

Gutwirth, Gila; Goffin, Evelyn; Vanhoof, Jan

Sensemaking unraveled: how teachers process school performance feedback data

Studia paedagogica. 2021, vol. 26, iss. 4, pp. [67]-97

ISSN 1803-7437 (print); ISSN 2336-4521 (online)

Stable URL (DOI): <https://doi.org/10.5817/SP2021-4-4>

Stable URL (handle): <https://hdl.handle.net/11222.digilib/144769>

Access Date: 28. 11. 2024

Version: 20220831

Terms of use: Digital Library of the Faculty of Arts, Masaryk University provides access to digitized documents strictly for personal use, unless otherwise specified.

SENSEMAKING UNRAVELED: HOW TEACHERS PROCESS SCHOOL PERFORMANCE FEEDBACK DATA

GILA GUTWIRTH,
EVELYN GOFFIN,
JAN VANHOOF

Abstract

The present study investigates how Flemish middle school mathematics teachers make sense of school performance feedback data from low-stakes, external standardized tests. We take an in-depth look into the interpretive steps they take, based on a conceptual model that integrates intuitive and rational aspects of individual and collective sensemaking and empirical data collected in semi-structured interviews. We describe the nature of these sensemaking processes and consider the impact of influencing factors. Our findings demonstrate that the mere availability of school performance feedback data does not spontaneously spark sensemaking, nor does it necessarily lead to improvements in instructional practice. Teachers' sensemaking of school performance feedback data appears to be a largely intuitive process, grounded in external attributions and absent of triangulation. Challenges regarding expertise and lack of inquiry-based attitude and commitment result in superficial and often incorrect interpretations of the data that tend to remain uncorrected as teachers barely engage in collaborative professional dialogue about the data.

Keywords

sensemaking, data-driven decision making, school performance feedback, intuition, rationality

Problem Statement

In recent decades, extensive decentralization and deregulation have resulted in a growing degree of autonomy for schools worldwide in terms of shaping their school policies. Policymakers assume that schools possess sufficient policy-making capacity to implement high-quality policies and to investigate and monitor their internal quality in a systematic way. Moreover, both in educational research and from a societal point of view, there is an increasing emphasis on informed school development that is based on objective, reliable, and valid data and not just on intuition and experience (Lai et al., 2014; Van Gasse et al., 2017). Educational professionals might use a host of data to inform policy and practice, including formal data such as school performance feedback or student achievement results from assessments, informal data such as classroom observations, and research evidence or even big data (Schildkamp, 2019). However, research shows that sustainable school development cannot be achieved by merely collecting and providing data (Vanlommel & Schildkamp, 2019). What raw data mean—what data mean in relation to pre-defined goals and how the data might serve to inform decisions and actions that effectively address school and student needs—is seldom self-explanatory: transforming data into information and subsequently into actionable knowledge requires recipients to *make sense* of the data within their own specific setting (Schildkamp et al., 2019).

Sensemaking is regarded as a crucial phase in the systematic, goal-oriented, and iterative process that is data-based decision making (DBDM) (Schildkamp et al., 2013; Verhaeghe et al., 2010). It involves actively analyzing data, forming interpretations, and making inferences (Coburn & Turner, 2011; Mandinach & Gummer, 2016; Schildkamp, 2019; Spillane & Miele, 2007; Vanlommel & Schildkamp, 2019). This interpretive process is neither straightforward nor exclusively rational (Bertrand & Marsh, 2015; Vanlommel & Schildkamp, 2019). Moreover, since individual sensemakers each have their own specific frame of reference and prior experiences that form personal, subjective lenses, the same data can come to hold different meanings for different educational professionals. Research therefore attests to the importance of collective sensemaking in DBDM in schools (Mandinach & Schildkamp, 2021). Collective sensemaking is a process of co-construction that takes shape in social interactions. With regard to teachers' team-level sensemaking, Bolhuis et al. (2016) referred to the positive impact of cognitive conflict: the tension created by divergent knowledge and assumptions between participants and by discrepancies between known and new information. Provided that cognitive conflict is embedded in a constructive collaboration based on openness and trust, it can lead to possible adjustments of instruction and learning, but this requires a profound professional dialogue among teachers.

However, research shows that this is the exception rather than the rule, resulting in superficial short-term solutions and quick fixes (Mausethagen et al., 2019).

The present study is a response to researchers' calls for more insight into the sensemaking process in DBDM (Mandinach & Schildkamp, 2021; Schildkamp, 2019). Our primary research aims are to investigate the interpretive steps teachers take when they make sense of data and to tap into the way teachers' individual sensemaking is embedded in collective sensemaking processes within their schools. In addition, we want to explore the impact of a number of explanatory variables on individual and collective sensemaking. A wide range of factors have been found to impact DBDM, including factors at the user level, the organization level, and the level of the data (systems) themselves (Bolhuis et al., 2016; Schildkamp et al., 2019; Van Gasse et al., 2015; Vanhoof et al., 2011; Verhaeghe et al., 2010). Personal characteristics such as data users' attitudes, self-efficacy, and data literacy have been discussed (Bolhuis et al., 2016; Prenger & Schildkamp, 2018; Van Gasse et al., 2015, 2017; Verhaeghe et al., 2010), as has the impact of collaboration on teachers' individual data use (Van Gasse et al., 2017) and contextual expectations such as the stakes associated with assessment or accountability structures in educational systems (Datnow & Park, 2018; Van Gasse et al., 2014; Vanlommel & Schildkamp, 2019). The relational factors that may affect data use in general and the sensemaking phase in particular have been less extensively explored.

With a qualitative, interview-based inquiry into teachers' individual and collective sensemaking, we want to address these knowledge gaps. We will do so by exploring a phase of sensemaking in relation to school performance feedback data—formal data, often achievement-based, that is confidentially provided to schools by an external party for self-evaluation (Schildkamp & Teddlie, 2008; Visscher & Coe, 2003). Specifically, we zoom in on teachers' use of school performance feedback data from low-stakes standardized testing aimed at internal quality assurance in Flanders (the northern, Dutch-speaking region of Belgium). Many studies addressing aspects of individual and/or collective sensemaking were set in high-accountability educational systems (e.g., Datnow et al., 2012) or (other) high-stakes decision-making contexts (e.g., Vanlommel & Schildkamp, 2019). However, expectations regarding DBDM are also becoming more salient in contexts and systems traditionally marked by a lower degree of accountability, such as the research context of Flanders. Performing our research in a low-accountability setting will allow us to unravel sensemaking processes from a school improvement logic that is minimally conflated with external (accountability) expectations.

In summary, with this qualitative study, we intend to examine how teachers make sense of school performance feedback data of standardized tests, and how this sensemaking process unfolds within the complexity of teachers'

own school (team) context. In addition, we investigate whether certain explanatory variables have an impact on individual and collective sensemaking. This translates into the following research questions:

- 1) How do teachers individually and collectively make sense of school performance feedback data?
- 2) Which factors promote or hinder teachers' sensemaking of school performance feedback data?

Theoretical Framework

We intend to investigate what happens in terms of sensemaking when teachers are confronted with school performance feedback data; specifically, we examine how they engage with personalized feedback reports that provide output results. For the purposes of this investigation, we conceptualize sensemaking by teachers within the cycle of improvement-oriented data use as *a continuous process in which teachers, from their own frame of reference and (school) context, individually and in interaction, notice and interpret information from school performance feedback data in a manner that will enable them to transform their schools' results into decisions and actions aimed at improving instructional practice and student achievement.*

In the following paragraphs, we unpack and substantiate the different components of our proposed conceptualization. We start by generally situating the sensemaking construct and a number of its central tenets, grounded in insights from cognitive and social psychology as well as literature on organizational change and knowledge management. Next, we distinguish between individual and collective sensemaking. In terms of individual sensemaking, we will look at the interpretive steps teachers take when engaging with the data, i.e., how they notice specific elements or cues and subsequently interpret these cues by framing them and by forming a judgment (Coburn & Turner, 2011; Spillane & Miele, 2007). In order to account for the fact that (individual) sensemaking involves intuitive as well as rational processes, and in line with prior sensemaking research, we will employ a dual-processing perspective (Kahneman & Frederick, 2005; Vanlommel et al., 2017). In terms of collective sensemaking, we explore the nature and affordances of professional dialogue (Gergen et al., 2004; Tsoukas, 2009). After all, discussing sensemaking in the context of DBDM in schools needs to take into account the fact that data use is rarely an isolated activity (Schildkamp, 2019). Moreover, individual data processing is shaped by the context and social environment of the sensemaker (Coburn & Turner, 2011; Spillane, 2012). Finally, our selection of influencing factors to explore is focused primarily on the relational factors that might shape sensemaking

and sensemaking outcomes in school teams, as the object of sensemaking under scrutiny in this study is formed by school performance feedback data. Our selection is based on the premise that collective inquiry is supported by the presence of sufficient human capital (pertaining to knowledge and expertise in participants) and social capital (pertaining to interaction between participants) (Christman et al., 2016).

Defining Sensemaking

The sensemaking construct has roots in literature on organizational change, crisis situations, information processing and workplace learning. Organizational psychologist Karl E. Weick (1995), historically regarded as one of the most influential sensemaking theorists, defines sensemaking as a social and continuous process in which people, from their own identities, retrospectively give meaning to cues and uncertainties in their environment, and proposes that (true) sensemaking leads to changes in beliefs or action. Broadly speaking, sensemaking thus has “outcomes” in the sense that it leads to some type of change in thought or behavior; it has cognitive, interpretive properties, as well as social, discursive properties; and it has a temporal dimension (Maitlis & Christianson, 2014).

Many theorists have elaborated on different aspects of sensemaking. Klein et al. (2007) discussed how sensemaking can be aimed at abstract or functional understanding of a situation or a novel set of data. In a DBDM discourse, we could connect those ideas of abstract vs. functional understanding to types of data use, such as to the distinction between conceptual and instrumental data use. Klein et al. (2007) regarded sensemaking primarily as a cognitive process, and discussed individual judgment, interpretation, and the role of internal mental models or the explanatory frames people use to interpret cues. Other authors, such as Cook and Gregory (2019), focused more on the discursive aspects of sensemaking, stating that sensemaking predominantly manifests itself through conversations and stories. Maitlis (2005) confirmed this social constructivist nature of the sensemaking process. In contrast, Klein et al. (2007) defined sensemaking as a cognitive process and focused on individual judgment, interpretation, and internal mental models. Hill and Levenhagen (1995) connected both views by stating that sensemaking consists of developing and formulating an individual vision or mental model that subsequently may serve to create support for one's views (sensegiving). Finally, regarding the temporal aspect of sensemaking, and in contrast to the original Weickian definition, authors such as Gephart et al. (2010) emphasized the prospective nature of sensemaking. They suggested that sensemaking is an ongoing and shared process in which meaning is

produced, negotiated, and maintained through verbal and nonverbal communication, adding that sensemaking uses the past to give meaning to future actions in the present. Similarly, Weick et al. (2005) stressed the necessity of sensemaking in organizations in order to achieve long-term goals.

*Interpretive steps in individual sensemaking,
from a dual-processing perspective*

Noticing

While reading school performance feedback reports, each individual teacher will—from their personal prior knowledge, expertise, previous (work) experiences, their beliefs about high quality instruction and about specific students and class groups, their identity and emotions—quickly and unconsciously notice certain cues and relate them to information already stored in memory (Klein et al., 2007; Kudesia, 2017; Maitlis et al., 2013; Weick et al., 2005). Intuitively noticing, recognizing, and selectively paying attention to information in school performance feedback data occurs automatically and reactively. These selective perceptions focus attention and thought, causing other information to be ignored or less relevant information to be amplified. Since interpretations resulting from this fast, unconscious “System 1” thinking (Kahneman, 2011) are based on incomplete information and information intuitively selected from data—whether scientifically collected or not—they are not always accurate (Shleifer, 2012). Additionally, this partly explains differences in interpretations between teachers (Spillane et al., 2002). Noticing more complex or unexpected cues in a feedback report, however, requires mental effort: the activation of teachers’ cognitive abilities through rational “System 2” thinking (Kahneman, 2011). Teachers will notice additional cues if they review and reflect on school performance feedback data more consciously and systematically (Cook & Gregory, 2019).

Interpreting

Framing

Several theories and models attempt to describe the process of individual sensemaking, including the internal conceptual changes that it entails and the cognitive mechanisms associated with it (Zhang & Soergel, 2016; Zhang & Soergel, 2019). According to the data-frame theory of sensemaking (Klein et al., 2007), everything revolves around the connection of data with cognitive frames or, put differently, categorizing cues or stimuli from data and connecting them with individuals’ pre-existing and internalized cognitive frames.

Intuitively noticing and recognizing specific information and specific cues in data (such as school performance feedback data) may feed into a variety of cognitive activities: elaborating an existing frame based on experience and

advanced understanding, questioning a frame when the selected information is inconsistent with it, maintaining a frame by dismissing anomalous information, comparing alternative frames and searching for connections within the selected information, shaping a new frame if existing frames are insufficient, and searching for a frame by looking in the available information for more cues that might have been previously ignored. According to Attfield and Baber (2017), multiple frames are activated simultaneously when one processes data in order to form a personal account. These cognitive frames contain both general knowledge and more specific and situational knowledge and representations (Attfield et al., 2018).

Calabretta et al. (2017) argued that framing occurs both intuitively and rationally. Intuitive framing entails both the fast and unconscious activation of all cognitive frames related to the cues noticed in the data as well as unconsciously seeking holistic connections between these cognitive frames (Dane & Pratt, 2007). Rational framing, on the other hand, is an explicit, analytical reasoning process that requires more time and involves structuring limited information according to logical, substantive, and personal or context-specific criteria before being able to arrive at judgment (Calabretta et al., 2017).

Judgment

Research by Vanlommel et al. (2018) showed that teachers use data less rationally and objectively than one might expect. They often judge student outcomes intuitively, based on perceptions, personal criteria, and various non-cognitive indicators. Because they rarely seek any other data sources or consider alternative explanations, inferences based on feelings and personal beliefs arise, which can in turn lead to unsound interpretations of student outcomes.

Kahneman and Frederick (2005) argued that intuition and rationality are closely intertwined and influence each other. Dual process models discuss how the fast and intuitive system on the one hand and the more thoughtful and rational system on the other interplay (Whittaker, 2018). Whereas some researchers have postulated that intuition precedes rationality by serving as an input to deliberate and rational thought processes (Salas et al., 2010), Calabretta et al. (2017) emphasized the integration of both systems, considering intuition not subordinate to rationality, but rather complementary. They suggested alternating rational judgment (step-by-step, thoughtful cognitive evaluation) with intuitive judgment (unconscious, rapid, and affectively charged evaluation) and subsequently evaluating the product of this process rationally, and they argued that allowing and even encouraging a balanced integration of both systems would result in effective strategic decision-making. Cook and Gregory (2019) also emphasized the interplay between cognition, emotion, and judgment.

Kahneman and Tversky (1979) argued that judgment is always made in comparison to a status quo. This reference point will thus determine the risk a teacher is willing to take with respect to a possible adjustment of their own instructional practice. We propose that individual judgment of school performance feedback will always be grounded in the broader context of accountability and/or development (i.e., the summative and/or formative purposes for educational testing and the stakes involved), the prevailing school culture, and interrelationships within the team, as well as the specific background of the students and class groups who participated in the tests.

Professional dialogue and its affordances for collective sensemaking

Sensemaking does not only occur internally, but also in a dynamic process of co-construction in which individuals' selective perception and intersubjective interpretation is embedded in the environment through verbal and nonverbal communication (Cecez-Kecmanovic, 2004; Kudesia, 2017). Like individual sensemaking, collective sensemaking occurs both intuitively through the use of frames and heuristics, and rationally and consciously (Avby, 2015; Cook & Gregory, 2019). Both noticing and interpreting information takes place within a particular context in which personal, professional, organizational, and social influences interact. Cognitive, affective, and political aspects all influence which and how much information is shared by whom and with whom (Cook & Gregory, 2019).

The interaction between individual and collective sensemaking is a non-linear and iterative process that can ultimately lead to shared sensemaking, in this case sensemaking of school performance feedback data and, ideally, further improvement of educational quality. In this regard, Tsoukas (2009) and Gergen et al. (2004) emphasized the importance of in-depth professional dialogue. This is a conversation between professionals that is characterized by an open exchange of ideas, assumptions, and experiences and by the explication of tacit knowledge (Cook & Gregory, 2019; Nonaka & Takeuchi, 1995). Collective sensemaking, as in data discussions, broadens the interpretive lens through which data are viewed, stimulates debate between participants, and has the potential to create new knowledge, both on an individual level and on a shared level (Coburn & Turner, 2011; Datnow et al., 2012; Spillane, 2012). Voicing individual interpretations and inferences in data interactions helps to expose assumptions and ambiguities (Bertrand & Marsh, 2015; Christman et al., 2016). Successful and productive professional dialogue tends to benefit from participants' adoption of a non-judgmental, curious attitude and their (readiness to engage in) active listening (Gergen et al., 2004).

Relational Factors as Proposed Predictors

Expertise

Teacher expertise influences in-depth sensemaking of school performance feedback data. The literature emphasizes the expertise of veteran colleagues in this regard. Amidu et al. (2019), Chudnoff (2019), and Sinclair (2010) referred to intuitive expertise in order to elaborate on this. According to these scholars, more experienced colleagues possess more elaborate cognitive structures or frames to assess information quickly and intuitively without prior deliberation or an explicitly rational approach. Experts will discuss a problem of practice based on underlying principles rather than superficial features (Chi et al., 1981). Moreover, their well-developed metacognitive skills allow them to continuously monitor and evaluate their own reasoning and to switch between intuition and rationality without much effort (Amidu et al., 2019).

Based on a comparative study of professional dialogue in two departments, Horn and Little (2010) found that individual knowledge, skills, and experience contribute to the depth of professional dialogue. According to this research, “normalizing, specifying, revising, and generalizing” problems relating to concrete instructional practices of new teachers fosters adaptive expertise within a department. This generic expertise can subsequently be deployed to quickly solve similar problems, through extrapolation and the integration of data literacy with subject-specific, pedagogical-didactic, and instructional expertise (Mandinach & Gummer, 2016; Shleifer, 2012). Thus, in-depth professional dialogue is only possible if teachers and departments have sufficient individual and collective expertise. Expertise serves as both input and output of in-depth professional dialogue.

Inquiry-based attitude

In terms of making sense of (educational) data, the notion of expertise is closely related to that of data literacy: the capacity to identify problems, transform data into actionable knowledge, and evaluate outcomes (Beck & Nunnaley, 2021). However, Krüger (2010) stated that not every teacher needs to be data literate—teachers need to deploy an inquiry habit of mind to be able to use and handle data effectively. Amels et al. (2019) added that data literacy only has a small impact on teacher capacity for change with respect to instructional practices. They propose that inquiry-based working is much more important. Characteristics of an inquiry-based attitude include curiosity, critical (self) reflection, asking questions, willingness to change perspectives without judgment, openness, honesty, willingness to share with others, and a focus on data, accuracy, and thorough understanding (Krüger, 2018; Uiterwijk-Luijk et al., 2017). An inquiry-based and problem-solving school

culture ensures that teachers—in the sensemaking phase—shift from external attributions to the acknowledgement of their own contribution (in interaction with their colleagues) and, in the process, also question their own practice.

Trust

Van Gasse et al. (2017) emphasized the importance of meaningful and authentic data use interactions between teachers. This requires a climate of psychological safety, in which teachers are confident and willing to be vulnerable. They dare to express doubts, raise problems, admit mistakes, and hold each other accountable for errors (Edmondson, 1999). Sufficient trust or confidence that taking interpersonal risks will not have negative (relational) consequences fosters cognitive and emotional commitment and mobilization of expertise (Edmondson & Lei, 2014).

Tsoukas (2009) argued that the productivity of professional dialogue is determined by the degree of relational commitment: co-workers taking collective responsibility for sensemaking of information and for reciprocal relationships within the team. They are open to being influenced and, through self-reflection, consciously scrutinize their usual ways of thinking and acting. After all, the meaning attributed to individual verbal and nonverbal expressions in a professional dialogue depends to a large extent on the response of the recipient and on the prevailing (school) culture. Additionally, thoughts and feelings are influenced by the (perceived) presence of others (Gergen et al., 2004). Roesch-Marsh (2018) also noted that facilitative relationships are indispensable for achieving deep professional dialogue.

Commitment

(Collective) sensemaking is not only a question of capacity (expertise and inquiry-based attitude) and confidence (trust) but also one of commitment: being willing to engage. Committed teachers consciously and voluntarily take individual and collective responsibility for student learning by participating in in-depth professional dialogue (Cameron & Lovett, 2015). They apply their subject-specific and pedagogical-didactic expertise to further develop their instructional practices within their own context (Sammons et al., 2007).

A study by Fransson and Frelin (2016) found that highly committed teachers exhibit a strong sense of professionalism. Challenging situations and complex problems encourage them to search for potential solutions. They feel responsible for their students' well-being and learning as well as for their own professional development. This commitment results in in-depth sensemaking of school performance feedback data and a positive impact on learning gain and student achievement (Day & Gu, 2007).

Methods

In order to gain in-depth insight into the complexity of sensemaking of school performance feedback data from standardized tests, we conducted semi-structured interviews aimed at exploring and describing current sensemaking practices (Alase, 2017). Based on teachers' perceptions and personal experiences, we aimed to better understand how they make sense of school performance feedback data and which factors promote or hinder teachers' sensemaking of school performance feedback data. The following sections discuss the research context and the more technical methodological choices and the approach we set forward to achieve this aim.

Research context

In the absence of central examinations, Flemish secondary schools have few standardized instruments at their disposal to measure student achievement or learning gain. They can, however, voluntarily participate in the Flemish national assessments, if they have been randomly selected to be part of the representative reference sample, or proactively decide to administer freely available parallel tests from these national assessments. The Flemish national assessments and parallel tests cover a range of learning subjects and measure the extent to which attainment targets are met: a set of formal learning objectives formulated by the government for the end of certain grades. The representative sample needed in these national assessments to conduct a valid assessment typically ranges from 10 to 20% of eligible Flemish schools, depending on the research design. On a yearly basis, parallel tests are typically used by (under) 10% of Flemish secondary schools. Overall, Flemish schools do not have a strong tradition of using externally generated output information (Van Gasse et al., 2015).

After participating in a national assessment or administering parallel tests, schools receive a school performance feedback report that gives statistical information about the proportion of students that reached the attainment targets, as well as value-added information based on a comparative analysis with schools that are similar in a number of input and context characteristics. These results are presented in graphical representations with an extensive reading guide, but do not include personalized recommendations. The reports of the parallel tests also include individual student results, displayed by attainment level. School performance feedback reports are strictly confidential and schools may only use the data for internal quality assurance. There are no stakes involved for schools in these tests.

Participants

For the present study, we recruited teachers from secondary schools that had recently (i.e., in the past school year) received school performance feedback after taking parallel tests of their own accord or after having agreed to participate in a national assessment. In the interest of homogeneity, and because the national assessment of mathematics in the second year had recently taken place, we focused on middle school mathematics teachers. No other (stratification) variables were taken into account when selecting participants. We started by recruiting participants in parallel test schools but moved to national assessment schools when responses proved insufficient.

In total, 11 teachers were interviewed, all of whom gave their informed consent to participate in the study. The majority of these teachers (9 out of 11) were female, and participant ages ranged from 24 to 55 years old. Two participants held master's degrees; the rest held bachelor's degrees. Participants had on average 12 years of experience working in education, with one participant being a first-year teacher, and the most veteran participant having 34 years of educational experience.

Interviews and procedure

All semi-structured interviews were conducted between early and mid-March 2020. The teachers were interviewed with the school performance feedback report that they had previously received on hand. A combination of interview questions and a think-aloud section enabled us to study the sensemaking process in depth, both from participants' retrospective accounts and perceptions, as well as from our own observations during the interviews (Eccles & Aarsal, 2017). The interviews were audio recorded and transcribed verbatim.

To enhance construct validity, a pilot interview was conducted with a fictitious school feedback report. To ensure content validity, the interview guide was grounded in the concepts identified in the theoretical framework, with concrete open-ended questions attached to each concept, aimed at gauging participants' thoughts, experiences and perceptions. Examples of such questions include "What was the first thing you noticed when you went through this feedback report?" (intuitive noticing); "How did your department colleagues explain the results, and do you agree with them?" (professional dialogue); and "Do you feel that school results can be freely discussed within your team?" (trust). Additionally, in order to enrich our understanding of how the different steps of the sensemaking process take shape, questions were included regarding participants' sensemaking of the data as it occurred *during* the interviews. The think-aloud method offered a way to explore participants' thought processes.

Analysis

In order to get a general overview of the main research findings and to facilitate coding, the most salient results for each participant were summarized under each group of questions in the interview guide (Alase, 2017). Next, the 11 interviews were coded deductively in NVivo 12, using a coding tree based on the theoretical framework. To ensure the reliability and validity of the coding, the conceptual foundations of the main codes were written out (see Table 1). Some codes were further divided, adding sub-codes afterwards, in order to facilitate the analyses. In addition to the codes as listed in Table 1, there was also a code that focused specifically on the collective aspects of sense making. This code was added to the data if elements of noticing, framing, or judging also involved social interaction or forms of professional dialogue. The distinction between individual and collective sense making (cf. research question 1) was thus brought into the analyses. In order to analyze differences and similarities, the data were analyzed both horizontally and vertically (Cohen et al., 2011; Donche, 2015).

Table 1
(*Non-exhaustive*) *Description of Key Codes Used*

| Codes | Conceptual foundations |
|--------------------|--|
| Sensemaking | |
| Intuitive noticing | The teacher mentions things that immediately struck them in the feedback reports without conscious thought. The teacher mentions cues or things that stood out, which they immediately recognized from prior experiences, personal prior knowledge, beliefs about certain students or class groups, etc. The teacher states that they are not aware of what they noticed immediately during the initial reading of the report. |
| Rational noticing | The teacher mentions things that stand out when they review the feedback reports and think about them intentionally, consciously, and systematically (based on guiding questions). |
| Intuitive framing | The teacher mentions their first impressions of or the ideas they formed about certain graphs or tables at first glance, from intuitive expertise, prior experiences, personal beliefs, knowledge about certain students or class groups, etc. |
| Rational framing | The teacher mentions their impressions of or their ideas about certain graphs or tables after having consciously and systematically thought about them (based on guiding questions), having triangulated them with other data sources, having reflected on them further or having explicitly related them to other parts of the feedback reports through an analytical reasoning process. |

| | |
|---------------------------------|--|
| Intuitive judgment | The teacher mentions their initial judgment, based on prior experiences (with particular students or class groups), feelings, perceptions, assumptions, etc. The teacher states that they are not aware of having formed a judgment during the initial reading of the report. |
| Rational judgment | The teacher mentions additional potential causes for the results presented in the feedback reports, identified through a step-by-step and conscious cognitive evaluation of the data. |
| Influencing factors | |
| Trust | The teacher mentions things that indicate a climate of psychological safety and trust in each other's abilities, such as being open about mistakes, daring to express doubts, raising problems and seeking solutions together regarding instruction and student outcomes, giving each other feedback, etc. |
| Participant expertise | The teacher understands the concepts used in the feedback report and can interpret the graphs and tables correctly. The teacher is aware of misinterpretations and makes adjustments when needed. |
| Expertise within the department | The teacher mentions that there is sufficient expertise within the department and is able to illustrate this with examples. |
| Inquiry-based attitude | The teacher mentions elements that indicate an inquiry-based attitude such as questioning one's own practice, raising questions for reflection, etc. (Distinction was made between "Participant level" and "Department level") |
| Commitment | The teacher mentions things that indicate participation and taking responsibility for student learning based on student outcomes. (Distinction was made between "Participant level" and "Department level") |

Findings

In order to answer the first research question, we describe participants' experiences, approaches, underlying thoughts, and feelings while making sense of school performance feedback data. We consider our participating mathematics teachers' recollections of their initial sensemaking upon reception of the feedback reports, as well as their sensemaking process as it took place when discussing the reports during the interview. We do this successively for the different steps of individual sensemaking and for collective sensemaking.

Individual sensemaking of school performance feedback data

Our findings with regard to individual sensemaking will be presented according to the theoretical distinction we made between noticing, framing, and judging. We discuss the intuitive or rational manifestations of these steps. Table 2 provides an overview of how individual sensemaking took shape for each of the participants.

Table 2
Coding for Initial Individual Sensemaking

| Participant | Noticing | Interpreting | |
|-------------|----------|--------------|---------|
| | | Framing | Judging |
| 1 | I | I | I |
| 2 | C | C | I |
| 3 | I | I | I |
| 4 | C | C | I |
| 5 | I | I | I |
| 6 | C | I | I |
| 7 | I | I | I |
| 8 | I | C | C |
| 9 | I | I | I |
| 10 | I | C | I |
| 11 | I | I | I |

Note: I = intuitive, R = rational, C = combination

Intuitive and rational noticing

The initial reading of the feedback report appears to have been a predominantly intuitive process for all participants. All participants also stated that, in that initial intuitive phase, they overlooked some (important) aspects of the feedback reports. Intuitively, they had particularly remarked how many of their students were not meeting the attainment targets, which for many participants corresponded to their expectations for certain class groups or individual students.

Out of a certain curiosity, you spot the things you want to know. ... I'm pretty sure I overlooked a lot of important things. After all, there's a lot of text to read, a lot of graphs to interpret. You focus on things that immediately stand out, for example, the red color. – Participant 9

Two participants stated that they were not (or were no longer) aware of what they had intuitively noticed when they first went through the feedback reports.

Only three participants indicated that after an initial intuitive reading, they had reread the reports with a more rational, step-by-step approach, comparing students and subject domains. In such cases, we speak of “combined” noticing.

When we asked participants to review the feedback reports during the interviews, three of them stated that they did not notice anything new or additional. Other participants indicated that they had taken a closer look at specific parts of the report because time had now been explicitly made available for it and because they had been explicitly requested to do so.

A number of them stated that they had now looked at the individual student results more thoroughly and step-by-step, thus rationally, in consideration of a potential adjustment of their own instructional practice or to inform discussions of individual students' educational progress at teacher council meetings.

Interpreting: Intuitive and rational framing

Almost all participants indicated that they had not or had only barely questioned their initial intuitive framing of the results. The teachers stated that they knew their students well, that they knew what to expect from their students, and that they based their framing of the results on this knowledge. They mainly focused on their own class groups and intuitively compared their own students' achievement to that of other class groups or study options, to the statistically expected scores, or to the national average. For instance, they said they immediately observed how some attainment targets were achieved by only a limited number of students. Some participants got this information from the tables; others found the visual representations a clearer way to get a general idea of the results from a broader perspective.

We prompted participants to think rationally about their initial sensemaking of the results, and had them reflect on their thought process, approach, and underlying thoughts and on potential questions this had raised for them, possibly also when triangulating the results with information from other sources. Only three participants indicated that they had thought about the feedback after their initial reading and sensemaking. As revisiting the feedback reports and looking for additional data for triangulation at a later stage during interpretation indicate a conscious and systematic process of sensemaking, we labelled this as "rational framing." Only one participant had compared the school performance feedback data with results from their own classroom assessments. This means that only four of the 11 participants engaged in combined (intuitive and rational) framing during the interpretation phase. We labelled two of these participants' noticing process as "combined".

During the interviews, we asked participants to review the feedback reports, in an attempt to encourage purposeful, systematic, and rational thinking. However, for more than half of the participants, this did not yield any additional insights. We also note that, even after having been stimulated to adopt a more rational approach, some participants continued to have difficulties forming an understanding of the data and indicated that they were unable to make any inferences based on the data. A large number of participants had difficulties with correctly interpreting the population scores and correctly comparing them to the achievement of their own students. We note that for at least eight of the participants, their interpretation of the information at hand was severely compromised.

One participant realized during the interview that their own initial framing of the information in the tables contradicted the graphs that essentially conveyed the same information. They tried to correct this by rationally reflecting on their earlier intuitive framing.

Yeah, well, in that case something is off, right [laughs]. ... From the table I would infer that we did not do well and from the graph I would infer that we did. ... Here you are not compared to the rest, but you are compared to the attainment targets and here you are compared to the other participants, right. ... That's how I interpret the difference. I don't know if that's correct [laughs]. – Participant 7

Interpreting: Intuitive and rational judgment

In our conceptual logic, framing is followed by judging. Our analyses of the interview data show that, with one exception, all participants initially only made fast and intuitive judgments about the school performance feedback data and based these judgments on individual perceptions and personal criteria.

Among our students, we have tremendous diversity. Many different home situations and native languages. Our students are also not motivated... Most of the time, they stop processing the subject matter instantly when the bell rings. ... Also, most of them do not show up for refresher classes and re-sit tests. – Participant 9

Two participants indicated that they did not remember how they arrived at judgments during their initial reading of the feedback reports. Other participants' initial judgments appear to have been largely intuitive. Teachers almost exclusively took into account input and context factors: student characteristics on the one hand, more specifically study attitude and motivation, specific educational needs, language proficiency, and mathematical knowledge and skills, and on the other hand the timing of and practicalities associated with test administration. Potential causes for the results that they put forward intuitively included the absence of support for students with special educational needs, the fact that some domains had not or had only just recently been covered at the time of the test and the fact that only final answers were scored instead of taking into account the solution strategy.

During the interviews, we asked the participants to reflect further on their judgments. Several of the initially and intuitively mentioned causes were explored further during the rational judgment phase—but in this case by participants who said they had not yet thought about these elements during their intuitive judgment phase. Additional explanations for performing well or not well were also provided. These included school-specific organizational features such as the availability of tutoring hours for mathematics and external

context factors such as insufficient parental support, students' excessive use of social media, and overloaded curricula. Process factors and curriculum-related factors were also mentioned, such as a lack of classroom management and strong collaboration within the department on the one hand, and on the other hand the difficulty or abstract nature of certain domains, automaticity development in primary education, and the low difficulty level of test questions for certain attainment targets.

Collective sensemaking of school performance feedback data

In this section we discuss how the teachers made sense of the school performance feedback data together with their colleagues. It is striking that all participants, with one exception, indicated that the feedback reports were only distributed to the 7th and 8th grade mathematics teachers (involved in the test administration) and not to other staff members. Moreover, for four of the 11 participants, this only happened in the run-up to this interview. In addition, all participants indicated that the school leader or (internal quality assurance) coordinator did not link any specific instructions or reflection questions to the distribution of the feedback reports, other than that the teachers were expected to go over the reports and discuss them within the department. However, this only occurred in one school and this discussion remained short and superficial in nature. Two participants from another school indicated that during a recent department meeting the “discussion” was limited to a short announcement that the results were poor. A few schools had taken spontaneous initiatives to be able to interpret the results better, such as participating in a workshop or consulting with teachers from other schools to discuss the data together. However, according to these teachers, this had contributed only minimally to a better understanding of the terminology used in the reports and of the tables and graphs. They did not feel it had led to an in-depth professional dialogue or any additional rational sensemaking of the school performance feedback data.

In all other schools, the feedback reports had not been discussed as a team at all. Participants attributed this to factors at the policy level, for instance to the fact that they had been obliged to participate and therefore did not feel the need to discuss the results with their colleagues, a lack of sufficiently explicit expectations from the school leader, changes in the school leadership, and absence of the school leader due to illness. They also referred to factors at the teacher level and to the results themselves. Almost all participants indicated that their school currently had other priorities. The fact that most schools achieve average results, at least in the participants' interpretation, does not encourage collective discussion either. Most participants presumed that there would have been a collective formal or informal discussion if their results had been very disappointing.

In summary, therefore, we can conclude that collective sensemaking did not occur in most schools. In those cases where it did, the participants indicated that there was hardly any interaction and no meaningful exchange of individual interpretations from different perspectives. As a result, no new insights were created, so we cannot speak of true collective sensemaking.

Factors influencing sensemaking processes

In order to answer the second research question, we probed participants' perceptions of factors that influence sensemaking of school performance feedback data. Since the feedback reports had only been discussed during a department meeting in one school, we will look at factors influencing sensemaking of student outcomes and data use in general, without distinguishing between the individual and the collective sensemaking process. Table 3 provides an overview of influencing factors as mentioned by the participants.

Table 3
Coding for Influencing Factors

| Participant | Trust | Expertise | | Inquiry-based attitude | | Commitment | |
|-------------|-------|-------------|------------|------------------------|------------|-------------|------------|
| | | Participant | Department | Participant | Department | Participant | Department |
| 1 | yes | yes | yes | no | no | yes | yes |
| 2 | yes | no | no | no | yes | yes | no |
| 3 | no | / | yes | no | no | no | no |
| 4 | yes | yes | yes | yes | no | yes | no |
| 5 | yes | no | no | no | yes | no | no |
| 6 | yes | yes | yes | / | yes | no | yes |
| 7 | yes | yes | yes | yes | yes | yes | yes |
| 8 | yes | yes | yes | yes | yes | no | / |
| 9 | yes | yes | yes | yes | no | yes | no |
| 10 | yes | / | yes | yes | yes | yes | no |
| 11 | yes | no | no | no | yes | no | no |

Note: / = insufficient data

Trust

With one exception, all participants indicated that individual opinions, thoughts, and feelings regarding student outcomes could be discussed freely, both formally and informally, with other team members at their schools. These participants sensed an atmosphere of openness and trust (in each other's abilities). Teachers felt broadly supported and felt there to be ample and open communication and advice with the aim to improve instructional practice and, consequently, student achievement. Moreover, in several

departments, sensitive issues tended to be discussed personally with the colleague in question, which further strengthened teachers' sense of psychological safety. However, although there appeared to be a strong sense of trust overall, this hardly (if at all) resulted in collective sensemaking of the school performance feedback data.

Expertise

Just over half of the participants appeared to have sufficient personal expertise to interpret the school performance feedback data correctly or to be able to correct intuitive misinterpretations by way of an in-depth, step-by-step, rational approach. Some participants explicitly stated during the interview that they did not understand the concepts or visualizations in the reports. They attributed this to the form and content of the feedback reports and to their own limited statistical literacy, indicating their training did not prepare them for this. Even during the interviews, several participants misinterpreted concepts or visualizations. These misinterpretations caused them to (un)consciously form incorrect inferences, particularly regarding the school level results in the data.

Almost all participants indicated they were quite confident that there was sufficient expertise within their team to interpret the school performance feedback data correctly. They often linked this to formal training and stated that teachers with a master's degree are more familiar with handling data and interpreting statistical analyses. Remarkably, one teacher who held a master's degree indicated that they found the feedback reports difficult to process and that they were convinced that there was insufficient expertise within their department to interpret these data correctly.

I've taken a master's myself and I think it's difficult too so never mind someone who's never dealt with data. Um, give this to four out of five of my colleagues and they won't understand a thing. ... They should make those reports a little simpler. – Participant 11

While a lack of expertise had a negative impact on the accuracy of interpretations, it did not appear to have a fundamental impact on the sensemaking process per se. Even when sufficient expertise is present, (collective) sensemaking remained (extremely) limited and superficial in nature.

Inquiry-based attitude

Five out of the 11 participants indicated that they reflected on the data individually, seeking suggestions on how to improve their instructional practice. However, three of them stated that, much to their regret, they were unable to derive from the data what adjustments or improvements would be necessary, which resulted in their self-reflection remaining only general and

superficial in nature. They specifically felt the lack of a report per test item or an overview of those units that students fail most, which could have informed and fostered their sensemaking process. They would have also appreciated some hands-on tips for how to improve instructional practice and general conclusions for the school level.

I would really add conclusions for the school itself. General conclusions. ... and what they need to work on because I can't find that anywhere. ... So that people can also get to work on something concrete, even if they don't understand the report, so they still have something. – Participant 11

In the one school where the feedback reports had been discussed, one participant indicated that no follow-up questions had been asked and that the data were not looked at from different perspectives. In contrast, their colleague from the same school stated that the data had been reflected upon very briefly during the first discussion of the reports. However, because of the positive interpretation of the results and because the team had other priorities, this had not led to any adjustments of instructional practices. Both participants did indicate spontaneously that they had gained more insight into the school performance feedback data thanks to the reflection questions asked during the interview, because this forced them, so to speak, to adopt an inquiry-based attitude and to question their own practice.

Finally, one participant indicated that, in general, their department lacked an inquiry-based attitude, especially when seeking potential instructional adjustments. This participant added that they personally regularly question their own practice—just not based on the data at hand, since the data were in line with their expectations.

Overall, our findings suggest that teachers' inquiry-based attitudes may be an important predictor of the depth of their individual and collective sensemaking of school performance feedback data. More specifically, we found that the lack of an inquiry-based attitude stopped teachers in their sensemaking efforts when moving from the phase of basic reading of the data to making more complex reflections.

Commitment

Without being prompted, five out of the 11 participants explicitly stated that they had initially only superficially read the feedback reports (because this interview was scheduled), had scanned through the introduction, and had mainly searched for a general trend and their overall position compared to other schools. One of these participants had not read the school feedback report at all and two other participants had not looked at the graphs prior to the interview. Some schools only took the test because the school leadership had required it.

Some people don't care about these kinds of inquiries and are only focused on their teaching. ... And I think that if it was not imposed upon us—that we have to engage in this—that it would fade into the background for everyone. ... This is something that would then slide into my drawer of “I'll do it one day,” but only when all my other work is finished. ... But, really, like studying documents and things like that, that's just where they lose me.
– Participant 5

For all of these participants, this lack of commitment resulted in incomplete, often partially inaccurate, and superficial sensemaking of the school performance feedback data.

Only a few participants indicated that they had individually looked at the reports again afterwards, which had led to corrections of initial sensemaking or additional and more specific interpretations of the data. The participants from the school who had discussed the reports collectively indicated that during the team meeting, questions were asked about the content of the feedback reports and that everyone or nearly everyone participated, which they regarded as a sign of commitment. Four participants, two from the same school, indicated that their department was highly committed anyway when it comes to discussing student outcomes. Yet eight of the 11 participants mentioned to a greater or lesser extent that they felt little personal commitment or involvement regarding the data. They also expected this to be the case among their colleagues in the department. They attributed this, among other things, to the (as they interpreted them) fairly good results, to a lack of commitment in general, and to a lack of commitment regarding educational research in particular. This lack of commitment thus appears to be a fundamental explanatory factor for the lack of deep and accurate (collective) sensemaking.

Conclusion and Discussion

Our aim in this study was to investigate how Flemish teachers make sense of school performance feedback data from standardized tests, to describe the steps they do or do not take in this sensemaking process, and to understand how these steps take shape within the complex context of secondary education. In addition, we wanted to investigate the potential influence of a selected number of explanatory variables on the sensemaking process.

How do individual and collective sensemaking processes take shape?

In general, we can conclude that the mere availability of school performance feedback data from standardized tests does not automatically give rise to sensemaking of these data. Within the Flemish educational system, which

gives schools great autonomy in shaping school policy and internal quality assurance, does not mandate central examinations or other forms of standardized testing in secondary education, and grants educators absolute authority over decisions regarding student educational progress, this is a striking finding. We would expect schools that voluntarily choose to take standardized tests, from a school development perspective, to also intensively make use of this information. However, this study shows that hardly any collective sensemaking occurs. None of the school teams involved in the study engaged in in-depth professional dialogue to make sense of the reported outcomes. There is also food for thought in the observation that almost half of the participating teachers were only sent the feedback reports because they would be interviewed about it. Moreover, in all schools, communication from the school leadership had been limited to a request to discuss the feedback reports, without there being any specific expectations tied to this request and without any initiatives to manage, support, or follow-up the sensemaking process.

When we take a closer look at the different steps of the sensemaking process and consider the relationship between intuition and rationality, our findings are consistent with those of prior studies (Datnow et al., 2012; Vanlommel et al., 2017, 2018, 2019). Teachers rely heavily on their intuition during the sensemaking process. What is more, during initial sensemaking, none of the participants employed a purely rational approach to make sense of the school performance feedback data. Their selective attention was focused on things they recognized and understood. According to Chudnoff (2019) and Sinclair (2010), this is not necessarily a problem, provided that teachers have sufficient intuitive expertise to interpret cues correctly. However, our findings that only one teacher triangulated the data with other sources, that the participants hardly reflected on the test results, and that concepts and visualizations were often misunderstood or misinterpreted, called this idea of intuitive expertise into question. In the course of this study, it also turned out that teachers' first impressions often needed to be adjusted and, in some cases, were downright inaccurate compared to what could be objectively determined on the basis of the feedback report. Surprisingly, most teachers were not even aware of this, and it ultimately led to inaccurate or questionable framing. Furthermore, at no point during the judgment phase was any reference made to other data sources in order to support initial and often also intuitive judgments, and attributions for the schools' results were exclusively external.

Since all teachers indicated that they had only superficially read the feedback reports and had looked for general trends and things they intuitively recognized, we may wonder whether their selective attention was unconsciously steered by specific assumptions and feelings about students and class groups. This would be in line with earlier empirical findings in research on sensemaking

in educational settings (Bertrand & Marsh, 2015; Vanlommel et al., 2017, 2019) and with research stating that intuition is not subordinate but complementary to rationality (Calabretta et al., 2017). We can probably also speak of confirmation bias here (Kahneman & Klein, 2009; Vanlommel et al., 2017).

However, in line with Mandinach & Gummer (2016), we argue that sensemaking is an essential part of purposeful and efficient data-driven decision-making processes. If rational data use is indeed this scarce during all the steps taken in initial individual sensemaking, and if individual intuitive expertise is not used to achieve a dynamic interaction process of co-creation and collective sensemaking, the resulting decision-making processes and actions will not adequately address existing gaps and will not have the intended or desired effect on the schools' internal quality assurance (Schildkamp, 2019) and decisions about individual students' learning processes (Vanlommel et al., 2017).

Which factors influence sensemaking?

To answer the second research question, we investigated the impact of trust, expertise, inquiry-based attitude, and commitment on sensemaking of school performance feedback data from standardized tests. We have no indication that a lack of trust is related to the lack of collective sensemaking in schools with respect to school performance feedback data. With one exception, all teachers sensed an atmosphere of openness and trust in their schools and felt their school culture was based on collaboration and knowledge sharing in a safe climate. Likewise, according to our findings, teacher expertise does not appear to have a fundamental impact on the occurrence of sensemaking. In every secondary school, at least some team members will have the necessary and appropriate expertise to correctly interpret the data, which is sufficient according to Krüger (2010). Our findings also show that even in schools where sufficient expertise was present, this had little impact on the depth and rational nature of sensemaking or on the time spent on sensemaking.

A lack of inquiry-based attitude and particularly a lack of commitment within school teams do, however, appear to impede in-depth individual and collective sensemaking of school performance feedback data. A lack of commitment can be explained to a certain extent by teachers' lack of interest in (participating in) research in general. Our study suggests that another and even more important explanatory factor is teachers' feeling that the data presented in the reports are not relevant to their day-to-day instructional practice. Consequently, most teachers did not feel the need to adopt an inquiry-based attitude, and sensemaking remained limited to a general, superficial, and largely intuitive interpretation of the data. A lack of guidance and clear expectations from the school leadership also appeared to contribute to the lack of commitment we found. Since the school performance feedback

is to be used for internal quality assurance, and since mathematics instruction is characterized by a cyclic approach and iterations throughout the different grades, we would have expected that the feedback reports would at least have been distributed to all teachers in the mathematics department. However, this was not the case, leaving teachers with insufficient information to engage in in-depth professional dialogue as a team. These findings are consistent with previous research regarding the link between relevance and instrumental data use among teachers on the one hand and the crucial role of the school leader on the other (Jimerson, 2014; Van Gasse et al., 2015). Finally, lack of commitment to adopting an inquiry-based attitude may explain some teachers' misunderstandings and misinterpretations. We found that teachers who had barely read the introductory sections of the feedback reports, for instance, misinterpreted central concepts and visualizations.

Discussion

In order to increase our understanding of and insight into the phase of sensemaking within the cycle of data use, we unraveled processes of sensemaking of school performance feedback data by secondary school teachers and integrated notions of intuition and rationality into the different steps they undertook in this process. In doing so, we addressed a knowledge gap, since sensemaking of school performance feedback data had not been previously studied in-depth (Schildkamp, 2019). Thus, one of the main scientific contributions of this study is the way we interpreted and conceptualized teachers' intuitive and rational sensemaking processes when making sense of data from standardized testing.

The distinctions we make between intuitive and rational processes pertain to timing, pace, depth, thought processes, and options. These are always dependent on the context in which the teacher operates (Coburn & Tunner, 2011; Abrams et al., 2020). When the teachers in our study intuitively made sense of school performance feedback data, this happened immediately during the initial reading of the feedback reports. Teachers quickly and superficially make sense of the data by looking only at the big picture, at what they recognize, expect, and understand. They are unaware of alternative explanations, and unequivocally choose one particular interpretation of the results. In contrast, rational sensemaking happens over time, when teachers revisit the feedback reports and think about them more deeply, possibly led by guiding questions. This thought process is slower and deliberate. By way of analytical and step-by-step reasoning, they broaden and deepen the sense they previously made of the data and consciously make a choice from among several alternative interpretations.

From a methodological point of view, this study was also an attempt to investigate an unexplored area in the literature. Since there was no ready-made theoretical framework available and we wanted to gain in-depth insight into sensemaking processes of school performance feedback data from standardized tests, we focused specifically on aspects of intuition and rationality throughout the different steps of the sensemaking process. Brock (2015) as well as Dane and Pratt (2007) argued that measuring intuition and even describing what happens during intuitive processes is complex because these processes are fast and often unconscious. We addressed this issue by having participants go through the different steps identified in the theoretical framework via the think-aloud method during the interviews. Nevertheless, our findings could be further enriched with observations. By observing teachers and departments and by discussing those observations with individual teachers, the validity of our findings regarding sensemaking of school performance feedback data from standardized tests can be further enhanced. Observations offer the opportunity to map the iterative and non-linear nature of the different steps, as well as the way intuition and rationality are intertwined, as described in research by Kahneman and Frederick (2005) and suggested by our own findings. In a later phase, quantitative research can also be carried out on a large scale in order to investigate the generalizability of our findings. In addition, we propose that future research should investigate how the different steps of the sensemaking process influence each other.

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, 69*, 1–13 . <https://doi.org/10.1016/j.stueduc.2020.100868>
- Alase, A. (2017). The Interpretative Phenomenological Analysis (IPA): A guide to a good qualitative research approach. *International Journal of Education and Literacy Studies, 5*(2), 9–19. <https://doi.org/10.7575/aiac.ijels.v.5n.2p.9>
- Amels, J., Krüger, M., Suhre, C., & van Veen, K. (2019). Impact of inquiry-based working on the capacity to change in primary education. *Journal of Educational Change, 20*(3), 351–374. <https://doi.org/10.1007/s10833-019-09337-3>
- Amidu, A. R., Boyd, D., & Gobet, F. (2019). A study of the interplay between intuition and rationality in valuation decision making. *Journal of Property Research, 36*(4), 387–418. <https://doi.org/10.1080/09599916.2019.1687572>
- Attfield, S., & Baber, C. (2017, June 20). *Elaborating the frames of data-frame theory*. 13th International Conference on Naturalistic Decision Making, Bath, United Kingdom.
- Attfield, S., Fields, B., & Baber, C. (2018). A resources model for distributed sensemaking. *Cognition, Technology and Work, 20*(4), 651–664. <https://doi.org/10.1007/s10111-018-0529-4>

- Avby, G. (2015). Professional practice as processes of muddling through: A study of learning and sense making in social work. *Vocations and Learning*, 8(1), 95–113. <https://doi.org/10.1007/s12186-014-9123-y>
- Beck, J. S., & Nunnaley, D. (2021). A continuum of data literacy for teaching. *Studies in Educational Evaluation*, 69, 1–8. <https://doi.org/10.1016/j.stueduc.2020.100871>
- Bertrand, M., & Marsh, J. A. (2015). Teachers' sensemaking of data and implications for equity. *American Educational Research Journal*, 52(5), 861–893. <https://doi.org/10.3102/0002831215599251>
- Bolhuis, E., Schildkamp, K., & Voogt, J. (2016). Data-based decision making in teams: Enablers and barriers. *Educational Research and Evaluation*, 22(3–4), 213–233. <https://doi.org/10.1080/13803611.2016.1247728>
- Brock, R. (2015). Intuition and insight: Two concepts that illuminate the tacit in science education. *Studies in Science Education*, 51(2), 127–167. <https://doi.org/10.1080/03057267.2015.1049843>
- Calabretta, G., Gemser, G., & Wijnberg, N. M. (2017). The interplay between intuition and rationality in strategic decision making: A paradox perspective. *Organization Studies*, 38(3–4), 365–401. <https://doi.org/10.1177/0170840616655483>
- Cameron, M., & Lovett, S. (2015). Sustaining the commitment and realising the potential of highly promising teachers. *Teachers and Teaching: Theory and Practice*, 21(2), 150–163. <https://doi.org/10.1080/13540602.2014.928132>
- Cecez-Kecmanovic, D. (2004). A sensemaking model of knowledge in organisations: A way of understanding knowledge management and the role of information technologies. *Knowledge Management Research & Practice*, 2(3), 155–168. <https://doi.org/10.1057/palgrave.kmrp.8500040>
- Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5(2), 121–152. https://doi.org/10.1207/s15516709cog0502_2
- Christman, J. B., Ebby, C., & Edmunds, K. (2016). Data use practices for improved mathematics teaching and learning: The importance of productive dissonance and recurring feedback cycles. *Teachers College Record*, 118(11), 1–32.
- Chudnoff, E. (2019). In search of intuition. *Australasian Journal of Philosophy*, 98(3), 465–480. <https://doi.org/10.1080/00048402.2019.1658121>
- 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>
- Cohen, L., Martin, S., McCulloch, G., O'Sullivan, C., Manion, L., Morrison, K., & Bell, R. (2011). *Research methods in education* (7th ed.). Routledge.
- Cook, L., & Gregory, M. (2019). Making sense of sensemaking: Conceptualising how child and family social workers process assessment information. *Child Care in Practice*, 26(2), 182–195. <https://doi.org/10.1080/13575279.2019.1685458>
- Dane, E., & Pratt, M. G. (2007). Exploring intuition and its role in managerial decision making. *Academy of Management Review*, 32(1), 33–54. <https://doi.org/10.5465/AMR.2007.23463682>
- 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>
- Datnow, A., Park, V., & Kennedy-Lewis, B. (2012). High school teachers' use of data to inform instruction. *Journal of Education for Students Placed at Risk*, 17(4), 247–265. <https://doi.org/10.1080/10824669.2012.718944>

- Day, C., & Gu, Q. (2007). Variations in the conditions for teachers' professional learning and development: Sustaining commitment and effectiveness over a career. *Oxford Review of Education*, 33(4), 423–443. <https://doi.org/10.1080/03054980701450746>
- Donche, V. (2015). *Inleiding in methoden en technieken. Een open leerpakket [Introduction to methods and techniques. An open learning package]*. University of Antwerp.
- Eccles, D. W., & Arsal, G. (2017). The think aloud method: What is it and how do I use it? *Qualitative Research in Sport, Exercise and Health*, 9(4), 514–531. <https://doi.org/10.1080/2159676X.2017.1331501>
- Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, 44(2), 350–383. <https://doi.org/10.2307/2666999>
- Edmondson, A. C., & Lei, Z. (2014). Psychological safety: The history, renaissance, and future of an interpersonal construct. *Annual Review of Organizational Psychology and Organizational Behavior*, 1(1), 23–43. <https://doi.org/10.1146/annurev-orgpsych-031413-091305>
- Fransson, G., & Frelin, A. (2016). Highly committed teachers: What makes them tick? A study of sustained commitment. *Teachers and Teaching: Theory and Practice*, 22(8), 896–912. <https://doi.org/10.1080/13540602.2016.1201469>
- Gephart, R. P., Topal, C., & Zhang, Z. (2010). Future-oriented sensemaking: Temporalities and institutional legitimation. In T. Hernes & S. Maitlis (Eds.), *Process, sensemaking and organizing* (pp. 275–312). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199594566.003.0013>
- Gergen, K. J., Gergen, M. M., & Barrett, F. J. (2004). Dialogue: Life and death of the organization. In D. Grant, C. Oswick, & L. Putnam (Eds.), *The Sage handbook of organizational discourse* (pp. 39–60). Sage. <https://doi.org/10.4135/9781848608122.n2>
- Hill, R. C., & Levenhagen, M. (1995). Metaphors and mental models: Sensemaking and sensegiving in innovative and entrepreneurial activities. *Journal of Management*, 21(6), 1057–1074. <https://doi.org/10.1177/014920639502100603>
- Horn, I. S., & Little, J. W. (2010). Attending to problems of practice: Routines and resources for professional learning in teachers' workplace interactions. *American Educational Research Journal*, 47(1), 181–217. <https://doi.org/10.3102/0002831209345158>
- 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>
- Kahneman, D. (2011). *Thinking, fast and slow*. Business Contact.
- Kahneman, D., & Frederick, S. (2005). A model of heuristic judgment. In K. Holyoak & B. Morrison (Eds.), *Cambridge handbook of thinking and reasoning* (pp. 267–293). Cambridge University Press.
- Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: A failure to disagree. *American Psychologist*, 64(6), 515–526. <https://doi.org/10.1037/a0016755>
- Kahneman, D., & Tversky, A. (1979). Prospect Theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–291.
- Klein, G., Phillips, J. K., Rall, E. L., & Peluso, D. A. (2007). A data-frame theory of sensemaking. In *Expertise out of context: Proceedings of the sixth international conference on naturalistic decision making* (pp. 113–155). Lawrence Erlbaum Associates.
- Krüger, M. L. (2018). *Onderzoeksmatig leidinggeven: Handreiking voor schoolleiders en bestuurders [Inquiry-based leadership: Guide for school leaders and administrators]*. ACCO.

- Krüger, M. L. (2010). *De invloed van schoolleiderschap op het onderzoeksmatig handelen van leraren in veranderingsprocessen* [The impact of school leadership on teachers' inquiry-based work in change processes]. Hogeschool van Amsterdam.
- Kudesa, R. S. (2017). Organizational sensemaking. In *Oxford Research Encyclopedia of Psychology* (pp. 1–42). Oxford University Press. <https://doi.org/10.1093/acrefore/9780190236557.013.78>
- Lai, M. K., Wilson, A., McNaughton, S., & Hsiao, S. (2014). Improving achievement in secondary schools: Impact of a literacy project on reading comprehension and secondary school qualifications. *Reading Research Quarterly, 49*(3), 305–334. <https://doi.org/10.1002/rrq.73>
- Maitlis, S. (2005). The social processes of organizational sensemaking. *Academy of Management Journal, 48*(1), 21–49. <https://doi.org/10.5465/AMJ.2005.15993111>
- Maitlis, S., & Christianson, M. (2014). Sensemaking in organizations: Taking stock and moving forward. *The Academy of Management Annals, 8*(1), 57–125. <https://doi.org/10.1080/19416520.2014.873177>
- Maitlis, S., Vogus, T. J., & Lawrence, T. B. (2013). Sensemaking and emotion in organizations. *Organizational Psychology Review, 3*(3), 222–247. <https://doi.org/10.1177/2041386613489062>
- 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., & Schildkamp, K. (2021). Misconceptions about data-based decision making in education: An exploration of the literature. *Studies in Educational Evaluation, 69*, 1–10. <https://doi.org/10.1016/j.stueduc.2020.100842>
- Mausethagen, S., Proitz, T. S., & Skedsmo, G. (2019). School leadership in data use practices: Collegial and consensus-oriented. *Educational Research, 61*(1), 70–86. <https://doi.org/10.1080/00131881.2018.1561201>
- Nonaka, I., & Takeuchi, H. (1995). *The knowledge-creating company*. Oxford University Press.
- 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>
- Roesch-Marsh, A. (2018). Professional relationships and decision making in social work: Lessons from a Scottish case study of secure accommodation decision making. *Qualitative Social Work, 17*(3), 405–422. <https://doi.org/10.1177/1473325016680285>
- Salas, E., Rosen, M. A., & DiazGranados, D. (2010). Expertise-based intuition and decision making in organizations. *Journal of Management, 36*(4), 941–973. <https://doi.org/10.1177/0149206309350084>
- Sammons, P., Day, C., Kington, A., Gu, Q., Stobart, G., & Smees, R. (2007). Exploring variations in teachers' work, lives and their effects on pupils: Key findings and implications from a longitudinal mixed-method study. *British Educational Research Journal, 33*(5), 681–701. <https://doi.org/10.1080/01411920701582264>
- 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., Lai, M. K., & Earl, L. (Eds.) (2013). *Data-based decision making in education. Challenges and Opportunities*. Springer.
- Schildkamp, K., Smit, M., & Blossing, U. (2019). Professional development in the use of data: From data to knowledge in data teams. *Scandinavian Journal of Educational Research, 63*(3), 393–411. <https://doi.org/10.1080/00313831.2017.1376350>

- 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>
- Shleifer, A. (2012). Psychologists at the gate: A review of Daniel Kahneman's Thinking, Fast and Slow. *Journal of Economic Literature, 50*(4), 1080–91. <http://doi.org/10.1257/jel.50.4.1080>
- Sinclair, M. (2010). Misconceptions about intuition. *Psychological Inquiry, 21*(4), 378–386. <https://doi.org/10.1080/1047840X.2010.523874>
- Spillane, J. P. (2012). Data in practice: Conceptualizing the data-based decision-making phenomena. *American Journal of Education, 118*(2), 113–141. <https://doi.org/10.1086/663283>
- Spillane, J. P., & Miele, D. B. (2007). Evidence in practice: A framing of the terrain. *Teachers College Record, 109*(13), 46–73.
- Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research, 72*(3), 387–431. <https://doi.org/10.3102/00346543072003387>
- Tsoukas, H. (2009). A dialogical approach to the creation of new knowledge in organizations. *Organization Science, 20*(6), 941–957. <https://doi.org/10.1287/orsc.1090.0435>
- 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 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., Vanhoof, J., & de Vos, W. (2014). Informatiegebruik in verantwoordingsgerichte en schoolontwikkelingsgerichte onderwijssystemen: Een Nederlands-Vlaams perspectief [Information use in accountability oriented and school development oriented education systems: A Dutch-Flemish perspective]. *Pedagogiek, 34*(2), 84–106.
- 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., 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., & Schildkamp, K. (2019). How do teachers make sense of data in the context of high-stakes decision making? *American Educational Research Journal, 56*(3), 792–821. <https://doi.org/10.3102/0002831218803891>
- Vanlommel, K., Van Gasse, R., Vanhoof, J., & Van Petegem, P. (2018). Teachers' high-stakes decision making. How teaching approaches affect rational and intuitive data collection. *Teaching and Teacher Education, 71*, 108–119. <https://doi.org/10.1016/j.tate.2017.12.011>
- 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*, 75–83. <https://doi.org/10.1016/j.ijer.2017.02.013>
- 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., & 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>

- Weick, K. E. (1995). *Sensemaking in organizations*. Sage Publications.
- Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (2005). Organizing and the process of sensemaking. *Organization Science*, 16(4), 409–421. <https://doi.org/10.1287/orsc.1050.0133>
- Whittaker, A. (2018). How do child-protection practitioners make decisions in real-life situations? Lessons from the psychology of decision-making. *British Journal of Social Work*, 48(7), 1967–1984.
- Zhang, P., & Soergel, D. (2016). Process patterns and conceptual changes in knowledge representations during information seeking and sensemaking: A qualitative user study. *Journal of Information Science*, 42(1), 59–78. <https://doi.org/10.1177/0165551515615834>
- Zhang, P., & Soergel, D. (2019). Cognitive mechanisms in sensemaking: A qualitative user study. *Journal of the Association for Information Science and Technology*, 71(2), 158–171. <https://doi.org/10.1002/asi.24221>

Corresponding authors

Gila Gutwirth

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

E-mail: g.gutwirth@yahoo.be

Evelyn Goffin

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

E-mail: evelyn.goffin@uantwerpen.be

Jan Vanhoof

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

E-mail: jan.vanhoof@uantwerpen.be

