

School self-evaluation: self-perception or self-deception? The impact of motivation and socially desirable responding on self-evaluation results

Jerich Faddar, Jan Vanhoof & Sven De Maeyer

To cite this article: Jerich Faddar, Jan Vanhoof & Sven De Maeyer (2018) School self-evaluation: self-perception or self-deception? The impact of motivation and socially desirable responding on self-evaluation results, *School Effectiveness and School Improvement*, 29:4, 660-678, DOI: [10.1080/09243453.2018.1504802](https://doi.org/10.1080/09243453.2018.1504802)

To link to this article: <https://doi.org/10.1080/09243453.2018.1504802>



© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 20 Aug 2018.



Submit your article to this journal [↗](#)



Article views: 403



View Crossmark data [↗](#)

ARTICLE

 OPEN ACCESS

 Check for updates

School self-evaluation: self-perception or self-deception? The impact of motivation and socially desirable responding on self-evaluation results

Jerich Faddar , Jan Vanhoof and Sven De Maeyer

Department of Training and Education Sciences, Faculty of Social Sciences, University of Antwerp, Antwerp, Belgium

ABSTRACT

In order to enhance the quality of education, school self-evaluation (SSE) has become a key strategy in many educational systems. During an SSE process, schools describe and evaluate their own functioning, often by administering questionnaires among teachers. However, it is unknown to what extent SSE questionnaire results are distorted by respondents' tendencies towards socially desirable responding and their motivation to fill in an SSE questionnaire. This study reports on a path analysis, performed on the results of an authentic SSE with 382 participants. Results indicate that socially desirable responding and motivation have indeed an impact on SSE results. However, the effects are differential and depend on the variable of interest. These findings can have serious implications, and should be taken into account when drawing conclusions and taking (school) policy decisions within the framework of an SSE process.

ARTICLE HISTORY

Received 20 February 2017;
Accepted 19 July 2018

KEYWORDS

School self-evaluation; data quality; socially desirable responding; motivation; path analysis

Problem statement

In many educational systems, school self-evaluation (SSE) has become a key strategy, next to external evaluation, in efforts to ensure the quality of education (Organisation for Economic Co-operation and Development [OECD], 2013). SSE is a mechanism by which the school itself takes the initiative to systematically describe its functioning by stakeholders. Drawing on this description, an evaluation is made of the school, leading to a consideration of policy decisions and to the undertaking of actions (Vanhoof & Van Petegem, 2010). In order to create a description of a school as an organisation, there is the need to measure constructs at the organisational level, an activity which sets a methodological challenge. Because a school as such cannot speak for itself, SSE often relies on collecting data from teachers or other stakeholders. They are, it is argued, involved in the everyday functioning of the school and, therefore, are highly eligible when it comes to providing insightful information with regard to their school. In order to capture this information, several instruments have already been developed. Often, these take the form of a questionnaire that probes for respondents' perceptions regarding

CONTACT Jerich Faddar  jerich.faddar@uantwerpen.be

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

school processes and are commonly found in the literature on school effectiveness (e.g., Hendriks, Doolaard, & Bosker, 2002; MacBeath, Schratz, Meuret, & Jakobsen, 2000).

The use of SSE questionnaires is stimulated for several reasons, although the literature on survey research points to methodological concerns with regard to the use of questionnaires as a data collection method. Several factors can distort the answers of respondents to questionnaire items, such as the mode of administration (e.g., paper-pencil vs. computer-assisted) or the difficulty of items (Belson, 1981; Krosnick & Presser, 2010). Also, the respondents' characteristics may come into play, affecting the quality of the questionnaire results. There are indications that respondents' tendencies towards socially desirable responding (SDR), a phenomenon where individuals give overly favourable self-descriptions, can impact data quality (Lam & Bengo, 2003; Wayne & Liden, 1995). Moreover, individuals who score highly on SDR scales are "faking good". As a result, individuals' (self-)reports in terms of different concepts in the broad field of SSE research and beyond could be considered distorted. Thomas and Kilmann (1975) argue that this mechanism is expected to operate in any study in which ratings are used to assess variables with evaluative overtones. Until now, it is unknown to what extent SSE results and other self-reported measures in the context of SSE are affected by SDR. Next to the issue of SDR, Bateson (1984) points to respondents' willingness to provide an answer as a crucial condition for quality data. When respondents are not motivated to provide an accurate response, they could start relying on response strategies that lead to acceptable, yet inaccurate responses (Krosnick, 1991). In order to be able to make valid conclusions out of SSE questionnaire responses, respondents should be motivated to engage in the cognitive processes that are required to produce an accurate answer to the items. This seems to suggest that only the quantity or amount of motivation matters, but quality is also an important dimension of motivation. The quality of motivation refers to the underlying attitudes and goals that lead to the action (Ryan & Deci, 2000). It is expected that SSE participants vary in the quality of their motivation when it comes to engaging in the SSE process and filling in the SSE questionnaires. However, it is not known to what extent this quality of motivation can explain respondents' answers on SSE questionnaires. At this moment, it is readily assumed that both respondent characteristics (their tendency towards SDR and the quality of motivation to fill in the SSE questionnaire) have no influence on SSE data quality. Nevertheless, valid conclusions on SSE data are of key importance, especially in an era where there is a strong emphasis on data-based or evidence-based decision making (OECD, 2007; Schildkamp, Lai, & Earl, 2013). As gathering data on a school's functioning is part of the SSE procedure, it meets the current tendency towards more attention being paid to data-based decision making in education (Campbell & Levin, 2009; Schildkamp et al., 2013), especially because the collected data are used as a basis for school development plans and policy decisions. In the light of the implications at the school policy level, it is of the utmost importance that the conclusions drawn from SSE data are valid.

This study aims to explore how SDR and the motivation of a respondent to fill in an SSE questionnaire can cause distortions in the different self-reported scores that are obtained in an SSE. As a result, this study will examine how these different variables relate to each other in all their complexity. The manuscript will focus on the extent to which SSE respondents differ in their tendency towards socially desirable responding and their motivation to fill in an SSE questionnaire. Furthermore, the extent to which SSE

questionnaire self-report data are affected by respondents' tendency towards socially desirable responding and motivation is examined.

Conceptual framework

The following sections explore and elaborate the key concepts of this study. It sets off with a clear depiction of what school self-evaluation is, and what can be the subject of an SSE process. An exploration is then made of what is understood by socially desirable responding, and the quantity and quality of respondents' motivation.

School self-evaluation and self-reporting

School self-evaluation (SSE) is a form of internal evaluation counterbalancing a tendency in many educational systems to rely on external evaluations to guarantee educational quality (OECD, 2013). In this study, SSE is defined as "a systematic process, largely initiated by the school itself, where participants describe and evaluate the functioning of the school for the purposes of making decisions or undertaking initiatives in the context of (aspects of) overall school (policy) development" (Vanhoof & Van Petegem, 2010, p. 20).

SSEs require the value judgements of participants regarding specific indicators that relate to the quality schools deliver, or the functioning of a school as referred to in the above definition. These indicators often have their origin in the school effectiveness literature (Scheerens, 1991, 2008; Van Petegem, 1998). Indicators can be situated at different levels and/or stages of the educational process. Generally, the following categories can be discerned: input indicators, process indicators, output indicators, and context indicators (Ikemoto & Marsh, 2007; Scheerens, Glas, & Thomas, 2003). An SSE's focus could, for instance, be narrowed down to hard output indicators such as pupils' results in standardised assessments in reading skills. However, it has already been demonstrated that focussing on school process indicators can lead to a greater impact on the enhancement of school improvement and school effectiveness as these are more easily manipulated (Scheerens, 1991). Typically, this involves concepts such as the quality of instruction, being a learning organisation or being characterised by distributed leadership (Muijs, Harris, Chapman, Stoll, & Russ, 2004; Reynolds, Sammons, Stoll, Barber, & Hillman, 1996; Scheerens, 1991). In this study, there are two central SSE process variables of interest. Although both are process variables, they are situated at a different level within the school. Distributed leadership, as an indicator for policymaking capacity in schools (Van Petegem, Devos, Mahieu, Dang Kim, & Warmoes, 2006), is typically situated at the school level, and can be described as a form of collective leadership where each team member is empowered to share their expertise to lead collectively (Harris, 2004). In this study, however, differentiation taps into the classroom level, and focuses on the extent to which teachers adapt their instruction to the particular situation and needs of their students (Tomlinson et al., 2003). As these process variables cannot be measured directly, since schools cannot speak for themselves as such, this conceals a methodological challenge. Consequently, SSE relies on the perceptions of stakeholders, or other well-chosen participants, who can provide insightful information on the topic under review (MacBeath, 1999). The literature points, next to factors at the instrumental

level, to respondents' characteristics as sources of measurement error. To what extent these characteristics influence the results of SSEs is as yet unknown.

Socially desirable responding

The phenomenon of socially desirable responding is known to influence individuals' behaviour in many different contexts. It has been found that individuals tend to over-report engaging in behaviour that could be described as socially desirable; for example, when teachers are asked to self-report on change in their instructional practices during mathematics classes (Lam & Bengo, 2003). But they may also under-report when they are asked about socially undesirable behaviours such as the extent to which they are confronted with discipline difficulties in their classrooms. In essence, individuals vary in the extent to which they depict themselves as overly positive (Paulhus, 2002). As such, when SDR interferes with providing an accurate response with regard to self-report items, it is considered to be a source of response bias. Holtgraves (2004) identifies three different stages in the cognitive process where SDR can take place. First, SDR can operate during the editing of a response. After having formatted a response, respondents make an evaluation of their response in terms of social desirability. Second, the retrieval stage can simply be bypassed due to SDR. Respondents are basing their answer only on the socially desirable implications of the answer, and are not making any attempt to retrieve relevant or accurate information about the item. A last possibility is that respondents are indeed retrieving information, but in a biased way. They selectively retrieve information which places them in a more favourable light. This retrieval is then aimed at confirming one's inquiry and ignoring contradictory information.

The phenomenon of SDR slipping into the answering process of respondents at different stages also raises the question as to whether different kinds of response behaviour can be discerned. Although SDR has been seen as a unidimensional concept (e.g., Crowne & Marlowe, 1960), research on SDR has already demonstrated that a more nuanced view of this phenomenon is advocated (Helmes & Holden, 2003). For a further operationalisation of the concept of SDR, it is important to acknowledge that response bias can be caused by a response style and/or a response set (Paulhus, 2002). A response style is found to bias individuals' responses over time and, as such, across different questionnaires. A response set, in contrast, occurs only temporarily and is a reaction to a particular question or questionnaire. This distinction is followed in the further conceptualisation of SDR by which several authors have suggested a two-component model (Millham & Kellogg, 1980; Paulhus, 1984). Paulhus (1984, 2002) identifies a first component as *impression management*. It is characterised by an individual, deliberately and consciously, describing her- or himself in an overly positive way in response to a certain question or questions, this being a response set. The second component is *self-deception*, where a respondent unconsciously and honestly reports an overly positive self-image across questionnaires and time, which follows the logic of the response style.

Although there is a debate on the extent to which SDR is a problem when using self-report measures, Thomas and Kilmann (1975) argue that it is likely that SDR occurs in any context in which variables are measured with an evaluative overtone. When SSE is performed either in a developmental or in an accountability-oriented context, an evaluation aspect is involved, and implications on the validity of SSE results are to be expected.

Furthermore, the literature has also demonstrated that the specific traits that are under review can influence the way in which respondents answer them. One could argue that indicators that are closely related to participants' individual responsibilities, such as differentiation, could be perceived as being more sensitive questions, and consequently evoke socially desirable responses compared to indicators that are more remote from the participants' individual responsibilities, such as in the case of distributed leadership (Moorman & Podsakoff, 1992). It could also be argued that both differentiation and distributed leadership are prone to SDR. Respondents who are asked about their own competences, in this case a skilled one such as differentiation, might trigger a socially desirable response because the respondent might want to be seen as a highly competent teacher. The same might be true for distributed leadership, which is a characteristic not commonly found in schools. This might prompt respondents to depict the school favourably. Also, the evaluative aspect of the SSE might contribute to the occurrence of SDR in the reported SSE scores (Thomas & Kilmann, 1975).

Respondents' motivation

A key condition for quality data is the willingness of the respondents to fill in the SSE questionnaire (Bateson, 1984; Krosnick, 1991). If people do not feel an impulse to act, they are described as unmotivated. When people are activated to an end, they are considered to be motivated. Much research has addressed the concept of motivation as a unitary construct, ranging from little motivation to a great deal of motivation. However, self-determination theory (SDT) adds, next to the *quantity* of motivation, another dimension: the *quality* of motivation (Deci & Ryan, 2002; Vansteenkiste, Sierens, Soenens, Luyckx, & Lens, 2009). The quality of motivation refers to the reasons why individuals engage in a particular behaviour, and these can vary substantially. According to SDT, this variation in the quality of motivation is due to the extent to which reasons or motives to engage in behaviour are internalised (Deci & Ryan, 2002). This internalisation is a process by which initial external values as a reason to regulate behaviour are becoming part of the self. Drawing on the dimensions of quantity and quality of motivation which are integrated by SDT, many studies make the distinction between the following types of motivation: *a-motivation* and autonomous and controlled motivation (Deci & Ryan, 2002; Vansteenkiste, Lens, & Deci, 2006).

Considering the quantity of motivation, the first concept is discerned when an individual lacks motivation. *A-motivation* refers to the lack of motivation to engage in filling in a questionnaire or having no intention to do so. A person can have no trust in achieving a desired outcome, may have no feeling of competence when it comes to executing the task, or may perceive the task as irrelevant (Deci & Ryan, 1985).

Autonomous motivation is characterised by a feeling of freedom, and a person's reasons for engaging in filling in a questionnaire can be described as being more or less self-determined. They engage in this activity because of sincere interest, and perceive it as inherently enjoyable and satisfying (Ryan & Deci, 2000). This type of motivation, with its reasons for engaging in filling in a questionnaire, is situated rather at the higher end of the continuum of internalisation. Often, autonomous motivation is further subdivided into intrinsic motivation, with the highest amount of internalisation,

and identification, where individuals engage in certain behaviours as they believe such behaviours help to attain their personal goals.

Individuals who, by contrast, experience a pressure to fill in a questionnaire are driven by *controlled motivation*. A subdivision can be made according to the attribution of this pressure. When the experienced pressure is the result of internal feelings of shame or guilt, then it is referred to as introjected regulation. When pressure is external to the self, as in the case of receiving incentives, avoiding punishments, or meeting the expectations of others, it is referred to as external regulation (Vansteenkiste et al., 2006). Obviously, external regulation is at the lower end of the continuum of internalisation.

Method

Context and participants

In order to examine the aforementioned research questions, the study was embedded in a self-evaluation performed in an educational service organisation in Flanders (Belgium). This organisation provides education and training in several disciplines, ranging from general education over technical education to vocational education. Students can enrol from the age of 16, but the organisation's main student population consists of adults. The SSE dealt with different topics of the organisation's functioning. However, this study will focus on two constructs of interest as typical cases. One variable is a typical organisational construct that is often mentioned in the school effectiveness literature and is widely debated in the field of school improvement: distributed leadership (e.g., Hallinger & Huber, 2012; Muijs & Harris, 2003). It is a way of thinking about leadership that focuses on the engagement of existing expertise scattered within the school, rather than sticking to formal or hierarchical positions and roles (Harris, 2004). The second variable is about teachers' individual practices within their classrooms, and focuses on differentiation. Differentiated classroom instruction can be defined as a systematic, proactive way of providing instruction tailored to the specific needs of students, and taking account of individual differences (Tomlinson et al., 2003).

The teaching staff was familiar with the notion of self-evaluation as they were asked to evaluate their own teaching on a yearly basis. These self-evaluations are performed from a rather developmental perspective in order to use the results as input for further development in terms of the educational quality provided. Previous self-evaluations had been conducted by means of a self-report questionnaire, which means that the teaching staff has experience in this method. All teaching staff were invited to participate in the self-evaluation, which was administered by means of an online questionnaire. A response rate of 58% was achieved, resulting in 382 completed questionnaires.

Instruments

For all concepts in this study, items from existing instruments were used and brought together in the SSE questionnaire. However, the instruments were adopted in a new context which required us to verify their psychometric qualities for this study. To measure socially desirable responding, we relied on the Paulhus Deception Scales and selected items that allowed us to tap into concepts of both impression management

and self-deception (Paulhus, 1998). All 40 items were translated from English into Dutch. As these scales were developed on a body of literature that overarches different contexts and disciplines, an explorative factor analysis (with oblique rotation) was performed to ensure that the two factors were retained in the data. Table 1 includes a sample item of each subscale, and reports on the scales' Cronbach's alpha.

The items of the Academic Self-Regulation Questionnaire (SRQ-A) were adopted to tap into the concepts of intrinsic motivation, identified, introjected, and external regulation (Vansteenkiste et al., 2009). As these original items were situated in the context of learning behaviour, the items were rephrased so that the behaviour of filling in a questionnaire became central in each item. A-motivation is, however, not integrated into the SRQ-A, and therefore we included and rephrased items from the Academic Motivation Scale (AMS) (Vallerand, Blais, Brière, & Pelletier, 1989). As these items were administered in a new context, we started with an explorative factor analysis (with oblique rotation) to identify factors based on the data. With regard to the dimension of the quality of motivation, the results indicated that only two factors could be discerned in the data. Clearly, two components were retained in the data: autonomous motivation, on the one hand, and controlled motivation, on the other. Some items were omitted from further analyses because they did not load well on the factors. Possibly, these items did not fit the context of answering a questionnaire very well as the original instrument was constructed to serve in an academic context (e.g., "I fill in this questionnaire because I find it a pleasant activity"). Cronbach's alphas for these scales were satisfied and showed no problematic inconsistencies (see Table 1).

The dependent variables in this study – distributed leadership and differentiation – were measured by means of a scale specifically designed for the field of education. For distributed leadership, a scale consisting of six items was adopted from an instrument that taps into the policymaking capacities of schools (Vanhoof, Deneire, & Van Petegem, 2011). The extent to which teachers differentiate during their lessons is measured by means of a scale that is borrowed from an existing instrument which aims to examine whether or not teachers demonstrate basic teaching competences (Meynen, Struyf, & Adriaensens, 2011). Cronbach's alphas for these scales were satisfied (see Table 1).

Table 1. Instrument's scales, reliabilities, and sample items.

	N° Items	Cronbach's alpha	Sample items
Socially desirable responding			
Impression management	20	0.79	I sometimes tell lies if I have to.*
Self-deception	20	0.72	I am fully in control of my own fate.
Motivation			
Autonomous motivation	6	0.87	I fill in this questionnaire because I personally find this very valuable.
Controlled motivation	7	0.76	I fill in this questionnaire because others expect me to do so.
A-motivation	4	0.82	Honestly, I don't know. I really feel that I am wasting my time when I'm filling in this questionnaire.
Distributed leadership	7	0.93	On this campus everyone has sufficient possibilities to engage in policymaking procedures.
Differentiation	5	0.74	I take into account the different pace of students.

*negatively phrased item.

Analyses

The conceptual framework identified that SDR can have an impact on any kind of self-report. Consequently, it must be acknowledged that it could also affect the scores obtained for motivation in our study. Therefore, we decided to identify the effect of SDR on motivation, and the SSE variables in a direct and indirect way. This enabled us to control for an SDR effect on the respondents' statements about their motivation when it came to filling in the questionnaire, on the one hand, and to the relationship between motivation and the scores on the SSE variables of interest, on the other. In order to accurately estimate the relationships between all these variables, we ran a path analysis by making use of structural equation modelling (SEM). This technique allowed us to run complex models with latent variables, making use of measures at item level and of multiple indicators for one latent variable. In this sense, the strength of this technique lies in the fact that it combines a measurement model, using confirmatory factor analysis, and a regression model (Kline, 2015; Ullman & Bentler, 2003). Several indices for model fit were consulted to ensure the quality of our analysis. The comparative fit index (CFI), which makes a comparison between the assumed model and a null model without relationships, indicates an acceptable model fit when higher than 0.90 (Hu & Bentler, 1995; Schreiber, Nora, Stage, Barlow, & King, 2006). The root mean squared error of approximation (RMSEA), which gives an indication of how well the model would fit the population, should not be higher than 0.06. In terms of the standardised root mean squared residual (SRMR), which gives an indication of the difference between the predicted and actual matrix, below 0.08 is considered to be acceptable (Hooper, Coughlan, & Mullen, 2008; Hu & Bentler, 1999). Modification indices were consulted in order to optimise the model if the initial model did not fit. Respondents who missed out on an item, or did not respond to one of the variables in the questionnaire, were retained in the analysis by estimating missing data using full information maximum likelihood (FIML). This technique performs well in comparison to other techniques for handling missing data (Enders & Bandalos, 2001).

Given the high number of parameters in our model, especially for the scales tapping into SDR, we parcelled the items of impression management and self-deception into four parcels of five items each. Each parcel served as an indicator for the respective latent constructs (Little, Cunningham, Shahar, & Widaman, 2002; Matsunaga, 2008). The allocation of items to parcels was done randomly and repeated 20 times. For each allocation solution, a satisfying level of fit was obtained for the measurement model of the SDR scale.

The strategy of random parcel allocation for the SDR scale was also adopted in the comprehensive path analysis, leading to 20 SEM models. The fit indices for all estimated models can be found in Table 2. Only slight differences were found across these different models. One of these 20 models was selected and presented in the results section. It is representative of all other models, as all reported significant relationships were found in every estimated model.

Relationships in a path model between independent and dependent variables can be of a direct or indirect nature, because of the presence of one or more mediating variables (Alwin & Hauser, 1975). First, the direct relationships will be discussed in the

Table 2. Average and range of fit indices of 20 path models using parcelling for social desirability scales.

Fit indices	Average	Min	Max
CFI	.912	.904	.920
TLI	.903	.894	.911
RMSEA	.044	.042	.046
SRMR	.073	.070	.076

results section. Path analysis, however, also enabled us to calculate the indirect effects which can be added up to the direct effects, resulting in a total effect parameter. These total effects show the overall effect of the independent variables on the dependent variables, including the effect these have on the mediating variables. The analyses were run using the R-package lavaan (Rosseel, 2012).

Results

In the first part of the results section, some descriptive information is used to discuss the respondents' tendency towards socially desirable responding, and their motivation with regard to filling in the SSE questionnaire. The second part focuses on the explanatory analyses.

Descriptive results

With regard to SDR, the respondents were found to have a rather moderate tendency to describe themselves as overly positively. The mean score for self-deception was 3.50 (see Table 3). The standard deviation ($SD = .58$) is found to be rather small, which means that the difference between the respondents in our sample is also small. The respondents scored slightly higher for their tendency towards impression management ($M = 3.70$), and the results also show a greater spread in responses in terms of this concept among the respondents ($SD = .77$).

The respondents' motives when it comes to filling in the SSE questionnaire were not highly internalised. Although there were some differences between the respondents, they do not see the administration of the SSE questionnaire as an interesting activity, or at least as a valuable activity with regard to achieving their personal goals ($M = 3.42$; $SD = .77$). To a lesser extent, the respondents experienced a feeling of pressure ($M = 2.24$; $SD = .76$) with

Table 3. Range, mean, and standard deviation for scales on socially desirable responding, motivation, distributed leadership, and differentiation.

CONCEPT	MIN.	MAX.	MEAN	SD
Socially desirable responding				
Self-deception	1	5	3.50	.58
Impression management	1	5	3.70	.77
Motivation				
Autonomous motivation	1	5	3.42	.79
Controlled motivation	1	5	2.24	.76
A-motivation	1	5	2.33	.88
Distributed leadership	1	5	3.42	.92
Differentiation	1	5	3.73	.67

regard to the administration of the SSE questionnaire. However, the mean score for a-motivation ($M = 2.33$) was higher than for controlled motivation. This means that the respondents tended to be more unmotivated when it came to filling in the questionnaire than they were controlled motivated. In terms of a-motivation, the spread among the respondents was highest with regard to all the motivation scales ($SD = .88$).

Regarding distributed leadership, the respondents were quite critical about it, with a mean of 3.42. However, of all the administered scales, the respondents differed most in their opinion about distributed leadership in their school ($SD = .92$). The respondents were slightly more positive about the extent to which they differentiate in their classrooms ($M = 3.73$), and they were more unanimous about this judgement with a standard deviation of .67.

Explanatory results

This section reports on the path analysis that was undertaken, with both distributed leadership and differentiation included as SSE variables of interest. The path model presented in Figure 1 has an acceptable fit with the data (CFI = .914; TLI = .905 RMSEA = .043; SRMR = .072).

Direct effects

The model illustrates that SDR has indeed an impact on the SSE variables of interest. Impression management has a significantly positive effect on the respondents' self-reported differentiation in the classroom (estimate = .270; $p < .001$). Consequently, respondents who scored higher on their tendency to deliberately describe themselves in a more favourable way provided a more positive picture of the extent to which they differentiate in their classroom. In contrast, impression management has no statistically significant effect on how respondents report on the extent of distributed leadership in their school (estimate = .127; $p = .062$). Impression management has an opposite effect

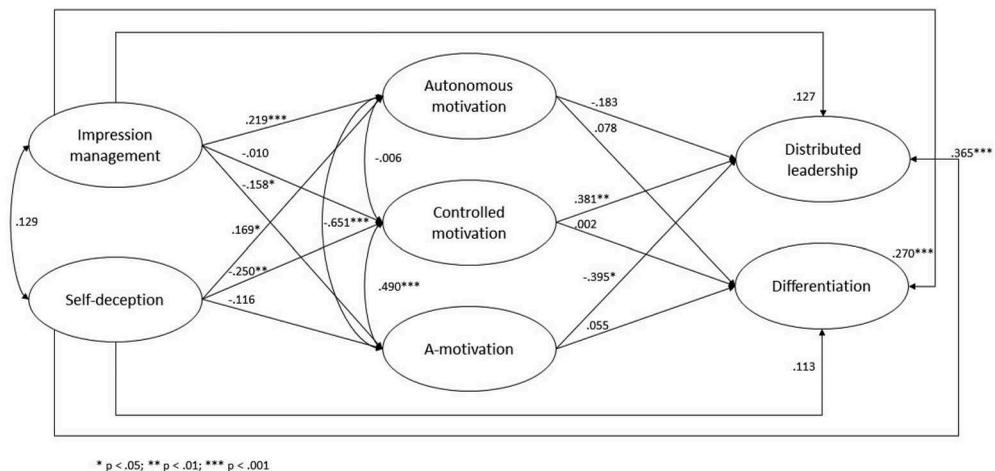


Figure 1. Relationships between impression management, self-deception, autonomous motivation, controlled motivation, a-motivation, and school self-evaluation scores based on the path analysis.

on the variables in comparison to self-deception. Self-deception has a positive significant effect (estimate = .365; $p < .001$) on the respondents' reported distributed leadership in the school. This means that the respondents who have a higher tendency to – unconsciously – give an overly positive self-description were found to have a more positive view of distributed leadership in their school. However, differentiation cannot be explained by self-deception (estimate = .113; $p = .137$).

The results of the path model also show a differential picture for the relationship between the different SSE variables of interest and the subscales of motivation. Autonomous motivation has no statistically significant effect on how respondents report on the extent of distributed leadership in their school (estimate = $-.183$; $p = .145$). This means that the extent to which the respondents indicate that they see the administration of the SSE questionnaire as a personally valuable or interesting task has no effect on how they report on distributed leadership in their school. Nor has this motivation dimension an effect on the reported amount of differentiation in the classroom (estimate = .078; $p = .489$). However, the extent to which the respondents feel an external or internal pressure to fill in the SSE questionnaire does indeed have an effect. There is a statistically significant effect (estimate = .381; $p = .002$) between controlled motivation and the perception of distributed leadership. The parameter is positive, which means that the more the respondents experience an external or internal pressure to fill in the SSE questionnaire, the more positively they report on distributed leadership in their school. With regard to the reported differentiation, no significant effect was found for controlled motivation (estimate = .002; $p = .986$). A-motivation, identifying whether respondents are motivated to fill in the SSE questionnaire anyway, has a negative relationship (estimate = $-.395$; $p = .012$) with distributed leadership. This would suggest that the more respondents are a-motivated, the more negatively they perceive distributed leadership in their school.

Indirect and total effects

The path analysis shows that there are not only direct effects on the SSE variables of interest. It is clear that SDR also has an indirect effect on the ultimate SSE measurements. As the model takes into account that the self-reported motivation scales could also be distorted by SDR, it is of interest to look at the extent to which they are. Autonomous motivation is affected both by impression management (estimate = .219; $p < .001$) and by self-deception (estimate = .169; $p = .011$). The effect is positive, which means that the higher the respondents score on the SDR scales, the more they state that they are motivated by sincere interest, or see the questionnaire as a valuable means of achieving personal goals. Furthermore, self-deception is a predictor for the variance in controlled motivation (estimate = $-.250$; $p = .002$). The found relationship is negative, which means that the higher the respondents' tendency towards self-deception, the lower they score for controlled motivation. Impression management, in contrast, has no significant explanatory effect on controlled motivation (estimate = $-.010$; $p = .883$). With regard to a-motivation, the opposite explanatory effects are found. The variance in a-motivation is not significantly impacted by self-deception (estimate = $-.116$; $p = .093$). Impression management, however, does indeed influence the a-motivation score in a negative way (estimate = $-.158$; $p = .018$). Therefore, those respondents scoring higher on impression management tended to score lower for a-motivation, which in fact has a reducing effect on the extent to which they perceive filling in the SSE questionnaire as being a useless task.

Table 4. Standardized parameters and *p* values for indirect effects and total effects of impression management and self-deception on SSE variables.

	Parameter	<i>p</i>
Indirect effects		
Impression management → Distributed leadership	0.018	0.547
Self-deception → Distributed leadership	-0.081	0.072
Impression management → Differentiation	0.008	0.546
Self-deception → Differentiation	0.006	0.817
Total effects		
Impression management → Distributed leadership	0.145	0.029
Self-deception → Distributed leadership	0.284	0.000
Impression management → Differentiation	0.278	0.000
Self-deception → Differentiation	0.120	0.083

Structural equation modelling enables us to calculate the indirect effects and direct effects with regard to the total effect parameters of the adopted impression management and self-deception scale. The indirect effects, via the motivational scales, seem to have no statistically significant effect on the dependent variables (see Table 4). The total effects include both the direct effects, which were discussed above, and the indirect effects, which are due to the path structure in the analysis. The total effect of impression management on distributed leadership shows a remarkable result. Although there is no significant direct or indirect effect of impression management on distributed leadership, the total effect of impression management on distributed leadership actually turns out to be statistically significant. This means that when both indirect and direct effects are taken into consideration, impression management still has an effect on the obtained result for distributed leadership. The total effect of self-deception on distributed leadership and the total effect of impression management on differentiation could be expected, as the earlier results already point out that there is a strong direct effect between these variables. Still, this indicates that the indirect effects and direct effects do not level each other out. The total effect of self-deception on differentiation is, after the inclusion of direct and indirect effects, not statistically significant.

Conclusion and discussion

The objectives of this study are threefold. First, the need to identify how respondents differ regarding their tendency to respond in a socially desirable way, together with the quality and quantity of their motivation to fill in a school self-evaluation (SSE) questionnaire. Second, this study aims to identify to what extent SSE questionnaire data are affected by socially desirable responding (SDR). Third, the study explores to what extent the quantity and quality of respondents' motivation affects SSE questionnaire data.

The results show that there is indeed variation in the respondents' tendency towards SDR. The respondents scored more highly for impression management in comparison to self-deception. However, the spread among the respondents is also higher for impression management than for self-deception. With regard to motivation, the respondents scored rather low for autonomous motivation. Still, although somewhat on the lower end of the 5-point Likert scale, a-motivation obtained a higher average score than did controlled motivation.

Furthermore, this study shows that SDR has both direct and indirect effects on the SSE variables of interest. However, the picture for both dependent variables adopted in our model differs. Whereas there is a direct effect of self-deception on the respondents' opinions about distributed leadership, there is none for the extent of the respondents' self-reported extent of differentiation in the classroom. In contrast, there is a direct effect in terms of impression management on the extent of the respondents' self-reported extent of differentiation, but not for respondents' opinions about distributed leadership in the school. A significant direct effect of self-deception was found with regard to distributed leadership. The literature describes impression management as a deliberate response behaviour operating for specific questionnaires or questions, in the form of a temporary reaction (Paulhus, 2002). Our study, in the context of SSE, complies with earlier research, suggesting that this mechanism indeed depends on the item's subject or the variable under consideration. The fact that the extent of self-reporting with regard to differentiation is affected by impression management may be explained by how the respondent relates to the reported behaviour. Differentiation in the classroom is situated more in their own control, and teachers may feel more responsible for it, whereas the extent to which their school is characterised by distributed leadership is not solely their own responsibility, nor a description of their own behaviour exclusively. This connects to the literature that deals with the question of what could be understood as sensitive, and consequently vulnerable for SDR. Tourangeau, Rips, and Rasinski (2000) identified that concerns with regard to possible consequences, or the perceived intrusiveness of questions, could trigger the respondents' tendency towards SDR. Tapping into differentiation in the classroom might be viewed as more intrusive in comparison to distributed leadership, or respondents might think that they will be held accountable. The direct effect of self-deception on the respondents' opinions about distributed leadership means that teachers tend to over-report characteristics at the school level in a genuine or unconscious way. Possibly they have an overly positive picture of their school or school-level characteristics, because they may believe that they are doing a good job. They might consider all the well-intended efforts of their colleagues and of management regarding distributed leadership and the schooling they provide in general, and may have a genuinely positive perception of it. Further in-depth research should look into this phenomenon, in order to uncover what is at play in this situation.

The indirect effects of impression management and self-deception via the path structure of the model are not significant. However, combining the direct and indirect effects into total effects indicates that there is also a significant effect of impression management on the reported distributed leadership, although no significant direct and indirect effects were found. This stresses the importance of considering a path-model approach as conducted in this study (Alwin & Hauser, 1975).

The motivation of the respondents to fill in an SSE questionnaire has indeed, even after correction for the impact of SDR, an impact on an SSE. The results demonstrate that this also depends on the variable under consideration. No impact in terms of the quantity or quality of motivation was found on the reported extent of differentiation. With regard to reported distributed leadership, motivation has an impact. In terms of quantity of motivation, this study points out that unmotivated respondents (scoring high for a-motivation) are evaluating distributed leadership in their school less positively. In terms of quality of motivation, this study finds that respondents who experience pressure when it comes to filling in an SSE

questionnaire, are attributing a higher score for distributed leadership. It remains unclear why the effect of motivation is different for differentiation and distributed leadership. Possibly, a-motivated respondents put less effort into thinking about positive examples of distributed leadership in their organisation, leading to a more negative picture of distributed leadership. Respondents who are reporting a higher extent of controlled motivation may feel an internal or external pressure to think of positive examples of distributed leadership, leading to a more positive score. Furthermore, there might also be a connection with the difficulty of the concept that is the subject of the SSE. As distributed leadership is not common within schools, it might be more difficult for the respondents to recall positive examples or indications thereof. That could make the eventual score more liable to be affected by the respondents' motivation to fill in the questionnaires. Making statements about differentiation, which is related more closely to their daily activities, might require less effort on the part of the respondents when it comes to recalling examples or indications. Possibly, this explains why motivation to fill in the questionnaire has no impact on the reported score in terms of differentiation. Nonetheless, further research should look into possible explanations for these findings. This study sketches a more nuanced picture than is generally found in the field of self-report methodology, which commonly states that the respondents' amount of motivation has an effect on the accuracy of their answers (e.g., Cannell, Miller, & Oksenberg, 1981; Kessler, Wittchen, Abelson, & Zhao, 2000). By discerning the quality in addition to the quantity of motivation, and identifying that there is a differential effect on the SSE variable of interest, the current study contributes to theory building in this area of research.

The most important outcome of this study is that it must be acknowledged that data gathered within the process of SSE are not free of distortion. The respondents' self-perceptions or their perceptions of the school are indeed influenced by self-deception, or socially desirable responding in general. Also, motivation has an impact on SSE results. This raises the question of to what extent SSE practitioners can rely on such questionnaire data in order to arrive at sound conclusions, or indeed make valid policy decisions. Moreover, the impact of SDR and motivation is not univocal, and depends on the SSE variable of interest. Possibly, the extent to which the variables are under the control and responsibility of the respondents involved makes it more or less vulnerable to the influence of the respondents' tendency towards impression management or self-deception. The same applies with regard to the respondents' motivation to engage in filling in the SSE questionnaire. These differential findings suggest that turning SSE results into valid interpretations is far from self-evident (Kane, 2013).

This study generates important insights about the conceptualisation of SDR and motivation. The factor analyses (exploratory and confirmatory) conducted in this study support the division of SDR into two sub-concepts. Discerning impression management and self-deception is not only theoretically underpinned but is also supported by the data. Moreover, it can be seen to be a necessary approach, since effects on the SSE variables are found in a differential fashion, both in a direct and in an indirect way. Conceptualising motivation into such sub-concepts as autonomous, controlled, and a-motivation also proved to be important. Although no significant effects are found on the reported extent of differentiation, distributed leadership is affected by controlled motivation and a-motivation. Autonomous motivation has no significant impact on either of the SSE variables. At the level of measurement, this study contributes to the

field by exploring the concepts of SDR and motivation. This study makes a first attempt to translate the Paulhus Deception Scales into Dutch. This instrument makes use of 40 items, which requires a great deal of effort on the part of the respondents. Further research could focus on the psychometric qualities of this instrument and focus on the feasibility of shortening the questionnaire. Concerning motivation, this study was not able to discern a further subdivision of autonomous motivation into intrinsic motivation and identified regulation. Nor was there evidence for subdividing controlled motivation into introjected and external regulation. Possibly, the translation of the instrument into the context of filling in an SSE questionnaire needs further testing and refinement. Furthermore, identifying the interplay between autonomous and controlled motivation, through person-oriented profile analyses, might also provide more insight into motivational profiles of respondents (Vansteenkiste et al., 2009).

This study has important implications for researchers and practitioners who want to use SSE questionnaire data in order to inform their decisions, actions, and policies. It is vital to avoid distortion in SSE questionnaires as much as possible. For SSE practice, it could be advised to motivate participants in an autonomous way, meaning that they engage in filling in the SSE questionnaire out of sincere interest, or that they at least identify it as a means of achieving their personal goals. By doing so, the risk of distorted SSE results is reduced. Autonomous motivation can be stimulated by fostering feelings of autonomy among the respondents. This can be enhanced by letting them decide on the focus of the SSE and developing an interest in the SSE by rousing their curiosity in the matter (Hidi & Renninger, 2006). By stimulating the respondents' involvement during the whole evaluation process, it might be possible to enhance the respondents' motivation to cooperate by filling in the SSE questionnaire (Cousins & Earl, 1995; Earl & Katz, 2006). Also, our findings suggest that calling for respondents to be honest in their answering is not sufficient when it comes to obtaining higher quality data. As self-deception occurs unconsciously, an awareness of their behaviour should be raised among SSE respondents. Triggering and stimulating their critical thinking is an important aspect in dealing with SSE. It is a central feature of self-evaluation capacity building (programmes) (Labin, 2014). This could be stimulated by creating a safe climate among staff, characterised by an openness to constructive critique and feedback (Vanhoof, Van Petegem, Verhoeven, & Buvens, 2009). Respondents can be asked to be hard on themselves when giving their opinion. Furthermore, practitioners are advised to supplement SSE data gathered from questionnaires with other data and data sources (MacBeath & McGlynn, 2002). Individual interview data or information obtained from focus group interviews could provide a deeper and/or broader insight into what could be derived from SSE questionnaires.

An SSE can be performed in varying contexts. When an SSE is performed in a context which is strongly characterised by accountability, the respondents might behave in a very different way in comparison to a rather development-oriented context. This study took place in a setting where teachers were familiar with the administration of this SSE within the framework of their personal development. An interesting extension to this study would be to examine how the respondents behave in contexts that have a different focus (accountability versus development) and what effects this might have on the SSE results.

In conclusion, it is noteworthy that this study is unique within the field of school self-evaluation. It makes a first attempt to identify how socially desirable responding, and the

quality and quantity of respondents' motivation to fill in a questionnaire, affects the quality of data. This study fits into a trend towards paying more attention to the quality of data that are gathered in the process of school self-evaluation, or where data are gathered to describe schools' own performance or functioning within the framework of quality assurance in general.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Jerich Faddar is a PhD candidate at the Department of Training and Education Sciences (Faculty of Social Sciences) of the University of Antwerp and a member of the Edubron research unit. His current work focuses on methodological issues of school self-evaluations.

Jan Vanhoof is associate professor on the staff of the Department of Training and Education Sciences of the University of Antwerp. He is a member of the Edubron research unit. His current research activities focus on school policy and quality care in general and on school self-evaluation and data-driven school policy in particular.

Sven De Maeyer is professor at the Department of Training and Education Sciences of the University of Antwerp. He is a member of the Edubron research unit. His research focuses on educational measurement and methodological issues in educational sciences.

ORCID

Jerich Faddar  <http://orcid.org/0000-0001-5465-8615>

References

- Alwin, D. F., & Hauser, R. M. (1975). The decomposition of effects in path analysis. *American Sociological Review*, 40(1), 37–47. doi:10.2307/2094445
- Bateson, N. (1984). *Data construction in social surveys*. London: Allen & Unwin.
- Belson, W. A. (1981). *The design and understanding of survey questions*. Hampshire: Gower Aldershot.
- Campbell, C., & Levin, B. (2009). Using data to support educational improvement. *Educational Assessment, Evaluation and Accountability (Formerly: Journal of Personnel Evaluation in Education)*, 21(1), 47–65. doi:10.1007/s11092-008-9063-x
- Cannell, C. F., Miller, P. V., & Oksenberg, L. (1981). Research on interviewing techniques. *Sociological Methodology*, 12, 389–437. doi:10.2307/270748
- Cousins, J. B., & Earl, L. M. (Eds.). (1995). *Participatory evaluation in education: Studies in evaluation use and organizational learning*. Washington, DC: Falmer Press.
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24(4), 349–354. doi:10.1037/h0047358
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York, NY: Plenum.
- Deci, E. L., & Ryan, R. M. (Eds.). (2002). *Handbook of self-determination research*. Rochester, NY: The University of Rochester Press.

- Earl, L. M., & Katz, S. (2006). *Leading schools in a data-rich world: Harnessing data for school improvement*. Thousand Oaks, CA: Corwin Press.
- Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 8(3), 430–457. doi:10.1207/S15328007SEM0803_5
- Hallinger, P., & Huber, S. (2012). School leadership that makes a difference: International perspectives. *School Effectiveness and School Improvement*, 23(4), 359–367. doi:10.1080/09243453.2012.681508
- Harris, A. (2004). Distributed leadership and school improvement: Leading or misleading? *Educational Management Administration & Leadership*, 32(1), 11–24. doi:10.1177/1741143204039297
- Helmes, E., & Holden, R. R. (2003). The construct of social desirability: One or two dimensions? *Personality and Individual Differences*, 34(6), 1015–1023. doi:10.1016/S0191-8869(02)00086-7
- Hendriks, M., Doolaard, S., & Bosker, R. J. (2002). Using school effectiveness as a knowledge base for self-evaluation in Dutch schools: The ZEBO-project. In A. J. Visscher & R. Coe (Eds.), *School improvement through performance feedback* (pp. 114–142). Lisse: Swets & Zeitlinger.
- Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, 41(2), 111–127. doi:10.1207/s15326985ep4102_4
- Holtgraves, T. (2004). Social desirability and self-reports: Testing models of socially desirable responding. *Personality and Social Psychology Bulletin*, 30(2), 161–172. doi:10.1177/0146167203259930
- Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, 6(1), 53–60.
- Hu, L.-t., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modelling: Concepts, issues, and applications* (pp. 76–99). Thousand Oaks, CA: Sage.
- Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. doi:10.1080/10705519909540118
- Ikemoto, G. S., & Marsh, J. A. (2007). Cutting through the data-driven mantra: Different conceptions of data-driven decision making. In P. A. Moss (Ed.), *Evidence and decision making* (pp. 105–131). Malden, MA: Wiley-Blackwell.
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1–73. doi:10.1111/jedm.12000
- Kessler, R. C., Wittchen, H.-U., Abelson, J., & Zhao, S. (2000). Methodological issues in assessing psychiatric disorders with self-reports. In A. A. Stone, J. S. Turkkan, C. A. Bachrach, J. B. Jobe, H. S. Kurtzman, & V. S. Cain (Eds.), *The science of self-report: Implications for research and practice* (pp. 229–255). Mahwah, NJ: Erlbaum.
- Kline, R. B. (2015). *Principles and practice of structural equation modeling* (4th ed.). New York, NY: The Guilford Press.
- Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, 5(3), 213–236. doi:10.1002/acp.2350050305
- Krosnick, J. A., & Presser, S. (2010). Question and questionnaire design. In P. V. Marsden & J. D. Wright (Eds.), *Handbook of survey research* (2nd ed., pp. 263–314). Bingley, UK: Emerald Group.
- Labin, S. N. (2014). Developing common measures in evaluation capacity building: An iterative science and practice process. *American Journal of Evaluation*, 35(1), 107–115. doi:10.1177/1098214013499965
- Lam, T. C. M., & Bengo, P. (2003). A comparison of three retrospective self-reporting methods of measuring change in instructional practice. *American Journal of Evaluation*, 24(1), 65–80. doi:10.1177/109821400302400106
- Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 151–173. doi:10.1207/S15328007SEM0902_1
- MacBeath, J. (1999). *Schools must speak for themselves: The case for school self-evaluation*. London: Routledge.
- MacBeath, J., & McGlynn, A. (2002). *Self-evaluation: What's in it for schools?* London: Routledge.
- MacBeath, J., Schratz, M., Meuret, D., & Jakobsen, L. (2000). *Self-evaluation in European schools: A story of change*. London: RoutledgeFalmer.

- Matsunaga, M. (2008). Item parceling in structural equation modeling: A primer. *Communication Methods and Measures*, 2(4), 260–293. doi:10.1080/19312450802458935
- Meynen, K., Struyf, E., & Adriaensens, S. (2011). Is the beginning teacher ready for the job? The validation of an instrument to measure the basic skills of the beginning teacher in secondary education. *Pedagogische Studien*, 88(4), 266–282.
- Millham, J., & Kellogg, R. W. (1980). Need for social approval: Impression management or self-deception? *Journal of Research in Personality*, 14(4), 445–457. doi:10.1016/0092-6566(80)90003-3
- Moorman, R. H., & Podsakoff, P. M. (1992). A meta-analytic review and empirical test of the potential confounding effects of social desirability response sets in organizational behaviour research. *Journal of Occupational and Organizational Psychology*, 65(2), 131–149. doi:10.1111/j.2044-8325.1992.tb00490.x
- Muijs, D., & Harris, A. (2003). Teacher leadership – Improvement through empowerment? An overview of the literature. *Educational Management Administration & Leadership*, 31(4), 437–448. doi:10.1177/0263211X030314007
- Muijs, D., Harris, A., Chapman, C., Stoll, L., & Russ, J. (2004). Improving schools in socioeconomically disadvantaged areas – A review of research evidence. *School Effectiveness and School Improvement*, 15(2), 149–175. doi:10.1076/sesi.15.2.149.30433
- Organisation for Economic Co-operation and Development. (2007). *Evidence in education: Linking research and policy*. Paris: Author.
- Organisation for Economic Co-operation and Development. (2013). *Synergies for better learning: An international perspective on evaluation and assessment*. Paris: Author.
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46(3), 598–609. doi:10.1037/0022-3514.46.3.598
- Paulhus, D. L. (1998). *Paulhus deception scales (PDS): The balanced inventory of desirable responding-7*. Toronto: Multi-Health Systems.
- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. I. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 49–69). Mahwah, NJ: Erlbaum.
- Reynolds, D., Sammons, P., Stoll, L., Barber, M., & Hillman, J. (1996). School effectiveness and school improvement in the United Kingdom. *School Effectiveness and School Improvement*, 7(2), 133–158. doi:10.1080/0924345960070203
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36. doi:10.18637/jss.v048.i02
- Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25(1), 54–67. doi:10.1006/ceps.1999.1020
- Scheerens, J. (1991). Process indicators of school functioning: A selection based on the research literature on school effectiveness. *Studies in Educational Evaluation*, 17(2–3), 371–403. doi:10.1016/S0191-491X(05)80091-4
- Scheerens, J. (2008). *Review and meta-analysis of school and teaching effectiveness*. Berlin: Bundesministerium für Bildung und Forschung (BMBF).
- Scheerens, J., Glas, C. A. W., & Thomas, S. M. (2003). *Educational evaluation, assessment, and monitoring. A systemic approach*. London: Taylor & Francis.
- Schildkamp, K., Lai, M. K., & Earl, L. (Eds.). (2013). *Data-based decision making in education: Challenges and opportunities*. Dordrecht: Springer.
- Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research*, 99(6), 323–338. doi:10.3200/JOER.99.6.323-338
- Thomas, K. W., & Kilmann, R. H. (1975). The social desirability variable in organizational research: An alternative explanation for reported findings. *The Academy of Management Journal*, 18(4), 741–752. doi:10.2307/255376
- Tomlinson, C. A., Brighton, C., Hertberg, H., Callahan, C. M., Moon, T. R., Brimijoin, K., ... Reynolds, T. (2003). Differentiating instruction in response to student readiness, interest, and learning profile in academically diverse classrooms: A review of literature. *Journal for the Education of the Gifted*, 27(2–3), 119–145. doi:10.1177/016235320302700203

- Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge, UK: Cambridge University Press.
- Ullman, J. B., & Bentler, P. M. (2003). Structural equation modeling. In I. B. Weiner, J. A. Schinka, & W. F. Velicer (Eds.), *Handbook of psychology: Research methods in psychology* (Vol. 2, pp. 607–634). Hoboken, NJ: John Wiley & Sons.
- Vallerand, R. J., Blais, M. R., Brière, N. M., & Pelletier, L. G. (1989). Construction et validation de l'échelle de motivation en éducation (EME) [Construction and validation of the motivation scale in education]. *Canadian Journal of Behavioural Science*, 21(3), 323–349. doi:10.1037/h0079855
- Vanhoof, J., Deneire, A., & Van Petegem, P. (2011). *Waar zit beleidsvoerend vermogen in (ver) scholen? Aanknopingspunten voor zelfevaluatie en ontwikkeling* [Where is policymaking capacity hidden in schools? Cruxes for self-evaluation and development]. Mechelen: Plantyn.
- Vanhoof, J., & Van Petegem, P. (2010). Evaluating the quality of self-evaluations: The (mis)match between internal and external meta-evaluation. *Studies in Educational Evaluation*, 36(1–2), 20–26. doi:10.1016/j.stueduc.2010.10.001
- Vanhoof, J., Van Petegem, P., Verhoeven, J. C., & Buvens, I. (2009). Linking the policymaking capacities of schools and the quality of school self-evaluations: The view of school leaders. *Educational Management Administration & Leadership*, 37(5), 667–686. doi:10.1177/1741143209339653
- Van Petegem, P. (1998). *Vormgeven aan schoolbeleid: Effectieve-scholenonderzoek als inspiratiebron voor de zelfevaluatie van scholen* [Shaping school policy: School effectiveness research as a source of inspiration for school self-evaluation]. Leuven: Acco.
- Van Petegem, P., Devos, G., Mahieu, P., Dang Kim, T., & Warmoes, V. (2006). *Hoe sterk is mijn school? Het beleidsvoerend vermogen van Vlaamse scholen* [How strong is my school? The policy-conducting capacity of Flemish schools]. Mechelen: Wolters Plantyn.
- Vansteenkiste, M., Lens, W., & Deci, E. L. (2006). Intrinsic versus extrinsic goal contents in self-determination theory: Another look at the quality of academic motivation. *Educational Psychologist*, 41(1), 19–31. doi:10.1207/s15326985ep4101_4
- Vansteenkiste, M., Sierens, E., Soenens, B., Luyckx, K., & Lens, W. (2009). Motivational profiles from a self-determination perspective: The quality of motivation matters. *Journal of Educational Psychology*, 101(3), 671–688. doi:10.1037/a0015083
- Wayne, S. J., & Liden, R. C. (1995). Effects of impression management on performance ratings: A longitudinal study. *The Academy of Management Journal*, 38(1), 232–260. doi:10.2307/256734