The interplay of user beliefs and situated characteristics in explaining school performance feedback use

The Interplay of User Beliefs and Situated Characteristics in Explaining School Performance Feedback Use

The present study explores predictors of school performance feedback (SPF) use. In total, 470 Flemish educational professionals were surveyed about their use of SPF from school-external, low-stakes standardized assessments. A path analysis was conducted in order to investigate how individual user beliefs impact SPF use on school level and how those beliefs mediate the effects of school-level features pertaining to school organization, performance and voluntariness. Findings include that users' cognitive attitude and perceived expectations of others have a small effect on engagement with SPF in schools, and that these predictors mediate the effects of certain organizational characteristics. Whereas performance levels do not impact school-level feedback use, voluntariness in feedback pursuit and particularly an SPF-oriented school culture emerge as drivers. Implications for practice include the need for stimulating ownership in data-based decision making. Suggestions for further research are also discussed.

Keywords: school performance feedback; school improvement; data-based decision making; data culture; theory of planned behavior

1. Introduction

The past decades have been marked by an increasing awareness of the importance of data use in education. Analogous to evidence-based approaches in medicine (Schildkamp, 2019), researchers find that educational professionals endeavoring to improve student achievement need to fully exploit all information sources available to them in order to shape their policy and practice. However, the literature on data-driven decision making (DDDM), or data-based decision making (DBDM), has also established that data themselves do not necessarily drive (Dowd, 2005; Lockton et al., 2020). In order to foster informed school improvement, for instance through interventions, it is not sufficient to make high quality data available (Hulpia & Valcke, 2004; Schildkamp & Kuiper, 2010). It is also crucial to be conscious of the factors that trigger, accommodate or inhibit efficient data use in schools.

Research has identified a wide range of such influencing factors. For one, data use requires human capacity (Mandinach & Gummer, 2013; Mandinach & Schildkamp, 2020b). A fundamental prerequisite to DDDM is that educators are sufficiently data literate. Data literate educators possess the knowledge, skills and dispositions that enable them to transform information into actionable knowledge (Mandinach & Gummer, 2016). Instead of solely relying on intuition, they confidently and critically approach a wide range of information sources, interpret and contextualize this information, and use it to shape their policy and practice in a responsible and appropriate manner (Mandinach & Gummer, 2016; Vanlommel et al., 2017). They are also willing and able to engage in collective sensemaking, as collaboration and co-construction are key in effective DDDM (Mandinach et al., 2011; Mandinach & Gummer, 2016).

Since data use processes always take place within a certain setting and structure, situational characteristics invariably influence how individual data users engage with data (Abrams et al., 2020). For instance, like other organizational processes, the data use process in schools is influenced by the school context (e.g. staff, expertise, professional capacity and resources) and the school organization (e.g. leadership, innovation climate, collaboration) (Abrams et al., 2020; Bryk, 2010; Jimerson et al., 2020; Visscher, 2020).

Moreover, the educational context determines to a great extent how data use processes take shape (Coburn & Turner, 2011; Mandinach & Schildkamp, 2020a). An educational system tends to be characterized by its inclination towards accountability or improvement, two purposes of data use between which there is a duality and often a tension (Datnow & Park, 2018; Schildkamp et al., 2014; Visscher & Coe, 2003). Several authors take a passionate stance against accountability-driven systems, which they propose corrupt the processes they intend to monitor (Nichols & Berliner, 2007), weaken schools because of the pressure they puts on educators (Nichols & Harris, 2016), and do not (or at least not conclusively) enhance

student achievement (Nichols et al., 2006, 2012; Nichols & Berliner, 2007). Such systems often rely on high-stakes testing: standardized forms of assessment that serve school accountability or student accountability goals (or both), and typically only cover a limited range of topics, which in turn also raises equity concerns (Datnow & Park, 2018). "High stakes testing" and "standardized testing" are frequently used as synonyms, also in many research accounts on DDDM. The issue lies, however, not entirely with the standardized nature of these tests and assessments, but rather in the stakes, or consequences attached to the outcomes (Nichols & Harris, 2016). Low-stakes external standardized assessments can provide valuable information for school improvement and can meaningfully contribute to a picture of student achievement, when they are part of a balanced system of testing, assessment and process monitoring (Nichols & Berliner, 2007) and when aimed at identifying areas for support and improvement (Datnow & Park, 2018; Nichols & Harris, 2016). In several educational contexts, insights like these are bringing about a gradual shift in focus from data use for accountability to data use for (continuous) improvement and organizational development (Mandinach, 2012; Mandinach & Schildkamp, 2020a).

Nevertheless, data used in DDDM should not be limited to assessment data and test scores (Mandinach & Schildkamp, 2020a). And above all, in data use, alignment of the data with the goals is paramount (Mandinach, 2012; Mandinach & Gummer, 2016). Using data inappropriately or for unintended purposes, raises issues of validity and may lead to poor, unfit or undesirable decisions (Mandinach, 2012; Mandinach & Gummer, 2016; Visscher & Coe, 2003).

Focus of this study

A growing number of descriptive studies is providing in-depth insight into the mechanics of how the aforementioned factors influence data use. However, the field is in need of more explanatory research (Van Gasse et al., 2017) to shed light on the relative impact of

influencing factors (Schildkamp et al., 2017) and their interplay (Coburn & Turner, 2011, 2012). In the present study, we address this knowledge gap by investigating the use of school performance feedback (SPF) from school-external, low-stakes standardized assessments as a case. SPF is conceptualized as data about a school's functioning or performance, provided confidentially to the school by an external agent for self-evaluation, intended to inform the school's decision making process (Visscher & Coe, 2003). This definition entails a clear school development orientation.

We investigate factors that enable or hinder SPF use in schools by adopting a quantitative approach, and by taking on a dual perspective. We include user-level predictors in order to acknowledge that SPF use, like other forms of data use, takes shape in the hands of individual actors (Coburn & Turner, 2012; Datnow & Hubbard, 2016; Prenger & Schildkamp, 2018; Schildkamp et al., 2014). By also including school-level predictors, we account for the fact that these data users do not operate in isolation (Abrams et al., 2020; Coburn & Talbert, 2006; Coburn & Turner, 2011, 2012; Schildkamp, 2019; Schildkamp et al., 2014).

In order to select user-level predictors of SPF use and in order to hypothesize how these interact with school-level predictors, we take inspiration from the Theory of Planned Behavior (TPB, Ajzen, 1991). The TPB states that the intention to perform a certain behavior is shaped by the strength and favorability of the agent's behavioral, normative and control beliefs (Ajzen, 1991). Operationally, these beliefs form three distinct constructs: attitude, subjective norm and perceived control. Consequently, we hypothesize that educational professionals' attitude, subjective norm and perceived control regarding SPF influence engagement with SPF reports in schools. The present study focuses on users' self-efficacy when investigating perceived control.

The TPB acknowledges that the relative impact of its central predictors varies across different settings (Ajzen, 1991, 2002, 2011; Armitage & Conner, 2001), which justifies

situating SPF use and individual user beliefs within in a specific school context. In the present study, we hypothesize that organizational, performance-related and contextual school-level features affect school-level SPF use (cf. Verhaeghe et al., 2010; Visscher & Coe, 2003) because they influence the beliefs of individual SPF users. More specifically, we hypothesize that users' beliefs about SPF will be more favorable or salient when they perceive their school culture to be accommodating of SPF use. Due to the nature of SPF, we also presume users in a coordinating role regard SPF more favorably than those who teach. Furthermore, we hypothesize that a higher performance level of the school has a positive impact on user beliefs regarding the SPF, and consequently, will positively relate to school-level SPF use. Finally, we propose that SPF that was actively and voluntarily requested, kindles more positive user perceptions and boosts engagement to a larger extent.

These hypotheses give rise to the following research questions:

(1) (RQ1) To what extent do users' attitudes, subjective norm and self-efficacy impact SPF use on school level?
(2) (RQ2a) To what extent do SPF-oriented school culture, users' work role, feedback sign and voluntariness in feedback pursuit impact SPF use on school level, and
(3) (RQ2b) Are those effects mediated by users' attitudes, subjective norm and self-efficacy?

2. Theoretical Framework

In the following paragraphs, we will briefly present theoretical insights pertaining to each of the predictors included in our conceptual model, as pictured in Figure 1, and discuss their operationalization in this study.

[Figure 1 near here]

2.1 School-level SPF use

Since implementation is key in effective SPF use (Hellrung & Hartig, 2013; Visscher & Coe, 2003), we investigate school-level SPF use in a tangible, concrete manner. We focus on the policy-making cycle that corresponds to effective use of SPF data for school improvement (Verhaeghe et al., 2010) and on the systematic process of transforming data into actionable knowledge which is central to different data use theories of action (Mandinach, 2012; Marsh, 2012). First of all, SPF reports need to be discussed and analyzed within the school team in order to turn raw data into information. Next, this information needs to be triangulated with prior knowledge and other sources to turn it into knowledge. Finally, this process ideally leads to a decision or action (Visscher & Coe, 2003), which, in the larger policy cycle, is then evaluated. In line with this take on data use, we conceptualize school-level SPF use as SPF having been discussed, thoroughly analyzed, used as an impulse for further inquiry and a basis for formulating actions.

2.2 User Beliefs Relating to SPF

2.2.1 Attitude

Attitudes are formed by an agent's behavioral beliefs: their judgements of prospective outcomes of the behavior (Ajzen, 1991, 2002). Attitudes have an affective and a cognitive dimension, as beliefs can be based on emotions elicited by the behavior or on an evaluation of its attributes (Prenger & Schildkamp, 2018; Sanbonmatsu & Fazio, 1990). Applied to data use, affective attitude may refer to users feeling comfortable with, excited about, or apprehensive of using data (Jimerson, 2014; Vanhoof et al., 2013) and it has been found that data use anxiety contributes to DDDM resistance in educators (Dunn, Airola, & Garrison, 2013). Cognitive attitude, which is influenced by educators' general views on DDDM (Dunn

et al., 2019; Jimerson, 2014), is concerned with "buy-in" (Schildkamp & Kuiper, 2010) or the extent to which users regard data as useful (van der Kleij & Eggen, 2013; Vanhoof et al., 2013). In this study, we will relate perceived usefulness of the SPF to potential outcomes of SPF use. Usefulness for school development refers to acknowledging the potential of SPF for formulating concrete actions or decisions (instrumental use) and for inspiring the decision making process (conceptual use), whereas usefulness for accountability pertains to regarding SPF as a tool for supporting prior decisions (symbolic use) and for self-promotion and legitimatization (strategic use) (Hellrung & Hartig, 2013; Rossi et al., 2004; Visscher & Coe, 2003).

2.2.2 Subjective Norm

Subjective norm is produced by an agent's normative beliefs: the way they feel others expect them to engage in the behavior (Ajzen, 1991, 2002). Normative beliefs do not necessarily reflect perceived coercion, but can also refer to perceived encouragement. Data use expectations can emanate from actors within the school, e.g. a school leader advocating the use of a certain instrument, or from external parties, e.g. in systems where data use is associated with compliance (Uiterwijk-Luijk et al., 2017; Vanhoof et al., 2013). School-external expectations can also prompt school leaders to formulate expectations towards teachers (Abrams et al., 2020). Some scholars propose that social pressure is detrimental to data use, because it compromises educators' autonomous motivation to use data (Vanlommel et al., 2016) or their sense of ownership (Schildkamp & Teddlie, 2008). These concerns are raised in particular in contexts with an emphasis on school or student accountability, which typically involve high-stakes standardized testing (Nichols & Harris, 2016). Nevertheless, subjective norm has also been found to positively affect data use in certain circumstances. Social pressure to work in an inquiry-based manner is positively related to teachers displaying an inquiry habit of mind (Uiterwijk-Luijk et al., 2017) and both development-oriented and

accountability-based external expectations have been found to motivate principals to make use of data (Vanhoof et al., 2013).

2.2.3 Self-Efficacy

Perceived control is rooted in an agent's control beliefs: their perception of factors that help or inhibit them in engaging in the behavior (Ajzen, 1991, 2002). We will operationalize perceived control as self-efficacy. In the context of data use, an educational professional's self-efficacy expresses the extent to which they feel capable of engaging in data use because they possess the necessary competences to do so (Bandura, 1997; Van Gasse et al., 2017). In other words, self-efficacy expresses confidence over one's data literacy, i.e. one's knowledge and skills for processing data and formulating responses accordingly (Mandinach & Gummer, 2016). The literature paints a rather pessimistic picture with regard to educators' data literacy (e.g. van der Kleij & Eggen, 2013; Vanhoof et al., 2011) and finds that many still profess to feeling insecure in this respect (Datnow & Hubbard, 2016; Earl & Fullan, 2003). Nevertheless, a sense of self-efficacy is an important determinant of data use (Dunn, Airola, Lo, et al., 2013a). In this study, we conceptualize self-efficacy as SPF users feeling they are able to understand SPF, interpret it, and translate it into concrete actions. We thus acknowledge that data literacy surpasses mere statistical literacy, but also entails the transformation of information into actionable knowledge (Mandinach & Gummer, 2016).

2.3 School-level Features Relating to SPF

2.3.1 SPF-Oriented School Culture

A strong data culture in schools is a DDDM enabler and fosters data literate educators. Strong data cultures are grounded in a clear vision and common goals, and collaborative structures (Bryk, 2010; Hamilton et al., 2009; Jimerson, 2014; Jimerson et al., 2020; Jimerson & Wayman, 2015). Leadership is key in shaping and facilitating these data cultures (Bryk, 2010;

Hamilton et al., 2009; Jimerson, 2014; Jimerson et al., 2020) as is a mindset of continuous improvement (Sutherland, 2004). In order to assess the prevailing school culture regarding SPF, we will first focus on users' perception of a shared goal orientation within the team. This refers to sharing a common vision and understanding about SPF use, including collective norms and objectives (Hoogland et al., 2016; Jimerson, 2014; Mandinach, 2012; Schildkamp et al., 2019). In essence, it refers to the collective frame of reference on how and why to use SPF (cf. Schildkamp et al., 2014). Additionally, we will gauge users' experience of internal support and collaboration regarding SPF use. Support is associated with networking, brokerage and coaching (Jimerson, 2014; Schildkamp et al., 2019). Both support and collaboration foster educators' actual data use competences as well as their confidence in using data (Abrams et al., 2020; Datnow & Hubbard, 2016; Schildkamp & Kuiper, 2010; Van Gasse et al., 2017; Vanhoof et al., 2011).

2.3.2 Users' Work Role

Individual school team members have different data use competences (van der Kleij & Eggen, 2013), needs (Coburn & Talbert, 2006) and objectives (Schildkamp & Kuiper, 2010). Formal work roles play a part in these differences, as the nature of DDDM also differs according to an educator's position (Mandinach et al., 2011). For teachers, data use and data literacy are oriented predominantly towards instructional decision making rather than school development. Whereas school leaders are often the directors of how data use takes shape within their team by functioning as culture builders and modeling good practices (Jimerson, 2014; Schildkamp et al., 2014; Schildkamp & Teddlie, 2008; Vanhoof et al., 2012), teachers tend to need encouragement to use data (Uiterwijk-Luijk et al., 2017; Vanhoof et al., 2012) and hold the school leader responsible for setting up a data policy (Hoogland et al., 2016). Considering that school leaders are often former teachers, and teachers in turn learn from

school leaders, there is a certain amount of reciprocity among these educator roles in DDDM (Jimerson, 2014).

2.3.3 Feedback Sign

Positive feedback has a positive impact on feedback acceptance (Ilgen et al., 1979) while negative feedback can negatively impact recipients' perceptions (Lechermeier & Fassnacht, 2018). Data use research has indeed found that teachers are reluctant to engage with data that challenge their efficacy (Coburn & Turner, 2011; Dunn, Airola, Lo, et al., 2013a; Lockton et al., 2020). Although we know that a school's performance in an external assessment influences the way the resulting SPF is received (Verhaeghe et al., 2010; Visscher & Coe, 2003), few studies have specifically zoomed in on how the "sign" of SPF relates to usage, as we intend to do here. Corresponding to the most prevalent frames of reference employed in external standardized assessments (AERA et al., 2014; Hellrung & Hartig, 2013), we will consider both criterion- and norm-referenced SPF results. Criterion-referenced measures are absolute and compare achievement to a standard. Norm-referenced measures are relative and compare achievement to that of a reference group.

2.3.4 Voluntariness

SPF systems are inherently focused on school improvement, since they are a self-evaluation tool, but in practice they can also serve accountability goals (Visscher & Coe, 2003). While a certain degree of accountability pressure stimulates engagement with SPF (Vanhoof et al., 2012), the self-evaluation purpose of a system needs to be explicit in order to foster engagement with the data for school improvement (Maier, 2010). Thus, a careful balance needs to be struck. As voluntariness corresponds to the degree of "free will" in adopting a certain system (Wu & Lederer, 2009), we will account for voluntariness in feedback pursuit by taking into account whether or not a school takes purposeful action to acquire SPF. In

general, SPF systems tend to be successful when they address a perceived information deficiency from the recipients themselves (Hendriks et al., 2002), when they are adopted rather than imposed, and when the SPF recipients feel they have sufficient ownership over the implementation (Visscher & Coe, 2003).

3. Method

We developed an online survey that was completed by 470 Flemish educational professionals whose schools had recently been presented with an SPF report. We performed a path analysis on the survey data.

3.1 Research Context

The study took place in Flanders, the Dutch-speaking part of Belgium. The Flemish educational system is largely decentralized (OECD, 2017). Government-issued attainment targets describe minimum goals for different stages in primary and secondary education, but schools enjoy great autonomy. The inspectorate monitors whether schools comply with regulations and pay sufficient attention to internal quality, but to date, Flanders does not organize central examinations.

On system level, achievement of attainment targets is periodically measured with large-scale national assessments (NA) that cover a wide range of topics. NA take a snapshot of the performance of a population at a certain point in time by testing representative samples of students. Schools cannot volunteer for participation, nor can participation be enforced onto them, but sample schools do receive a personalized and strictly confidential SPF report. Similar SPF reports can be requested by administering parallel tests (PT) free of charge. PT are parallel versions of the tests administered in the NA, released after the national averages have been made public. NA and PT are highly standardized tests in terms of content,

administration and scoring (AERA et al., 2014), but they are low-stakes for schools and pupils. Results are not reported to the educational government nor are they made public.

The analyses used in the Flemish NA are grounded in item response theory (IRT). The standard that corresponds to achieving the attainment targets is determined by a panel of educational professionals and experts in a process based on the Bookmark procedure (Mitzel et al., 2001). The SPF reports describe the extent to which the attainment targets were reached within the school (criterion-referenced feedback) as well as the schools' performance relative to the national average (norm-referenced feedback). Note that in the PT reports, the representative NA sample constitutes the reference group. The SPF also contains value added information that corrects performance for input characteristics as a measure for "fair comparison" (Visscher & Coe, 2003).

3.2 Data Collection

3.2.1 Participants

An online survey was sent out to 427 schools which had participated in an NA of French in Grade 6 or Technology in Grade 8 (148 schools), or had voluntarily taken PT on various subjects (279 schools). These schools had received SPF about five months prior to the administration of the survey. Because we aimed to illuminate SPF use from multiple perspectives, we asked for responses from the school leadership and from (the) teacher(s) involved.

On school level, a response rate of 72% was achieved. In total, 470 online surveys were completed in full by educational professionals in both primary (60%) and secondary education (40%). Overall, respondents' average age was 46, and the number of females surpasses the number of males with 69% to 31%. Some 22% percent of respondents hold a master's degree; for the majority this is a bachelor's degree (76%).

As a result of convenience sampling there is some nesting of participants within schools, but this nesting is very limited. A majority of respondents were single observations within their school, as we also need to take into account that PT respondents from one and the same school did not all focus on the same test subject report to discuss SPF use. Consequently, in the analyses, respondents were treated as not nested within schools.

3.2.2 Instrument

The online survey assessed users' perceptions regarding the (type of) SPF that their schools had been presented with, in the case of NA participation, or had actively collected, by taking PT. One scale measured the extent to which the report had actually been put to use in the school, 'SPF use' for short. Five scales measured user beliefs: 'Affective attitude', 'Cognitive attitude: Usefulness for school development', 'Cognitive attitude: Usefulness for accountability', 'Subjective norm' and 'Self-efficacy'. Finally, two scales addressed the perceived presence of 'Shared goals' and 'Support and collaboration' with regard to SPF use .

Items were selected from other studies on data use and inspired by literature on the construction of TPB-based questionnaires (Ajzen, 2002; Francis et al., 2004; Pierce et al., 2013; Van Gasse et al., 2015; Vanhoof et al., 2013; Wayman et al., 2017). They were adapted to particularly tap into perceptions about SPF from NA and PT. All items were statements to be scored on a 5-point Likert scale (1 – entirely disagree, 2 – disagree, 3 – neither agree nor disagree, 4 – agree, 5 – entirely agree) with a possibility to opt out (I don't know/This statement does not apply). In order to establish face validity, the items were submitted to peer review.

The construct validity of each scale was examined with a confirmatory factor analysis (CFA) (Brown, 2006). This approach allowed us to take covariance between items into account and to optimize each model based on modification indices. Moreover, it provided an opportunity to handle missing data in an advanced manner by employing 'full information

maximum likelihood' (FIML) as an estimator. The analyses were conducted in R 3.5.1 'Feather Spray' with the lavaan-package (Rosseel, 2012). In order to remodel and ultimately assess the validity of each scale, we considered the comparative fit indices (CFI), the Tucker-Lewis indices (TLI), the root mean square errors of approximation (RMSEA) and the standardized root mean square residuals (SRMSR). For the CFI and TLI, a cutoff of .95 was exceeded in all models, and good fit was confirmed by the RMSEA and SRMSR values which were all smaller than or close to the cutoff of .08 (Schreiber et al., 2006). Reviewing the factor loadings of each item on the corresponding latent concept, we found that the loading of one item pertaining to 'Subjective norm' did not meet a cutoff set at .400. Based on content validity considerations, we decided to retain the item in the scale. As shown in Table 1, all scales show adequate reliability (Nunnally & Bernstein, 1994) with Cronbach's alpha values ranging from .67 to .87.

[Table 1 near here]

3.3 Data Analysis

3.3.1 Measures

For all scales, i.e. the dependent variable as well as user beliefs and characteristics of school culture, we moved forward with the factor scores predicted by the CFA. Respondents' work roles were coded into a dummy variable. Respondents exercising a predominantly coordinating function (52%) were assigned value 1, while those who mainly teach (48%) constituted the reference group.

The feedback sign of the SPF was retrieved from the (focal) SPF reports. We aimed to express the school's result in straightforward, continuous variables. In order to capture the

school's criterion-referenced result, we calculated the average percentage of pupils that reached the attainment targets over all tests in the (focal) report. These average percentages were then standardized into Z-scores. For the norm-referenced result, we compared the proportion of pupils that reach the attainment targets within the school to the proportion that had done so in the full NA sample. We made this comparison for all tests in the (focal) report. The average differences were, again, standardized into Z-scores.

In order to assess the effect of voluntariness, we took into account whether a response pertained to a PT school, where SPF was acquired voluntarily through active participation, or to an NA school, where the SPF resulted from more passive test participation. This information was coded into a dummy variable. The full sample of 470 complete survey responses consisted of 330 responses from schools that had taken PT (70%), assigned value 1, and 140 responses from schools that had participated in an NA (30%), which constituted the reference group.

3.3.2 Path Analysis

A path analysis was conducted with the R-package lavaan (Rosseel, 2012). In correspondence with the conceptual model, we started out with a model in which user beliefs mediate the effect on SPF use of school-level features. We assumed no covariance between variables. Based on the modification indices we gradually added covariance and meaningful regressions, and eliminated non-significant parameters in pursuit of a parsimonious model with optimal fit. With FIML as an estimator we were able to use 410 responses on a total of 470 responses. The final model fits the empirical data well (RMSEA = 0.029; SRMR = 0.015; CFI = 0.993; TLI = 0.961) and significantly better than the starting model ($\chi^2(7)=139.7$, p <.001).

4. Findings

In this section we will present the results of the path analysis. In order to provide some

perspective to these explanatory findings, we will first briefly discuss descriptive findings for the scale variables from the survey.

4.1 Descriptive Findings

As shown in Table 2, overall, Flemish educational professionals report fairly limited school-level use of SPF from NA and PT (M = 3.24) although responses vary considerably (SD = 1.10). A closer examination of the individual items revealed that in general, SPF reports are formally discussed with the team, but much less thoroughly analyzed, used as an impetus for further inquiry, or as input for formulating actions.

We find that users' affective attitude towards the use of NA and PT feedback is neutral (M = 3.15). Cognitively, they do not take an outspokenly positive or negative stance towards its usefulness for accountability (M = 3.41) but do regard it as a rather useful tool for school development (M = 3.97). There is little indication that respondents feel pressured to make use of SPF from NA and PT (M = 2.92). Note, however, that the subjective norm scale comprises quite some missing data: only 363 responses out of the overall 470 could be used in the descriptive analysis of this variable. Looking at users' perception of control over SPF use, we see they gauge their self-efficacy as rather high (M = 3.92). So, they are quite confident they possess the necessary competences to process the results fed back to them from NA and PT. A relatively small standard deviation indicates that users are relatively united in this perception (SD = 0.69).

Lastly, the data indicate that users do not experience a strong school culture towards SPF use from NA and PT. On average, users neither agree nor disagree with the thesis that their school team shares a common goal-orientation regarding SPF from NA and PT (M = 3.17). Their perception of support and collaboration is somewhat more positive but still situated towards the center point of the scale (M = 3.30). Note, however, that there are relatively large differences between individual users in this respect (SD = 1.03).

[Table 2 near here]

4.2 Explanatory Findings

The final path model is graphically represented in Figure 2. Full line arrows indicate regression and double-headed dashed arrows depict covariance. The standardized coefficients and the significance level of the effects are included. For user beliefs as mediating variables, and for school-level SPF use as a dependent variable, the R^2 values are mentioned in bold. For the sake of clarity, the figure only comprises those effects that are statistically significant ($p < .05$). A full overview of all parameters is given in Appendix 1 and Appendix 2.

[Figure 2 near here]

4.2.1 The (Mediating) Effect of User Beliefs on School-Level SPF Use

Users' perception of the SPF's 'Usefulness for school development', and their 'Subjective norm' or the extent to which they feel it is expected of them to engage with the SPF, both bear a small positive relationship with school-level 'SPF use' (β = .17 and .13 respectively). Thus, users' cognitive attitude and normative beliefs have a statistically significant impact on school-level SPF use (cf. RQ1). These beliefs mediate the effects of SPF-oriented school culture and users' work roles to a certain extent (cf. RQ2b). Users' perception of 'Shared goals' explains about 6% of the variance in perceived 'Usefulness for school development' (β = .24) and some 13% of the variance in 'Subjective norm' (β = .36). A regression coefficient of .13 shows that only a further 2% of variance in the former is explained by whether or not the user holds a coordinating role at the school (as opposed to primarily being a teacher).

Coordinators report a higher awareness of the feedback's potential for school development purposes, which is in turn associated with higher levels of reported school-level SPF use.

No other (mediating) effects of user beliefs on SPF use were identified (cf. RQ1 and RQ2b). The fact that users' affective attitude does not have a statistically significant impact on SPF use is in line with other studies exploring data use in Flanders (Van Gasse et al., 2015) but contradicts other findings that show that a favorable affective attitude can outweigh a favorable cognitive attitude (Vanhoof et al., 2013).

User beliefs do covary to varying extents. For instance, all factors on the attitudinal level covary positively, meaning that users' cognitive attitudes towards SPF (use) appear to run parallel to their affective attitude to a certain extent. 'Affective attitude' also covaries positively with 'Self-efficacy' (*cov* = .23), while negative covariance is found between 'Self-efficacy' and 'Subjective norm' (*cov* = -.11). Thus, users' level of enjoyment in engaging with SPF appears to correspond to a certain extent to the way they feel capable of this engagement, a finding we can relate to the fact that self-efficacy is often inversely related to anxiety (Dunn, Airola, Lo, et al., 2013b). On the other hand, users who experience more pressure to put SPF to use, appear to feel less capable of effectively doing so. Finally, we see that all user beliefs are impacted to varying extents by one or more school-level features we have taken into account.

4.2.2 The Direct and Indirect Impact of School-Level Features on School-Level SPF Use

Overall, the full path model explains about 26% of the variance in school-level SPF use (R^2 = .26). An SPF-oriented school culture, and more specifically users' perception of a shared goal orientation, proves to be the most salient determinant of school-level SPF use when controlling for other factors (cf. RQ2a). A feedback user's perception of 'Shared goals' regarding the use of SPF from NA and PT affects the actual school-level use of this feedback

directly (β = .31) as well as indirectly through perceived 'Usefulness for school development' and 'Subjective norm', as elaborated on above. 'Shared goals' also have a statistically significant impact on users' 'Affective attitude' (β = .21) and their 'Self-efficacy' (β = .30). The fact that this variable positively influences most of the user beliefs we measured, confirms that beliefs take shape within professional communities (Coburn & Talbert, 2006; Datnow & Hubbard, 2016; Jimerson, 2014). Users' perception of the usefulness of the SPF for accountability purposes is the only belief-driven predictor not affected by 'Shared goals'. However, accountability is not a dominant goal in this specific case and research context. Users' 'Self-efficacy', or their conviction of being capable of engaging with SPF, is further enhanced when they sense internal 'Support and collaboration' (β = .15). This is in line with our hypothesis and with other research (Abrams et al., 2020). As elaborated above, however, there is no ensuing statistically significant effect of 'Self-efficacy' on school-level SPF use.

Secondly, we find that users' formal work roles play a part in predicting their attitude towards SPF from NA and PT, which is in line with prior research (Van Gasse et al., 2015). Overall, coordinators tend to have a more positive 'Affective attitude' towards SPF use (β = .22) than teachers and they value the SPF's 'Usefulness for school development' somewhat higher (β = .13). As established above, this dimension of cognitive attitude is in turn a predictor for school-level SPF use (cf. RQ 2b). Work role does not have a statistically significant impact on users' 'Subjective norm' nor their 'Self-efficacy'. The latter is particularly counterintuitive, since we would expect coordinators to be more familiar with SPF-type data and thus more confident in handling those data. However, perhaps coordinators are more aware of the complexity of this feedback, precisely because they have more experience in dealing with it.

Thirdly, we see that feedback sign has no direct or indirect effect on SPF use (cf. RQ 2a and RQ2b). The school's criterion-referenced result bears no statistically significant

relationship at all to other variables in our model: the extent to which the school reaches the attainment targets is not proportionate to the favorability of user beliefs associated with the SPF. The norm-referenced result, however, is positively related to self-efficacy beliefs (β = .17) and to users' assessment of the SPF's usefulness for accountability purposes (β = .14). Thus, a more positive comparative result is associated with a stronger sense of being able to 'make sense' of the result. This suggests that, when a school scores along the average, it can be difficult to base a conclusion on that result, but when the comparative result is more saliently positive it is easier to process. The fact that it does not correspond to higher levels of perceived usefulness for school development nor to higher levels of SPF use, however, may entail that the conclusion is that no action needs to be taken when performance is satisfactory. However, higher norm-referenced performance is deemed more useful for accounting for oneself, which is perhaps not surprising. When a school has scored markedly better than the population, the team is more inclined to use that information to account for their policy or practice to other parties.

Fourthly, we find a clear and direct effect of voluntariness on school-level SPF use (cf. RQ 2a). SPF actively requested by schools is associated with a higher level of use (β = .13) than feedback simply presented to schools. There is no mediation: whether the SPF was collected actively or not, has no statistically significant impact on user beliefs.

5. Discussion and Conclusion

Feeding back output indicators to schools is at the nexus of school effectiveness and school improvement (Hulpia & Valcke, 2004; Visscher & Coe, 2003). The present study approached engagement with SPF from low-stakes standardized assessments as both a belief-driven phenomenon and a situated phenomenon. It sheds new light on the relative impact and the interplay of predictors of SPF use. By conducting quantitative research on a large dataset, we

explored how school-level SPF use is affected by individual users' attitudes, subjective norm and self-efficacy (RQ1). The analyses show that engagement with SPF in schools is only explained by these user beliefs to a limited extent when we control for other factors. We also examined the effects on SPF use of selected school-level features, namely SPF-oriented school culture, users' formal work roles, feedback sign and voluntariness (RQ2a). In addition, we investigated whether those effects are mediated by user beliefs (RQ2b). Our hypothesis that a user-centered outlook on SPF use needs to be situated within a school is confirmed to a certain extent. User beliefs regarding SPF are influenced by school-level features. However, not all of these relationships lead to a heightened use of SPF on school level. Thus, the mediating role of user beliefs is modest.

On the level of the individual user, our most salient findings include that SPF use will increase when users recognize the utility of SPF for school improvement. We also established that users report a higher level of engagement with SPF at their schools when they have a stronger sense that they are expected to use SPF. On a school organizational level, users' work role plays a small part in explaining perceptions about SPF use. Above all, however, our analyses demonstrate that a strong data culture is an important precondition for engagement with SPF. Users' perception of a shared goal-orientation within the team emerges as the most prominent predictor in the model. As in certain other studies, our analyses point out that it has a greater positive impact on data use than individual user characteristics (e.g. Van Gasse et al., 2015). On a contextual level, we found that data from a voluntarily adopted SPF system are used more intensively than data not actively collected. This suggests that also in an improvement-oriented context, ownership is an important driver.

Our findings attest to the fact that data themselves "do not drive" (Dowd, 2005; Lockton et al., 2020) and entail implications for practice. As with other types of educational innovations and interventions, and organizational change in general, data should meet users'

improvement needs and address what they find important (Ketelaar et al., 2012; Schildkamp & Teddlie, 2008). Research suggests that a positive cognitive attitude towards DDDM can be stimulated by providing instruction to educators. Instruction serves to address concerns regarding data use, and serves to identify and challenge views that might (otherwise) lead to reluctance to engage in DDDM (Dunn et al., 2019).

In the interest of ownership and putting educational professionals in the driver's seat of data use, inquiry-based working should be stimulated in schools. Programs and interventions can also explicitly promote a stronger data orientation in school teams. The purpose of data such as SPF should be clear to the educational professionals expected to make use of them. Those expectations should be made explicit, so that individuals sense data use is an integral part of their job and that their contribution is vital to the organizational dynamic of data use.

In the spirit of fostering continuous quality monitoring and continuous improvement, the introduction of professional learning communities such as data teams (Schildkamp et al., 2019) or networked improvement communities (LeMahieu et al., 2017) can strengthen data use practices and data cultures in schools. Such initiatives foster shared goals and provide a collaborative space in which individual perspectives and expertise are appreciated. School improvement becomes continuous improvement, a form of organizational learning (Datnow & Park, 2018; Dolle et al., 2018). A data culture is a setting in which data use is done by, not to the school (Sutherland, 2004).

Concerning schools' performance, we found that feedback sign has no statistically significant effect on school-level SPF use in our research context. The literature suggests that, whereas the effect of feedback sign on perceptions is rather straightforward, its effect on actual behavior is much more complex (Kluger & DeNisi, 1996; Lechermeier & Fassnacht, 2018). In empirical studies on SPF use, being confronted with lower performance has indeed

been found to prompt action as opposed to receiving positive feedback, but very negative feedback tends to be brushed off (Hellrung & Hartig, 2013; Verhaeghe et al., 2010). Drawing on social cognitive theory, Visscher and Coe (2003) addressed tensions surrounding SPF sign as follows: "Although negative feedback is necessary to motivate the need for improvement, without positive feedback, individuals are unlikely to believe themselves capable of achieving it" (p. 326). We contend that SPF systems can help educational professionals in making sense of their results by providing more sense in the message itself. In line with feedback intervention theory (Kluger & DeNisi, 1996, as discussed by Visscher & Coe, 2003) SPF systems should provide sufficient cues in the feedback message – also when that message is positive. Positive feedback tends to be automatically processed and is therefore more easily ignored, but providing more complexity to positive feedback turns it into guidance and not "just" praise (Geddes & Linnehan, 1996).

Limitations and Suggestions for Further Research

A TPB-approach assumes a linear perspective on behavior. In practice, however, data use is not a linear process. Therefore, further research should incorporate effects in other directions, such as the effect of school-level data use on individual beliefs (cf. Datnow & Hubbard, 2016) or the effect of individual users' attitude and self-efficacy on collaboration (cf. Dunn, Airola, Lo, et al., 2013a; Van Gasse et al., 2017). The TPB framework we applied can also be further extended. Researchers could consider including a measure for behavioral intention as a motivational mediator between psychological factors and behavior (cf. Prenger & Schildkamp, 2018), as well as a measure for actual behavioral control. Actual competences have featured in other studies about the influence of personal characteristics on data use (e.g. Vanhoof et al., 2011) and we should investigate how they complement the present model. While educators' data literacy is a fundamental prerequisite for effective DDDM, we know that educators' self-efficacy regarding data use, one of the central predictors in our conceptual

model, is often not in line with their actual knowledge and skills (Dunn, Airola, & Garrison, 2013).

Additionally, the construct of perceived control in itself can be revisited. We limited this variable to a self-efficacy measure, in line with a strand of other TPB-based studies (Armitage & Conner, 2001). However, perceived control over data use can also pertain to users' sense of having straightforward and timely access to the data (Pierce et al., 2013) or to their perceived autonomy in the decision-making process (Prenger & Schildkamp, 2018). Therefore we recommend exploring self-efficacy and other operationalizations of perceived control as separate constructs (cf. Prenger & Schildkamp, 2018). Furthermore, we point out that self-efficacy, like data literacy, is a multilayered construct. We captured and combined users' confidence in interpretation and their confidence in transforming information, in one overarching 'Self efficacy' measure. Future research should distinguish between different dimensions of efficacy, as prior research has established that these are indeed separate constructs and are regarded as such by data users (Dunn, Airola, Lo, et al., 2013b).

Further research could also consider fine-tuning other factors. We attempted to explain the mechanics of SPF use and not its actual effects, but did not differentiate between different phases. It would be beneficial to look at interpretation, analysis, and translation into action separately in order to account for the individuality of each phase and the mechanisms at play in each of them. Furthermore, we regarded feedback sign as an objective attribute of the SPF. It would interesting to reconceptualize this variable into feedback valence. A message's valence pertains to its attractiveness, i.e. whether it is perceived as positive or negative by the recipient (Geddes & Linnehan, 1996). Since several other predictors in our model are based on perceptions of SPF users, it would make sense to do the same for the feedback sign variable.

Because of the specificity of the research context, it is advisable to replicate this study in order to explore the generalizability of our findings. Replication studies could consider pursuing a multilevel design in order to further unravel individual differences between users and further enrich the information on prevailing data cultures in schools. Nevertheless, setting up this study in Flanders, an educational context low in outcome accountability, has allowed us to retain a clear focus on improvement-oriented data use. Moreover, SPF from Flemish NA and PT constituted a particularly suitable case for exploring our research questions. For one, because of the inclusion of feedback sign, an understudied predictor of data use. Results on Flemish NA and PT are research-based, refer to explicit standards and offer a clear normative benchmark. Additionally, because of the inclusion of voluntariness in feedback pursuit as a potential determinant of SPF use. A comparative analysis of NA participants and PT takers captured differences in engagement between two conditions in which the instrument was the same or very similar. Overall, Flemish schools provided a fruitful context for hypothesizing differential effects as they show great variability when it comes to data use (Vanhoof et al., 2012), which was confirmed by our descriptive analysis of school-level SPF use.

References

Abrams, L. M., Varier, D., & Mehdi, T. (2020). The intersection of school context and teachers' data use practice: Implications for an integrated approach to capacity building. *Studies in Educational Evaluation*. https://doi.org/10.1016/j.stueduc.2020.100868

AERA, APA, & NCME. (2014). *Standards for educational and psychological testing*. American Educational Research Association.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Ajzen, I. (2002). *Constructing a TpB questionnaire: Conceptual and methodological considerations*. https://pdfs.semanticscholar.org/6074/b33b529ea56c175095872fa40798f8141867.pdf

Ajzen, I. (2011). The theory of planned behaviour: Reactions and reflections. *Psychology & Health*, *26*(9), 1113–1127. https://doi.org/10.1080/08870446.2011.613995

Armitage, C. J., & Conner, M. (2001). Efficacy of the theory of planned behaviour: A meta-analytic review. *British Journal of Social Psychology*, *40*(4), 471–499. https://doi.org/10.1348/014466601164939

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W.H. Freeman.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. Guilford Publications.

Bryk, A. S. (2010). Organizing schools for improvement. *Phi Delta Kappan*, *91*(7), 23–30. https://doi.org/10.1177/003172171009100705

Coburn, C. E., & Talbert, J. E. (2006). Conceptions of evidence use in school districts: Mapping the terrain. *American Journal of Education*, *112*(4), 469–495. https://doi.org/10.1086/505056

Coburn, C. E., & Turner, E. O. (2011). Research on data use: A framework and analysis. *Measurement: Interdisciplinary Research & Perspective*, *9*(4), 173–206. https://doi.org/10.1080/15366367.2011.626729

Coburn, C. E., & Turner, E. O. (2012). The practice of data use: An introduction. *American Journal of Education*, *118*(2), 99–111. https://doi.org/10.1086/663272

Datnow, A., & Hubbard, L. (2016). Teacher capacity for and beliefs about data-driven decision making: A literature review of international research. *Journal of Educational Change*, *17*(1), 7–28. https://doi.org/10.1007/s10833-015-9264-2

Datnow, A., & Park, V. (2018). Opening or closing doors for students? Equity and data use in schools. *Journal of Educational Change*, *19*(2), 131–152. https://doi.org/10.1007/s10833-018-9323-6

Dolle, J., White, M. E., Evans-Santiago, B., Flushman, T., Guise, M., Hegg, S., Myhre, O., Ramirez, E., & Won, N. (2018). *Improvement science in teacher preparation at California State university. How teacher preparation partnerships are building capacity to learn to improve*. SRI International and WestEd.

Dowd, A. C. (2005). *Data don't drive: Building a practitioner-driven culture of inquiry to assess community college performance.* Lumina Foundation for Education Research Report. https://eric.ed.gov/?id=ED499777

Dunn, K. E., Airola, D. T., & Garrison, M. (2013). Concerns, knowledge, and efficacy: An application of the teacher change model to data driven decision-making professional development. *Creative Education*, *04*(10), 673–682. https://doi.org/10.4236/ce.2013.410096

Dunn, K. E., Airola, D. T., Lo, W.-J., & Garrison, M. (2013a). Becoming data driven: The influence of teachers' sense of efficacy on concerns related to data-driven decision making. *The Journal of Experimental Education*, *81*(2), 222–241. https://doi.org/10.1080/00220973.2012.699899

Dunn, K. E., Airola, D. T., Lo, W.-J., & Garrison, M. (2013b). What teachers think about what they can do with data: Development and validation of the data driven decision-making efficacy and anxiety inventory. *Contemporary Educational Psychology*, *38*(1), 87–98. https://doi.org/10.1016/j.cedpsych.2012.11.002

Dunn, K. E., Skutnik, A., Patti, C., & Sohn, B. (2019). Disdain to acceptance: Future teachers' conceptual change related to data-driven decision making. *Action in Teacher Education*, *41*(3), 193–211. https://doi.org/10.1080/01626620.2019.1582116

Earl, L., & Fullan, M. (2003). Using data in leadership for learning. *Cambridge Journal of Education*, *33*(3), 383–394. https://doi.org/10.1080/0305764032000122023

Francis, J. J., Eccles, M. P., Johnston, M., Walker, A., Grimshaw, J., Foy, R., Kaner, E. F. S., Smith, L., & Bonetti, D. (2004). *Constructing questionnaires based on the theory of planned behaviour: A manual for health services researchers*. Centre for Health Services Research, University of Newcastle upon Tyne. https://doi.org/0-9540161-5-7

Geddes, D., & Linnehan, F. (1996). Exploring the dimensionality of positive and negative performance feedback. *Communication Quarterly*, *44*(3), 326–344. https://doi.org/10.1080/01463379609370021

Hamilton, L. S., Halverson, R., Jackson, S., Mandinach, E. B., & Supovitz, J. A. (2009). *Using student achievement data to support instructional decision making*. United States Department of Education. http://repository.upenn.edu/gse_pubs

Hellrung, K., & Hartig, J. (2013). Understanding and using feedback – A review of empirical studies concerning feedback from external evaluations to teachers. *Educational Research Review*, *9*, 174–190. https://doi.org/10.1016/j.edurev.2012.09.001

Hendriks, M. A., 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. 115–142). Swets & Zeitinger.

Hoogland, I., Schildkamp, K., van der Kleij, F., Heitink, M., Kippers, W., Veldkamp, B., & Dijkstra, A. M. (2016). Prerequisites for data-based decision making in the classroom: Research evidence and practical illustrations. *Teaching and Teacher Education*, *60*, 377–386. https://doi.org/10.1016/j.tate.2016.07.012

Hulpia, H., & Valcke, M. (2004). The use of performance indicators in a school improvement policy: The theoretical and empirical context. *Evaluation & Research in Education*, *18*(1–2), 102–119. https://doi.org/10.1080/09500790408668311

Ilgen, D. R., Fisher, C. D., & Taylor, M. S. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology*, *64*(4), 349–371. https://doi.org/10.1037/0021-9010.64.4.349

Jimerson, J. B. (2014). Thinking about data: Exploring the development of mental models for "data use" among teachers and school leaders. *Studies in Educational Evaluation*, *42*, 5–14. https://doi.org/10.1016/j.stueduc.2013.10.010

Jimerson, J. B., Garry, V., Poortman, C. L., & Schildkamp, K. (2020). Implementation of a collaborative data use model in a United States context. *Studies in Educational Evaluation*. https://doi.org/10.1016/j.stueduc.2020.100866

Jimerson, J. B., & Wayman, J. C. (2015). Professional learning for using data: Examining teacher needs & supports. *Teachers College Record*, *117*(4), 1–36.

Ketelaar, E., Beijaard, D., Boshuizen, H. P. A., & Den Brok, P. J. (2012). Teachers' positioning towards an educational innovation in the light of ownership, sense-making and agency. *Teaching and Teacher Education*, *28*(2), 273–282. https://doi.org/10.1016/j.tate.2011.10.004

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*(2), 254–284. https://doi.org/10.1037/0033-2909.119.2.254

Lechermeier, J., & Fassnacht, M. (2018). How do performance feedback characteristics influence recipients' reactions? A state-of-the-art review on feedback source, timing, and valence effects. *Management Review Quarterly*, *68*(2), 145–193. https://doi.org/10.1007/s11301-018-0136-8

LeMahieu, P. G., Grunow, A., Baker, L., Nordstrum, L. E., & Gomez, L. M. (2017). Networked improvement communities. *Quality Assurance in Education*, *25*(1), 5–25. https://doi.org/10.1108/QAE-12-2016-0084

Lockton, M., Weddle, H., & Datnow, A. (2020). When data don't drive: Teacher agency in data use efforts in low-performing schools. *School Effectiveness and School Improvement*, *31*(2), 243–265. https://doi.org/10.1080/09243453.2019.1647442

Maier, U. (2010). Accountability policies and teachers' acceptance and usage of school performance feedback – A comparative study. *School Effectiveness and School Improvement*, *21*(2), 145–165. https://doi.org/10.1080/09243450903354913

Mandinach, E. B. (2012). A perfect time for data use: Using data-driven decision making to inform practice. *Educational Psychologist*, *47*(2), 71–85. https://doi.org/10.1080/00461520.2012.667064

Mandinach, E. B., & Gummer, E. S. (2013). A systemic view of implementing data literacy in educator preparation. *Educational Researcher*, *42*(1), 30–37. https://doi.org/10.3102/0013189X12459803

Mandinach, E. B., & Gummer, E. S. (2016). What does it mean for teachers to be data literate: Laying out the skills, knowledge, and dispositions. *Teaching and Teacher Education*, *60*, 366–376. https://doi.org/10.1016/j.tate.2016.07.011

Mandinach, E. B., Gummer, E. S., & Muller, R. D. (2011). *The complexities of integrating data-driven decision making into professional preparation in schools of education: It's harder than you think*. CNA Education, Education Northwest, and WestEd.

Mandinach, E. B., & Schildkamp, K. (2020a). Misconceptions about data-based decision making in education: An exploration of the literature. *Studies in Educational Evaluation*. https://doi.org/10.1016/j.stueduc.2020.100842

Mandinach, E. B., & Schildkamp, K. (2020b). The complexity of data-based decision making: An introduction to the special issue. *Studies in Educational Evaluation*. https://doi.org/10.1016/j.stueduc.2020.100906

Marsh, J. A. (2012). Interventions promoting educators' use of data: Research insights and gaps. *Teachers College Record*, *114*(11), 1–48.

Mitzel, G. H., Lewis, D. M., Patz, R. J., & Green, D. R. (2001). The Bookmark procedure: Psychological perspectives. In G. J. Cizek (Ed.), *Setting performance standards. Concepts, methods and perspectives* (pp. 249–281). Lawrence Erlbaum Associates.

Nichols, S. L., & Berliner, D. C. (2007). *Collateral damage: How high-stakes testing corrupts America's schools*. Harvard Education Press.

Nichols, S. L., Glass, G. V., & Berliner, D. C. (2006). High-stakes testing and student achievement: Does accountability pressure increase student learning? *Education Policy Analysis Archives*, *14*(1). https://doi.org/10.14507/epaa.v14n1.2006

Nichols, S. L., Glass, G. V, & Berliner, D. C. (2012). High-stakes testing and student achievement: Updated analyses with NAEP data. *Education Policy Analysis Archives*, *20*(20). https://doi.org/10.14507/epaa.v20n20.2012

Nichols, S. L., & Harris, L. R. (2016). Accountability assessment's effects on teachers and schools. In G. T. L. Brown & L. R. Harris (Eds.), *Handbook of human and social conditions in assessment* (pp. 40–56). Routledge.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.

OECD. (2017). *Education policy outlook: Belgium*. Organisation for Economic Co-operation and Development. www.oecd.org/edu/profiles.htm

Pierce, R., Chick, H., & Gordon, I. (2013). Teachers' perceptions of the factors influencing their engagement with statistical reports on student achievement data. *Australian Journal of Education*, *57*(3), 237–255. https://doi.org/10.1177/0004944113496176

Prenger, R., & Schildkamp, K. (2018). Data-based decision making for teacher and student learning: A psychological perspective on the role of the teacher. *Educational Psychology*, *38*(6), 734–752. https://doi.org/10.1080/01443410.2018.1426834

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: A systematic approach.* (7th ed.). Sage Publications.

Sanbonmatsu, D. M., & Fazio, R. H. (1990). The role of attitudes in memory-based decision making. *Journal of Personality and Social Psychology*, *59*(4), 614–622. https://doi.org/10.1037/0022-3514.59.4.614

Schildkamp, K. (2019). Data-based decision-making for school improvement: Research insights and gaps. *Educational Research*, *61*(3), 257–273. https://doi.org/10.1080/00131881.2019.1625716

Schildkamp, K., Karbautzki, L., & Vanhoof, J. (2014). Exploring data use practices around Europe: Identifying enablers and barriers. *Studies in Educational Evaluation*, *42*, 15–24. https://doi.org/10.1016/j.stueduc.2013.10.007

Schildkamp, K., & Kuiper, W. (2010). Data-informed curriculum reform: Which data, what purposes, and promoting and hindering factors. *Teaching and Teacher Education*, *26*(3), 482–496. https://doi.org/10.1016/j.tate.2009.06.007

Schildkamp, K., Poortman, C. L., Ebbeler, J., & Pieters, J. M. (2019). How school leaders can build effective data teams: Five building blocks for a new wave of data-informed decision making. *Journal of Educational Change*, *20*(3), 283–325. https://doi.org/10.1007/s10833-019-09345-3

Schildkamp, K., Poortman, C. L., Luyten, H., & Ebbeler, J. (2017). Factors promoting and hindering data-based decision making in schools. *School Effectiveness and School Improvement*, *28*(2), 242–258. https://doi.org/10.1080/09243453.2016.1256901

Schildkamp, K., & Teddlie, C. (2008). School performance feedback systems in the USA and in The Netherlands: A comparison. *Educational Research and Evaluation*, *14*(3), 255–282. https://doi.org/10.1080/13803610802048874

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. https://doi.org/10.3200/JOER.99.6.323-338

Sutherland, S. (2004). Creating a culture of data use for continuous improvement: A case study of an Edison Project school. *American Journal of Evaluation*, *25*(3), 277–293. https://doi.org/10.1016/j.ameval.2004.05.009

Uiterwijk-Luijk, L., Krüger, M., Zijlstra, B., & Volman, M. (2017). The relationship between psychological factors and inquiry-based working by primary school teachers. *Educational Studies*, *43*(2), 147–164. https://doi.org/10.1080/03055698.2016.1248901

van der Kleij, F. M., & Eggen, T. J. H. M. (2013). Interpretation of the score reports from the computer program LOVS by teachers, internal support teachers and principals. *Studies in Educational Evaluation*, *39*(3), 144–152. https://doi.org/10.1016/j.stueduc.2013.04.002

Van Gasse, R., Vanhoof, J., Mahieu, P., & Van Petegem, P. (2015). *Informatiegebruik door schoolleiders en leerkrachten. [Information use by school leaders and teachers]*. Garant.

Van Gasse, R., Vanlommel, K., Vanhoof, J., & Van Petegem, P. (2017). The impact of collaboration on teachers' individual data use. *School Effectiveness and School Improvement*, *28*(3), 489–504. https://doi.org/10.1080/09243453.2017.1321555

Vanhoof, J., Vanlommel, K., Thijs, S., & Vanderlocht, H. (2013). Data use by Flemish school principals: Impact of attitude, self-efficacy and external expectations. *Educational Studies*, *40*(1), 48–62. https://doi.org/10.1080/03055698.2013.830245

Vanhoof, J., Verhaeghe, G., Van Petegem, P., & Valcke, M. (2012). Flemish primary teachers' use of school performance feedback and the relationship with school characteristics. *Educational Research*, *54*(4), 431–449. https://doi.org/10.1080/00131881.2012.734726

Vanhoof, J., Verhaeghe, G., Verhaeghe, J. P., Valcke, M., & van Petegem, P. (2011). The influence of competences and support on school performance feedback use. *Educational Studies*, *37*(2), 141–154. https://doi.org/10.1080/03055698.2010.482771

Vanlommel, K., Van Gasse, R., Vanhoof, J., & Van Petegem, P. (2017). Teachers' decision-making: Data based or intuition driven? *International Journal of Educational Research*, *83*(March 1994), 75–83. https://doi.org/10.1016/j.ijer.2017.02.013

Vanlommel, K., Vanhoof, J., & Van Petegem, P. (2016). Data use by teachers: The impact of motivation, decision-making style, supportive relationships and reflective capacity. *Educational Studies*, *42*(1), 36–53. https://doi.org/10.1080/03055698.2016.1148582

Verhaeghe, G., Vanhoof, J., Valcke, M., & Van Petegem, P. (2010). Using school performance feedback: Perceptions of primary school principals. *School Effectiveness and School Improvement*, *21*(2), 167–188. https://doi.org/10.1080/09243450903396005

Visscher, A. J. (2020). On the value of data-based decision making in education: The evidence from six intervention studies. *Studies in Educational Evaluation*. https://doi.org/10.1016/j.stueduc.2020.100899

Visscher, A. J., & Coe, R. (2003). School performance feedback systems: Conceptualisation, analysis, and reflection. *School Effectiveness and School Improvement*, *14*(3), 321–349. https://doi.org/10.1076/sesi.14.3.321.15842

Wayman, J. C., Cho, V., Mandinach, E. B., Supovitz, J. A., & Wilkerson, S. B. (2017). *Teacher data use survey : Teacher version*. REL Appalachia and the Institute of Educational Sciences, United States Department of Education.

Wu, J., & Lederer, A. (2009). A meta-analysis of the role of environment-based voluntariness in information technology acceptance. *MIS Quarterly*, *33*(2), 419. https://doi.org/10.2307/20650298

Tables

Table 1. Overview of the Survey Scales

Scale *Example item*	Number of items	Cronbach's Alpha
SPF use *At our school, this feedback report was thoroughly analyzed.*	4	0.80
Affective attitude *I enjoy engaging with feedback from NA/PT.*	4	0.77
Cognitive attitude: Usefulness for school development *I consider this feedback report useful for supporting vision development within the school.*	4	0.83
Cognitive attitude: Usefulness for accountability *I consider this feedback report useful for justifying our methods to outsiders.*	4	0.77
Subjective norm *People whose opinion I value, expect me to engage with feedback from NA/PT.*	5	0.67
Self-efficacy *I feel I have the necessary skills to understand the content of (the) feedback reports (from an NA).*	5	0.82
Shared goals *At our school, there is a clear vision about how to use feedback from NA/PT.*	4	0.82
Support and collaboration *At our school, people make optimal use of each other's skills in order to engage with feedback from NA/PT.*	4	0.87

Table 2. Descriptive Statistics of the Survey Scales

Scale	n	M [a]	SD
SPF use	426	3.24	1.10
Affective attitude	420	3.15	0.78
Cognitive attitude: Usefulness for school development	445	3.97	0.78
Cognitive attitude: Usefulness for accountability	417	3.41	0.83
Subjective norm	363	2.92	0.79
Self-efficacy	448	3.92	0.69
Shared goals	416	3.17	0.84
Support and collaboration	420	3.30	1.03

Note. [a] Mean values ranging from 1 (entirely disagree) to 5 (entirely agree).

Figures

Figure 1. Conceptual Model

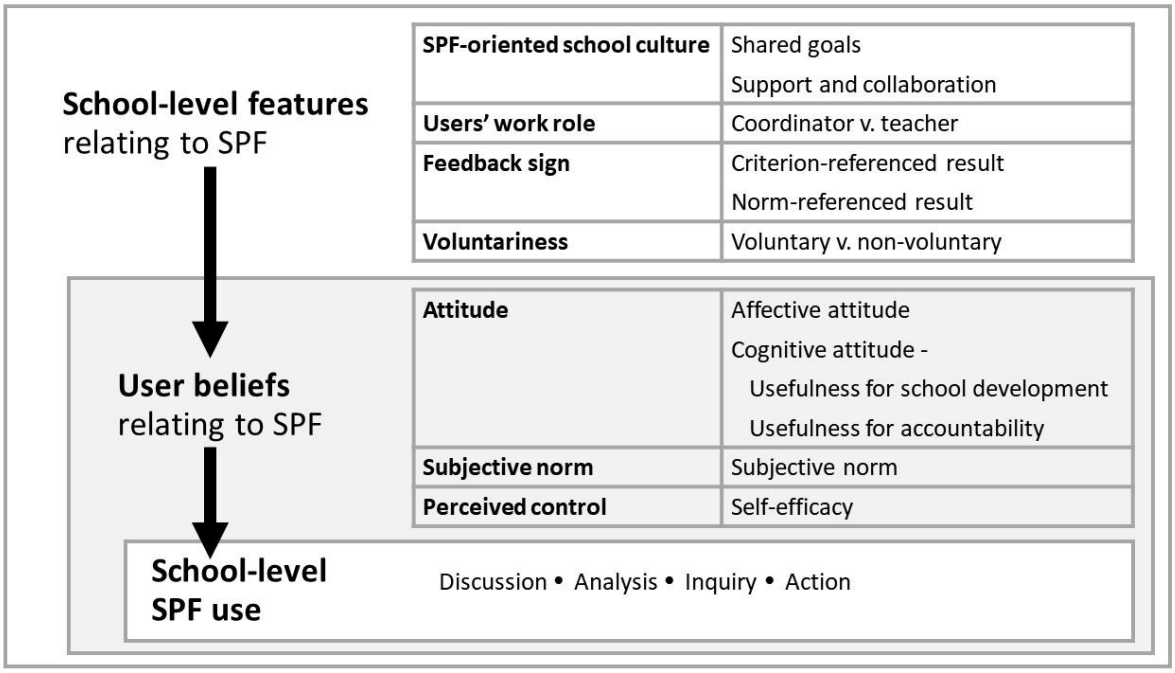
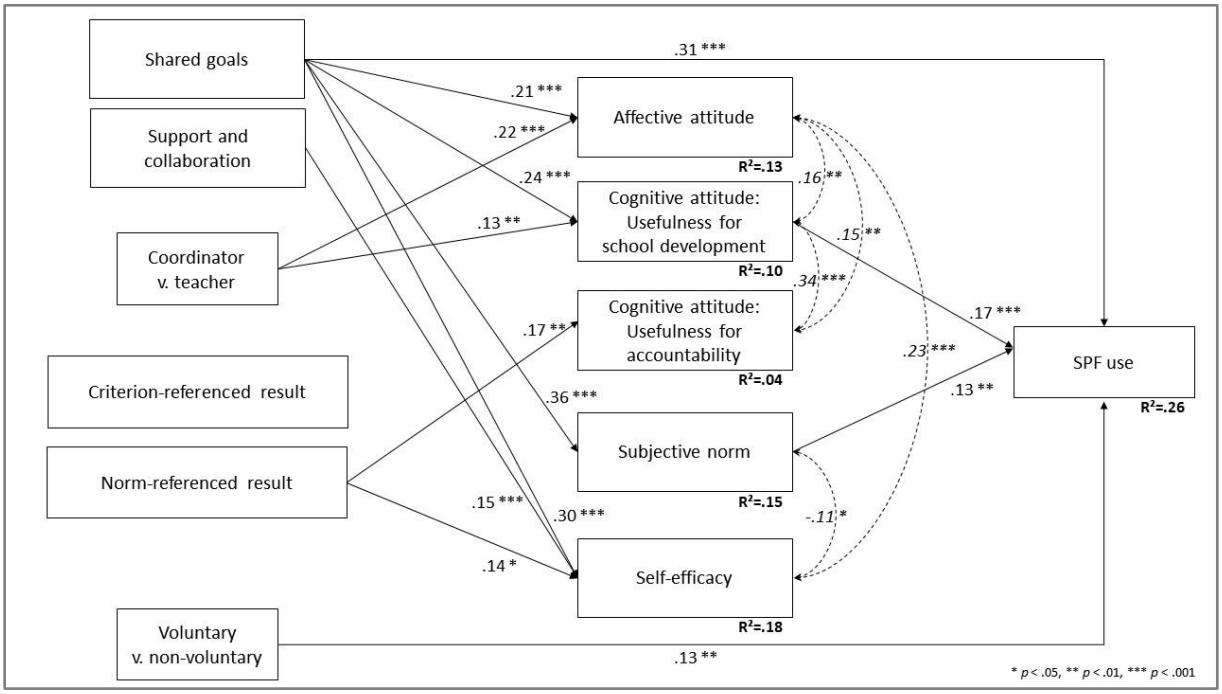


Figure 2. Path Model



Appendices

Appendix 1. Regression Parameters of the Path Model

Regression	B [a]	SE [b]	Z [c]	β [d]	p [e]	Sig [f]	R^2 [g]
SPF use ~							.264
Cognitive attitude: Usefulness for school development	0.173	0.047	3.656	0.175	<.001	***	
Cognitive attitude: Usefulness for accountability	-0.047	0.047	-1.009	-0.046	.313	ns	
Affective attitude	0.047	0.046	1.034	0.048	.301	ns	
Subjective norm	0.136	0.048	2.839	0.132	.005	**	
Self-efficacy	0.038	0.047	0.827	0.039	.408	ns	
Shared goals	0.305	0.050	6.035	0.312	<.001	***	
Voluntary v. non-voluntary	0.253	0.087	2.906	0.127	.004	**	
Affective attitude ~							.125
Shared goals	0.211	0.052	4.077	0.210	<.001	***	
Support and collaboration	-0.005	0.049	-0.109	-0.006	.913	ns	
Coordinator v. teacher	0.411	0.088	4.699	0.221	<.001	***	
Criterion-referenced result	0.027	0.056	0.482	0.029	.630	ns	
Norm-referenced result	0.087	0.056	1.566	0.095	.117	ns	
Voluntary v. non-voluntary	0.025	0.097	0.256	0.012	.798	ns	
Cognitive attitude: Usefulness for school development ~							.100
Shared goals	0.233	0.051	4.523	0.237	<.001	***	
Support and collaboration	0.056	0.048	1.161	0.059	.245	ns	
Coordinator v. teacher	0.234	0.087	2.689	0.129	.007	**	
Criterion-referenced result	0.008	0.055	0.141	0.009	.888	ns	
Norm-referenced result	-0.007	0.055	-0.122	-0.007	.903	ns	
Voluntary v. non-voluntary	0.101	0.096	1.053	0.050	.293	ns	
Cognitive attitude: Usefulness for accountability ~							.043
Shared goals	0.095	0.051	1.846	0.100	.065	ns	
Support and collaboration	0.016	0.048	0.325	0.017	.745	ns	
Coordinator v. teacher	-0.158	0.087	-1.817	-0.090	.069	ns	
Criterion-referenced result	-0.067	0.055	-1.214	-0.076	.225	ns	
Norm-referenced result	0.150	0.055	2.717	0.172	.007	**	
Voluntary v. non-voluntary	-0.146	0.096	-1.516	-0.075	.129	ns	
Subjective norm ~							.150
Shared goals	0.344	0.048	7.107	0.360	<.001	***	
Support and collaboration	0.058	0.045	1.268	0.063	.205	ns	
Coordinator v. teacher	-0.054	0.082	-0.653	-0.030	.514	ns	
Criterion-referenced result	-0.075	0.052	-1.434	-0.085	.152	ns	
Norm-referenced result	0.039	0.053	0.748	0.045	.454	ns	
Voluntary v. non-voluntary	-0.059	0.090	-0.649	-0.030	.516	ns	
Self-efficacy ~							.175
Shared goals	0.299	0.050	5.978	0.299	<.001	***	
Support and collaboration	0.140	0.047	2.992	0.146	.003	**	
Coordinator v. teacher	-0.073	0.085	-0.862	-0.039	.389	ns	
Criterion-referenced result	-0.019	0.054	-0.357	-0.021	.721	ns	
Norm-referenced result	0.128	0.054	2.372	0.139	.018	*	
Voluntary v. non-voluntary	0.013	0.093	0.141	0.006	.888	ns	

Note. [a] unstandardized coefficient; [b] standard error; [c] z-value; [d] standardized coefficient; [e] p-value; [f] significance; [g] explained variance.

ns $p > .05$, * $p < .05$, ** $p < .01$, *** $p < .001$

Appendix 2. Covariance Parameters of the Path Model

Covariance	B [a]	SE [b]	Z [c]	β [d]	p [e]	Sig [f]
.Cognitive attitude: Usefulness for school development ~~ .Cognitive attitude: Usefulness for accountability	0.250	0.039	6.446	0.337	<.001	***
.Cognitive attitude: Usefulness for school development ~~ .Affective attitude	0.117	0.037	3.168	0.156	.002	**
.Cognitive attitude: Usefulness for accountability ~~ .Affective attitude	0.110	0.037	3.000	0.147	.003	**
.Affective attitude ~~ .Self-efficacy	0.168	0.036	4.613	0.231	<.001	***
.Self-efficacy ~~ .Subjective norm	-0.078	0.033	-2.326	-0.114	.020	*

Note. [a] unstandardized coefficient; [b] standard error; [c] z-value; [d] standardized coefficient; [e] p-value; [f] significance.

* $p < .05$, ** $p < .01$, *** $p < .001$