

Mondelinge toetsing in het natuurkundeonderwijs in de bovenbouw van het havo en vwo

Wander Schmitz
Universiteit Twente
Januari 2025

Inhoud

1	Inleiding	4
1.1	Literatuur	4
1.1.1	Randvoorwaarden	4
1.1.2	Literatuur	5
1.1.3	Samenvatting van de bevindingen	8
1.2	Onderzoeksvragen	8
1.2.1	Onderzoeksvragen	9
1.2.2	Theorie	9
2	Methode	12
2.1	Interviews	13
2.2	Toetsen	14
2.3	Lerarenervaringen mondelinge toetsen	15
3	De interviews	16
3.1	Resultaten	16
3.2	De ideale toets volgens deskundigen	18
3.3	Ontwerpeisen voor een mondelinge toets	19
4	De toetsen	21
4.1	Vorbereidingen	21
4.2	Loop van de mondelinge toetsen	22
4.3	Analyse	23
5	Conclusie	26
6	Discussie	29
6.1	Bevindingen uit Duitstalige literatuur	29
6.2	Limitaties en aanbevelingen	32
7	Bronnen	33
8	Appendix	35
8.1	Appendix 1: interviewvragen	35
8.2	Appendix 2: Schriftelijke toets havo	37
8.3	Appendix 3: schriftelijke toets vwo	41
8.4	Appendix 4: vragenlijst lerarenervaringen	45
8.5	Appendix 5: resultaten interviews uitgezet tegen de vragen	46
8.5.1	Respondent 1	46
8.5.2	Respondent 2	47
8.6	Appendix 6: vragenlijst mondelinge toetsen	49
8.7	Appendix 7: foto's mondelinge toetsen	51

8.8	Appendix 8: Beschrijving van elke mondelinge toets	52
8.9	Appendix 9: Ervaringen van de docenten met de mondelinge toetsafname.....	55
8.9.1	Ervaringen van de vakdocent/gespreksleider	55
8.9.2	Ervaringen van de tweede corrector/notulist	56

1 Inleiding

Toen ik op de universiteit Twente een vak over supergeleiding ging afronden, kreeg ik als eindtoets een mondeling. Mijn kennis van de stof werd gedurende een half uur door twee professoren geëvalueerd. De docenten stelden open vragen en lieten mij veel aan het woord. Soms stelden ze sturende vragen, om specifiekere kennis te testen, of om te zorgen dat ik niet te lang over hetzelfde onderwerp bleef praten. Deze manier van toetsen was mij compleet onbekend bij een natuurkundevak. Mondelinge toetsing heb ik alleen meegemaakt bij talen op de middelbare school. Vanzelfsprekend vroeg ik me af waarom dat was; deze toets leek de docenten namelijk een heel goed beeld te geven van mijn begrip van de stof.

Een korte zoektocht in de Nederlandstalige wetenschappelijke literatuur hielp me niet verder opweg. Ik kon geen reden vinden waarom er geen mondelinge toetsen afgenomen worden bij natuurkunde op de middelbare school. Sterker nog, ik kon helemaal geen Nederlandse onderzoeken vinden over mondelinge toetsen bij niet-taalvakken. Hoog tijd om daar verandering in te brengen!

1.1 Literatuur

De literatuur die van toepassing is op mondelinge overhoringen in het natuurkundeonderwijs op de Nederlandse middelbare scholen is schaars. In deze paragraaf beschrijf ik aan wat voor aspecten een stuk literatuur moet voldoen om van toepassing te zijn op dit onderwerp. Dan bespreek ik het meest aangrenzende onderzoek wat ik in het Nederlands of Engels kon vinden. Ik rond de paragraaf af met enkele bijdragen over methodes voor mondelinge toetsing.

1.1.1 Randvoorwaarden

Het onderwerp is vierdelig: mondelinge overhoringen in het natuurkundeonderwijs op de Nederlandse middelbare school. Verander één van die variabelen, en onderzoeksresultaten kunnen niet betrouwbaar worden ingezet om voorspellingen te doen over het onderwerp. Voor mondelinge toetsen verwachten we andere onderzoeksresultaten dan voor niet-mondelinge toetsen, omdat er variabelen spelen die bij andere toetsen niet van invloed zijn, zoals spreekvaardigheid, en non-verbale communicatie.

Het belang van het natuurkundeonderwijs-aspect is minder strikt. Wel is het verschil met mondelinge toetsen bij taalvakken belangrijk, omdat leerlingen daarin onder andere worden beoordeeld op hun gespreksvaardigheid, waar gespreksvaardigheid bij niet-taalvakken een storende variabele is.

Middelbare schoolleerlingen zitten in een periode van sociale neurologische ontwikkeling, waarin hun communicatievaardigheden veranderen. Daarom is het van belang dat literatuurstukken op middelbare schoolleerlingen focussen. Het meest toepasselijk zijn onderzoeken bij leerlingen met dezelfde leeftijd (16-18 jaar oud).

Het feit dat het onderzoek in Nederland gedaan moet zijn, is relevant door de Nederlandse cultuur. Het is aannemelijk dat de verhouding tussen docent en leerling juist bij mondelinge toetsing een sterk effect kan hebben op de prestaties van de leerlingen, en op de beoordeling van de docent. De gemiddelde verhouding tussen leerling en docent is afhankelijk van de

cultuur (Bear et al. (2014), Chen et al. (2019), Xu et al. (2023)), dus onderzoek uit andere landen kan één-op-één aangenomen worden voor de Nederlandse situatie.

Een ideaal stuk literatuur over dit onderwerp zou dus voldoen aan die vier aspecten: mondelinge overhoringen, niet-taaltvakken, in Nederland, op de middelbare school. Zo'n stuk heb ik in het Nederlands of Engels niet kunnen vinden. Sterker nog, de literatuur die ik kon vinden (waarbij het mondeling-aspect natuurlijk een vereiste is) voldeed aan maximaal twee van die aspecten, met uitzondering van één Maltees onderzoek. Hieronder zal ik voor de meest relevante literatuurstukken de context uitleggen, gevolgd door de resultaten die relevant zijn voor dit onderzoek. Daarbij zal ik ook helder maken aan welke van de eerder genoemde aspecten elk van de onderzoeken voldoet.

1.1.2 Literatuur

In deze paragraaf licht ik voor elk stuk literatuur eerst de context uit, omdat de relevantie van het stuk sterk afhankelijk is van de vier aspecten uit de vorige paragraaf. Daarna beschrijf ik de meest relevante bevindingen voor dit onderzoek. In de volgende paragraaf vat ik de bevindingen samen.

Resultatenvergelijking mondelinge en schriftelijke toets

Op een middelbare school in Malta hebben Cauchi en Musumeci (2016) getest of leerlingen betere resultaten behaalden bij mondelinge natuurkundetoetsen dan bij schriftelijke. Het onderzoek betreft 26 leerlingen uit form 4 (14 jaar oud), die mondeling en schriftelijk getest zijn op stof over impuls. Een week na een 30 minuten durende schriftelijke toets over het onderwerp, kregen de studenten een mondeling van 10-15 minuten die heel erg leek op de schriftelijke toets.

De leerlingen scoorden significant beter op de mondelinge toetsen: 41,5%, tegenover de 20,7% op de schriftelijke toetsen. De schrijvers noemen dat het verschil bij slecht scorende leerlingen hoger is dan dat bij goed scorende leerlingen. In een vragenlijst, afgenomen direct na de mondelinge toetsen, geeft 53% van de leerlingen aan comfortabel te zijn met mondelinge toetsing.

De leerlingen mochten hun mondeling in Engels of Maltees doen, waar de schriftelijke toetsen in het Engels moesten. Het is de vraag, hoe een soortgelijk onderzoek in Nederland uit zou pakken, zonder sprake van eventuele invloed van toetsing in verschillende talen op de resultaten.

Mondelinge toetsing van verbale vaardigheden

In Noorwegen worden middelbare schoolleerlingen op 15-jarige en 18-jarige leeftijd mondeling getoetst op hun verbale vaardigheden. Het examen bestaat uit een presentatie, gevolgd door een mondelinge discussie tussen de leerling en diens leraar, in aanwezigheid van een externe examiner. De externe examiner neemt het eindbesluit wat het cijfer betreft. In 'Managing students' insufficient answers in oral examinations' onderzoeken Vonen et al. (2022) hoe leraren reageren als leerlingen tekort schieten in hun antwoorden.

Ze vonden dat leraren normaal gesproken kiezen tussen het najagen van een antwoord en het veranderen van onderwerp. Bij belangrijke vragen gebruiken leraren een paar methodes voor het najagen van een antwoord. Ze stellen bijvoorbeeld specifiekere vragen, geven hints, of moedigen de leerling aan een antwoord te geven. Als leraren besluiten om van onderwerp te veranderen, gebruiken ze strategieën om het gesprek op gang te houden. Dat doen ze bijvoorbeeld door te lachen, of door het antwoord te geven en dan te doen alsof de leerling het antwoord al wist.

Een betoog voor mondelinge toetsing bij statistiek

In “Oral Exams: A More Meaningful Assessment of Students’ Understanding” beschrijft Theobold (2021), een docent statistiek aan een Amerikaanse universiteit, diens ervaring met het afnemen van afsluitende mondelinge toetsen. Theobold heeft diens vak volledig mondeling getoetst (3 toetsen op verschillende momenten in de module), en heeft in het proces geleerd hoe je mondelinge toetsing kan inrichten. Het onderzoeksverslag neemt je mee in het proces van de docent.

Theobold betoogt dat de mondelinge toets een waardevolle manier van toetsing is, om een breed scala aan redenen. De studenten kunnen op een flexibelere wijze hun begrip uitleggen, wat sommige studenten hun cijfers flink ten goede kwam. Bij de laatste toetsen merkte Theobold ook dat de studenten beter waren geworden in het verwoorden van hun denken. Er worden ook een aantal strategieën besproken, die ik in de volgende paragraaf zal uitlichten.

Mondelinge toetsen in groepen voor het vormen van een gemeenschapsgevoel

Tijdens de Covid-pandemie zocht Goodman (2020) een manier om de studenten met elkaar in contact te houden tijdens de lockdowns, evenals een plagiaatbestendige toetsvorm. De docent onderzocht daarvoor het gebruik van mondelinge groepsexamens. De studenten van het biochemievak werden uitsluitend mondeling getoetst.

De onderlinge interactie die studenten voor de voorbereiding en uitvoering van de toetsen moesten hebben zorgde ervoor dat de toetsmethode haar sociale doel haalde. De toetsen gaven de docent ook de kans om directe feedback en ondersteuning te bieden, om het leerproces van de studenten verder te bevorderen. Hoewel de studenten soms stress ondervonden van de nieuwe toetsvorm, gelooft Goodman dat de toetsen ervoor zorgden dat de studenten in een moeilijke tijd op diep niveau kunnen hebben leren.

De invloed van mondelinge toetsvormen op de leerstrategie van studenten

Iannone et al. (2020) onderzochten op een Britse universiteit hoe derdejaars wiskundestudenten hun leerstrategie aanpasten als ze een “high stakes” mondelinge toets moesten doen. De studenten vergeleken hun aanpak bij een schriftelijke toets met gesloten boek, met die bij de mondelinge toets, in een vragenlijst.

De studenten gaven er blijk van meer te hebben ingezet op het ontwikkelen van begrip gedurende de loop van het vak. Dit uit zich in de mate waarin ze meededen tijdens de colleges, maar ook in de manier waarop ze na afloop over de mondelinge toetsen spraken. De studenten

gaven ook aan meer van de informatie die tijdens de mondelinge toets aan bod was gekomen op lange termijn te onthouden, dan bij de schriftelijke.

De gepastheid van mondelinge toetsing bij geneeskunde- of tandheelkundestudies

Rahman (2011) doet in “Appropriateness of using oral examination as an assessment method in medical or dental education” een aantal suggesties om de mondelinge toetsingsstrategieën die traditioneel gebruikt worden in Saudi Arabië te optimaliseren. Hij stelt dat de toetsen, in hun huidige vorm, niet geschikt zijn voor het geneeskunde- en tandheelkundeonderwijs. Hij geeft daarbij niet duidelijk aan waar hij zijn beeld van de huidige vorm vandaan haalt.

Rahman geeft de volgende richtlijnen voor het afnemen van goede mondelinge toetsen:

- Bereid de studenten voor. Geef ze de kans om van tevoren antwoorden te bedenken op algemene en specifieke vragen. Zorg dat ze weten wat ze kunnen verwachten van de structuur van het mondeling, en geef ze eventueel mondelinge oefentoetsen.
- Examinatoren moeten van tevoren weten hoe ze hun taak moeten doen tijdens het mondeling, en wat de criteria voor beoordeling zijn.
- Gebruik meerdere examinatoren, zodat de cijfers minder onderhevig zijn aan vooroordelen over de leerlingen en fraude.
- Examineer op meerdere momenten.
- Vragen moeten duidelijk en helder zijn. Vragen moeten in een paar zinnen gesteld kunnen worden en van tevoren bedacht zijn. De examinatoren moeten een duidelijk idee hebben van belangrijke punten die ze in het antwoord willen horen. In een ideaal scenario worden de vragen getest op collega's.
- Baseer de vragen op scenarios waar de leerling zich iets bij kan voorstellen. Dit kan goed inzicht geven in de kennis, interpretatievaardigheden, probleemoplossingsvaardigheden, en houding. Dat op zijn beurt verhoogt validiteit en betrouwbaarheid.

1.1.3 Samenvatting van de bevindingen

Goodman (2020) en Theobold (2021) schrijven lovend over mondelinge toetsing. Ze noemen de flexibiliteit waarmee studenten hun begrip kunnen uiten en het diepe niveau van leren wat de studenten door de toetsvorm behalen. Naast vakinhoudelijke leerdoelen behaalt de mondelinge toetsing ook vakoverstijgende leerdoelen: studenten worden beter in hun denken verwoorden, en worden in het geval van groepsmondelingen gedwongen tot constructieve sociale interactie. De stof die wordt behandeld bij mondelinge toetsen blijft ook beter onthouden, volgens studenten (Iannone, 2020).

Leerlingen in Malta (Cauchi en Muschemi, 2016) scoren significant hoger voor mondelinge dan voor schriftelijke toetsen over dezelfde stof, maar bij de mondelinge toetsen hadden ze de kans om de toets in het Maltees te maken, waar de schriftelijke toetsen in het Engels waren.

Verder geven Rahman (2011) en Vonen et al. (2022) een aantal tips voor het organiseren en afnemen van mondelinge toetsen.

Relevantie van de literatuurstukken

Zoals beschreven in de paragraaf 1.1.1, zijn er een aantal belangrijke aspecten van dit onderzoek, waaraan de relevantie van literatuurstukken nauw verbonden is. In tabel 1 is voor elk van de literatuurstukken getoond, aan welke aspecten het voldoet.

	Mondelinge toetsing	Middelbare school	Nederland	Natuurkunde (V voor niet-taalkvak)
Cauchi en Muschemi (2016)	X	X		X
Vonen et al. (2022)	X	X		
Theobold (2021)	X			V
Goodman (2020)	X			V
Iannone (2020)	X			V
Rahman (2011)	X			V

Tabel 1: relevantie van elk literatuurstuk.

Tabel 1 laat zien, dat geen van de relevante gevonden onderzoeken in Nederland zijn uitgevoerd. De vraag is dus, hoe de bevindingen in het Nederlandse onderwijsklimaat passen. Verder zijn de onderzoeken die de meerwaarde van de toetsvorm uitlichten uitgevoerd in het hoger onderwijs. Het zal interessant zijn om te zien of diezelfde aspecten ook bij een jongere doelgroep naar voren komen.

1.2 Onderzoeksvragen

In deze paragraaf introduceer ik de vragen die ik met dit onderzoek wil beantwoorden, gevolgd door enkele theorie over toetsing, die gebruikt kan worden om mondelinge en schriftelijke toetsen te vergelijken.

1.2.1 Onderzoeksvragen

De hoofdvraag die ik wil beantwoorden luidt: kan een mondelinge toets een vervanging van een schriftelijke toets zijn?

De vragen die ik in dit onderzoek wil beantwoorden over mondelinge toetsen bij natuurkunde in 4 havo-vwo zijn als volgt:

- 1 Hoe verzamel je zo veel mogelijk informatie over de kennis van je leerling tijdens een mondelinge toets?
- 2 Wat zijn de voor- en nadelen van mondelinge toetsing ten opzichte van schriftelijke toetsen wat betreft evaluatie van vakinhoudelijk begrip?
- 3 Wat zijn de voor- en nadelen van mondelinge toetsing ten opzichte van schriftelijke toetsen wat betreft evaluatie van vaardigheden?
- 4 Hoe verhouden de resultaten van mondelinge toetsing zich tot die van schriftelijke toetsing bij dezelfde leerlingen over dezelfde stof?
- 5 Hoe verhoudt de werklast van de leraar zich bij een mondelijke toetsafname ten opzichte van die bij een schriftelijke afname?

1.2.2 Theorie

Welnu, om verschillende toetsvormen te vergelijken, moeten we hun sterke en zwakke kanten uitlichten. Dat doe ik hier op twee gebieden. Ten eerste op het gebied van de soorten kennis die getest kunnen worden, ten tweede op het gebied van toetsvorminherente storende variabelen.

Soorten kennis

De taxonomie van Bloom is een sterk stuk gereedschap om verschillende soorten kennis georganiseerd te evalueren. In zijn taxonomie categoriseert Bloom verschillende niveaus van begrip, die door Krathwohl verder zijn uitgewerkt in zijn revisie van de taxonomie. (Krathwohl, 2002) Deze revisie, waarmee een completer beeld geschetst kan worden, zal ik gebruiken om de toetssoorten te vergelijken.

Krathwohl deelt de taxonomie op in twee dimensies, met elk een aantal categorieën en subcategorieën. De dimensies zijn te zien in figuren 1 en 2.

Structure of the Cognitive Process Dimension of the Revised Taxonomy

- 1.0 Remember** – Retrieving relevant knowledge from long-term memory.
 - 1.1 Recognizing**
 - 1.2 Recalling**
- 2.0 Understand** – Determining the meaning of instructional messages, including oral, written, and graphic communication.
 - 2.1 Interpreting**
 - 2.2 Exemplifying**
 - 2.3 Classifying**
 - 2.4 Summarizing**
 - 2.5 Inferring**
 - 2.6 Comparing**
 - 2.7 Explaining**
- 3.0 Apply** – Carrying out or using a procedure in a given situation.
 - 3.1 Executing**
 - 3.2 Implementing**
- 4.0 Analyze** – Breaking material into its constituent parts and detecting how the parts relate to one another and to an overall structure or purpose.
 - 4.1 Differentiating**
 - 4.2 Organizing**
 - 4.3 Attributing**
- 5.0 Evaluate** – Making judgments based on criteria and standards.
 - 5.1 Checking**
 - 5.2 Critiquing**
- 6.0 Create** – Putting elements together to form a novel, coherent whole or make an original product.
 - 6.1 Generating**
 - 6.2 Planning**
 - 6.3 Producing**

Figuur 1: de cognitieve procesdimensie (Krathwohl, 2002)

Structure of the Knowledge Dimension of the Revised Taxonomy

- A. Factual Knowledge** – The basic elements that students must know to be acquainted with a discipline or solve problems in it.
 - Aa. Knowledge of terminology**
 - Ab. Knowledge of specific details and elements**
- B. Conceptual Knowledge** – The interrelationships among the basic elements within a larger structure that enable them to function together.
 - Ba. Knowledge of classifications and categories**
 - Bb. Knowledge of principles and generalizations**
 - Bc. Knowledge of theories, models, and structures**
- C. Procedural Knowledge** – How to do something; methods of inquiry, and criteria for using skills, algorithms, techniques, and methods.
 - Ca. Knowledge of subject-specific skills and algorithms**
 - Cb. Knowledge of subject-specific techniques and methods**
 - Cc. Knowledge of criteria for determining when to use appropriate procedures**
- D. Metacognitive Knowledge** – Knowledge of cognition in general as well as awareness and knowledge of one's own cognition.
 - Da. Strategic knowledge**
 - Db. Knowledge about cognitive tasks, including appropriate contextual and conditional knowledge**
 - Dc. Self-knowledge**

Figuur 2: de kennisdimensie (Krathwohl, 2002)

Meestal begint men schriftelijke natuurkundetoetsen met een aantal korte vragen over cognitieve procesdimensies 1 en 2 (onthouden en begrijpen) (fig. 1), gevolgd door een grotere hoeveelheid problemen waarin leerlingen meerdere cognitieve procesdimensies van 1 tot 5 moeten combineren. De toetsen vragen de leerlingen kennis uit kennisdimensies A t/m C (fig. 2). Je ziet zelden vragen waarvoor metacognitieve kennis nodig is.

In een literatuurstudie over mondelinge toetsen, beschrijft Joughin (1998) dat de literatuur ‘verrassend eensgezind’ is, wat betreft de kennis en vaardigheden die getoetst wordt met mondelinge toetsen. Ten eerste: kennis en begrip (cognitieve procesdimensies 1 en 2), waarbij verwezen wordt naar Blooms eerste cognitieve procesdimensie. Ten tweede: toegepast probleemoplossen (‘probleemoplossings-, toepassings-, en interpretatievaardigheden’), die duiden op cognitieve procesdimensies 3 tot 5. Ten derde wordt interpersoonlijke competentie genoemd, en ten vierde intrapersoonlijke competentie. Die laatste is belangrijk, omdat intrapersoonlijke competentie metacognitieve kennis is, wat zelden getest wordt met schriftelijke toetsen.

Mondelinge toetsen worden volgens deze literatuurstudie dus niet alleen gebruikt om alle vier de kennisdimensies te testen, maar kunnen ook worden ingezet voor het evalueren van metacognitieve kennis.

Storende variabelen

Een van de eisen voor een kwalitatief goede toets, is validiteit (*Eisen Aan een Goede Toets*, z.d.). Bij validiteit wordt er onderscheid gemaakt tussen inhoudsvaliditeit en begripsvaliditeit. Inhoudsvaliditeit betreft de dekking van de leerstof. Een toets die de behandelde leerstof op verantwoorde wijze afdekt, heeft een hoge inhoudsvaliditeit. Begripsvaliditeit betreft de mate waarin de toets toetst wat hij moet toetsen. Een veelgebruikt voorbeeld is een geschiedenistoets, waarin gebruik wordt gemaakt van een Engelstalige bron. Het begrijpend lezen van Engelstalige teksten wordt in die toets een zogenaamde storende variabele (*Eisen Aan een Goede Toets*, z.d.).

Storende variabelen zijn bijzonder interessant wanneer je verschillende toetsvormen vergelijkt, omdat elke toetsvorm inherent een set storende variabelen met zich meebrengt. De mate waarin de storende variabelen van invloed zijn, is afhankelijk van de leerling. Zo is de Engelstalige bron in het voorbeeld hierboven niet van invloed op het cijfer van een leerling die vloeiend Engels spreekt, maar desastreus voor een leerling die geen Engels spreekt.

Er zijn een aantal storende variabelen die slechts bij één van de toetsvormen een rol spelen. Bij mondelinge toetsen zijn dat bijvoorbeeld: welbespraaktheid, accent, zelfverzekerdheid, interpersoonlijke competenties, en uiterlijk. Bij schriftelijke toetsen kunnen begrijpend lezen, schrijfvaardigheid, handschrift, dyslexie, en schrijfstijl storende variabelen zijn. Ook zijn er storende variabelen die de twee toetsvormen delen, maar waarbij de storing van verschillende grootte is. Een voorbeeld daarvan is een taalachterstand van iemand met een andere moedertaal. Het kan zijn dat diegene zich beter mondeling uitdrukt, met eventueel doorvragen, maar het kan ook zijn dat diegene liever elke zin rustig uitdenkt en formuleert, en daardoor liever schriftelijk toetsantwoorden geeft. Toetsangst kan ook bij een van de toetsvormen sterker zijn dan bij de andere. Ook kan het zijn dat leerlingen hun leerstrategie aanpassen aan de toetsvorm. Samengevat: toetsresultaten worden altijd vertekend door storende variabelen, en de mate van vertekening is afhankelijk van de toetsvorm en de leerling.

2 Methode

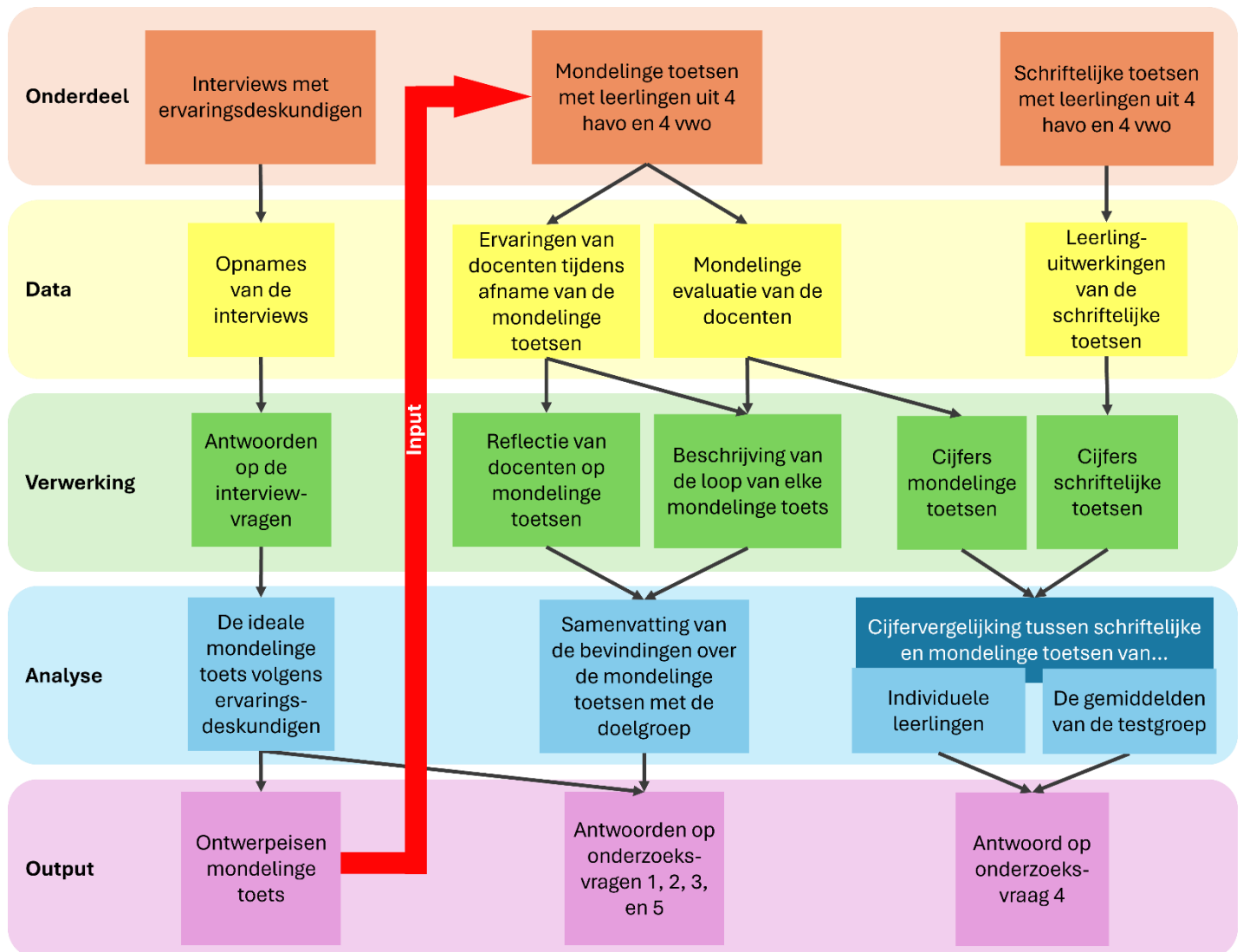
Om de onderzoeksvragen te beantwoorden, wordt de dataverzameling opgedeeld in drie onderdelen. Ten eerste neem ik interviews af met twee docenten die ervaring hebben met het afnemen van mondelinge toetsen over natuurkundestof in Nederland. De interviews met ervaringsdeskundigen zijn bedoeld om antwoorden te vinden op onderzoeksvragen 1, 2, 3, en 5, maar ook om een beeld te krijgen van hoe een goede mondelinge toets in het Nederlandse onderwijsklimaat ingericht is. Ten tweede neem ik bij de doelgroep (4 havo en 4 vwo) mondelinge toetsen af, waarvan de methode mede gebaseerd is op de resultaten van de interviews. Ten derde neem ik bij dezelfde leerlingen schriftelijke toetsen van dezelfde stof af. In tabel 2 wordt voor elke onderzoeksvraag zichtbaar gemaakt welke data gebruikt zal worden.

	Hoe verzamel je zo veel mogelijk informatie over de kennis van je leerling tijdens een mondelinge toets?	Wat zijn de voor- en nadelen van mondelinge toetsing ten opzichte van schriftelijke toetsen wat betreft evaluatie van vakinhoudelijk begrip?	Wat zijn de voor- en nadelen van mondelinge toetsing ten opzichte van schriftelijke toetsen wat betreft evaluatie van vaardigheden?	Hoe verhouden de resultaten van mondelinge toetsing zich tot die van schriftelijke toetsing bij dezelfde leerlingen over dezelfde stof?	Hoe verhoudt de werklast van de leraar zich bij een mondelijke toetsafname ten opzichte van die bij een schriftelijke afname?
Interviews met ervaringsdeskundigen	X	X	X		X
Docent-ervaringen mondelinge toetsen 4 havo en 4 vwo	X	X	X		X
Cijfers mondelinge toetsen 4 havo en 4 vwo				X	
Cijfers schriftelijke toetsen 4 havo en 4 vwo				X	

Tabel 2: onderzoeksvragen gekoppeld aan de data waarmee ze beantwoord worden.

In dit hoofdstuk bespreek ik het onderzoeksplan zoals dat ligt vóór het afnemen van de interviews. Details van de inrichting van de mondelinge toetsen bespreek ik in het hoofdstuk 4, omdat de ontwerpeisen van de mondelinge toetsen volgen uit de resultaten van de interviews met ervaringsdeskundigen.

Een schematische weergave van het onderzoek is te zien in figuur 1. De volgorde van de onderdelen is belangrijk: uit de interviews met de ervaringsdeskundigen volgen ontwerpeisen voor de mondelinge toetsen. Ook moeten de schriftelijke toetsen pas nagekeken worden nadat de cijfers voor de mondelinge toetsen bepaald zijn, om subjectiviteit te minimaliseren bij het evalueren van de mondelinge toetsen. Daarom is de onderzoeksvolgorde: eerst de interviews met de ervaringsdeskundigen, dan de mondelinge toetsen met de leerlingen uit 4 havo en 4 vwo, en tot slot de schriftelijke toetsen.



Figuur 1: Schematische weergave van het onderzoek. De onderdelen worden van links naar rechts uitgevoerd.

2.1 Interviews

De interviews zijn belangrijk, om een beeld te krijgen van de manier waarop mondelinge toetsafname bij natuurkunde er in het Nederlandse onderwijsklimaat uit ziet. Specifiek moet het perspectief van de docent helder worden. Daarom is het belangrijk dat de respondenten ervaring hebben met het afnemen van mondelinge toetsen bij natuurkundevakken in Nederland.

Voor de interviews heb ik in eerste instantie de docenten van de Universiteit Twente benaderd die de mondelinge toets afnamen, die de aanleiding voor dit onderzoek was. Verder heb ik contact opgenomen met natuurkundevakdidactici werkzaam aan de Universiteit Twente. Tussen de benaderde docenten waren er twee, elk met jaren ervaring met mondelinge toetsafname, bereid om een interview af te geven.

Om de dataverzameling zo objectief mogelijk te houden, worden beide interviews afgenomen op basis van dezelfde vragen. Het uitvoeren van een gestructureerd interview is hier echter een gemiste kans, omdat er meer kwalitatieve data gewonnen kan worden uit een semi-gestructureerd interview. In een semi-gestructureerd interview kan er immers doorgevraagd worden, wat de ervaringsdeskundigen de ruimte geeft om een groter deel van hun kennis te delen.

De vragen die ik als leidraad gebruik voor de interviews zijn te vinden in appendix 1. Het doel is om de kennisbasis van de experts zo uitgebreid mogelijk te onderzoeken gedurende het interview, dus waar me dat gepast lijkt zal ik doorvragen.

2.2 Toetsen

Op basis van de interviews worden mondelinge toetsen voorbereid. Het onderzoek is uitgevoerd op de school waar ik zelf les geef in Oost-Nederland. De school werkt met kleine klassen en uren van 85 minuten, met het doel om de leerlingen persoonlijke aandacht te bieden. Er zijn geen tussenuren in het rooster, wat van 9 tot half 5 duurt, en de leerlingen hebben één dag in de week vrij. Ik geef les aan alle vierdeklassers die natuurkunde volgen op deze school. Het zijn 7 leerlingen: drie havo'ers en vier vwo'ers. Zij vormen de testgroep voor de vergelijking van mondelinge en schriftelijke toetsen. De inrichting van de mondelinge toetsen wordt samengesteld na de interviews, zodat er gebouwd kan worden op de expertise van de ervaringsdeskundigen. De schriftelijke toetsen zijn te vinden in appendix 2 en appendix 3.

Bij een onderzoek als dit komen praktische uitdagingen kijken. Ik zal eerst ingaan op de ideale manier om twee toetsvormen te vergelijken. Daarna geef ik een overzicht van de werkelijke omstandigheden.

Ideale testomstandigheden

Om de toetscijfers van verschillende toetsvormen te vergelijken, moeten leerlingen op beide manieren getoetst worden over dezelfde stof, en beide toetsen moeten dezelfde weging hebben. Idealiter zou je leerlingen eigenlijk tegelijkertijd beide toetsen laten doen: zo zou je voorkomen dat de resultaten beïnvloed worden door storende variabelen als de tijd van de dag, de dag van de week, het humeur van de leerling, et cetera. Dat kan natuurlijk niet, maar we kunnen dichtbij komen, door dezelfde stof snel achter elkaar te toetsen. Het ideale scenario is moeilijk vast te stellen: als de toetsen op dezelfde dag zijn wordt vermoeidheid een factor, als ze op verschillende dagen zijn kan de stof wegvallen, of kunnen leerlingen juist nog wat extra gaan oefenen.

Werkelijke testomstandigheden

In dit onderzoek wordt de knoop door externe omstandigheden doorgesneden: er is maar één

manier om het zo te verroosteren, dat de mondelinge toetsen voor alle leerlingen op dezelfde dag vallen. Deze dag is een week na de afname van de schriftelijke toets. De schriftelijke toetsen worden pas na het afnemen van de mondelinge toetsen nagekeken, om onnodige subjectiviteit bij de evaluatie van de mondelinge toetsen te voorkomen.

Ook de cijferweging wordt door externe omstandigheden beïnvloed. De weging is als volgt: de schriftelijke toets is de basis voor het SE-cijfer van het hoofdstuk van straling en energie. Als de leerling voor de mondelinge toets hoger scoort dan voor de schriftelijke, dan is het eindcijfer het gemiddelde van de twee. Deze methode is een compromis met de scholengemeenschap, waarin normaal gesproken alleen een schriftelijke toets gebruikt wordt. Op deze manier is de mondelinge toets summatief, terwijl de leerlingen geen nadeel ondervinden van het experiment.

2.3 Lerarenervaringen mondelinge toetsen

De mondelinge toetsen worden afgenomen door mijzelf en een tweede corrector. De tweede corrector is een docent natuurkunde van een andere school, die ik vertrouw als persoon en als professional. Diegene kent de leerlingen niet, en zal daarom een objectiever perspectief kunnen bieden, dan mijn collega die hen vorig jaar les gegeven heeft. De ervaringen van de docenten die de mondelinge toetsafname uitvoeren in dit onderzoek worden gebruikt voor het beantwoorden van vier van de vijf onderzoeksvragen. Om hun inzichten goed te kunnen vergelijken, worden hun ervaringen door middel van een vragenlijst achterhaald. Die vragenlijst wordt ingevuld na de afname van de mondelinge toetsen. De vragenlijst is te vinden in appendix 4.

3 De interviews

In dit hoofdstuk behandel ik de resultaten van beide interviews. In appendix 5 zijn transcripties van de interviews te vinden. Een van de respondenten neemt mondelinge toetsen af ter afsluiting van natuurkundevakken op een universiteit. De andere respondent heeft niet alleen ervaring met het afnemen van mondelinge toetsen bij natuurkundevakken in de rol van docent en vakdidacticus natuurkunde, maar is in eigen woorden ook geslaagd voor diens eindexamen, doordat diegene mondeling getoetst werd.

Op basis van de interviews presenteer ik een analyse van de ideale mondelinge toets volgens deze ervaringsdeskundigen, in het licht van de eerste drie onderzoeksvragen. Daaruit volgt ook een set richtlijnen voor het organiseren en afnemen van mondelinge toetsen voor docenten, die in hoofdstuk 4 gebruikt wordt om mondelinge toetsen te ontwerpen en af te nemen bij de doelgroep.

3.1 Resultaten

Aan de hand van de interviewvragen uit appendix 1 heb ik met beide ervaringsexperts een semi-gestructureerd interview afgenomen. De bevindingen zijn hieronder samengevat.

Redenen om mondelinge toetsen af te nemen

Beide respondenten noemen dat leerlingen leren tijdens de toets: over de inhoud, over wat ze beter kunnen doen qua leren, en over hun gedachten goed articuleren. Ook zeggen ze dat leerlingen het altijd eens zijn met hun cijfer. Een van de respondenten gebruikt mondelinge toetsen ook om diens universiteitsstudenten beter te leren kennen. Een leraar op de middelbare school zou het op soortgelijke manier kunnen inzetten om de band met diens leerling te verbeteren.

Methoden voor de afname van mondelinge toetsen

Beide respondenten geven aan dat er twee docenten nodig zijn: één als notulist (die ook de tijd bijhoudt), en één als gespreksleider. Dit zorgt voor een afname in subjectiviteit bij de beoordeling. De mondelinge toets kan volgens de respondenten het best ingericht worden als een semi-gestructureerd interview. Een van de respondenten wijdt uit dat 10 minuten genoeg is voor een mondelinge toets op de middelbare school. De leerling moet dan het lokaal verlaten zodat de leraren onderling kunnen nabespreken, wat ongeveer 2 minuten duurt. Daarna komt de leerling weer binnen, en wordt de toets gedurende ongeveer 5 minuten nabesproken. De respondent raadt aan 20 minuten per toets in te roosteren, zodat er wat speling is.

Leerdoelen

De respondenten geven aan dat leerdoelen die te maken hebben met conceptueel begrip goed te testen zijn met mondelinge toetsen. Voor uitgebreide rekenvragen, leesvaardigheid, en dingen waarover de leerling even rustig moet kunnen nadenken lenen mondelinge toetsen zich minder goed, aldus de respondenten. De methodiek van probleemoplossen is wel goed te testen volgens hen, omdat je met de leerling het proces kunt doorlopen. Als de leerling vastloopt, kan

je diegene een beetje hulp geven, om de rest van de methodiek te testen, waar vastlopen bij een schriftelijke toets er meestal voor zorgt dat de hele rest van de toetsvraag verloren gaat.

De voorbereiding van een mondelinge toets

Ter voorbereiding van een mondelinge toetsafname bereiden de ervaringsdeskundigen een breed scala aan vragen over de stof voor, zodat geen twee mondelinge toetsen dezelfde vragen bevatten. Eventueel wordt dit aangevuld door plaatjes, diagrammen, of practicummateriaal gerelateerd aan de stof. Ook brengt een van de respondenten een aantal A4'tjes mee, zodat leerlingen kunnen schrijven of tekenen als ze dat graag willen. Voor de mondelinge toetsen overleggen de ervaringsdeskundigen ook met hun collega de slagingscriteria, en verdelen ze de rollen van notulist en gespreksleider.

Het leiden van het gesprek

De ervaringsdeskundigen gebruiken dezelfde strategieën om het gesprek te leiden. Ten eerste maken ze de leerling duidelijk wat het plan is, en hoeveel onderwerpen ze minimaal willen aanstippen. Dan testen ze eerst of de slagingscriteria behaald zijn, en van daaruit gaan ze wat meer de diepte in, afhankelijk van het niveau van de leerling. Ze geven beiden aan dat je je als docent flexibel moet opstellen en het gesprek een beetje de vrije loop moet laten gaan. Wel is het belangrijk dat een leerling niet te lang over hetzelfde onderwerp blijft praten. Een van de ervaringsdeskundigen begint graag met de vraag wat de leerling het leukst of meest interessant vindt aan het onderwerp.

Het bepalen van een cijfer

De ervaringsdeskundigen hanteren ook hier soortgelijke strategieën. Beide toetsafnemers bepalen onafhankelijk een cijfer, en overleggen daarna. Volgens de deskundigen gaat dat in de praktijk vaak op gevoel, met de slagingscriteria in het achterhoofd. Het maken van een rubric heeft voordelen en nadelen: als voordeel kan je je cijfer goed beargumenteren, maar als nadeel kan je niet alles wat in een mondelinge toets aan bod komt voorspellen met een rubric. Een rubric strak handhaven kan zorgen voor minder flexibiliteit tijdens het mondeling, maar als je hem als algemene richtlijn gebruikt, kan hij nuttig zijn.

De voordeel van een mondelinge toets ten opzichte van een schriftelijke

Een van de ervaringsdeskundigen benoemt dat het verschil in storende vaardigheden kan voor sommige leerlingen (bijvoorbeeld dyslecten) een voordeel zijn. Die leerlingen worden immers niet belemmerd door hun leesvaardigheden. Beide ervaringsdeskundigen zeggen dat je de vragen kan aanpassen op het niveau van de leerling, wat ervoor zorgt dat je beter kan inschatten wat het werkelijke niveau van de leerling is. Als het op een specifiek deel van de vraag fout gaat, kan de leerling een beetje geholpen worden, zodat de examinerator nog informatie over de rest van de vraag kan winnen. Ook verwijzen de ervaringsdeskundigen naar de eerder genoemde redenen om mondelinge toetsen af te nemen: leerlingen leren van de toets, en zijn ze het altijd eens met hun cijfer. Verder wordt er genoemd dat de manier van toetsen veel vrijer is, omdat de interactie tussen leerling en docent ervoor zorgt dat de leerling op diens eigen manier duidelijk kan maken hoe diegene de stof begrijpt. Er is meteen feedback, en leerlingen met een

onvoldoende lopen altijd weg met verbeterpunten. Voor de ervaringsdeskundigen is het ook fijn dat de leraren na de toetsafname geen nakijkwerk hebben.

De nadelen van een mondelinge toets ten opzichte van een schriftelijke

Volgens de ervaringsdeskundigen ervaren leraren mondelinge toetsen als tijdrovender en organisatorisch lastiger dan schriftelijke toetsen. Voor sommige leerlingen zullen de storende vaardigheden die bij een mondelinge toets komen kijken, zoals gespreksvaardigheden en sociale vaardigheden, meer problemen vormen dan die bij een schriftelijke toets. Ook noemt een van de ervaringsdeskundigen dat sommige leerlingen meer stress kunnen ervaren bij mondelinge toetsen. Daar wordt aan toegevoegd, dat het element van stress tijdens een mondelinge toets kan verminderd worden door vroeg in de schoolloopbaan te starten met mondelinge overhoringen. Ook kan het helpen om de leerlingen in de les goed voor te bereiden op mondelinge toetsen, door ze te laten merken dat hun input gewaardeerd wordt. In het verlengde van dit onderwerp raadt de ervaringsdeskundige aan, om de leerlingen van tevoren uit te leggen waarom mondelinge toetsing zo waardevol en leuk is.

De kenmerken van een goed mondeling

Na een goede mondelinge toets zijn leraar en leerling tevreden, aldus beide ervaringsdeskundigen. Ze zijn het eens over het cijfer, onafhankelijk van wat het cijfer is, en de leerling loopt weg met het idee iets geleerd te hebben. Een leerling met een onvoldoende weet waar de verbeterpunten zitten.

3.2 De ideale toets volgens deskundigen

In deze paragraaf bekijk ik de resultaten van de interviews in het licht van de eerste drie onderzoeksvragen.

Efficiënte verzameling van informatie

Uit beide interviews blijkt dat het belangrijk is dat de leerling op diens gemak is. Enkele strategieën daarvoor zijn de leerling voorbereiden op mondelinge overhoringen; de inhoud van het de mondelinge toets bij aanvang goed duidelijk maken; naast de leerling zitten; de leerling pauze of uitstel aanbieden in geval van blackout.

Verder is het flexibel aanpassen van het niveau van de vragen belangrijk, vertellen beide ervaringsdeskundigen. Als leraar is het je taak in de eerste paar minuten van de mondelinge toets uit te vinden waar het niveau van de leerling ongeveer ligt, zodat je vragen de grenzen van de kennis van de leerling onderzoeken. Te makkelijke of te moeilijke vragen geven de docent geen nieuwe informatie over het niveau van de leerling. Een van de respondenten benadrukt dat een leraar met goede vraagstelling een accurater beeld kan krijgen van het niveau van de leerling dan bij een schriftelijke toets.

Mondelinge overhoring van verschillende soorten leerdoelen

Beide ervaringsdeskundigen geven aan dat de mondelinge toets een beter middel is voor het

testen van begrip dan het schriftelijk, omdat de vragen toegespitst kunnen worden op het niveau van de leerling. Daardoor kan er met veel meer precisie worden bepaald waar de grenzen van het begrip van de leerling liggen.

Uit de interviews blijkt dat mondelinge toetsen zich niet zo goed lenen voor rekenvragen of vragen waarbij leerlingen rustig een tijdje moeten kunnen nadenken. Mondelinge overhoringen zijn echter uitermate geschikt om de methodiek van probleemoplossing te testen, aangezien het denkproces van de leerling beter te zien is dan bij schriftelijke toetsen.

Tijdsinvestering van de docent

Een van de ervaringsdeskundigen zegt dat het afhangt van de groepsgrootte, of een mondelinge toets meer tijd kost dan een schriftelijke toets. Diegene noemt ook dat het heel fijn is, dat er na de toetsafname geen nakijkwerk meer is. De andere ervaringsdeskundige geeft aan bij groepen groter dan 20 leerlingen liever schriftelijke toetsen af te nemen, omdat het te lang duurt om die groepen mondeling te toetsen.

3.3 Ontwerpeisen voor een mondelinge toets

Uit de interviews heb ik een aantal ontwerpeisen samengesteld waarnaar de mondelinge toetsen ingericht worden, opgedeeld in de categorieën organisatie, inrichting, gesprekstechnieken. De ontwerpeisen zijn hieronder uiteengezet.

Organisatie

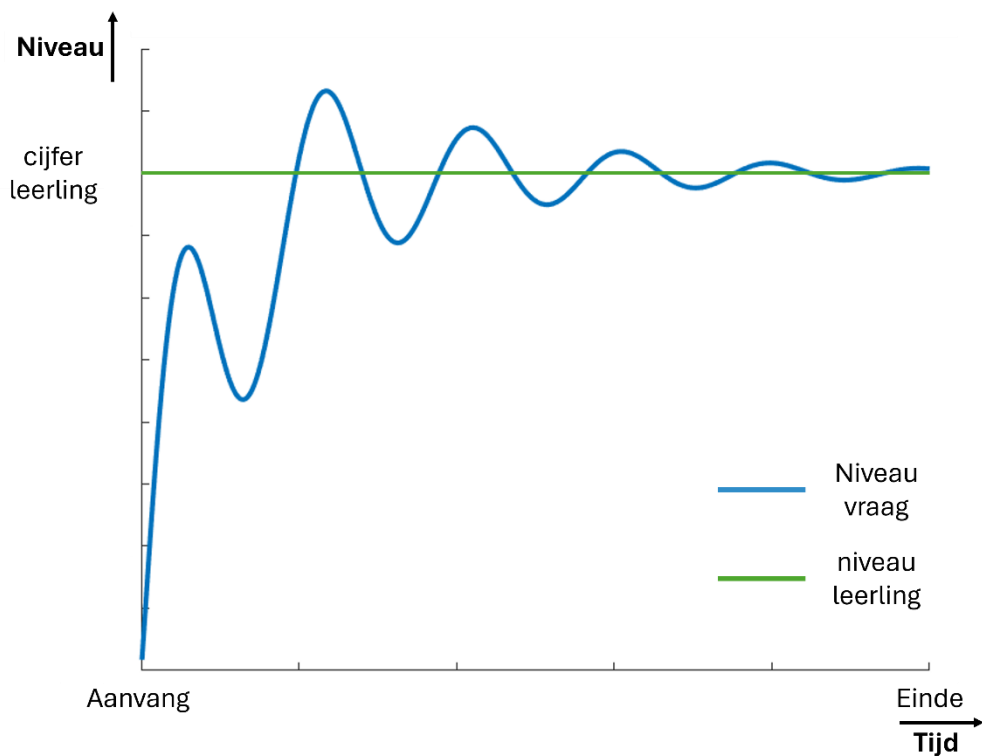
- De mondelinge toetsen worden afgenomen door twee docenten die bekend zijn met de stof, waarvan de een het gesprek leidt, en de ander notuleert en de tijd bijhoudt
- De docenten overleggen van tevoren de slagingscriteria
- De docenten bereiden een aantal vragen en figuren voor, die als gespreksstarters ingezet kunnen worden
- De docenten zorgen voor papier en schrijfmateriaal, die leerlingen tijdens de toets mogen gebruiken om zichzelf te verduidelijken

Inrichting

- De gespreksleider past het niveau van de vragen flexibel aan aan het niveau van de leerling (zie figuur 2)
- Niet alle vragen hoeven uit de vragenlijst te komen
- Het gesprek vloeit organisch van de ene naar de andere vraag
- De notulist geeft één minuut voor het einde van de toets een seintje aan de gespreksleider
- Na afloop van de toets verlaat de leerling de ruimte, en overleggen de docenten over het cijfer
- De leerling wordt weer binnen gelaten, en de toets wordt nabesproken

Gesprekstechnieken

- Het is belangrijk dat de leerling op diens gemak is
- De gespreksleider maakt aan de aanvang van de mondelinge toets goed duidelijk wat de bedoeling is, en stelt de leerling gerust als dat nodig is
- De gespreksleider zit naast de leerling
- De gespreksleider geeft de leerling de ruimte om even pauze te nemen buiten het lokaal als het te veel wordt
- De gespreksleider zorgt ervoor dat alle slagingscriteria getest worden
- De gespreksleider kan de loop van het gesprek onderbreken, als er te lang bij hetzelfde onderwerp gebleven wordt, eventueel met behulp van een nieuwe vraag van de lijst



Figuur 2: de gespreksleider past gedurende de mondelinge toets het niveau van de vragen aan op basis van het niveau van de leerling. Het niveau van de vragen begint laag, om de leerling zelfvertrouwen te geven. Daarna convergeert het naar het niveau van de leerling.

4 De toetsen

In dit hoofdstuk bespreek ik de toetsen die ik bij de doelgroep afgenomen heb. De schriftelijke toetsen zijn gebaseerd op exemplaren uit het archief van de scholengemeenschap. Ze zijn te vinden in appendix 2 en appendix 3. In dit hoofdstuk ligt de nadruk op de mondelinge toetsen, waarbij ik de resultaten van de schriftelijke toetsen als data voor vergelijking presenteer.

Met de ontwerpbeisen uit de interviews heb ik, in overleg met een tweede corrector, een plan samengesteld voor het afnemen van de mondelinge toetsen. In dit hoofdstuk bespreek ik eerst de voorbereiding en de geplande indeling voor de mondelinge toetsen. Dan vat ik de loop van de mondelinge toetsen samen, zoals ervaren door de tweede corrector en mij. Het hoofdstuk wordt afgesloten met een analyse van de bevindingen, in het licht van de onderzoeksvragen.

4.1 Voorbereidingen

In deze paragraaf bespreek ik wat er vooraf ging aan de mondelinge toetsen. Eerst bespreek ik hoe we de mondelinge toetsen inrichten, daarna wat voor materialen we gebruiken.

Materialen

Voor de mondelinge toetsen hebben we een lijst met vragen voorbereid. Ter inspiratie is een combinatie van schoolboeken (Overal Natuurkunde, Stevin), en ChatGPT geraadpleegd. De selectie is verfijnd en aangevuld met vragen van eigen ontwerp. De lijst is te vinden in appendix 6. Verder zijn er enkele foto's van scenario's omtrent energie die de leerlingen uit de lessen bekend voor komen geprint. De foto's zijn te vinden in appendix 7. De leerlingen krijgen een paar lege A4'tjes om evt. op te tekenen of schrijven, met een pen, potlood, gum, en geodriehoek. Er ligt ook een BiNaS, die ze mogen raadplegen.

De tweede corrector, een externe natuurkundedocent, en ik hebben mondeling kortgesloten wat de stof was en globaal slagingscriteria besproken.

Inrichting

Voor elke leerling plannen we 20 minuten in, per advies van respondent 2. De tweede corrector is notulist en houdt de tijd bij. We zitten aan drie tafels in een T-vorm, de notulist met computer tegenover de leerling, en ik schuin naast allebei, per advies van respondent 1.

De eerste halve minuut tot minuut gebruiken we om de leerling uit te leggen wat de bedoeling is. Dan mag de leerling kiezen met welk onderwerp diegene wil starten, waarna we een willekeurige vraag van de lijst over dat onderwerp vragen. Vervolgens voeren we een gesprek over de stof, waarin ik steeds vragen stel die aangrenzend aan het gespreksonderwerp zijn, om de breedte of diepte van de kennis van de leerling te onderzoeken. Als een leerling te lang bij een onderwerp probeert te blijven plakken, wissel ik het onderwerp.

Rond het midden van de 10 minuten van de toets stel ik een willekeurige vraag van de lijst over het andere hoofdstuk, waarna we het proces herhalen. Een minuut voor het eind geeft de tweede corrector me een signaal, waarna mijn volgende vraag is of de leerling nog iets kwijt wil.

Na afloop van de toets gaat de leerling weg en bespreken de tweede corrector en ik in twee minuten de loop van het mondeling, en een indicatie van het cijfer. We halen de leerling terug en houden een korte nabespreking (5 minuten), waarin we de toets en het verwachte cijfer met de leerling bespreken. Daarna begint de volgende toets.

4.2 Loop van de mondelinge toetsen

In deze paragraaf bespreek ik de loop van de mondelinge toetsen, zoals ervaren door de docenten (de externe tweede corrector met de rol van notulist, en ik met de rol van gespreksleider). Daarbij stip ik een aantal nieuwe inzichten aan die we gedurende het proces hebben opgedaan. Ik zal steeds naar de docenten refereren als 'wij'. Daarna presenteer ik de cijfers van de mondelinge en schriftelijke toetsen.

De beantwoorde vragenlijsten van de lerarenervaringen zijn te zien in appendix 8, en beschrijvingen van de loop van elk mondeling in appendix 9.

Inrichting

De inhoudelijke planning van de mondelinge toetsen werd goed gevolgd: we hadden beiden het idee dat de leerlingen op alle slagingscriteria accuraat getest waren. De leerlingen gaven na de nabespreking allemaal aan dat ze het eens waren met de cijfers, wat dat idee van de docenten versterkte. Ook de tijdsindeling bleek accuraat: we hadden beiden het idee dat er voor elke fase van de toetsprocedure genoeg tijd was gepland. Van de start tot het einde van de procedure bleek 18-20 minuten gemiddeld een goede inschatting.

Het flexibel aanpassen van het niveau van de vragen bleek een uitdaging, het afnemen van een mondelinge toets blijkt een hele specifieke vaardigheid te zijn. We merkten allebei dat we van elke toetsafname veel leerden. Daarvoor was de ingeplande 2 minuten speling zeer nuttig: we bespraken steeds wat verbeterpunten, die we in de volgende toetsafnames konden implementeren.

Evaluatie

Het evalueren van de mondelinge toetsen hebben we als makkelijker ervaren dan de uitvoering. We waren het steeds snel eens over welk cijfer de leerling moest krijgen. We beloonden leerlingen meer voor vragen waarbij ze het goede antwoord met meer zekerheid brachten. Als er wat doorvragen nodig was om ze tot het goede antwoord te laten komen, kregen ze een lagere evaluatie. We spraken de notities door, en besloten het cijfer als een gemiddelde van hoe goed de leerling gedurende de toets de vragen beantwoordde. Daar zat enige mate van *fingerspitzengefühl* bij, maar de docenten werden het onderling steeds eens, en de leerlingen konden zich tijdens de nabesprekingen ook steeds in hun cijfer vinden.

Bij het becijferen van de mondelinge toetsen hebben we ook enigszins gecompenseerd voor het karakter van de leerling. Zo kreeg een leerling die door mij in de klas als onzeker ervaren werd wat minder puntenaftrek voor onzekere antwoorden. Een leerling die vol overtuiging foute antwoorden gaf, kreeg echter minder pluspunten voor zelfvertrouwen bij goede antwoorden.

Overige opvallendheden

We kregen het idee met veel meer nauwkeurigheid het begripsniveau van de leerling te kunnen bepalen. Niet alleen dat, het denkproces van de leerling kwam heel goed in beeld, op een manier die we bij schriftelijke toetsen nooit gezien hebben. Dat was niet alleen fijn als aspect om een cijfer mee te bepalen, maar ook als informatie die ik als vakdocent kon gebruiken om mijn lessen naar in te richten in de toekomst. We hadden het idee dat deze toetsvorm ons veel meer feedback gaf over wat de leerlingen van me geleerd hadden, dan een schriftelijke toets.

Tijdsinvestering

Omdat mondelinge toetsen niet met veel leerlingen tegelijk kunnen, ben je veel tijd kwijt aan het afnemen. Dat is moeilijk in te roosteren. Het is een best vermoeiend proces voor de leraren, dus je kan niet een hele werkdag vullen met mondelinge toetsen. Omdat je ongeveer drie mondelinge toetsen per uur af kan nemen, schat ik dat ik er 10 tot 15 op een dag kan doen, waarvan de helft als gespreksleider en de andere helft als notulist. Er moeten wel een of twee flinke pauzes tussen zitten.

Met een jaarlaag van 60 leerlingen ben je als leraar dus wel een werkweek zoet, en in die week vallen al je lessen uit, dus dat moet je in je jaarplanning verwerken. Dat geldt ook voor de tweede corrector. Mondelinge toetsen in groepen afnemen zou kunnen helpen, maar het doet ook af aan de individuele aandacht voor de leerling. Dat kan de versterkende werking die de mondelinge toetsen kunnen hebben op de band met de leerling verkleinen. Ook kan de groepsdynamiek een effect hebben op de prestaties van de leerlingen. De mate van dat effect is een goed onderwerp voor vervolgonderzoek.

Cijfers

De cijfers van de leerlingen zijn uiteengezet in tabel 3. Ze worden geanalyseerd in de volgende paragraaf.

	Mondelinge toets	Schriftelijke toets
Leerling 1	6	6,5
Leerling 2	4,8	5,9
Leerling 3	3	3,3
Leerling 4	7	8,4
Leerling 5	6,5	5,4
Leerling 6	8	6,5
Leerling 7	5,1	2,5
Gem. $\pm$ std.	5,8 $\pm$ 1,5	5,5 $\pm$ 1,9

Tabel 3: leerlingcijfers voor de mondelinge en schriftelijke toetsen

4.3 Analyse

In deze paragraaf bespreek ik de bevindingen van de mondelinge toetsen in de context van de onderzoeksvragen.

Efficiënte verzameling van informatie

Tijdens de mondelinge toetsen hebben we de gesprekstechnieken voor efficiënte

informatieverzameling uit de interviews toegepast. Een leerling die achteraf aangaf niet zo op diens gemak te zijn geweest tijdens het mondeling, scoorde meer dan een punt lager dan op diens schriftelijke toets. Dit ondersteunt de uitspraken van beide ervaringsdeskundigen, dat het van belang is dat de leerling op diens gemak is.

Gedurende de mondelinge toetsen hebben we kunnen doorvragen, en hebben we het niveau van de vragen kunnen afstellen op het niveau van de leerling. Door het inzetten van die technieken kregen we het idee dat we heel nauwkeurig het niveau van de leerling konden bepalen. Het feit dat de leerlingen zich konden vinden in hun cijfers, ondersteunt dat idee.

Mondelinge overhoring van verschillende soorten leerdoelen

Door middel van doorvragen kunnen we met hogere nauwkeurigheid testen waar de grenzen van de kennis en het begrip van de leerling ligt. Het feit dat de leerling je meeneemt in diens denkproces, stelt de leraar in de gelegenheid om dat denkproces in elke fase te evalueren, in plaats van alleen de stappen die de leerling op papier zet. Ook kan de leerling verder geholpen worden als diegene vastloopt, wat ervoor zorgt dat latere stappen in het denkproces ook geëvalueerd kunnen worden.

We merken echter wel, dat er geen tijd is om een wijd scala aan ingewikkelde rekenproblemen te behandelen. Er is ook weinig tijd voor leerlingen om rustig na te denken, voor één leerling leidde dat tot een lager cijfer.

De cijfers

Om die vraag te beantwoorden zijn in tabel 4 de cijfers uiteengezet. We zien dat de cijfers gemiddeld 1,2 punten van elkaar verschillen, met een standaarddeviatie van 0,7 punt. Je zou kunnen overwegen om de data van leerling 7 niet mee te nemen, omdat het een uitschieter is vergeleken met de andere data. Dan is het ook logisch om de data van leerling 3 niet mee te nemen, omdat bij hem een andere tweede corrector aanwezig was. In het geval waarin je hun data weglaat, is het gemiddelde verschil 1,1 punt, met een standaarddeviatie van 0,4 punt.

Uit de afwijking kunnen we opmaken dat de hele groep gemiddeld net iets hoger scoort voor de mondelinge dan voor de schriftelijke toetsen, maar het resultaat heeft een hele hoge standaarddeviatie. Drie van de zeven leerlingen scoort hoger bij het mondeling, de andere vier lager. Of een leerling hoger of lager scoort bij de mondelinge toets lijkt onafhankelijk van het cijfer wat diegene haalt. Ook lijkt er bij de evaluatie van mondelinge toetsen geen voorkeur voor hogere of lagere cijfers te zijn: drie leerlingen scoren hoger bij het mondeling, en vier scoren lager. Dit kan komen door allerlei factoren, bijvoorbeeld: de toetsvorm kan de leerling beter of slechter liggen, de stof kan meer of minder weggezakt zijn, de leerling kan slechter of beter geslapen hebben, et cetera.

	Gem. voorgaande hoofdstukken	Mondelinge toets	Schriftelijke toets	Vershil	Afwijking (mondelinge-schriftelijke toets)
Leerling 1	6,5	6	6,5	0,5	-0,5
Leerling 2	7,3	4,8	5,9	1,1	-1,1
Leerling 3	5,4	3	3,3	0,3	-0,3
Leerling 4	7,3	7	8,4	1,4	-1,4
Leerling 5	6,8	6,5	5,4	1,1	1,1
Leerling 6	6,9	8	6,5	1,5	1,5
Leerling 7	6,1	5,1	2,5	2,6	2,6
Gem. $\pm$ std.	6,6 $\pm$ 0,6	5,8 $\pm$ 1,5	5,5 $\pm$ 1,9	1,2 $\pm$ 0,7	0,3 $\pm$ 1,4

Tabel 4: toetsresultaten van de leerlingen.

Tijdsinvestering

In het geval van een erg laag aantal leerlingen kost mondeling toetsen minder werk, maar als het aantal leerlingen boven de 20 uitkomt, zit je al ruim boven het kantelpunt. In dit onderzoek heb ik zo'n kantelpunt niet nauwkeurig kunnen bepalen, maar het hangt natuurlijk af van hoe snel je nakijkt, en welke materialen er al beschikbaar zijn. Het is dus in elke situatie anders. Een voorbeeldberekening van het afnemen van mondelinge toetsen bij een jaarlaag van 60 leerlingen komt uit op een tijdsinvestering van één werkweek van twee leraren. Voorbereidingen zijn niet meegenomen in die berekeningen, maar het is noemenswaardig dat er praktisch geen nawerk is.

5 Conclusie

In dit onderzoek heb ik verschillende aspecten van mondelinge toetsing van natuurkunde in de bovenbouw van het havo en vwo belicht. Bestaand wetenschappelijk onderzoek naar het onderwerp is schaars, en wat er is, mist meestal twee van de volgende vier aspecten van dit onderzoek: het onderzoek is naar mondelingen, in Nederland, op de middelbare school, bij natuurkundevakken gedaan. Onderzoeken die niet aan die eisen voldoen, hebben geen betrouwbare voorspellende waarde voor dit onderzoek. Om toch op bestaande kennis te bouwen, heb ik twee ervaringsdeskundigen geïnterviewd. Een van hen voldoet aan alle vier de aspecten, de ander aan drie van de vier. Op basis van die interviews heb ik mondelinge toetsen ontworpen en, met hulp van een tweede corrector, afgenomen bij de doelgroep: een 4 havo-klas van drie leerlingen, en een 4 vwo-klas van vier leerlingen. De resultaten van die mondelinge toetsen zijn vergeleken met schriftelijke toetsen over hetzelfde onderwerp, die een week eerder zijn afgenomen. De bevindingen van de interviews, de mondelinge, en de schriftelijke toetsen heb ik gebruikt om de onderzoeksvragen te beantwoorden.

Onderzoeksvraag 1: het efficiënt verzamelen van informatie

Uit de interviews blijkt dat het van belang is dat de leerling op diens gemak is. Daarvoor zijn enkele strategieën gepresenteerd: de leerling voorbereiden op mondelinge overhoringen; de inhoud van de mondelinge toets bij aanvang goed duidelijk maken; naast de leerling zitten; de leerling pauze of uitstel aanbieden in geval van blackout. Dit belang van gemak hebben we bij de toetsafname gemerkt: een leerling die aangaf tijdens de mondelinge toets niet op diens gemak te zijn scoorde meer dan een punt onder haar schriftelijke toetscijfer.

Ook geven de ervaringsdeskundigen aan, dat het flexibel aanpassen van het niveau van de vragen belangrijk is. Als de leraar dit goed doet, kan diegene met hogere precisie bepalen waar de grenzen van de kennis van de leerling ligt, dan met schriftelijke toetsing. Elke vraag moet de leraar een beter beeld geven van die grenzen. Tijdens de mondelinge toetsafname hebben we dit ook gemerkt: de leerlingen en wij waren het erover eens dat we een eerlijke beoordeling van de leerling konden maken. Ook hadden wij het idee dat de mogelijkheid tot doorvragen meerwaarde bood voor de precisie waarmee we het niveau van de leerling konden bepalen.

Onderzoeksvragen 2&3: de waarde van de mondelinge toets bij verschillende soorten leerdoelen

De ervaringsdeskundigen stellen dat dat goede mondelinge toetsen beter zijn dan schriftelijke toetsen op het gebied van begrip testen. Dat geldt ook voor vaardigheden, omdat een mondelinge toets het denkproces van de leerling goed in beeld kan brengen. Mondelinge toetsen zijn volgens de respondenten kwalitatief even goed als schriftelijke toetsen voor het testen van feitelijke kennis, maar er wordt aangegeven dat dat beter met schriftelijke toetsing getest kan worden, omdat de tijd tijdens een mondelinge toets beter benut kan worden door begrip en vaardigheden te testen. Ook noemen de respondenten dat mondelinge toetsen slechter dan schriftelijke toetsen zijn voor het testen van rekenvragen en vragen waarbij de leerling een tijdje in alle rust moet kunnen nadenken.

De ervaringen die wij hadden bij de mondelinge toetsafname waren in lijn met de uitspraken van de ervaringsdeskundigen. De mogelijkheid tot doorvragen en hints geven, stond ons toe om de

denkprocessen van de leerlingen te onderzoeken op een niveau dat we met schriftelijke toetsing nooit hebben bereikt. Het flexibel aanpassen van het niveau van de vragen zorgde ervoor dat we een (naar ons idee) nauwkeurige bepaling van het begripsniveau van de leerlingen konden doen. We hebben ook gemerkt dat er weinig tijd is om uitgebreide rekenvragen te stellen.

Onderzoeksvraag 4: de cijfersverhouding tussen schriftelijke en mondelinge toetsen

De mondelinge toetsen verschillen gemiddeld 1,2 punt van de schriftelijke toetsresultaten, met een standaarddeviatie van 0,7 punt. Tussen het gemiddelde van de mondelinge en de schriftelijke toetsen zit bijna geen verschil (0,3 punt), maar de standaarddeviatie bij die uitspraak is erg groot (1,4 punt). Met andere woorden, de toetsvormen hebben gemiddeld over de leerlingen dezelfde moeilijkheidsgraad, maar leerlingen hebben op individueel niveau wel baat bij beide testvormen, omdat hun goede eigenschappen bij de één anders naar voren komen dan bij de ander. Hierdoor concluderen we dat de toetsvormen nadruk leggen op verschillende vaardigheden van leerlingen, en dat een combinatie van de toetsvormen het best werkt om elk individu goed te beoordelen.

Onderzoeksvraag 5: tijdsinvestering van de docent

Een van de geïnterviewde ervaringsdeskundigen geeft aan liever schriftelijke toetsen af te nemen bij groepen groter dan 20 studenten, omdat dat minder tijd kost. In dit onderzoek is niet zo'n hard kantelpunt bepaald, voor hoeveel leerlingen er nodig zijn om een mondelinge toetsafname meer tijd in beslag te laten nemen dan een schriftelijke. Tijdsinvestering is natuurlijk onderhevig aan allerlei variabelen, bijvoorbeeld het aantal leerlingen, de hoeveelheid vooraf beschikbare materialen, en de nakijksnelheid van de docent. Ook moet de docent een afweging maken of de toegenomen precisie bij het testen van vaardigheden en begrip een bepaalde hoeveelheid extra tijdsinvestering waard is.

Hoofdvraag: de mondelinge toets als vervanging voor een schriftelijke

Het hoofddoel van dit onderzoek was om uit te vinden of mondelinge toetsen als vervanging voor schriftelijke toetsen gebruikt kunnen worden. Mijn conclusie luidt: soms. Organisatorisch bieden mondelinge toetsen vaak een grotere uitdaging dan schriftelijke toetsen, zeker in het geval van grote groepen leerlingen. Echter, mondelinge toetsen hebben veel waarde, omdat leerlingen gemiddeld meer dan een punt hoger of lager scoren, afhankelijk van de toetsvorm. Een leerling kan er dus veel baat bij hebben om mondeling getoetst te worden.

Er zijn een aantal verschillen tussen de toetsvormen, waardoor de beste keuze afhankelijk is van de situatie. Elke toetsvorm heeft een eigen profiel van leerdoeltypen waarvoor hij zich goed leent. In dit onderzoek zijn beide toetsen zó ingericht, dat de leerdoeltypen waarvoor de toetsvorm het meest geschikt is, het grootste deel van het eindcijfer bepaalden. Toen deze methode gevolgd werd, verschilden de gemiddelde cijfers over de gehele populatie niet veel van elkaar. Wat dat betreft, suggereert dit onderzoek dus dat mondelinge toetsen als vervanging van schriftelijke toetsen gebruikt kunnen worden.

De belangrijkste conclusie van dit onderzoek is de volgende: voor een compleet beeld van de competenties van de leerling kan beter een combinatie van mondelinge én schriftelijke toetsen gebruikt worden, dan slechts één van de twee.

6 Discussie

In een late fase van dit onderzoek (na de afname van de mondelinge toetsen) attendeerde een van de ervaringsdeskundigen me op het bestaan van Duitstalige literatuur over het afnemen van mondelinge toetsen bij natuurkunde. Ik zal deze discussie starten, door daarvan de sluier op te lichten, met behulp van google en chatgpt voor het vertalen van de bronnen. Daarop volgend bespreek ik de limitaties van dit onderzoek, aangevuld met aanbevelingen voor toekomstig onderzoek.

6.1 Bevindingen uit Duitstalige literatuur

Verdere motivatie voor mondelinge toetsen wordt gegeven door Koch en Oesterreicher (1985), die het volgende schrijven: "Niet alleen bij goed opgeleide leken blijft hardnekkig de opvatting bestaan dat gesproken taal als een tekortkomende modus van de 'eigenlijke' taal, namelijk de geschreven taal, moet worden beschouwd." Koch en Oesterreicher maken het onderscheid tussen twee soorten taal, het formelere geschreven en het informelere gesproken. Ze volgen op met de notie dat die categorisaties niet heel sluitend zijn, omdat soms gesproken taal formeel is (bijvoorbeeld bij een toespraak) en geschreven taal informeel (bijvoorbeeld in een dagboek). Een betere onderscheiding, zo schrijven ze, is een taal van afstand tegenover een taal van nabijheid. De kenmerken die ze aan beide soorten toewijzen, zijn uiteengezet in tabel 5. Dat mooie citaat van Koch en Oesterreicher vindt dus de volgende herformulering: in tegenstelling tot algemeen geloof (ook onder goed opgeleiden), is de taal van afstand niet superieur ten opzichte van de taal van nabijheid.

Taal van Nabijheid	Taal van Afstand
Communicatievoorwaarden	
Dialogoog	Monoloog
Vertrouwdheid tussen de gesprekspartners	Onbekendheid tussen de gesprekspartners
Face-to-face-interactie	Ruimte- en tijdsafstandsinteractie
Vrije ontwikkeling van thema's	Vaststelling van thema's
Geen publieke aanwezigheid	Publieke aanwezigheid
Spontaniteit	Reflectie
Betrokkenheid	Afstandelijkheid
Situaties verbonden	Situaties losgekoppeld
Expressiviteit, affectiviteit	Objectiviteit
Taalstrategieën	
Procesmatigheid	Verdinglijking
Voorlopiegheid	Definitiviteit
Lagere informatie-dichtheid	Hogere informatie-dichtheid
Lagere compactheid	Hogere compactheid
Lagere integratie	Hogere integratie
Lagere complexiteit	Hogere complexiteit
Lagere elaboratie	Hogere elaboratie
Minder planning	Meer planning

Tabel 5: Communicatievoorwaarden en taalstrategieën bij de taal van afstand en de taal van nabijheid.

Als deze stelling klopt, is het eigenlijk heel raar dat we onze toetsen met die taal van afstand indelen. Schriftelijke toetsen en practica richten we normaal gesproken in als een communicatievorm, waarbij beide partijen de taal van afstand spreken. Toetsen in de taal van nabijheid is een logische aanvulling, gegeven dat de taal van nabijheid minstens zo waardevol is als die van afstand. De mondelinge toets leent zich daar natuurlijk uitstekend voor.

We moeten dus de mondelinge toets zien als een kans om te toetsen in de taal van nabijheid. Om zo'n kans goed aan te grijpen, is het belangrijk dat de mondelinge toets wordt ingericht met de eigenschappen van de taal van nabijheid in het achterhoofd. De belangrijkste aspecten zijn de volgende: er moet dialoog zijn; vertrouwdheid tussen gesprekspartners; vrije ontwikkeling van thema's; spontaniteit; betrokkenheid; en minder planning.

Als we terugkijken naar de interviews, zien we dat de ervaringsdeskundigen er een punt van maken om de mondelinge toetste gebruiken als toetsvorm in de taal van nabijheid. Ze richten het in als een dialoog, waarbij onderwerpen van de ene naar de andere leiden. Ze gebruiken eigenschappen van de taal van nabijheid (betrokkenheid, vertrouwen tussen gesprekspartners) in de leerling op diens gemak te brengen, wat ervoor zorgt dat het onderste uit de kan gehaald kan worden. We raden docenten die mondelinge toetsen willen afnemen aan om de eigenschappen van de taal van nabijheid in het achterhoofd te houden, zodat de toetsvorm optimaal benut kan worden.

Verder zijn er suggesties voor mondelinge natuurkundetoetsen door Pommeranz (z.j.) gedaan, waarbij ingegaan wordt op algemene methoden, maar ook op wat gedetailleerder niveau op elk van de eindexamenonderwerpen. Er wordt onderscheid gemaakt tussen drie moeilijkheidsniveaus: reproductie, toepassing, en probleemoplossing. De niveaus vergelijk ik in tabel 6 met de herziene cognitieve procesdimensies van Bloom (Krathwohl, 2002).

Niveau volgens Pommeranz	Benodigde herziene cognitieve procesdimensies van Bloom (Krathwohl, 2002)
1: Reproductie: het reproduceren van eenvoudige feiten en methoden, het beschrijven van eenvoudige verbanden, en informatie presenteren op een gevraagde wijze.	1: onthouden 2: begrijpen 3: toepassen (op simpel niveau: kleine omrekeningen of berekeningen, of datavisualisatie met grafieken e.d.)
2: toepassing: het herorganiseren van kennis, het toepassen van methoden in nieuwe situaties, en het geven van eenvoudige verklaringen.	1: onthouden 2: begrijpen 3: toepassen 4: analyseren (op middelmatig niveau, er hoeft geen kennis uit verschillende gebieden gebruikt te worden)
3: Probleemoplossen: het toepassen en overdragen van complexe concepten en methoden, evenals het selecteren en evalueren van relevante informatie.	1: onthouden 2: begrijpen 3: toepassen 4: analyseren 5: evalueren

Tabel 6: de moeilijkheidsgraden van Pommeranz (Pommeranz, z.j.) vergeleken met de taxonomie van Bloom (Krathwohl, 2002)

Pommeranz licht bij al deze niveaus uit op welke manier de vier competentiegebieden (vakinhoudelijke kennis, verwerving van kennis, communicatie, en evaluatie) naar voren kunnen komen. Bij de mondelinge toetsen legt hij de nadruk op communicatie en evaluatie. Over toetsvormen met nadruk op de andere competentiegebieden zegt Pommeranz niks, we veronderstellen dat bij schriftelijke toetsen de nadruk op vakinhoudelijke kennis en op evaluatie ligt. Voor toetsing van kennisverwerving kunnen practica zich goed lenen.

In zijn suggesties legt Pommeranz veel nadruk op hoe vragen ontworpen moeten worden. Hij geeft een aantal voorbeeldvragen over verschillende thema's, elk met deelvragen van verschillende niveaus. Hij geeft aan dat vragen aan de volgende vier eisen moeten voldoen:

- de fysische competenties in alle drie de moeilijkheidsniveaus controleren;
- meerdere vakinhoudelijke methoden behandelen;
- verschillende stofgebieden met elkaar verbinden;
- een vakinhoudelijke communicatie mogelijk maken waarin een discussie wordt gevoerd.

Daarnaast levert hij een rubric aan, waarmee hij de leerling beoordeelt. De rubric is te vinden in tabel 7.

Tabel: Formuleringen voor beoordeling mondelinge prestaties

Beoordelingscriteria	1 - Uitstekend	2 - Zeer goed	3 - Goed	4 - Voldoende	5 - Onvoldoende	6 - Slecht
Thema's begrijpen	Zeer volledig, creatief.	Volledig en accuraat.	Met hulp redelijk volledig.	Oppervlakkig, beperkt begrip.	Geen echt begrip.	Volledig onbegrip.
Presentatiestijl	Zeer samenhangend, overtuigend.	Samenhangend, meestal overtuigend.	Redelijk overtuigend.	Onsamenhangend, onzeker.	Rommelig, geen focus.	Niet overtuigend, chaotisch.
Vakinhoudelijke kennis	Diepgaand begrip, creatieve toepassing.	Goed begrip, zekere toepassing.	Basiskennis aanwezig.	Beperkte kennis, onzeker in toepassing.	Nauwelijks basiskennis.	Geen aantoonbare kennis.
Gebruik van vaktaal	Zeer correct, consistent gebruik.	Correct en meestal consistent.	Redelijk correct.	Onzeker gebruik van vaktaal.	Vaak fout en inconsistent.	Onjuist gebruik, niet herkend als vaktaal.
Argumentatie	Logisch, overtuigend, goed onderbouwd.	Meestal logisch, overtuigend.	Met enige fouten in logica.	Moeizaam, vaak onlogisch.	Nauwelijks onderbouwd.	Geen argumentatie aanwezig.
Interactie-vermogen	Uitstekend reageren op vragen, flexibel.	Meestal goed reageren, redelijk flexibel.	Soms stug, maar kan enkele vragen beantwoorden.	Vaak geen passend antwoord, onzekerheid.	Niet in staat om adequaat te reageren.	Reacties zijn volledig ongepast of afwezig.

Tabel 7: Rubric bij mondelinge overhoringen (Pommelan, z.j.)

Samengevat levert Pommelanaz suggesties voor mondelinge toetsen die enigszins afwijken van de suggesties die de ervaringsdeskundigen leveren. Zijn mondelinge toetsen zijn minder flexibel dan de mondelinge toetsen die in dit onderzoek, op basis van de interviews met de ervaringsdeskundigen, zijn uitgevoerd. Hij bepaalt alle vragen van tevoren, en test ze allemaal aan dezelfde rubric, waar in dit onderzoek de vragen aangepast worden op de situatie tijdens de toetsafname, en getest worden aan de vooraf besproken slagingsdoelen. Het verschil in aanpak zou aan een cultuurverschil tussen Nederland en Duitsland kunnen liggen.

Een mogelijk voordeel van de methode van Pommelanaz, is dat de toetsen makkelijker met een rubric te beoordelen zijn. Het gesprek zal minder onverwachte wendingen aannemen. Echter, die onverwachte wendingen, en de flexibiliteit van het mondelinge, zijn eigenschappen waaraan de ervaringsdeskundigen enorm veel waarde hechten. Ze kunnen namelijk gebruikt worden om het onderste uit de kan te halen. Ook ik ben gedurende dit onderzoek een voorstander geworden van het semi-gestructureerde mondelinge. De mondelinge toets van Pommelanaz is, vooral door de uitgebreide voorbereiding en de aangebrachte structuur, meer in de taal van afstand. Dat terwijl we eerder in deze literatuuranalyse de waarde van de mondelinge toets in de taal van nabijheid hebben uitgelicht. Wel is de rubric van Pommelanaz bij de flexibelere mondelinge toets toepasbaar. We raden dan ook aan om deze in vervolgonderzoek als bron te gebruiken.

6.2 Limitaties en aanbevelingen

Dit onderzoek poogt met relatief kleine testgroepen (twee ervaringsdeskundigen, zeven leerlingen, één docent, en een notulist) een kwalitatieve evaluatie van een aantal aspecten van mondelinge toetsing van natuurkunde in de bovenbouw van het havo en vwo te doen. Over alle onderzoeksvragen zijn zinnige conclusies getrokken, maar door de grootte van de testgroepen zijn die conclusies niet significant. Het onderzoek licht wel de waarde van de mondelinge toets voor deze doelgroep uit. Daarom adviseer ik een grootschaliger onderzoek, waarbij verschillende docenten mondelinge toetsen afnemen bij verschillende klassen, zodat de resultaten meer algemeen interpreteerbaar zijn. In sommige gevallen moeten de mondelinge toetsen vóór de schriftelijke toetsen afgenomen worden, in andere gevallen erna. Ook is een grondigere evaluatie van de Duitstalige literatuur aanbevolen.

Er kan ook geëxperimenteerd worden met mondelinge toetsen in groepen. Het verschil tussen mondelinge toetsen met individuen en mondelinge toetsen met groepen voor de nauwkeurigheid van de toetsing is een interessant onderzoeksonderwerp, omdat mondelinge toetsen in groepen afnemen een manier zou zijn om de tijdsinvestering en organisatorische uitdaging van de mondelinge toetsen te verminderen.

7 Bronnen

Bear, G. G., Yang, C., Glutting, J., Huang, X., He, X., Zhang, W., & Chen, D. (2014). Understanding Teacher-Student Relationships, Student-Student Relationships, and Conduct Problems in China and the United States. *International Journal Of School & Educational Psychology*, 2(4), 247–260. <https://doi.org/10.1080/21683603.2014.883342>

Chen, M., Zee, M., Koomen, H. M., & Roorda, D. L. (2019). Understanding cross-cultural differences in affective teacher-student relationships: A comparison between Dutch and Chinese primary school teachers and students. *Journal Of School Psychology*, 76, 89–106. <https://doi.org/10.1016/j.jsp.2019.07.011>

Goodman, A. L. (2020). Can Group Oral Exams and Team Assignments Help Create a Supportive Student Community in a Biochemistry Course for Nonmajors? *Journal Of Chemical Education*, 97(9), 3441–3445. <https://doi.org/10.1021/acs.jchemed.0c00815>

Hayhoe, R. (2020). Book Series: Intercultural Reciprocal Learning in Chinese and Western Education. *Frontiers Of Education in China*, 15(3), 526–529. <https://doi.org/10.1007/s11516-020-0022-8>

Iannone, P., Czichowsky, C., & Ruf, J. (2020). The impact of high stakes oral performance assessment on students' approaches to learning: a case study. *Educational Studies in Mathematics*, 103(3), 313–337. <https://doi.org/10.1007/s10649-020-09937-4>

Joughin, G. (1998). Dimensions of Oral Assessment. *Assessment & Evaluation in Higher Education*, 23(4), 367–378. <https://doi.org/10.1080/0260293980230404>

Koch, P., & Oesterreicher, W. (1985). Sprache der Nähe – Sprache der Distanz: Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. In O. Deutschmann, H. Flasche, B. König, M. Kruse, W. Pabst, & W.-D. Stempel (Eds.), *Romanistisches Jahrbuch* (Vol. 36, pp. 15–43). Berlin, New York: Walter de Gruyter.

Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice*, 41(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2

Pommeranz, H.-P. (z.j.). *Anregungen zur Gestaltung mündlicher Abschlussprüfungen*. Bildungsserver Sachsen-Anhalt. Geraadpleegd op 05/01/2025, van https://www.bildungs-lsa.de/files/fd502a5faa8f6d33e54892f7476e9184/Muendliche_Abschlusspruefungen.pdf

Rahman, G. (2011). Appropriateness of using oral examination as an assessment method in medical or dental education. *Journal Of Education And Ethics in Dentistry*, 1(2), 46. <https://doi.org/10.4103/0974-7761.103674>

Rincke, K., & Krabbe, H. (2020). Sprache im Physikunterricht. In *Springer eBooks* (pp. 367–401). https://doi.org/10.1007/978-3-662-59490-2_10

Eisen aan een goede toets. (z.d.). Stichting leerplanontwikkeling. Geraadpleegd op 21 januari 2025, van <https://www.slo.nl/handreikingen/vmbo/handreiking-se-gesch-vmbo/toetsen-schoolexamen/eisen-goede-toets>

Theobold, A. S. (2021). Oral Exams: A More Meaningful Assessment of Students' Understanding. *Journal Of Statistics And Data Science Education*, 29(2), 156–159. <https://doi.org/10.1080/26939169.2021.1914527>

Vonen, M. N., Solem, M. S., & Skovholt, K. (2022). Managing students' insufficient answers in oral examinations. *Classroom Discourse*, 14(3), 258–280.
<https://doi.org/10.1080/19463014.2022.2079694>

Xu, C., Huizinga, M., De Luca, G., Pollé, S., Liang, R., Sankalaite, S., Roorda, D. L., & Baeyens, D. (2023). Cultural universality and specificity of teacher-student relationship: a qualitative study in Belgian, Chinese, and Italian primary school teachers. *Frontiers in Psychology*, 14.
<https://doi.org/10.3389/fpsyg.2023.1287511>

8 Appendix

8.1 Appendix 1: interviewvragen

1. Kunt u kort iets vertellen over uw ervaring met mondelinge toetsen?
2. Waarom neem je mondelinge toetsen af?
3. Hoe neem je een mondeling af?
4. Wat voor soort leerdoelen kan je goed testen met een mondeling?
5. Wat voor soort leerdoelen kan je minder goed testen met een mondeling?
6. Hoe bereidt u een mondeling voor als docent?
7. Hoe leidt u het gesprek tijdens een mondeling?
8. Hoe bepaalt u het cijfer van de leerling bij een mondeling?
9. Wat is in uw ervaring het voordeel van een mondelinge toets t.o.v. een schriftelijke?
10. Wat is in uw ervaring het nadeel van een mondelinge toets t.o.v. een schriftelijke?
11. Wat zijn in uw ogen de kenmerken van een goed mondeling?
12. Is er verder nog iets wat u kwijt wil?

8.2 Appendix 2: Schriftelijke toets havo

4 Havo PTA-toets Hoofdstuk 5&6

Straling en Energie

Instructies:

- vul in ieder geval je naam, achternaam en klas in op het toetsbladje.
- denk om je 'gereedschap': eerst de gegevens en formules opschrijven, dan pas invullen en uitrekenen.
- schrijf het eindantwoord in de juiste SIGNIFICANTIE en met de juiste hoeveelheid DECIMALEN.

1. Geef van de volgende beweringen aan of ze waar of onwaar zijn. (5 pt.)

- Het doordringend vermogen van bètastraling is groter dan dat van alfastraling.
- Het ioniserend vermogen van gammastraling is groter dan dat van bètastraling.
- Energie blijft behouden in een situatie met wrijvingskrachten.
- De halfwaardetijd is de tijd die het kost voordat een stralingsdeeltje de helft van haar energie heeft verloren aan de omgeving.
- De arbeid die het kost om iets (wrijvingsloos) omhoog te bewegen, is gelijk aan de zwaarte-energie die het object krijgt.

2. Geef de vervalvergelijking van Rn-218 (3 pt.)

3. Een kruisboog is een soort kruising tussen een boog en een geweer. De gespannen pees wordt achter een blokkeerstaaf gehaakt die de boog op spanning houdt. Met de trekker kan de blokkeerstaaf worden weggetrokken, zodat een pijl kan worden weggeschoten. Zie de figuur hiernaast. Als de kruisboog volledig uitgerekt is, dan is de pees 30 cm naar achteren getrokken. Je kan een pijl van 20 gram uit de boog afschieten met een maximale snelheid van 87 m/s. Ga ervan uit dat de boog hetzelfde werkt als een veer.



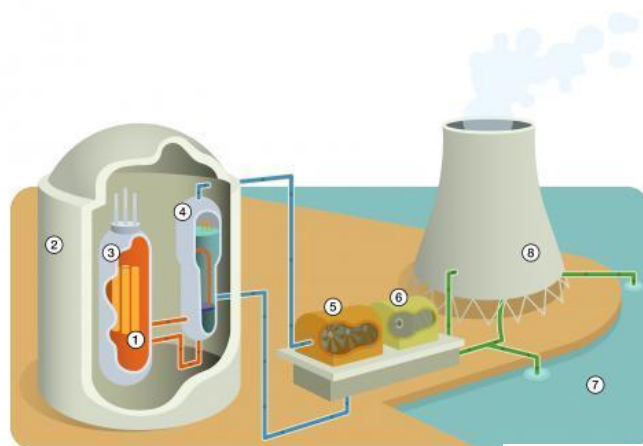
- Welke** energie-overgangen vinden er plaats als je de pijl verticaal de lucht in schiet? Begin bij het moment dat de trekker wordt overgehaald, en stop als de pijl in rust is.

(2 pt.)

- Bereken** de maximale kinetische energie van de pijl.

(3 pt.)

4. In een kerncentrale wordt elektrische energie opgewekt door uraniumkernen te splijten. Als een uraniumkern wordt beschoten met neutronen, splitst de uraniumkern in andere atoomkernen en neutronen. De warmte die bij deze reactie ontstaat, wordt gebruikt om elektrische energie op de wekken.



Figuur 1

In een kerncentrale wordt een mengsel van de isotopen uranium-235 en uranium-238 gebruikt. Deze isotopen hebben niet ieder dezelfde eigenschappen ondanks dat het beide vormen van uranium zijn. Zo is de massa van een enkel atoom bij de beide stoffen niet identiek.

- a) **Leg uit** hoe dit massaverschil wordt veroorzaakt.

(1 pt.)

Uranium-238 is onder andere een gammastraler. De gammastraling van deze bron draagt een energie met zich mee van 100 keV. Het is wenselijk dat alle straling wordt gebruikt voor het opwekken van energie. Met deze reden zal zoveel mogelijk straling moeten worden geabsorbeerd door het waterbad waar de radioactieve stof zich in bevindt. Dit waterbad met een diameter van 18,0 m is zichtbaar in figuur 1 gemarkeerd met 1.

- b) **Toon aan met behulp van een berekening** dat zo goed als alle gammastraling wordt geabsorbeerd door het reactorbad als het uranium in het midden wordt geplaatst.

(3 pt.)

Als een neutron door U-238 wordt ingevangen, treedt er geen kernsplijting op. Er ontstaat dan U-239 dat in twee stappen vervalst tot Pu-239

- c) **Leg uit** of U-239 een α - of β -straler is.

(1 pt.)

5. In het meer van Genève bevindt zich een van de grootste fonteinen ter wereld. Bij de fontein hangt een informatiebordje. De tekst op dit bordje staat, vertaald, weergegeven in figuur 2.



Fontein van Genève

Elke seconde wordt er 450 liter water de lucht in gestuwd tot een hoogte van 140 m.

Het water wordt met twee pompen door een spuitmond gespoten met een snelheid van 200 km/h. De twee elektrische pompen hebben elk een vermogen van 500 kW.

Na zonsondergang wordt de straal verlicht door een aantal lampen met een gezamenlijk vermogen van 13,5 kW.

Fontein in werking:
maandag tot vrijdag: 10.00 - zonsondergang
vrijdag tot en met zondag: 10.00 – 22.30 uur.

Figuur 2

De twee elektrische pompen hebben elk een vermogen van 500 kW.

Het water wordt met een snelheid van 200 km/h uit de spuitmond gespoten

- a) **Bereken** het rendement van de elektrische pompen. Neem hierbij voor de dichtheid van water 1,00 kg/L. **(3 pt.)**
- b) **Toon met een berekening aan** of het water de maximale hoogte die op het bordje staat kan halen. **(4 pt.)**

6. Technetium-99m wordt in ziekenhuizen gebruikt als tracer. Het Tc-99m dat daarvoor nodig is, wordt in het ziekenhuis zelf geproduceerd. Tc-99m is een vervalproduct van molybdeen-99. Tc-99m is metastabiel. Dit betekent dat de protonen en neutronen in de kern van een Tc-99m atoom zich nog kunnen herschikken tot een toestand met minder energie, die ze dan als gammafoton uitstralen.

a) **Geef** de vervalvergelijking waarbij Tc-99m ontstaat. (3 pt.)

b) **Leg uit** of er bij het gebruik van een tracer voor de patiënt sprake is van bestraling of van besmetting. (1 pt.)

Seth krijgt een dosis Tc-99m met een activiteit van 800 MBq toegediend. Daarna moet hij nog 10,0 minuten wachten tot zijn scan begint.

c) **Bereken** de activiteit van het Tc-99m op het moment dat de scan begint. (3 pt.)

In het ziekenhuis wordt het Tc-99m van het Mo-99 gescheiden. Tijdens dit scheidingsproces blijft er Mo-99 in de oplossing van het Tc-99m achter. Deze verontreiniging van het Tc-99m is ongewenst, omdat Mo-99 bètastraling uitzendt. Volgens wettelijke eisen mag de activiteit van de Mo-99 verontreiniging maximaal 0,15 kBq zijn per 1,0 MBq activiteit van het Tc-99m.

$$\frac{A(t)_{\text{Mo-99}}}{A(t)_{\text{Tc-99m}}} = \frac{t_{\frac{1}{2}\text{Tc-99m}} \cdot N(t)_{\text{Mo-99}}}{t_{\frac{1}{2}\text{Mo-99}} \cdot N(t)_{\text{Tc-99m}}}$$

De verhouding van activiteiten is te berekenen met:

d) **Bereken** hoeveel Mo-99-kernen er maximaal per miljoen Tc-99m-kernen mogen voorkomen. (2 pt.)

Het geproduceerde Tc-99m wordt bewaard in potten gemaakt van 6,0 mm dik lood. In een pot ontstaat de volgende straling:

- Gammafotonen, met een energie van 0,1 MeV, uitgezonden door Tc-99m;
- bètadeeltjes, uitgezonden door Mo-99;
- gammafotonen, met een energie van 1,0 MeV, uitgezonden door Mo-99.

e) **Geef een reden** waarom de bètadeeltjes, uitgezonden door Mo-99, niet buiten de pot gedetecteerd kunnen worden. (1 pt.)

8.3 Appendix 3: schriftelijke toets vwo

4 Vwo PTA-toets Hoofdstuk 5&6

Straling en Energie

Instructies:

- vul in ieder geval je naam, achternaam en klas in op het toetsbladje.
- denk om je 'gereedschap': eerst de gegevens en formules opschrijven, dan pas invullen en uitrekenen.
- schrijf het eindantwoord in de juiste SIGNIFICANTIE en met de juiste hoeveelheid DECIMALEN.

7. Geef van de volgende beweringen aan of ze waar of onwaar zijn. (5 pt.)

- f) Het doordringend vermogen van bètastraling is groter dan dat van alfastraling.
- g) Het ioniserend vermogen van gammastraling is groter dan dat van bètastraling.
- h) Energie blijft behouden in een situatie met wrijvingskrachten.
- i) De halfwaardetijd is de tijd die het kost voordat een stralingsdeeltje de helft van haar energie heeft verloren aan de omgeving.
- j) De arbeid die het kost om iets (wrijvingsloos) omhoog te bewegen, is gelijk aan de zwaarte-energie die het object krijgt.

8. Geef de vervalvergelijkingen van de volgende isotopen:

- a) Al-26 (3 pt.)
- b) Rn-218 (3 pt.)

9. Een kruisboog is een soort kruising tussen een boog en een geweer. De gespannen pees wordt achter een blokkeerstaaf gehaakt die de boog op spanning houdt. Met de trekker kan de blokkeerstaaf worden weggetrokken, zodat een pijl kan worden weggeschoten. Zie de figuur hiernaast. Als de kruisboog volledig uitgerekt is, dan is de pees 30 cm naar achteren getrokken. Je kan een pijl van 20 gram uit de boog afschieten met een maximale snelheid van 87 m/s. Ga ervan uit dat de boog hetzelfde werkt als een veer.



- c) **Welke** energie-overgangen vinden er plaats als je de pijl verticaal de lucht in schiet? Begin bij het moment dat de trekker wordt overgehaald,

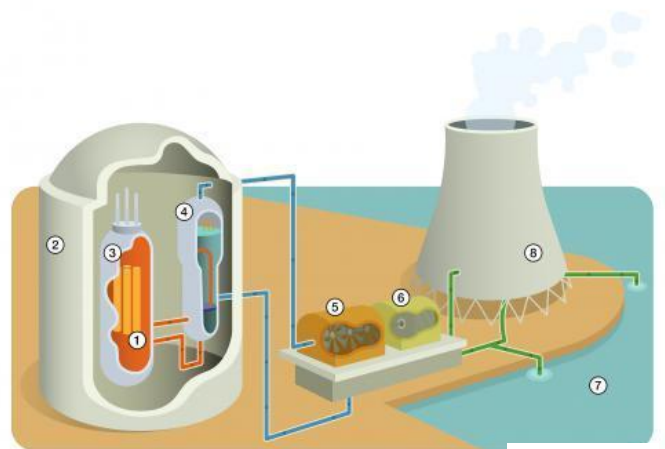
en stop als de pijl in rust is.

(2 pt.)

d) **Bereken** de veerconstante van de boog.

(4 pt.)

10. In een kerncentrale wordt elektrische energie opgewekt door uraniumkernen te splijten. Als een uraniumkern wordt beschoten met neutronen, splitst de uraniumkern in andere atoomkernen en neutronen. De warmte die bij deze reactie ontstaat, wordt gebruikt om elektrische energie op te wekken.



Figuur 1

In een kerncentrale wordt een mengsel van de isotopen uranium-235 en uranium-238 gebruikt. Deze isotopen hebben niet ieder dezelfde eigenschappen ondanks dat het beide vormen van uranium zijn. Zo is de massa van een enkel atoom bij de beide stoffen niet identiek.

d) **Leg uit** hoe dit massaverschil wordt veroorzaakt.

(1 pt.)

Uranium-238 is onder andere een gammastraler. De gammastraling van deze bron draagt een energie met zich mee van 100 keV. Het is wenselijk dat alle straling wordt gebruikt voor het opwekken van energie. Met deze reden zal zoveel mogelijk straling moeten worden geabsorbeerd door het waterbad waar de radioactieve stof zich in bevindt. Dit waterbad met een diameter van 18,0 m is zichtbaar in figuur 1 gemarkeerd met 1.

e) **Toon aan met behulp van een berekening** dat zo goed als alle gammastraling wordt geabsorbeerd door het reactorbad als het uranium in het midden wordt geplaatst.

(3 pt.)

Als een neutron door U-238 wordt ingevangen, treedt er geen kernsplijting op. Er ontstaat dan U-239 dat in twee stappen vervalst tot Pu-239

f) **Leg uit** of U-239 een α - of β -straler is.

(1 pt.)

11. In het meer van Genève bevindt zich een van de grootste fonteinen ter wereld. Bij de fontein hangt een informatiebordje. De tekst op dit bordje staat, vertaald, weergegeven in figuur 2.



Fontein van Genève

Elke seconde wordt er 450 liter water de lucht in gestuwd tot een hoogte van 140 m.

Het water wordt met twee pompen door een spuitmond gespoten met een snelheid van 200 km/h. De twee elektrische pompen hebben elk een vermogen van 500 kW.

Na zonsondergang wordt de straal verlicht door een aantal lampen met een gezamenlijk vermogen van 13,5 kW.

Fontein in werking:
maandag tot vrijdag: 10.00 - zonsondergang
vrijdag tot en met zondag: 10.00 – 22.30 uur.

Figuur 2

De twee elektrische pompen hebben elk een vermogen van 500 kW.

Het water wordt met een snelheid van 200 km/h uit de spuitmond gespoten

- c) **Bereken** het rendement van de elektrische pompen. Neem hierbij voor de dichtheid van water 1,00 kg/L. **(3 pt.)**
- d) **Toon met een berekening aan** of het water de maximale hoogte die op het bordje staat kan halen. **(4 pt.)**

12. Technetium-99m wordt in ziekenhuizen gebruikt als tracer. Het Tc-99m dat daarvoor nodig is, wordt in het ziekenhuis zelf geproduceerd. Tc-99m is een vervalproduct van molybdeen-99. Tc-99m is metastabiel. Dit betekent dat de protonen en neutronen in de kern van een Tc-99m atoom zich nog kunnen herschikken tot een toestand met minder energie, die ze dan als gammafoton uitstralen.

f) Geef de vervalvergelijking waarbij Tc-99m ontstaat. **(3 pt.)**

g) Leg uit of er bij het gebruik van een tracer voor de patiënt sprake is van bestraling of van besmetting. **(1 pt.)**

Seth krijgt een dosis Tc-99m met een activiteit van 800 MBq toegediend. Daarna moet hij nog 10,0 minuten wachten tot zijn scan begint.

h) Bereken de activiteit van het Tc-99m op het moment dat de scan begint. **(3 pt.)**

In het ziekenhuis wordt het Tc-99m van het Mo-99 gescheiden. Tijdens dit scheidingsproces blijft er Mo-99 in de oplossing van het Tc-99m achter. Deze verontreiniging van het Tc-99m is ongewenst, omdat Mo-99 bètastraling uitzendt. Volgens wettelijke eisen mag de activiteit van de Mo-99 verontreiniging maximaal 0,15 kBq zijn per 1,0 MBq activiteit van het Tc-99m.

$$\frac{A(t)_{\text{Mo-99}}}{A(t)_{\text{Tc-99m}}} = \frac{t_{\frac{1}{2}\text{Tc-99m}} \cdot N(t)_{\text{Mo-99}}}{t_{\frac{1}{2}\text{Mo-99}} \cdot N(t)_{\text{Tc-99m}}}$$

De verhouding van activiteiten is te berekenen met:

i) Bereken hoeveel Mo-99-kernen er maximaal per miljoen Tc-99m-kernen mogen voorkomen. **(2 pt.)**

Het geproduceerde Tc-99m wordt bewaard in potten gemaakt van 6,0 mm dik lood. In een pot ontstaat de volgende straling:

- Gammafotonen, met een energie van 0,1 MeV, uitgezonden door Tc-99m;
- bètadeeltjes, uitgezonden door Mo-99;
- gammafotonen, met een energie van 1,0 MeV, uitgezonden door Mo-99.

j) Geef een reden waarom de bètadeeltjes, uitgezonden door Mo-99, niet buiten de pot gedetecteerd kunnen worden. **(1 pt.)**

8.4 Appendix 4: vragenlijst lerarenervaringen

1. Wat was je beeld van mondelingen vooraf? Denk aan het doel/nut ervan, en of je ze zou gebruiken
2. Waren er achteraf dingen opgevallen?
3. Is je beeld van mondelingen veranderd? In welke zin? Waarom?
4. Wat is je huidige beeld van mondelingen?
5. Zou je ze zelf gebruiken in de toekomst? Waarom?
6. Wat voor praktische consequenties zie je?

8.5 Appendix 5: resultaten interviews uitgezet tegen de vragen

8.5.1 Respondent 1

Respondent 1 is een professor, die bij master- en capita selectavakken binnen de natuurkunde mondelinge toetsing gebruikt.

Respondent geeft aan dat mondelinge toetsen nuttig zijn, omdat ze de leraar binnen korte tijd inzicht van kunnen en inzicht geven, en omdat het niveau tijdens de mondelinge toets aan te passen is aan de student. Zo valt bij een klein begripsgat niet de hele toets uit elkaar. Ook noemt respondent dat het binnen de functie als professor handig is om de studenten een beetje te leren kennen, waarvoor de mondelinge toetsen zich lenen.

Respondent neemt modelingen af met één andere docent. Dat zorgt voor verlichting voor de docenten, en voor meer objectiviteit. Van tevoren sluiten de docenten kort wat de slagingscriteria zijn. Die zijn meestal gericht op inzicht. Er wordt niet gerekend, maar stappenplannen moeten studenten soms wel uitleggen. Eerst worden er vragen gesteld over de slagingscriteria, daarna worden vragen moeilijker. Er wordt geprobeerd om alle aspecten aan bod te laten komen. Er wordt aan het begin of aan het einde van een mondeling gevraagd waarin ze geïnteresseerd zijn bij het vak [onderwerp van toetsing]. Beide docenten geven eerst zelf een cijfer, en beslissen later samen wat het cijfer wordt. Ze zitten zelden meer dan een halve punt van elkaar af.

Respondent geeft aan graag naast de student te zitten, zodat er geen ‘twee-tegen-eengevoel’ ontstaat bij de leerling. Zo kan de docent ook makkelijker meekijken als de leerling iets wil tekenen, of als de docent dat van de leerling vraagt.

Respondent geeft aan dat mondelinge toetsen meestal 30-40 min duren, waarbij de laatste 10 min het cijfer meestal niet lager maken. Cijfers komen de studenten zelden als verrassing.

Respondent geeft aan dat de ‘hogere leerdoelen’ beter getoetst kunnen worden met een mondeling, waar een mondeling weinig voordelen heeft t.o.v. een schriftelijk bij het toetsen van parate kennis. Met ‘hogere leerdoelen’ bedoelt de respondent aan inzichtsvragen, verduidelijkt diegene later. Bij rekenvragen leent een schriftelijke toets zich beter, omdat de student het dan in alle rust kan uitpluizen. Je kan echter wel testen of de student de methodiek onder controle heeft.

Respondent geeft aan dat afkijkgevaar laag is, mits de docent een flink arsenaal aan vragen paraat heeft.

Respondent geeft de tip om je als examinerator flexibel op te stellen en de stijl aan te passen aan de student.

Volgens de respondent zijn de leukste mondelinge toetsen, die waarbij de student naar buiten gaat met het idee dat ze iets geleerd hebben.

Respondent vindt een goed mondeling een mondeling waarin docent en student het eens zijn over het cijfer. Na een goed mondeling kan je benoemen wat er fout ging tijdens het leerproces, of waar een gebrek aan kennis zit. De kwaliteit van de mondelinge toets is compleet onafhankelijk van het cijfer.

8.5.2 Respondent 2

Respondent 2 is vakdidacticus natuurkunde, die naar eigen zeggen geslaagd is voor de middelbare school in Duitsland, omdat diegene daar bij de centrale examens mondelinge toetsen ter vervanging van schriftelijke toetsen aanbieden. Dit maakte de dyslexie van de respondent een non-factor. Daarna heeft respondent in Duitsland en in Nederland bij lerarenopleidingen ervaringen met mondelinge toetsen, meest recentelijk in de rol van docent bij een lerarenopleiding in Nederland.

Respondent 2 geeft de komende voorbeelden van voordelen van mondelinge toetsing t.o.v. schriftelijke toetsing:

- Voor dyslecten is het beter als de toetsafname niet schriftelijk is.
- In het verlengde daarvan: je test altijd meerdere dingen, en de dingen die je naast de vakinhoud test [storende vaardigheden] zijn anders.
- De student leert diens gedachten goed te articuleren, wat in het bijzonder waardevol is op de lerarenopleiding.
- De docent komt dieper in het bepalen van iemands niveau, door herformuleren of navragen, of door het niveau van de toets aan te passen.
- Je kan tot een consensus komen over een cijfer met de student.
- De manier van toetsen is heel mooi: je kan het gebruiken om vanuit meerdere perspectieven te toetsen, en het is veel vrijer dan schriftelijke toetsing.
- Leerlingen leren van de toets.
- Feedback komt meteen.
- Leerlingen zijn het altijd eens met het cijfer
- Leerlingen met een onvoldoende lopen altijd weg met verbeterpunten.

Ter voorbereiding van een mondeling maakt de respondent een bak met kaarten, met daarop vragen over de stof. De bak wordt hergebruikt over verschillende jaarlagen. Aan de hand van een paar willekeurige vragen uit de bak, richt respondent de mondelinge toetsen in als een semi-gestructureerd interview. Mondelinge toetsen worden afgenomen met twee docenten, om subjectiviteit te minimaliseren en om een sparringspartner te hebben. Tegen het eind van de mondelinge toets vraagt de respondent of de leerling nog iets kwijt wil. Na de mondelinge toets is het eerste wat je doet overleggen met de andere docent. Voor een mondeling is 10 minuten genoeg, respondent rekent 2 minuten voor het nabespreken met de collega en 5 minuten voor het nabespreken met de leerling. Voor wat speling roostert de respondent 20 minuten per leerling in.

Respondent geeft aan dat conceptueel begrip goed te testen is met mondelinge toetsen, en dat mondelinge toetsen zich niet zo lenen voor rekenvaardigheden. Ook leesvaardigheid is niet goed te testen met mondelinge toetsen. Over het algemeen stelt respondent geen vragen waar leerlingen liever een tijdje over na zouden denken.

Respondent bereidt de mondelinge toets voor met de eerder benoemde kaartenbak. Evt. neemt diegene ook plaatjes/schetsen/diagrammen mee, of practicummaterialen die met het onderwerp te maken hebben. Ook verdeelt diegene de rollen met de andere docent: een van de twee is notulist, die houdt de tijd bij, en geeft een signaal als de laatste minuut ingaat.

Respondent laat de mondelinge toets graag open in het begin, door bijvoorbeeld een plaatje te laten zien en te vragen wat de leerling daarin ziet in de context van de toetsstof. Dat is een goede antenne voor het niveau van de leerling. Evt. vraagt de leraar of de leerling het liever ergens anders over heeft. Daarna kunnen de vrager wat concreter gesteld worden, met het geschatte niveau van de leerling in het achterhoofd. Als een leerling te lang bij hetzelfde onderwerp probeert te blijven, verandert respondent het onderwerp, bijvoorbeeld met behulp van een ander kaartje. Tijdens de inleiding van de mondelinge toets vertelt respondent de leerling ook graag over hoeveel onderwerpen de respondent het minstens wil hebben. Vaak is dat 2 of 3, afhankelijk van de leerling.

Cijfers bepaalt de respondent altijd in overleg met de notulist. Respondent geeft aan dat het vaak op gevoel gaat. Diegene draagt een optie aan om een rubric maken met de notulist, maar je wordt altijd verrast door de loop van het mondeling. Je kan een rubric naar de buitenwereld ook gebruiken om “de schijn van nauwkeurigheid op [te] houden”. Je kan hem ook gebruiken om het cijfer aan de leerling uit te leggen, maar in de tijd dat respondent mondelinge toetsen afneemt is het nog nooit gebeurd dat een leerling het oneens was met het cijfer.

Respondent noemt de vrijheid waarmee je vragen kunt stellen, en die op de leerlingen aan kan passen of na kunt vragen, als een voordeel van mondelinge toetsen. Het zorgt ervoor dat je een veel eerlijker cijfer kunt bepalen.

Respondent noemt dat mondelinge toetsen ervaren worden als tijdrovend en organisatorisch lastiger. Respondent noemt ook dat het nakijken van toetsen voor dyslectische leraren erg tijdrovend is, dus dat dat aspect bij diegene zelf niet echt van toepassing is. Er zijn waarschijnlijk ook leerlingen die meer stress ervaren in een mondeling. Dit kan evt. verminderd worden door er vroeg in de schoolloopbaan mee te beginnen. Verder noemt respondent geen nadelen van mondelinge toetsing.

Respondent vindt een goed mondeling een mondeling waarbij de leerling aan het eind tevreden is en diens cijfer begrijpt. In het geval van een onvoldoende begrijpt de leerling wat er verbeterd kan worden. Dat zorgt voor een tevreden gevoel, een gevoel dat hun begrip is gezien. Voor de docent is er dan ook een tevreden gevoel, doordat het begrip goed is getoetst (en doordat er geen nakijkwerk is).

Op de vraag: sla je jezelf wel eens voor de kop als je dan zo’n mondeling doet, dat je denkt van: dat is mijn fout, dat de leerling dat denkt, of zegt, of...? Zegt de respondent: ‘Ja, je wordt er wel echt een betere docent van. Dat je denkt van: oh, je kon dat [die vraag] ook zo begrijpen.’

Respondent geeft aan dat het goed is als de leerlingen in de les al enigzins voorbereid zijn op mondelinge toetsing. Dat kan de docent doen, door veel vragen te stellen tijdens de les, en leerlingen het gevoel te geven dat input gewaardeerd wordt, onafhankelijk van of het goed is. Verder is het goed als je de leerlingen van tevoren uitlegt wat de zin van de mondelinge toets is, en waarom het leuk is.

8.6 Appendix 6: vragenlijst mondelinge toetsen

Energie

1. Leg het verschil uit tussen kinetische en potentiële energie.
2. Geef een voorbeeld van elke vorm van energie.
3. Wat is het principe van energiebehoud?
4. Hoe wordt energie gemeten?
5. Wat is de eenheid van energie?
6. Leg het verschil uit tussen warmte- en mechanische energie.
7. Wat is het verschil tussen energie en vermogen?
8. Wat is het verband tussen arbeid en energie?
9. Hoe kan energie worden omgezet van de ene vorm naar de andere?
10. Wat zijn de verschillende vormen van energieopwekking?
11. Geef voorbeelden van hernieuwbare energiebronnen.
12. Welke natuurlijke hulpbronnen worden gebruikt voor energieopwekking?
13. Wat is de belangrijkste eigenschap van energie?
14. Hoe wordt energie overgedragen in een gesloten systeem?
15. Wat is de efficiëntie van een energieomzetting?
16. Wat zijn de voordelen van energiebesparing?
17. Leg het begrip "energieomzetting" uit.
18. Hoe beïnvloedt wrijving de energieoverdracht in een systeem?
19. Wat is het principe van energieomzetting in een machine?
20. Hoe beïnvloeden massa en snelheid kinetische energie?
21. Wat is de rol van energie in ecosystemen?
22. Wat is de relatie tussen energiegebruik en klimaatverandering?
23. Hoe wordt energie-efficiëntie gemeten?

Straling

1. Wat is straling?
2. Wat zijn de verschillende soorten straling?
3. Wat is ioniserende straling?
4. Hoe wordt straling geproduceerd?
5. Wat zijn de eigenschappen van alfastraling?
6. Wat zijn de eigenschappen van bètastraling?
7. Wat zijn de eigenschappen van gammastraling?
8. Wat is het verschil tussen ioniserende en niet-ioniserende straling?
9. Hoe beïnvloedt straling levende organismen?
10. Wat is stralingsziekte?
11. Wat is kernverval?
12. Wat is radioactief verval?
13. Wat zijn radioactieve isotopen?
14. Hoe wordt de activiteit van een radioactieve bron gemeten?
15. Wat is de halveringstijd van een radioactieve stof?
16. Wat is het verband tussen halveringstijd en radioactief verval?
17. Wat is het verschil tussen alfa-, bèta- en gamma-verval?
18. Wat zijn de risico's van blootstelling aan radioactieve straling?
19. Hoe wordt radioactieve straling gebruikt in de geneeskunde?
20. Wat zijn de veiligheidsmaatregelen bij het werken met radioactieve materialen?
21. Hoe wordt straling gedetecteerd en gemeten?

22. Wat zijn de voordelen van kernenergie?
23. Wat zijn de nadelen van kernenergie?
24. Wat zijn de risico's van kernrampen?
25. Hoe kan kernafval veilig worden opgeslagen?
26. Hoe beïnvloedt kernenergie het klimaat?

8.7 Appendix 7: foto's mondelinge toetsen



Fontein van Genève

Elke seconde wordt er 450 liter water de lucht in gestuwd tot een hoogte van 140 m.

Het water wordt met twee pompen door een spuitmond gespoten met een snelheid van 200 km/h. De twee elektrische pompen hebben elk een vermogen van 500 kW.

Na zonsondergang wordt de straal verlicht door een aantal lampen met een gezamenlijk vermogen van 13,5 kW.

Fontein in werking:

maandag tot vrijdag: 10.00 - zonsondergang
vrijdag tot en met zondag: 10.00 – 22.30 uur.



8.8 Appendix 8: Beschrijving van elke mondelinge toets

In deze paragraaf geef ik bij elke leerling wat context over hoe ik diegene ervaar, om de lezer kort te informeren over mijn beeld bij elke leerling bij ingang van de mondelinge toetsen.

Leerling 1 (vwo, 16 M):

Leerling 1 is een rustige jongen die gemiddeld een 6,5 staat voor schriftelijke natuurkundetoetsen. Hij stelt zich vrij passief op tijdens lessen; hij reageert niet vaak op uitnodiging tot interactie met de uitleg. Met ingang van de mondelinge toets verwacht ik dat hij tussen de 6 en 7,5 zal scoren.

Tijdens de mondelinge toets geeft de leerling goede antwoorden, maar hij is langzaam en niet overtuigd. Er is ook af en toe doorvragen nodig voordat hij tot een goed antwoord komt. Bij een inhoudelijk ingewikkelde vraag loopt hij helemaal vast, en het lukt hem ook met hulp niet om in eruit te komen. Hij had er niet heel veel tijd voor, omdat hij te lang met simpelere vragen bezig was, maar een snellere leerling had er wel uit kunnen komen.

Wij (de docenten) komen op een 6 uit, wat hij wel ongeveer verwachtte. Zijn schriftelijke toets was een 6,5.

Leerling 2 (havo, 16 M):

Leerling 2 is een stille jongen met een neiging voor 'fixed mindset'. Als hij ergens niet uitkomt, komt hij geregeld tot stilstand. Zijn schriftelijke toetscijfers zijn gemiddeld 7,3. Met ingang van de mondelinge toets verwacht ik een cijfer tussen de 6,5 en de 7,5.

Tijdens de mondelinge toets blijkt dat leerling 2 bijna niks kan reproduceren over straling, maar vrij goed in de stof van energie zit.

Omdat de twee onderwerpen even zwaar wegen, komen wij op een 4,8 uit. De leerling kan zich daarin vinden. Voor zijn schriftelijke toetshad hij een 5,9

Leerling 3 (havo, 17 M):

Leerling 3 is een wat actievere jongen die een 5,4 staat voor schriftelijke natuurkundetoetsen. Hij raakt makkelijk afgeleid door zijn klasgenoten, maar heeft gedurende het jaar voortgang geboekt in zijn algemene studievaardigheden, evenals in natuurkunde. Hij doet zijn huiswerk vaker en met hogere kwaliteit. Daardoor heeft hij een zekere gun-factor. Ik hoop dat hij een voldoende krijgt voor het mondeling.

Tijdens de mondelinge toets bleek dat de leerling van allebei de hoofdstukken bijna niks wist. Hij kon geen basisvragen over straling beantwoorden, en toonde geen begrip van hoe je het concept van energiebehoud kon toepassen.

Tijdens (alleen) dit mondeling was de notulist geen natuurkundeleraar, dus ik kon niet uitgebreid met diegene overleggen. Het cijfer werd een 3. De leerling snapte dat wel. Voor zijn schriftelijke toetshad hij een 3,3.

Leerling 4 (havo, 16 V):

Leerling 4 is een stil, hardwerkend meisje, dat gemiddeld een 7,3 staat voor schriftelijke natuurkundetoetsen. Ze stelt veel vragen tijdens zelfwerkijd. Meestal kost het weinig moeite om haar met wedervragen naar het goede antwoord te leiden. Mede daaruit heb ik het idee dat ze haar eigen niveau van bekwaamheid wat betreft natuurkunde niet goed inschat. Ik verwacht dat haar mondeling tussen de 6 en 7 zal uitkomen.

Tijdens de mondelinge toets blijkt de leerling zit vrij goed in de stof van energie te zitten. Er is veel twijfel, maar met wat doorvragen komt ze vaak op het goede antwoord uit. De stof van straling beheerst ze een stuk minder goed.

Wij komen op een 7 uit voor dit mondeling. De leerling is het daar wel mee eens. We hebben hierbij ons eerste resultaat (6,5) een halve punt hoger gemaakt, om te compenseren voor de storende variabele die hier speelt: haar stille karakter. Voor de schriftelijke toets bleek later dat dat misschien een goed idee was: daarvoor haalde ze een 8,4.

Leerling 5 (vwo, 16 M):

Leerling 5 is een kalme jongen met weinig motivatie, die gemiddeld een 6,8 staat voor schriftelijke natuurkundetoetsen. Tijdens de lessen gedraagt hij zich vergelijkbaar met leerling 1: weinig deelname, maar ook niet storend. Hij heeft weinig inzet nodig om goed te presteren op dat soort toetsen. Ik verwacht dat zijn mondeling tussen de 6 en 7 zal uitkomen.

Tijdens de mondelinge toets laat de leerling een degelijk begrip zien van de basis van de stof. Hij is een beetje langzaam en twijfelachtig, waardoor er weinig tijd is om de diepte in te gaan.

Wij komen uit op een 6,5, waar de leerling zich goed in kan vinden. Hij had zelf lager verwacht. Voor de schriftelijke toets had hij een 5,4.

Leerling 6 (vwo, 16 M):

Leerling 6 is een leergierige jongen, die veel meedoet tijdens de les. Voor de schriftelijke natuurkundetoetsen staat hij gemiddeld een 6,9. Zijn grootste struikelpunt is het omschrijven van formules. Hij doet erg enthousiast mee met de lessen, wat de andere leerlingen ook een beetje losser maakt. Ik verwacht dat hij tussen de 6,5 en 7,5 zal scoren op het mondeling.

Tijdens de mondelinge toets steekt deze leerling met kop en schouders boven de anderen uit. Hij kan zijn antwoorden goed articuleren. Als hij iets niet wist, was hij zich daar goed van bewust, en kon hij benoemen waar het kennisgat zat.

Wij kwamen uit op een 8, maar hij had zelf een 6 verwacht. Hij was blij verrast, en begreep het nadat we uitlegden hoe we tot het cijfer waren gekomen. Hij gaf ook aan dat er een heel ander soort vragen gesteld werd bij mondelinge toetsen, en dat je het veel beter moet begrijpen, in contrast met het vele rekenwerk van schriftelijke toetsen. Voor de schriftelijke toets had hij een 6,5.

Leerling 7 (vwo, 16 M):

Leerling 7 is een begaafde jongen, die zonder te werken de meeste van zijn toetsen haalt, hij staat gemiddeld een 6,1 voor schriftelijke toetsen. Hij is erg zelfverzekerd en ligt goed in de groep. Onder het docententeam baart deze combinatie zorgen, omdat het ervoor zorgt dat hij niet hoeft te leren. Hij hoeft geen discipline te leren, en geen goede werkhouding. Hij ziet zelf het probleem niet, omdat hij nergens vast loopt. Vanuit pedagogisch perspectief is de consensus in het docententeam, dat het goed voor hem zou zijn om lage cijfers te halen, zodat hij gedwongen wordt om aan die punten te werken. Bij de mondelinge toets zou dat best nog wel eens kunnen gebeuren: zijn wiskundige vaardigheden kunnen hem niet uit de brand helpen als hij niet geleerd heeft. Voor de mondelinge toets hoop ik dat hij goed geleerd heeft, zodat ik hem kan belonen met een hoog cijfer. Mijn verwachting is echter dat hij niet goed geleerd heeft, wat ik vanuit pedagogische overweging graag afstraf met een laag cijfer.

Tijdens de mondelinge toets blijkt inderdaad dat hij niet goed in de stof zit. Hij is goed in zich articuleren, en hij weet nog dingen van de lessen. Er ligt dus enige basis. Als de vragen ook maar een klein beetje de diepte in gaan, loopt hij vast. Tips helpen hem nauwelijks.

Wij komen uit op een 5,1. Ik had hem zelf liever een 4 gegeven, in de hoop dat hij daar iets van zou leren, maar mijn collega compenseert voor mijn beeld van de leerling om hem een eerlijker cijfer te geven. Ook zij kwam niet op een voldoende uit. De leerling kan zich hier wel in vinden. Voor de schriftelijke toets had hij een 2,5. Hij heeft hem niet herkanst, en in het volgende kwartiel heeft hij het cijfer opgehaald.

8.9 Appendix 9: Ervaringen van de docenten met de mondelinge toetsafname

In deze paragraaf bespreken de tweede corrector en ik allebei onze ervaringen met het afnemen van de mondelinge toetsen.

8.9.1 Ervaringen van de vakdocent/gespreksleider

Wat was je beeld van mondelinge toetsen vooraf? Denk aan het doel/nut ervan, en of je ze zou gebruiken

Dit onderzoek heeft me voor het afnemen van de mondelinge toetsen al veel geleerd over de toetsvorm. Mijn beeld van mondelinge toetsen bij ingang was, dat het een goede manier zou zijn om het niveau van de leerling, in samenwerking met de leerling, nauwkeurig te meten. Verder verwachtte ik dat het een prettige toetsvorm zou zijn voor beide partijen.

Waren er achteraf dingen opgevallen?

Het afnemen van de mondelinge toetsen was voor mij een positief en leerzaam proces. Naarmate ik langer bezig was, ging ik betere vragen stellen. Ook werd ik beter in het reageren op de leerling. Wat me opviel tijdens het afnemen van de mondelinge toetsen, was dat ik verwacht had dat de leerlingen beter begrip voor de stof zouden hebben. Dat is iets wat ik normaal gesproken minder goed kan toetsen, terwijl ik het wel een belangrijk onderdeel van het vak vind. De leerling wiens cijfers normaal lijdten onder zijn slechte rekenvaardigheden, kreeg hier de kans om zijn goede begrip voor de stof te laten zien. Hij vond het ook fijn dat dat gezien werd. Een aantal andere leerlingen uitte ook dat ze het gevoel hadden dat hun niveau beter getest werd dan bij een schriftelijke toets.

Is je beeld van mondelinge toetsen veranderd? In welke zin? Waarom?

Zeker. Ik had niet verwacht dat het voor de leraar zo'n leerzaam proces zou zijn. Het afnemen van mondelinge toetsen is echt een hele specifieke vaardigheid, waar je mee moet oefenen. Verder had ik niet verwacht dat de toegenomen precisie van begrip testen heel waardevol zou zijn voor de leraar, omdat de lessen aangepast kunnen worden op basis van de bevindingen. Ik denk dat de mondelinge toets veel beter werkt dan het schriftelijke toetsals feedback voor de leraar.

Wat is je huidige beeld van mondelinge toetsen?

Het beeld wat ik gedurende het onderzoek heb opgebouwd is versterkt. Ik heb veelal dezelfde ervaringen als de ervaringsdeskundigen die ik geïnterviewd heb, wat voor mij de waarde van die ervaringen als basis voor een mening versterkt. Om het anders te zeggen, heb ik het idee dat ik met zekerheid kan zeggen hoe mondelinge toetsen goed ingericht kunnen worden. Ook kan ik zeggen hoe de ervaring van de leraar, en tot in zekere mate die van de leerling, is.

Op het moment zou ik het zo samenvatten: de mondelinge toets is een persoonlijke manier om in samenwerking met de leerling nauwkeurig te bepalen waar diens niveau van beheersing van de stof ligt. Het is voor beide partijen een leerzaam proces, de leerling leert iets tijdens de toetsafname, en de leraar leert indirect over de gevolgen van diens lespraktijken. Allebei leren ze meer dan bij schriftelijke toetsen. Het is een goed middel om de band tussen leerling en leraar

te versterken, en om de leerling wat beter te leren kennen. De leerling voelt zich gezien door de leraar, wat daaraan bijdraagt.

Zou je ze zelf gebruiken in de toekomst? Waarom?

Ik zou mondelinge toetsen graag gebruiken in de toekomst, omdat ik denk dat ze in veel opzichten meerwaarde hebben vergeleken met schriftelijke toetsen, zoals ik bij de vorige vraag beschreef. Ik zie het niet als vervanging voor alle schriftelijke toetsen, omdat leerlingen ook met schriftelijke toetsen voorbereid moeten worden op de eindexamens. Verder is het onmogelijk om een grote hoeveelheid ingewikkelde rekenproblemen te toetsen met een mondeling.

Wat voor praktische consequenties zie je?

Omdat mondelinge toetsen niet met veel leerlingen tegelijk kunnen, ben je veel tijd kwijt aan het afnemen. Dat is moeilijk in te roosteren. Het is een best vermoeiend proces voor de leraren, dus je kan niet een hele werkdag vullen met mondelinge toetsen. Omdat je ongeveer drie mondelinge toetsen per uur af kan nemen, schat ik dat ik er 10 tot 15 op een dag kan doen, met een flinke pauze in het midden. Met een jaarlaag van 60 leerlingen ben je als leraar dus wel een werkweek zoet, en in die week vallen al je lessen uit, dus dat moet je in je jaarplanning verwerken. Dat geldt ook voor de tweede corrector. Mondelinge toetsen in groepen afnemen zou kunnen helpen, maar het doet ook af aan de individuele aandacht voor de leerling. Dat kan de versterkende werking die de mondelinge toetsen kunnen hebben op de band met de leerling verkleinen. Ook kan de groepsdynamiek een effect hebben op de prestaties van de leerlingen. De mate van dat effect is een goed onderwerp voor vervolgonderzoek.

8.9.2 Ervaringen van de tweede corrector/notulist

Wat was je beeld van mondelinge toetsen vooraf? Denk aan het doel/nut ervan, en of je ze zou gebruiken

Als student TN [technische natuurkunde] heb ik twee keer een mondeling gehad. De eerste keer beviel het me goed omdat ik toen de concepten goed begreep en de berekeningen een stuk lastiger vond. Ook waren de twee docenten die het toen afnamen erg gezellig en was de sfeer goed. Bij het tweede mondeling wat ik had gehad vond ik het een stuk minder prettig, omdat de moeilijkheidsgraad van de mondelinge toets toen een stuk hoger lag dan verwacht. Ook was die docent toen erg streng en was aan de lichaamstaal en intonatie op te merken dat hij niet zo tevreden was over de antwoorden die ik gaf, waardoor ik juist nog onzekerder en paniekeriger werd met antwoord geven.

Ik denk dat een mondeling een mooi instrument is om te testen hoe de parate kennis is van leerlingen en of ze concepten goed begrijpen. Dat is ook hoe ik het zelf zou inzetten.

Waren er achteraf dingen opgevallen?

Ik vond het verrassend lastig om mondelinge toetsen op een objectieve manier te beoordelen. Het scheelde dat we met z'n tweeën waren, dus één iemand kon notuleren, maar dat is ook wel echt nodig anders is het niet te doen.

Is je beeld van mondelinge toetsen veranderd? In welke zin? Waarom?

Om als docent de mondelinge toetsen te coördineren vond ik het een stuk lastiger dan verwacht.

Wat is je huidige beeld van mondelinge toetsen?

Nog steeds goed, alleen zie ik nu ook de andere kant ervan als docent en had ik van tevoren niet gedacht dat het voor de docenten zoveel werk kost. Ook was ik bang om te subjectief te beoordelen, ook al hadden we genotuleerd. Je krijgt toch ook een beeld van de leerlingen als ze bijvoorbeeld heel onzeker overkomen, wat je al dan niet onbewust mee kan nemen in de beoordeling. Met een schriftelijke toets heb ik dat zelf wat minder.

Zou je ze zelf gebruiken in de toekomst? Waarom?

Ik denk het wel, omdat ik er wel meerwaarde in zie om juist de concepten te testen bij leerlingen in plaats van (toch vaak) de berekeningen op een schriftelijke toets. Ook kan het een mooi moment zijn om individueel contact te maken met de leerlingen. Wat je op een schriftelijke toets ook minder goed merkt, is hoe snel ze op de antwoorden komen en welke denkstappen ze tussendoor maken. Met een mondeling kan je “real time” erachter komen hoe ze denken, en zie je niet alleen het uiteindelijke antwoord wat ze op zouden schrijven. Dit kan je als docent ook waardevolle informatie geven over hoe leerlingen denken, wat je in het vervolg mee kan nemen bij het ontwerpen van je lessen.

Ik kan alleen niet goed inschatten of een mondeling juist fijner zou zijn voor leerlingen met een taalachterstand, of dat het het alleen maar moeilijker maakt.

Wat voor praktische consequenties zie je?

Er gaat veel tijd zitten in het afnemen van het mondeling, met name ook omdat je het niet kan paralleliseren. En je moet goed nadenken wat je dan de andere leerlingen laat doen: laat je die voor zichzelf werken? Hoe zorg je ervoor dat ze zo min mogelijk de vragen naar elkaar doorspelen? Ook moet je ervoor zorgen dat je een collega erbij hebt, wat ook logistiek een uitdaging kan zijn.