengths of your school?

Never ☐ seldom ☐ sometimes ☐ usually ☐ always ☐

8- Do KPIs help identify the main weakness of your school?

Never ☐ seldom ☐ sometimes ☐ usually ☐ always ☐

SECTION D: Workload and stress

This section is designed to explore workload and stress. To be clear Stress in this questionnaire refers to the emotional, physiological and/or behavioural response to demands and pressures.

A- Workload

1- I do not feel pressure due to workload resulting from SPIS evaluation.

Strongly disagree [＿＿] Disagree [＿＿] undecided [＿＿] Agree [＿＿] strongly agree [＿＿]

2- I work long hours due to SPIS evaluation.

Strongly disagree [＿＿] Disagree [＿＿] undecided [＿＿] Agree [＿＿] strongly agree [＿＿]

B- Stress

1- In general, how stressful do you find being a headteacher of scholl evaluating by SPIS

Not at all stressful [＿＿] Mildly stressful [＿＿] . Moderately stressful [＿＿]

Very stressful [＿＿] . Extremely stressful [＿＿]

If you experienced stress, please explain what aspects of SPIS evaluation caused the stress in your opinion

[＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿]

2- In the six months after the SPIS evaluation did you have time off due to stress?

Never [＿＿] seldom [＿＿] sometimes [＿＿] usually [＿＿] always [＿＿]

235

SECTION C: headteachers' morale

Section E. This section is designed to explore the morale of headteachers.

1- **Overall, I am satisfied with the administration provided by the SPIS evaluation team.**

Strongly agree ☐   Agree ☐   Disagree ☐   Strongly disagree ☐

2- **The demands placed on me by the SPIS evaluation were reasonable.**

Strongly agree ☐   Agree ☐   Disagree ☐   Strongly disagree ☐

3- **The SPIS evaluation provided appropriate opportunities for the headteacher to work productively with the evaluation team.**

Strongly agree ☐   Agree ☐   Disagree ☐   Strongly disagree ☐

4- **The oral feedback and the written report from SPIS were consistent.**

Strongly agree ☐   Agree ☐   Disagree ☐   Strongly disagree ☐

5- **The SPIS evaluation identified clear recommendations for improvement.**

Strongly agree ☐   Agree ☐   Disagree ☐   Strongly disagree ☐

6- **I will use the SPIS evaluation's recommendations to move the school/ my teaching forward.**

Strongly agree ☐   Agree ☐   Disagree ☐   Strongly disagree ☐

7- **I am satisfied that the views of pupils were explored by SPIS inspectors.**

Strongly agree ☐   Agree ☐   Disagree ☐   Strongly disagree ☐

6

SECTION D: School improvement

This section is designed to explore the quality of education and what pupils have achieved after evaluation, in terms of educational standards, pupils' behaviour and pupils' attendance. In addition, it will look at professional development opportunities, professional opportunities for teachers and the working environment in the schools.

Since the SPIS evaluation has there been a change in:

1. **The quality of education provided**

   Improvement ☐ No change ☐ Deterioration ☐ Unable to say ☐

2. **Educational standards achieved by pupils**

   Improved ☐ No change ☐ Deteriorated ☐ Unable to say ☐

3. **Pupils' behaviour**

   Improved ☐ No change ☐ Deteriorated ☐ Unable to say ☐

4. **Pupils' attendance**

   Improved ☐ No change ☐ Deteriorated ☐ Unable to say ☐

Would you be willing to take part in a 30-40-minute research interview to discuss your experience of SPIS evaluation in more detail?

YES ☐ NO ☐

Please add your email contact address if you would like to be contacted for the interviews:

☺ Thank you very much for your time and cooperation. ☺

8

Questionnaire for teachers

SECTION A: BACKGROUND INFORMATION

The section below is designed to determine the length of your work experience at the school, position, qualifications and date of last SPS evaluation.

Instruction: Please, tick (√) the response that corresponds with your background information.

1- What is the highest level of formal education you have completed?

Diploma ☐   Degree ☐   Master's ☐   PhD ☐

Any other? (Please specify) ....·

2- How many years have you been working as a teacher?

Less than 1 year ☐   1 - 2 years ☐   3- 6 years ☐   7 and more years ☐

3- How many years have you been working as a teacher in this school?

Less than 1 year ☐   1 - 2 years ☐   3- 6 years ☐   7 and more years ☐

SECTION B: school monitoring

This section is designed to explore frequency of SPIS monitoring; and, monitored teaching, teachers' awareness of when the SPIS monitoring will take place, as well as the various monitoring techniques used in SPIS. Also, it is designed to explore the use and awareness of key performance indicators (KPIs) of SPS evaluation; which include four questions. They are related to the ability of teacher to read performance tables, KPI and its ability to identify the strengths and weakness of school performance. To be clear, key performance indicators (KPIs) in this questionnaire refer to a tool that measures the achievement scores in schools and describe their performance

School monitoring in this questionnaire focuses on the SPIS evaluation

1- **How often monitoring is conducted by each of the following people? :**

By Head teacher

Never ☐   Once to three times per year ☐   four times and more per year ☐

By head teacher's assistant

Never ☐   Once to three times per year ☐   four times and more per year ☐

By other teaching staff

Never ☐   Once to three times per year ☐   four times and more per year ☐

By school advisor

Never ☐   Once to three times per year ☐   four times and more per year ☐

By SPS

Never ☐   Once to three times per year ☐   four times and more per year ☐

SECTION D: Workload and stress

This section is designed to explore workload and stress. To be clear stress in this questionnaire refers to the emotional, physiological and/or behavioural response to demands and pressures.

A- Workload

1- I do not feel pressure due to workload resulting from SPIS evaluation.

Strongly disagree ☐ Disagree ☐ undecided ☐ Agree ☐ strongly agree ☐

2- I work long hours due to SPIS evaluation.

Strongly disagree ☐ Disagree ☐ undecided ☐ Agree ☐ strongly agree ☐

B- Stress

1- In general, how stressful do you find being a teacher? in school evaluated by SPS

Not at all stressful ☐ Mildly stressful ☐ . Moderately stressful ☐

Very stressful ☐ . Extremely stressful ☐

If you experienced stress, please explain what aspects of SPS evaluation caused the stress in your opinion:

```



```

2- In the six months after the SPS evaluation did you have time off due to Stress?

Never ☐ seldom ☐ sometimes ☐ usually ☐ always ☐

3- Do you feel that your illness was linked to the evaluation?

Major contributing factor ☐ contributing factor ☐ Minor contributing factor ☐

Not connected ☐

SECTION E: teachers' morale

This section is designed to explore the morale of teachers.

1- Overall, I am satisfied with the administration provided by the SPS evaluation team.

Strongly agree ☐ Agree ☐ Disagree ☐ Strongly disagree ☐

2- The demands placed on me by the SPS evaluation were reasonable.

Strongly agree ☐ Agree ☐ Disagree ☐ Strongly disagree ☐

3- The oral feedback and the written report from SPS were consistent.

Strongly agree ☐ Agree ☐ Disagree ☐ Strongly disagree ☐

4- The SPS evaluation identified clear recommendations for improvement.

Strongly agree ☐ Agree ☐ Disagree ☐ Strongly disagree ☐

5- I will use the SPS evaluation's recommendations to move the school/ my teaching forward.

Strongly agree ☐ Agree ☐ Disagree ☐ Strongly disagree ☐

6 - I am satisfied that the views of pupils were explored by SPS inspectors

Strongly agree ☐ Agree ☐ Disagree ☐ Strongly disagree ☐

7- The benefits of the SPS evaluation outweigh the negative aspects.

Strongly agree ☐ Agree ☐ Disagree ☐ Strongly disagree ☐

8- Overall, I am satisfied with the way the SPS evaluation was carried out.

Strongly agree ☐ Agree ☐ Disagree ☐ Strongly disagree ☐

SECTION F: School improvement

This section is designed to explore the quality of education and what pupils have achieved after evaluation, in terms of educational standards, pupils' behaviour and pupils' attendance. In addition, it will look at professional development opportunities, professional opportunities for teachers and the working environment in the schools.

Since the SPS evaluation has there been a change in:

1. **The quality of education provided**

   Improvement ☐ No change ☐ Deterioration ☐ Unable to say ☐

2. **Educational standards achieved by pupils**

   Improved ☐ No change ☐ Deteriorated ☐ Unable to say ☐

3. **Pupils' behaviour**

   Improved ☐ No change ☐ Deteriorated ☐ Unable to say ☐

4. **Pupils' attendance**

   Improved ☐ No change ☐ Deteriorated ☐ Unable to say ☐

Appendix III

<u>**Provisional Semi-Structured Interview's Matrix for teachers, tailored to teachers'**</u>
<u>**experience of leading SPIS evaluation and monitoring at the school.**</u>

Thank you for taking the time to complete both the teacher Questionnaire and agreeing to take part in this 40-minute interview. The interview should take about 30-40 minutes. Are you available to respond to some questions now?

	Questions
1	I would like to ask you some questions about your background, your education and some experience you have had with the SPIS evaluation process.
	What professional and academic qualifications do you hold? How long have you been a teacher in this school?
	Tell me about your role at this school.
	Have you worked as a teacher in another school before this? If yes, where and for how long?
	How often do you get involved in SPIS evaluation?
2	Thank you, it was helpful. Could we now talk about Public Education Evaluation Commission (SPIS)?
	In general, what do you think about SPIS and its influence on the improvements in education?
	Now let us talk about your experience with SPIS evaluation.
	I am interested in learning about what happens when you have a SPIS evaluation. First, tell me how are you informed that a SPIS evaluation will take place.
	Then, what happens?
	As a teacher, to what extent you are committed to this evaluation?
3	Thank you for your explanation, now we can talk about the first step of SPIS evaluation. Can you as a teacher in school evaluated by SPIS describe your experience of being monitored by SPIS ?
	Do you think there is any difference between monitoring by SPIS and by Ministry of education supervisors?
	From your experience, what would you avoid/remove in SPIS's monitoring and what do you think would be useful to add and why?
4	Thank you for all information. It will help my research. Now I want to talk about Key performance indicators (KPIs) in PEEC. Do you think that the KPIs reflect best practice in school performance and why? In your opinion, as a

teacher, does SPIS indicator system help identify your key success factors? Why do you think SPIS can help your school achieve its aims and objectives? In your opinion, are there KPIs in PEEC that cannot be achieved and why? If you had the opportunity to add some KPIs to SPIS KPIs what would they be and why? If you have nothing to add, please tell me why.

Appendix IV

Tick one:

Staff project: _____   PhD ✓   EdD _____

Name of applicant (s): Azzah Alsubaie

Title of project: The impact of PEEC evaluation on private secondary schools in Jeddah from the perspective of teachers and head teachers.

Name of supervisor (for student projects): Dr. Karen Jones.

Please complete the form below including relevant sections overleaf.

	YES	NO	
Have you prepared an Information Sheet for participants and/or their parents/carers that:			
a) explains the purpose(s) of the project	✓		
b) explains how they have been selected as potential participants	✓		
c) gives a full, fair, and clear account of what will be asked of them and how the information that they provide will be used	✓		
d) makes clear that participation in the project is voluntary	✓		
e) explains the arrangements to allow participants to withdraw at any stage if they wish	✓		
f) explains the arrangements to ensure the confidentiality of any material collected during the project, including secure arrangements for its storage, retention, and disposal	✓		
g) explains the arrangements for publishing the research results and, if confidentiality may be affected, for obtaining written consent for this	✓		
h) explains the arrangements for providing participants with the research results if they wish to have them	✓		
i) gives the name and designation of the member of staff with responsibility for the project, together with contact details, including email. If any of the project investigators are students at the IoE, then this information must be included and their names provided	✓		
k) explains, where applicable, the arrangements for expenses and other payments to be made to the participants		✓	
j) includes a standard statement indicating the process of ethical review undergone by the project at the University, as follows: 'This project has been reviewed following the procedures of the University Research Ethics Committee and has been given a favourable ethical opinion.'	✓		
k) includes a standard statement regarding insurance: 'The University has the appropriate insurances in place. Full details are available on request.'	✓		
Please answer the following questions			
1) Will you provide participants involved in your research with all the information necessary to ensure that they are fully informed and not in any way deceived or misled as to the purpose(s) and nature of the research? (Please use the subheadings used in the example information sheets on Blackboard to ensure this)	✓		
2) Will you seek written or other formal consent from all participants, if they are able to provide it, in addition to (1)?	✓		
3) Is there any risk that participants may experience physical or psychological distress in taking part in your research?		✓	
4) Have you taken the online training modules in data protection and information security (which can be found here: http://www.reading.ac.uk/internal/imps/Staffpages/imps-training.aspx)?	✓		
5) Have you read the Health and Safety Booklet (available on Blackboard) and completed a Risk Assessment Form to be included with this ethics application?	✓		
6) Does your research comply with the University's Code of Good Practice in Research?	✓		
	YES	NO	N.A.
7) If your research is taking place in a school, have you prepared an information sheet and consent form to gain the permission, in writing, of the head teacher or other relevant supervisory professional?	✓		
8) Has the data collector obtained satisfactory DBS clearance?			✓
9) If your research involves working with children under the age of 16 (or those whose special educational needs mean they are unable to give informed consent), have you prepared an		✓	

information sheet and consent form for parents/carers to seek permission in writing or give parents/carers the opportunity to decline consent?			
10) If your research involves processing sensitive personal data[1] or audio/video recordings, have you obtained the explicit consent of the participants/parents?	✓		
11) If you are using a data processor to subcontract any part of your research, do you have a written contract with that contractor that (a) specifies that the contractor is required to act only on your instructions and (b) provides for appropriate technical and organisational security measures to protect the data?			✓
12a) Does your research involve data collection outside the UK?	✓		
12b) If the answer to question 12a is 'yes', does your research comply with the legal and ethical requirements for performing research in that country?	✓		
13a) Does your research involve collecting data in a language other than English?	✓		
13b) If the answer to question 13a is 'yes', please confirm that information sheets, consent forms, and research instruments, where appropriate, have been directly translated from the English versions submitted with this application.	✓		
14a. Does the proposed research involve children under the age of 5?		✓	
14b. If the answer to question 14a is 'yes': My Head of School (or authorised Head of Department) has given details of the proposed research to the University's insurance officer, and the research will not proceed until I have confirmation that insurance coverage is in place.			
If you have answered YES to Question 3, please complete Section B below			

Please complete **either** Section A or Section B and provide the details required in support of your application. Sign the form (Section C) then submit it with all relevant attachments (e.g., information sheets, consent forms, tests, questionnaires, and interview schedules) to the Institute's Ethics Committee for consideration. Any missing information will result in the form being returned to you.

A: My research goes beyond the 'accepted custom and practice of teaching' but I believe that this project has **no** significant ethical implications (please tick the box).	✓

Purpose of project and its academic rationale

This study investigates the impact of the PEEC process on private schools in Jeddah. PEEC is a quality process similar to the UK Ofsted process. The main purpose of the study is to investigate the views and perceptions of teachers and head teachers in private schools in Jeddah, as well as to explore the effects of the Public Education Evaluation Commission (PEEC) process, specifically the effect of the use and awareness of the KPIs of the PEEC and grading outcomes on various aspects of school life, which include the following:

1- School monitoring

2- Head teacher morale

3- Teacher morale

4- School improvement (as defined by head teachers and the PEEC)

Brief description of methods and measurements

This research will be designed using a mixed-methods approach (quantitative and then qualitative), which includes a sequential explanatory design, beginning with quantitative data and following up with qualitative data. The quantitative data include an online questionnaire, and the qualitative data include a face-to-face semi-structured interviews.

Document analysis will also be undertaken in twelve schools participating in the qualitative study. The documents reviewed will include, for example, PEEC reports and documents about the school, its staff and students, its curriculum, and so forth. In addition, any documents reporting staff absenteeism will also be collected.

Participants: recruitment methods, number, age, gender, and exclusion/inclusion criteria

A total of 112 girls' secondary schools in Jeddah will be involved in the study. The age of students in

---

[1] Sensitive personal data consist of information relating to the racial or ethnic origin of a data subject or his or her political opinions, religious beliefs, trade union membership, sexual life, physical or mental health or condition, criminal offences, or criminal record.

researcher and reviewed by a professional translator.

Consent and participant information arrangements, debriefing (attach forms where necessary)

Informational letters and consent forms are contained in the Appendix, as follows:

1. Ethics form
2. Risk assessment form
3. Permission to Conduct Research Study from the Director of the Jeddah Directorate of Education.
4. Information letter and consent form for head teachers (to allow schools to take part in the study)
5. Information letter and consent form for head teachers who will be interviewed
6. Information letter and consent form for teachers who will be interviewed
7. Information email/consent statement for the questionnaire(Headteacher)
8. Information email/consent statement for the questionnaire(Teacher)
9. A copy of the interview questions for head teachers
10. A copy of the interview questions for teachers
11. A copy of the questionnaire for head teachers
12. A copy of the questionnaire for teachers
13. My results for a Data Protection Act course in 1998
14. My results for a Freedom of Information Act course in 2000.

In addition, all respondents will be required to agree to a consent statement in the online questionnaire so that consent can be obtained from the survey respondents before they take the survey. This will prevent respondents from taking the survey if they do not provide consent.

Estimated start date and duration of project
Need to plan a pilot study.
I will start on the 1st of May 2017 and finish on the 30th of October 2017, as follows:
- Pilot online questionnaire: May-June 2017
- Roll out online questionnaire: June-August 2017
- Pilot then roll out qualitative face-to-face semi-structured interviews: September/October 2017

B: I believe that this project **may** have ethical implications that should be brought before the Institute's Ethics Committee.	

Please state the total number of participants that will be involved in the project and give a breakdown of how many there are in each category, e.g., teachers, parents, pupils, etc.

Give a brief description of the aims and methods (participants, instruments, and procedures) of the project in up to 200 words.
1. Title of project
2. Purpose of project and its academic rationale
3. Brief description of methods and measurements
4. Participants: recruitment methods, number, age, gender, and exclusion/inclusion criteria
5. Consent and participant information arrangements, debriefing (attach forms where necessary)
6. A clear and concise statement of the ethical considerations raised by the project and how you intend to deal with then.
7. Estimated start date and duration of project

C: SIGNATURE OF APPLICANT:

Note: a signature is required. Typed names are not acceptable.

I declare that all relevant information regarding my proposed project has been provided and confirm that ethical good practices will be followed during the project.

Signed: .   ..   Print Name…Azzah Alsubaie………. Date…06/03/2017…

STATEMENT OF ETHICAL APPROVAL FOR PROPOSALS SUBMITTED TO THE INSTITUTE ETHICS COMMITTEE

This project has been considered using agreed-upon Institute procedures and is now approved.

Signed: ..   ....   Print Name…Xiao Lan Curdt-Christiansen.   Date10/05/2017…….
(IoE Research Ethics Committee representative)*

* A decision to allow a project to proceed is not an expert assessment of its content or of the possible risks involved in the investigation, nor does it detract in any way from the ultimate responsibility that students/investigators must themselves have for these matters. Approval is granted on the basis of the information declared by the applicant.

Appendix V

University of Reading

Researcher:
Name Azzah Alsubaie
Phone: +44 7979924199
Email: A.H.H.Alsubaie@student.reading.ac.uk

Supervisors:
Name Dr. Karen Jones
Phone: +44 (0)118 378 2603
Email: Karen.Jones@reading.ac.uk
Name Dr. Chris Turner
Email:c.k.turner@reading.ac.uk

Dear Head teacher,

You have been asked to participate in a research study because of your experience in school teaching. A total of approximately 24 people have been asked to participate in this study, including twelve head teachers and twelve teachers. The purpose of this study is to investigate the views and perceptions of teachers and head teachers in private schools in Jeddah, as well as to explore the effects of the PEEC process and grading outcomes on various aspects of school life, including the following:

- head teacher stress levels, work load, and morale

- teacher stress levels, work load, and morale

- school improvement (as defined by the head teachers and the PEEC)

- school monitoring

The results of this study will be used for research purposes, within my thesis and as part of external research publications in the future. If you agree to participate in this study, you will be asked to take part in an interview with the researcher, which will last approximately 30 minutes. The interview will be recorded and transcribed with your permission. The information gathered will be used by the student researcher for data analysis .If you agree to participate in this study, you will be asked to take part in an interview with the researcher, which will last approximately 30 minutes. The interview will be recorded and transcribed with your permission. The information gathered will be used by the student researcher for data analysis. These data may subsequently be published in a report or other research publications. If you would like to see the publications resulting from the research, please contact me and I will be happy to send them to you. All data arising from the research will be kept securely in a locked filing cabinet and on a password-protected computer. In line with the University policy on the management of research data, anonymised data gathered in this research may be preserved and made publicly available for others to consult and re-use. These data may be used in future publications in appropriate academic journals and/or books. All participants will have access to a copy of the published research upon request. Your decision to participate is entirely voluntary. Also, you are free to withdrawal your consent at any time, without giving a reason, by contacting the student researcher, Azzah Alsubaie, at (+44 7979924199) or via e-mail at A.H.H.Alsubaie@student.reading.ac.uk if you wish to withdraw from the study. This application has been reviewed inlight of the procedures of the University Research Ethics Committee and has been given a favourable ethical opinion. The University has the appropriate insurances in place. Full details are available on request. If you have any queries or wish to clarify anything about the study, please feel free to contact me via email at A.H.H.Alsubaie@student.reading.ac.uk

Appendix VI

University of Reading

Researcher: **Supervisors:**

Name Azzah Alsubaie   Name Dr. Karen Jones

Phone: +44 7979924199   Phone: +44 (0)118 378 2603

Email: A.H.H.Alsubaie@student.reading.ac.uk Email: Karen.Jones@reading.ac.uk

Name Dr. Chris Turner

Email:c.k.turner@reading.ac.uk

Dear Teacher,

You have been asked to participate in a research study because of your experience in school teaching. A total of approximately 24 people have been asked to participate in this study, including twelve head teachers and twelve teachers. The purpose of this study is to investigate the views and perceptions of teachers and head teachers in private schools in Jeddah, as well as to explore the effects of the PEEC process and grading outcomes on various aspects of school life, including the following:

- head teacher stress levels, work load, and morale

- teacher stress levels, work load, and morale

- school improvement (as defined by the head teachers and the PEEC)

- school monitoring

The results of this study will be used for research purposes, within my thesis and as part of external research publications in the future. If you agree to participate in this study, you will be asked to take part in an interview with the researcher, which will last approximately 30 minutes. The interview will be recorded and transcribed with your permission. The information gathered will be used by the student researcher for data analysis. If you agree to participate in this study, you will be asked to take part in an interview with the researcher, which will last approximately 30 minutes. The interview will be recorded and transcribed with your permission. The information gathered will be used by the student researcher for data analysis. These data may subsequently be published in a report or other research publications. If you would like to see the publications resulting from the research, please contact me and I will be happy to send them to you. All data arising from the research will be kept securely in a locked filing cabinet and on a password-protected computer. In line with the University policy on the management of research data, anonymised data gathered in this research may be preserved and made publicly available for others to consult and re-use. These data may be used in future publications in appropriate academic journals and/or books. All participants will have access to a copy of the published research upon request.

Your decision to participate is entirely voluntary. Also, you are free to withdrawal your consent at any time, without giving a reason, by contacting the student researcher, Azzah Alsubaie, at (+44 7979924199) or via e-mail at A.H.H.Alsubaie@student.reading.ac.uk if you wish to withdraw from the study. This application has been reviewed in light of the procedures of the University Research Ethics Committee and has been given a favourable ethical opinion. The University has the appropriate insurances in place. Full details are available on request. If you have any queries or wish to clarify anything about the study, please feel free to contact me via email at A.H.H.Alsubaie@student.reading.ac.uk

Signed:

Appendix VII (Teachers' interview example)

<u>المقابلة الشخصية للمعلمين</u>

شكراً لك لإكمالك استبيان المدرسين وقبولك أن تخصص 40 دقيقة من وقتك لإجراء المقابلة. هل أنت متفرغ للإجابة عن بعض الأسئلة الآن؟

سأطرح بعض الأسئلة عن خلفيتك وتعليمك وبعض الخبرات التي حصلت عليها خلال عملية تقويم المنظومة.

ما المؤهلات العلمية والمهنية التي تملكها؟

بكالوريوس وعدة دورات في طرق التدريس

منذ متى كنت مدرسًا في هذه المدرسة؟

منذ ثلاث سنوات في هذه المدرسة.

أخبرني عن دورك في هذه المدرسة.

تعليمي

هل عملت مدرسًا من قبل في مدرسة أخرى؟ إن أجبت بنعم، حدد المكان والمدة الزمنية؟

عملت سنة داخل جدة وخارجها

شكراً لك، لقد كان هذا مفيدًا. هل يمكن أن نتحدث الآن عن المنظومة؟

بشكل عام، ما رأيك فيها وما تأثيرها برأيك على تحسين عملية التعليم؟

عرفت من المنظومة من المديرة، وبحثت عن المعلومات عنها وأظن أنها تضمن لك ما عليك وما لك؛ بمعنى أن مؤشراتها أخبرتني بما هو مطلوب مني وحددت لي ما هو دوري كمعلم، وأظن لها دور في مساءلتي عن تطوير المدرسة أو تطوير نفسي؛ ربما يرجعون لها عند كتابة تقرير أدائي الوظيفي، كما أن مؤشرات المنظومة تساعدني في بعض النقاط التي تعينني في تحسين أدائي.

مثل ماذا؟

مثل مطالبتي بتطبيق استراتيجيات الأداء.

الآن دعينا نتحدث عن تجربتك مع تقويم المنظومة.

أخبريني كيف يتم إعلامك بموعد إجراء التقويم؟

المديرة أخبرتني أنهم سيحضرون في هذه الفترة؛ ليس اليوم بالتحديد لكن قالت أنهم سيحضرون قريبًا.

ثم ماذا حدث؟

زيارة عادية؛ دخلوا الصف لمدة دقائق وطلبوا ملفاتي.

هل ترين نفسك ملتزمة بالعمل بالمنظومة؟

ألتزم بها لأنها جيدة أولًا، وثانيًا لأن المديرة تساءلني عن تطبيق المؤشرات وتحثنا دائمًا، هي تقول إنه سيساعد في تطوير المدرسة، أظن فعلًا المؤشرات تساعد في تطوير أدائي.

زيارة عادية؛ دخلوا الصف لمدة دقائق وطلبوا ملفاتي.

هل ترين نفسك ملتزمة بالعمل بالمنظومة؟

ألتزم بها لأنها جيدة أولًا، وثانيًا لأن المديرة تساءلني عن تطبيق المؤشرات وتحثنا دائمًا، هي تقول إنه سيساعد في تطوير المدرسة، أظن فعلًا المؤشرات تساعد في تطوير أدائي.

أريد إزالة كل ما هو مستحيل تطبيقه في بيئة مدرستنا؛ مثلا يطالبوننا باستخدام التكنولوجيا ونحن لا نملك أجهزة مساعدة في المدرسة، بصراحة لا أجد من المنطقي أن يحكموا علينا في شيء لا نملك القدرة على تنفيذه! إذا أردت تطوير المدرسة أعطني الأدوات التي أستطيع فيها تحسين المدرسة، أنا لا أعرف إذا كان ذلك يؤثر على أدائنا، لكن لو حدث ذلك فسيكون خطأ كبيرًا.

ولا يوجد ما أضيفه.

في تجربتك، ما التحديات التي تواجهها أثناء القيام بهذا التقويم؟

ضغط العمل بسبب أعداد طالبات المدرسة؛ فكل طالبة لها ملف خاص بها؛ مما جعل أعباء كبيرة علي، لدي جدول حصص كبير ومتعب، ليس لدي ساعات مكتبية للتفرغ لهذا العمل المكتبي؛ لذا آخذه معي في البيت، إذا كانت هذه الوثائق تساعد في تطوير أداء طالباتي ولو لم أفعله ينقص أدائي أعطني وقتًا له.

أخبرني عن تأثير عملية تقويم المنظومة عليك شخصيًا.

أصابتني بالإحباط خاصة أن ما كان مطلوبًا مني كان فوق احتمالي.

كيف شعرت خلال تقويم المنظومة؟

ليس شيئًا مهمًا بالنسبة لي لأنهم لم يأتوا لي أنا، لقد جاءوا للمدرسة، ربما كنت محظوظة لأنهم لم يدخلوا صفي، أعتقد أني كنت سأتحدث معهم صراحة عن هذه الضغوط.

أخبرني عن تأثير عملية تقويم المنظومة عليك شخصيًا. كيف شعرت كمدرس؟

شعرت بمشاعر سلبية؛ تراكم عمل ومطالبات بأمور مستحيلة؛ فلم أجد لدي قدرة على تنظيم وقتي.

ما الذي يمكن القيام به في رأيك لجعل تقويم المنظومة أقل إرهاقًا؟

زيادة أعداد المعلمات، وتوفير بيئة مناسبة ممكن أحقق فيها ما يريدونه، معلمة مساعدة لي أيضًا وقت للعمل المكتبي مثل التوثيق.

كنت تحدثت عن ضغط العمل ـ ما الأشياء الأخرى التي تعتقد أنها تحتاج إلى تحسين لجعل التقويم أكثر ملاءمة للمدرسين، إن أردت، من حيث التوتر؟

وجود مساعدين للمعلم، وتقليل عدد الطلاب، ووضع معايير منطقية تناسب بيئة المدرسة، كما أرجو أن لا يحاسبونا على عدم توفير المستلزمات.

مثل ماذا؟

يعني معيار يكون بيئة المدرسة مؤهلة لتطبيقه.

برأيك ما الإجراءات التي يجب على المدارس والمنظومة استخدامها للحد من التوتر؟

مثل ما قلت: تخفيف العمل وإعطاؤنا وقتًا ودعمًا ماليًا مناسبًا.

شكرًا لك على جميع المعلومات. سوف تساعد بحثي. الآن أريد أن أتحدث عن مؤشرات الأداء المفتاحية (KPI) في تقويم المنظومة. هل تعتقد أن KPI تعكس أفضل الممارسات في الأداء المدرسي ولماذا؟

نعم، لكن في البيئة الجيدة برأيي بيئة مدرستي سيئة ولا تستطيع تطبيق هذه المعايير فيها.

كيف ذلك؟

مثلا؛ لا توجد سبورات ذكية أو مساحات في الفصول لتطبيق الاستراتيجيات المناسبة.

برأيك كمدرس، هل يساعد نظام مؤشرات المنظومة على تحديد عوامل النجاح الرئيسية؟

نعم إلى حد ما، لكن كما قلت سابقًا نحتاج لبيئة مدرسية مساعدة.

لماذا تظن أن المنظومة يمكن أن تساعد مدرستك تحقيق أهدافها وغاياتها؟

ربما، لكن ليس كل المدارس.

لماذا؟

لأن بيئة المدرسة وأعداد الطلاب أحيانًا بل دائمًا تؤثر على أهداف المدرسة، والمعايير في المنظومة جعلت الأمر يبدو أن المدارس كلها عالية التجهيز، هذا الأمر غير عادل.

برأيك، هل هناك KPI في تقويم المنظومة لا يمكن تحقيقها ولماذا؟
نعم الخاصة باستراتيجيات التعليم كما قلت الفصول صعب تحقيق ذلك مع ازدحامها.

إن أتيحت لك الفرصة لإضافة بعض المؤشرات إلى مؤشرات المنظومة ماذا ستكون ولماذا؟
بودي كل مؤشر يناسب المدرسة بيئتها ويكون واقعيًا.

إذا كان لديك أي شيء لإضافته، من فضلك أخبرني لماذا.

لا شكرًا.

Appendix VIII

A Reviewed Sample of a Transcript of a Teacher's Interview

Me: Thank you for accepting to answer our questions for 40 minutes. First, what are your qualifications?

Teacher: I have a bachelor's degree in addition to some courses on teaching methods.

Me: How long have you spent at this school?

Teacher: About three years.

Me: What do you do in this school?

Teacher: I work as a teacher.

Me: Have you worked in other schools? If yes, please tell me where and when.

Teacher: Yes, I worked for a year outside Jeddah city.

Me: What do you think about CCE (Continuous and Comprehensive Evaluation), and do you think these KPIs will affect the educational process?

Teacher: I have read about KPIs and I think that these accurately detect the duties and rights of every teacher; because the committee constantly assesses our educational performance and our efforts to develop our school, after we present our work performance report, in addition to giving us some advice on enhancing our performance.

Me: What does this advice consist of?

Teacher: They have advised me to apply performance strategies.

Me: How do you know the time of the KPI assessment?

Teacher: The manager tells me that this assessment will start soon.

Me: And then, what happens after she has notified you of the time?

Teacher: The committee enters the classroom and asks me for my papers.

Me: Are you interested in dealing with these KPIs?

Teacher: Yes; because I think that benefiting from KPI recommendations will develop our school, in addition to developing my performance.

Me: Tell me about your impressions after the SPIS visit to your school.

Teacher: I didn't notice that they were there during my work; I hope they are successful.

Me: There are many committees like the Educational Supervision Committee, which comes to your school to assess your performance. Do you think there is a difference between these committees and the SPIS?

Teacher: Of course; because SPIS comes to assess the school, but other committees come to assess the targeted teacher only; so I wish that they would co-ordinate with each other, because there are a lot of committees that visit our school, and we find ourselves being forced to change our timetables to help these committees.

Me: From your experience of these KPIs, do you think that we should add something to them, or remove something from them?

Teacher: I don't wish to add anything, but if you wanted to remove something, you should remove unrealistic standards from them; for example, SPIS asks us to use technology, but we don't have it; so if the committee wants us to use technology, it should support us with technology.

Me: What are the difficulties that you face while this assessment is being performed?

Teacher: The amount of work is the main thing that bothers me, because there are a lot of students and I don't have enough time to read any additional documents at school, so I take them home to read. Therefore, I would like more time to do that, because these additional documents would enhance my performance a lot.

Me: How has this assessment affected you?

Teacher: I have become frustrated, because I can't meet all these standards.

Me: What are your feelings towards this assessment?

Teacher: It doesn't matter; because SPIS don't only come to see me, but also to see the school… they haven't entered my classroom, and if they had, I would have told them about the problems that face us.

Me: You are a teacher; how did you feel after they came to your school?

Teacher: I had negative feelings; because there is a lot of work, in addition to impossible demands, so I couldn't find time to implement the KPI recommendations.

Me: What do you suggest to make assessment easier?

Teacher: I would like to see more teachers employed at our school, in addition to finding more time to do office work, and I want an assistant to work with me.

Me: You talked about stress at work; what are your suggestions to solve this problem, and do you think that employing more teachers would be a solution?

Teacher: Yes, I do, but student numbers should also be reduced in the classrooms; KPIs should also be realistic, instead of unrealistic.

Me: What do you mean by realistic KPIs?

Teacher: Realistic KPIs are any standards that we can achieve.

Me: What procedures do you think should be taken to stop stress at work?

Teacher: As I said before; schools and CCE committees should do three things: reduce the number of KPIs, give us more time to achieve these KPIs, give us financial support.

Me: Do you think that KPIs offer the best suggestions for achieving the best educational performance, and why?

Teacher: Yes, but only if these KPIs are applied in good environments, not in my school which suffers from several problems.

Me: How could we enable your school to meet the KPI requirements?

Teacher: If classrooms were extended and we could work with smart blackboards, it would be a good start for meeting KPI requirements

Me: Do you think that KPIs can detect the principle factors of success?

Teacher: Yes, for other schools, but not for our school.

Me: Do you think that KPIs can help schools to achieve their goals?

Teacher: Yes, but not for all schools.

Me: Why?

Teacher: Because many factors like the number of students in the classroom correspond to schools' goals, but the real problem is in the KPI requirements that deal with all schools in the same way, without taking the differences between schools into considerations. Therefore, the KPIs are not fair.

Me: Do you think that the KPI requirements are unrealistic standards?

Teacher: Yes, especially the standards that are related to education strategies; because these strategies are aimed at lower numbers of students in the classroom, in order to ensure success.

Me: If you could add new standards to the KPIs, what would you add?

Teacher: I would add new KPIs to replace the older ones, which would be compatible with schools' capabilities.

Me: Do you have anything else to say?

Teacher: No, thank you.

Appendix IX (Headteacher interview example in Arabic)

مقابلة مديرة باللغة العربية

أنا: مرحبًا بك عزيزتي، وأشكرك على قبولك مقابلتي والتي لا شك ستزيد دراستي وبحثي غنًى.
سأطرح بعض الأسئلة عن خلفيتك وتعليمك وبعض الخبرات التي حصلت عليها خلال عملية تقويم المنظومة، بإمكانك الاحتفاظ باسمك وبإمكانك التصريح به، لكنه لن يتم عرضه أبدًا .

أولًا لدي سؤال حول خلفيتك العلمية وهو التالي:

ما المؤهلات العلمية والمهنية التي تملكينها؟

المديرة : اسمي هو.... وأنا لدي بكالوريوس قسم تاريخ، وخبرتي 23 سنة، تعينت مبكرًا؛ تخرجت ثم تعينت مباشرة؛ لذا اكتسبت خبرة منذ كنت صغيرة نسبيًا .
أنا: كم لك سنة مديرة هنا؟

المديرة: لي 7 سنوات مديرة هنا تقريبًا
أنا: هل لديك تجربة مع مدارس أخرى قبلها؟
المديرة: لدي خبرات متعددة في مدارس أخرى؛ عملت وكيلة مدرسة هنا في جدة، وشغلت مهام إدارية خلال هذه الفترة أيضًا ثم مديرة، اكتسبت خبرات في التعامل مع الجميع على اختلافهم والفروق الفردية بينهم؛ سواء طالبات أو معلمات، كما اكتسبت خبرة في التعامل مع المشرفات والوزارة وتقويم الأداء خاصة تقويم أداء المعلمات.
أنا: هل يمكن أن تصفي لي دورك في هذه المدرسة؟
المديرة: أنا هنا أم بكل شفافية، وأنا أفضل هذا الوصف على الرغم من وجود الكثير من العبارات عن القائدة المدرسية، لكني أعتقد أن قائدة المدرسة في النهاية أم، والأم هي دائمًا القائدة والمسؤولة عن كل شيء، وفي النهاية الأم دائمًا تخاف على أبنائها وتحرص أن تؤثر عليهم بشكل إيجابي حتى يكونوا جيدين، والمفترض من تكون قائدة المدرسة أن تعتبر نفسها أمًا بالدرجة الأولى.
أنا: لكن هذا لا يلغي دورك كمديرة المدرسة
المديرة: بالطبع، أنا أرى نفسي حازمة ورسمية كثيرًا ومهنية في عملي.
أنا: شكرًا لك، لقد كان هذا مفيدًا. هل يمكن أن نتحدث الآن عن المنظومة؟ وعن تجربتك مع تقويم المنظومة؟
أنا: أخبريني كيف يتم إعلامك بموعد إجراء التقويم؟
المديرة: علمت من خلال اجتماعات ولقاءات مختلفة.
أنا: مع من؟
المديرة: مع مكتب التعليم في شمال المدينة، وهو المكتب المسؤول عن مدرستنا مع المشرفات اللواتي قمن بشرحها لنا وإخبارنا بأننا سيتم تقييمنا وسيسبق ذلك دورات وتفاصيل كثيرة، وهناك أداة تقويم جديدة تم تزويدنا بها قبل ذلك من قبل الوزارة توضّح المعايير وآلية العمل، كما تم تزويدنا بقائمة بالمؤشرات في دليل اشتمل على كل شيء تقريبًا.

أنا: بشكل عام، ما رأيك فيها وما تأثيرها برأيك على تحسين عملية التعليم؟
المديرة: دعيني أقول بكل صراحة ووضوح، إن مدرستنا من قبل كان لديها خطة للتطوير ومنهج للرقي بالتعليم بغض النظر عن برامج الوزارة، في المدرسة عندنا سياسة وتوجه للنهوض بالتعليم، وتقسيم المعلمات والطالبات، وملفات الإنجاز، وكل الأمور التي جاءت بها المنظومة هي أصلًا كانت مطبقة لدينا من قبل من عدة سنوات؛ فسياستنا كانت قائمة على تفاصيل نجدها الآن في هذه المنظومة.
فلما جاءت المنظومة كانت أصلًا لدينا معاييرها إلا ببعض الفروقات، فقط احتجنا للترتيب والتنظيم؛ لذا دعيني أقول: إن المنظومة ساعدتنا في إظهار منجزنا الذي اعتدنا على فعله من قبل بشكل موثق، لكننا بدونها كنا نسير في نفس الاتجاه.
أنا: هل تعرفين إذا كان غيركم من المدارس يمارس ذلك أم مدرستكم فقط؟
المديرة: لا أعرف، لكن كان هذا توجهي أنا وفريقي رغبة في مسايرة التطوير في التعليم.

أنا: شكرًا لكم على التفسير، والآن يمكننا أن نتحدث عن الخطوة الأولى من تقويم المنظومة، هل يمكنك كمديرة في المدرسة التي تم تقويمها أن تصف تجربتك مع زيارة المنظومة لمدرستك؟

المديرة: هم أرسلوا لنا جدولًا بحضور الفريق الخاص بالمنظومة لنجهز عملنا ونرتب لهم جولة داخل المدرسة أثناء اليوم الدراسي، ومن الجدول عرفت متى سيحضرون؛ لذا أمضينا اليوم بشكل عادي، دخلوا وحضروا بعض الفصول التي تم اختيارها بشكل عشوائي، وكنت قد جهزت جميع الملفات لهم وقاموا بالاطلاع عليها بشكل دقيق ومراجعة المعايير معي وما حققته، كان يومًا طويلًا.

أنا: اليوم اللي حضروا فيه..؟

المديرة: حضروا مبكرًا، وأنا لا أنسى هذا اليوم لأن إحدى المعلمات أصيبت بنزف وكانت حاملًا.

أنا: هل كان ذلك بسبب ضغط التقويم؟

المديرة: لا... لا أظن بسبب ذلك، لكن هي كانت مريضة.

أنا: هل أنت متأكدة؟

المديرة: صعب إثبات ذلك، عمومًا لا يمكن أنسى هذا اليوم.

أنا: دعيني أسألك هل تعتقدين أن هناك أي فرق بين الزيارة من قبل المنظومة من قبل مشرفة المدرسة المعتادة؟

المديرة: الفرق أن الجهد المبذول منا يكون أكبر؛ لأن مشرفة التقويم العادية تأتي لعمل خاص يخص مادتها أو تخصصها، لكن فريق تقويم المنظومة يبحث في التفاصيل في كل شيء؛ لذا الجهد مضاعف عدة مرات، مثلًا لدي خطة تفصيلية يجب أن يراجعوها ثم يراجعوا ملفات إنجاز المعلمات والطالبات.

أنا: هل تشعرين بأن ذلك ضاعف عملك وأرهقك؟

المديرة: في الحقيقة هو جهد عظيم، لكن الهدف أن يكون التعليم مختلفًا وأن تلميذاتي يكن جيدات، أنا لست أملأ ورق، أنا أريد أن يلمس الناس في أولادهم ما أبذل من جهد لصناعته في المدرسة، والتعليم كله عبء وهو مسؤولية عظيمة، وأنا رضيت أن أكون قائدة؛ لذا يجب أن أبذل جهدًا وأوثّق ما فعلته حقًا لأرضي الله عز وجل.

أنا: حسنًا، هل بعد اطلاعك على المعايير تشعرين بأنك ترغبين بالالتزام بها؟

المديرة: في الحقيقة هناك صعوبات في بعض المعايير؛ مثلًا وضعوا درجات على حضور بعض المشاركات، مثلًا يطلبون أن يحضر مسؤول لفاعلية عندك؛ فكيف إذا كان لدى المسؤولة أعمال وغير متفرغة للحضور عندي؟ كيف يقلل ذلك من درجتي؟

أنا: من تجربتك، ما الذي ترغبين في تجنبه/ إزالته من تقويم المنظومة؟

المديرة: هناك نقطة مرهقة أتمنى أن تُجنب وهي أن يوضع من ضمن التقويم حضور قيادات لفاعليات المدرسة، في الحقيقة إن القياديات مشغولات؛ فكيف تلزمني بدعوتهم حتى ترفع درجة تقويمي وهم ليسوا متوفرين؟ أنا أقدم تجربة وأدعو هذه القيادية لكنها لا تحضر؛ وبالتالي تنقص درجتي في التقويم بلا ذنب مني! هذا لابد من إزالته لأنه لو وضعنا مؤشرًا لابد أن يكون قابلًا للتطبيق.

أنا: شكرًا لمشاركتك القيمة؛ وهذا مما يدفعني لسؤالك: ما الذي تعتقدين أن إضافته إلى التقويم ستكون مفيدة ولماذا؟

المديرة: أن يكون هناك معايير مناسبة لكل مدرسة اعتمادًا على بيئة المدرسة وموقعها وعدد طلابها، التعامل مع المدارس بأسلوب واحد لا يصح، فليس كل المدارس يصلح لها نفس الإجراء، هناك مدارس لا تملك بيئة مساعدة؛ فكيف أقومهم على عدم وجود استخدام تقنيات معينة وهي لا تتوافر أو لا يوجد لديهم مكتبة أو معامل؟! أيضًا هناك الكثير من المدارس قد تجاوزت هذه المعايير وتحتاج معايير أعلى وتحديات أعمق لتطبيقها كل عام جديد، وهذا مطلبي: معايير ومؤشرات خاصة بكل مدرسة عبر تحليل نقاط القوة والضعف لكل مدرسة، ومنها نستخرج المعايير والمؤشرات التي نتطلع إليها مع المحافظة على الأساسيات.

أنا: من تجربتك، ما التحديات التي تواجهينها أثناء القيام بهذا التقويم؟ أخبريني عن تأثير عملية تقويم المنظومة عليك شخصيًا.

المديرة: ليست صعوبة أو تحديًا، لكن وقتها يكون فيه توتر سائد داخل المدرسة، أنا واقعية لكن هم يأتون بصورة مثالية؛ مثلًا في فترة الفسحة يسمعون أصوات الطالبات؛ أنا كمديرة أرى أن صراخهم ولعبهم شيء طبيعي وطريقة للتعبير وفرصتهم الوحيدة الحرة، لكن قد يعتبرونه أثناء التقويم شيئًا غير طبيعي أو منقصة للمدرسة، أو يدخلون على فصل فيه معلمة اختارت اليوم تلقي درسًا بطريقة ليست فيها ورقة عمل أو لم تستخدم استراتيجية معينة؛ هل هذا غير طبيعي؟ لكن في الحقيقة هذه الأمور تحدث، لكن عندما يعتبرونها منقصة على الرغم من أن ملف المعلمة ممتلئ بالأدلة على استخدامها لكل استراتيجيات التعليم؛ لذا تصرفهم مثالي ومحرج ومبعث للتوتر والإرهاق؛ فمهما عملت فأنت لا تصل للكمال.

أنا: حسنًا، هل يمكن لي سؤالك: ما الذي يجعل عملهم أقل إرهاقًا لك؟

المديرة: في الحقيقة كانوا فريقًا جيدًا في التعامل، لكن هيبة نزول لجنة ومحاولة تكونين دقيقة في عملك مرهق.

أنا: ماذا تقصدين؟

المديرة: أقصد يكونون واقعيين قليلًا.

258

أنا: شكرًا لك على جميع المعلومات
..
الآن أريد أن أتحدث عن مؤشرات الأداء المفتاحية (KPI) في تقويم المنظومة .
هل تعكس أفضل الممارسات في الأداء المدرسي ولماذا؟
المديرة: إلى حد ما.
أنا: لماذا قلت "إلى حد ما" وليس "تمامًا"
المديرة: لأن بعض المؤشرات لا تقيس الواقع؛ فكل مؤشر المفترض يكون واقعيًا وليس دليلًا على نجاح وفشل،
مجموعات المدارس لا تتفق في غنى بيئاتها.
أنا: في رأيك كمديرة، هل يساعد نظام مؤشرات الهيئة على تحديد عوامل النجاح الرئيسية؟ لماذا تظنين أن الهيئة يمكن أن
تساعد مدرستك في تحقيق أهدافها وغاياتها برأيك؟

المديرة: أذكر بعض المؤشرات تتحدث عن قياس مشاعر شيء وجداني؛ مثل مؤشر يتحدث عن أداء الصلاة للطالبات، في
رأيي هذه العلاقة بين العبد وربه؛ فكيف أقيسها أو أعتبرها مؤشرًا، وهذه أمور دينية لا أعرف علاقتها بالمؤشرات أو
كيف نقيسها وتجعلني مراقبة للطالبات في شيء لا يقاس!
أنا: إن أتيحت لك الفرصة لإضافة بعض المؤشرات إلى مؤشرات الهيئة ماذا ستكون ولماذا؟
المديرة: أتمنى أن تكون هناك معايير مرتبطة بحرية الرأي والمشاركة في إدارة المدرسة من قبل الطالبة والمعلمة.

أنا: إذا كان لديك أي شيء لإضافته، من فضلك أخبريني .

المديرة: لا، وشكرًا لك.
أنا: العفو، وشكرًا لك.

Appendix X

A Sample of a Reviewed Transcript of an Interview with a Head Teacher

Me: Welcome my dear, thank you for coming and meeting me; undoubtedly, you will enhance my research, making it more valuable and useful.
I will ask you some questions about your background, education, and some of the experience that you have gained of the organisation during the evaluation process. You can decide whether you withhold or give me your name, but it will never be disclosed by me in this study.
Firstly, I have a question about your academic background, as follows:
What educational and professional qualifications do you have?

Head teacher: My name is… and I have a bachelor's degree in History. I have 23 years of professional experience… I was employed immediately after my graduation, so I have gained a great deal of experience over the course of my career.

Me: How many years have you been a head teacher here?

Head teacher: I have worked here as a head teacher for seven years.

Me: What other school experience do you have?

Head teacher: I have a great deal of experience from other schools. I have worked here in Jeddah as a school principal. I have also performed administrative tasks during this period. I later worked as a head teacher and gained experience in dealing with all kinds of people: their diversity, their individual differences, whether students or teachers. I have also gained experience of dealing with supervisors, the Ministry, and of evaluating performance, especially the assessment of teachers.

Me: Can you describe your role in this school for me?

Head teacher: Here, I am like a mother, with transparency… I prefer this description, although many statements have been made about the role of the school principal, I see the school principal as a mother, and she is always the head teacher and responsible for everything. In the end, the mother is always afraid for her children and trying to affect them in positive ways for their own good. The school head teacher must first and foremost consider herself to be a mother.

Me: This does not nullify your role as the head teacher, surely?

Head teacher: Of course not, I see myself as very assertive, formal, and professional in my work.

Me: Thank you, this has been very useful. Can we now talk about the organisation, your experience of evaluation? And please tell me how you learn about the time of this evaluation.

Head teacher: I find out about it through various meetings and conventions.

Me: Can you specify who provides this knowledge?

Head teacher: The supervisors of the Education Offices in the north of the city. They are in charge of our school and they clarify it to us. They also tell us that we will be evaluated and then we will take many courses and receive lots of details, as there will be a new evaluation method implemented for us by the Ministry, which works on clarifying the standards and mechanism, and a list of indicators to cover almost every aspect.

Me: In general, what do you think about this approach and its impact on improving the education process?

Head teacher: Let me tell you frankly that our school has a plan for development and a new curriculum for the advancement of education, regardless of the Ministry's programmes in the school. It is probably already included in the policy and guidelines for the advancement of education. Moreover, classifying the teachers, students, achievement files, and everything else that has been introduced by the organisation was already applied by us several years ago. Our policy was based on the details that we can now find in the organisation. Thus, when the organisation appeared, we originally met its standards, except for some differences that just needed to be rearranged. Let me say that the organisation has helped us display our achievements, which we used to do before through documentation, but even without it we were walking in the same direction.

Me: Can you tell me if you know that other schools are practicing this, or is it just your school?

Head teacher: I do not know, but this way is my direction, insofar as my team keep up with developments in education.

Me: Thank you for the explanation, and now we can talk about the first step in evaluation by the organisation. Can you, as the head teacher of the school being evaluated, describe your experience of the organisation visiting your school?

Head teacher: They sent us a schedule, which included details of the organisation's team who would come to prepare our work. They arranged a tour for them at the school during school hours. By looking at the schedule, I knew that when they came, we would spend this day as normal. They entered and attended some of the classes, which were randomly selected, and I prepared all the files for them. Then, they read them carefully, reviewed the standards with me, and asked me about what I had achieved. Oh, it was a long day.

Me: The day when they attended?

Head teacher: They came early, and I cannot forget this day, because one of the teachers was pregnant and bleeding.

Me: Was it because of pressure due to the evaluation?

Head teacher: I do not think it was because of that, but she was sick.

Me: Are you sure?

Head teacher: It is difficult to prove that, but generally, I cannot forget this day.

Me: Let me ask you, do you think there is any difference between the visits made by the organisation and those made by the usual school supervisor?

Head teacher: The difference is that the usual supervisor does not make much effort, because she comes to do specific work, which involves a subject, but the organisation team makes a lot of effort, because they examine every detail. For example, I have a detailed plan that must be reviewed by them, and then they review the teachers' and students' achievement files.

Me: Do you feel that it is the additional effort exerted in your work that tires you?

Head teacher: In fact, it is a great effort, but the goal is to change education and my students are good. I want parents to feel and see in their children that the effort I make with them in the school is very tiring. Education is a huge responsibility and I am willing to be the head teacher, so I must exert all efforts and maintain my sense of purpose solely according to Almighty Allah.

Me: Well, after reviewing the standards, do you feel that you still want to work in this capacity?

Head teacher: In fact, there are difficulties in some of the standards, such as ranking some of the participants; for example, they ask active, responsible people about their ability to attend, but what if they cannot attend, because of their business or work? How does this reduce my ranking?

Me: According to your experiments, what do you want to avoid/remove from your evaluation by the organisation?

Head teacher: There is a stressful point that I wish to avoid, which is to require the leaders to come to school events. In fact, the leaders are very busy, so how can I be obligated to invite them to raise my evaluation grade, if they are not available? I have already attempted this as an experiment and invited these leaders, but they did not attend. Therefore, my ranking in the evaluation will be reduced through no fault of my own. This obligation should be eliminated, because if there is an indicator, it should be applicable.

Me: Thank you for your valuable advice, which prompts me to ask you what you would find useful, and what should therefore be introduced into the evaluation and why?

Head teacher: There should be appropriate standards for each school, depending on the school environment, location, and number of students. Using one treatment approach with every school is not right, because not all schools are suitable for the same procedure. There are schools with no helpful environment. Therefore, how can I evaluate them for not using facilities that are unavailable to them, where there is no library or laboratory? However, there are also many schools that have exceeded these standards and need higher ones, with steeper challenges to apply each year. This is my request, which includes specific indicators and standards for each school by analysing their strengths and weaknesses. From this, we can extract the standards and indicators that we anticipate by preserving the basics.

Me: In your experience, what challenges do you face while this evaluation is being undertaken? Please tell me about the impact of the organisation's evaluation on you, personally.

Head teacher: It is not difficult or challenging, but at this time, there is some tension in the school. When they come, I try to be realistic, but their evaluation is not fair and requires ideal manners. For example, during breaktimes, I consider the students' shouting and playing to be normal; it is the only way for them to enjoy some freedom. In contrast, when they (the organisation) hear these shouts, they probably see them as abnormal, which impairs the school evaluation. Moreover, when they go to a teacher's classroom, where she has chosen on that particular day to undertake the lesson without a worksheet or has not used a particular strategy, they see it as abnormal. In fact, these things happen to us, because they see it as neglect by the teacher, even if her file is full of evidence of using these strategies in class, but their way is too ideal, embarrassing, leading to tension and exhaustion. In fact, whatever you do, it never leads you to perfection.

Me: Well, can I ask you why their work fatigues you?

Head teacher: In fact, they were a good team in their treatment of us and to deal with, but their prestige and our efforts to be careful in the work are exhausting.

Me: What do you mean?

Head teacher: I mean, they must be more realistic.

Me: Thank you for all this information. Now I want to talk about the key performance indicators in the organisation's evaluation. Do they reflect the best practices in school performance and why/why not?

Head teacher: To some extent.

Me: What do you mean by 'to some extent', as opposed to 'a great deal'?

Head teacher: Because some indicators do not measure reality. Each indicator is supposed to be realistic, not to be evidence of whether a group has been successful or failed, because it does not correspond, due to the richness of its environment.

Me: In your opinion, as the head teacher, does the organisation's system of indicators help to identify the main success factors and why do you think the organisation can help your school achieve its goals and objectives?

Head teacher: I remember some indicators that measured feelings about very personal things; for example, there was an indicator that mentioned the students' prayer performance. In my opinion, this involves the relationship between the worshiper and Allah, so how can I measure this or consider it as an indicator? These things are regarded as religious affairs, and I do not know about their relationship to the indicators or how they are measured. Accordingly, I am meant to observe students on a matter that cannot be measured.

Me: If you had the opportunity to add some indicators to the ones that already exist, what would you add and why?

Head teacher: I would like to see standards related to freedom of opinion and participation in the school administration by students and teachers.

Me: If you have anything to add, please don't hesitate to do so.

Head teacher: No, thank you.

Me: Pardon me and thank you too.

Appendix XI

Example of Coding the themes

The headteacher: It is not difficult or challenges, but at this time there are some tension in the school. When they come I try to be realistic but their evaluation do not fair and require ideal manners. For example, during the break, I see the shootings of students and their playing are normal and it is the only way to get their free opportunities. On the other side, when they hear their shouts, they probably see these things are not normal way which decrease the school evaluates. When they go to a teacher classroom whose chose in this day to practice the lesson without the worksheet or did not use a particular strategy, they see that are non normal way. In fact, these things happen to us because they see that as neglect occurred by the teacher although her file is full evidences of

- Like what?

Like community partnership, this supported me and helped me get financial support to help poor students as well as benefit from the experiences of mothers to help us train students.

- In your opinion as a manager, does the method of the system indicators help identifying the key success factors?

Yes, by a large percentage, it makes a good manager

- Why do you think that the system can help your school to achieve its goals and objectives?

Because it organized the work and provide me with a guide to follow. It also taught me how to arrange myself and my work at school.

Appendix XII

Example of Coding the Qualitative Data

Examples of Quotes Used	Round Codes	Final Codes
• I know that through various meetings with my supervisor. • I knew about it through courses and from the Organisational Guide. The head teacher told us about the system, its criteria, and the indicators.	• The government's dominance over organisations. • Poor knowledge of the SPIS system. • Lack of information.	• SPIS monitoring awareness. • Decentralisation. • The quality and stress of the process.