

Aligning Assessment with Purpose to Support the Learning of Children in Dutch Primary Schools with Different Pedagogical Approaches

Femke Wassenaar, 1153545

Windesheim Honors College

Global Project and Change Management: Bachelor Internship

Radostina Sharenkova-Toshkova

5 July 2024

## Table of contents

Abstract .....	3
Glossary.....	5
1. Introduction .....	6
1.1 Background .....	6
1.2 Problem statement .....	7
1.3 Internship organization: Operation Education.....	8
2. Literature Review.....	10
3. Conceptual Framework .....	19
4. Research Questions and Objective .....	21
5. Research Design .....	22
5.1 Data collection methods .....	23
5.2 Sampling methods.....	25
5.3 Data analysis .....	27
5.4 Quality assurance .....	29
6. Research results and discussion.....	30
6.1 Sub-question 1 .....	31
6.2 Sub-question 2 .....	40
6.3 Sub-question 3 .....	48
6.4 Sub-question 4 .....	52
6.5 Other relevant observations .....	56
7. Limitations.....	57
8. Conclusion .....	59
9. Relevance.....	61
10. Contribution to the PP .....	62
11. References.....	63
Appendix A – Conceptual framework.....	70
Appendix B - Participants overview .....	71
Appendix C - Interview guide assessment experts.....	74
Appendix D - Interview guide school administrators and teachers .....	77

## Abstract

This study examines how Dutch primary schools with different pedagogical approaches align their assessment methods with their educational purposes to support children's learning. The current emphasis in Dutch education on standardized testing and measurable outcomes often overlooks broader educational objectives such as creativity, social skills, and personal development. The frequent use of summative assessments has led to increased stress among students and a decline in intrinsic motivation, impacting their overall learning experience and well-being. Furthermore, this focus on standardized testing perpetuates educational inequality, as wealthier families can afford additional tutoring to improve test scores. Therefore, there is a growing need for a balanced assessment approach that supports learning and aligns with educational purposes, requiring intentional leadership and curriculum redesign to achieve constructive alignment between learning activities, objectives, and evaluation methods.

The conceptual framework is based on three main themes identified by Schellekens et al. (2021) on how assessment supports learning. These themes are student-teacher roles and relationships, assessment learning environments and educational outcomes. Focusing on these themes (and the nine underlying characteristics) helps avoid confusion from overlapping assessment concepts and emphasizes how assessment can support learning.

The research adopts a qualitative, exploratory design to understand the problem's nature and formulate future research questions. It involves three phases: ongoing desk research, interviews with assessment experts, and interviews with school administrators and teachers. The study uses semi-structured interviews and a combination of purposive- and snowball sampling to gather diverse perspectives. Data analysis combines deductive and inductive methods to identify themes and patterns, employing comparative analysis and colour coding. Quality assurance measures include triangulation, consistency checks, expert review, and reflexivity to address potential biases and enhance credibility, dependability, confirmability, and transferability.

Eventually, three assessment experts and eight education professionals (one school administrator and teacher per school) from four schools with different pedagogical approaches have been interviewed. While all participants agree that assessments should enhance learning, practical application often falls short due to external pressures and lack of ownership. Schools with development-based, nature-inclusive, and Montessori approaches showed better alignment and more supportive assessment characteristics, whereas 'regular' schools faced greater challenges.

Moreover, according to this research a big issue hindering the alignment of assessment with purpose is the impact of top-down assessment and accountability approaches. This limits the involvement of students, teachers, and administrators, hindering effective, bottom-up assessment practices. There also seems to be a lack of awareness about various assessment functions. Despite the potential of self- and peer-assessment to foster student ownership and improve outcomes, practical challenges like administrative burdens and insufficient assessment literacy often limit their use. The system's theoretical flexibility to adapt to student needs is rarely realized in practice. Some schools rely too heavily on standardized tests, focusing narrowly on measurable outcomes rather than overall educational quality.

To address these issues, assessment experts recommend empowering schools with assessment autonomy and increasing assessment literacy among educators and students. This would help better align assessment practices with educational goals, effectively supporting children's learning.

## Glossary

Some common terms used in Dutch primary education (that are relevant for this study) are described in Table 1.

*Table 1 - Common terms in Dutch primary education (relevant for this study)*

<b>LVS (LeerlingVolgSysteem)</b>	An LVS ('Student Following System') monitors student's developments. Dutch primary schools are obliged to use such a system, consisting of a series of tests taken twice a year. The test results are compared normatively to gain insight into the current arithmetic and language level of a student.
<b>Arithmetic</b>	Arithmetic is the basic field of mathematics taught at primary school, such as addition, subtraction, multiplication, and division (in Dutch called 'rekenen').
<b>Transition test (doorstroomtoets)</b>	The 'transition test' is taken at the end of primary school by (almost) all Dutch children and shows what level students have in the areas of language and arithmetic. The test is intended to provide a fitting secondary school advice and to measure the learning outcomes of a primary school. There are several providers of the transition test in the Netherlands.
<b>CITO (Centraal Instituut voor Toetsontwikkeling)</b>	CITO ('Central Institute for the Development of Tests') is the founder of the transition test and is still one of the main providers. Moreover, they provide a LVS (including the corresponding biennial tests).
<b>IEP (Inzicht Eigen Profiel)</b>	IEP ('Insight into Own Profile') is a relatively new provider of the transition test, as well as an LVS. IEP currently (2024) is the second most popular provider of both, after CITO.
<b>ParnasSys</b>	ParnasSys is an online data collection program, for monitoring the entire student administration. Test results, as well as non-numerical data, such as notes from conversations and observations regarding social-emotional development can be entered here.
<b>Method-related test</b>	Most Dutch primary schools use 'method-related tests' (in Dutch known as 'methode toetsen'). These are tests linked to a certain teaching method, for example to train children's reading skills. After each period or instruction block, children take a method-related test to determine whether they have sufficiently mastered what they have just learned. Different than LVS-tests, the scores of method-related tests are not compared nationally.
<b>SLO (Sectie Leeropbouw Onderwijs)</b>	The SLO ('Section Learning Structure Education') is an organisation developing the (base of) the national curriculum in the form of goals for primary, secondary and special education in the Netherlands.
<b>De Onderwijsinspectie ('De Inspectie')</b>	The 'Onderwijsinspectie', or abbreviated 'Inspectie' ('Inspectorate of Education') monitors the quality of education on behalf of the Ministry of Education, Culture and Science (OCW). Every year they report on the 'current state of education'.

# 1. Introduction

## 1.1 Background

With the establishment of mass education, the need for national standardized assessments grew. Initially, these assessments served as a passive control mechanism for the Inspection but eventually evolved into tools for actively improving education (Mellink, 2016). A significant development in Dutch education was the introduction of the CITO-test in 1966, a standardized arithmetic and language test for all children transitioning from primary to secondary education. This assessment aimed to increase objectivity and increase equality of opportunities (Heij, 2021; Pointer, 2023).

Nowadays, this transition test (now called 'doorstroomtoets') still functions as a reliable tool for a fitting secondary school advice, an objective 'second opinion,' and one of the measurements indicating the quality of schools, according to the website of the Dutch government (Ministerie van Algemene Zaken, 2024). The 'doorstroomtoets' is an example of the normative, standardized tests most Dutch children from the age of six take twice a year, measuring their language and arithmetic levels and comparing these to the national norms (Mellink, 2016).

The Dutch education system places significant emphasis on test outcomes and the frequency of testing (Mellink, 2016; Heij, 2021). Consequently, national quality standards are based on school results (outputs) rather than educational ideology or a shared educational purpose (which is part of a school's mission or vision) (inputs) (Boonstra, 2023; Biesta, 2012). This focus has led to an education system predominantly measured through objectively assessed summative tests (Heij, 2021), compelling schools to prioritize measurable language and math skills over other important attributes such as creativity, social skills, democratic understanding, and personal development (Biesta, 2012; Didactief & CITO, 2014).

## 1.2 Problem statement

The current Dutch education system faces multiple challenges, including stress and mental health issues among students due to the pressure of summative tests, which increase cortisol levels (Reeve & Tseng, 2011; HBSC, 2021). Additionally, there is a decline in learning outcomes as students focus on passing tests rather than truly understanding the material (Visser, 2023). This is compounded by a drop in intrinsic motivation, as the emphasis on grades reduces students' autonomy, competence, and connection to their studies (Heylen, 2022; Deci, 1971). For instance, no European country has as little enjoyment in mathematics as in the Netherlands (OECD, 2016), and nearly half of Dutch children has never read for pleasure (OECD, 2016). Furthermore, Dutch students' language and math scores have been declining over the past years. PISA's (Programme for International Student Assessment) latest research shows that only Greece scored lower in reading skills among 14 participating European countries (PISA, 2022).

Thus, there is a need to make Dutch education more engaging and improve academic outcomes. The focus on standardized testing also contributes to inequality, as wealthier families can afford extra tutoring (Scheider et al., 2023; Heij, 2021), disadvantaging late bloomers and non-native speakers, and perpetuating educational disparities (Boonstra, 2023; HUMAN, 2022). Lastly, standardized tests often misjudge students' abilities by concentrating on limited cognitive skills and ignoring diverse forms of intelligence (Biesta, 2016; Gardner, 1995).

This raises the question to what extent Dutch primary schools (consciously) choose methods of assessment that support the learning of children and are in line with their purposes. In other words: do we treasure what we (can) measure? Or do we measure what we find valuable in ways that actually support the learning of children (Biesta, 2012)?

Sneider et al. (2023) argue that the Dutch educational assessment system is unable to properly detect students' potential and fully support their learning, due to a lack of alignment and absolute standards. A clear vision from the central government and educational institutions is necessary for a balanced approach to testing and assessment (Ministerie voor Onderwijs, Cultuur en Wetenschap, 2019). Sluijsmans (2023) agrees, noting that a shift toward assessment as a crucial and more meaningful component of a school's curriculum requires strong and intentional leadership at the school level.

Moreover, according to Sluijsmans (2023), working towards a different vision on assessment demands a lot from the involved actors. First, it requires an increase in literacy in assessment – the different forms, functions, and effects on children's learning - and how to align this with curriculum. Secondly, re-designing assessment procedures involves reconsidering the curriculum as a whole: having a clear idea about a school's purpose and what to transfer to the children. This is linked to the 'constructive alignment' principle, which is widely used in (higher) education. The goal of constructive alignment is to consciously establish connections between learning activities, learning objectives, and evaluation (Biggs & Tang, 2020).

### 1.3 Internship organization: Operation Education

Going back to the purpose is also what Operation Education – the organization where this internship took place - stands for. They are looking for new ways to view education and development and how to organise this differently together. Moreover, Operation Education contributes to the educational transition by offering innovative training programs for formal and informal leaders.

One example of their services is 'Expedition Leadership' – a profound and transformative annual program for current and future leaders in education who realize that they are the ones who must be the change they want to see in the world. These people are guided and coached by practical experts and school administrators, who have led the way in this transformation (Operation Education, 2024). Other examples of their services are a keynote and workshop based on the 'Education Questions book' that they wrote. In this book, they questioned and researched many components of the Dutch education system, like 'Why do we focus on cognitive developments?' and 'Why do we work with standardized teaching methods?'.

## 2. Literature Review

Assessment can be defined as “a wide range of methods for evaluating student performance and attainment” (Gipps, 2011, p.11). Assessment is a crucial component of the educational process, serving as a critical tool for evaluating student learning and a measurement instrument for quality education (Gezer et al., 2021). When people think of ‘tests’ or ‘assessments’, they often associate this with ‘sweating, knowing, forgetting’ (LAKS, 2021; Visser, 2023). Contrary to serving predominantly as a tool for accountability or being a stressful moment of performance for students, assessment can take a variety of different forms and truly support children’s learning (Sluijsmans, 2020).

In this literature review, a few common concepts of assessment (summative, formative, and assessment of-, for-, and as learning) are explored. Additionally, the effects of different types of assessment on children’s learning are described, specifically in primary education. In the next section, an overview of assessments in the Dutch primary education system is provided, as well as some points of discussion. Finally, the importance of more purposeful assessment to support children’s learning is explained.

### **Different notions of assessment**

#### *Summative and formative assessment*

Summative and formative assessment have been common terms used by educators since Scriven distinguished these concepts in 1967. These two concepts can generally be assigned distinct purposes: assessments with a formative function help students learn more and encourage them to learn (through ongoing feedback), whereas assessments with a summative function rate or certify students’ competency based on their performance after an instructional period or course unit. (Gezer et al., 2021; Taras, 2009).

Some examples of summative assessments include final state exams, term essays, and presentations at the end of a semester (Gezer et al., 2021). Asking a diagnostic question in class to determine who needs more instruction, in-class discussions, peer feedback, and self-assessment are examples of formative assessment (Kneyber et al., 2022). Moreover, according to Hattie (2013), formative assessment is most effective when the student provides feedback to the teacher. This increases the students' feelings of ownership of their learning process and therefore they engage more actively. Moreover, the teaching and learning can be more aligned and adapted to the students' needs due to the direct feedback provided.

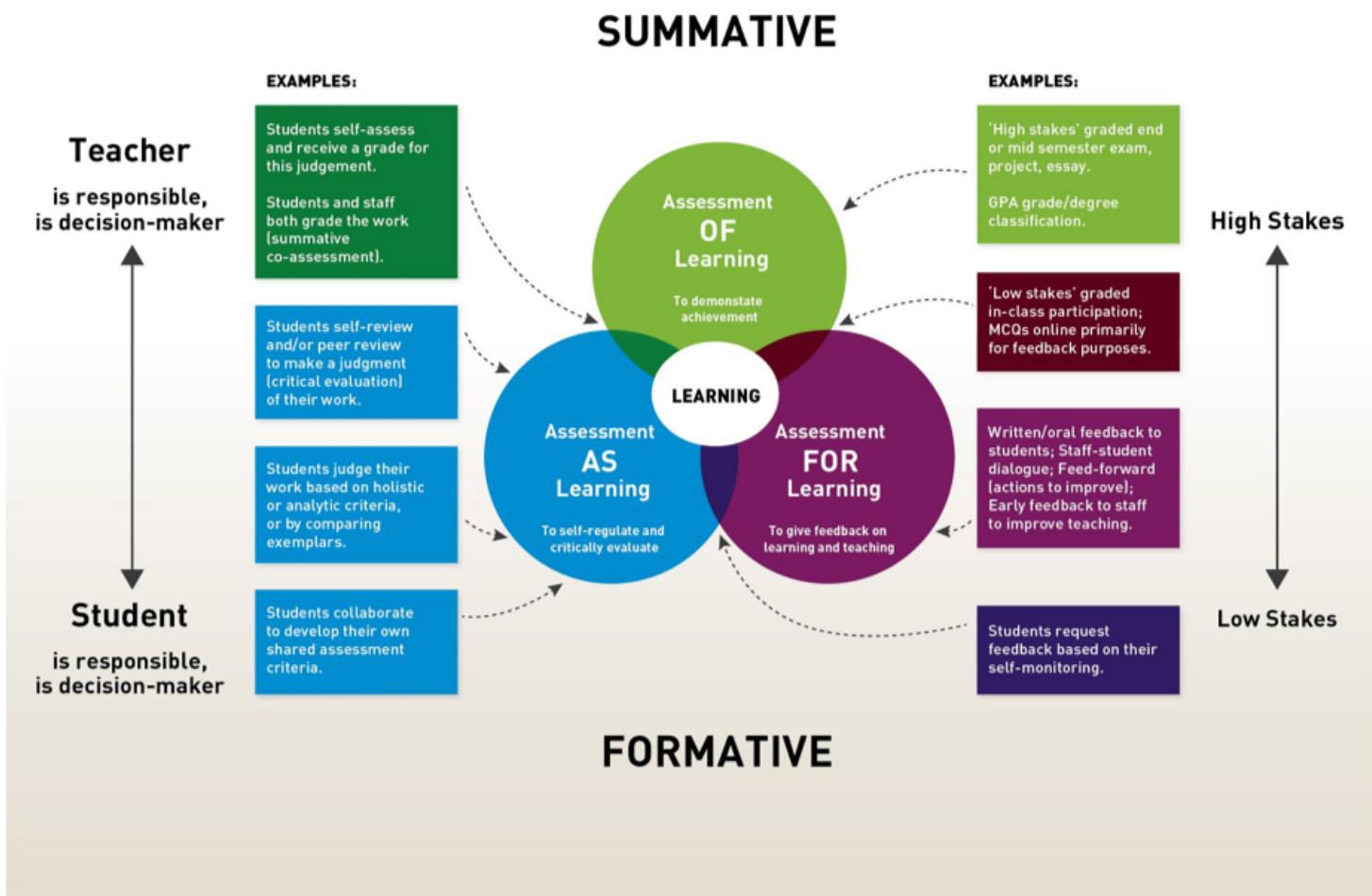
A delusion often portrayed in the literature is that summative assessment is 'bad' for students' learning and formative tests are 'good' (Lau, 2015). Research shows that in order to support student's learning, summative and formative assessment should both be used in connection with each other. The problem is often the disconnection between the assessment method and the overall learning environment (Lau, 2015).

#### *Assessment of-, for- and as learning*

According to Vasileiadou and Karadimitriou (2021), traditional assessment methods are often unable to respond to learning environments, in which students are expected to engage in active learning, critical thinking, self-monitoring, and self-regulation. This has led to new notions of assessment over the past decades: assessment of-, for- and as learning. The National Forum for the Enhancement of Teaching and Learning in Higher Education (The National Forum, 2017), which is based in Northern-Ireland, further explains these concepts in a visual (see Figure 1 on the next page). Generally, assessment of learning (AoL) resembles the summative assessment in the 'old' model and assessment for learning (AfL) is similar to formative assessment (Earl & Katz, 2006). Assessment as learning (AaL) is added to this model, which is characterized by the active participation of students in assessment and their learning (Earl & Katz, 2006; Schellekens et al., 2021).

The visual created by The National Forum (2017), does not only give examples of these different types of assessment, but also show how they overlap, demonstrating the difficulty in assigning a fixed label to some assessment forms. Furthermore, they show how AoL, AfL and AaL relate to low/high stakes testing. High-stakes tests are generally used for the purpose of accountability and used to make critical decisions about students, educators or schools (Sluijsmans, 2020). Whereas a low-stake test would be used to support students' learning by measuring academic achievement, identifying learning gaps and informing educators about instructional adjustments. The difference between these two types of tests is not their form, but its function, in other words: how the results are used (The National Forum, 2017). Lastly, the visual shows the relationship between the assessment types and the amount of student/teacher ownership.

Figure 1 - Assessment of-, as- and for learning (The National Forum for the Enhancement of Teaching and Learning in Higher Education, 2017)



## **Moving away from assessment notions**

The assessment notions and the way they are defined, connected, and visualized by the National Forum (2017) are appealing in theory. However, owing to the range of educational settings, the variety of definitions used in academic literature, and the number of misconceptions, these notions are still unclear constructs to put into practice (McDowell et al., 2011; Schellekens et al., 2021). As a result, educators might not be fully aware of how assessment should be used to support learning and instead become preoccupied with attempting to comprehend various notions (Tan, 2016). Therefore, Schellekens et al. (2021) did a scoping review based on different notions of assessment (AoL, AfL, and AaL), to research the common characteristics of assessment that support learning. As a result, they found nine characteristics (see '3. Conceptual Framework' for an overview of all nine characteristics), that they grouped into three different categories 1) student-teacher roles and relationships within assessment; 2) assessment learning environment; and 3) educational outcomes of assessment. Afterwards, they researched how these characteristics contribute to students' learning in practice. With this scoping review, Schellekens et al. (2021) offer a more nuanced overview of the connection between learning and assessment than the separate definitions and descriptions that the assessment notions offer. Moreover, their research is better suited to be operationalized for this research, as compared to, for example, the assessment concepts offered by the National Forum (2017) (see '3. Conceptual Framework').

## **The connection between assessment and learning**

Assessment significantly influences how children learn. It affects their understanding of the importance of learning activities, their engagement level, and their ability to apply insights to future learning (Schellekens et al., 2021). In this research, 'learning' refers to the process of anchoring knowledge and skills in long-term memory, and to students being able to use this anchored knowledge and skills in different situations, also outside the school context (Soderstrom & Bjork, 2015).

According to Sluijsmans (2020), it is insufficient to determine what students have learned by taking one test at one moment directly after an instruction period. Moreover, this can lead to the experience of extra pressure and stress, which reduce the ability to learn (HBSC, 2021; Reeve & Tseng, 2011). A focus on passing tests rather than internalizing knowledge also leads to decreased learning outcomes (Visser, 2023). This is intensified by a decline in intrinsic motivation for learning, as the focus on grades diminishes students' autonomy, competence, and connection to their studies (Deci, 1971; Heylen, 2022). Finally, standardized tests misjudge students' abilities by focusing on limited cognitive skills and failing to account for diverse forms of intelligence (Biesta, 2016; Gardner, 1995).

So, how can there be a more meaningful connection between assessment and learning? The scoping review by Schellekens et al. (2021), which will be elaborated on later in this study, already offers a valuable framework for the characteristics of assessment to support learning (see '3. Conceptual Framework'). Additionally, Sluijsmans (2020) emphasizes that effective learning can only happen based on quality information about each student's learning process. To obtain this information, many short evaluation moments ('Where does the student stand?', 'What is needed to move forward?' and 'Where is the student working towards?'), peer feedback and encouraging students to take responsibility for their own learning are helpful (Sluijsmans, 2020). Moreover, using other non-graded learning activities instead of standardized tests stimulates teachers to focus more on the learning process of children and differentiate their instruction methods (Baarda et al., 2020).

Self-assessment is also a useful tool to receive useful information for teachers on how students think and learn. It can also assist students in identifying areas in which they require additional support, leading to more (self-)directed and focused learning (Vasileiadou & Karadimitriou, 2021). However, students must be assisted in becoming more assessment- and feedback-literate in order for self-assessment to be used effectively (Schellekens et al., 2021; Sluijsmans, 2023).

## **Different pedagogical approaches and methods of assessment in Dutch primary education**

### *Free choice of pedagogical approach in Dutch education*

In the Netherlands there is a range of different pedagogical approaches, like democratic education, Dalton, Jenaplan, Montesorri, OGO ('ontwikkelingsgericht onderwijs', meaning 'development-based education'), EGO ('ervarings-gericht onderwijs', meaning 'experience-based education'), nature-inclusive and Vrije School (Steiner school) (Wouda, 2022). However, the majority of the primary schools in the Netherlands do not have a 'special educational concept' and could be labelled as 'regular'. This does not mean they do not have their own vision or pedagogical approach.

### *Dutch assessment at primary schools in a nutshell*

Children enroll in primary education at the age of four in the Netherlands. According to the Ministerie van Algemene Zaken (2024) (the Dutch Ministry), primary schools are obliged to use one of the four government-approved student-tracking-systems, namely Boom, IEP, CITO, or Dia, to track the development of at least Dutch language and calculating-mathematics skills. Primary schools use these student-tracking systems to keep records of the results of standardized tests, as well as other observation instruments. In the past, Dutch children had to take standardized tests from the start of school (so from the age of four), but a new education law (2023) requires schools to only use observation tools for the first two years (Ministerie van Algemene Zaken, 2024).

At the end of (on average) eight years of primary education, Dutch children take the CITO-test, or since 2023 the 'doorstroomtoets' [transition test] (Ministerie van Algemene Zaken, 2024). Children must complete four multiple-choice tests to measure their language and calculating-mathematics capabilities. As a result of the test, each child receives a score, ranging from 500 to 550, which corresponds with a secondary school level.

Since the renaming of the test in 2023, the test only functions as a ‘second opinion’ on top of the teacher’s advice. According to the website of the Dutch government, the transition test functions as a reliable tool for fitting secondary school advice, an objective ‘second opinion,’ and one of the measurements indicating the quality of schools (Ministerie van Algemene Zaken, 2024).

#### *Points of discussion regarding assessment in Dutch primary education*

In the past years, there has been increasing turmoil about this transition test, as many children taking this test experience a lot of stress and pressure (Pointer, 2023). Moreover, there is no country in the world where children are being tested and selected at such a young age, as in the Netherlands (Heij, 2021). The transition test is more and more seen as a symbol of the early sorting of children in the Dutch education system, which also affects the equality of opportunity (HUMAN, 2020; Boonstra, 2023).

The value placed on the outcomes of tests, as well as the number of tests (an average of 102 tests per year at secondary schools (Visser, 2023)) is relatively high in the Dutch education system (Mellink, 2016; Heij, 2021). In an article from Didactief and CITO (2014), Scheerens explains that this is a logical consequence of the amount of ‘school autonomy’ that Dutch schools have. Here, Scheerens refers to Article 23 – Freedom of Education -, which has been part of the constitution since 1917. According to this law, everyone could potentially establish a school (e.g., with their own religious focus or pedagogic way of working), as long as they adhere to governmental quality standards (De Rooy, 2018). As a result, these national quality standards are based on school results (outputs) rather than ideology or a common educational purpose (inputs) (Boonstra, 2023; Biesta, 2012).

Therefore, Dutch education has been narrowed down to one that can be measured with objectively assessed summative tests (Heij, 2021), which forces schools to prioritize measurable language and math skills over other essential qualities, such as creativity, social skills, understanding of democracy, or personal development (Biesta, 2012; Didactief & CITO, 2014).

Based on the increasing debates and discussions on assessment, The Education Council of the Netherlands (Onderwijsraad) wrote an advisory report with the main focus on primary and secondary education (2019). They conclude three things: firstly, there is too little room for formative assessment due to the great emphasis on high-stakes testing. Secondly, Dutch education predominantly uses quantitative testing, which means that qualitative testing methods are not sufficiently addressed.

Finally, there is a lot of emphasis on standardized testing, which means that there is too little ownership and expertise locally (in the educational institutions themselves) (Ministerie van Onderwijs, Cultuur en Wetenschap, 2019).

### **Back to the purpose: assessment as an opportunity for learning**

A balanced testing and assessment practice requires a clear vision from both the central government and educational institutions (Ministerie van Onderwijs, Cultuur en Wetenschap, 2019). However, according to Scheider et al, the Dutch educational assessment system lacks alignment and fails to recognize the full potential of students (2023).

Biggs introduced the term 'constructive alignment' in 1996. The goal of constructive alignment is to consciously establish connections between learning activities, learning objectives, and evaluation. This triangular relationship should lead to a more aligned education that supports students' learning by making sure that all aspects of education work towards the same objective (Biggs & Tang, 2020).

However, according to Loughlin et al. (2020), the step-by-step simplicity can also decrease its usefulness as an educational tool when imposed top-down for accountability purposes or when utilized as a quality assurance tool. They argue that the systematic application of learning objectives and alignment can give the appearance of quality control, even though this is not really the case (Loughlin et al., 2020). Moreover, constructive alignment is an output-based principle, whereas for example Biesta (2012) arguments a more input-based approach. This means that educational design starts with defining one or more educational purposes and from there creating the curriculum, the learning goals and lastly, the appropriate means of assessments, which are eventually all in line with the purpose(s) (Biesta, 2012).

According to De Rooy (2023), there are three main 'purposes' laying the foundation of education. A classical tradition, also known as 'Bildung', which emphasizes personal development through cultural transmission. A tradition rooted in a rational, modern, and industrial society, where the focus is on language and arithmetic. And a tradition based on the individualized society, where the emphasis is on developing interdisciplinary skills and where students are primarily expected to demonstrate flexibility. And there is not one 'best' or 'most effective' purpose. In fact, defining the purpose(s) of education will always be highly complex, and there will always be different perspectives (De Rooy, 2023; Visser, 2018).

However, Dutch schools have a unique amount of freedom in determining their own purpose, as well as what and how children learn. Therefore, they also have the individual responsibility to align their assessment methods with educational objectives and curriculum in a way that fully supports the learning of children (Boonstra, 2023).

### 3. Conceptual Framework

The base of the conceptual framework (see 'Appendix A – Conceptual Framework' for a visual) is formed by the three main themes that Schellekens et al. (2021) found during their scoping review of how assessment relates to learning.

The first theme is student-teacher roles and relationships within assessment. Within this theme, five characteristics of assessment supporting student's learning are briefly described. First of all, students who are given the opportunity to learn how to evaluate both themselves and their peers, are more likely to actively participate in assessment practices (1). Moreover, educational assessment refers to a collaborative relationship between students and teachers where they can switch roles (2) and where they can both consistently gather, analyse, and reflect on diverse information sources to track developments and utilize data for better learning (3). Lastly, the importance of developing an understanding of quality assessment and being able to communicate about learning (4) and the teacher being able to adapt to students' needs (5) are regarded as important characteristics of assessment for learning within student-teacher relationships.

The second theme refers to the assessment learning environment. This should be a safe and supportive space that motivates students to participate in the learning process and boosts their self-confidence (6). Additionally, alignment of a program- and classroom-based learning environment in which teaching, learning, and assessment are interdependent (7) is part of the assessment supporting learning.

The third theme is about the educational outcomes of assessment. Assessment is about enhancing students' learning and is aimed at improving student achievement, the level of their work, and the quality of teaching (8). Finally, assessment refers to assessing and rating the accomplishments of students, teachers, and educational schools to make well-informed choices (internal and external accountability) (9).

For this research, it is more meaningful to focus on these themes rather than using the common notions for assessment (summative, formative, AoL, AfL, and AaL) because the descriptions and definitions of different assessment concepts overlap in meaning and are not applied consistently in practice. Moreover, confusion regarding these common notions of assessment could distract the subjects from the focus of the research: how assessment can be used to support the learning of children.

Furthermore, as described in the literature review, Dutch schools have a unique amount of freedom in determining their educational purpose as well as shaping their curriculum and teaching methods due to Article 23, freedom of education (De Rooy, 2018; Boonstra, 2023). Therefore, the 'school's purpose(s)' is also included in this conceptual framework because every Dutch school has to decide for themselves what, how, and why their children learn certain things (as long as they adhere to governmental quality standards). Therefore, it is important to clarify how a school's purpose statement approaches learning before going into the way that assessment supports this process.

## 4. Research Questions and Objective

Based on the problem statement, literature review, and conceptual framework, the following research objective is formulated:

The research objective is to explore how Dutch primary schools with different pedagogical approaches align their assessment methods with their own purposes to support the learning of children.

The main research question is:

How do Dutch primary schools with different pedagogical approaches align their assessment methods with their own purposes in order to support children's learning?

The main question consists of the following four sub-questions:

1. How do Dutch primary schools with different pedagogical approaches choose which assessment methods support different aspects of learning that align with their own purpose(s)?
2. How do Dutch primary schools with different pedagogical approaches establish the relationship between student and teacher within assessment to support children's learning?
3. How do Dutch primary schools with different pedagogical approaches create assessment environments that support children's learning?
4. How do the educational outcomes of assessment in Dutch schools with different pedagogical approaches encourage the learning of children?

The three main themes of sub-questions 2 through 4 are taken from Schellekens et al.'s research on the characteristics of assessment to support learning (2021), as explained in the conceptual framework.

## 5. Research Design

This study is a qualitative, exploratory research. The aim of this exploratory research was to provide an understanding of the nature of a problem and formulate questions for more thorough research (Casula et al., 2020). Moreover, Marlow (2023) emphasizes that exploratory research is a process wherein inductive and deductive activities can take place concurrently or back-and-forth, especially while reviewing the literature and developing the research design.

This research consisted of three different phases. Throughout the whole research desk research has been conducted ('phase 0'). Additionally, assessment experts were questioned during phase 1 and during phase 2A and 2B school administrators and teachers were interviewed. See Table 2 for an overview of the phases with additional information on the selection criteria, number of participants, the sub-research question(s) linked to each phase, and the focus. In the upcoming sections, the data collection methods, sampling methods, data analysis and quality assurance will be explained.

*Table 2 – Overview of data collection & sampling*

Phase	Phase 0 (ongoing)	Phase 1	Phase 2A	Phase 2B
<b>Description</b>	Desk research about 1) Dutch primary school assessment methods in general and 2) the participating primary schools	Interviews with Dutch assessment experts, with extra knowledge on the Dutch (primary) education system	Interviews with Dutch primary school administrators on how they align their assessment methods with their purpose to support the learning of children	Interviews with Dutch primary school teachers on how the assessment methods support the learning of children in practice
<b>Data collection method</b>	Desk research	Semi-structured interviews	Semi-structured interviews	Semi-structured interviews

<b>Selection criteria</b>	1) Only recommended by participants  2) Only documents regarding purpose, children's learning and/or assessment from the participating primary schools	Only Dutch assessment experts who did a study on assessment and/or wrote a book or publication about assessment	A variety of Dutch primary schools, varying in:  1) Pedagogical Approaches  2) Assessment methods  3) Assessment purpose (purposefully stating 'assessment to support children's learning or development' or not)	
<b>Number of participants</b>	-	3	4 different schools  1 administrator per school (4 in total)	4 different schools  1 teacher per school (4 in total)
<b>Sub-research question</b>	All sub-questions.	All sub-questions.	Sub-question 1 (as well as 2, 3 and 4)	Sub-questions 2, 3 and 4 (as well as 1)
<b>Focus</b>	1) Expanding the (online) sources on the purposes and assessment methods used by the different participating Dutch primary schools  2) Broadening knowledge on Dutch primary school assessment methods based on recommendations	1) Understanding how Dutch primary schools are generally (not) aligning their assessment with purpose and supporting children's learning. 2) Understanding how assessment could potentially support children's learning  3) Understanding the process of aligning assessment with purpose	Understanding a variety of primary school's visions and processes on how they align their purpose with assessment methods to support the learning of children	Understanding how a variety of primary schools use different characteristics of assessment to support the learning of children in practice

## 5.1 Data collection methods

For this research, semi-structured interviews were the main data collection method in combination with desk research.

During the first phase, the researcher interviewed three assessment experts. The first goal was to gather more in-depth information about Dutch assessment methods and - systems and their effects on children's learning. The second goal was to gain more understanding of the process of (re-)aligning assessment with a school's purpose. This building of knowledge on the topic, together with additional desk research, has allowed the researcher to add more relevant and in-depth questions to the interviews in the second phase (see appendices C and D with the interview protocols).

During the second phase, four Dutch primary schools with different pedagogical approaches have been researched. The goal was to interview at least one school administrator and one teacher per school (apart from each other) through semi-structured interviews. In the end, the interviewer managed to do this, with the exception of one school where the school administrator and teacher were interviewed simultaneously (due to their availability). The researcher assumed that the school administrator could provide information about the school's purpose statement, as well as their pedagogical approach, and what these state about what and how children learn at their particular school. Moreover, the researcher assumed that the school administrator could give more insight into the process of choosing their assessment methods and how they align these with their purpose and pedagogical approach. Additionally, the researcher assumed that the teacher could provide more information about how different characteristics of the school's assessment (student-teacher roles and relationships within assessment, learning environment; and educational outcomes) support the learning of children in practice.

Throughout the interviews, the researcher found out that both the school administrator and the teacher could provide valuable information about all topics discussed in all the sub-questions. The teachers also knew a lot about the school's purpose and pedagogical approach and how the assessment methods were chosen for example, and the school administrators also had a lot of ideas about which characteristics of assessment to support children's learning were happening in the classroom in practice.

Hence, the school administration and teacher were asked the same questions as opposed to a different set, and they each responded with a viewpoint based on their respective roles and areas of expertise.

Regarding research ethics, all participants have been asked for permission to use the answers given during the interviews as results of the research. Moreover, all participants gave consent to use their full name, without anonymization of the data, in the research.

Finally, desk research ('phase 0') has been conducted continually throughout the research to acquire a deeper understanding of the participating schools' purposes and their pedagogical approaches, what this indicates about the learning of their children, and their assessment methods. Furthermore, the desk research was conducted to expand the (online) sources of different Dutch assessment methods and systems and their effects on children's learning in general.

## 5.2 Sampling methods

Purposive sampling and snowball sampling have been used to select the assessment experts, school administrators, and teachers for this study.

The assessment experts have been selected based on relevant (academic) literature they have written about assessment in Dutch (primary) education (see 'Appendix B – Participants overview'). In this way, the researcher had the ability to choose participants that were most beneficial for the study (Gill, 2020), based on their knowledge, experience and expertise. As some of the selected and contacted assessment experts were not available, snowball sampling was used as well. This was less cost-efficient, but also increased the risk of bias, because the interviewed assessment experts might have recommended experts with a similar vision or background for example.

This is similar to the selection of teachers by asking the school administrators for recommendations, which is also a form of snowball sampling. Here the risk of bias could for example be that a school administrator would ask the teacher who values assessment the most, instead of a teacher with a more 'average' view on assessment.

Initially, purposive sampling was used for the selection of schools (school administrators and teachers) as the aim was to interview a maximum variety of schools with different pedagogical approaches, rather than a sampling of schools which would be representative for Dutch primary education in general. In this way the researcher aimed to get a broad picture of how different Dutch primary schools align their assessment with purpose to support children's learning, which is in line with the explorative character of the study. During this research, a 'regular', a Montessori-, a OGO- and a nature-inclusive Dutch primary school have participated (see 'Appendix B – participants overview') for more information about their different pedagogical approaches and purposes. .

Moreover, the schools varied regarding their assessment methods. Some of the schools used portfolios for example, whereas other schools did not. Most of the schools used IEP as their mandatory LVS, whereas one school used DIA. Most schools made use of method-related tests, whereas one did not. A last example is that one school used self-assessment on a daily basis, whereas other schools barely used this at all.

Lastly, the schools varied regarding what they wrote on their website and/or in their school plans about the purpose of their assessment. Some of the schools wrote that the main goal of their assessment is to support children's learning or development, whereas another school wrote that they use assessment to evaluate the effectiveness of their education. Additionally, another school wrote that they use assessment to gain an accurate picture of the children's development to be able to give them fitting secondary school advice.

### 5.3 Data analysis

The gathered primary data was examined using qualitative data analysis. A combination of (primarily) deductive, as well as inductive data analysis was applicable to this process, which fits the explorative character of this study.

#### **Deductive analysis**

First, all data (from phases 0, 1, 2A and 2B) related to Schellekens et all's nine characteristics of assessment to support children's learning (see '3. Conceptual framework) have been evaluated deductively (answering sub-questions 2 – 4). During this process, there was a focus on comparative data analysis: finding similarities and differences of the (lack of) appearance of the nine themes of Schellekens et al. (2021) when comparing data from different sources. In this study, data from the different schools have been compared, as well as data from the different participant groups (assessment experts, teachers and school administrators). This is a form of data analysis triangulation. Lastly, the four phases of qualitative data analysis described by Verhoeven (2015, p. 293) have been applied to organize this process.

The first step involved examining the ideas presented in the transcriptions. This required a thorough review of the content to understand the key themes and insights offered by the participants. By analyzing the transcriptions, significant patterns, recurring concepts, and unique perspectives that were crucial for the research were identified.

Next, concepts were created and referred to based on the conceptual framework using color coding. This step involved mapping the identified ideas, linking them to one or more of the nine characteristics of assessment to support children's learning and assigning specific colors to different concepts. Color coding allowed for visual differentiation between various themes and made it easier to track how each idea fit within the broader framework.

During the third phase, the ideas were arranged, and connections to the sub-questions of the study were established.

This involved systematically organizing the concepts in a way that directly addressed the research sub-questions. Moreover, data from different sources were compared (comparative data analysis). By doing so, it ensured that each sub-question was thoroughly explored and supported by relevant data from the transcriptions. This structured approach helped in building a cohesive narrative that aligned with the research objectives.

Finally, a discussion per sub-question was included. This discussion synthesized the findings from the previous steps, highlighting how the different ideas and themes interrelated and contributed to the understanding of the main topic.

### **Inductive analysis**

The inductive approach for analyzing qualitative data by Thomas (2006) has been used for analyzing the data linked to sub-question 1. Here follows a short description of this process:

All data files were formatted uniformly in terms of font size, margins, and highlighting key elements. The text was read closely to understand the themes and events. Next, categories were created by identifying general themes and specific categories from multiple readings of the data. The inductive coding involved using actual quotes from the text.

In the overlapping coding phase, it was acknowledged that one text segment could belong to multiple categories and that much of the text might not be relevant. The category system was continuously revised, searching for subtopics and selecting quotations that captured the essence of each category. Categories with similar meanings were combined under superordinate categories.

Finally, a discussion was included. In this discussion, the results from the earlier phases were summarized, emphasizing the connections between the many concepts and themes and how they advanced knowledge of the primary subject.

## 5.4 Quality assurance

The quality assurance of this study is based on the quality assurance indicators in qualitative studies that Lincoln and Guba developed in 1985.

Firstly, the *credibility* of the study was increased with the use of person triangulation – as mentioned in ‘5.3 Data Analysis’. There has been data collected from independent assessment experts, as well as from two different perspectives inside the schools (school administrators and teachers). In this way, the data could be validated through multiple perspectives on the research topic.

Secondly, the risk of *dependability* was lowered by interviewing both a school administrator and a teacher. By interviewing multiple people from one school separately, the reliability of the data has increased, as the results were to some extent consistent.

Thirdly, the risk of *confirmability* has been lowered by asking an external researcher – who is an assessment expert himself – for a second opinion on the interpretation of the research results.

Lastly, there has been attention to the *transferability* of the study. The researcher has provided a descriptive research context of the schools that participated (see ‘Appendix B – Participants overview’), so readers can assess the applicability of the data to other contexts.

In addition to these quality indicators, the researcher has continuously practiced the process of reflexivity to resolve the researcher’s bias. In short, reflexivity refers to the ongoing analysis and justification of how one’s own actions have impacted a study endeavor by qualitative researchers (Dowling, 2006). Some insights from these analyses have been written down in ‘7. Limitations’. The researcher paid special attention to the reflexivity process by considering how she might have been biased throughout the data collecting, before starting with the data analysis.

## 6. Research results and discussion

During this study, eleven interviews were conducted with assessment experts, school administrators and teachers. Together they provided perspectives on how Dutch primary schools with different pedagogical approaches align their assessment methods with their own purpose(s) to support the learning of children.

In short, regarding the perspectives of the assessment experts, Nicole Hanegraaf has expertise in supporting schools with educational transitions ‘starting from the inside’. Moreover, Dominique Sluijsmans published a lot of articles and literature with a focus on formative assessment and Karen Heij recently published a critical PhD about the transition test (or ‘selection test’ according to her PhD). More information on the expertise and published works of the assessment experts can be found in ‘Appendix B – Participants overview’.

The teachers and school administrators were asked a bit about their own ideas regarding the research topic, but mainly represented perspectives of one of the four participating schools where they worked. More information about each’s schools pedagogical approach, purpose(s) and assessment methods, can be found in ‘Appendix B- Participants overview’.

In the first sub-question, the relationship between all the main concepts described in the research objective (‘assessment’, ‘purpose’ and ‘children’s learning) has been explored, with a focus on how schools do (not) align these in reality. Moreover, some of the participants (especially the assessment experts) gave advice on how schools could improve this (process of) alignment. The focus of the remaining sub-questions (2, 3, and 4) was on the extent to which Dutch primary schools adhere to Schellekens et al.’s (2021) description of assessment to support children’s learning.

All the quotes have been translated from Dutch to English by the researcher.

## 6.1 Sub-question 1

*How do Dutch primary schools with different pedagogical approaches determine which forms of assessment support different aspects of learning that are in line with their own purposes?*

### **School's purpose(s)**

All participating teachers and school administrators could clearly describe their school's purpose(s) (or 'mission' or 'vision'), as well as their pedagogical approach. The descriptions of these from the teacher and school administrator working at the same school were largely in line with each other. Moreover, they were largely in line with the purpose(s) and pedagogical approaches as described on the schools' websites and/or in their school guides. This could indicate that education professionals working at schools generally have a common understanding of their own purpose(s). Two examples of how two of the participants described the purpose(s) of their schools are given here:

Well, the vision is 'actually happy children learn more'. So, we really focus on that.

And that's why we find it important that children have the space to move. And the space to choose what makes them happy. (Miranda Scharff, teacher at Montessori Arcade, 15-05-24).

...especially to be completely yourself. Being truly yourself also means that each child has their own development, their own pace. It is not a linear process; it really goes in waves. [...] And spending plenty of time outdoors. (Rianne Spin, school administrator Buitenvijs, 13-05-24).

More information on the schools' purposes can be found in 'Appendix B – Participants overview'.

## **The potential of assessment for supporting children's learning**

According to the participants (from all three groups: assessment experts, teachers and school administrators), assessment in Dutch primary education is generally not seen as a tool to (directly) support children's learning. However, when asking about the possible potential of assessment as an opportunity to support children's learning, the participants were more positive. They pointed out two main ways in which assessment can contribute to this.

First of all, assessment is often seen as a measurement or feedback instrument for teachers to gain insights in the knowledge and skills of their students. Assessment can help to make the learning of children more visible and give teachers information about the knowledge and skills of their students. In this way, teachers can evaluate their teaching continuously and adapt their instruction to the needs of the children, which supports their learning directly. The teacher from the Montessori school said the following about this:

I always tell the children, 'You are taking a test for me, so I can see if I need to adjust something or if I haven't covered something well enough. This way, I know if I need to give you more instruction.' So now they are very relaxed about it. They understand that they are not doing it for themselves and are not focused on grades. I never put a total score on it either. (Miranda Scharff, teacher at Montessori Arcade, 15-05-24).

Secondly, giving children more freedom and ownership in assessment was mentioned several times as a contributing factor to supporting children's learning. Standardized LVS-or method-tests are something that children 'undergo', without choosing it or knowing why it is done. After doing an LVS-test, a child is not suddenly better at arithmetic or language, this is merely a measurement moment. Whereas for example with discussing a portfolio, it is easier, more visible and more fun to reflect together with the student on his or her learning.

When there is more freedom for children's own vision and their growth is visible for themselves, assessment can directly support their learning, as indicated by the teacher from the 'regular' school:

I think if it's a test where there isn't just one correct answer, but where children really have to think hard and can express their own views, it's very beneficial. I'm particularly thinking of something like writing an essay, which is a good example. Not where their work just ends up in the teacher's drawer and nothing more is done with it, but where they have a final product they have thought about thoroughly, can be proud of, and can look at later and say, 'Wow, I made this.' This approach supports learning more effectively because they can see their own growth. (Nicky Dijks, teacher Het Mozaïek, 29-04-24).

The school administrator from the development-based school also expressed her enthusiasm about making children's learning visible. At their school they work with 'development goal cards', which they primarily fill in themselves. This increases their motivation and their interest in their own learning. Increasing children's ownership by making their learning visible can also be done in-class on a public list, as indicated by a teacher of the development-based school:

I wrote down the arithmetic learning goals for this domain on a list. For example, 'I can perform operations up to 10,' 'I can perform operations up to 20,' 'I can perform operations up to 100.' They practice with these goals using worksheets and games. When they are practicing, they draw one line next to their name. When they feel they've mastered a goal, they turn the line behind their name into a cross and move on to the next goal. (Daymi Bakker, teacher Matthieu Wiegman school, 24-05-24).

## **Assessment in practice**

Even though all of the participants did underline the potential of assessment for supporting children's learning, they stated that in reality assessment is often done for other reasons that are not necessarily beneficial for the students. Most of the teachers and school administrators admitted that they sometimes take standardized tests, even though they would rather use more formative assessment, because this is expected from the government or the Inspectorate of Education. The school administrator from the 'regular' school said the following about this:

My heart goes out to formative assessment. Because you can look back and evaluate processes, what you have learned from them, and how you can do better in the future. This is, of course, much more educational than just saying, 'Now you have an 8,' which is just a rigid number. Yet we do that because the government requires it from us, because that is the system we are in. (Erik van Faassen, school administrator Het Mozaïek, 29-04-24).

The school administrator from the development-based school also expressed frustration about the limited amount of information the Inspectorate of Education gains about the children's development, as well as the performance of a school:

It's bizarre that an image, such a snapshot, determines how you, as a school, are rated as insufficient. And that doesn't equate to how you assess children every day based on the texts they write, the poems they create, and the oral conversations that take place in the talking circle. It doesn't account for the sense of self-worth they have gained or the reflective skills they've learned either. (Marlies Pepping, school administrator Matthieu Wiegman school, 24-05-24).

Another problem with standardized LVS-tests is that children are not able to check the results of standardized LVS-tests and learn from their mistakes, which makes the tests a lot less useful. Moreover, the analyses of the results can take too long, resulting in the outdated outcomes, as indicated by multiple participants.

### **Assessment often not in line with purpose(s)**

Even though all schools' purpose(s) are clear, according to the participating teachers and school administrators, there is a discrepancy between school's purposes and their assessment methods. The assessment experts also all expressed that schools' assessment methods are often not in line with their purpose(s) and/or pedagogical approaches. One of the assessment experts describes this as follows:

I see a kind of schizophrenia between what schools actually want and aspire to and the context they are forced to operate in due to the use of tests. In a mission and vision, it's often about the pedagogical relationship and how they want to interact with children, but tests put just that relationship under pressure. Tests are rarely intended to improve the relationship between child and teacher or to help with growth. Instead, tests become a system separate from the teacher and the student, where the student is judged and, at the same time, the teacher as well. If the results are disappointing, the teacher is told they are doing something wrong, making children essentially part of the teacher's performance. (Karen Heij, assessment expert, 06-05-24).

Some of the participants stated that aligning purpose and assessment often goes as far as choosing an LVS that fits their purpose the best. Here, participants from the school with the 'regular' pedagogical approach experienced the least alignment, the nature-inclusive and Montessori school generally experienced some more alignment, and the development-based school was most positive.

Moreover, most of the interviewed schools are using IEP and they did specify that this LVS is more in line with their purpose than CITO, as each child is compared to him- or herself and because IEP is based on 'heart, hands and head' instead of only cognitive abilities. On the other hand, some of the teachers also said that they do not feel like IEP is that much different than CITO and still is a limited, standardized test not in the interest of children's learning.

### **Causes of the lack of alignment**

The lack of alignment in assessment can be attributed to several key factors. The cause mentioned by most of the participants is that the goals of assessment are often not internally decided by the schools themselves. Schools have minimal control over the assessment process and lack the necessary expertise, leading to most assessments being determined externally by the LVS-test-creators and method-makers.

Additionally, education professionals frequently lack awareness of the different functions of assessment. They often confuse assessments meant for accountability purposes with those intended to provide relevant information for teachers or to motivate children to learn and grow. This confusion leads to a mix-up of assessment purposes and methods, as indicated by one of the assessment experts:

And what often happens is that under the guise of being a helpful tool, very important decisions about people are being made. Whereas a test that helps with learning looks very different from a test with which you want to make decisions. (Karen Heij, assessment expert, 06-05-24).

Parental pressure also plays a significant role in this misalignment. Parents want to know how well their children are performing through scores and expect teachers to aim for the highest possible scores at the end of primary school, particularly when their children take the transition test.

Lastly, assessments are often associated with negative experiences, such as stress for both children and teachers. This makes assessment a sensitive topic to discuss within the school team, further complicating efforts to align assessment practices with educational goals.

### **The process of re-aligning assessment with purpose**

Especially the assessment experts could answer questions about how schools with different pedagogical approaches could re-align their assessment with their purposes. All assessment experts would agree that schools should redesign their assessment procedures by starting with their own goals, pedagogical approach and curricula. Instead of using the externally supplied assessment (method-related tests, LVS-tests and the transition test) as a starting point, they should start asking questions like 'What knowledge do you wish the kids to acquire?' and 'What do you hope stays with them?', as indicated by one of the assessment experts:

I am truly convinced that if schools have this first part well-established among themselves—the core, the purpose, the goals—then the method of assessment and teaching becomes much simpler. You don't have to search for all sorts of method tests or things, as you already have a clear idea. 'If we consider this important, what do I want to know first about my students to determine if they are proficient in it? What is the first step they take?' Yes, and that becomes the curriculum you build, that's your design for learning and your design for teaching. (Dominique Sluijsmans, assessment expert, 25-04-24).

It could be that they referred to the 'constructive alignment' principle by Biggs (1996) multiple times, without using that term specifically.

Moreover, half of the participants (including teachers and school administrators) expressed that a 'continuous learning line' ('doorlopende leerlijn' in Dutch) is also helpful with the re-alignment of purpose and assessment.

According to one of the assessment experts, the standardized tests are deeply ingrained in the education system, making it difficult to break away from them and go back to the initial purpose. Schools with a strong pedagogical foundation can more easily distance themselves from this system and develop alternatives, although this requires an intensive investment in the craftsmanship of teachers.

Lastly, the school administrator of the Montessori school expressed how their positive pedagogical approach relates to their attitude towards assessments:

We see 'mistakes' as opportunities to learn and grow. Our motto truly is, 'It's not a failure; it's just not successful yet.' And this applies not only to assessing the students but also to assessing the teachers. (Anita Nijland, school administrator Montessori Arcade, 15-05-24).

### **Brief discussion of the results of sub-question 1**

It is remarkable that all participants agreed on assessment as an opportunity to support children's learning, whereas all of them also recognized that in reality, assessment is often used for other reasons, that are not necessarily beneficial for the students. The school with the 'regular' pedagogical approach experienced the least alignment between their purpose and assessment. This could indicate a connection between having a less strong pedagogical foundation, as mentioned as a cause of lack of alignment by one of the assessment experts

The cause of the lack of alignment between assessment and purpose which was mentioned the most was the lack of ownership that schools have, as many tests are imposed by the government or method-makers.

This results in a top-down approach rather than a bottom-up one. However, this top-down approach to accountability and assessment has diminished the voice and participation of parents and students in shaping education quality, as schools prioritize meeting government standards (Kneyber & Evers, 2013). Furthermore, this bureaucratic accountability hinders education professionals from taking responsibility for their actions and their intended outcomes (Biesta, 2012).

Furthermore, there appears to be a lack of awareness about the different functions and forms of assessment and how to effectively align these, which has also been addressed by Boonstra (2024). One example is the transition test, which is intended to measure the child's knowledge and abilities, profile the child and act as an important accountability instrument. When all of these functions are combined, the transition test ultimately becomes inappropriate for any of these functions (Boonstra, 2024).

Lastly, it is remarkable how much the answers to this first sub-question overlapped, regardless of the school's pedagogical approach or purpose and the different backgrounds of the assessment experts..

## 6.2 Sub-question 2

*How do Dutch primary schools with different pedagogical approaches establish the relationship between student and teacher within assessment to support children's learning?*

### **Actively involved students: self- and peer-assessment**

Most of the participants agree that there is room for improvement regarding the implementation of self- and peer assessment. Half of the participants expressed the importance of increasing the students' feedback literacy to practice self- and peer assessment effectively. Practicing this on a regular base is currently not the norm. According to one of the assessment experts, this is partly due to not realizing the possible potential of using students as source of information for themselves and for each other. However, one of the school administrators explained how their school (development-based approach) does create an environment where students can use each other as source of information:

We always assumed that all areas of development are equally important. This creates a culture where everyone needs to make progress every day, but each person in a completely different area of development. This, in turn, makes people less judgmental. ... We also don't have any level groups. This allows children to assess each other in a positive way. (Marlies Pepping, school administrator Matthieu Wiegman school, 24-05-24).

The interviewed teachers from Montessori and the 'regular' approach indicated they barely practice self- or peer assessment with their children. They do use some form of self-assessment provided in the arithmetic working book of the children where they for example need to rate how well they understand a certain concept or do a 'test yourself' exercise. The teacher from the 'regular' school explained that this is only effective if the teacher also actively guides the children in such self-assessment:

There will always be children who assess themselves on their own, but the majority really need guidance, which is very logical because self-reflection is quite difficult.

So, as a teacher, if you discuss it extensively and talk about its importance and how to do it, then the children will definitely learn something from it. (Nicky Dijks, teacher Het Mozaiek, 29-04-24).

This is in line with the study of Vasileiadou & Karadimitriou (2021) on self-assessment in primary education where they concluded that training and guidance is very important for the integration of effective self-assessment.

The teacher from the nature-inclusive school stated that he uses self- and peer-assessment in his pedagogical practice daily and illustrated this with various examples. However, according to him this is not something that is structurally integrated and differs a lot per teacher.

Lastly, multiple participants expressed their enthusiasm towards portfolios as a form of self-assessment. Most of the interviewed schools are using these to a greater or lesser extent and would like the portfolios to fulfil a bigger role in their assessment methods. One of the school administrators explained her positive attitude towards portfolios:

...I just want the whole school to do it this way eventually. Not because I think it's so important, but because the child builds up something over the years. And we also want to move away from reports and make the portfolio the main topic of discussion. So, what the child provides themselves becomes the guiding factor. (Anita Nijland, school administrator Montessori Arcade, 15-05-24).

### **Collaborative relationship: teacher-centered to student-centered approach**

A more collaborative relationship between teacher and student is a concept that is often mentioned in vision statements written in school guides.

However, in reality it requires a lot of guidance and instruction from the teacher when it comes to handling more freedom and taking a leading role. The teacher from the Montessori school explained that her school has also struggled with this:

Once, we had a very student-centred approach, and we failed. The inspection rated us 'very weak'. They said we did not have a good view of the children's development. We tried to be student-centred, but instead of supporting the children, we were merely following them. (Miranda Scharff, teacher Montessori Arcade, 15-05-24).

All of the participants stated they observe a trend towards more student-centred education. Participants from two of the participating schools mentioned that ownership is an important factor in switching towards a more student-centred approach. Hattie (2013) confirmed that students' feelings of ownership of their learning process lead to more active engagement.

However, one of the assessment experts explained that she also notices some schools moving back towards a more traditional teacher-centred approach:

There is a lot of talk about how children are falling behind in reading and arithmetic these days. And you see a group of schools that are diving even deeper into that system - making it more rigid, an increased focus on testing, and holding the children's hands even more. Meanwhile, there is a movement in society towards a more human- and world-oriented approach. This involves much more collaboration, awareness, and doing things together. And it changes the role you have as a professional. (Nicole Hanegraaf, assessment expert, 11-04-24).

## **Using various information sources to act upon**

In order to track progress and apply the knowledge to support further learning, teachers and students should constantly gather, analyze, and consider data from a variety of sources. Practices consist of low-stakes tests, suitable and helpful criticism, and chances for rehearsal and practice (Schellekens et al., 2021). According to some of the participants, most of this data is coming from the method-related tests as this data can easily be entered in one of the standard LVS's. One of the assessment experts stated the following about the relationship between method-tests and LVS's:

The method-related tests are very similar to the LVS tests. And that is not without reason, because method makers only sell their methods if they prepare well for those tests. And the ambition is to score high on the tests. (Karen Heij, assessment expert, 06-05-24).

Another assessment expert is more positive about more schools using a 'richer data collection', but warns for an increase in (unnecessary) administration:

I am quite positive that the concept of rich data collection is slowly becoming part of teachers' vocabulary. This includes using not only 'measurable' data but also 'noticeable' data. However, teachers are still figuring out what this actually means. Which data should I collect, and which shouldn't I? Which data should I administer, and which not? Because before you know it, you're maintaining a kind of shadow administration, and I believe that should be avoided in all cases. (Dominique Sluijsmans, assessment expert, 25-04-24).

This is in line with the frustration the teacher from the 'regular' school expressed about the number of lists she has to complete for each student in multiple online platforms. She mentioned the student-administration system ParnasSys as an example. In contrary, the teacher from the nature-inclusive school expressed his enthusiasm about this system, as it gives a more complete picture of each child.

He confirmed that the teacher does need to fill in a wide collection of questions, for instance about the child's family situation or whether the child really likes moving or not. In this way he truly has a complete image of each student, and he can adapt his education even more to support each unique child.

Lastly, the teacher from the development-based school clarified why many schools still focus on traditional LVS-tests as main data sources:

As a school, you are given a lot of freedom regarding testing and how you design your education, but you are still measured by traditional standards. As long as the results from the traditional LVS are satisfactory, the Inspectorate of Education can be enthusiastic about the way you achieved those results with alternative systems. However, if the LVS results are insufficient, they automatically disapprove of the rest of the education as well. (Daymi Bakker, teacher Matthieu Wiegman school, 24-05-24).

### **Assessment literacy amongst children**

Half of the participants stated that there is too little attention for assessment literacy among children in Dutch primary education. Moreover, the little assessment literacy that is provided generally is intended to increase the test results, rather than improving children's learning, as indicated by one of the assessment experts:

They do get familiar with tests. Yes. But mainly because this will increase their scores. So, there's a lot of training on tests, practicing with tests. But not with the purpose of supporting their own learning. In fact, they never receive feedback like 'this question is correct' and 'this question is incorrect'. That's not even allowed with LVS tests. (Karen Heij, assessment expert, 06-05-24).

Furthermore, all of the interviewed schools try to decrease the pressure as much as possible when giving children summative tests. They mainly do this by emphasizing that the test results are not used in order to judge the children, but to provide information for the teacher. The teacher from the nature-inclusive school said the following about this:

Before we start an assessment, I always used to say, 'Dear children, this is not about who you are or whether you're doing it right or wrong. It's about me wanting to know what you still need to learn and what I need to do as a result. In other words, did I teach it well? (Wouter Reitsema, teacher Buitenwijs, 13-05-24).

Finally, the school administrator from the nature-inclusive school shared that she hopes that future tests will not only provide valuable information for the coaches, but also for the children themselves. And that the children will be so curious about their progress that they will ask the teachers if they may take a test.

### **Flexibility to change and adapt lessons to individual needs and preferences of children**

Most participants agree that there is a lot of flexibility in Dutch primary education to adapt to students' needs, but that this flexibility is barely used and/or experienced by teachers in practice. One reason for this is the focus on outputs and that some teachers are scared to deviate from the prescribed program as the students' test results might be lower.

Moreover, three teachers also mentioned that too much deviation can lead to unrest in the classroom. For example, because this can be too unpredictable, whereas most students need a clear structure to support their learning.

One of the assessment experts expressed that the flexibility to adapt to students' needs can only effectively be used if the education professionals are aware of the underlying vision and purpose of their curriculum:

Yes, that flexibility is indeed significant, but it's not perceived as such because they still have too little grasp of the purpose and the curriculum. If they deepen their understanding about this, they will dare to apply that flexibility and think, 'Well, if I see that my students need something else now, I'll take the time for that instead of rushing to lesson three just because it needs to be done this week.' (Dominique Sluijsmans, assessment expert, 25-04-24).

The school administrators from the 'regular' and nature-inclusive school explained that they both experience freedom in adapting their lessons to the (individual) needs of students. However, according to the school administrator from the 'regular' school, this freedom is limited as they do have to follow the SLO learning goals. Whereas the school administrator from the nature-inclusive school described the SLO learning goals as a starting point, where the rest of their curriculum is built around

### **Brief discussion of the results of sub-question 2**

Regarding the first theme of Schellekens et al. (2021) – the relationship between student and teacher – there seems to be room for improvement in Dutch primary education. Again, the results from the assessment experts and the participants from the different schools were largely in line and barely contradicted each other. However, it is remarkable that the participants from the 'regular' school recognized the least characteristics of assessment to support children's learning regarding this theme.

The potential of self- and peer-assessment for children's learning is recognised by most participants. Self- and peer-assessment lead to more student ownership, which Hattie (2013) describes as one of the most important factors for effective assessment. More specifically, some of the participants expressed their enthusiasm towards portfolios, which was also confirmed by the ongoing desk research.

For example, according to a study by López-Crespo et al. (2021) where the effects of summative assessment and portfolios were compared, the students' self-efficacy and engagement increased more through the use of portfolios. Another example is the research of Tiwari and Tang (2003), which showed that positive academic results were obtained from portfolios, as well as an improved interest in learning for those students who lacked motivation before.

Secondly, all of the participants identified a shift from a teacher-centred to a student-centred approach. Thirdly, using various information sources to act upon appeared to be harder in practice, for example due to the possible extra administrative burden. Fourthly, there seems to be a lack of children's assessment literacy to support their learning. Instead, some schools advice children on how to score higher on tests and make an effort to relieve stress by highlighting the fact that the tests are done to provide feedback for the teachers.

Finally, most participants agreed that there is a lot of flexibility in Dutch primary education to adapt to students' needs, but that this flexibility is barely used and/or experienced by teachers in practice. According to Eerkens (2017), this is also explainable by the effects of the 'lumpsum financing', which is an educational law introduced in 2006 for Dutch primary education. According to this regulation, a school's funding largely depends on the number of students a school has. Therefore, attracting students suddenly became crucial for school boards. One of the consequences was an increased fixation on the numbers that ensure a high ranking on school lists, as well as a good result from the Inspectorate of Education.

### 6.3 Sub-question 3

*How do Dutch primary schools with different pedagogical approaches create assessment environments that support children's learning?*

#### **Safe and supportive assessment environment that encourage children's learning**

At three of the schools, apart from the 'regular' school, there is a lot of attention for creating safe and supportive assessment environments, according to the participants. At the Montessori school for example they have a library, an atelier and a big coop with chickens at the roof terrace which all arouse children's curiosity. And at the nature-inclusive school they try to actively encourage the children in the learning process by adapting the physical space to the theme they are working with:

We work with themes, each lasting six weeks and focusing on a particular subject.

For this period, it's about flying, so we create an enriching learning environment. We ensure there are plenty of activities, things that captivate them in terms of indoor and outdoor setups. It stimulates and excites them to get involved. We also involve the children, so they can come up with their own activities or bring items from home related to the theme. (Rianne Spin, school administrator Buitenvijs, 13-05-24).

Moreover, they use a method called 'De Stad van Axen', to support their children in recognizing their behaviour and giving words to their feelings. This contributes to the feeling of social safety.

However, both the interviewed teacher and school administrator of the 'regular' school were not so positive about the safe and supportive environment at their school. They think this is partly due to a challenging student population with various (unsafe) backgrounds, but also because there are for example little opportunities to display children's work or re-arrange the classroom.

The other two assessment experts also expressed a more critical attitude towards this theme.

One of them pointed out that according to recent research by the OECD (from PISA), Dutch children are generally happy, but do score low on the feeling of being competent and having confidence about their own learning. Additionally, the other assessment expert addressed that this feeling of incompetence and insecurity could be related to the Dutch assessment system. She thinks standardized testing from the age of six and hearing that you belong to the weakest of the group could enhance a feeling of an unsafe learning environment. And this could also be the case for teachers feeling unsafe, as they might be judged based on bad testing results.

The teacher and school administrator from the development-based school explained that they find it important to prevent these feelings of incompetence and unhelpful comparing:

We really want to be a safe place where children can learn, instead of feeling like you need extra support three times a week just because you're not reading fast enough.  
(Daymi Bakker, teacher Matthieu Wiegman school, 24-05-24).

Some children do find it comforting to know what they've scored, and they actually thrive on that. And we do focus on growth. So, it's not about comparing with others, but it's perfectly okay to know what score you had before as long as you know what score you have right now and how your development is progressing. (Marlies Pepping, school administrator Matthieu Wiegman school, 24-05-24).

### **Aligned learning environment: aligning teaching, learning and assessment**

There was not much correspondence from most teachers or school administrators when they were asked about this theme. The teacher of the nature-inclusive school did state that trying to align teaching, learning and assessment is the core of his profession and he did not fully understand why this was a question.

The school administrator from the nature-inclusive school also stressed how much she values this alignment and that the starting point should be clearly formulating all learning goals.

All assessment experts expressed stronger and more critical attitudes towards this topic. They were for example critical about what school's assessment methods exactly measure:

I believe that in most schools, people are very capable of creating a pleasant climate that fosters learning. But then the question arises: how do you measure the outcome of that? Do you measure it through the system, using tests? (Nicole Hanegraaf, assessment expert, 11-04-24).

Are you gathering information about your curriculum and whether you are taking the right steps in teaching children what you want them to learn? (Dominique Sluijsmans, assessment expert, 25-04-24).

Karen Heij expressed that she does think that schools do their best to align their teaching with their assessment, in order to achieve high scores. But aligning this with children's learning is forgotten in this process. This assumption can be illustrated with a quote from the school administrator of the 'regular' school about how this alignment is going in practice:

If the results in all groups are disappointing in terms of spelling, then you focus on that for the next six months. So, you revisit the teaching methods for spelling. And we summarize this in so-called quality cards. So, if the teacher is unsure about the teaching methods for spelling, they can refer to these quality cards. (Erik van Faassen, school administrator Het Mozaïek, 26-04-24).

In this explanation the results of assessment, the period until the next test and the effect on adapting the teaching is explained. But the learning of children is not mentioned.

### **Brief discussion of the results of sub-question 3**

In this section, Schellekens et al.'s (2021) theme of 'assessment environments' has been explored. The participants were generally positive about the safe and supportive assessment environments that Dutch primary schools provide. Only the 'regular' school experienced more difficulty with ensuring safe and supportive environments, due to various reasons.

Furthermore, the 'regular' school works the most with standardized teaching methods. According to one of the assessment experts this could decrease the ownership experienced by the teacher, also leading to taking less initiative in creating a supportive learning environment.

## 6.4 Sub-question 4

*How do the educational outcomes of assessment in Dutch schools with different pedagogical approaches encourage the learning of children?*

### **Assessment to enhance children's development and achievements**

According to all assessment experts and the teacher and school administrator from the 'regular' school, assessment is not (directly) used to enhance children's development and achievements in Dutch primary education. They stated that assessment is (too) often used to evaluate how a whole class, cohort or school scores. Although teachers are meant to use test findings to improve their instruction, the teacher from the "regular" school stated that in reality, they rarely have the time or flexibility to act on the data, thus the children barely benefit in the end. When the school administrator from the 'regular' school was asked if their assessment methods were intended to enhance student's development and achievement, he responded the following:

Well, we have protocols for testing, so to speak. But they are more focused on ensuring that the tests are done well. In the classroom itself, we also create an atmosphere and environment where the children can do those tests well. So, it's separate; it's quiet during the test. The preparation is good, so the teacher explains what the test looks like, takes them through it. And then, at some point, the children can work hard. (Erik van Faassen, school administrator Het Mozaïek, 26-04-24).

On the contrary, the teachers from the other schools expressed that their assessment is clearly intended to improve the children's development, for example with the feedback or 'feedforward' they gain from tests.

The school administrator from the Montessori school concluded that it depends on the kind of test. Summative tests are for objective measuring and provide information for teachers, for herself as administrator and for accountability purposes. Whereas for example portfolios are directly aimed at enhancing children's development and achievements.

They help the children to celebrate their achievements and increase their motivation to further develop. The school administrator from the nature-inclusive school concluded something similar:

Yes, I think that when children become owners of their learning process, it enhances the learning. When it's teacher-driven, it's not the same because they have no say in it themselves; they just have to take a test at a certain point, and that's that. It's something they undergo. But when they have more choice in it and are much more aware of its purpose; when they reverse it so that it's about what they can master and demonstrate in some way, whether through a portfolio or providing evidence, then I believe that process stimulates and supports learning. (Rianne Spin, school administrator Buitenwijs, 13-05-24).

### **Assessment for internal and external accountability**

All participants agree on the fact that assessment plays a big role in internal and external accountability. There were mentioned several positive and negative sides of this.

First of all, assessment can offer valuable input for the internal evaluation of the quality of a school. This could be about the school as a whole or about how a certain class and the main teacher or coach is functioning. Assessment results can for example be a good starting point for the biannual school evaluations with the whole team or as input during a conversation between a teacher and an internal supervisor ('IB'er' in Dutch) about the teacher's performance. The teacher from the nature-inclusive school emphasized the importance of objective data from assessment for internal accountability:

If you rely solely on intuition and your own personal observations, there will still be blind spots. At my previous school, everyone thought the language education was going well. But I discovered through taking an objective test that spelling and reading instruction in all the other units were not being done properly.

So, that was the driving force for me to become a language coordinator. I started providing instruction to all my colleagues on how to teach language skills and you can see that afterwards, the results start to improve again. Because initially, I was able to objectively determine the issue. (Wouter Reitsema, coach Buitenvijs, 13-05-24).

The participants also expressed their understanding about the importance of external accountability. However, multiple teachers stated that even in the event that the students of their school learned a lot and were happy, the Inspectorate of Education would nonetheless award them an insufficient rating based on the poor performance on standardized tests. This limited way of measuring the quality of a school and external accountability is something one of the assessment experts is also frustrated about:

You can see quality in practice and in the actions of the professional. And if you have to measure output at all, I think it's about 'is the intended purpose being fulfilled here?' And how do we demonstrate that? I find that especially important for society and also for the Inspection. But it needs to be in a comprehensive palette, not just those standardized tests, and that's the situation now. (Nicole Hanegraaf, assessment expert, 11-04-24).

#### **Discussion of the results of sub-question 4**

In this last section, the theme of 'educational outputs' of Schellekens et al. (2021) has been researched. Regarding 'assessment to enhance children's development and achievements', the results were mixed. The assessment experts and participants from the 'regular' school stated that assessment is not (directly) used to enhance children's development and achievements. On the contrary, the teachers from the other schools expressed that their assessment is clearly intended to improve the children's development.

Moreover, two of the school administrators concluded that different forms of assessment have different functions, as Boonstra (2024) also implied. Tests with more student ownership, like portfolios, are often more directly aimed at enhancing children's development and achievements. Hattie (2013) also mentioned 'student ownership' as one of the most important factors of effective assessment.

Lastly, the responses about assessment for internal and external accountability purposes were mixed. Most of the participants were positive about the internal accountability function of assessment. Especially when this is used as a starting point for internal evaluation and as it provides more objective data. However, the participants were less positive about the external accountability function of assessment. Some of the education professionals stated that the Inspectorate of Education judges their education quality too much based on results from standardized tests, which is a limited way of measuring educational quality. This is in line with some of the criticism Heij (2021) expressed about the transition test, as well as Biesta's (2012) concern on focussing education too much on what is 'most measurable'.

## 6.5 Other relevant observations

Almost all of the participants' responses have been connected to a sub-question by the primary themes that emerged from the data analysis. However, there are a few additional observations that the researcher felt were notable and significant to the study.

Even though a broad definition of assessment was given during the first part of the interview (see the interview guides in appendices C and D), and the researcher reminded most participants at least once or multiple times about the definition used throughout the rest of this research, the teachers and school administrators mainly talked about traditional, summative tests. This could indicate that the participating teachers and school administrators have less knowledge and/or experience with using other forms of assessment or perhaps that they assess less value to these forms. Another explanation could be that the Inspectorate of Education focusses their attention (initially) on the outcomes of the standardized, summative tests, so the teachers and school administrators are more familiar with answering questions about these types of assessments.

Moreover, the participants mainly talked about how to assess language and/or arithmetic, unless the researcher directly asked a follow-up question about how they would use assessment for other parts of the curriculum. This could have similar reasons.

Finally, it was remarkable that the school administrators were in most cases more optimistic about to what extent their assessment currently is in line with their vision and supporting children's learning as compared to the teachers. This could be because it is harder to align teaching, learning and assessment in practice in the classroom than it sounds on paper.

## 7. Limitations

This study has several limitations that need to be taken into account. First, the interview guide was relatively long for the interview time available, so during a few of the interviews the researcher had to rush a little bit through some of the questions where she would have liked to ask more follow-up questions.

Additionally, the participants may have taken slightly varied interpretations of certain concepts (like 'assessment' or 'purpose'), even though the researcher did explain what was meant with these concepts in the context of this research. Perhaps this was mostly the case with the assessment experts, as they are more used to working with their own definitions and (academic) background in education. This could have influenced the consistency of the responses, which might have affected the findings' overall reliability.

Thirdly, certain phrases may have lost meaning once the Dutch quotes were translated into English, even though the researcher followed up with the participants to confirm that the quotes retained their meaning. However, most of the respondents expressed beforehand they were less comfortable speaking English. Therefore, doing the interviews in English might have further constrained the results.

Fourthly, in order to improve the accuracy, credibility, validity, and transferability of this research, it would have been helpful to have conducted a member check. The researcher did check the descriptions of the assessment experts' expertise (see 'Appendix B – participant list') with the experts. However, due to time constraints, it was not possible to share the research results with the participants to gain their feedback.

Fifthly, contextual factors, for example that the 'regular' school is based in a more challenging neighbourhood and has a big variety of students with different backgrounds, could have influenced the results, rather than their pedagogical approach or purpose.

It is also important to note that the participating schools do not fully represent the pedagogical approach with which they were labelled. The Matthieu Wiegman School, for instance, employs a development-based approach; however, other schools that follow a similar approach most likely have a different curriculum and different methods of assessment. Therefore, more schools should be included for each pedagogical approach to improve comparisons between schools. Nonetheless, one school per pedagogical approach was adequate given the exploratory nature of this study.

Lastly, the researcher would have preferred interviewing more different schools with a wider range of assessment methods. Even though each participating school did have their own ways of working (with assessment), three out of the four schools worked with IEP for example. And none of the schools used CITO as their LVS, even though this is used by many schools in the Netherlands. However, the combination of interviewing both school administrators and teachers, as well as assessment experts did offer a range of perspectives on the research topic, which was one of the goals of this explorative research.

## 8. Conclusion

In conclusion, this research has revealed significant insights into how Dutch primary schools with varying pedagogical approaches align their assessment methods with their educational purposes to support children's learning. Despite the common understanding among participants that assessment should support children's learning, practical application often diverges from this ideal, primarily due to external pressures and lack of ownership. The development-based-, nature-inclusive- and Montessori school showed better alignment between their assessment purposes and methods, as well as more characteristics of assessment to support children's learning (Schellekens et al., 2021), while 'regular' schools faced greater challenges in these regards.

A recurring theme was the impact of top-down approaches to assessment and accountability, which overshadow the involvement of students, teachers and school administrators in shaping educational quality. As a result of this limited autonomy, the development of more effective, bottom-up assessment in line with school's own purposes and pedagogical approach, is hindered. Additionally, a widespread lack of awareness about the various functions and forms of assessment further complicates efforts to align assessment with educational goals.

Regarding Schellekens et al.'s (2021) characteristics of assessment to support children's learning, participants highlighted the potential of self- and peer-assessment to foster student ownership and improve learning outcomes. However, practical challenges, such as increased administrative burdens and insufficient assessment literacy among children, often limit the effective implementation of these methods. The flexibility within the Dutch primary education system to adapt to students' needs exists in theory but is rarely experienced in practice, due to various reasons.

The research also identified mixed outcomes regarding the use of assessment to enhance children's development and achievements.

While some schools effectively utilize assessments like portfolios to promote student development, others, particularly the 'regular' school, do not directly use assessment for this purpose. Concerns about relying too much on standardized tests for external accountability were common, reflecting criticisms of focusing too narrowly on measurable outcomes instead of looking at the overall quality of education.

To address these challenges and enhance the alignment between assessment practices and educational purposes to support children's learning, the following recommendations were proposed by the assessment experts:

According to Nicole Hanegraaf and Dominique Sluijsmans, schools should be empowered with assessment autonomy. Schools should be encouraged to choose and/or develop their assessment methods based on their educational goals and curriculum, reducing reliance on externally imposed standardized tests and increasing constructive alignment.

Moreover, according to Karen Heij, assessment literacy should be increased. Professional development for teachers and school administrators should be provided to enhance their understanding of different assessment functions and how to effectively integrate them into their teaching practices.

In summary, while there is a broad consensus on the ideal role of assessment in supporting learning, practical implementation in Dutch primary schools with different pedagogical approaches is limited by external pressures, lack of ownership, and systemic constraints. Greater alignment between assessment practices and educational purposes requires addressing these challenges, promoting bottom-up approaches, and enhancing assessment literacy among educators and students.

## 9. Relevance

Throughout this research numerous sources (both during interviews and in literature (Heij, 2021; Mellink, 2016; Ministerie van Onderwijs, Cultuur en Wetenschap, 2019)) stated that the current emphasis on high-stake, standardized testing in the Netherlands is moving away the focus from assessment that supports children's learning. Therefore, it is relevant that this research explored how Dutch primary schools can (re-)align their assessment methods with their own purposes to put the interest of the children and their learning processes in the centre again.

Hence, this research is relevant for anyone working in the education sector wishing to refocus attention on the children's interests and/or more purposefully use assessment in education to do so. This could for example be school administrators, teachers, policy makers or activists. This research could provide them with an increased sense of urgency, more understanding about (the effects of) assessment or (concrete) inspiration about how assessment could be used more meaningfully and in line with a school's purpose.

Moreover, almost all participants have underlined (without the interviewer asking about it) the relevance and importance of this topic and that there still is much room for improvement. Most of their confirmation came from their expertise with assessment and working experience in education, but also from personal experiences and motivations.

## 10. Contribution to the PP

This research has revealed significant insights into how Dutch primary schools with varying pedagogical approaches align their assessment methods with their educational purposes to support children's learning. Despite the common understanding among participants that assessment should support children's learning, practical application often diverges from this ideal.

Two of the main causes of the lack of alignment between educational purposes and assessment are a lack of ownership and a lack of awareness about the different functions of assessment. According to the assessment experts, addressing these challenges requires enhancing assessment literacy among educators and students, as well as more bottom-up assessment autonomy.

These research results, as well as an exploration into useful resources and platforms currently available, inspired to the idea of an online assessment library. In the online assessment library, education professionals can find a collection of assessment methods that are more in line with schools' purposes and more in the interest of children's learning. Moreover, they can find a range of informative sources about assessment, for example about the different functions of assessment and which forms could be suitable, as well as legal information about the Dutch education system. Inspiring narratives and interviews about more purposeful assessment (for example with school administrators) will be shared as well. Finally, the concept for the online assessment library contains a number of useful tools that could assist schools in realigning assessment and purpose and open up the conversation about this topic.

Together with Operation Education the researcher has developed a video where the concept for an online assessment library is presented. This video will be helpful to share this first concept with stakeholders and gather more feedback, before the online assessment library will actually be developed.

## 11. References

Apple, M. W., Biesta, G., Bright, D., Giroux, H. A., McKay, A., McLaren, P., Riddle, S., & Yeatman, A. (2022). Reflections on contemporary challenges and possibilities for democracy and education. *Journal of Educational Administration and History*, 54(3), 245–262. <https://doi.org/10.1080/00220620.2022.2052029>

Au, W. (2007). High-Stakes testing and curricular Control: a qualitative metasynthesis. *Educational Researcher*, 36(5), 258–267. <https://doi.org/10.3102/0013189x07306523>

Baarda, J., De Geus, M., & Reinders, L. (2020). *De praktijk van online leren in het Gooi*. <https://www.voion.nl/media/3728/de-praktijk-van-online-leren-in-het-gooi-augustus-2020.pdf>

*Basisschool Buitenwijs in Zwolle*. (n.d.). Retrieved 12 June, 2024, from <https://www.onderwijstransformeert.nl/basisschool-buitenwijs>

Biesta, G. J. J. (2012). *Goed onderwijs en de cultuur van het meten: ethiek, politiek en democratie*. Boom Lemma uitgevers.

Biesta, G. J. J. (2013). *Beautiful risk of education*. Routledge. <https://doi.org/10.4324/9781315635866>

Boonstra, C. (2024, January 29). *De doorstroomtoets is een hoog - laag sorteermachine*. Retrieved June 11, 2024, from <https://www.linkedin.com/pulse/de-doorstroomtoets-een-hoog-laag-sorteermachine-claire-boonstra-735dc/?trackingId=F07toN9OROSwODGHoBC5Hg%3D%3D>

Boonstra, C. (2023, June 19). *Dag van het praktijkonderwijs - Waarom we de dingen doen, zoals we ze doen?* [Video]. YouTube. Retrieved February 27, 2024, from <https://www.youtube.com/watch?v=hQJVJg4CeyY>

Boonstra, C., & Verpalen, C. (2019). *Het onderwijsvragenboek: Waarom doen we de dingen zoals we ze doen?* Amsterdam University Press.

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32(3), 347–364. <https://doi.org/10.1007/bf00138871>

Casula, M., Rangarajan, N., & Shields, P. (2020). The potential of working hypotheses for deductive exploratory research. *Quality and Quantity*, 55(5), 1703–1725.  
<https://doi.org/10.1007/s11135-020-01072-9>

De Rooy, P. (2023). *Een geschiedenis van het onderwijs in Nederland*. Nieuw Amsterdam.

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, 18(1), 105–115.  
<https://doi.org/10.1037/h0030644>

Didactief & CITO. (2016). Toetswoede: Rondetafelgesprek over toetsen. Didactief. Retrieved 01 April, 2024, from <https://didactieffonline.nl/artikel/rondatafelgesprek-over-toetsen-toetswoede>

Dowling, M. (2006). Approaches to reflexivity in qualitative research. *Nurse Researcher*, 13(3), 7–21. <https://doi.org/10.7748/nr2006.04.13.3.7.c5975>

Earl, L. M., & Katz, M. S. P. (2006). *Rethinking Classroom Assessment with Purpose in Mind: Assessment for Learning, Assessment As Learning, Assessment of Learning*. Manitoba Education, Citizenship and Youth.

Eerkens, M. (2017, May 8). *Scholen mogen zelf hun geld uitgeven. Nu weten we dat leraren en leerlingen daar de dupe van zijn*. De Correspondent. Retrieved 11 June, 2024, from <https://decorrespondent.nl/6664/scholen-mogen-zelf-hun-geld-uitgeven-nu-weten-we-dat-leraren-en-leerlingen-daar-de-dupe-van-zijn/ed3133f5-306a-0d30-0eae-2bf2884d2498>

Expeditie leiderschap. (2024, February 12). OE Operation Education. Retrieved February 20, 2024, from <https://operation.education/expeditie-leiderschap/>

Gardner, H. E. (1995). *The unschooled mind: How Children Think And How Schools Should Teach*. Basic Books.

Gezer, T., Wang, C., Polly, A. B., Martin, C., Pugalee, D., & Lambert, R. G. (2021). The Relationship between Formative Assessment and Summative Assessment in Primary Grade Mathematics Classrooms. *International Electronic Journal of Elementary Education*, 13(5), 673–685. <https://doi.org/10.26822/iejee.2021.220>

Gipps, C. (2011). *Beyond testing (classic edition): Towards a theory of educational assessment*. Routledge.

Gill, S. L. (2020). Qualitative sampling methods. *Journal of Human Lactation*, 36(4), 579–581. <https://doi.org/10.1177/0890334420949218>

Hattie, J. (2013). *Visible learning for teachers: Maximizing Impact on Learning*. Corwin.

HBSC. (2021). *HBSC 2021. Gezondheid en Welzijn van jongeren in Nederland*. Retrieved February 27, 2024, from <https://www.trimbos.nl/wp-content/uploads/2022/09/AF2022-HBSC-2021-Gezondheid-en-welzijn-van-jongeren-in-Nederland.pdf>

Heij, K. (2021). Van de kat en de bel: Tellen en vertellen met de eindtoets basisonderwijs. [Doctoral Thesis, Tilburg University].

Het Mozaïek. (n.d.). *Mozaïek in het kort | Kindcentrum Het Mozaïek*. Kindcentrum Het Mozaïek. Retrieved 12 June, 2024, from <https://www.hetmozaiek-lelystad.nl/onderwijs/kc-het-mozaiek-in-het-kort>

Heylen, L. (2022). *De school van je leven: waar elke les een levensles is*. Uitgevrij LannooCampus.

HUMAN. (2020, December 28). *Klassen: Het advies dat bij je past*. 2Doc.nl. Retrieved February 15, 2024, from <https://www.2doc.nl/documentaires/2020/12/klassen-aflevering-5.html>

Kneyber, R., & Evers, J. (2013). *Het alternatief - Weg met de afrekencultuur in het onderwijs!* Uitgeverij Boom. [https://www.boomfilosofie.nl/media/1/het\\_alternatief\\_inkijkex.pdf](https://www.boomfilosofie.nl/media/1/het_alternatief_inkijkex.pdf)

Kneyber, R., Sluijsmans, D., Devid, V., & López, B. W. (2022). *Formatief handelen*. Uitgeverij Phronese.

LAKS. (2021, September 28). *Toetscultuur - Landelijk Aktie Komitee scholieren*. Landelijk Aktie Komitee Scholieren. Retrieved February 15, 2024, from <https://www.laks.nl/projecten/toetscultuur>

Lau, A. (2015). 'Formative good, summative bad?' – A review of the dichotomy in assessment literature. *Journal of Further and Higher Education*, 40(4), 509–525. <https://doi.org/10.1080/0309877x.2014.984600>

López-Crespo, G., Blanco-Gandía, M. C., Valdivia-Salas, S., Fidalgo, C., & Sánchez-Pérez, N. (2021). The educational e-portfolio: preliminary evidence of its relationship with student's self-efficacy and engagement. *Education and Information Technologies*, 27(4), 5233–5248. <https://doi.org/10.1007/s10639-021-10827-2>

Marlow, C. R. (2023). *Research methods for generalist social work*.

Matthieu Wiegman. (n.d.). Retrieved 12 June, 2024, from <https://matthieuwiegman-saks.nl/waarom>

McDowell, L., Wakelin, D., Montgomery, C., & King, S. (2011). Does assessment for learning make a difference? The development of a questionnaire to explore the student response. *Assessment and Evaluation in Higher Education*, 36(7), 749–765. <https://doi.org/10.1080/02602938.2010.488792>

Mellink, B. (2016, January 25). *Waarom maatwerk altijd tot meetwerk leidt*. De Correspondent. Retrieved February 28, 2024, from <https://decorrespondent.nl/3926/waarom-maatwerk-altijd-tot-meetwerk-leidt/0b4e5e99-de2e-0b27-0c27-e5acbb632bc4>

Ministerie van Algemene Zaken. (2024, February 6). *Hoe legt de basisschool de ontwikkeling van mijn kind vast?* Rijksoverheid.nl. Retrieved 30 March, 2024, from <https://www.rijksoverheid.nl/onderwerpen/basisonderwijs/vraag-en-antwoord/hoe-legt-de-basisschool-de-prestaties-van-mijn-kind-vast>

Ministerie van Algemene Zaken. (2023, December 7). *Vernieuwen curriculum primair onderwijs en voortgezet onderwijs*. Toekomst Van Het Onderwijs | Rijksoverheid.nl. Retrieved March 1, 2024 from <https://www.rijksoverheid.nl/onderwerpen/toekomst-onderwijs/toekomstgericht-curriculum>

Ministerie van Algemene Zaken. (2024, February 2). *Waarom maakt mijn kind de doorstroomtoets?* Rijksoverheid.nl. Retrieved February 27, 2024, from

<https://www.rijksoverheid.nl/onderwerpen/schooladvies-en-doorstroomtoets-basisschool/vraag-en-antwoord/waarom-doorstroomtoets>

Ministerie van Onderwijs, Cultuur en Wetenschap. (2019, September 9). *Toets wijzer*.

Advies | Onderwijsraad. Retrieved March 30, 2024, from

<https://www.onderwijsraad.nl/publicaties/adviezen/2018/12/13/toets-wijzer>.

Montessori Arcade. (n.d.). *Daarom zijn wij een goede school*. Retrieved 12 June, 2024, from

<https://www.montessoriarcade.nl/onze-school/daarom-zijn-wij-een-goede-school/>

National Forum for The Enhancement of Teaching and Learning in Higher Education (2017).

*Expanding our Understanding of Assessment and Feedback in Irish Higher Education*. Forum Insights. Retrieved 9 March, 2024, from [95.-NF-2017-Expanding-our-Understanding-of-Assessment-and-Feedback-in-Irish-Higher-Education.pdf](https://teachingandlearning.ie/95.-NF-2017-Expanding-our-Understanding-of-Assessment-and-Feedback-in-Irish-Higher-Education.pdf) (teachingandlearning.ie)

OECD. (2016). Netherlands 2016. In *Reviews of national policies for education*.

<https://doi.org/10.1787/9789264257658-en>

PISA. (2022). *Prestaties Nederlandse 15-jarigen drastisch gedaald*. PISA 2022. Retrieved February 27, 2024, from <https://www.pisa-nederland.nl/resultaten2022/>.

Pointer. (2023, April 3). *Jonge kinderen zitten in de stress over toetsen. Waar komt die druk vandaan?* KRO-NCRV. Retrieved March 1, 2024 from <https://pointer.kro-ncrv.nl/jonge-kinderen-zitten-in-de-stress-over-toetsen-waar-komt-die-druk-vandaan>

Reeve, J., & Tseng, C. (2011). Cortisol reactivity to a teacher's motivating style: the biology of being controlled versus supporting autonomy. *Motivation and Emotion*, 35(1), 63–74. <https://doi.org/10.1007/s11031-011-9204-2>

Scheerens, J., Witziers, B., & Steen, R. (2013). A Meta-analysis of school effectiveness Studies. *Revista De Educacion*, 2013(361), 619–645. <https://doi.org/10.4438/1988-592x-re-2013-361-235>

Scheider, S., Rosenfeld, S., Bink, S., & Lecina, N. (2023). Educational inequality due to lack of validity: A methodological critique of the Dutch school system. *International*

*Journal of Educational Research*, 117, 102097.

<https://doi.org/10.1016/j.jjer.2022.102097>

Schellekens, L. H., Bok, H. G. J., De Jong, L. H., Van Der Schaaf, M. F., Kremer, W., & Van Der Vleuten, C. (2021). A scoping review on the notions of Assessment as Learning (AaL), Assessment for Learning (AfL), and Assessment of Learning (AoL). *Studies in Educational Evaluation*, 71, 101094. <https://doi.org/10.1016/j.stueduc.2021.101094>

Scriven, M. (1967). The methodology of evaluation. In R. E. Stake (Ed.), *Perspectives of curriculum evaluation* (Vol. 1, pp. 39–55). Chicago: Rand McNally.

SLO. (2020). *Kerndoelen primair onderwijs 2006*. Retrieved February 27, 2024, [https://www.slo.nl/publish/pages/17358/kerndoelen-primaireonderwijs2006-overzicht\\_1.pdf](https://www.slo.nl/publish/pages/17358/kerndoelen-primaireonderwijs2006-overzicht_1.pdf)

Sluijsmans, D., & Leenheer, P. (2023). Een herbezinning op de plek van toetsing in het onderwijs. In Sardes (Ed.), *De Nieuwe Meso: Vol. nummer 26*. <https://sluijsmans.net/wp-content/uploads/2024/01/dnm-2023-02-inleidiing.pdf>

Sluijsmans, D. (2020). *Toetsing als kans voor leren*. <https://www.nro.nl/sites/nro/files/media-files/Toetsing-als-kans-voor-leren-formatief-toetsen-en-evaluieren.pdf>

Soderstrom, N. C., & Bjork, R. A. (2015). Learning versus performance. *Perspectives on Psychological Science*, 10(2), 176–199. <https://doi.org/10.1177/1745691615569000>

Stop nu het verplichten van toetsen - Leve het Onderwijs. (2023, November 9). Leve Het Onderwijs. Retrieved February 28, 2024, <https://levehetonderwijs.nl/thema/stop-nu-het-verplichten-van-toetsen/>

Tan, K. (2016). Asking questions of (what) assessment (should do) for learning: the case of bite-sized assessment for learning in Singapore. *Educational Research for Policy and Practice*, 16(2), 189–202. <https://doi.org/10.1007/s10671-016-9196-5>

Taras, M. (2009). Summative assessment: the missing link for formative assessment. *Journal of Further and Higher Education*, 33(1), 57–69. <https://doi.org/10.1080/03098770802638671>

Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation/the American Journal of Evaluation*, 27(2), 237–246. <https://doi.org/10.1177/1098214005283748>

Tiwari, A., & Tang, C. (2003). From process to outcome: the effect of portfolio assessment on student learning. *Nurse Education Today*, 23(4), 269–277.  
[https://doi.org/10.1016/s0260-6917\(03\)00012-1](https://doi.org/10.1016/s0260-6917(03)00012-1)

Vasileiadou, D., & Karadimitriou, K. (2021). Examining the impact of self-assessment with the use of rubrics on primary school students' performance. *International Journal of Educational Research Open*, 2, 100031. <https://doi.org/10.1016/j.ijedro.2021.100031>

Verhoeven, P. S. (2015). Doing research: The hows and whys of applied research (4th ed.). Boom Lemma.

Visser, J. (2023). *Is het voor een cijfer?: hoe jongeren meer kunnen leren met minder stress* (2nd ed.). de Correspondent.

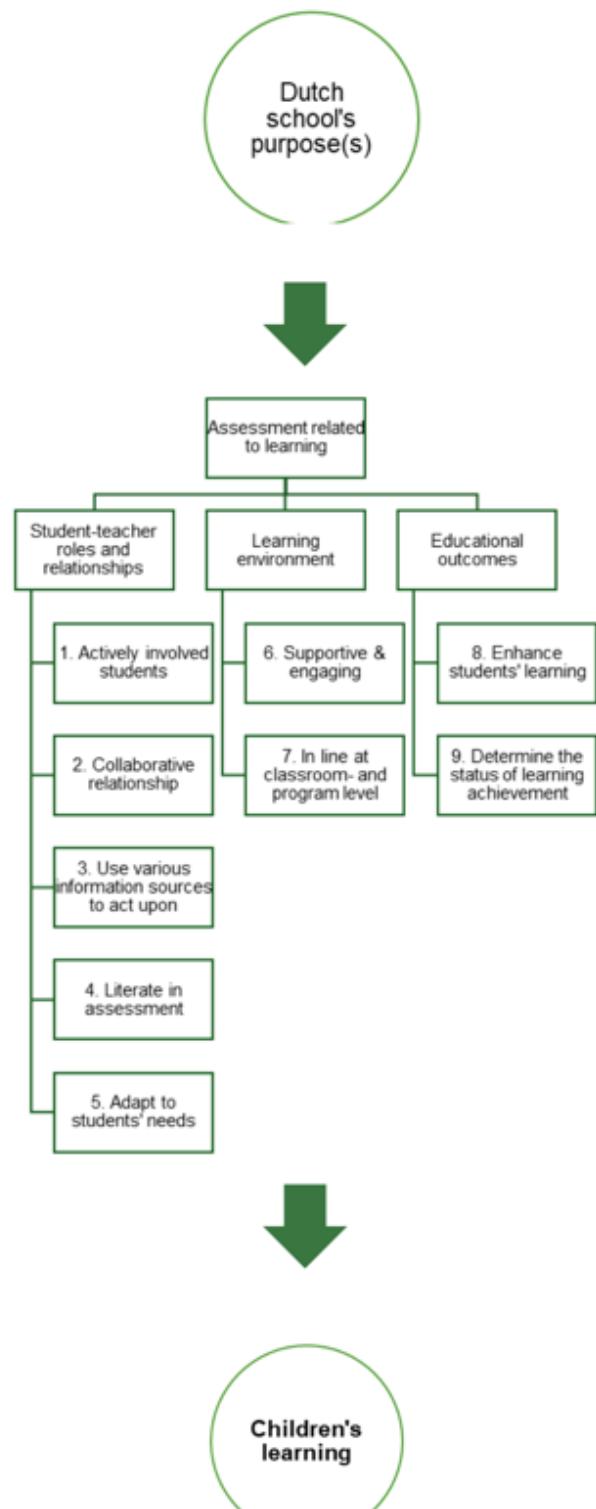
Visser, J. (2018, July 26). *Dit boek vertelt je waarom juist niet alles in het onderwijs anders moet*. De Correspondent. Retrieved February 15, 2024, from <https://decorrespondent.nl/8533/dit-boek-vertelt-je-waarom-juist-niet-alles-in-het-onderwijs-anders-moet/de63b208-e1a9-0b54-25bb-53399b61fc06>

Wouda, J. (2022, March 14). *Welke soorten basisscholen zijn er Nederland en wat is het verschil?* Heutink Voor Thuis. Retrieved 8 June, 2024, from <https://www.heutinkvoorthuis.nl/nl/welke-soorten-basisscholen-zijn-er-nederland-en-wat-is-het-verschil/news/451/>

Zoet, M. (2022, July 16). Zes kansen om het personeelstekort aan te pakken. VPRO. Retrieved March 1, 2024, from <https://www.vpro.nl/programmas/tegenlicht/lees/artikelen/2022/6-kansen-om-het-personeelstekort-aan-te-pakken.html>

## Appendix A – Conceptual framework

(Wassenaar, 2024, based on Schellekens et al., 2021)



## Appendix B - Participants overview

### Assessment experts

Name	Expertise	Published work	Interview technicalities
Nicole Hanegraaf	Involved in improving and creating alternative assessment since 2004. Currently supporting schools with educational transitions ‘van binnenuit’ (from the inside) and implementation of ‘IK Ontwikkel’	Most famous: “Onderwijs van binnenuit” (book)  Own LVS: <a href="#">Borden   IK Ontwikkel PO-SBO-SO-opvang</a>	11/04, online
Dominique Sluijsmans	Educationalist (PhD.) with a focus on formative action, assessment, pedagogy and curriculum design. Researcher, independent educational consultant, speaker and author	Most famous: “Toetsrevolutie” (book)  <a href="#">Publicaties - Dominique Sluijsmans</a>	25/04, online
Karen Heij	Director of Bureau ICE (2002 – 2016). Published PhD in 2021 about the ‘doorstroomtoets’. Currently independent testing expert at Parrhesia Onderwijsadvies.	Most famous: “Van de Kat en de Bel” (PhD)  <a href="#">Mijn publicaties   Parrhesia Onderwijsadvies</a>	06/05, offline

### School administrators and teachers

School	Educational concept	Location	Name	Function	Interview technicalities
Het Mozaïek	‘Regular’	Lelystad	Erik van Faassen	Interim administrator	26/04, offline
			Nicky Dijks	Teacher in ‘groep 2’	29/04, offline
Buitenwijs	More nature-inclusive	Zwolle	Rianne Spin	Founder & administrator	13/05, offline
			Wouter Reitsema	Teacher in unit ‘Waterwijs’	13/05, offline
Montesorri Arcade	Montesorri	Utrecht	Anita Nijland	School administrator	15/05, offline
			Miranda Scharff	Teacher in ‘middenbouw’	15/05, offline
	Ontwikkelingsgericht	Bergen	Marlies Pepping	School administrator	24/05, online

Matthieu Wiegman school			Daymi Bakker	Teacher in 'onderbouw'	24/05, online
-------------------------	--	--	--------------	------------------------	---------------

**More information on school's pedagogical approaches, purposes and assessment methods<sup>1</sup>**

	<b>Pedagogical approaches (according to desk research)</b>
Het Mozaïek	<b>'Regular' (without specific pedagogical approach)</b>
Buitenwijs (nature-inclusive)	Natuurinclusief onderwijs biedt een vruchtbare bodem voor kinderen en jongeren om goed te leren zorgen voor onszelf, voor elkaar en voor de wereld om ons heen. (from: <a href="#">Natuurinclusief Onderwijs: Duurzaam Leren voor een Betere Wereld — Collectief Natuurinclusief</a> )
Montesorri Arcade (Montesorri)	<b>"Het uitgangspunt van het montessorionderwijs is dat een kind een natuurlijke en noodzakelijke drang tot zelfontwikkeling heeft. De pedagogisch medewerker of leerkracht volgt deze ontwikkeling en speelt hierop in door de juiste omgeving en materialen aan te bieden. Op deze manier leert een kind zo veel mogelijk in zijn eigen tempo."</b> (from: <a href="#">Wat is het onderwijsconcept van montessori? - OCO (onderwijsconsument.nl)</a> )
Matthieu Wiegman school (development-based school)	<b>"Ontwikkelingsgericht onderwijs wordt gestuurd door ontwikkelingspatronen, behoeften, motivatie en tempo van kinderen en leerlingen. Iedereen heeft de potentie in zich om zichzelf te ontwikkelen en dit kan leiden tot processen van autonoom leren."</b> (from: <a href="#">Ontwikkelingsgericht onderwijs (OGO) - uitleg (wij-leren.nl)</a> )

	<b>Purpose(s) (according to desk research)</b>
Het Mozaïek	"KC het Mozaïek wil een plaats zijn waar je jezelf mag zijn en waar je ook rekening houdt met de ander. De focus van KC Het Mozaïek ligt op ontwikkeling. Wat ons onderscheidt is dat we in alles authentiek, verschillig, opbouwend zijn." (from their website)
Buitenwijs	"Alle kinderen uitzwaaieren aan het einde van hun schoolloopbaan, vol zelfvertrouwen en zin in het leven. Met een rugzak vol aan eigenzinnigheid, vaardigheden en ervaringen; de ingrediënten om het eigen leven verder vorm te geven. Dat is de missie van Buitenwijs." (from their website)
Montesorri Arcade	"Montessori Arcade biedt kinderen een prettige, fijne plek om te leren. Een plek om jezelf te zijn. Het onderwijs op Arcade bereidt kinderen voor op de toekomst. Zodat kinderen alles in huis hebben om een bijdrage te leveren aan de maatschappij en in het bezit zijn van de juiste kennis en vaardigheden voor het vervolgonderwijs." (from their website)

<sup>1</sup> This information is not necessarily complete and has not been checked or confirmed by the participating schools.

Matthieu Wiegman school	“De Matthieu Wiegmanschool heeft gekozen voor onderwijs waarbij de ontwikkeling van het kind centraal staat. Waarom? Omdat wij geloven dat het de beste resultaten oplevert. Eigen. Samen. Wijzer. Het kan echt.” (from their website)
-------------------------	--

	<b>Assessment methods (according to desk research and interviews):</b>	
	<b>(Mandatory) LVS</b>	<b>Additional assessment methods</b>
Het Mozaïek	IEP	Method-related tests
Buitenvijs	IEP	Some coaches use method-related tests, some not. The same for (administrated) observations. Portfolios. Stad van Axen.
Montessori Arcade	IEP	Most teachers use method-related tests, apart from one group where they experiment with domain-related tests. Some teachers use portfolios, others not, which is the same for observations.
Matthieu Wiegman school	DIA	IK Ontwikkel. Sometimes method-related tests, but often domain-related (self-)assessment.

## Appendix C - Interview guide assessment experts

### Introductie

*“Beste [naam], ten eerste bedankt voor het vrijmaken van je tijd om mee te werken aan dit interview. Ik waardeer het enorm. Ik zal mezelf eerst kort even voorstellen. Mijn naam is Femke, ik studeer ‘Global Project and Change Management’ en ik loop stage bij Operation Education. In samenwerking met hen ben ik ook bezig met mijn afstudeerscriptie. Hiervoor onderzoek ik hoe Nederlandse basisscholen hun toetsing meer in lijn kunnen brengen met hun visie om het leren van kinderen te ondersteunen. Voordat wij beginnen, heb jij nog vragen voor mij?”*

*Ik zou het interview graag opnemen. Is dit oké voor jou?”*

### Deel 1 – Kennismaken met de expert

1. Zou je jezelf kort willen voorstellen?
2. Wat is je achtergrond in werken met thema's als 'beoordeling' en 'toetsing'?
  - a. Welke (specifieke) kennis en/of ervaring heb je op het gebied van toetsing?
3. Wat heeft jou er persoonlijk toe gebracht om expert in toetsing te worden?
4. Waarom vind je het onderwerp toetsing belangrijk?
5. Wat versta je zelf onder 'toetsing'?
6. Wat zijn volgens jou de doelen van toetsing?
7. Hoe kan toetsing volgens jou het leren van kinderen ondersteunen?

### Deel 2 – Verdiepende vragen over hoe scholen doelgerichter kunnen toetsen in lijn met de visie van hun school

*“Tijdens de rest van dit interview wordt met ‘toetsing’ het volgende bedoeld: Toetsen is het continu ophalen van informatie over waar leerlingen staan ten opzichte van de vooraf bepaalde leerdoelen. Hiermee worden dus zowel papieren methode- of LVS toetsen, als observaties, als portfolio’s en alles daartussenin bedoeld. Heb je hier nog vragen over?”*

*Daarnaast zal ik vanaf nu steeds vragen stellen over hoe Nederlandse basisscholen bepaalde dingen doen. Ik begrijp dat elke school anders is en het lastig is om te generaliseren. Toch wil ik je vragen dit te doen, op basis van al jouw ervaring met Nederlandse basisscholen.”*

Vragen gerelateerd aan deelvraag 1 - Hoe beslissen Nederlandse basisscholen welke vormen van assessment verschillende aspecten van leren ondersteunen die aansluiten bij hun eigen doeleinden?:

8. Hoe bepalen Nederlandse basisscholen over het algemeen de doelstellingen van toetsing?
9. Wat is de relatie tussen de purpose/ visie van Nederlandse basisscholen en hoe zij hun toetsmethoden kiezen?

10. Wordt toetsing vaak gezien als een kans om het leren van kinderen in het Nederlandse basisonderwijs te ondersteunen?
11. Hoe kunnen Nederlandse basisscholen beslissen welke vormen van toetsing verschillende aspecten van het leren ondersteunen die in lijn zijn met hun eigen purpose of visie?
12. Ken je voorbeelden van Nederlandse basisscholen waar zij dit doen? Licht toe, alsjeblieft.

### **Deel 3 – bevragen op de negen thema’s rondom ‘assessment to support learning’ van Schellekens**

*“Dit volgende deel is gebaseerd op een onderzoek van Schellekens uit 2021 en vormt de basis van mijn conceptuele framework. Schellekens en de andere onderzoekers hebben geconstateerd dat we veel te veel blijven hangen in allerlei concepten rondom toetsing. Dus wat is precies summatief versus formatief, assessment for, of en as learning, et cetera. En dat we in het definiëren van de toetsvorm soms vergeten wat eigenlijk de achterliggende functie van de toets is. Daarom hebben zij een meta-analyse gedaan waarbij ze 200 studies naast elkaar hebben gelegd. Al die studies hebben onderzoek gedaan naar allerlei verschillende assessment vormen, maar met als gezamenlijk doel om het leren van kinderen te ondersteunen. Zij hebben gekeken wat, los van de soort toetsing, de kenmerken zijn van ‘toetsing om het leren te ondersteunen’. Hier zijn negen kenmerken uitgekomen en ik ga je hier nu op bevragen. Heb je daar nog vragen over, voordat we doorgaan?”*

Vragen met betrekking tot deelvraag 2 - Hoe brengen Nederlandse basisscholen de relatie tussen leerling en leraar binnen de beoordeling tot stand om het leren van kinderen te ondersteunen?:

13. Hoeveel kansen bieden Nederlandse basisscholen kinderen om zelfbeoordeling en beoordeling van klasgenoten te oefenen?
14. Zie je in het Nederlandse basisonderwijs een verschuiving van een leraargerichte naar een meer leerlinggerichte aanpak? Met andere woorden: is de leerkracht meer een soort gids voor het leerproces van kinderen en zijn de kinderen deels partners, in plaats van passieve ontvangers van de beslissingen en handelingen van de leerkracht?
15. Zijn er in het Nederlandse basisonderwijs veel mogelijkheden voor leerlingen en docenten om op consistente wijze gegevens uit verschillende bronnen te verzamelen, analyseren en overwegen om hun ontwikkeling te volgen en de gegevens toe te passen om het leren te verbeteren?
16. Hoeveel tijd wordt besteed aan kinderen wegwijs maken over hoe toetsing kan worden gebruikt om hun leerproces te ondersteunen?
17. Hoeveel flexibiliteit hebben Nederlandse leraren op basisscholen om hun lessen te veranderen en aan te passen aan de individuele behoeften van de kinderen?

Vragen gerelateerd aan deelvraag 3 - Hoe creëren Nederlandse basisscholen beoordelingsomgevingen om het leren van kinderen te ondersteunen?:

18. In hoeverre zijn Nederlandse basisscholen veilige en ondersteunende ruimtes die kinderen motiveren om deel te nemen aan het leerproces en hun vertrouwen vergroten?
19. Hoe goed stemmen Nederlandse basisscholen het lesgeven, leerproces van kinderen en beoordelen op elkaar af?

Vragen gerelateerd aan deelvraag 4 - Hoe bevorderen de onderwijsresultaten van toetsing op Nederlandse scholen het leren van kinderen?

20. In hoeverre proberen beoordelingsprocedures op Nederlandse basisscholen de ontwikkeling te bevorderen?
21. In welke mate speelt beoordeling een rol bij het evalueren en beoordelen van de prestaties van studenten, docenten en onderwijscholen om goed geïnformeerde keuzes te maken? Met andere woorden: hoe speelt toetsing een rol bij het afleggen van verantwoording zowel intern als extern?

### **Afsluiting**

*“Bedankt voor je waardevolle antwoorden. Ik heb genoten van het gesprek. Mag ik je volledige naam gebruiken in de uitwerking of wil je liever dat ik dit anoniem doe? En vind je het leuk om het uiteindelijke onderzoek doorgestuurd te krijgen?”*

## Appendix D - Interview guide school administrators and teachers

### Introductie

*“Beste [Naam], ten eerste bedankt voor het vrijmaken van je tijd om mee te werken aan dit interview. Ik waardeer het enorm. Ik zal mezelf eerst kort even voorstellen. Mijn naam is Femke, ik studeer ‘Global Project and Change Management’ en ik loop stage bij Operation Education. In samenwerking met hen ben ik ook bezig met mijn afstudeerscriptie. Hiervoor onderzoek ik hoe Nederlandse basisscholen hun toetsing meer in lijn kunnen brengen met hun visie om het leren van kinderen te ondersteunen. Voordat wij beginnen, heb jij nog vragen voor mij?”*

*Ik zou het interview graag opnemen. Is dit oké voor jou?”*

### Deel 1 – Kennismaken met de schoolleider of leerkracht

1. Zou je jezelf kort willen voorstellen?
2. Wat is je functie bij [naam school]?
3. Waarom ben je hier schoolleider/ leerkracht geworden?
4. Heb je nog specifieke kennis en/of ervaring, naast je ervaring als schoolleider/ leerkracht? Zo ja, wat?
5. Wat is de visie van [naam school]?
6. Wat versta je zelf onder ‘toetsing’?

### Deel 2 – Verdiepende vragen over doelgericht toetsen in lijn met visie van de school

*“Tijdens de rest van dit interview wordt met ‘toetsing’ het volgende bedoeld: Toetsen is het continu ophalen van informatie over waar leerlingen staan ten opzichte van de vooraf bepaalde leerdoelen. Hiermee worden dus zowel papieren methode- of LVS toetsen, als observaties, als portfolio’s en alles daartussenin bedoeld. Heb je hier nog vragen over?”*

*Daarnaast wil ik benadrukken dat ik hier ben om te onderzoeken en niet om jullie te beoordelen of te bekritiseren. Het zou fijn zijn als je de vragen zo eerlijk mogelijk zou willen beantwoorden op basis van wat er in de praktijk gebeurt.”*

Vragen gerelateerd aan deelvraag 1 - Hoe beslissen Nederlandse basisscholen welke vormen van assessment verschillende aspecten van leren ondersteunen die aansluiten bij hun eigen doeleinden?

7. Wat voor toetsing wordt er bij jullie op school gedaan?
8. Wat zijn de doelen van jullie toetsing?
9. Hoe hebben jullie die doelen bepaald?
10. Is er een relatie tussen jullie visie en hoe jullie je toetsmethoden kiezen? Zo ja, zou je die relatie willen omschrijven?
11. Wordt toetsing op jullie school gezien als een kans om het leren van kinderen te ondersteunen? Zo ja, hoe zorgen jullie ervoor dat toetsing het leren van kinderen ondersteunt?

### **Deel 3 – bevragen op de negen thema’s rondom ‘assessment to support learning’ van Schellekens**

*“Dit volgende deel is gebaseerd op een onderzoek van Schellekens uit 2021 en vormt de basis van mijn conceptuele framework. Schellekens en de andere onderzoekers hebben geconstateerd dat we veel te veel blijven hangen in allerlei concepten rondom toetsing. Dus wat is precies summatief versus formatief, assessment for, of en as learning, et cetera. En dat we in het definiëren van de toetsvorm soms vergeten wat eigenlijk de achterliggende functie van de toets is. Daarom hebben zij een meta-analyse gedaan waarbij ze 200 studies naast elkaar hebben gelegd. Al die studies hebben onderzoek gedaan naar allerlei verschillende assessment vormen, maar met als gezamenlijk doel om het leren van kinderen te ondersteunen. Zij hebben gekeken wat, los van de soort toetsing, de kenmerken zijn van ‘toetsing om het leren te ondersteunen’. Hier zijn negen kenmerken uitgekomen en ik ga je hier nu op bevragen. Heb je daar nog vragen over, voordat we doorgaan?”*

Vragen met betrekking tot deelvraag 2 - Hoe brengen Nederlandse basisscholen de relatie tussen leerling en leraar binnen de beoordeling tot stand om het leren van kinderen te ondersteunen?

12. Hoeveel kansen biedt jullie school de kinderen om zelfbeoordeling en beoordeling van klasgenoten te oefenen?
13. Zie je bij jullie school een verschuiving van een leraargerichte naar een meer leerlinggerichte aanpak? Met andere woorden: is de leerkracht meer een soort gids voor het leerproces van kinderen en zijn de kinderen deels partners, in plaats van passieve ontvangers van de beslissingen en handelingen van de leerkracht?
14. Zijn er bij jullie op school veel mogelijkheden voor leerlingen en docenten om op consistente wijze gegevens uit verschillende bronnen te verzamelen, analyseren en overwegen om hun ontwikkeling te volgen en de gegevens toe te passen om het leren te verbeteren?
15. Wordt er tijd besteed aan kinderen wegwijs maken over hoe toetsing kan worden gebruikt om hun leerproces te ondersteunen? Zo ja, hoeveel en hoe doen jullie dat dan?
16. Hoeveel flexibiliteit hebben leraren op jullie school om hun lessen te veranderen en aan te passen aan de individuele behoeften van de kinderen?

Vragen gerelateerd aan deelvraag 3 - Hoe creëren Nederlandse basisscholen beoordelingsomgevingen om het leren van kinderen te ondersteunen?

17. In hoeverre is jullie school een veilige en ondersteunende plek die kinderen motiveert om deel te nemen aan het leerproces en hun vertrouwen vergroten?
18. Hoe goed stemt jullie school het lesgeven, leerproces van kinderen en beoordelen op elkaar af?

Vragen gerelateerd aan deelvraag 4 - Hoe bevorderen de onderwijsresultaten van toetsing op Nederlandse scholen het leren van kinderen?

19. In hoeverre probeert toetsing op jullie school de prestaties van leerlingen te verhogen?
20. In welke mate speelt toetsing op jullie school een rol bij het evalueren en beoordelen van de prestaties van studenten, docenten en onderwijscholen om goed geïnformeerde keuzes te maken? Met andere woorden: hoe speelt toetsing een rol bij het afleggen van verantwoording zowel intern als extern?

### **Afsluiting**

*“Bedankt voor je waardevolle antwoorden. Ik heb genoten van het gesprek. Mag ik je volledige naam gebruiken in de uitwerking of wil je liever dat ik dit anoniem doe? En vind je het leuk om het uiteindelijke onderzoek doorgestuurd te krijgen?”*