Art. #2472, 12 pages, https://doi.org/10.15700/saje.v45n2a2472

Whole or skewed: Implementing whole-school evaluation policy in South Africa

Thabo Happy Peloyahae and Ke Yu
Department of Educational Leadership and Management, Faculty of Education, University of Johannesburg, Johannesburg, South Africa
kyu@uj.ac.za

Abstract
In this article we examine the wholeness (or skewness) of implementing South Africa's whole-school evaluation (WSE) policy. Drawn from a doctoral study in which the focus was on internal whole-school evaluation (IWSE), we investigate the relationship between the intended and the realised dual purpose of WSE for school improvement and accountability. We investigate the relationship between IWSE and other school evaluation instruments, the dual purpose within IWSE and any potential tension and/or cooperation between the 2 purposes, as well as how the tension and/or cooperation manifest. Through exploratory sequential design, data were collected from both semi-structured interviews and online questionnaires in close to 20 schools. The findings highlight a generally greater emphasis on accountability. However, many contradictions also emerge regarding whether the perceived emphasis on accountability prevails.

Keywords: accountability; internal whole-school evaluation; school improvement; South Africa; whole-school evaluation

Introduction
School effectiveness, which often lies at the heart of any school evaluation, is an essential indication of a school's ability to achieve its mission and objectives (Scheerens, 2000). However, the mission of schools (or education in general) and what is regarded as school effectiveness can be conceived and operationalised differently. During the apartheid era in South Africa, the dominant ideology encompassing compliance, control, exclusion and suppression meant that school evaluation was compliance-driven, where obedience was the main criterion (Kumalo & Skosana, 2014). External evaluation alone sufficed for this purpose. In other countries, such ideology was less prevalent. However, historically, an over-reliance on external inspection systems to monitor the quality of education and school effectiveness was also the norm around the world (Hossain, 2017).

A major shift occurred in the late 1990s (Scheerens, 2002) in the form of a move towards school self-evaluation (SSE) and corresponding to a broader acceptance of democracy, decentralisation and responsiveness to local demands (Fushimi, 2014). The aim with SSE is not to replace external inspection but is often conceptualised as a complementary process for school monitoring and evaluation (Madikida, 2016; Organisation for Economic Co-operation and Development [OECD], 2013). Typically, external evaluation remains an important snapshot exercise that emphasises objectivity and accountability; SSE, on the other hand, tends to be an ongoing process that focuses more on reflection and improvement (Chapman & Sammons, 2013; Swaffield & MacBeath, 2005). The situation in South Africa is no exception. Since the country embraced democracy in 1994, the transition towards democracy and redress has also resulted in a re-conceptualisation of school effectiveness away from an adherence to subservience and bureaucratic controls towards school improvement and performance. In 2001, a policy on whole-school evaluation (WSE) was introduced, officially endorsing both external whole-school evaluation (EWSE[i]) and internal whole-school evaluation (IWSE,[ii] which is equivalent to SSE).[iii] The main aim with WSE is to empower schools (Govender, Grobler & Mestry, 2016; Setlalentoa, 2014) through an emphasis on inclusion and improvement (Steyn, 2002). The "whole" agenda reflected in the policy title is conceived and manifests in the following interrelated ways:

1) Internal and external evaluation. This refers to the inclusion of both internal and external evaluation components aligning with the international trend mentioned earlier.
2) Dual purpose of SSE. Although SSE and external evaluation are generally regarded as suitable for improvement or accountability purposes, SSE itself can also serve for both accountability and improvement purposes (MacBeath, 2005). Between these purposes, however, SSE for improvement tends to be highlighted in both policies and practice (Baker, Curtis & Benenson, 1991; Cobbinah & Eshun, 2021; MacBeath, Schratz, Meuret & Jakobsen, 2003; O'Brien, McNamara & O'Hara, 2014).
3) Stakeholder involvement. This refers to the wide range of stakeholders (for example, school management team [SMT] members, teachers, support staff, school governing bodies [SGBs], parents and learners) and the various ways (for example, interviews, discussions, classroom observations, and SWOT [strengths, weaknesses, opportunities and threats] analyses) in which stakeholders are involved in WSE.
4) Comprehensive school functions and areas. WSE proposes nine focus areas in its evaluation for both EWSE and IWSE.[iv] This wide range of coverage of school evaluation is unique to South Africa, as other countries such as the United Kingdom (UK) and Ireland (on whose policies WSE in South Africa is modelled), tend to focus on fewer aspects.[v]

Since the enactment of this policy, however, implementation gaps have been repeatedly reported. The most frequently reported implementation gap is skewed stakeholder participation (point 3 in the list above, Govender et al., 2016; Mncube, 2009; Nwosu & Chukwuere, 2017; Phukubje, 2019; Setlalentoa, 2011, 2014; Siafwa, Manchishi & Cheyeka, 2019). Despite the policy prescription of the wide inclusion of stakeholders in the process,

the dominance of SMT members and educators remains. The conceived wide range of areas that WSE seeks to address and the areas it actually covers (point 4 in the list above) has never been explored in the literature, thus requiring exploration. In this article we focus on the first two points (internal and external components and the dual purpose of SSE), by drawing on a doctoral study in which the four aspects mentioned were examined. The primary focus of the doctoral study was on IWSE, but it also touched on EWSE and school improvement plans (SIPs) in a search to determine schools' realisation of their accountability, improvement and effectiveness agenda. Fulfilling the purpose of improvement or accountability is reported in some WSE studies, but many are only in the form of unpublished dissertations and theses. The findings from these studies reveal that the schools tended to adopt a compliance mentality. This is reflected, for example, in many schools only conducting IWSE when selected for EWSE, instead of doing it annually (Phukubje, 2019; Sehlapelo, 2021) or not following through on IWSE recommendations and action plans (Govender et al., 2016). However, published works and dissertations seldom explore the relationship between the purposes of improvement and accountability in terms of the relationship between IWSE and other school evaluation instruments. Publications also do not examine the dual purposes within IWSE, any potential tension and/or cooperation between these two purposes, or how the tension and/or cooperation manifests. The aim with this article was to examine these factors.

We begin with a brief review of the literature relating to the history of WSE, its articulated and potential purpose for improvement or accountability, and the potential tension between the intended and realised purposes. This is followed by an introduction to systems theory which provides the theoretical framework and methodology of the study, especially the study's exploratory sequential mixed method design and sampling. In the results section we present the data, highlighting a seemingly greater emphasis on accountability mentality compared to improvement but with the presence of contradictions that actually dominate. We further expand and explain this contradiction and inclusiveness in the discussion.

Literature Review

The WSE's aim for inclusivity and wholeness emanates from South Africa's history. The autocratic and punitive external inspection process prior to 1994 was associated with intimidation, resistance, fear and negativism where a punitive mentality dominated the system (Amoako, 2014; Reddy, 2005). School visits were unannounced and the outcome of the inspections often resulted in the transfer or dismissal of non-compliant teachers (WITS Education Policy Unit, 2005). As a result, schools, educators and educators' unions rejected these inspections, which led to a culture of resistance and suspicion (Jansen, 2004). From early 1990, neither educators nor schools in South Africa were externally evaluated (Department of Basic Education [DBE], 2000).

Given this history, the WSE policy introduced in 2001 (Department of Education [DoE], Republic of South Africa [RSA], 2001) deliberately emphasises the purpose of improvement in its rationale:

- To recognise good schools and support under-performing schools; and
- To improve overall education quality "to ensure that all our children are given an equal opportunity to make the best use of their capabilities" (p. 7).

This WSE policy, which highlights the term "whole", is modelled after that of Ireland, through the stewardess and pioneering of Professor Kadar Asmal, who spent over 25 years in Dublin during exile under apartheid and was appointed as the South African Minister of Education in 1999.

SSE is typically promoted in other countries as a result of the improvement agenda highlighted in WSE which stresses schools' conscious efforts to achieve school effectiveness through self-examination (DoE, RSA, 2001), reflection on and modification of teaching or management practices (Kobola, 2020; Ryan, 2011). Together with other tools such as SIPs, the findings from the doctoral study show that IWSE is often conceived as the key to improving school effectiveness and performance. In contrast to apartheid inspections, which had no criteria, WSE pinpoints specific areas with evaluation guidelines. The inspectors are called supervisors and they are reminded to focus on developmental feedback. Overall, Steyn (2002) asserts the importance of primarily regarding WSE as a supportive and developmental tool in the new dispensation.

Even though accountability does not feature prominently or directly in the text of the WSE policy itself, it probably should not come as a surprise that the accountability purpose remains crucial in the WSE conceptualisation, especially considering the multitude of challenges facing the South African education system, including under-performance, a tendency to skirt around regulations, corruption, and poor financial management (DBE, RSA, 2016; Govender et al., 2016). In addition, schools, like any other organisations, operate within a regulatory environment and context. Therefore, compliance with the relevant policies and laws that govern the sector is not unexpected.

The manifestation of the purpose of accountability or improvement is often seen in the implementation itself. For example, if IWSE is only implemented when schools are selected for EWSE (instead of implementing IWSE irrespective of whether they are selected for EWSE or not), this

could be an indication of a compliance mentality. So too is the practice of participating in or implementing WSE for record-keeping and submission only, instead of progressing to its natural next step of drawing, implementing, and monitoring an SIP. An emphasis on objectivity might also downplay the value of IWSE in terms of self-reflection, deliberation, self-learning and ownership. Other more subtle forms of downplaying the value of IWSE might include delegating IWSE implementation to lower ranking personnel instead of actively seeking and taking leadership over the process, might also indicate a conscious or unconscious playing down of the improvement purpose. In this article we report on the detailed and nuanced manifestation of IWSE implementation to determine whether any potential tension and/or cooperation between these two purposes exist and to discover whether and how one of these purposes prevails.

Theoretical Framework

Systems theory was the guide to the theoretical framework of this study. Systems theory is derived from multiple disciplines including biology, economics and engineering (Yoon & Kuchinke, 2005). Concepts like wholeness, differentiation, domination, growth, hierarchical order, competition and control are characteristics of an organisation and considered by the systems theory through its primary principles, including the principle of system holism, the principle of system orderliness, the principle of hierarchical levels, and the principle of dynamic interrelation (Von Bertalanffy, 1968:47). Among these, the principle of system holism is most closely related to this study. This principle highlights the necessity to examine a phenomenon holistically (a forest) instead of through individual elements alone (each tree). It asserts that interlinked subsystems within the system constantly influence each other and affect the behaviour of the whole (Martinelli, 2001; Wagner, Kegan, Lahey, Lemons, Garnier, Helsing, Howell & Rasmussen, 2006). Von Bertalanffy explains that systems theory "is a general science of 'wholeness'… [where] constitutive characteristics are not explainable from the characteristics of isolated parts" (1968:37, 55). In addition, all processes are continuous where feedback within and between different tiers of the system exists and their impact on the system as a whole is considered to be crucial.

Systems theory can be applied to all four aspects of the wholeness focused on in the larger study. The corroboration of IWSE with EWSE helps to provide schools with a more comprehensive assessment that incorporates both insider and outside views, allowing one to explain an organisation's behavioural patterns from both inside and outside (Ivancevich & Matteson, 2002). To achieve the aim that "whole-school evaluation is not an end in itself, but the first step in the process of school improvement and quality enhancement" (DoE, RSA, 2001:8), both IWSE and EWSE findings should be used to corroborate SIPs and school improvement. IWSE's comprehensive inclusion of nine areas is echoed by its acknowledgment that schools' various areas of functioning influence each other and collectively determine school effectiveness. Finally, the wider range of stakeholder involvement ensures that participants do not automatically blame each other or other elements outside the organisation (Senge, 1990).

Methodology

The research question in the larger doctoral study was why IWSE was implemented in the way it was. For this article, the research question is more specific on how the dual purpose of WSE was conceived and negotiated. To achieve both the broad and narrower research aim, an exploratory sequential mixed method design was used where qualitative interviews were conducted prior to the completion of quantitative questionnaires (Cohen, Manion & Morrison, 2011). In other words, the data analysis of the qualitative phase informed the design of the quantitative instrument "to test or generalise the initial qualitative results" (Schoonenboom & Johnson, 2017:117). The planned and actual sample in terms of both the schools and participants within the schools are presented in Table 1.

For the qualitative research, purposive sampling was used to select schools and participants. Applying maximum variation sampling, the goal was to capture as many different perspectives as possible. For schools, school characteristics, such as whether they had been externally evaluated in the previous 3 years (2017–2019) prior to data collection in 2020, whether they were primary or secondary schools, and whether they were township or urban schools were also considered. The difference between planned and realised was largely due to coronavirus disease (COVID-19) regulations, which coincided with the data collection. This implied that only the former two criteria were realised for schools in the qualitative phase; in the quantitative phase, the latter two were realised while all schools were externally evaluated in the past 3 years (see Table 2). It was originally planned to triangulate quantitative data with qualitative data by including the same schools for both instruments. However, this also didn't happen due to COVID-19 regulations to minimise contact with individual schools.

Table 1 Planned and actual sample

	Schools		Participants/Respondents (per school)	
	Planned	Actual	Planned	Actual
Qualitative	10	4	12 each	5 each
Quantitative	20	11	12 each	141 in total, average 13 each, varying per category and school

Table 2 Actual sample distribution: schools

	External evaluation in the past 3 years or not	Primary or secondary	Township or urban
Qualitative	2 yes: 2 no	2 primary: 2 secondary	All township
Quantitative	All yes	4 primary: 7 secondary	5 township: 6 suburb

The school samples consisted of SMT members, educators, support staff and SGB parents. The planned participants for both qualitative and quantitative phases were 12 from each school: three SMT members, seven educators, at least one support staff member and at least one SGB parent. Varied successes were achieved in the data collection. Ultimately, five participants at each school were interviewed during the qualitative phase: three SMT members (one principal, one deputy principal, one head of department) and two educators. For the quantitative phase, all four categories were targeted. This greater inclusion was realised in seven schools, although two schools missed one category; one school missed two categories; and only one category (SMT) was realised in a third school.

We also considered to include district officials and secondary school learners. However, district officials were excluded due to expected bias as a result of their vested interest in positive findings (Vaganay, 2016). Although the inclusion of secondary school learners was initially planned, the learners participated in WSE through the representative council of learners (RCL) whose members were often in Grade 12. This meant that the RCL with knowledge of IWSE would have graduated by the time they were approached, while the current RCL would not have been exposed to the WSE. Grade 12 learners also had heavier academic schedules; thus, upon further discussion with the school principals, they were ultimately excluded.

Permission was requested and obtained in writing from the provincial DoE, school principals and SGBs. The participants were assured that the names of the schools and individuals would not be mentioned in the research report and that the responses would not be traced back to individual participants. The participants were then requested to sign consent forms before participating. The online questionnaire was distributed by the school principals.

In the data analysis, content analysis was applied to the qualitative data and the key issues emerging from the responses were identified. The questionnaires were validated through exploratory factor analysis, tests of normality, and one-way analysis. In this article, results from factor analysis, cross-tabulation (cross-tab) and group comparison (using the Statistical Package for the Social Sciences [SPSS] with the assistance of a statistician) are reported.

Results

The main finding from this study was a seemingly greater emphasis on an accountability mentality compared to one of improvement. However, several contradictions also emerged, thereby making the drawing of a conclusive conclusion impossible. This finding was substantiated through the various themes discussed below.

Improvement Purpose was Acknowledged but, Overall, Utterances on Improvement Tended to be Normative

A small number of the interviewed participants, who tended to be SMT members, were aware that the school could or should become a learning organisation, which, crucially, aligned with the improvement purpose. They were prepared to learn from WSE in order for their school to improve. For example, Schools A and B did not relegate the tasks of providing support or IWSE training solely to external entities but took the initiative to train stakeholders on IWSE themselves. However, the sentiment was often not shared among other participants, including other SMT members.

A few participants (especially those in SMTs) also indicated their preference for EWSE due to its perceived thoroughness, better objectivity, accuracy or fairness. This preference indicated their yearning for improvement but also raised questions as to why IWSE, which bore greater relevance to improvement, failed. The following quotations reflect the opinions of various participants:

> *The EWSE process was a thorough, detailed and most valuable process. EWSE went well and all the stakeholders were involved. EWSE was accurate. There was no deviation* [from WSE policy procedures] *and the recommendations were detailed. Even the request of the school was included in the recommendations.* (Principal, School B)
>
> *EWSE was a fair procedure as compared to IWSE. EWSE is objective and honest, in-depth and fair. The recommendations are specific, and the evaluators are well-trained and knowledgeable* (Principal, School C).

> *EWSE was a fair process and the recommendations were positive and encouraging. The way EWSE was conducted was professional … EWSE highlighted some of the issues that were overlooked by IWSE* (Deputy principal, School C).
>
> *The process of evaluation by the EWSE team was intensive and spot-on. The process was developmental and identified the weaknesses that were overlooked by IWSE* (Head of department, School D).
>
> *IWSE is not important since it does not identify the schools' shortcomings and does not provide a plan to improve the schools' performance* (Educator, School A).

Overall, the participants agreed on the necessity for an improvement agenda, although their utterances tended to be normative; they used words like must, should and need.

Although what could or should have been done to move forward (for example, in the form of a vision, a hope or a desired state) was needed – likely through implementation itself – this transition did not happen automatically. Similarly, quite a few participants spoke about the value of identifying strengths and weaknesses in order to move towards improvement; again, these tended to be largely normative. One educator at School D thought that IWSE was only to "assist the schools that underperformed to improve." A few others especially pointed to the issue of accountability.

> *The purpose of IWSE is to ensure that all schools are managed and well-resourced. It is also to make schools accountable for the performance of learners and staff* (Educator, School A).
>
> *The purpose of IWSE is to evaluate the school to monitor if things are done* (Educator, School B).
>
> *IWSE is a useful tool to assist to gauge the productivity and effectiveness of the school* (Educator, School C).

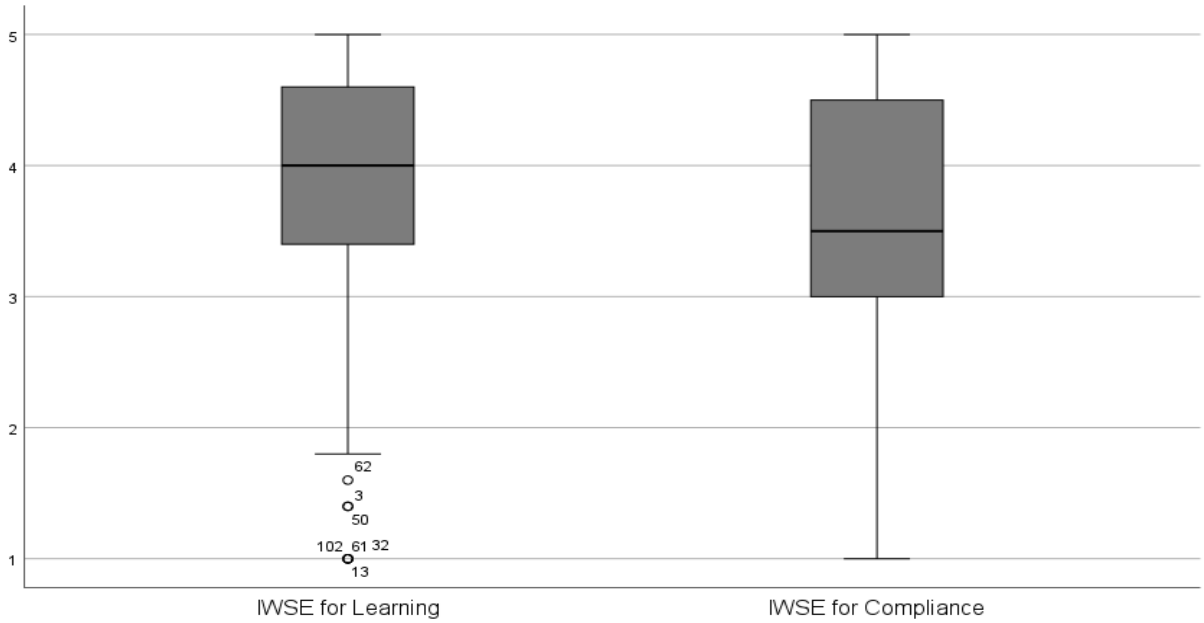

Figure 1 Box-plot on the IWSE for learning and IWSE for compliance

Factor analysis in the quantitative phase revealed additional contradictions.[vi] Firstly, while qualitative data suggests more accountability, the composite factor for IWSE for learning was higher than that for compliance (see Figure 1). This figure also shows that both items covered the full range, from strongly agree to strongly disagree, indicating a wide range of discrepancies for both items. The quantitative data hosted another set of contradictions. On the one hand, the positive benefits that IWSE brought to schools were overwhelmingly agreed upon (78%).[vii] Simultaneously, opposite statements also received a high level of agreement (all over 50%). This was particularly high for statements such as "the school can do well without IWSE" and "school focus should be more on teaching and learning than on IWSE" (58.9% and 61.7% respectively).

The participants' observations on the implementation of IWSE (both in interviews and on the questionnaire), as well as their concerns and overall negative attitudes towards the implementation, suggest that the suboptimal implementation compromised the improvement purpose.

Compliance Mentality was Evident but with Contradictions

Similar to what was reported in other studies (Bayeni & Bhengu, 2018; Carlson, 2009; Govender et al., 2016; Mathaba, 2014; Phukubje, 2019), we found evidence of the implementation of IWSE for compliance. The deputy principal at School D

frankly admitted that "*the principal will always remind the school that the staff must comply with the instructions of the district to avoid being on the bad side of the law*." The head of department at School B similarly cautioned that "*the school SMT is more concerned about the image of the school and therefore develops the IWSE document to satisfy the district instruction*." Giving the example of the last time the school prepared to conduct IWSE, an educator at School A concurred: "*IWSE is only conducted when the school is about to be evaluated externally by a team from the head office … When we were informed that the evaluation was postponed, the IWSE process was stopped*."

Others at the same school admitted that "*IWSE was not done since the school was not expecting EWSE*" (Educator, School A). The school principal explained: "*The school was not evaluated by EWSE in the past 3 years, therefore, we did not conduct IWSE in 2018 and 2019 since the school only conducts IWSE when informed of EWSE visits*."

One of the educators at School D recalled the occasion when "*IWSE was not conducted since there were no planned EWSE visits*" and, on other occasions, "*it was done in a hurry since it was required by the district on a specific day*" (Head of department, School D).

The participants also observed an overall defensive attitude towards criticism at their schools, similar to Clift, Nuttall and McCormick's (1987) finding, which could be the reason for their negative attitudes towards participation in and implementation of IWSE. What seemed to be happening was a vicious circle where the participants' views on issues like management style, resource provision, the role of the SGB and the use of finances were rejected or silenced during staff meetings. Therefore, IWSE meetings were regarded and used as the only opportunity to have their voices heard (particularly at Schools C and D). The participants further pointed to conflicts where different opinions or anything that was not in line with the SMT's views were often automatically disregarded and reacted to negatively. In a milder version, "*negative comments and complaints*" at the IWSE meetings were labelled as a "*waste of time*" (Deputy principal, School B), an irritation or an attack on leadership and management qualities. Overall, it seems that many schools had not developed a school culture that embraced critical deliberation and debate that prioritised reflection and learning.

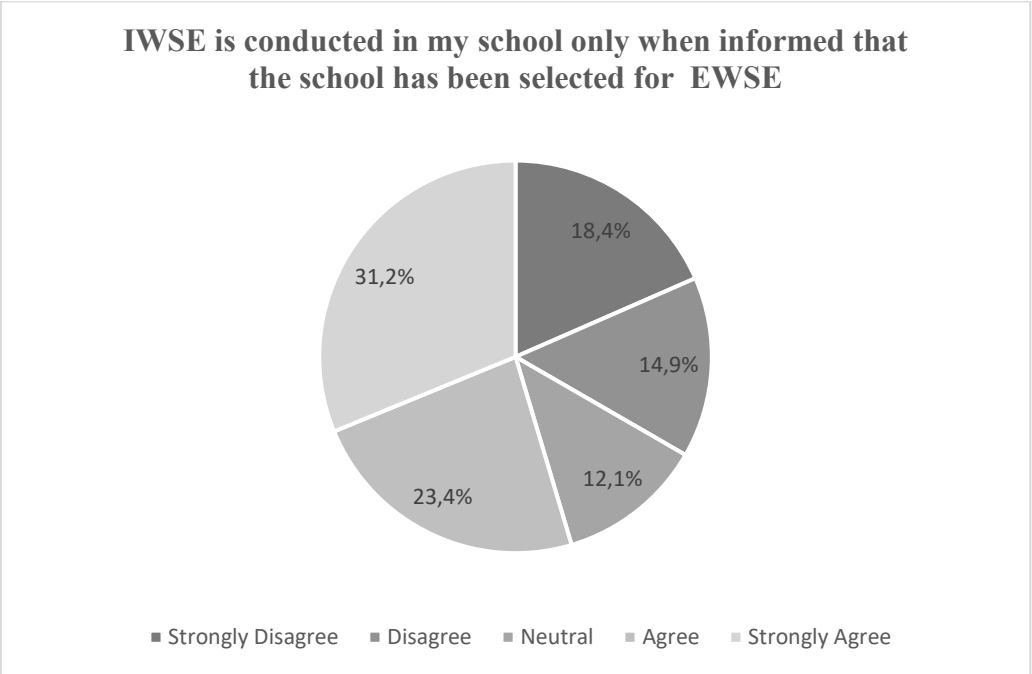

Figure 2 Degree of agreement on the statement "IWSE is conducted in my school only when informed that the school has been selected for EWSE"

In the quantitative data (see Figure 2), slightly over half of the participants (54.6%) agreed with the statement, "IWSE is conducted in my school only when informed that the school has been selected for EWSE." Those who disagreed with the statement, that is, those who responded with disagree or strongly disagree, represented less than one-third of the sample (33.3%). A further indication of a compliance mentality can be seen in the low frequency of IWSE implementation. If the schools conducted IWSE annually, all the years should have been marked when the participants were asked to tick the number of years in which IWSE was implemented (that is, 100% for each year in

Figure 3). However, this was clearly not the case. Instead, the extent of IWSE implementation was at approximately 30%.[viii]

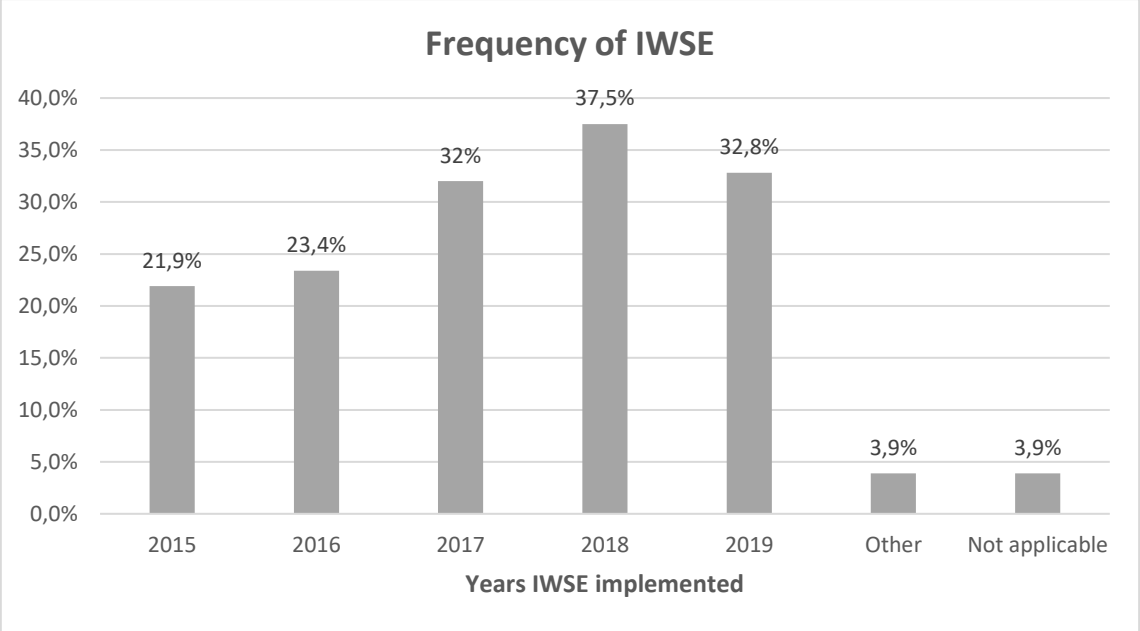

Figure 3 Frequency of IWSE in schools

One observed challenge relates to the objectivity and subjectivity concern where 35.5% of the participants did not think that the IWSE findings truly reflected what was happening at their schools. This is similar to the finding by Sehlapelo (2021).

In addition, a cross-tab with participant demographics showed extreme variance across the schools without a concrete pattern. However, a clear pattern emerged where agreement with the statement "IWSE is conducted in my school only when informed that the school has been selected for EWSE" was higher among educators and support staff. The high agreement from the support staff might have been due to the small number of participants in this category. For SMTs, their opinions were almost equally split. For SGB parents, approximately one third were neutral, indicating a much lower level of knowledge on this matter. This could suggest that a stronger improvement purpose was acknowledged by the SMTs, but it could also signal a potential tendency to positively portray the school or reveal a communication gap between SMTs and other members in the schools regarding the purpose of IWSE.

Detecting Improvement or Accountability Mentality through SIPs

Some of the participants held extreme views: "*SIP is a waste of taxpayers' money*" (Educator, School B). The findings further indicate that most of the schools' SIPs were not based on IWSE findings, in line with the findings of Phukubje (2019) and Sehlapelo (2021). At School A, only the principal claimed that their SIP was developed from the IWSE findings while all the other participants at the same school disagreed. Similar situations were observed at Schools B and D. School C was the only school in the qualitative sample where most of the participants agreed that IWSE resulted in the development of their SIP.

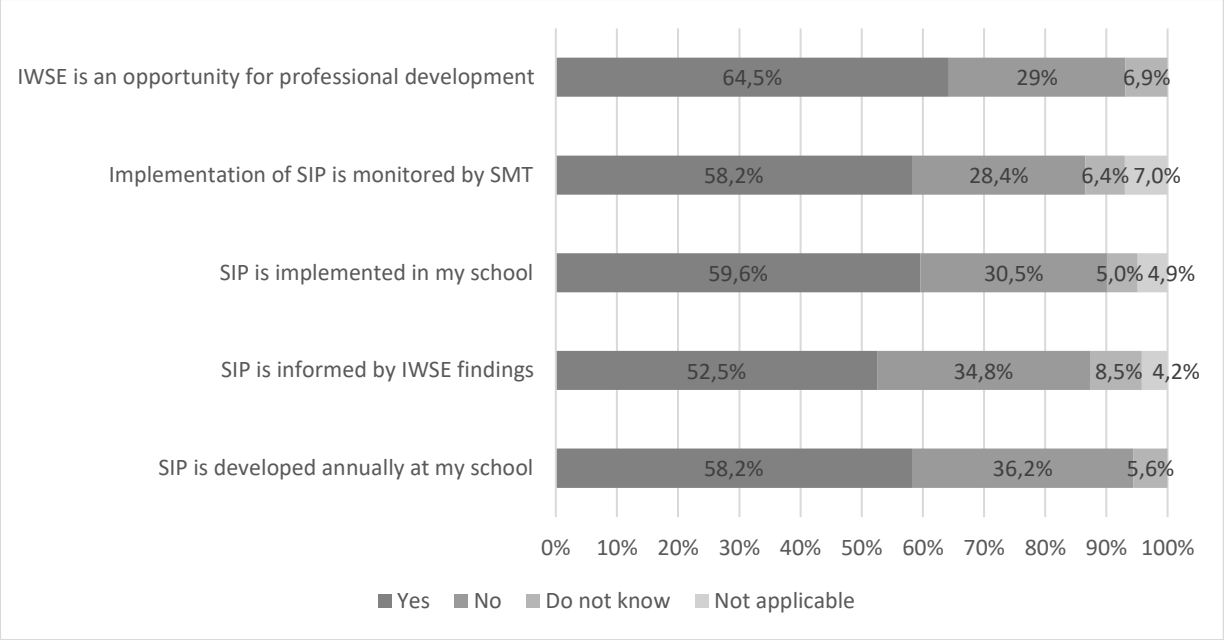

Figure 4 Agreement with statements on SIP implementation and monitoring

The quantitative data in Figure 4 indicate strong agreement of SIP implementation in a school (59%), followed by annual SIP development (58.2%), and an SIP that was monitored by the SMT (58.2%). Of all the participants, 52.5% agreed that an SIP was informed by IWSE findings, the latter being better than Mbalati's (2010) finding in Mpumalanga that most schools' SIPs were not informed by IWSE recommendations. In addition, over 64.5% of participants agreed that "IWSE is an opportunity for professional development." However, 29% of participants disagreed with the statement – suggesting that a sizeable number of participants did not associate IWSE with development.

A cross-tab with the schools again showed a vast but unclear pattern regarding the position, although the SMT was clearly the group that was most likely to claim annual SIP development, with almost one-third disagreeing, which was similar to the result showing implementation of IWSE only when a school was selected for EWSE. This is cause for concern. The overall situation for educators is similar to that of SMTs. Overall, the trend is a decrease in agreement with this statement, accompanied by an increase of disagreement and a do-not-know response.

Discussion

Stakeholder exclusion is not unique to South Africa (Katsuno & Takei, 2008; O'Brien, McNamara, O'Hara & Brown, 2019) nor is the inadequate consideration of the improvement agenda reported in this article. For example, this tendency towards a more accountable and compliance-driven mentality has also been reported in England (Dangerfield, 2012; MacBeath, 2005), Ireland (Brown, McNamara, O'Brien, Skerritt, O'Hara, Faddar, Cinqir, Vanhoof, Figueiredo & Kurum, 2020) and elsewhere. The data from this study indicates a complex picture regarding the tendency towards either improvement or accountability – a similar finding as in studies by Phukubje (2019) and Sehlapelo (2021). The qualitative data collected suggest that the participants favoured accountability more than improvement (for example, the relationship between IWSE and the SIP and attitudes towards learning, development, and criticism). The necessity of IWSE tends to be mentioned primarily in a normative sense, while attitudes suggest more negativity. Some participants struggled to articulate the purpose of IWSE besides any general and normative utterances and few related the IWSE's purposes to identifying strengths and weaknesses. However, other quantitative and qualitative data (such as taking the initiative for IWSE training) present an encouraging contradiction. The quantitative data show higher agreement on IWSE for learning, although other contradictions complicated the conclusion. Although the participants agreed with the potential of IWSE, they also thought that schools should focus more on teaching and learning than on IWSE or that schools could do well without IWSE. In addition, there was strong disagreement that the IWSE finding truly reflected what was happening in schools, potentially further undermining the improvement potential of IWSE.

Considering these in relation to the research question on how the dual purpose of IWSE were conceived and negotiated, what emerges is confirmation of the existence of both (contrasting)

purposes, potentially pulling towards inconsistent implementation. The observed dominance of a normative and other contradictory stance on IWSE further compromises their endorsement for this improvement purpose. This is in line with some participants' struggle to articulate the purpose of IWSE reflecting their ambiguity or resistance toward the process. When participants are unclear about the direct benefit of IWSE or they feel that it is overly bureaucratic, they may have difficulty engaging with it meaningfully, resulting in a limited sense of ownership, which in turn makes it difficult for them to view IWSE as a positive, actionable tool rather than just another compliance task. This could explain why participants reverted to generalities rather than discussing specifics of IWSE. However, whatever the reason, such an ambiguous attitude might lead to a tendency to fall back to the default short-term focus dominated by accountability, potentially at the expense of real longer-term improvement needs. For example, "teach for test" would produce an immediate return on better assessment results without contributing to real learning (Tóth & Csapó, 2022:463).

This is not to say that the participants did not agree with the school's objective of effectiveness or the need to improve. In fact, overall they subscribed to this aim (for example, as shown in their preference for EWSE), however, they did not necessarily see the value of IWSE implementation in assisting school effectiveness. In other words, this negativity or inability to realise the improvement agenda seemed to arise from the suboptimal implementation itself. In addition, good intention and emphasis on improvement with inadequate recognition of the need of, for example, accountability without clear or deliberate deliberation thereof, might also have exacerbated the impression and perception that the policy was merely (or more) about talk the talk (not walk the walk) (Rapoo, 2016).

An interesting point was some discernible differences in terms of the participants' positions in relation to their orientation towards improvement or accountability. The SMT members seemed to lean more towards an improvement and learning purpose, while educators and support staff leaned more towards accountability. Some SMT members tended to portray a better implementation at their school while other schools seemed to be more likely to retain a critical and honest stance. Combined with the over-representation of educators in the sample, this portrayal of better implementation could explain the overall tendency towards accountability. School characteristics such as external evaluation, location or level did not seem to play any role in how improvement or accountability was conceptualised or perceived.

Limitations

One main limitation of the study was the size of the sample. Not only was the study limited to only two of the 15 districts in the Gauteng province, but events in one province can also not be said to represent what is happening in the country as a whole. In addition, both realised samples for the qualitative and quantitative components were impacted by COVID-19, which further reduced originally intended variations and planned triangulation between the qualitative and quantitative data. Furthermore, schools' self-selection (willingness to participate) as well as the school principals' gatekeeping role (influenced access to other stakeholders in the schools) might also have influenced the data.

Another set of limitations arises from the nature of the research design and instrument. Using a survey (with items such as level of agreement) inherently limits engagement in probing further why participants choose a certain option. In addition, the normative (dis)agreement may only represent surface-level (dis)agreement where participants, for whatever reasons, were unwilling to be forthcoming with their true perspectives. Although attempts were made to ensure anonymity, no tracking back to participants, and an indication that there were no right/wrong answers, the actual effects of such efforts are hard to determine. This might have been further hampered by the nature of the online questionnaire where it was impossible to go back to the same participants for member checking or verifying interpretations. All these resulted in responses from the quantitative component taken at face value without possibilities for further investigation.

Conclusion

Drawn from a doctoral study that focused on IWSE, we investigated whether WSE in South Africa demonstrates a healthy balance between the two purposes of school improvement and accountability or whether IWSE implementation shows skewness towards one.

The overall finding suggests a generally greater emphasis on accountability opposed to improvement. However, many contradictions also emerged, making the drawing of a conclusive conclusion impossible. This suggests a complex picture regarding the tendency towards either improvement or accountability: there might not always be a clear distinction between these two purposes of IWSE. In terms of the contradictions that we detected, it might have been caused by the dynamics within individual schools or different positions. It is also possible that the contradictions might have arisen from a difference between what participants hoped (normative) and what actually

happened (actual). Although this was a relatively small-scale study, it provides much-needed knowledge on WSE, particularly in terms of the relationship between IWSE and other school evaluation instruments, the dual purpose within IWSE, potential tension and/or cooperation between these two purposes, and how the tension and/or cooperation manifest. The finding on the tension and ambiguity between the two purposes highlights a need for policymakers to reconsider the approach in WSE. It will not be easy to balance these potentially opposing purposes, but a more deliberate and conscious awareness of this could pave the way to a conscious effort to decide on the focus for the short term and the long term, mitigate the tension between the two in either the short term or the long term, as well as search for ways to tap into potential (if limited) synergies between the two so that they may work in the same direction. For example, accountability measures, such as standardised tests and performance metrics, can be used to provide quantitative data on learner achievement, teacher effectiveness, and overall school performance. This has been one of the most typical ways that these two purposes come together in practice. Feedback from such accountability purpose is crucial for creating continuous improvement cycles, where schools can reflect on their practices, implement changes, and measure progress over time. The key, and this is also what we found in this study, is a necessity to frame accountability as a supportive rather than a punitive process, both in rhetoric and practice. In addition, schools might also consider using accountability data to advocate for additional support or interventions that target the root causes of underperformance.

Due to its limited scale, however, future studies expanding the scale of this study to other parts of the province and country, and broadly more studies on the topic of IWSE and its "failure" to fulfil its improvement agenda are needed. In addition, ethnographic descriptions documenting exactly how various stakeholders think and negotiate the two purposes during WSE implementation are needed. This can also include their capacity and need.

Authors' Contributions

THP conceptualised the study and conducted all interviews. KY supervised the study. Both authors contributed to the writing and reviewing of the manuscript.

Notes

i. Schools are randomly sampled for EWSE by the Department of Education every 3 or 5 years; once every 5 years for primary schools and once every 3 years for secondary schools. EWSE is facilitated by a supervisory team from the Department, usually after IWSE.

ii. An annual evaluation led by the school's internal stakeholders.

iii. IWSE and SSE are used interchangeably in this article.

iv. The nine areas include basic function, leadership, governance, quality and teaching and learning and educator development, curriculum, learner achievement, school safety, infrastructure and parents and community.

v. WSE in the UK and Ireland focuses on four aspects: management, school planning, teaching and learning and pupil support (Department of Education and Science, 2006).

vi. Two factors were discernible from the quantitative data: corresponding to improvement (IWSE for learning) and accountability (IWSE for compliance). The questionnaire items for the improvement factor included: IWSE is a necessary and integral part of WSE; IWSE evaluates teaching and learning at public schools; IWSE helps my school to be effective; IWSE leads to school improvement; and participation in the IWSE process is an opportunity for professional development. The statements relating to accountability factor included: IWSE is conducted in my school only when informed that the school has been selected for EWSE; IWSE is implemented for compliance in my school.

vii. This figure is from agreement with the questionnaire item "the process of IWSE is of benefit to the school."

viii. It should be noted that this information is per participant, not per school, so it could be that their schools did not implement IWSE in those years, or that the school had implemented IWSE, but that the participants did not know about it (although in the latter case, one would have expected them to select "not applicable", which is 4% in Figure 3).

x. DATES: Received: 8 May 2023; Revised: 9 October 2024; Accepted: 23 February 2025; Published: 31 May 2025.

References

Amoako S 2014. Teacher unions in political transitions: The South African Democratic Teachers' Union (SADTU) and the dying days of apartheid, 1990–1993. *Journal of Asian and African Studies*, 49(2):148–163. https://doi.org/10.1177/0021909613479302

Baker P, Curtis D & Benenson W 1991. *Collaborative opportunities to build better schools*. Bloomington, IL: Illinois Association for Supervision and Curriculum Development.

Bayeni SD & Bhengu TT 2018. Complexities and contradictions in policy implementation: Lived experiences of three school principals in South Africa. *SAGE Open*, 8(3):1–12. https://doi.org/10.1177/2158244018792037

Brown M, McNamara G, O'Brien S, Skerritt C, O'Hara J, Faddar J, Cinqir S, Vanhoof J, Figueiredo M & Kurum G 2020. Parent and student voice in evaluation and planning in schools. *Improving Schools*, 23(1):85–102. https://doi.org/10.1177/1365480219895167

Carlson B 2009. School self-evaluation and the 'critical friend' perspective. *Educational Research and Review*, 4(3):78–85. Available at https://academicjournals.org/journal/ERR/article-full-text-pdf/D66A0D83679. Accessed 28 May 2025.

Chapman C & Sammons P 2013. *School self-evaluation for school improvement: What works and why?* Berkshire, England: CfBT Education Trust. Available at https://files.eric.ed.gov/fulltext/ED546801.pdf. Accessed 17 July 2023.

Clift PS, Nuttall DL & McCormick R (eds.) 1987. *Studies in school self-evaluation*. London, England: Falmer Press.

Cobbinah JE & Eshun ES 2021. School self-evaluation and the nature of support system for basic schools in Ghana. *SAGE Open*, 11(2):1–13. https://doi.org/10.1177/21582440211022732

Cohen L, Manion L & Morrison K 2011. *Research methods in education* (7th ed). New York, NY: Routledge.

Dangerfield AS 2012. School self-evaluation in action: A case study in the North-West of England. PhD thesis. Huddersfield, England: University of Huddersfield. Available at https://eprints.hud.ac.uk/id/eprint/14043/1/Anthony_Dangerfield_Final_Thesis.pdf. Accessed 28 May 2025.

Department of Basic Education 2000. *Quality education for all: Statement of public service commitment*. Pretoria, South Africa: Author.

Department of Basic Education, Republic of South Africa 2016. *Revised five-year strategic plan 2015/16-2019/20*. Pretoria: Author. Available at https://www.education.gov.za/Portals/0/Documents/Reports/Revised%20DBE%20Strategic%20Plan%20March%202016.pdf. Accessed 28 May 2025.

Department of Education, Republic of South Africa 2001. The national policy on whole-school evaluation. *Government Gazette*, 433(22512):1–24, July 26. Available at https://www.gov.za/sites/default/files/gcis_document/201409/225120.pdf. Accessed 17 July 2023.

Department of Education and Science 2006. *A guide to whole-school evaluation in primary schools*. Dublin, Ireland: Department of Education and Skills Ireland.

Fushimi A 2014. School self-evaluation for quality improvement: Investigating the practice of the policy in Kenya. EdD thesis. Falmer, England: University of Sussex.

Govender N, Grobler B & Mestry R 2016. Internal whole-school evaluation in South Africa: The influence of holistic staff capacity. *Educational Management Administration & Leadership*, 44(6):996–1020. https://doi.org/10.1177/1741143215595414

Hossain M 2017. *School inspection challenges: Evidence from six countries*. Paris, France: United Nations Educational Scientific and Cultural Organization. Available at https://lib.ncdd.gov.kh/storage/app/public/test_library/CAT_22367_1/2017-SchoolInspection-en.pdf. Accessed 28 May 2025.

Ivancevich J & Matteson M 2002. *Organizational behavior and management* (6th ed). New York, NY: McGraw-Hill.

Jansen JD 2004. Autonomy and accountability in the regulation of the teaching profession: A South African case study. *Research Papers in Education*, 19(1):51–66. https://doi.org/10.1080/0267152032000176972

Katsuno M & Takei T 2008. School evaluation at Japanese schools: Policy intentions and practical appropriation. *London Review of Education*, 6(2):171–181. https://doi.org/10.1080/14748460802185219

Kobola MW 2020. Managing for improved school effectiveness at selected primary schools in the Gauteng province. PhD thesis. Pretoria, South Africa: University of South Africa.

Kumalo V & Skosana D 2014. *A history of the South African Democratic Teachers' Union (SADTU)*.

MacBeath J 2005. *Background, principles and key learning in self-evaluation: A guide for school leaders*. Nottingham, England: National College for School Leadership.

MacBeath J, Schratz M, Meuret D & Jakobsen L 2003. Self-evaluation in European schools: Case examples. In M Preedy, R Glatter & C Wise (eds). *Strategic leadership and educational improvement*. London, England: Paul Chapman Publishing.

Madikida PPP 2016. Challenges in the implementation of whole school evaluation at secondary schools in the Libode district, Eastern Cape Province. PhD thesis. Pretoria, South Africa: University of South Africa. Available at https://uir.unisa.ac.za/server/api/core/bitstreams/873ab449-5c77-4622-9030-3bad2d20dab8/content. Accessed 26 May 2025.

Martinelli DP 2001. Systems hierarchies and management. *Systems Research and Behavioral Science*, 18(1):69–81. https://doi.org/10.1002/sres.390

Mathaba RS 2014. External Whole School Evaluation of underperforming secondary schools in Mpumalanga province. DEd thesis. Durban, South Africa: Durban University of Technology. Available at https://web.archive.org/web/20200318162808id_/https://openscholar.dut.ac.za/bitstream/10321/1073/1/MATHABA_2014.pdf. Accessed 26 May 2025.

Mbalati T 2010. A critique on the implementation of whole school evaluation policy in Limpopo Province. PhD thesis. Sovenga, South Africa: University of Limpopo. Available at http://196.21.218.27/bitstream/handle/10386/494/TIMOTHY%20MBALATI.pdf?sequence=1&isAllowed=y. Accessed 26 May 2025.

Mncube V 2009. The perceptions of parents of their role in the democratic governance of schools in South Africa: Are they on board? *South African Journal of Education*, 29(1):83–103. https://doi.org/10.15700/saje.v29n1a231

Nwosu LI & Chukwuere JE 2017. The roles and challenges confronting the school governing body in representing schools in the digital age. *Journal of Economics and Economic Education Research*, 18(2):1–24.

O'Brien S, McNamara G & O'Hara J 2014. Critical facilitators: External supports for self-evaluation and improvement in schools. *Studies in Educational Evaluation*, 43:169–177. https://doi.org/10.1016/j.stueduc.2014.08.002

O'Brien S, McNamara G, O'Hara J & Brown M 2019. Irish teachers, starting on a journey of data use for school self-evaluation. *Studies in Educational Evaluation*, 60:1–13. https://doi.org/10.1016/j.stueduc.2018.11.001

Organisation for Economic Co-operation and Development 2013. *Synergies for better learning: An international perspective on evaluation and assessment*. Paris, France: OECD Publishing.

https://doi.org/10.1787/9789264190658-en
Phukubje L 2019. Leading and managing the school self-evaluation process in rural schools. MEd dissertation. Johannesburg, South Africa: University of Johannesburg. Available at https://www.proquest.com/docview/2505376369?pq-origsite=gscholar&fromopenview=true&sourcetype=Dissertations%20&%20Theses. Accessed 22 May 2025.
Rapoo T (ed.) 2016. *Talking the talk, not walking the walk: Assessing gender mainstreaming in South Africa's mining sector*. Braamfontein, South Africa: Commission for Gender Equality. Available at http://cge.org.za/wp-content/uploads/2021/01/talking-the-talk-not-walking-the-walk.pdf. Accessed 17 July 2023.
Reddy NM 2005. The assessment of educator competence: Implications for whole school evaluation. MEd dissertation. Johannesburg, South Africa: University of Johannesburg. Available at https://ujcontent.uj.ac.za/esploro/outputs/graduate/The-assessment-of-core-educator-competence/9913752907691. Accessed 22 May 2025.
Ryan J (ed.) 2011. *Education reform in China: Changing concepts, contexts and practices*. New York, NY: Routledge.
Scheerens J 2000. *Improving school effectiveness*. Paris, France: United Nations Educational, Scientific and Cultural Organisation. Available at https://ris.utwente.nl/ws/portalfiles/portal/5154343/Improving-122424e.pdf. Accessed 17 July 2023.
Scheerens J 2002. School self-evaluation: Origins, definition, approaches, methods and implementation. In D Nevo (ed). *School-based evaluation: An international perspective* (Vol. 8). Bingley, England: Emerald Group Publishing.
Schoonenboom J & Johnson RB 2017. How to construct a mixed methods research design. *Kölner Zeitschrift fur Soziologie und Sozialpsychologie*, 69(Suppl 2):107–131. https://doi.org/10.1007/s11577-017-0454-1
Sehlapelo TW 2021. Teachers' perceptions of school self-evaluation on quality teaching and learning in Mankweng, Limpopo Province. PhD thesis. Pretoria, South Africa: University of South Africa. Available at https://uir.unisa.ac.za/server/api/core/bitstreams/51b4dc83-11d4-4d00-b689-7bc96df13788/content. Accessed 21 May 2025.
Senge PM 1990. *The fifth discipline: The art and practice of the learning organization*. New York, NY: Doubleday.
Setlalentoa WN 2011. Barriers to the implementation of Whole School Evaluation at selected primary schools in the southern Free State. *Interim: Interdisciplinary Journal*, 10(1):74–85. Available at https://journals.co.za/doi/epdf/10.10520/EJC120880. Accessed 21 May 2025.
Setlalentoa WN 2014. School self-evaluation: How involved are educators in the process? *International Journal of Educational Sciences*, 7(3):525–533. Available at http://ir.cut.ac.za/bitstream/handle/11462/1538/Setlalentoa%2cW.N%3bPages%20p.525-533.pdf?sequence=1&isAllowed=y. Accessed 20 May 2025.
Siafwa SN, Manchishi PC & Cheyeka AM 2019. Involvement of stakeholders in school-based self-evaluation in assuring quality management. *International Journal of Humanities Social Sciences and Education (IJHSSE)*, 6(10):1–11. https://doi.org/10.20431/2349-0381.0610001
Steyn GM 2002. The changing principalship in South African schools. *Educare*, 31(1&2):251–274.
Swaffield S & MacBeath J 2005. School self-evaluation and the role of a critical friend. *Cambridge Journal of Education*, 35(2):239–252. https://doi.org/10.1080/03057640500147037
Tóth E & Csapó B 2022. Teachers' beliefs about assessment and accountability. *Educational Assessment, Evaluation and Accountability*, 34:459–481. https://doi.org/10.1007/s11092-022-09396-w
Vaganay A 2016. Outcome reporting bias in government-sponsored policy evaluations: A qualitative content analysis of 13 studies. *PLoS ONE*, 11(9):e0163702. https://doi.org/10.1371/journal.pone.0163702
Von Bertalanffy L 1968. *General system theory: Foundations, development, applications*. New York, NY: Braziller.
Wagner T, Kegan R, Lahey L, Lemons RW, Garnier J, Helsing D, Howell A & Rasmussen HT 2006. *Change leadership: A practical guide to transforming our schools*. San Francisco, CA: Jossey-Bass.
Wits Education Policy Unit 2005. *The state of teacher professionalism in South Africa*. Commissioned to the EPU by the South African Council for Educators (SACE).
Yoon SW & Kuchinke KP 2005. Systems theory and technology. Lenses to analyze an organization. *Performance Improvement*, 44(4):15–20. https://doi.org/10.1002/pfi.4140440406