Vanlommel, Kristin; Pepermans, Elke

Validation of the Teacher Decision-Making Inventory (TDMI): measuring data-based and intuitive dimensions in teachers' decision Process

VALIDATION OF THE TEACHER DECISION-MAKING INVENTORY (TDMI): MEASURING DATA-BASED AND INTUITIVE DIMENSIONS IN TEACHERS' DECISION PROCESS

KRISTIN VANLOMMEL,
ELKE PEPERMANS

Abstract

Teacher decision making has a great impact on the quality of education in schools, yet we know little about how teachers make decisions in practice. It is assumed that teachers use both intuition and data in the different steps of the decision process. No reliable, valid scales are available to research both dimensions during the different steps of teachers' decision process (problem definition, data collection, sense making, and evaluation of alternatives). Building on the integrated framework we constructed in earlier research, the main aim of this study was to develop and validate a Teacher Decision-Making Inventory (TDMI). One hundred and one teachers in adult education participated voluntarily in a web-based survey. Based on the good EFA factor loadings, the CFA fit indices, and the internal consistency (Cronbach's alpha), we conclude that the TDMI is a valid psychometric tool that can be used to assess the intuitive and data-driven dimensions of teachers' decisions in large-scale quantitative research.

Keywords

teachers' decision-making, evidence-informed, intuition, data, validation, professional judgment, dual process

Introduction

Decision making is an important topic in education, since teachers' decisions greatly influence pupils' trajectories, especially when the stakes are high (e.g., passing or failing, moving on to the next educational track; Bonvin, 2003). Therefore, it is important for teachers to make wise, professional decisions. However, we know little about how teachers actually make decisions (Earl & Katz, 2006; Eurydice, 2011; Harteis et al., 2008). This leads to important questions: what are professional decisions, and how can we better understand decision making in practice?

In education, research that studies teacher judgment has shifted from a personal knowledge perspective based on expertise within the teaching profession towards an emphasis on data-based decision making (DBDM). Following disappointing findings with regard to the accuracy of teacher judgment (Urhahne & Wijnia, 2020), policymakers and researchers expected educational decision making to become more data informed (Mandinach, 2006). Data use models describe optimal teacher judgment as based on a systematic inquiry cycle: problem definition, data collection, analyses, and interpretation to evaluate alternatives before a decision is made (Datnow et al., 2007; Schildkamp & Lai, 2013; Strayhorn, 2009).

In the broader field of decision theory, many scholars have agreed that human judgment is guided by both data and intuition, which may influence the different steps of the decision process to a greater or lesser extent (Blackwell et al., 2006; Evans, 2008; Kahneman & Frederick, 2002). Although educational research mostly studies teacher judgment from either a data use or a teacher knowledge perspective, it seems appropriate to assume that both dimensions will influence teacher judgment in practice (Evans, 2008; Klein, 2008; Tversky & Kahneman, 1981). The question is not whether teachers make intuitive or data-based decisions. It is more interesting to grasp the extent to which teachers use data or intuition in the different steps of the decision process.

In the past, studies have indicated that teacher judgment shows much variability at the level of the individual teacher (Kaiser et al., 2013). In earlier research (Vanlommel et al., 2017, 2018, 2020), we therefore used a qualitative case study design to explore how teachers differ in the extent to which they use data or intuition before the final decision is made. Based on the level of data and intuition use by teachers, we identified four different approaches to decision making: (a) rational (high on data, low on intuition); (b) intuitive (high on intuition, low on data); (c) professional (a combination of both); and (d) arbitrary (a restricted decision process involving little use of data or intuition).
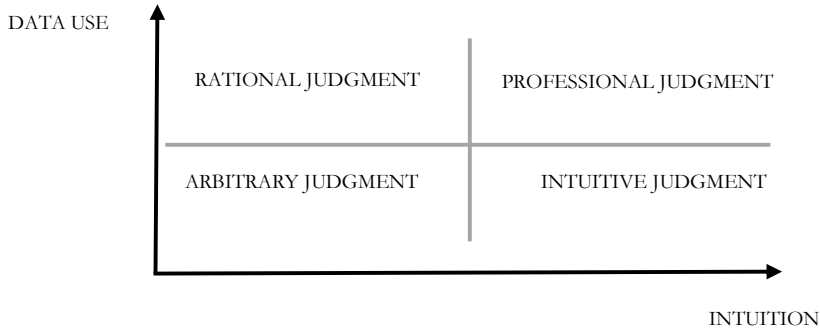
DATA USE

RATIONAL JUDGMENT	PROFESSIONAL JUDGMENT
ARBITRARY JUDGMENT	INTUITIVE JUDGMENT

INTUITION

Figure 1   (Vanlommel, 2018)
Teachers' approaches to decision making

When we wanted to take the next step, to quantitatively study the use of data and intuition in the decision process, we encountered two major problems. First, there was no framework available that integrated data and intuition in the different steps of the decision process, and thus no instruments to study teacher decision making on a large scale. Second, there was conceptual haziness about intuition in the context of teacher judgment. We needed a transparent definition that disentangled the confusion stemming from a lack of insight and that permitted empirical research on this topic in education. We tackled both obstacles in previous research (see Vanlommel, 2018) and we build on those insights to develop and validate a Teacher Decision-Making Inventory (TDMI) that allows large-scale research in order to advance the field.

Theoretical Framework

The decision process

Both data-based and intuitive processes are considered to be valuable parts of teacher judgment that each have their own merits and pitfalls; thus, they need to be combined to make the best professional decisions possible. For example, intuitive recognition can allow teachers to recognize a problem quickly at an early stage, and expert knowledge is important to understand what data mean in a specific context. At the same time, research has shown that intuition may be vulnerable to different sources of bias (Burgess et al., 2009; Kahneman & Frederick, 2002). For one thing, teachers may mainly pay attention to indicators that confirm what they already believe and often ignore data that indicate the contrary (Goldstein & Hogarth, 1997; Harteis

et al., 2008; Klein, 2008; Tversky & Kahneman, 1981). This may lead to self-fulfilling prophecies that perpetuate social or economic disparities. In order to prevent the pitfalls of confirmation bias, data use is crucial for questioning assumptions. To make good decisions, teachers are required to use both data and intuition in the different steps of the decision process. Information deriving from one source can complement information from another (Earl & Katz, 2006; Kahneman & Klein, 2009); the complexity of conclusions related to pupil competence also requires a detailed and balanced view drawn from more than one data source (Cohen et al., 2017.)

Professional decision making from an integrated perspective

In the field of decision making, theories on dual-process approaches to decision making indicate that data-based and intuitive processes both influence human judgment (Hogarth, 2001; Klein, 2008). In the field of education, theories that approach decision making as a dual process influenced by the use of both data and intuition are scarce. In earlier research (Vanlommel et al., 2017), we combined theories of DBDM that are commonly used within education (Datnow & Hubbard, 2015; Mandinach & Jimerson, 2016; Schildkamp et al., 2016) with the theory of naturalistic decision making that studies intuition as expertise (Klein, 2008). The recognition-primed decision model describes how experts can use their professional knowledge of subject and context to make accurate decisions, based on their expertise (Klein, 2008). We will elaborate on both dimensions in the next paragraphs and integrate them in the theoretical framework that will be used to develop our questionnaire.

What is DBDM and how can it contribute to teacher judgment?

In a movement away from the era in which research primarily studied how teachers' intuitive knowledge influenced the outcomes of teacher judgment, the initial body of data use research mainly conceptualized data as quantitative indicators of pupils' cognitive output (Hubbard et al., 2014). This was based on the assumption that the quality of educational decisions would increase to the extent that they were based on objective measures, such as standardized tests.

More recently, scholars have criticized this narrow view because it inhibits a full understanding of pupil competences and it has led to undesirable practices such as 'teaching to the test' (Brown, 2017; Ehren & Swanborn, 2012). Therefore, broadening the concept of data to include all indicators that inform some aspect of schooling has been advocated (Schildkamp & Lai, 2013). These definitions of data include quantitative measures, such as results from (standardized) tests or attendance rates, but also qualitative

indicators, such as observations in the classroom or conversations with colleagues, pupils, or parents. To differentiate the formal use of data from incidental, spontaneous gathering of indicators, data collection needs to be initiated based on a clear goal or question and to follow an inquiry cycle (Earl & Louis, 2013; Schildkamp & Lai, 2013). Data are collected systematically and deliberately (Bromme et al., 2014).

DBDM can then be defined as a systematic process in which (1) a problem or question is diagnosed using at least one type of data collected deliberately and systematically, (2) data are collected systematically with the aim of exploring the question or problem, (3) data are interpreted by objective criteria, and (4) evaluative arguments are based on data use in steps 1–3 (Coburn & Turner, 2012).

What is intuition and how can it contribute to professional teacher judgment?

Theories of naturalistic decision making focus on the value of expert intuition, originating from early research on master chess players who were able to make accurate decisions because they recognized cues and complex patterns (Chase & Simon, 1973). This led to the definition of intuition as recognition, and was elaborated further in the recognition-primed decision model (Klein, 2008). Klein (2008) described how subject-matter experts are able to make good decisions in complex contexts because they recognize cues and patterns based on the expert knowledge stored in their memory, without a deliberate and systematic search. Applied to teacher judgment, this means that teachers are able to recognize a problem spontaneously, without using data in their diagnosis, and to make a decision without a deliberate and systematic collection and analysis of data (Kahneman, 2003; Klein, 2008). Teachers hold patterns in their memory, based on learning and experience, that draw attention to cues without a deliberate search for answers to a question. The spontaneous recognition of elements in a given situation triggers expectancies for the future based on similar cases in the past, and thereby informs decision making without deliberate analyses. In our study, we define intuition as a personal knowledge base that consists of patterns and mental models teachers have acquired through learning and experience, enabling them to recognize cues and solutions spontaneously without deliberate attention or a systematic approach.

Intuitive processes of decision making refer to (1) spontaneous recognition of a problem without further diagnosis, (2) automatic collection of information without a deliberate or systematic approach, (3) interpretation based on personal criteria, and (4) evaluative arguments based on evidence collected through intuitive processes in steps 1-3.

*To an integrated framework: Data use and intuition integrated
in the different steps of the decision process*

In step 1, a problem or goal is defined when the actual situation is weighed against personal or objective standards for the desired situation (Mintzberg & Westley, 2001; Schildkamp et al., 2016).

A decision process may be initiated when a teacher recognizes a problem spontaneously without deliberately weighing the actual state of affairs against the standards. For example, a teacher spontaneously notices that a pupil is staring out the window during daily work. The teacher feels this might be a problem and keeps this information in mind. This intuitive problem recognition might or might not be followed by problem diagnosis. For example, the teacher can start observing this pupil using an observation protocol, focusing on pre-planned indicators.

Once teachers have defined the problem, this is expected to trigger a wider search for more data (Evans, 2008; Schildkamp et al., 2016). In step 2 (data collection), a data search may or may not be guided by the problem or question defined in step 1 or by a clear plan (e.g., Mandinach et al., 2006). For example, when the teacher defines the problem as a student's possible learning disorder in step 1, that teacher can develop a plan: what data do I need and how do I collect the data in order to gain fine-grained insight into the problem? Intuitive data collection might start from the same problem definition but is not guided by a plan. During teachers' daily practice, their attention is spontaneously drawn by elements that (mostly) confirm or (seldom) question their initial problem recognition. Independent of the rational or intuitive nature of teachers' data collection, in step 3, data need to be analyzed and interpreted before they can inform teachers' decision making (Bertrand & Marsh, 2015). In this sense-making process, it has been suggested that although data use models to prescribe optimal procedures for coming to valid conclusions (Bosker et al., 2007), in practice teachers might take mental shortcuts (heuristics) to reach quicker and easier conclusions (Evans, 2006; Kahneman, 2008; Klein, 2008). False inferences are often explained in terms of confirmation bias, when teachers frame the data to fit their existing beliefs (Harteis et al., 2008; Kahneman & Frederick, 2002). Therefore, it is important to look at the criteria used when teachers make inferences. While DBDM refers to the use of objective, pre-defined criteria, teachers might also trust in personal criteria to make sense of data (Vanlommel & Schildkamp, 2019).

In the fourth step, after teachers have run through steps 1-3, an important question concerns the extent to which they take data and intuition into account when they evaluate alternatives and make a decision. Information deriving from data-based and intuitive processes may coincide and thus strengthen

teacher judgment, or it may provide contrasting viewpoints. In that case, an important question to investigate is how teachers use data and intuition to reach their final decision.

Even decision processes that are predominantly led by data use processes may result in intuitive judgment when information deriving from one intuitive cue overrules all other evidence. Research has shown that the decisive criteria applied by teachers are often based on subjective beliefs about good learning and teaching (Allal, 2013; Rubie-Davies, 2010; Zanting et al., 2001). For example, despite test results, reports, or conversations with colleagues, a teacher may rely on their personal trust or distrust in the student's motivation.

We approach professional decision making as the combination of both dimensions in the different steps of the decision process. Figure 2 provides a static visual overview of what is, in practice, a complex, iterative process.
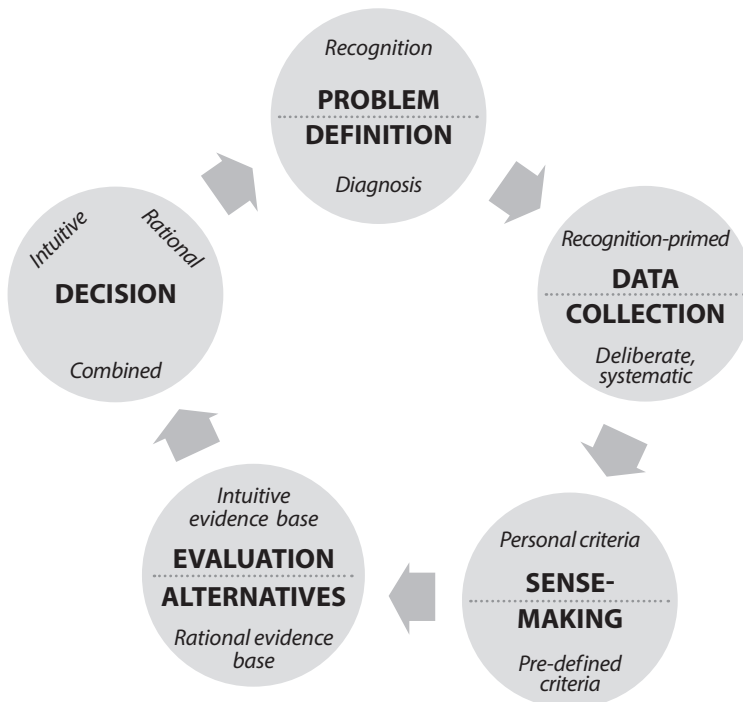

Figure 2
Theoretical model for the Teacher Decision-Making Inventory, based on Vanlommel et al. (2018)

Aim

Given the integrated framework we developed and tested in earlier research (Vanlommel et al., 2020), the main aim of this study is to develop and validate a questionnaire based on that theoretical framework that supports investigation of teachers' decision making in practice on a larger scale. This will contribute to our understanding of teacher decision making in practice and help strengthen the knowledge base on decision making in education.

Method

For the development and validation of the instrument, we used the five steps described by Hinkin (1998): (1) items were constructed based on a theoretical model, (2) the survey was administered to the target group, (3) the number of items was reduced by means of exploratory factor analysis (EFA), (4) the structure was confirmed by means of a confirmatory factor analyses (CFA), (5) the internal consistency of the final scales was measured.

Instrument

In this section we describe the step-by-step process of constructing and validating the instrument as suggested by Benson and Clark (1982). As a first step in phase 1 (planning), we formulated the purpose: we wanted to construct an instrument that allowed us to investigate decision-making in practice, starting from a dual process approach. Phase 1 was largely conducted in our previous research (Vanlommel, 2018) as it consisted of (a) a broad literature review and (b) semi-structured interviews. The literature review showed that no existing instruments were readily available and offered us guiding frameworks to study decision making (see Theoretical Framework). The semi-structured interviews offered us a rich qualitative insight in on how educators make decisions in practice. Both the literature review and interview results were used as input for phase 2 (Construction). In this phase, we constructed the instrument starting from in-depth, semi-structured interviews we had conducted in earlier research. In that research, we followed 32 individual cases (the decision process related to an individual student) during an academic year, using a case-study design with repeated interviews. This provided us with rich and dense descriptions we used to develop the questionnaire (inductive scale development; Hinkin, 1998). These items were derived from teachers' descriptions of what they did in the different steps of the decision process, how they used data, and how they used their intuition. For example, a teacher statement such as "In order to gain more understanding of the problem, I will observe my pupils during my daily practice and collect

information on the fly" was translated into an item for phase 2 (data collection). This was followed by deductive scale development. We used the existing literature described above (on data use and the recognition-primed decision model) to develop a table of specifications (see Table 1) and complement the items.

Table 1
Table of Specifications for Scale

Dimension	Intuitive	Data-driven
Phase		
Problem definition	(6 items)	(3 items)
Collection	(3 items)	(4 items)
Sense-making	(6 items)	(4 items)
Evaluation	(4 items)	(3 items)

This was followed by a content validation session. The items were typed as they would appear in the final instrument without being allocated to the different scales. All items that referred to deliberate and systematic processes of collection and analyses, starting from a pre-defined goal or question, were to be allocated to the data-driven dimension. All items referring to recognition—information gathering without a deliberate or systematic goal or plan—were to be allocated to the intuitive dimension.

The two authors and one colleague (researcher) independently matched the items with the scales. All items were allocated to the same scale except for two. The first was "When I make sense of data, I discuss this with colleagues." It was not clear to what extent this referred to data-driven or intuitive processes. During our collegial consultation, we changed the item to "When I make sense of data, I use shared criteria discussed with a colleague." The second item was "When I evaluate alternative decisions, I tend to rely most on what's in the student's best interest." This was not clear and was therefore changed to "most on *my feeling about* what is in the student's best interest." After this construction of the instrument, Hinkin (1998) stressed the necessity of a qualitative pre-test to establish construct validity. We ran a pre-test with four teachers in adult education and two peers who were fellow researchers (Cohen et al., 2017). While participants filled out the instrument, a think-aloud protocol was used to strengthen the cognitive validity: did teachers interpret the items in the same way that we as researchers intended (Field, 2009)? An interview was administered after the survey had been completed, to assess each individual item's suitability, face validity, and readability (Burgess et al., 1998).

After adjustments were made to the instrument's items based on the comments, the instrument was converted into a survey of 33 items with a 4-point Likert scale response format (ranging from "1=not important at all" to "4=extremely important"). Thus the higher the score, the more important the item for the respondent. A Likert scale is often used to measure respondent's attitudes by asking the extent to which they agree or disagree with a particular question or statement (Benson & Clark, 1982). The survey was structured around the 4 steps of the decision process (problem definition, data collection, sense making, evaluation of alternatives) for each of the two dimensions (data-based and intuitive). The Teacher Decision-Making Inventory (TDMI) can be found in Appendix 1.

Data analyses
All statistical analyses were carried out in open source R software, using the lavaan package (Rosseel, 2012).

Exploratory Factor Analyses
The factor structure was tested by carrying out exploratory factor analyses (EFA) for each of the steps of the decision process. Given our hypothesis that data-based and intuitive elements of teacher decision making mutually influence each other, we used oblique rotation, because it accounts for the expected relationship between the different factors (Loehlin, 2004).

Three elements were taken into account when defining the likely number of factors:
(1) applying the Kaiser criteria by calculating the number of factors with eigenvalues > 1;
(2) visually inspecting the scree plot;
(3) checking factor loadings and seeing whether there was a sound theoretical explanation.

After we had defined the number of factors, we inspected the loadings and deleted the items with unsatisfactory loadings (Field, 2009). In the final instrument, we kept only those items with high factor loadings on their own factor (≥ 0.30) and no/low loadings on the other factor. If an item did load on two factors, the difference between the two loadings should be > 0.15.

Confirmatory Analyses of the Structure of the TDMI
We performed confirmatory factor analyses (CFA) to test if the factorial structure was consistent with the theoretical model we developed for the instrument. We used the comparative fit index (CFI) and the root mean square error of approximation (RMSEA). CFI values ≥ 0.90 and an RMSEA value ≤ 0.05 were taken as indications that the data showed a relatively good fit with the model (De Maeyer & Kavadias, 2007).

<div align="center">Reliability Analysis</div>

The internal consistency of the instrument was measured by calculating its Cronbach's alpha coefficient. According to De Maeyer and Kavadias (2007), a scale with a Cronbach's alpha coefficient (α) in the range of 0.60 to 0.80 has reasonably good internal consistency, in the range of 0.80 to 0.90 has good internal consistency, and $\alpha > 0.90$ shows excellent internal consistency.

The scale has poor internal consistency if $\alpha < 0.60$ and unacceptable internal consistency if $\alpha < 0.50$.

<div align="center">*Participants*</div>

The web-based survey was administered to 101 teachers: 84 women (84%) and 17 men (16%) in adult education in Flanders (Belgium). The population of adult educators in Flanders consists of 665 teachers: 552 women (84%) and 113 men (16%). Although it is a small simple size for this pilot in validating the questionnaire, the sample is a good representation of the population in the distribution over men and women. Teacher participation was voluntary and all participants signed an informed consent form. They were informed about the purpose of the study, that they could decide to end their cooperation at any time, and that results could not be traced back to a single teacher's responses. Anonymity and preservation of the privacy of each participant was guaranteed.

<div align="center">**Results**</div>

<div align="center">*Exploratory Factor Analyses (EFA)*</div>

First, we conducted a data-driven approach. The factor structure was initially tested by carrying out exploratory factor analyses (EFA) with oblique rotation for each step of the decision process. The likely number of factors was found using the Kaiser criteria (eigenvalues > 1) and scree plot analyses (factors before the first inflection point). The analysis resulted in a two-factor solution for all four steps. The first factor referred to the intuitive dimensions of teacher decision making; the second factor referred to the data-based dimension of teacher decision making. Subsequently, factor loadings were checked and items were included if loadings were ≥ 0.30.

Table 2 shows the results of the factor analyses for the different phases of the decision process.

Table 2

Factor loadings (P = problem definition, D = data collection, S = sense making, E = evaluation of alternatives)

	Factor Loadings	
Items	Factor 1	Factor 2
The following aspects of the student's learning lead me to identify problems in relation to promotion…		
P1. Concentration-related behaviors that catch my attention	0.48	
P2. Motivation-related behaviors that catch my attention	0.54	
P3. Behaviors related to the student's interest in learning that catch my attention	0.61	
P4. Data that I analyze in the student tracking system		0.38
P5. Deficits related to literacy that I spontaneously recognize	0.37	
P6. Characteristics related to social status that I spontaneously recognize	0.59	
P7. Behaviors related to work ethic that catch my attention	0.47	
P8. Information from a regular meeting with a colleague		0.66
P9. Information from a team meeting		0.87
When I need more information in relation to the problem, I…		
D1. Observe the student using an observation protocol		0.49
D2. Search for information in the literature		0.45
D3. Read the notes I make during my daily practice		0.81
D4. Administer a targeted tests or assignment (e.g., to measure literacy)		0.32
D5. Retrieve information from memory of similar cases in the past	0.49	
D6. Feel what my intuition tells me	0.74	
When I make sense of data, I…		
S1. Take into account the effort a student makes	0.51	
S2. Take into account the student's socio-economic situation	0.72	
S3. Take into account the student's first language	0.61	
S4. Take into account the student's well-being	0.71	
S5. Take into account the student's social behavior	0.48	
S6. Adjust my evaluative criteria to meet the student's individual needs	0.51	
S9. Use shared criteria discussed with a colleague		0.82
S10. Use fixed criteria that apply for the school		0.41
S11. Weigh this result against earlier results		0.40
S12. Use criteria discussed with the students		0.30
When I evaluate alternative decisions, I tend to rely most on…		
E1. Information on the student's well-being gathered on the fly	0.85	
E2. Information on the student's social background gathered on the fly	0.71	
E3. Information on the student's motivation gathered on the fly	0.60	
E4. Results of the student's self-evaluation		0.53
E5. Information on the requirements of the future track		0.91
E6. My feeling about what is in the student's best interest	0.56	
E7. Test results		0.31

Confirmatory Factor Analyses (CFA)

Subsequently, we tested whether the two-factor structure fit within our theoretical model build around the four phases of the decision process. We subjected this model to confirmatory factor analyses to confirm our model, but the initial model did not fit the data well (CFI = 0.77; RMSEA = 0.76). We carefully studied modification indices, looking for a better fitting model, with theoretical considerations also being taken into account. Given the observed cross-loading of the item "conversation with colleagues" (evaluation of alternatives) with "data collection" and "sense making," we deleted this item. Further, error covariances were included.

Our final model is shown in Figure 3. Based on the goodness-of-fit indices, we concluded that our data show a good fit with the model (CFI = 0.90; RMSEA = 0.05).

Reliability: internal consistency

Subsequently, the internal consistency was calculated. The different subscales appear to show reasonably good to good internal consistency: problem recognition ($\alpha = 0.72$), problem analysis ($\alpha = 0.74$); intuitive collection ($\alpha = 0.69$), data-based collection ($\alpha = 0.60$); subjective interpretation ($\alpha = 0.84$), objective interpretation ($\alpha = 0.72$); intuitive evaluation ($\alpha = 0.82$), and data-based evaluation ($\alpha = 0.70$).

Discussion and Conclusion

The aim of the present study was to fill an important gap in the research: the lack of a validated instrument that allows the investigation of teacher decision making on a large scale, adopting a dual-process approach. Moreover, fine-grained insight into how teachers use data or intuition in the different steps of the decision process is scarce. In our study, we developed and validated a Teacher Decision-Making Inventory (TDMI) that measures two dimensions (data use and intuition) in the four different steps of the decision process: (1) problem definition, (2) data collection, (3) sense making, and (4) evaluation of alternatives. This questionnaire was built in two phases. The starting point was our theoretical model derived from earlier research (Vanlommel et al., 2020). Based on the rich and dense descriptions from teachers in our qualitative research, we developed items for each of the steps in the model. In a second phase, this survey was pre-tested in practice and discussed in the research team before it was administered.

Exploratory factor analyses identified two dimensions: data-based and intuitive, with good, unique factor loadings. Ideally, we would have split the sample in two, using the first half for exploratory factor analyses to identify the initial structure and using the other half for confirmatory factor analyses.
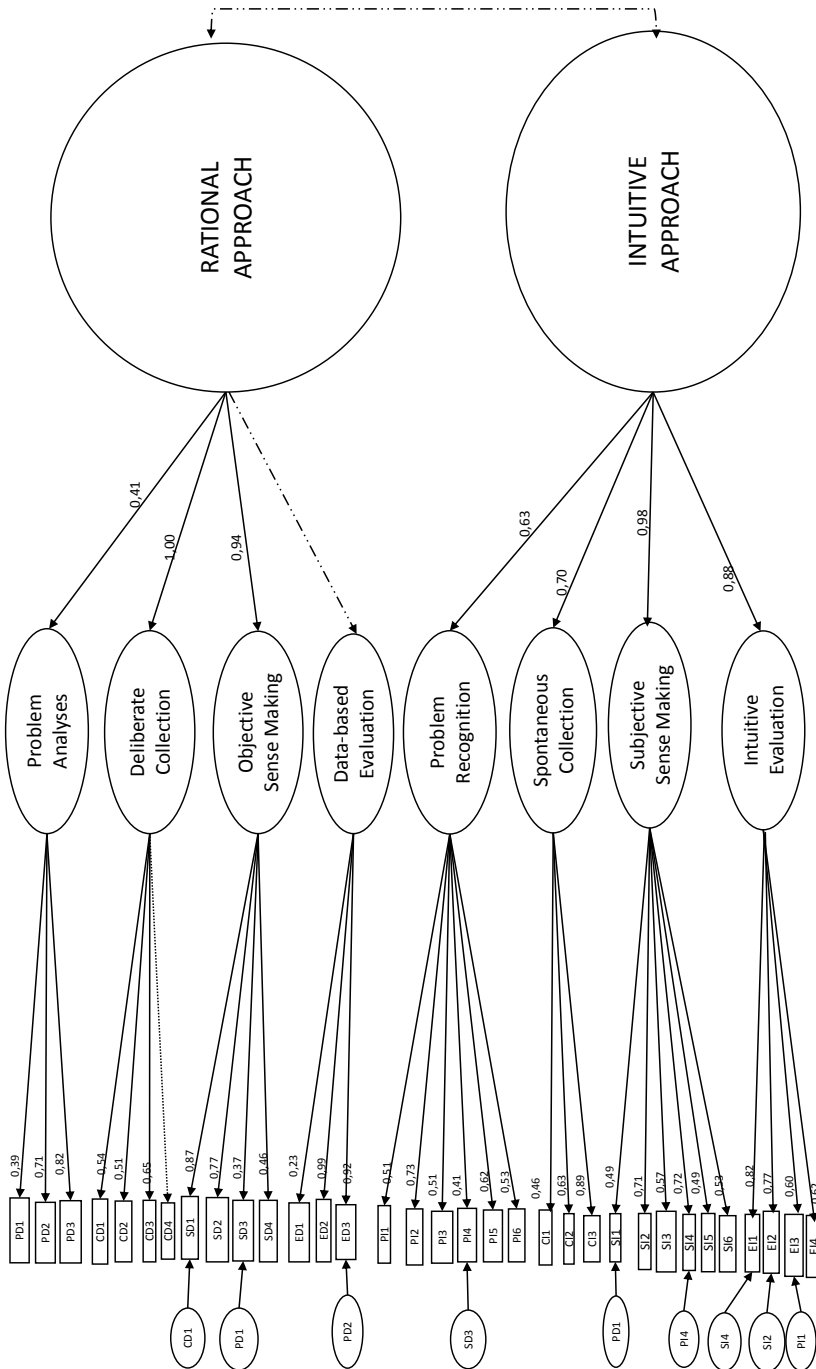
Figure 3

Findings of the confirmatory factor analyses (CFA) for the Teacher Decision-Making Inventory (TDMI). All factors with full arrows significant at the 0.05 level.

Our sample was, however, not large enough to follow this approach. This research can be considered as a pilot study in which the first, important steps were taken to develop and validate a teacher decision-making inventory. The scales measuring data-driven and intuitive collection of information need extra consideration.

Overall, our confirmatory factor analyses showed a good fit for our model, while reliability analyses showed reasonably good to good internal consistency of the scales. It can thus be concluded that the TDMI is both grounded in theory and a good psychometric tool that can be used to assess how data-based or intuitive a teacher's approach is in making decisions. This is an important step for research, policy, and practice to understand and support professional decision making in education.

Our starting point was that professional decision making begins with a wise combination of data and intuition, collected, analyzed, and weighed through an extensive decision process. The main question is not the extent to which teachers use either data or intuition; the crux of the matter is the extent to which teachers critically question problems they recognize, consciously search for answers, combine information, weigh alternatives, and conduct a decision process deliberately and skillfully. Our validated survey is a valuable step towards exploring, explaining, and strengthening this professional decision making in practice.

There are, of course, also limitations to this study. For one thing, we had a small sample size. Planning, constructing, and validating instruments requires large amounts of time, large funding, and large sample sizes (Benson & Clark, 1982). Therefore, validation should be seen as a continual process. We feel that our small scale pilot study delivers a valuable and important starting point for the next step in validating the TDMI. Future research is needed, on a larger scale, in different contexts such as different educational levels and in different educational cultures or political structures. For one thing, the educational system of Flanders (Belgium) is characterized by high decision-making autonomy and low accountability: there is, for example, no binding obligation to use the results of standardized test for streaming or tracking. Other decision-making processes or data may appear in other systems.

In order to contextualize and standardize the questions for all teachers to some extent, we also focused on a tough promotion decision. Given the high stakes related to promotion to a subsequent educational level/track or retention, we expected teachers to go through the decision process thoroughly and use a wide range of evidence before reaching a decision. Further research could test and further develop our survey for other decisions (e.g., student placement in groups, evaluation of creative projects, curriculum redesign, and so forth). In order to enhance the validity and reliability of teacher decisions, it is important to gain a fine-grained understanding of how teachers use data or

intuition in different contexts, for different decisions. These insights can help practitioners, researchers, and policymakers to develop targeted support and training that strengthens professional decision making in education.

We do hope that our research will be used as a starting point for further exploration and validation of decision-making processes in practice. Investigating, understanding, and enhancing the quality of teacher judgment is important as it highly influences student's educational trajectories, fairness, and equity.

References

Allal, L. (2013). Teachers' professional judgement in assessment: A cognitive act and a socially situated practice. *Assessment in Education: Principles, Policy & Practice, 20*(1), 20–34. https://doi.org/10.1080/0969594X.2012.736364

Benson, J., & Clark, F. (1982). A guide for instrument development and validation. *American Journal of Occupational Therapy, 36*(12), 789–800. https://doi.org/10.5014/ajot.36.12.789

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

Blackwell, R. D., Miniard, P. W., & Engel, J. F. (2006). *Consumer behavior.* Harcourt College Publishers.

Bonvin, P. (2003). The role of teacher attitudes and judgement in decision-making: The case of grade retention. *European Educational Research Journal, 2*(2), 277–294. https://doi.org/10.2304/eerj.2003.2.2.6

Bosker, R. J., Branderhorst, E. M., & Visscher, A. J. (2007). Improving the utilisation of management information systems in secondary schools. *School Effectiveness and School Improvement, 18*(4), 451–467. https://doi.org/10.1080/09243450701712577

Bromme, R., Prenzel, M., & Jäger, M. (2014). Empirische Bildungsforschung und evidenzbasierte Bildungspolitik. *Zeitschrift für Erziehungswissenschaft, 17*(4), 3–54. https://doi.org/10.1007/s11618-014-0514-5

Brown, C. (2017). Research learning communities: How the RLC approach enables teachers to use research to improve their practice and the benefits for students that occur as a result. *Research for All, 1*(2), 387–405.

Burgess, P. W., Alderman, N., Evans, J. O. N., Emslie, H., & Wilson, B. A. (1998). The ecological validity of tests of executive function. *Journal of the International Neuropsychological Society, 4*(6), 547–558. https://doi.org/10.1017/S1355617798466037

Burgess, S., Greaves, E., Vignoles, A., & Wilson, D. (2009). *What parents want: school preferences and school choice.* CMPO.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*(1), 55–81. https://doi.org/10.1016/0010-0285(73)90004-2

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

Cohen, L., Manion, L., & Morrison, K. (2017). Action research. In L. Cohen, L. Manion, & K. Morrison (Eds.), *Research methods in education* (8th ed., pp. 440–456). Routledge.

Datnow, A., & Hubbard, L. (2015). Teachers' use of assessment data to inform instruction: Lessons from the past and prospects for the future. *Teachers College Record, 117*(4), 1–26.

Datnow, A., Park, V., & Wohlstetter, P. (2007). *Achieving with data.* University of Southern California, Center on Educational Governance.

De Maeyer, S., & Kavadias, D. (2007). *Openleerpakket beschrijvende statistiek.* Academia Press.

Earl, L. M., & Katz, S. (Eds.). (2006). *Leading schools in a data-rich world: Harnessing data for school improvement.* Corwin Press.

Earl, L., & Louis, K. S. (2013). Data use: Where to from here? In K. Schildkamp, M. K. Lai, & L. Earl (Eds.), *Data-based decision making in education: Challenges and opportunities* (pp. 193–204). Springer.

Ehren, M. C., & Swanborn, M. S. (2012). Strategic data use of schools in accountability systems. *School Effectiveness and School Improvement, 23*(2), 257–280. https://doi.org/10.1080/09243453.2011.652127

Eurydice. (2011). *Grade retention during compulsory education in Europe: regulations and statistics.* OECD.

Evans, J. S. B. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology, 59,* 255–278. https://doi.org/10.1146/annurev.psych.59.103006.093629

Field, A. (2009). *Discovering statistics using SPSS.* SAGE Publications Limited.

Harteis, C., Koch, T., & Morgenthaler, B. (2008). How intuition contributes to high performance: An educational perspective. *US-China Education Review, 5*(1), 68–80.

Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods, 1*(1), 104–121. https://doi.org/10.1177/109442819800100106

Hogarth, R. M. (2001). *Educating intuition.* University of Chicago Press.

Hubbard, L., Datnow, A., & Pruyn, L. (2014). Multiple initiatives, multiple challenges: The promise and pitfalls of implementing data. *Studies in Educational Evaluation, 42,* 54–62. https://doi.org/10.1016/j.stueduc.2013.10.003

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist, 58*(9), 697–720. https://doi.org/10.1037/0003-066X.58.9.697

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49–81). 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

Kaiser, J., Retelsdorf, J., Südkamp, A., & Möller, J. (2013). Achievement and engagement: How student characteristics influence teacher judgments. *Learning and Instruction, 28,* 73–84. https://doi.org/10.1016/j.learninstruc.2013.06.001

Klein, G. (2008). Naturalistic decision making. *Human Factors, 50*(3), 456–460. https://doi.org/10.1518/001872008X288385

Loehlin, J. C. (2004). *Latent variable models: An introduction to factor, path, and structural equation analysis* (4th ed.). Lawrence Erlbaum.

Mandinach, E. B., & Jimerson, J. B. (2016). Teachers learning how to use data: A synthesis of the issues and what is known. *Teaching and Teacher Education, 60,* 452–457. https://doi.org/10.1016/j.tate.2016.07.009

Mandinach, E. B., Honey, M., & Light, D. (2006, April 9). *A theoretical framework for data-driven decision making.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Mintzberg, H., & Westley, F. (2001). It's not what you think. *MIT Sloan Management Review, 42*(3), 89–93.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *Journal of Statistical Software, 48*(2), 1–36.

Rubie-Davies, C. M. (2010). Teacher expectations and perceptions of student attributes: Is there a relationship? *British Journal of Educational Psychology, 80*(1), 121–135. https://doi.org/10.1348/000709909X466334

Schildkamp, K., & Lai, M. K. (2013). Conclusions and a data use framework. In K. Schildkamp, M. K. Lai, & L. Earl (Eds.), *Data-based decision making in education: Challenges and opportunities* (pp. 177–191). Springer.

Schildkamp, K., Poortman, C. L., & Handelzalts, A. (2016). Data teams for school improvement. *School Effectiveness and School Improvement, 27*(2), 228–254. https://doi.org/10.1080/09243453.2015.1056192

Strayhorn, T. L. (2009). Accessing and analyzing national databases. In T. Kowalski & T. J. Lasley (Eds.), *Handbook of data-based decision making in education* (pp. 121–138). Routledge.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*(4481), 453–458. https://doi.org/10.1007/978-1-4613-2391-4_2

Urhahne, D., & Wijnia, L. (2020). A review on the accuracy of teacher judgments. *Educational Research Review*, 32, 100374. https://doi.org/10.1016/j.edurev.2020.100374

Vanlommel, K. (2018). *Opening the black box of teacher judgement: The interplay of rational and intuitive processes* [Dissertation]. University of Antwerp. https://www.researchgate.net/publication/324909244_Opening_the_black_box_of_teacher_judgement_the_interplay_of_rational_and_intuitive_processes_Doctoral_thesis

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. (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

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*(1), 108–119.

Vanlommel, K., Van Gasse, R., Vanhoof, J., & Van Petegem, P. (2020). Sorting pupils into their next educational track: How strongly do teachers rely on data-based or intuitive processes when they make the transition decision? *Studies in Educational Evaluation*, 69, 100374. https://doi.org/10.1016/j.stueduc.2020.100865

Zanting, A., Verloop, N., & Vermunt, J. D. (2001). Student teachers' beliefs about mentoring and learning to teach during teaching practice. *British Journal of Educational Psychology, 71*(1), 57–80. https://doi.org/10.1348/000709901158398

Corresponding authors
Kristin Vanlommel
University of Applied Sciences Utrecht, the Netherlands
E-mail: kristin.vanlommel@hu.nl

Elke Pepermans
Institute for Adult Education Antwerp, Belgium
E-mail: elke.pepermans@cvoantwerpen.be

Appendix 1

Teacher Decision-Making Inventory (TDMI)

This inventory is designed to explore how you make a tough transition decision in relation to promotion to the next educational level/track. We ask you to picture a tough case, in which the decision is not straightforward. The questions investigate the different phases of the decision process during the year. How were you able to define the problem in relation to promotion at the start of the year? How did you collect more information during the year? How did you interpret test results for this student and on what evidence base would you rely most at the end of the year?

You can answer these questions on a scale ranging from (1) not important at all to (4) very important.

Statement	Assessment			
The following aspects of the student's learning lead me to identify problems in relation to promotion…				
Concentration-related behaviors that catch my attention	1	2	3	4
Motivation-related behaviors that catch my attention	1	2	3	4
Behaviors related to the student's interest in learning that catch my attention	1	2	3	4
Data that I analyze in the student tracking system	1	2	3	4
Deficits related to literacy that I spontaneously recognize	1	2	3	4
Characteristics related to social status that I spontaneously recognize	1	2	3	4
Behaviors related to work ethic that catch my attention	1	2	3	4
Information from a regular meeting with a colleague	1	2	3	4
Information from a team meeting	1	2	3	4
When I need more information in relation to the problem, I…				
Observe the student using an observation protocol	1	2	3	4
Search for information in the literature	1	2	3	4
Read the notes I make during my daily practice	1	2	3	4
Administer a targeted tests or assignment (e.g., to measure literacy)	1	2	3	4
Retrieve information from memory of similar cases in the past	1	2	3	4
Feel what my intuition tells me	1	2	3	4
When I make sense of data, I…				
Take into account the effort the student makes	1	2	3	4
Take into account the student's socio-economic situation	1	2	3	4
Take into account the student's first language	1	2	3	4
Take into account the student's well-being	1	2	3	4
Take into account the student's social behavior	1	2	3	4
Adjust my evaluative criteria to meet the student's individual needs	1	2	3	4
Use shared criteria discussed with a colleague	1	2	3	4

Use fixed criteria that apply for the school	1	2	3	4
Weigh this result against earlier results	1	2	3	4
Use criteria discussed with the students	1	2	3	4
When I evaluate alternative decisions, I tend to rely most on…				
Information on the student's well-being gathered on the fly	1	2	3	4
Information on the student's social background gathered on the fly	1	2	3	4
Information on the student's motivation gathered on the fly	1	2	3	4
Results of the student's self-evaluation	1	2	3	4
Information on the requirements of the future track	1	2	3	4
My feeling about what is in the student's best interest	1	2	3	4
Test results	1	2	3	4