

# MASTER'S THESIS

Oh nee, nog maar één kwartier....

**Een onderzoek naar het effect van een tijdslimiet op het probleemoplossend vermogen tijdens de toetsing van natuurkundige vraagstukken**

Brocaar, Jacomien

**Award date:**  
2019

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

## Take down policy

If you believe that this document breaches copyright please contact us at:

[pure-support@ou.nl](mailto:pure-support@ou.nl)

providing details and we will investigate your claim.

Downloaded from <https://research.ou.nl/> on date: 09. Sep. 2021

**Open Universiteit**  
[www.ou.nl](http://www.ou.nl)



# Oh nee, nog maar één kwartier....

Een onderzoek naar het effect van een tijdslimiet op het probleemoplossend vermogen tijdens de  
toetsing van natuurkundige vraagstukken

The effect of a timelimit on the problem solving ability of students during a physics test

Jacomien Brocaar

Master Onderwijswetenschappen

Open Universiteit

Datum : 24 December 2019

Begeleider : Dr. Desirée Joosten-ten Brinke / Drs. Maartje Henderikx

## Voorwoord

Voor u ligt de masterscriptie van het onderzoek ter afsluiting van mijn master Onderwijswetenschappen aan de Open Universiteit. Het onderzoek is uitgevoerd bij de opleiding Fontys Technische Natuurkunde in samenwerking met het Fontys lectoraat Technology-enhanced Assessment. Dit onderzoek en mijn studie had ik niet kunnen voltooien zonder de steun van familie, vrienden, collega's en docenten.

Allereerst wil ik Desirée Joosten-ten Brinke bedanken voor de mogelijkheid mijn masteronderzoek uit te voeren via het Fontys lectoraat Technology-enhanced Assessment en de begeleiding tijdens het onderzoek. Hierdoor kon ik tijdens mijn afstuderen mijn netwerk op het gebied van toetsing bij Fontys op een natuurlijke manier uitbreiden. De vergaderingen bij de kenniskring gaven mij als docent bij TNW zicht op wat er elders in de organisatie speelt op dit gebied. Daarnaast wil ik mijn collega's en de studenten van de opleiding Technische Natuurkunde bedanken. Mijn collega's voor de wisselwerking als ik hardop denkend mijn gedachten aan het ordenen was en voor de feedback op de conceptversies van het onderzoeksvoorstel en de masterscriptie. De studenten Technische Natuurkunde voor de deelname aan het surveyonderzoek en de leuke gesprekken op en rond het Da Vinciplein. Jullie input was een waardevolle aanvulling op mijn perspectief als docent en als onderzoeker.

Daarnaast gaat mijn grootste dank uit naar mijn gezin. Marco, Stijn, Tim en Luuk zonder jullie steun en support had ik de studie nooit kunnen voltooien. Regelmatig als ik zat te studeren zorgden jullie voor een bakje thee en wat mentale steun in de vorm van een knuffel of door mij de rust en de tijd te geven om weer een stap vooruit te zetten.

Jacomien Brocaar

Eindhoven, 16 december 2018

## Inhoudsopgave

Voorwoord .....	ii
Samenvatting .....	v
Summary .....	vi
1 Inleiding .....	1
1.1 Beoordelen van probleemoplossend vermogen in natuurkundige vraagstukken .....	1
1.2 Verschillen tussen expert en novice .....	3
1.3 Toetsing van probleemoplossend vermogen .....	4
1.4 Vraagstellingen en hypothesen.....	6
1.4.1 Deelonderzoek 1 - Surveyonderzoek .....	6
1.4.2 Deelonderzoek 2 - Toetsanalyse.....	8
2 Deelonderzoek 1 – Surveyonderzoek.....	9
2.1 Methode.....	9
2.1.1 Ontwerp.....	9
2.1.2 Onderzoeksgroep.....	9
2.1.3 Materialen.....	10
2.1.4 Procedure.....	11
2.1.5 Analyse.....	12
2.2 Resultaten .....	13
2.2.1 Interne consistentie van de onderzoekschalen.....	13
2.2.2 Onderzoeksschalen in relatie tot de populatie bij TN .....	15
2.2.3 Onderzoekschalen in relatie tot de ervaren tijdsdruk en het toetsresultaat.....	16
2.3 Conclusie en discussie.....	19
3 Deelonderzoek 2 – Toetsanalyse .....	22
3.1 Methode.....	22
3.1.1 Ontwerp.....	22
3.1.2 Onderzoeksgroep.....	23
3.1.3 Materialen.....	23

3.1.4	Procedure.....	23
3.1.5	Analyse.....	24
3.2	Resultaten.....	25
3.2.1	Vergelijking toets met extra tijd ten opzichte van een toets met reguliere toetstijd...25	
3.2.2	Besteding van de extra toetstijd.....	27
3.2.3	Aannames binnen het quasi-experiment.....	28
3.2.4	Vergelijking op basis van onderzoeksschalen.....	29
3.3	Conclusie en discussie.....	29
4	Algemene conclusie en discussie .....	31
	Referenties.....	33
	Appendix .....	38
A.	Vragenlijst Surveyonderzoek .....	38
B.	Survey .....	43
B.1	Analyse onderzoeksschaal Toetsvaardigheid .....	43
B.2	Vergelijking van groepen binnen de onderzoeksschalen.....	46
C.	Toets & Toetsanalyse .....	48
C.1	Toets .....	48
C.2	Vragen uit de evaluatie van de verlengde toets .....	50
C.3	Procedure van de analyse op basis van de Physics Problem Solving Rubric .....	51
C.4	Resultaten van de toetsanalyse op basis van eigenschappen van de populatie.....	56
D	Aannames in het quasi-experiment.....	59
D.1	Samenstelling toetsrooster samenstelling.....	59
D.2	Vergelijking resultaten 2018 en 2017 met het lange termijn gemiddelde .....	59

## Samenvatting

In het onderwijs wordt vaak gebruik gemaakt van toetsen om te bepalen of iemand voldoende kennis en vaardigheden bezit op een vakgebied. Zo'n toets is voorzien van een tijdslimiet, die kan leiden tot een ervaren tijdsdruk tijdens de toets. Ervaren tijdsdruk kan zorgen voor een betere focus op de belangrijkste prioriteiten, maar soms ook leiden tot zorgen om het gebrek aan tijd. Dit laatste is niet effectief. Zeker niet bij het toetsen van probleemoplossende vaardigheden zoals bijvoorbeeld fysische vraagstukken. In dit soort vraagstukken worden conceptuele inzichten gecombineerd met wiskundige methoden. Dat is reeds een flinke cognitieve belasting. De centrale vraag in dit onderzoek is: "Welk effect heeft een gelimiteerde toetstijd op het probleem oplossend vermogen van eerstejaars studenten?" Deze vraag is opgesplitst in de deelvragen: (1) Is de mate van ervaren tijdsdruk afhankelijk van studenteigenschappen zoals studiebenadering, toetsvaardigheden en toetsangst? (2) Is de mate van ervaren tijdsdruk van invloed op het toetsresultaat?

Het antwoord op deze vragen is gezocht met twee deelonderzoeken: een surveyonderzoek en een toetsanalyse. In het surveyonderzoek zijn eerste- en tweedejaars studenten ( $n=107$ ) bevraagd op hun studiebenadering, toetsvaardigheden, toetsangst en ervaren tijdsdruk. In het onderzoek is toestemming gevraagd om deze resultaten te mogen koppelen aan gegevens uit het studievoortgangssysteem. In het tweede onderzoek is een quasi-experiment met de toets elektromagnetisme uitgevoerd. Er is een toetssituatie met een reguliere toetstijd van 50 min ( $n=56$ ) vergeleken met een toetssituatie met een ongelimiteerde toetstijd ( $n=55$ ). De vergelijking is gemaakt met een toetsanalyse op basis van behaalde score, gebruikte toetstijd en een rubric voor het beoordelen van fysisch probleem oplossen (Docktor, 2009).

Uit het surveyonderzoek kan geconcludeerd worden dat ervaren tijdsdruk sterk positief samenhangt met toetsangst en toetsresultaatgericht werken. Ervaren tijdsdruk is binnen het survey onderzoek geen eenduidige voorspeller richting studieresultaat. Studenten geven aan dat zij ervaren dat de toetstijd de resultaten negatief ervaren. Op basis van behaalde resultaten is er geen effect in het eerste studiejaar geconstateerd. In de toetsanalyse is geconstateerd dat de tijdslimiet van invloed is op het toetsresultaat. De resultaten van de toets in de ongelimiteerde toetssituatie zijn significant beter op de belangrijkste aspecten binnen het fysisch probleem oplossen. Het advies aan de opleiding is om de toetstijd voor de toets van elektromagnetisme te verruimen en voor de andere propedeusevakken te bepalen in hoeverre toetstijd hier als beoordelingscriterium een rol mag spelen. Op wetenschappelijk gebied is het van belang te bepalen in hoeverre de resultaten van dit onderzoek te transfereren zijn naar andere kennisdomeinen.

## Summary

In education tests are used to determine whether someone possesses sufficient knowledge and skills in a discipline. Such a test is often provided with a time limit, which sometimes causes an experienced time pressure during the test. This experienced time pressure can lead to a better focus on the most important priorities, but it can also cause concerns about the lack of time. The latter is not effective. Certainly not when testing problem-solving skills, such as physical problems. In these kinds of problems conceptual insights in physics have to be combined with mathematical methods. Physics problem solving is already a considerable cognitive load. The central question in this study is: "What effect does a limited test time have on the problem solving capacity of first-year physics students?" This question is divided into the sub-questions: (1) Is the degree of time pressure dependent on student characteristics such as study approach, test skills and fear of the test? (2) Is the degree of perceived time pressure affecting the test result?

These questions are researched with (1) a survey and (2) a test analysis on a quasi-experiment. In the survey first and second year physics students ( $n = 107$ ) answered an questionnaire about their study approach, test skills, fear of testing and experienced time pressure. The study asked permission to link these results with data from the study progress system. In the second study, a quasi-experiment with a test 'Electromagnetism' was performed. A test situation with a regular test time of 50 min ( $n = 56$ ) is compared with a situation with unlimited test time ( $n = 55$ ). The comparison was made on score, used test time and aspects of the physical problem solving rubric (Docktor, 2009).

From the survey can be concluded that perceived time pressure correlates positively with test anxiety and test result-oriented work. Experienced time pressure is not a clear predictor for study results within the survey. The quasi-experiment shows that the time limit influences the test scores. The scores of the test in the unlimited time situation have a higher score and are significantly better on the most important aspects within physical problem solving. There is suggested to broaden the test time for electromagnetism and to determine for other tests in the first year to what extent a certain timelimit is part of the assessmentcriteria. In the field of educational research it is important to determine to what extent the results of this research can be transferred to other educational domains.

# Oh nee, nog maar één kwartier....

Een onderzoek naar het effect van een tijdslimiet op het probleemoplossend vermogen tijdens de toetsing van natuurkundige vraagstukken

## 1 Inleiding

Binnen het onderwijs worden modules veelal afgesloten met een toets waar een student moet laten zien dat de materie van die module begrepen wordt en toegepast kan worden. Voor beroep gerelateerde vaardigheden is het vaak van belang dat de student aantoont de kennis te kunnen toepassen in een zekere situatie (probleemoplossend vermogen) en dit ook nog binnen een gelimiteerde tijd. Het aanbrenge van een tijdslimiet kan zorgen voor een stresssituatie die de prestaties beïnvloedt. In dit geval is er sprake van een speededness (Lu & Sireci, 2007): beïnvloeding van de prestatie door tijdsdruk. Er is relatief weinig bekend over het effect van tijdsdruk bij toetsen die een sterk beroep doen op het probleemoplossend vermogen, zoals bijvoorbeeld toetsen met natuurkundige vraagstukken. Tijdens het oplossen van natuurkundige vraagstukken moeten conceptuele fysische inzichten gecombineerd worden met wiskundige vaardigheden. In de literatuur op het gebied van onderzoek naar natuurkundig onderwijs is er volop aandacht voor het effect van instructiestrategieën op basis van (concept) problem solving frameworks (Docktor, Strand, Mestre, & Ross, 2015; Kuo, Hallinen, & Conlin, 2017) of verschillen tussen natuurkundige experts en novices (Fakcharoenphol, Morpew, & Mestre, 2015; Morpew, Mestre, Ross, & Strand, 2015; Snyder, 2000). Doel van dit onderzoek is te achterhalen welk effect de ervaren tijdsdruk heeft op de prestaties van eerstejaars studenten HBO Technische Natuurkunde bij het oplossen van natuurkundige vraagstukken. Het onderzoek is uitgevoerd in twee delen: (1) een verkennend surveyonderzoek naar de ervaren tijdsdruk, studiebenadering, toetsvaardigheid en toetsangst van eerstejaars HBO-natuurkunde studenten in relatie tot behaalde toetsresultaten volgens de student en (2) een toetsanalyse van een toets waar mogelijk sprake is van een speededness-effect in vergelijking met dezelfde toets afgenomen zonder tijdslimiet op basis van de physics problem solving rubric (Docktor, 2009).

### 1.1 Beoordelen van probleemoplossend vermogen in natuurkundige vraagstukken

Voor een technisch natuurkundige is het van belang een goed inzicht te hebben in de fysische sleutelconcepten en die in praktijksituaties te kunnen toepassen (Landelijk Opleidingsprofiel Hbo-bachelor Technische Natuurkunde, 2016). In een natuurkunde opleiding is daarom veel aandacht voor het aanleggen van een brede kennisbasis van de verschillende fysische concepten en het aanleren van een systematische probleem aanpak. De systematische probleem aanpak (Tabel 1) geeft beginners in het vakgebied stapsgewijze ondersteuning (scaffolding) om de complexiteit in de vraagstukken te reduceren. De aanpak is gebaseerd op onderzoek naar verschillen tussen experts en novices. De



experts baseren hun probleemaanpak op de achterliggende concepten of fysische principes (diepe benadering) en novices kijken naar de gegevens en starten gelijk met wiskundige vergelijkingen (oppervlakte benadering) (Chi, Feltovich, & Glaser, 1981).

Tabel 1

*Verskillende stappen in een systematische aanpak voor fysische vraagstukken zoals beschreven in diverse lesmethodes, zoals problem-solving strategie hoofdstukken of kaders (Giancoli, 2014; Young & Freedman, 2012)*

	Beschrijving
1	Analyse van het probleem: Teken een fysisch diagram (bijv. een schematisch plaatje met relevante krachten)
2	Bepaal bekende en onbekende variabelen
3	Selecteer relevante fysische principes
4	Stel de vergelijkingen op passend bij de specifieke situatie
5	Uitrekenen
6	Controleren of de uitkomst realistisch is (bijv. klopt de ordegraote).

In de toetsing van het begrip van de fysische basisconcepten wordt veelal gebruik gemaakt van redelijk goed gestructureerde opgaves die een vereenvoudiging zijn van de werkelijkheid. De combinatie met de wiskundige vaardigheden die nodig zijn om de vraagstukken op te lossen, zorgt voor een hoge mate van complexiteit voor een beginner. Als een student niet komt tot een oplossing, dan kan dat verschillende oorzaken hebben. Jeniffer Docktor ontwikkelde een “physics problem solving rubric” om studenten gerichtere feedback te kunnen geven op de manier van het fysisch probleem oplossen (Docktor, 2009). Tabel 2 bevat een beschrijving van de verschillende deelgebieden uit de rubric (Effectieve beschrijving, natuurkundige aanpak, specifieke toepassing van de natuurkunde, wiskundige vaardigheden en logische voortgang). De rubric is bruikbaar om bij onsuccesvolle studenten te analyseren waar zij hun systematische probleem aanpak kunnen verbeteren. De rubric zegt echter niets over de effectiviteit van de gebruikte oplossingsstrategie (Hull, Kuo, Gupta, & Elby, 2013). Studenten met meer inzicht in de dieperliggende concepten komen soms tot alternatieve en/of efficiëntere oplossingen van het probleem. Deze grotere mate van expertise is niet terug te zien in de rubric score. Onderzoek naar het kijkgedrag tijdens het handelen van expert en novices op verschillende vakgebieden geven hier wel inzichten.

Tabel 2.

*De deelgebieden uit de physics problem solving rubric (Dockett, 2009)*

Deelgebied	Beschrijving	Score
<b>Effectieve beschrijving</b>	Feedback op de vaardigheid om de aangeboden informatie te ordenen of representeren in een passende vorm, die helpt bij het oplossen van het probleem (bijv. een tabel met belangrijkste informatie, een schematische weergave van de belangrijkste krachten, etc.).	0-5
<b>Natuurkundige aanpak</b>	Feedback op de vaardigheid van het selecteren van natuurkundige begrippen en principes die nodig zijn om het probleem op te lossen (bijv. het opstellen van een energiebehoudswet).	0-5
<b>Specifieke toepassing van de natuurkunde</b>	Feedback op de vaardigheid om de gekozen natuurkundige aanpak toe te passen in de specifieke situatie van het probleem (bijv. het verwaarlozen van warmteverliezen naar de omgeving in een geïsoleerd thermisch systeem).	0-5
<b>Wiskundige methodes</b>	Feedback op de wiskundige vaardigheden (bijv. rekenfouten, omschrijven van functies, oplossen differentiaal/integralen, gebruik van logaritmes etc.).	0-5
<b>Logische voortgang</b>	Feedback op de vaardigheid om een probleemoplossing duidelijk en logisch te structureren en te communiceren. Logisch wil zeggen dat het helder is welke stappen gezet zijn om tot de oplossing te komen.	0-5

## 1.2 Verschillen tussen expert en novice

Er is redelijk wat onderzoek uitgevoerd naar de verschillen in probleemaanpak tussen novices (beginners) en experts door met eye-tracking het kijkpatroon te analyseren (hoe en hoe lang kijkt iemand waarnaar). Onderzoeken op het gebied van vluchtcontrole (Van Meeuwen et al., 2014), medische beeld diagnoses (Gegenfurtner, Lehtinen, Jarodzka, & Säljö, 2017), geowetenschappen (Kastens, Shipley, Boone, & Straccia, 2016) en natuurkundevraagstukken (Morphew et al., 2015) laten een vergelijkbaar beeld zien: in het algemeen is een expert beter in staat taakrelevante onderdelen te selecteren en kan daardoor sneller en accurater beslissingen nemen met een lagere mentale load. Succesvolle studenten vertonen gedrag dat lijkt op dat van een expert, ze hebben kennis en vaardigheden die ervoor zorgen dat probleem oplossen makkelijker gaat. Studenten met meer expertise hebben minder tijd nodig om het probleem op te lossen, besteden meer aandacht aan taakrelevante gebieden en besteden minder aandacht aan taakirrelevante gebieden. Onderzoeken zijn uit gevoerd binnen de wiskunde (Hu, Wu, & Gu, 2017; Susac, Bubic, Kaponja, Planinic, & Palmovic, 2014), natuurkunde (Madsen, Larson, Loschky, & Rebello, 2012), chemie (Tang & Pienta, 2012) en science (Tai, Loehr, & Brigham, 2006). Dat experts zich meer richten op de dieperliggende fysische principes is ook zichtbaar in onderzoeken naar de reactietijd bij het zoeken naar verschillen tussen fysische diagrammen, zoals bijvoorbeeld een schematische weergave met de fysische krachten (Morphew et al., 2015) of als in een fysisch diagram een wijziging wordt aangebracht die het fysisch

basisprincipe beïnvloedt: “change blindness” (Feil & Mestre, 2010). Een expert ziet het snel als het fysisch basisprincipe anders is, terwijl een novice meer gericht is op de fysieke veranderingen.

Tijdens het eigen maken van een vakgebied is een beginner in ontwikkeling van novice naar expert. De manier waarop een student dit aanpakt is gedefinieerd als studiebenadering (Biggs, Kember, & Leung, 2001). Een diepe benadering van leren is volgens Ausubel een betekenisvolle benadering van het leren (Driscoll, 2005). Het leren is gericht op het begrijpen van de dieperliggende principes. Bij de oppervlakkige benadering staat het leren in het teken van onthouden en leren voor de toets (‘rote learning’ of ‘surface learning’). Een diepe benadering en dus betekenisvol leren kan enkel plaatsvinden als aan drie voorwaarden is voldaan (1) de lerende wil op zoek gaan naar de betekenis (regulatie van de lerende), (2) de leertaken en leeromgeving ingericht zijn op begrijpen (inrichten van de instructie en leeromgeving) en (3) het te leren aansluit bij de voorkennis van de lerende (uitbouwen van kennis). Een novice die zijn kennis op een betekenisvolle manier heeft uitgebreid (diepe studiebenadering) zal aangeboden vraagstukken oplossen vanuit de dieperliggende principes en waarschijnlijk meer expert-like gedrag vertonen.

### **1.3 Toetsing van probleemoplossend vermogen**

Een toets (formatief en summatief) wordt in het onderwijs vaak ingezet om de voortgang of ontwikkeling van studenten te meten. Binnen de natuurkunde toont een student veelal vakspecifieke kennis aan door een aantal vraagstukken op te lossen. De prestatie op de toets kan echter beïnvloed worden door allerlei factoren, zoals bijvoorbeeld de beschikbare toetstijd. Speededness is een maat voor de beïnvloeding van de prestatie door de tijdslimiet (Lu & Sireci, 2007). Het effect van de speededness op de validiteit van toetsen is middels een reviewstudie in kaart gebracht door Lu en Sireci (2007). Onderzoek op dit gebied is veelal uitgevoerd op basis van gegevens van computergebaseerde toetsen en richt zich bijvoorbeeld op de statistische analyse van responsetijden van items bij computergebaseerde toetsen om te bepalen of er sprake is van speededness (Van der Linden, 2011). Een andere insteek in het onderzoek naar tijdseffecten bij toetsen is meer gericht op verschillen tussen geëxamineerden, zoals geslacht (De Paola & Gioia, 2016), demografische verschillen (Y.-H. Lee & Haberman, 2016) of een studiehandicap (Sireci, Scarpati, & Li, 2005). Op het gebied van wiskundig probleem oplossen is geconstateerd dat het instellen van een tijdslimiet vooral effect heeft op studenten met een hoge mate van wiskundeangst, met name voor de vraagstukken met een hogere complexiteit (Hunt & Sandhu, 2017). De vraagstukken op een fysische toets zijn een combinatie van het kunnen toepassen van inzichten uit het fysisch vakgebied in een specifieke situatie in combinatie met wiskundige bewerkingen. In principe is het uitgangspunt dat er voldoende tijd is om de vraagstukken op te lossen, aangezien tijdens het toetsontwerp de beschikbare toetstijd een bekende randvoorwaarde is. Een student die langzamer werkt of meer moeite heeft met probleem oplossen kan echter toch tijdsdruk ervaren. Een student die tijdsdruk ervaart, kan zich

mogelijk zorgen maken over het effect van die tijdsdruk op de geleverde prestatie. Dit levert een extra belasting van het brein. Bij het werken aan wiskundige vraagstukken die veel van het werkgeheugen vragen, is geconstateerd dat studenten met meer wiskundeangst, meer last hebben van ervaren tijdsdruk (Hunt & Sandhu, 2017). Een onderzoek naar tijdsdruk tijdens toetsen bij een economie-opleiding laat zien dat de tijdsdruk meer effect lijkt te hebben op de prestatie van de vrouwen dan de mannen (De Paola & Gioia, 2016). Een eye-tracking onderzoek naar tijdsdruk en visuele aandacht bij spatiale probleem oplossing (vaardigheid om 2D of 3D beelden te kunnen roteren) toont aan dat onder tijdsdruk deelnemers met een lage mentale rotatie-vaardigheid de visuele aandacht niet goed kunnen verdelen, waardoor testcores significant lager zijn dan bij een testsituatie zonder tijdsdruk (Roach, Fraser, Kryklywy, Mitchell, & Wilson, 2017). De verwachting is dat studenten met een diepere studiebenadering meer gericht zijn op de achterliggende principes uit een vakgebied en daardoor efficiënter keuzes kunnen maken om de vraagstukken op te lossen (Madsen et al., 2012; Morpew et al., 2015). Dit leidt tot minder ervaren tijdsdruk.

Bij ervaren tijdsdruk moet er gekozen worden hoe binnen de beschikbare tijd het beste toetsresultaat gerealiseerd kan worden. Toetsvaardigheid speelt hierbij een belangrijke rol (Dodeen, 2008). Iemand met een goede toetsvaardigheid zal een betere aanpak hebben voor het time-management gedurende de toets, bijvoorbeeld door eerst de delen van de toets te maken waarvan men zeker is of die als makkelijk ingeschat worden. Een andere strategie is in gedachten de tijd per opdracht verdelen. Naarmate de expertise stijgt is men zich bewuster van de taakrelevante aspecten en achterliggende basisprincipes (Madsen et al., 2012) en kan men beter de moeilijkheid van vraagstukken inschatten (Fakcharoenphol et al., 2015). Dit leidt tot beter timemanagement tijdens de toets. De keuzes in aanpak van de toets kunnen leiden tot een ervaren tijdsdruk bij het oplossen van een vraagstuk. Dit heeft als consequentie niet of gedeeltelijk gemaakte vraagstukken of leidt tot keuzes in de informatieverwerking: versnellen, filteren of kiezen voor een strategieverandering in de informatieacquisitie (Pieters & Warlop, 1999). Versnellen betekent minder tijd besteden aan de verschillende aspecten die van belang zijn (kortere fixatietijden of minder fixatiegebieden) of een tekst scannend lezen (skimming) (Duggan & Payne, 2009). Filteren is het overslaan van bepaalde informatie omdat het de verwachting is dat het weinig toegevoegde waarde heeft voor de vraagstelling. Strategieverandering wil zeggen dat er een andere patroon ontstaat in de visuele aandacht. Kijkend naar verschillen tussen experts en novices op het gebied van het fysisch probleem oplossen, dan is te zien dat experts enkel taakrelevante regio's aandacht geven bij het oplossen van vraagstukken (Madsen et al., 2012). Dit is een indicatie van een betere filtering. Anderzijds is er te zien dat bij voldoende expertise ook andere oplossingsstrategieën toegepast worden, die minder formeel verlopen (Kuo et al., 2017) of meer flexibiliteit geven in het beoordelen van visuele informatiebronnen zoals medische beelden (Roach et al., 2017). Het flexibel kunnen wisselen van

oplossingsstrategie of informatieverwerkingsstrategie zijn typisch voorbeelden van strategieverandering.

In een toetsituatie is het goed mogelijk dat een student als gevolg van minder expertise of minder toetsvaardigheid een tijdsdruk ervaart. Als de ervaren tijdsdruk te hoog wordt, kan dat leiden tot panieksituaties, waarin een student niet meer helder kan denken. De geleverde prestatie is dan niet in overeenstemming met de werkelijke kennis en vaardigheden van de student. Een student met goede toetsvaardigheid weet wat er moet gebeuren vóór, tijdens en na een toets. De mate van toetsvaardigheid correleert positief met de houding tot toetsen en negatief met toetsangst (Dodeen, 2008). Een positieve houding tot toetsen en een lage dosis van toetsangst hebben een positief effect op prestaties (Dodeen, Abdelfattah, & Alshumrani, 2014). Andere onderzoeken zijn minder eenduidig. De relatie tussen toetsangst en prestatie is negatief bij een Zweeds onderzoek naar prestaties in een nationale science toets (Eklöf & Nyroos, 2013), echter niet significant voor meisjes en bij scheikunde vragen. Een onderzoek naar de relatie tussen toetsangst, epistemologische overtuigingen en probleemoplossend vermogen toont geen significant effect tussen toetsangst en probleem oplossend vermogen (Mehdinezhad & Bamari, 2015). Een onderzoek onder Taiwanese middelbare scholieren laat zien dat binnen de groep laagpresterende studenten de prestaties verbeteren als functie van toetsangst en binnen de groep goed goedpresterende studenten is er een negatieve correlatie tussen toetsangst en prestatie (Sung, Chao, & Tseng, 2016). De relatie tussen toetsangst en toetsresultaat is niet eenduidig in de literatuur gezien de verschillen op basis van gender (Eklöf & Nyroos, 2013) en algemene prestatie (Sung et al., 2016).

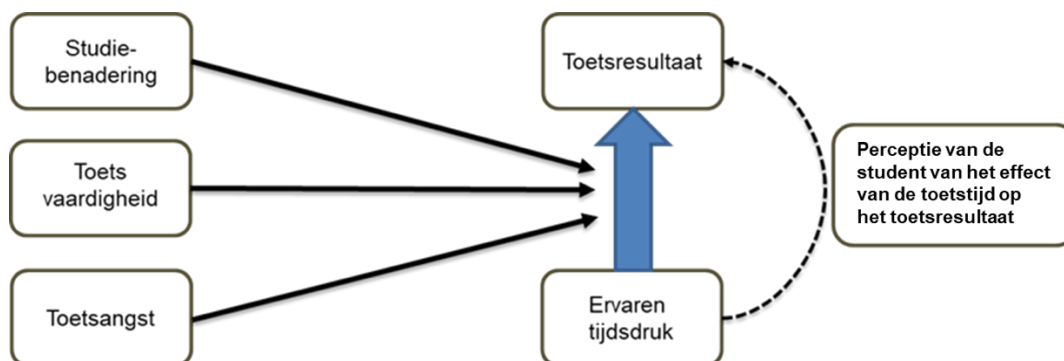
#### **1.4 Vraagstellingen en hypothesen**

Probleemoplossend vermogen wordt gemeten met een toets met gelimiteerde toetstijd. Deze gelimiteerde toetstijd kan leiden tot een ervaren tijdsdruk indien een student onvoldoende voorkennis heeft, een laag werktempo en/of te weinig toetsvaardigheid. De centrale vraag binnen dit onderzoek is: Welk effect heeft de gelimiteerde toetstijd op het aantonen van de probleemoplossende vaardigheden van eerstejaars HBO natuurkunde studenten? Deze vraag is op te splitsen in twee deelvragen: (1) Is de mate van ervaren tijdsdruk afhankelijk van studenteigenschappen zoals studiebenadering, toetsvaardigheid en toetsangst? en (2) Is de mate van ervaren tijdsdruk van invloed op het toetsresultaat? Dit wordt onderzocht op basis van een surveyonderzoek (deelonderzoek 1) en een toetsanalyse op basis van een quasi-experiment (deelonderzoek 2). De twee onderzoeken worden onafhankelijk van elkaar gerapporteerd.

##### **1.4.1 Deelonderzoek 1 - Surveyonderzoek**

Het surveyonderzoek heeft als doel om in kaart te brengen in hoeverre studenten tijdsdruk tijdens toetsen ervaren en of dit (volgens de student) van invloed is op hun toetsresultaat. Studenteigenschappen zoals studiebenadering, toetsvaardigheid en toetsangst zijn factoren die deze relatie

kunnen beïnvloeden. Figuur 1 geeft een visuele weergave van het onderzoeksmodel. Dit onderzoek sluit aan bij beide deelvragen.



*Figuur 1.* Visuele weergave van het onderzoeksmodel bij deelvraag 1. De variabelen “studiebenadering”, “toetsvaardigheid”, “toetsangst”, “ervaren tijdsdruk” en “perceptie van de student van het effect van toetsijd op het toetsresultaat” worden gemeten middels een survey onderzoek over studeren en toetsen. Het toetsresultaat in deelonderzoek 1 is geoperationaliseerd in drie indicatoren op basis van informatie uit het studievoortgangssysteem.

Hypothese 1: De verwachting is dat studenten met een diepere studiebenadering meer gericht zijn op de achterliggende principes uit een vakgebied en daardoor efficiënter keuzes kunnen maken om problemen op te lossen (Madsen et al., 2012; Morpew et al., 2015). Dit leidt tot minder tijdsdruk.

H1: Er is een negatieve samenhang tussen “de mate van diepe studiebenadering” (Biggs et al., 2001) en “ervaren tijdsdruk”.

Hypothese 2: In zijn algemeenheid zijn de studenten tevreden over de toetsing in studentevaluaties. In individuele gevallen wordt tijdsdruk wel genoemd als reden voor tegenvallende resultaten. In de vragenlijst wordt aan de studenten gevraagd welk effect zij denken dat de gelimiteerde toetsijd heeft op hun resultaten. De verwachting is dat de ‘gemiddelde’ student de toetsijd niet als een probleem ervaart. Studenten met een hoge score op ervaren tijdsdruk zullen wel een effect op het toetsresultaat verwachten

H2a : Er wordt geen ervaren tijdsdruk verwacht en daarmee geen “effect op het toetsresultaat (perspectief student)”, omdat de meeste studenten binnen de tijd de toetsen kunnen maken. De gemiddelde score op het item “Als de toetsen geen tijdslimiet hadden zou ik betere resultaten halen” is maximaal overeenkomstig met “soms”.

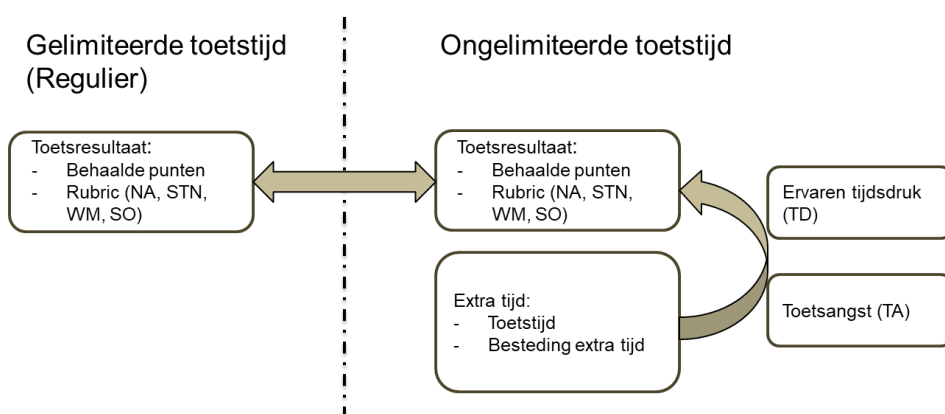
H2b : De groep studenten die veel tijdsdruk ervaart heeft een hogere score op “effect op het toetsresultaat (perspectief student)” dan de groep studenten die een lage of gemiddelde tijdsdruk ervaart.

Hypothese 3: De toetsvaardigheid is van invloed op de prestaties (Dodeen et al., 2014). Een lage score op de subschaal “time management” is een teken van tijdsproblemen gedurende de toets, dit heeft ook impact op de verwachte prestatie op de toets.

H3 Er is een negatieve samenhang tussen “toetsvaardigheid” (Dodeen, 2008) en “de perceptie van de student over het effect van de toetstijd op het behaalde toetsresultaat”.

#### 1.4.2 Deelonderzoek 2 - Toetsanalyse

In het tweede onderzoek wordt een quasi-experiment uitgevoerd. Een toets wordt aangeboden met twee verschillende toetstijden (regulier – ongelimiteerde toetstijd). Door de twee situaties te vergelijken met een toetsanalyse kan een uitspraak gedaan worden over het effect van de toetstijd op de prestaties van de studenten in deze toets. Figuur 2 is een weergave van het model dat getoetst wordt. Deelvraag 2 krijgt een antwoord op basis van deze toetsanalyse. Toetsresultaat wordt in deelonderzoek 2 beoordeeld in de vorm van een toetsscore (aantal behaalde punten op de toets) en met een score op de Physics Problem Solving Rubric (Docktor, 2009).



Figuur 2. Visuele weergave van het onderzoeksmodel behorende bij de toetsanalyse. Toetsresultaten worden vergeleken en er vindt een exploratief onderzoek plaats naar de besteding van de extra tijd.

De verwachting is dat studenten met een goede voorkennis (voldoende expertise in het vakgebied) de taakrelevante aspecten herkennen en dus weinig tijdsdruk ervaren. Het verwachte effect van de ongelimiteerde toetstijd is daardoor beperkt. Voor studenten die veel tijdsdruk ervaren tijdens toetsing is er mogelijk wel een effect. Dit leidt tot de volgende hypothesen:

- H4 Er is geen significant verschil in de gemiddelde toetsscore van de toetsen met en zonder tijdslimiet
- H5 Er is geen significant verschil in de rubricscore (Docktor, 2009) voor de deelgebieden natuurkundige aanpak (NA) en specifieke toepassing van de natuurkunde (STN)
- H6 Er is een hogere score in de rubric (Docktor, 2009) voor het deelgebied wiskundige methoden (WM) in de toets met ongelimiteerde toetstijd
- H7 De gemiddelde toetstijd van de deelnemers aan de toets met ongelimiteerde toetstijd zal hoger zijn dan de reguliere toetstijd. Het verschil is niet significant.

De toetsanalyse wordt afgesloten met een exploratief onderzoek naar de overige effecten van de extra tijd. Er wordt geregistreerd waaraan de student de extra tijd heeft besteed, welk effect dit had en de student wordt bevraagd op de mate van toetsangst, de besteding van de extra tijd en de perceptie van de student van het effect van de extra tijd op het behaalde toetsresultaat.

## **2 Deelonderzoek 1 – Surveyonderzoek**

### **2.1 Methode**

#### **2.1.1 Ontwerp**

In dit survey onderzoek zijn studenten bevraagd op “ervaren tijdsdruk” tijdens toetsen en “welk effect zij verwachten dat de tijdsdruk heeft op hun toetsresultaten”. In het surveyonderzoek zijn de variabelen “studiebenadering”, “toetsvaardigheid” en “toetsangst” meegenomen als mogelijk invloedsfactoren. Het toetsresultaat is geoperationaliseerd op basis van informatie uit het studievoortgangsregistratiesysteem.

#### **2.1.2 Onderzoeksgroep**

Voor de survey zijn studenten van de opleiding Technische Natuurkunde benaderd. Van de 261 ingeschreven studenten zijn er 205 aangeschreven om deel te nemen aan het surveyonderzoek. De overige studenten zijn uitgesloten, omdat ze bezig zijn met een bedrijfsstage, een externe minor of kenbaar hebben gemaakt te gaan stoppen met de opleiding. Tabel 3 bevat een overzicht van de beschrijvende statistieken. De onderzoeksgroep is na afname van de survey gelimiteerd tot de groep studenten uit cohort 2016 en cohort 2017. Studenten uit de latere studiefases gaven aan dat de toetsen van 50 minuten lang geleden waren.

De onderzoeksgroep bestaat daardoor uit 107 eerste en tweedejaars HBO natuurkundestudenten bij Fontys Technische Natuurkunde (cohort 2016 - 42 studenten / cohort 2017 – 65 studenten). Deze groep bestaat uit 81 mannen en 26 vrouwen (24,3% vrouw). De gemiddelde leeftijd is 19,9 jaar (17 - 26 jaar, SD=1.7 jaar). Ongeveer 33% van de studenten heeft voorzieningen toegekend in verband met een studiehandicap zoals bijvoorbeeld dyslexie, ASS of AD(H)D.



Tabel 3

Beschrijvende statistieken voor de populatie TN-studenten per april 2018, response van cohort 2016 en 2017 en de response van totale onderzoeksgroep.

Item		TN*	Response Cohort 2016- 2017	Response totaal
<b>Aantal aangeschreven voor surveyonderzoek</b>	Cohort 2016/2017 Totaal*	107 205	60 (R: 56%)	99 (R: 48%)
<b>Geslacht</b>	Man Vrouw	211 (81%) 50 (19%)	45 (75%) 15 (25%)	77 (78%) 22 (22%)
<b>Studiefase [toetstijd]</b>	P-fase [50 min] K-fase [100 min] A-fase [150 min]		35 24 1	35 30 34
<b>Cohort</b>	2017 2016 2015 2014 en ouder	65 42 46 52	39 21	39 (R: 60%) 21 (R: 50%) 24 (R: 52%) 15 (R: 29%)
<b>Studiehandicap</b>	Ja Nee		28 (47%) 32 (53%)	39 (39%) 60 (61%)
<b>Waarvan</b>	Dyslexie ASS AD(H)D		18 8 6	21 12 9
<b>Extra voorzieningen toetsing</b>	Ja Nee		20 (33%) 11	27 (27%) 12
<b>Leeftijd</b>	Gemid ± 2 SD	21,5 ± 2*2.6	19.3 ± 2*1.6	20.3 ± 2.5

\*Voetmoot : Het aantal ingeschreven studenten bij de opleiding was 261. Het aantal is gecorrigeerd voor de studenten die via hun studieloopbaanbegeleider kenbaar hebben gemaakt te gaan stoppen en studenten in een externe minor/bedrijfsstage.

### 2.1.3 Materialen

Het onderzoek start met een vragenlijst waarin de TN studenten bevraagd zijn op hun studiebenadering, toetsvaardigheid, toetsangst en ervaren tijdsdruk in toetsen. De vragen over de studiebenadering (SB) zijn afkomstig uit de R-SPQ-2F (Biggs et al., 2001). Op basis van deze vragenlijst wordt een score toegekend voor een diepe studiebenadering (SBD) en een oppervlakkige studiebenadering (SBO) met strategiekeuze en motivatie als onderliggende subschalen. Voor toetsvaardigheid (TV) is gebruik gemaakt van de Test-taking Strategy Scale (Dodeen, 2008) met daarin de subschalen “Voor de toets”, “Time management”, “Tijdens de toets” en “Na de toets”. De vragen met betrekking tot toetsangst (TA) zijn afkomstig uit een onderzoek naar de relatie tussen belang van een toets, de geleverde inspanning en toetsangst (Eklöf & Nyroos, 2013). Voor het meten van de ervaren tijdsdruk (TD) bij toetsen zijn acht items geformuleerd. Alle items in de vragenlijst zijn beoordeeld met een vierpunts Likert-schaal van zelden tot nooit [1] – soms [2] – regelmatig [3] – (bijna) altijd [4]. Tabel 4 geeft een voorbeeld van enkele vragen uit de vragenlijst en alle geformuleerde vragen voor de categorie tijdsdruk.

Tabel 4

*Voorbeelden van vragen uit de vragenlijst op gebied Studiebenadering (SB), Toetsvaardigheid (TV), toetsangst (TA) en ervaren tijdsdruk (TD). De vragen zijn voorzien van een Likertschaal: zelden tot nooit [1] – soms [2] – regelmatig [3] – vaak [4]*

Code	Voorbeeldvragen
SB2	Ik vind dat ik net zo lang moet studeren tot ik mijn eigen conclusies kan vormen over de kernbegrippen binnen het vakgebied.
SB3	Mijn doel is met zo min mogelijk werk een voldoende te halen.
TV11	Ik maak een schatting hoeveel tijd ik heb voor iedere vraag.
TV16	Ik lees iedere opgave zorgvuldig voor ik hem probeer te beantwoorden.
TA1	Ik ben kalm en zelfverzekerd als ik toetsen maak.
TA2	Voorafgaand aan toetsen ben ik bezorgd over hoe moeilijk de toets zal zijn.
TD1	Ik kan alle vragen beantwoorden binnen de geplande toetstijd.
TD2	Tijdens toetsen maak ik me zorgen of ik voldoende tijd heb om alles af te krijgen.
TD3	Door de tijdslimiet maak ik vaak rekenfouten.
TD4	Vaak moet ik opgaven sneller maken dan ik eigenlijk zou willen.
TD5	Ik laat vragen onbeantwoord, omdat daar onvoldoende tijd voor is.
TD6	Door de tijdslimiet heb ik onvoldoende tijd de vragen zorgvuldig te lezen.
TD7	Ik heb de tijd om mijn antwoorden na te kijken.
TD8	Als de toetsen geen tijdslimiet hadden zou ik betere resultaten halen.

De studenten is toestemming gevraagd voor het gebruik van resultaten uit het studievoortgangssysteem. Toetsresultaat wordt in deelonderzoek 1 geoperationaliseerd door drie indicatoren vastgesteld op basis van informatie uit het studievoortgangssysteem: (1) gemiddeld cijfer P-fase (Pgem) voor eerstejaars, (2) gemiddeld cijfer K-fase (Kgem) voor tweedejaars en (3) het aantal behaalde vakken in de propedeuse (Pn) voor de eerstejaars. Deze laatste is een norm in het toekennen van een positief studieadvies. Er wordt gebruik gemaakt van drie indicatoren, omdat toetsresultaten niet eenduidig te vergelijken zijn voor de verschillende cohorten. Tweedejaars hebben reeds meer mogelijkheden gehad om de propedeuse vakken te herkansen dan eerstejaars studenten. De variabele “perceptie van de student op het effect van de toetstijd op het toetsresultaat” is gemeten met item TD8 binnen het surveyonderzoek. Appendix A Vragenlijst Surveyonderzoek bevat de vragenlijst die de studenten hebben ontvangen.

#### **2.1.4 Procedure**

De afname van de vragenlijst heeft plaatsgevonden binnen 3 weken na de toetsperiode van kwartaal 3. Alle toetsinzages waren op dat moment afgerond. De studenten zijn via een E-mail geïnformeerd over het onderzoek. Gedurende de week van het surveyonderzoek zijn de studenten actief benaderd tijdens de projecturen om mee te werken aan het onderzoek. Het surveyonderzoek is afgenomen met het programma Parantion in verband met de privacy van de student. In het onderzoek is toestemming gevraagd om resultaten uit het studievoortgangssysteem te gebruiken. Bij toestemming moest de student de survey voorzien van naam en studentnummer.

### 2.1.5 Analyse

Allereerst zijn de items uit de vragenlijst omgezet naar de variabelen in dit onderzoek. Dit gebeurt door de items samen te voegen tot een continue schaal op basis van de gemiddelde score van de onderliggende items. Tevens wordt voor iedere schaal een ordinale variabele aangemaakt, zodat de variabele te groeperen is voor de groep studenten met een lage mate van (25%), de groep met een gemiddelde mate van (50%) en een groep met een hoge mate van (25%) die variabele. Deze groepsgewijze vergelijking geeft inzicht als de scores sterke verschillen laat zien in de extremen. Het toetsresultaat is gedefinieerd op drie manieren vanuit het studievoortgangssysteem en de student is gevraagd “tijdslimiet van de toets is van invloed op het toetsresultaat”. Tabel 5 bevat een overzicht van de gedefinieerde variabelen en hun codenaam.

Tabel 5

*Overzicht van de verschillende variabelen binnen het surveyonderzoek. De code geeft de benaming van de variabele bij gebruik als continue variabele en bij gebruik als ordinale variabele. De ordinale variant omvat 3 groepen lage mate (25%) – gemiddelde mate (50%) – hoge mate (25%).*

Variabele	Code (cont.)	Schaal	Code (ord.)
Studiebenadering (SB) – 20 items:			
• Diepe studiebenadering (SBD)	@SBD	1 – 4	SBD_LMH
• Oppervlakkige studiebenadering (SBO)	@SBO	1 – 4	SBO_LMH
Toetsvaardigheid (TV) – 31 items:	@TV	1 – 4	TV_LMH
Toetsangst (TA) – 4 items	@TA	1 – 4	TA_LMH
Ervaren tijdsdruk (TD) - 8 items	@TD	1 – 4	TD_LMH
Perceptie student “effect tijdslimiet op toetsresultaat”	TD8	1 – 4	
Toetsresultaat:			
• Gemiddelde cijfer eerstejaars	Pgem	1 – 10	
• Gemiddelde cijfer tweedejaars	Kgem	1 – 10	
• Aantal behaalde vakken in eerste jaar	Pn	0 – 15	

Vervolgens zijn de verschillende schalen uit de vragenlijst geverifieerd door het bepalen van de interne consistentie middels een factoranalyse voor de onderliggende subschalen en het bepalen van de bijbehorende cronbach alfa's. De schalen voor de studiebenadering zijn afkomstig uit de R-SPQ-2F hebben een verwachte  $\alpha$  van 0,73 (diepe benadering) en 0,64 (oppervlakkige benadering) en ongeveer 0,65 voor de vier onderliggende subschalen (Biggs et al., 2001). De schaal van studievaardigheden heeft een  $\alpha$  van 0,74 met een  $\alpha$  van ongeveer 0,6 voor de onderliggende subschalen (Dodeen, 2008). De schaal toetsangst heeft een verwachte  $\alpha$  van ongeveer 0,67 (Eklöf & Nyroos, 2013).

Vervolgens is de homogeniteit van de onderzoekspopulatie bepaald door de scores op de onderzoekschalen te vergelijken op basis van cohort (1<sup>e</sup> versus 2<sup>e</sup> jaars studenten), gender en

studiehandicap. Voor cohort wordt geen verschil verwacht, aangezien er ieder jaar vergelijkbare studenten starten bij de opleiding. In principe is de toetsing gender onafhankelijk. In de literatuur zijn er echter aanwijzingen dat de prestaties van vrouwen meer lijken te leiden onder tijdsdruk dan die van mannen (De Paola & Gioia, 2016). Voor studiehandicap worden geen verschillen verwacht, studenten met een studiehandicap krijgen extra toetstijd ter compensatie van hun studiehandicap, hierdoor hebben zij vergelijkbare toetsomstandigheden als de reguliere studenten (Sireci et al., 2005).

Tot slot wordt het effect van studiebenadering (SB), toetsvaardigheid (TV) en toetsangst (TA) op ervaren tijdsdruk (TD), perceptie van de student op het effect van de toetstijd (TD8) en toetsresultaat (Pgem, Kgem, Pn) onderzocht door de hypothesen te verifiëren. Dit gebeurt door het toetsen van hypothese 1, 2a en 3 op basis van de onderlinge correlaties tussen de continue variabelen. De verwachte correlatie tussen diepe studiebenadering (SB) en ervaren tijdsdruk (TD) is negatief (Hypothese 1). Voor de perceptie van de student over het effect van de toetstijd op het toetsresultaat (TD8) wordt een gemiddelde score van 2 of lager verwacht (hypothese 2a). Er is een negatieve samenhang tussen toetsvaardigheid en de perceptie van de student op het effect van de toetstijd TD8 (Hypothese 3).

Tevens is op basis van de ANOVA en bijbehorende post-hoc testen voor iedere gegroepeerde onderzoeksschaal bepaald of er significante verschillen zijn tussen de groepen met hoge en lage scores op de diverse onderzoeksschalen (@SBD, @SBO, @TVnieuw, @Tres, @TA, @TD), de perceptie van de student van het effect van de tijdslimiet op het toetsresultaat (TD8) en de indicatoren voor toetsresultaat (Pgem, Pn en Kgem). Voor studenten die een hoge mate van tijdsdruk ervaren (TD<sub>hoog</sub>) is de score op TD8 significant hoger dan voor studenten met lage of gemiddelde ervaren tijdsdruk (Hypothese 2b).

## 2.2 Resultaten

### 2.2.1 Interne consistentie van de onderzoekschalen

Tabel 6 bevat een overzicht van de betrouwbaarheid van de verschillende onderzoekschalen. De schalen diepe en oppervlakkige studiebenadering (Biggs et al., 2001) en toetsangst (Eklöf & Nyroos, 2013) hebben een  $\alpha > 0.7$ . Dit is een redelijke tot goede betrouwbaarheid (Field, 2013) en de waarden zijn hoger dan verwacht volgens de literatuur. De schaal studiebenadering maakt nog een onderscheid tussen items op basis van motivatie en strategie. De onderliggende schaal oppervlakkige studiebenadering (motivatie) heeft een matige betrouwbaarheid ( $\alpha=0.55$ ). Er is besloten enkel verder te werken met de hoofdschalen oppervlakkige en diepe studiebenadering. De voor dit onderzoek geformuleerde schaal ervaren tijdsdruk heeft een goede betrouwbaarheid ( $\alpha=0.88$ ).

De betrouwbaarheid van de hoofdschaal toetsvaardigheid en de onderliggende subschalen is zeer slecht en niet overeenkomstig met de literatuur (Dodeen, 2008). De items uit deze schaal zijn onderworpen aan een item-analyse en een factoranalyse (zie Appendix: B.1 Analyse onderzoeksschaal

Toetsvaardigheid). In de schaal zaten vijf items die niet relevant waren voor de toetssituatie bij Technische Natuurkunde. Deze zijn verwijderd. Tevens waren er vier items die een tweeledig effect hadden: (1) Ze zijn negatief voor het lange termijn leereffect, (2) ze kunnen echter op de korte termijn een positief effect hebben (Roediger & Karpicke, 2006a). Dit heeft geleid tot het opsplitsen van de variabele toetsvaardigheid in “Toetsvaardigheid - nieuw (@TVnieuw)” en “Toetsresultaatgericht werken (@Tres)”. De items behorend bij de schalen zijn terug te vinden in de factoranalyse (Tabel 16 in Appendix B Survey). De twee schalen zijn gedefinieerd als:

- Toetsvaardigheid – nieuw (@TVnieuw) : vaardigheden die te maken hebben met het zorgvuldig voorbereiden op en het maken van een toets, het maken van strategische keuzes tijdens een toets en het leren van toetsen.
- Toetsresultaatgericht werken (@Tres) : Studie en toetsgedrag gericht op het korte termijn resultaat, minder effectief voor het lange termijn leerrendement.

Tabel 6

*De betrouwbaarheid van de verschillende onderzoeksschalen op basis van cronbachalfa in vergelijking met de literatuur.*

Onderzoeksschaal	$\alpha$	$\alpha_{it}$	Bron
<b>Studiebenadering - SB:</b>			(Biggs et al., 2001)
Diepe Studiebenadering (10 items) - SBD	0.77	0.73	
• Diep Motivatie	0.63	0.62	
• Diep Strategie	0.59	0.63	
Oppervlakkige studiebenadering (10 items) -SBO	0.73	0.64	
• Opp. Motivatie	0.55	0.72	
• Opp. Strategie	0.60	0.57	
<b>Toetsvaardigheid (31 items) -TV:</b>	0.50	0.74	(Dodeen, 2008)
Voor de toets (8 items)	0.24	0.71	
Timemanagement (7 items)	0.35	0.75	
Tijdens de toets (11 items)	0.14	0.76	
Na de toets (5 items)	0.53	0.81	
<b>Toetsangst (4 items) - TA</b>	0.85	0.64	(Eklöf & Nyroos, 2013)
<b>Ervaren tijdsdruk (8 items) - TD</b>	0.88		
<b>Toetsvaardigheid nieuw (13 items) - TVnieuw</b>	0.66		Appendix B1
Zorgvuldig werken (9 items)	0.70		
Timemanagement (4 items)	0.61		
<b>Toetsresultaatgericht werken (4 items) Tres</b>	0.63		Appendix B1

## 2.2.2 Onderzoeksschalen in relatie tot de populatie bij TN

De onderzoeksschalen zijn vergeleken voor cohort (1e versus 2e jaars), gender en studiehandicap (Tabel 7). Er zijn geen significante verschillen in de onderzoeksschalen tussen de studenten uit het eerste (cohort 2017) en het tweede studiejaar (cohort 2016). Item TD8 vraagt de perceptie van de student in hoeverre de student denkt dat de tijdsduur van de toets heeft op het behaalde resultaat. Er is een significant verschil ( $F(1,49)=4.27, p=0.044$ ) tussen de twee cohorten. Studenten uit het eerste jaar verwachten vaker dat hun toetsresultaat beïnvloedt wordt door de tijdsduur van toetsen dan de tweedejaars studenten. TD8 is de vraag “Heeft de toetstijd een negatief effect op je toetsresultaten?”. Van de eerstejaars geeft 44% aan (bijna) altijd een negatief effect te verwachten ten opzichte van 8,3% van de tweedejaars. Dit kan komen doordat in het tweede jaar de toetstijd verdubbeld is ten opzichte van het eerste studiejaar (100 min versus 50 min).

Op basis van gender zijn er significante verschillen zichtbaar op het gebied van toetsangst en bereikte toetsresultaten. De vrouwelijke studenten ervaren significant meer toetsangst dan de mannelijke studenten ( $F(1,58)=13.32, p=0.001$ ). Op het gebied van toetsresultaten binnen de studievoortgang is te zien dat van de eerstejaars studenten de vrouwen significant minder vakken hebben behaald dan hun mannelijke collega's (Verskil 2.7 vakken,  $F(1,29)=4.63, p=0.040$ ). Het gemiddelde cijfer van de vrouwelijke studenten in de P-fase en K-fase is ook lager, deze verschillen zijn echter niet significant.

Tabel 7

*Gemiddelde scores en de standaardafwijking voor verschillende onderzoeksschalen. Uitgesplitst voor cohort, gender en studiehandicap.*

	Cohort		Gender		Studiehandicap	
	2017	2016	man	vrouw	nee	ja
<b>@SBD</b>	2.59 ± 0.52	2.63 ± 0.35	2.60 ± 0.46	2.61 ± 0.49	2.52 ± 0.42	2.70 ± 0.49
<b>@SBO</b>	2.00 ± 0.42	2.10 ± 0.45	2.08 ± 0.48	1.90 ± 0.21	2.03 ± 0.42	2.04 ± 0.46
<b>@TVnieuw</b>	2.58 ± 0.37	2.63 ± 0.42	2.55 ± 0.38	2.75 ± 0.38	2.62 ± 0.44	2.57 ± 0.34
<b>@Tres</b>	2.34 ± 0.76	2.60 ± 0.62	2.38 ± 0.71	2.58 ± 0.74	2.48 ± 0.70	2.38 ± 0.74
<b>@TA</b>	2.30 ± 0.76	2.43 ± 0.79	2.16 ± 0.67	2.92 ± 0.78	2.30 ± 0.76	2.39 ± 0.79
<b>@TD</b>	2.33 ± 0.68	2.20 ± 0.60	2.20 ± 0.62	2.58 ± 0.71	2.36 ± 0.54	2.23 ± 0.78
<b>TD8</b>	2.97 ± 1.06	2.25 ± 1.06	2.82 ± 1.14	2.77 ± 1.01	3.04 ± 0.94	2.54 ± 1.22
<b>Pgem (cohort 2017)</b>	6.51 ± 0.94		6.66 ± 0.88	5.99 ± 0.78	6.55 ± 0.90	6.52 ± 0.92
<b>Pn (cohort 2017)</b>			12.9 ± 2.5	10.2 ± 3.3	12.4 ± 3.1	12.6 ± 2.2
<b>Kgem (cohort 2016)</b>	5.28 ± 1.23		5.41 ± 1.33	5.01 ± 0.99	5.57 ± 1.22	5.04 ± 1.24

Voor de studenten die studeren met een studiehandicap zijn er geen verschillen zichtbaar op het gebied van studiebenadering, toetsvaardigheden, toetsangst, ervaren tijdsdruk en toetsresultaten ten opzichte van hun collega-studenten.

### 2.2.3 Onderzoekschalen in relatie tot de ervaren tijdsdruk en het toetsresultaat

De correlatieanalyse laat zien dat er negen significante correlaties zijn tussen de verschillende onderzoekschalen en de indicatoren voor toetsresultaat (Tabel 8). De negatieve samenhang tussen diepe studiebenadering (@SBD) en oppervlakkige studiebenadering (SBO) is overeenkomstig met de literatuur (Biggs et al., 2001). De sterk positieve correlatie tussen het gemiddelde cijfer in het eerste jaar (Pgem) en het aantal behaalde vakken in het eerste jaar (Pn) is volgens verwachting. Indien veel vakken niet behaald zijn, is het gemiddelde cijfer lager. De overige correlaties worden besproken vanuit de beantwoording van de deelvragen en het toetsen van de hypothesen.

Tabel 8

*Significante correlaties in de correlatieanalyse van de variabelen diepe studiebenadering (@SBD), oppervlakkige studiebenadering (@SBO), Toetsvaardigheden (@TVn), Toetsresultaatgericht werken (@Tres), Ervaren tijdsdruk (@TD), perceptie van effect toetstijd (TD8), gemiddeld cijfer P-fase (Pgem), Gemiddeld aantal behaalde vakken in de P-fase (Pn) en gemiddeld cijfer K-fase (Kgem) - \*  $p < .05$ , \*\*  $p < .01$  en \*\*\*  $p < .001$*

	@SBO	@TVn	@Tres	@TA	@TD	TD8	Pgem	Pn	Kgem
@SBD	-.380**					.332*			
@SBO									
@TVn						.373**			
@Tres				.428***	.375**				
@TA					.579***				
@TD						.697***			-.579*
TD8									
Pgem								0.811***	
Pn									
Kgem									

Met een variantieanalyse zijn de toetsresultaten en ervaren tijdsdruk vergeleken voor hoge, gemiddelde en lage mate van studiebenadering, toetsvaardigheid, toetsangst en ervaren tijdsdruk, zie Appendix “B.2 Vergelijking van groepen binnen de onderzoeksschalen”. De hoge en lage groepen van diepe en oppervlakkige studiebenadering (SBD en SBO) en toetsvaardigheid nieuw (TVnieuw) laten geen verschillen zien in relatie tot de ervaren tijdsdruk (@TD) of de indicatoren voor toetsresultaat (Pn, Pgem en Kgem). Tabel 9 bevat de resultaten van de variantieanalyse voor studenten met een hoge, gemiddelde en lage mate van ervaren tijdsdruk. **Fout! Verwijzingsbron niet gevonden.** geeft de scores op de indicatoren van het toetsresultaat voor de groepen met een hoge, gemiddelde en lage mate van ervaren tijdsdruk.

Tabel 9

Resultaten van de variantieanalyse voor studenten met een lage, gemiddelde en hoge mate van ervaren tijdsdruk voor studiebenadering (@SBD en @SBO), toetsvaardigheid (@TVnieuw en @Tres) toetsangst(@TA), perceptie van de student voor effect van de tijdslimiet (TD8) en de indicatoren voor toetsresultaat ( $P_{gem}$ ,  $P_n$ ,  $K_{gem}$ ).

Variabele	df	F	p
Diepe studiebenadering (@SBD)	(2,48)	1.468	.241
Oppervlakkige studiebenadering (@SBO)	(2,48)	0.576	.566
Toetsvaardigheid – nieuw (@TVnieuw)	(2,48)	3.060	.056
Toetsresultaatgericht werken (@Tres)	(2,48)	3.993	.025
Toetsangst (@TA)	(2,26.7)	15.360*	.000
Perceptie effect tijdslimiet op toetsresultaat (TD8)	(2,26.7)	61.671*	.000
Gemiddelde cijfer eerstejaars ( $P_{gem}$ )	(2,35)	0.740	.485
Aantal behaalde vakken eerstejaar ( $P_n$ )	(2,28)	5.561	.493
Gemiddelde cijfer tweedejaars ( $K_{gem}$ )	(2,10)	5.062	.030

Voetmoot: Resultaten zijn significant voor  $p < .05$ . \*Bij ongelijke varianties is F-welch bepaald.

#### *Deelvraag 1: Is de mate van ervaren tijdsdruk afhankelijk van studenteigenschappen?*

Ervaren tijdsdruk (@TD) hangt sterk samen met toetsresultaatgericht werken (@Tres) en toetsangst (@TA) (Tabel 8). Deze samenhang doet geen uitspraken over oorzaak en gevolg, echter een hoge mate van tijdsdruk wordt ervaren door studenten die een hoge mate van toetsangst en/of een hoge mate van toetsresultaatgericht werken hebben.

De samenhang tussen de diepe studiebenadering (@SBD) en ervaren tijdsdruk (@TD) is positief en niet significant ( $r=0.199$ ,  $p=.161$ ). De samenhang tussen “diepe studiebenadering” (@SBD) en “de perceptie op het effect van de toetstijd” (TD8) is positief en significant ( $r=.332$ ,  $p<.05$ ). Beide resultaten zijn tegengesteld aan de verwachting dat een diepe studiebenadering leidt tot goede inzichten, wat het efficiënt en snel problemen kunnen oplossen faciliteert (uitgangspunt voor hypothese 1). Ook op basis van de variantieanalyse is er voor ervaren tijdsdruk (@TD) geen significant verschil tussen studenten met een lage, gemiddelde en hoge mate van diepe studiebenadering ( $F(2,48)=1.468$ ,  $p=0.241$ ). Hypothese 1 is verworpen.

#### *Deelvraag 2 : Is de mate van ervaren tijdsdruk van invloed op het toetsresultaat?*

De student heeft deze vraag beantwoord in item TD8 van de survey :“Als de toetsen geen tijdslimiet hadden zou ik betere resultaten behalen”. Een positieve correlatie tussen ervaren tijdsdruk en TD8 komt overeen met een verwacht negatief effect op het toetsresultaat. TD8 is te beschouwen als de perceptie van de student. Het werkelijk behaalde toetsresultaat is gedefinieerd op basis van “gemiddelde cijfer in het eerst studiejaar” ( $P_{gem}$ ), “gemiddelde cijfer in het tweede studiejaar” ( $K_{gem}$ ) en “het aantal behaalde vakken in de propedeuse” ( $P_n$ ).

De “diepe studiebenadering” (@SBD), “toetsvaardigheid nieuw” (@TVnieuw) en “ervaren tijdsdruk” (@TD) hebben een significante samenhang met “het perspectief van de student op het effect



van toetstijd op de toetsresultaten” (TD8). Tabel 8 geeft de correlaties. Er is geen enkele significante correlatie tussen de onderzoeksschalen en de daadwerkelijk behaalde toetsresultaat in het eerste studiejaar ( $P_{gem}$  en  $P_n$ ). In het tweede studiejaar heeft de mate van ervaren tijdsdruk een negatieve correlatie met de behaalde cijfers in het tweede studiejaar,  $r=-.579$ ,  $p=0.038$ , dus meer ervaren tijdsdruk komt overeen met lagere cijfers in het tweede studiejaar.

Tabel 10

*De indicatoren voor toetsresultaat voor de verschillende mate van ervaren tijdsdruk. Het gemiddelde van alle deelnemers uit cohort 2016 en 2017 (@TD), groep met een lage mate van, gemiddelde maten en hoge mate van ervaren tijdsdruk ( $TD_{laag}$ ,  $TD_{gem}$  en  $TD_{hoog}$ )*

	@TD	$TD_{laag}$	$TD_{gem}$	$TD_{hoog}$
Evaren tijdsdruk (@TD)	$2.30 \pm 0.66$	$1.50 \pm 0.21$	$2.14 \pm 0.32$	$3.16 \pm 0.46$
Effect van tijdslimiet volgens student (TD8)	$2.80 \pm 1.10$	$1.40 \pm 0.52$	$2.85 \pm 0.95$	$3.71 \pm 0.47$
Gemiddeld cijfer 1 <sup>e</sup> jaar ( $P_{gem}$ )	$6.5 \pm 0.9$	$6.8 \pm 0.9$	$6.6 \pm 1.0$	$6.2 \pm 0.8$
Aantal behaalde vakken ( $P_n$ )	$12.5 \pm 2.7$	$13.2 \pm 1.8$	$12.7 \pm 2.8$	$11.4 \pm 3.1$
Gemiddeld cijfer 2 <sup>e</sup> jaar ( $K_{gem}$ )	$5.0 \pm 1.2$	$5.2 \pm 1.1$	$5.7 \pm 0.8$	$3.9 \pm 0.8$

Tabel 10 laat zien dat studenten die veel tijdsdruk ervaren op alle indicatoren van het toetsresultaat ( $P_{gem}$ ,  $P_n$  en  $K_{gem}$ ) het laagst scoren. Enkel voor het gemiddeld cijfer in het tweede jaar is er een significant verschil ( $F(2,10)=5.062$ ,  $p=.030$ ). Studenten met een hoge mate van ervaren tijdsdruk hebben gemiddeld lagere cijfers behaald in het tweede jaar, dan de studenten met een gemiddelde mate van ervaren tijdsdruk (verschil=-1.8,  $SE=0.5$ ,  $p=.030$ ) en lage ervaren tijdsdruk (verschil=-1.3,  $SE=0.7$ ,  $p=.275$ ). Het effect is enkele significant tussen de groepen hoge en gemiddelde ervaren tijdsdruk (Tabel 10). Het aantal deelnemers in deze groep is echter er laag ( $n=10$ ). Item TD8 (het perspectief van de student op het effect van de toetstijd op het toetsresultaat) volgt een zelfde beeld. De groepen studenten met een verschillende mate van ervaren tijdsdruk verschillen significant op item “effect van de toetstijd op het toetsresultaat volgens de student” ( $F(2,26.7)=61.671$ ,  $p=.000$ ). De varianties voor de verschillende groepen zijn ongelijk dus er is gebruik gemaakt van F-Welch.

Voor de groepen met een verschillende mate van toetsangst is er een significant verschil in het aantal behaalde vakken in de P-fase (Appendix B.2 Vergelijking van groepen binnen de onderzoeksschalen): Studenten met een lage mate van toetsangst hebben meer vakken behaald, dan studenten met een gemiddelde mate van toetsangst (verschil=2.6,  $SE=0.8$ ,  $p=.011$ ) en hoge toetsangst (verschil=1.54,  $SE=1.3$ ,  $p=.502$ ). Het effect is enkel significant tussen de groepen lage en gemiddelde toetsangst. De mate van toetsresultaatgericht werken (Tres) laat geen significant verschil zien op de indicatoren voor toetsresultaat ( $P_n$ ,  $P_{gem}$ ,  $K_{gem}$ ).

Hypothese 2a veronderstelt een gemiddelde score op “effect van toetstijd op het toetsresultaat (TD8)” lager dan 2. Dit geldt enkel voor de groep studenten met een lage mate van ervaren tijdsdruk

( $M=1.40$ ;  $SE=0.52$ ). Hypothese 2b veronderstelt dat studenten met een hoge mate van ervaren tijdsdruk een groter effect van de toetstijd verwachten dan studenten met een lage mate van ervaren tijdsdruk. Voor de perceptie van de student op het effect van de toetstijd op het toetsresultaat (TD8) is de variantie verschillend voor de verschillende mate van ervaren tijdsdruk. Op basis van Welch F is het effect van de mate van ervaren tijdsdruk op de perceptie van de student bepaald, de verschillen tussen de groepen zijn significant ( $F(2,26.7)=61.671$ ,  $p=.000$ ). Studenten met veel ervaren tijdsdruk ( $M=3.71$ ,  $SE=0.47$ ) verwachten meer nadelig effect van de beperkte toetstijd, dan studenten met weinig ervaren tijdsdruk ( $M=1.40$ ,  $SE=0.52$ ) of gemiddeld ervaren tijdsdruk ( $M=2.85$ ,  $SE=0.95$ ).

De schaal toetsvaardigheid is opgesplitst naar een positieve subschaal “toetsvaardigheid nieuw” (@TVnieuw) en negatief gecodeerde subschaal “toetsresultaatgericht werken” (@Tres). Hypothese 3 voorspelde een negatieve correlatie tussen de toetsvaardigheid (@TV) en de perceptie van de student op toetsresultaat (TD8). Hypothese 3 wordt daarmee aangepast aan de nieuwe schalen:

H3a Meer toetsvaardigheid leidt tot een beter toetsresultaat. Een student met betere toetsvaardigheden (@TVnieuw) zal minder effect verwachten van de toetstijd op het toetsresultaat (TD8) en zal een beter toetsresultaat ( $P_{gem}$ ,  $P_n$ ,  $K_{gem}$ ) behalen.

De schaal “toetsvaardigheid nieuw” (@TVnieuw) correleert positief met “de perceptie op het effect van de toetstijd” (TD8 -  $r=.373$ ,  $p<.007$ ), negatief en niet significant met “gemiddeld cijfer P-fase” ( $P_{gem}$ :  $r=-.259$ ,  $p=.106$ ), negatief en niet significant met “behaalde vakken P-fase” ( $P_n$ :  $r=-.280$ ,  $p=.128$ ) en negatief en niet significant met “gemiddeld cijfer K-fase” ( $K_{gem}$ :  $r=-.150$ ,  $p=.505$ ). Hypothese 3 is daarmee verworpen, de mate van toetsvaardigheden heeft geen positief effect op het toetsresultaat.

### 2.3 Conclusie en discussie

Het surveyonderzoek heeft als doel om in kaart te brengen in hoeverre eerstejaars HBO natuurkunde studenten tijdsdruk tijdens toetsen ervaren en of dit (volgens de student) van invloed is op het toetsresultaat. Studenteigenschappen zoals studiebenadering, toetsvaardigheid en toetangst zijn gemeten als factoren die deze relatie kunnen beïnvloeden. De HBO-natuurkunde studenten in dit onderzoek ervaren een matige tijdsdruk tijdens toetsen. De vragenlijst tijdsdruk levert een score van 2,30 als gemiddeld ervaren tijdsdruk. Dit ligt hoger dan de score 2 die overeen kwam met “soms ervaar ik tijdsdruk”. In het eerste studiejaar wordt meer tijdsdruk ervaren dan in het tweede studiejaar. Dit verschil is niet significant. Het percentage studenten dat de vraag: “Heeft de toetstijd een negatief effect op je toetsresultaten?” beantwoordt met (bijna) altijd verschilt sterk tussen de eerste en tweedejaars studenten. Van de eerstejaars geeft 44% aan (bijna) altijd een negatief effect te verwachten ten opzichte van 8% van de tweedejaars. Een belangrijk verschil tussen het eerste en het tweede studiejaar is de toetsduur (1<sup>e</sup> studiejaar: 50 minuten – 2<sup>e</sup> studiejaar: 100 minuten). Naast het verschil in toetstijd tussen eerste en tweede jaar is er ook een verschil in samenstelling van het cohort.

De groep eerstejaars studenten bevat ook nog studenten, die de opleiding gaan verlaten, zoals studenten met onvoldoende studievoortgang of studenten die na de propedeuse gaan verder studeren aan de universiteit. Beide groepen hebben een verhoogd risico richting dat ze verwachten dat de resultaten geremd worden door de toetstijd. Dit is bij de groep met onvoldoende studievoortgang, omdat ze veel tegenvallende resultaten hebben. De groep studenten met de ambitie verder te studeren aan de universiteit streeft vaak hoge resultaten na en had mogelijk de ambitie voor hogere cijfers. In het tweede jaar is daardoor de spreiding tussen de verschillende studenten wellicht kleiner, omdat deze twee groepen vertrokken zijn.

#### *Deelvraag 1 : Is de mate van ervaren tijdsdruk van invloed op het toetsresultaat?*

Op basis van de correlatieanalyse heeft de mate van ervaren tijdsdruk geen invloed op de toetsresultaten van de eerstejaarsstudenten. De eerstejaarsstudenten die een hoge mate van tijdsdruk ervaren hebben niet significant minder vakken behaald of significant lagere cijfers behaald dan de eerstejaars studenten die weinig tijdsdruk ervaren. Voor de tweedejaars studenten is er wel een significante negatieve correlatie tussen ervaren tijdsdruk en het gemiddelde cijfers in het tweede studiejaar. Vanuit het perspectief van de student is er wel een sterke correlatie tussen “ervaren tijdsdruk” en “het effect van de toetstijd op de toetsresultaten”. Op basis van de variantieanalyse is daarom onderscheid gemaakt tussen de studenten met een lage, gemiddelde en hoge mate van ervaren tijdsdruk. Daaruit blijkt dat de studenten met een hoge mate van ervaren tijdsdruk op alle indicatoren voor toetsresultaat lager scoren dan de studenten met een gemiddelde of lage mate van ervaren tijdsdruk. Alleen voor de tweedejaarsstudenten zijn de verschillen significant. Er kan dus geconcludeerd worden dat er samenhang is tussen de mate van ervaren tijdsdruk en toetsresultaat.

#### *Deelvraag 2: Zijn de onderzoeksschalen van invloed op de ervaren tijdsdruk?*

Voor de onderzochte studenteigenschappen studiebenadering (Biggs et al., 2001), toetsvaardigheden (Dodeen, 2008) en toetsangst (Eklöf & Nyroos, 2013) is bepaald welke impact ze hebben op “ervaren tijdsdruk”, “de indicatoren voor toetsresultaat” en “de perceptie van de student met betrekking tot het effect van de toetsduur op het toetsresultaat”. De onderzoeksschaal toetsvaardigheid (Dodeen, 2008) was niet consistent en daardoor niet bruikbaar. Met een factoranalyse zijn de items uit deze schaal ondergebracht in twee nieuwe schalen “toetsvaardigheid-nieuw” en toetsresultaatgericht werken. De schaal toetsresultaatgericht werken bevat de items die negatief zijn voor lange termijn leereffecten, maar op korte termijn wel een positief resultaat kunnen hebben (Roediger & Karpicke, 2006b).

“Toetsresultaatgericht werken”, “toetsangst” en “ervaren tijdsdruk” zijn onderling sterk positief gecorreleerd. Dat wil zeggen dat een student die meer toetsangst heeft en/of meer toetsresultaatgericht werkt een hogere tijdsdruk ervaart. Deze samenhang doet geen uitspraken over oorzaak en gevolg. Het is namelijk niet te onderscheiden of de tijdsdruk bijvoorbeeld ontstaat omdat een student angstig is

voor de toets of dat er angst voor de toets ontstaat, omdat er in eerdere toetsen te weinig tijd was om alles af te krijgen. Een voorbeeld van toetsresultaatgericht werken is in de voorbereiding op de toets te werken op basis van het testing-effect (H. S. Lee & Ahn, 2018; Roediger & Karpicke, 2006b; Rowland, 2014). Hierbij wordt het leren gericht vanuit het oefenen met toetsen. Een valkuil in het testing-effect is dat zaken die niet getoetst worden, minder goed onthouden worden (Butler, Karpicke, & Roediger, 2007; Yeo & Fazio, 2018). Als een student het leren richt vanuit de voorbeeldtoets in plaats van de leerstof, dan ontstaat er het risico dat niet alles geleerd wordt. Een eerdere toets is namelijk slechts een voorbeeld van een mogelijke toets en niet een weergave van alle leerdoelen. Het leren richten vanuit het zoveel mogelijk opgaven maken is niet per definitie de beste manier om natuurkundig probleem oplossen aan te leren (Byun & Lee, 2014).

#### *Overige bevindingen in het onderzoek*

“De mate van diepe studiebenadering” en “de mate van toetsvaardigheden nieuw” correleren beide positief met “de perceptie van de student dat de tijdsduur van toetsen een negatief effect heeft op het toetsresultaat”. Dit is tegengesteld aan verwachtingen voorafgaand aan het onderzoek. Hypothese 1 en 3 zijn daarmee verworpen. Een diepe studiebenadering wordt beschreven als een betekenisvolle manier van leren, waarbij nieuwe kennis en ervaringen worden gekoppeld aan reeds aanwezige voorkennis. Dit zorgt voor beter georganiseerde kennis (Driscoll, 2005). Op basis van expert-novice onderzoeken was verwacht dat bij beter georganiseerde voorkennis, sneller en adequater problemen opgelost worden (Fakcharoenphol et al., 2015; Feil & Mestre, 2010; Morpew et al., 2015). Dit effect van sneller en gericht problemen oplossen is ook gezien bij expertise verschillen tussen studenten (Madsen et al., 2012; Tai et al., 2006; Tang & Pienta, 2012). In eerdere onderzoeken is geconstateerd dat een lage toetsvaardigheid de prestaties negatief beïnvloedt (Dodeen, 2008; Dodeen et al., 2014). en dus was de verwachting dat iemand die weinig toetsvaardigheden heeft ook vaker tegenvallende prestaties zou ervaren, omdat het managen van de toets binnen de tijd niet gelukt was (hypothese 3). In dit onderzoek lijkt juist de student met een hoge mate van toetsvaardigheid meer negatieve verwachtingen te hebben over het toetsresultaat. De verwachtingen van de student komen niet overeen met de daadwerkelijk behaalde resultaten uit het studievoortgangssysteem. Er is geen significante correlatie tussen de studiebenadering, toetsvaardigheden en de werkelijk behaalde toetsresultaten.

Binnen de opleiding Technische Natuurkunde studeren relatief veel studenten met een studiehandicap (zoals dyslexie). Zij krijgen ter compensatie extra tijd toegekend tijdens toetsen (10 min extra per 50 minuten toetstijd). De resultaten van deze groep studenten is vergelijkbaar met hun reguliere collega-studenten en ook de scores voor studenteigenschappen zoals studiebenadering, toetsvaardigheden en toetsangst zijn vergelijkbaar. Dit komt overeen met de bevindingen dat studenten met een studiehandicap profiteren van extra voorzieningen, terwijl een reguliere student er geen voordeel aan heeft (Sireci et al., 2005).

In eerder onderzoek was gezien dat onder tijdsdruk vrouwen minder goed presteren dan mannen (De Paola & Gioia, 2016) en dat vrouwen meer tijd besteden aan het lezen en verwerken van informatie tijdens het oplossen van science vraagstukken (Huang & Chen, 2016). De toetsresultaten en studenteigenschappen zijn daarom ook vergeleken op basis van gender. Vrouwen ervaren meer tijdsdruk dan de mannelijke studenten, dit effect is echter niet significant. De vrouwelijke studenten hebben wel een significant hogere score op toetsangst en ze hebben significant minder vakken in de propedeuse behaald. De gemiddelde cijfers in de P-fase (1<sup>e</sup> jaar) en K-fase (2<sup>e</sup> studiejaar) zijn ook lager, maar dit is een niet significant effect. De hogere toetsangstscores voor vrouwen en de impact van toetsangst op toetsen komt overeen met andere onderzoeken (Devine, Fawcett, Szűcs, & Dowker, 2012; Eklöf & Nyroos, 2013)

### *Beperkingen van het onderzoek*

Een kanttekening bij dit survey-onderzoek is de gelimiteerde onderzoeksgroep met een klein percentage vrouwelijke studenten. De opleiding heeft slechts een beperkt aantal studenten binnen de opleiding technische natuurkunde. Hierdoor zijn alleen grote effecten meetbaar. Er is daarom ook een vergelijking gemaakt tussen de extremen van de verschillende onderzoeksschalen (lage mate van versus een hoge mate van). Deze vergelijkende analyses laten eenzelfde beeld zien als de correlatie-analyses. De populatie van TN heeft slechts 19% vrouwelijke studenten. De genderverschillen zijn daardoor gebaseerd op een klein aantal deelnemers. Voor het gebruik van informatie uit het studievoortgangssysteem heeft 73% toestemming gegeven. De relaties richting de indicatoren van toetsresultaten zijn daardoor bepaald op basis van een beperkt aantal deelnemers (eerste studiejaar:  $n=31$  / tweede studiejaar:  $n=13$ ).

Een andere beperking in het onderzoek is de niet consistente onderzoeksschaal voor toetsvaardigheid. De factoranalyse levert wel factoren die thematisch te koppelen zijn. Het aantal respondenten in het onderzoek is aan de lage kant om betrouwbare factoren te bepalen met een factoranalyse. De nieuwe groepering is minder eenduidig met de literatuur te vergelijken. De verschillende onderzoeksschalen zijn niet volledig onafhankelijk. Een voorbeeld van potentiaal afhankelijke situatie: Heeft een student toetsangst of heeft een student angst voor toetsen omdat er reeds veel toetsen niet behaald zijn. Zeker voor een student die de opleiding moet verlaten in verband met onvoldoende studievoortgang is dit nauwelijks uit elkaar te halen.

## **3 Deelonderzoek 2 – Toetsanalyse**

### **3.1 Methode**

#### **3.1.1 Ontwerp**

Het vak elektromagnetisme wordt door de studenten als complex en abstract ervaren. In studentevaluaties wordt de toetstijd van dit vak door een deel van de studenten als een belemmerende

factor benoemd. Met dit vak is een quasi-experiment opgezet door de toetstijd te variëren. De reguliere toetstijd is 50 minuten. Studenten met een erkend studiehandicap krijgen via de examencommissie extra toetstijd toegekend (+10 min). Dit studiejaar is dezelfde toets als vorig jaar aangeboden met ongelimiteerde toetstijd. Een vergelijking van de toetsresultaten moet inzicht geven in het effect van de gelimiteerde toetstijd op het resultaat.

### 3.1.2 Onderzoeksgroep

De onderzoeksgroep bestaat uit eerstejaars studenten van HBO Technische natuurkunde, die hebben deelgenomen aan de toets elektromagnetisme in juni 2017 en juni 2018. De toets is afgenomen aansluitend op het lesaanbod van dit vak. Het betreft studenten die reeds een vol jaar onderwijs bij de opleiding hebben gevolgd. In 2017 hebben 65 studenten aan de toets deelgenomen en in 2018 waren er 67 toetsdeelnemers. Een deel van de studenten maakt de toets in de herkansing. De groep herkansers is zeer divers en verschilt sterk van jaar tot jaar. Voor een zuivere vergelijking zijn alle analyses uitgevoerd op de groep reguliere toetsdeelnemers. Deze studenten maken voor het eerst de toets elektromagnetisme. Tabel 11 bevat een vergelijking van de toetssituatie in 2017 en 2018.

Tabel 11

*Vergelijking van de toetssituatie in juni 2017 en juni 2018*

	Reguliere toets (2017)	Verlengde toets (2018)
	Regulier	Regulier
Besteedde toetstijd [min]	52.5 ± 4.4	84.4 ± 22.9
Aantal deelnemers	56	55
Toegekende punten [max. 22p]	8.1 ± 4.0	12.4 ± 4.4

### 3.1.3 Materialen

De toets elektromagnetisme bestaat uit een viertal vragen, zie appendix C.1 Toets. Twee vragen waar kwantitatief aan gerekend moet worden en twee vragen waar een kwalitatief antwoord volstaat. Studenten zullen daarmee dus meer tijd kwijt zijn aan het uitwerken van de eerste twee vragen en minder aan het beantwoorden van de laatste twee vragen. De onderwerpen omvatten de wet van Coulomb, wet van Gauss en geladen deeltjes in elektrische en magnetische velden.

### 3.1.4 Procedure

Bij Technische Natuurkunde worden alle toetsen beoordeeld op basis van een toetsmatrijs en het vier-ogen-principe (tweede docent controleert de toets in relatie tot de vooraf opgestelde toetsmatrijs). In juni 2017 is de toets onder reguliere omstandigheden afgenomen. Dit wil zeggen met een toetstijd van 50 min. Studenten met een studiehandicap hebben 10 minuten extra. In 2018 is dezelfde toets aangeboden met een toetstijd van maximaal 150 minuten. Tijdens de start van deze toets is dit aan de studenten gecommuniceerd. Er is bewust gekozen, dit pas bij aanvang van de toets te communiceren

om het studeergedrag niet vooraf te beïnvloeden. Na de reguliere toetstijd (50 min) is aan alle studenten, die nog aanwezig waren in het toetslokaal, een andere kleur pen gegeven. Hierdoor is zichtbaar welke informatie buiten de reguliere toetstijd is toegevoegd. Van iedere deelnemer is het tijdstip van inleveren geregistreerd. Na afloop van de toets is aan de studenten een evaluatievragenlijst met negen vragen voorgelegd. In deze vragenlijst zijn de studenten bevraagd op (1) manier van werken tijdens deze toets zonder tijdsdruk, (2) de mate van toetsangst voor deze toets en (3) de perceptie op de ideale toetstijd voor de propedeusevakken van de opleiding. Tevens had de student via dit formulier de mogelijkheid om nog feedback te geven aan de onderzoeker.

De vakdocent heeft de toets beoordeeld en kent een cijfer toe op basis van de score op de verschillende toetsvragen (toetsresultaat). De toetsresultaten zijn besproken met de tweede docent en vervolgens teruggekoppeld aan de studenten via Progress (de studievoortgangsregistratie). Na de toetsinzage zijn de toetsuitwerkingen ter beschikking gesteld aan de onderzoeker voor een uitgebreide analyse van alle studentwerken. De studentuitwerkingen van beide toetsen zijn beoordeeld op basis van de physics problem solving rubric (Docktor, 2009). Deze rubric beoordeelt de deelgebieden (1) effectieve beschrijving, (2) natuurkundige aanpak, (3) Specifieke toepassing van de natuurkunde, (4) wiskundige methodes en (5) logische samenhang, elk met een score van 0 t/m 5. Een lage score op (2) natuurkundige aanpak of (3) specifieke toepassing van de natuurkunde is een indicatie dat de natuurkundige concepten niet begrepen zijn. Een lage score op (4) wiskundige methode geeft een indicatie dat de wiskundige vaardigheden ontbreken of dat er veel rekenfouten worden gemaakt. De rubric is niet onderscheidend als een student de basisprincipes zeer goed beheerst en daardoor alternatieve oplossingsstrategieën kan toepassen (Hull et al., 2013). In dit geval wordt er geregistreerd dat een student meer dan gemiddelde inzichten vertoont (wel / geen short cut oplossing toegepast). In appendix C Toets & Toetsanalyse staat de beschrijving van de beoordelingsaspecten in de rubric uitgewerkt voor de toets van elektromagnetisme.

### **3.1.5 Analyse**

De toetsanalyse is een quasi-experiment waarbij twee testcondities met een t-toets worden vergeleken op toetstijd en toetsresultaat (behaalde punten en de physics problem solving rubric (Docktor, 2009)). Tevens wordt de samenhang tussen toetstijd en toetsresultaat bepaald. De toets is ontworpen met de reguliere toetstijd als uitgangspunt en toetst het correct toepassen van de fysische basisconcepten. Een student met voldoende kennis (expertise) zal snel de taakrelevante aspecten herkennen en daaruit passende conclusies trekken. Dit is grotendeels bepalend voor het aantal behaalde punten. Daarnaast kan een student nog fouten maken in de wiskunde handelingen. De verwachting is dat er geen significant effect wordt ervaren op het aantal behaalde punten (hypothese 4) en de deelgebieden natuurkundige aanpak en specifieke toepassing van de natuurkunde (hypothese 5). Mogelijk is er wel een effect voor het deelgebied van de wiskundige methoden (hypothese 6).

Aangezien de toets is ontworpen met 50 minuten als uitgangspunt is de verwachting dat er geen tijdsprobleem zal zijn. Een student zal echter van de mogelijkheid gebruik maken, omdat het kan. In de ongelimiteerde toetstijd is een hogere toetstijd verwacht, echter niet significant (hypothese 7). Met een variantieanalyse is onderzocht of er verschillen zijn op basis van studiehandicap, gender en toetsangst. De score voor toetsangst is bepaald op basis van 4 items (Eklöf & Nyroos, 2013) in het toetsevaluatieformulier. De respondenten zijn vervolgens onderverdeeld in drie groepen (lage toetsangst (25%) – gemiddelde toetsangst (50%) – hoge toetsangst(25%)).

Tot slot is met exploratief onderzoek gekeken naar het effect van de verlengde toetstijd. Dit is gerealiseerd door te kijken naar (1) de besteedde toetstijd en naar (2) wat doet de student als de reguliere toetstijd voorbij is. Hiervoor wordt gebruik gemaakt van de evaluatie door de studenten (student perspectief) en de uitwerkingen van de studenten. Na afloop van de reguliere toetstijd van 50 minuten is de student overgestapt op een andere kleur pen. Op basis van de veranderde penkleur is in te schatten welk percentage van de toets gemaakt is binnen de reguliere toetstijd. Het werk na het verstrijken van de toetstijd wordt gecategoriseerd op (1) het maken van (de laatste) opgaven (positief voor het toetsresultaat), (2) aanpassen van reeds gemaakt werk (verbetering/positief voor het toetsresultaat) (3) een combinatie van afmaken en aanpassen met positief resultaat of (4) aanpassen van reeds gemaakt werk (neutraal of negatief effect).

In het quasi-experiment zit de aanname dat de toetssituatie identiek is voor 2017 en 2018. De digitale leeromgeving en de vakdocent is hetzelfde in beide studie jaren. Voor de volgende twee aannames in de toetssituatie is geverifieerd of ze vergelijkbaar zijn. De aanname dat de studentenpopulatie vergelijkbaar is, is geverifieerd door voor alle vakken uit periode 4 het toetsresultaat te vergelijken met het vijfjarig gemiddelde met behulp van een t-test. Tevens is op een kwalitatieve manier de positionering van elektromagnetisme in de toetsweek vergeleken om te bepalen of het moment van toetsafname van invloed is.

## **3.2 Resultaten**

### **3.2.1 Vergelijking toets met extra tijd ten opzichte van een toets met reguliere toetstijd**

In de toetssituatie zonder tijdsdruk gebruiken de studenten significant meer toetstijd dan in de reguliere toetssituatie (verschil gemiddeld bijna 32 minuten) en zijn er significant hogere toetsresultaten (Tabel 12). Er is gemiddeld 4,3 punt meer toegekend op de toets met de verlengde toetstijd dan in de reguliere toets ( $t(110)=5.389, p=.000$ ). Iedere opgave heeft een significante verbetering van de score, zelfs opgave 1 waar de score reeds initieel hoog was. In beide toetsen zijn geen short-cut oplossingen aangetroffen. Geen van de studenten heeft de opgaven opgelost met een andere oplossingsstrategie.

De uitwerkingen van beide toetsen zijn beoordeeld met de Physics Problem Solving Rubric (Docktor, 2009). Tabel 12 bevat de resultaten van de vergelijking. Er is te zien dat in de toets met



extra tijd de scores voor “Natuurkundige aanpak” (verschil: 0.77;  $t(110)=4.21, p=0.000$ ), “Specifieke toepassing van de natuurkunde” (verschil: 1.01;  $t(110)=5.41, p=0.000$ ) en “Logische voortgang” (verschil: 0,84;  $t(110)=4.26, p=0.000$ ) significant hoger zijn dan in de reguliere toetsituatie. De score voor “wiskundige methodes” is niet afhankelijk van de toetstijd.

Tabel 12

*Verlengde toetstijd vergeleken met de reguliere toetstijd met een onafhankelijke t-toets (M: gemiddelde, SE : standaarddeviatie,  $\mu$  : verschil, t : t-toets, p : probability, d : Cohen's d).*

	Reguliere toetstijd		Verlengde toetstijd		T-toets			
	M	SE	M	SE	$\mu$	t	p	d
Toetstijd [min]	52.5	4.4	84.4	22.9	31.9	10.12	0.000	7.25
Aantal behaalde punten [22]	8.1	4.0	12.4	4.4	4.30	5.39	0.000	1.08
Aantal pagina's	3.46	0.96	4.13	1.14	0.67	3.39	0.001	0.70
Effectieve beschrijving	2.79	0.89	2.62	0.89	-0.18	-1.06	0.291	0.20
Natuurkundige aanpak	2.67	0.97	3.43	0.96	0.77	4.21	0.000	0.79
Specifieke toepassing	2.04	0.98	3.05	1.00	1.01	5.41	0.000	1.03
Wiskundige methodes	2.57	1.31	2.71	1.10	0.14	0.63	0.53	0.11
Logische voortgang	2.42	1.05	3.26	1.04	0.84	4.26	0.000	0.80
Opg 1 [6p]	4.58	1.41	5.46	0.75	0.88	4.10	0.000	0.62
Opg 2 [7p]	1.64	1.88	3.26	2.08	1.62	4.32	0.000	0.86
Opg 3 [4p]	0.93	1.22	1.68	1.62	0.75	2.77	0.007	0.61
Opg 4 [5p]	0.96	1.13	2.03	1.70	1.06	3.90	0.000	0.93

*Voetnoot:* Cohen's d geeft indicatie van de grootte van het effect. Groot effect  $d>0.80$ ; Klein effect  $d<0.20$

In appendix C.4 Resultaten van de toetsanalyse op basis van eigenschappen van de populatie staan de resultaten voor een toetsanalyse als er vergeleken wordt op basis van studiehandicap, gender of mate van toetsangst. In de eerste analyse is er gekeken naar de studenten met een studiehandicap (studenten die extra tijd toegekend hebben gekregen van de examencommissie) in vergelijking met de reguliere studenten. Studenten met een studiehandicap gebruiken gemiddeld 18 minuten meer dan hun medestudenten. Dit verschil is significant volgens de geplande contrast vergelijking tussen reguliere student en student met studiehandicap binnen de verlengde toetstijdsituatie,  $t(31,12)=-2.95, p=.006$ . Voor alle andere aspecten zoals toegekende punten en de beoordelingsaspecten uit de Physics Problem Solving rubric zijn er geen significante verschillen tussen de reguliere studenten en de studenten met een studiehandicap. In de tweede analyse is gekeken naar genderverschillen. Hieruit blijken geen significante verschillen aanwezig te zijn tussen man en vrouw.

In de laatste analyse is voor de score in de toetsituatie met verlengde toetstijd gekeken naar verschillen tussen de groepen studenten met een lage, een gemiddelde en een hoge mate van toetsangst. Er is wel te zien dat de groep met hoge toetsangst overal lager scoort dan de groep met lage of gemiddelde toetsangst. De verschillen tussen de groepen zijn echter niet significant. Bij het

vergelijken van de groep met een hoge toetsangst met behulp van geplande contrasten ten opzichte van de andere twee groepen, dan is de conclusie dat studenten met een hoge mate van toetsangst significant minder punten scoren dan studenten met een lage mate van toetsangst ( $t(32)=2.248, p=.32, r=.37$ ) en studenten met een gemiddelde mate van toetsangst ( $t(32)=2.063, p=.047, r=.34$ ). Het is in beide gevallen een medium effect (Field, 2013). Voor de “gemiddelde specifieke toepassing van de Natuurkunde” is een medium effect aanwezig als de vergelijking wordt gemaakt tussen studenten met een hoge toetsangst en studenten met een lage toetsangst ( $t(32)=2.207, p=.035, r=.36$ ) en studenten met hoge toetsangst en een gemiddelde toetsangst ( $t(32)=2.319, p=.027, r=.38$ ).

### 3.2.2 Besteding van de extra toetstijd

Na afloop van de toets met verlengde tijd is er door de deelnemers aan de toets een korte vragenlijst ingevuld als evaluatie van de verlengde toetssituatie. Hierin hebben 37 van de 55 studenten toestemming gegeven om de resultaten uit de vragenlijst te gebruiken (67 % van de toetsdeelnemers). In deze vragenlijst zijn de studenten bevraagd op (1) de manier van werken tijdens deze toets zonder tijdsdruk, (2) de mate van toetsangst voor deze toets en (3) de perceptie op de ideale toets tijd voor onze propedeusevakken. Tabel 13 bevat een overzicht van de resultaten van de evaluatie. Voor de manier van werken is het percentage aangegeven van de respondenten die dit ervaren.

Tabel 13

*Overzicht van de resultaten uit de vragenlijst na afloop van de toets met onbeperkte toetstijd*

Categorie	Item	Score
Manier van werken	Ik heb de vragen aandachtiger gelezen	73% Beetje / volledig eens
	Ik heb netter en overzichtelijker gewerkt	75% Beetje / volledig eens
	Ik verwacht een beter resultaat	91% Beetje / volledig eens
Toetsangst*	Lage toetsangst (n=11)	1.84 ± 0.20
	Gemiddelde toetsangst (n=15)	2.48 ± 0.18
	Hoge toetsangst (n=9)	3.22 ± 0.36
Ideale toetstijd	50 - 65 min (max kwartier extra)	49 %
	70 – 100 min (tot 2x zo lang)	45%
	>100 min (meer dan 2x zo lang)	6%

*Voetnoot:* \* Toetsangst is uitgedrukt als de gemiddelde score op de vier items gerelateerd aan toetsangst [schaal 1-4].

Op het moment dat de reguliere toetstijd is afgelopen is aan alle studenten een andere kleur pen verstrekt. Van iedere toetsuitwerking is geregistreerd hoeveel pagina's er in de reguliere toetstijd zijn gemaakt en hoeveel pagina's zijn beschreven buiten de reguliere toetstijd (kwantiteit) en er is bepaald hoeveel punten zijn toegekend (kwaliteit) buiten de reguliere toetstijd (op basis van de uitwerking in penkleur 2). Tabel 14 bevat het overzicht voor deze factoren. Tevens is een indicatie geregistreerd van het uitgevoerde werk in de extra tijd: (1) maken van ontbrekende opgaven (2) aanpassen eerder gemaakt werk (positief resultaat) (3) combinatie van aanpassen en afmaken (4) aanpassen eerder

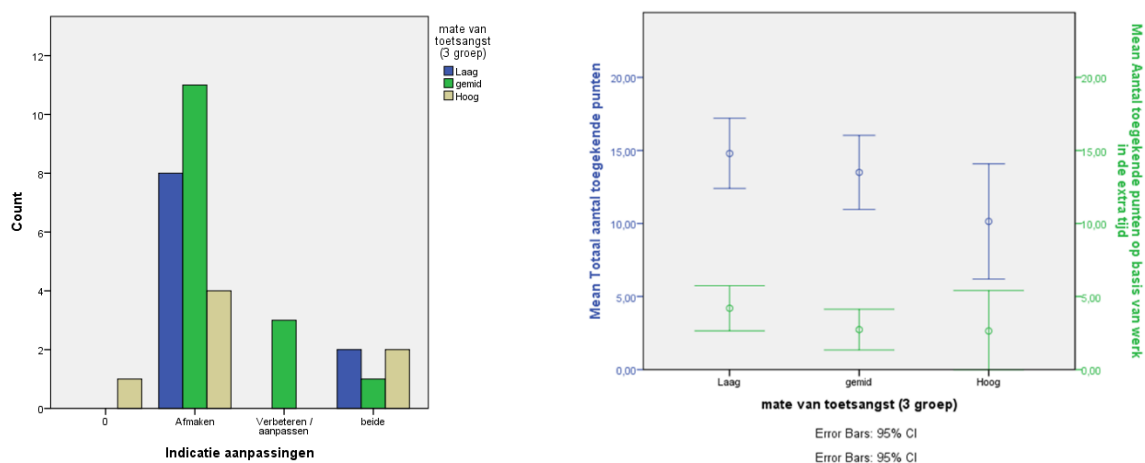
gemaakt werk (negatief resultaat). De besteding van de extra tijd is vergeleken voor de drie groepen binnen de mate van toetsangst. Figuur 3 geeft een visuele weergave van de vergelijking. Hierin is te zien dat de studenten de tijd voornamelijk gebruiken om het werk af te maken (linker figuur).

Daarnaast vergelijkt de rechter grafiek de verschillende groepen binnen mate van toetsangst op aantal behaalde punten en aantal behaalde punten buiten de reguliere toetstijd. De verschillen zijn niet significant (aantal behaalde punten:  $F(2,32)=2.949, p=.067$  – Toegekende punten op basis van werk in de extra tijd:  $F(2,29)=1.209, p=.313$ ). Ondanks dat de verschillen niet significant zijn, is wel de trend zichtbaar dat studenten met een hogere mate van toetsangst een lagere score hebben.

Tabel 14

*Kwantificering van het werk uitgevoerd in de extra tijd. Inschatting is gemaakt op basis van het werk in de tweede penkleur ten opzichte van het volledige werk. Kolom gemiddeld is bepaald op basis van alle toetsdeelnemers. Kolommen mate van toetsangst ( $TA_{laag}$ ,  $TA_{gem}$  en  $TA_{hoog}$ ) zijn bepaald op basis van de studenten die de evaluatie hebben in gevuld*

	Gemiddeld	$TA_{laag}$	$TA_{gem}$	$TA_{hoog}$
Aantal pagina's tot	4.1 ± 1.1	4.2 ± 1.1	4.2 ± 0.9	3.5 ± 1.2
Pagina's in extra tijd	1.3 ± 0.8	1.0 ± 0.6	1.0 ± 0.7	1.3 ± 0.8
Toegekende punten (totaal)	12.4 ± 4.4	14.1 ± 4.0	13.5 ± 4.6	9.8 ± 3.8
Toegekend in de extra tijd	3.1 ± 2.6	4.2 ± 2.2	2.7 ± 2.5	2.6 ± 3.0
Verschil (punten in reguliere tijd)	10.0 ± 3.4	10.6 ± 3.2	10.8 ± 3.6	7.5 ± 2.0
<i>n</i>	56	11	15	9



Figuur 3 (a - links) Visuele weergave van de besteding van het werk uitgevoerd in de extra tijd gegroepeerd voor de mate van toetsangst. (b - rechts) Vergelijking van het aantal behaalde punten op de toets en het aantal behaalde punten in de extra tijd voor de groepen lage, gemiddelde en hoge toetsangst.

### 3.2.3 Aannames binnen het quasi-experiment

De vergelijking tussen de twee toetsen is een quasi-experiment. Dit betekent dat er naast de aangebrachte manipulatie (reguliere toetstijd – ongelimiteerde toetstijd) ook een aantal andere variabelen een rol spelen die van invloed kunnen zijn op het toetsresultaat:

- **Toetsdeelnemers:** De reguliere toetsdeelnemers zijn ieder jaar de studenten die net hun vierde periode onderwijs bij Technische Natuurkunde hebben afgerond. Gemiddeld genomen zal deze groep ieder jaar vergelijkbaar zijn, het zijn echter fysiek andere studenten. De resultaten van cohort 2016 en 2017 zijn daarom voor alle vakken in periode 4 vergeleken met het lange termijn gemiddelde (5 jaar). Hieruit blijkt dat cohort 2016 presteert conform het gemiddelde (geen enkel vak wijkt significant af van het lange termijn gemiddelde) en cohort 2017 haalt voor alle vakken betere resultaten (voor 3 vakken significant verschil – voor 3 vakken niet significant).
- **Samenstelling toetsrooster:** De plek in de toetsweek kan van invloed zijn op de prestaties. Met 6 toetsen in 3 tot 4 toetsdagen binnen anderhalve week komt het voor dat er 2 toetsen op één dag plaatsvinden. In 2018 was de toets van elektromagnetisme ingepland op de woensdag in de eerste toetsweek en als enige toets die dag ingepland (voordeel). In 2017 vond de toets plaats op de dinsdag van de tweede toetsweek (voordeel: 3 voorafgaande studeerdagen) samen met een andere toets (nadeel).

Het lijkt erop dat het toetsrooster van 2018 en de fysieke samenstelling van deze lichte studenten een hogere verwachting geven voor het resultaat van de toets van elektromagnetisme. Appendix D Aannames in het quasi-experiment bevat de verantwoording van deze verwachting. Op basis van de beschikbare gegevens kan dit niet verder gekwantificeerd worden.

### 3.2.4 Vergelijking op basis van onderzoeksschalen

Door slechts 43 van de toetsdeelnemers is ook het volledige survey-onderzoek ingevuld. Dit aantal deelnemers is erg laag. Er zijn geen significante effecten in de variantieanalyse naar de verschillen tussen de studenten met een hoge en lage mate van de verschillende onderzoeksschalen voor de behaalde toetsresultaten gevonden. Een koppeling met deelonderzoek 1 is niet verantwoord.

## 3.3 Conclusie en discussie

De toetsanalyse is ingezet om het effect van de toetstijd op de probleem oplossende vaardigheden te bepalen. De student maakt dezelfde toets in een situatie met en zonder ervaren tijdsdruk. De toetstijd heeft een significant effect op de behaalde toetsresultaten.

### *Is er verschil in toetsresultaat tussen toetsen met en zonder tijdsdlimiet*

In het tweede deelonderzoek zijn de resultaten van een toets elektromagnetisme vergeleken voor een afname onder reguliere toetstijd (juni 2017 – 50 min) en zonder toetslimiet (juni 2018 – 150 min). De resultaten van de toets zijn beoordeeld op toegekende punten door de vakdocent. De besteedde toetstijd door de student en de uitwerkingen van alle toetsdeelnemers zijn door de onderzoeker beoordeeld. Hierbij is gebruik gemaakt van de Physics Problem Solving Rubric (Dockett, 2009). De toets met ongelimiteerde toetstijd is significant beter gemaakt dan de toets met reguliere toetstijd. Bij de toets zonder toetstijd besteden de studenten ruim 30 minuten meer aan de toets dan de reguliere

toetstijd van 50 minuten en ze scoren 20% meer punten. De beoordeling met de rubric problem solving rubric laat zien dat onder toetsomstandigheden zonder tijdsdruk de “natuurkundige aanpak”, “specifieke toepassing van de natuurkunde” en de “logische voortgang” significant verbeteren. Hiermee kan er geconstateerd worden dat het fysisch probleem oplossen verbetert als de student meer tijd krijgt.

De toetsen zijn ook vergeleken voor de studenten met een studiehandicap. Studenten met voorzieningen voor extra tijd gebruiken gemiddeld 19 min meer en behalen een vergelijkbare score als hun reguliere collega studenten. Dit is in lijn met eerder onderzoek naar het toekennen van extra voorzieningen aan studenten met een studiehandicap: Toekennen van extra tijd heeft positieve effecten voor studenten met een studiehandicap, maar levert nauwelijks iets op voor reguliere studenten (Sireci et al., 2005). Indien de resultaten worden opgesplitst op basis van gender zijn er geen significante verschillen geconstateerd tussen de resultaten van mannen en vrouwen. Ondanks dat het verschil niet significant is, hebben de vrouwelijke studenten een 13% lagere score dan hun mannelijke collega's. Dit is ook gezien in deelonderzoek 1.

Een onderzoek naar het gebruik van formatieve quizen met tijdsdruk bij het leren over constructieveiligheid (Jaeger & Adair, 2017) zorgt ervoor dat er minder met trial-en-error geleerd wordt en studenten zijn positief over de ervaren tijdsdruk die overeenkomt met de praktijksituatie. In dit onderzoek leidt de prestatie niet onder stress door de tijdsdruk. Echter de toets met tijdsdruk vond plaats tijdens het leren. De uiteindelijke toets op basis waarvan de leerresultaten werd beoordeeld was voor alle studenten identiek.

#### *Exploratief onderzoek naar de besteding van de extra tijd*

De studenten geven in de evaluatie zelf aan aandachtiger te lezen en netter/overzichtelijker te werken. Dit wordt ook geconstateerd in de analyse van de gemaakte werken. Tijdens de toets is iedere student voorzien van een andere kleur pen. Op basis van de kleurverschillen in de uitwerking is een inschatting gemaakt van de besteding in de extra tijd door de studenten. De studenten werken gemiddeld 60% langer aan de toets en er is circa 25% van de punten toegekend in de tweede penkleur (buiten de reguliere toetstijd). Op basis van mate van toetsangst zijn er geen significante verschillen waargenomen, Er is echter wel een duidelijk verschil in score zichtbaar. Studenten met een hoge mate van toetsangst (bepaald na afloop van de toets) behalen met extra tijd een vergelijkbaar aantal punten dan de studenten met een gemiddelde en een lage toetsangst binnen de reguliere toetstijd.

#### *Beperkingen van het onderzoek*

De toetsanalyse is een quasi-experiment. De studenten van dit jaar zijn niet dezelfde studenten onder dezelfde omstandigheden als de deelnemers aan de toets van vorig jaar. Op basis van resultaten in andere vakken lijkt het aannemelijk dat de studenten, die in juni 2018 aan de toets hebben deelgenomen gemiddeld iets beter presteren dan het vijfjarig lange termijn gemiddelde. Dit is in het

voordeel van de toets met de interventie. Tevens waren in juni 2018 de omstandigheden iets gunstiger. Er was die dag slechts één toets opgenomen in het toetsrooster. Het effect als gevolg van de toetstijd is daardoor mogelijk overschat. Het effect van een 20% hogere score is echter groot.

De Physics problem Solving Rubric is ontworpen om feedback te geven aan individuele studenten. In dit onderzoek is het gebruikt als beoordelingsinstrument. Dat is alleen mogelijk als er goed wordt vastgelegd hoe de verschillende dimensies uit de rubric worden gedefinieerd. In appendix C Toets & Toetsanalyse is vastgelegd hoe er beoordeeld is. Twee collega docenten hebben deze uitwerking gebruikt en komen tot vergelijkbare beoordelingen voor de deelgebieden, nadat de beoordelingsmethodiek samen besproken is.

Een toets met een onbeperkte toetstijd is geen reële situatie. De student weet vanaf de aanvang van de toets dat de toetstijd verlengd wordt. Het werk in het eerste deel is daardoor geen “normaal” werk. Het aantal gescoorde punten binnen de toetstijd is daardoor niet gelijk aan de score binnen de reguliere toetstijd. In dit onderzoek wordt ervan uitgegaan dat de studentenpopulatie van de opleiding technische natuurkunde over de jaren heen vergelijkbaar is. Dit is natuurlijk niet het geval. Er is een design overwogen waarbinnen de huidige populatie de toets deels onder tijdsdruk en deels zonder tijdsdruk werd afgenomen. De toetsomstandigheden voor de studenten worden dan echter te veel verstoord en het zorgt voor ongewenste ongelijkheidsverschillen. Het afnemen van dit onderzoek zorgt voor een iets andere toetsprocedure dan gebruikelijk. Dit zou voor sommige studenten het toetsproces kunnen verstoren. Er is echter sprake van een ongelimiteerde toetstijd, waardoor de studenten een voordeel hebben ten opzichte van de reguliere toetssituatie.

## 4 Algemene conclusie en discussie

De centrale vraag binnen dit onderzoek is: Welk effect heeft de gelimiteerde toetstijd op de probleem oplossende vaardigheden van eerstejaars HBO natuurkunde studenten? Dit is onderzocht op basis van een surveyonderzoek en een quasi-experiment met een toetsafname. In het surveyonderzoek geven de studenten aan dat zij vinden dat de toetstijd van invloed is op de behaalde resultaten. Dit is niet terug te zien in de toetsresultaten in het studievoortgangssysteem. In het experiment bij elektromagnetisme is de toetstijd verdrievoudigd en zijn de toetsresultaten significant hoger. Dit lijkt aan te tonen dat de tijdslimiet van invloed is op de prestaties bij het oplossen van fysische vraagstukken voor het vak elektromagnetisme. Studenten scoren significant beter op de natuurkundige aanpak en de specifieke toepassing van de natuurkunde in de toets waar ze onbeperkt de tijd kregen.

De volgende stap die binnen de opleiding gezet moet worden is bepalen in hoeverre het gewenst is dat de toetstijd zoveel invloed heeft op het resultaat. Het is belangrijk te bepalen of iemand de kennis en vaardigheden heeft om de opleiding te voltooien. Enerzijds heeft de toetstijd een nadelige invloed op de resultaten in de toets. Anderzijds zijn de eerstejaarsvakken het fundament van

de opleiding, waarop toekomstige vakken voortborduren. Het is dus de vraag in hoeverre de snelheid waarmee men deze eerstejaars fysieke vraagstukken kan oplossen, een onderdeel is van het gewenste expertise niveau. In dat geval is snelheid en dus toetstijd een bewust criterium in de toetsing. Voor het vak elektromagnetisme is het aan te raden om de toetstijd op te rekken. Het effect in dit onderzoek is te groot. Studenten gebruiken 60% meer tijd en scoren 20% hoger. De toetstijd volledig loslaten is niet realistisch, want een zekere snelheid van werken is noodzakelijk om goed te functioneren als natuurkundig onderzoeker. In een beperkte tijd moet men dan een goede inschatting kunnen maken van de belangrijkste parameters binnen een complex probleem. Elektromagnetisme is een abstract vak. De resultaten van dit onderzoek zijn niet zonder meer vertaalbaar naar andere vakken. Echter in het survey onderzoek geeft 44% van de eerstejaars studenten aan regelmatig tijdsdruk te ervaren tijdens de toetsen. Tevens geeft de groep studenten met een hoge mate van diepe studiebenadering en een hoge mate van toetsvaardigheden aan dat de tijdslimiet een negatief effect heeft op de behaalde toetsresultaten. Van deze groep zou je verwachten dat juist zij snel de taakrelevante aspecten in een opdracht herkennen en daarna kunnen handelen. Het is dus tegenstrijdig dat deze groep tijdsdruk ervaart tijdens een toets. Het advies aan de opleiding is om voor elektromagnetisme de toetstijd te verruimen van 50 naar 60 minuten en voor andere vakken binnen het eerste jaar van de opleiding in kaart te brengen of het van belang is dat de snelheid van probleem oplossing een toetscriterium is.

Indien de snelheid van oplossen een beoordelingscriterium is, moet de student tijdens het leren voor zichzelf vast kunnen stellen beheers ik de inhoud van het vakgebied voldoende (kwaliteit) en ben ik snel genoeg (voldoende expertise). Digitale ondersteuning bij de lesmaterialen, zoals bijvoorbeeld Mastering Physics (Education, 2018), bevat daarvoor mogelijkheden. Ook digitale toetsing geeft mogelijk om een tijdslimiet mee te nemen in de beoordelingscriteria (langer dan x-tijd dan iets van cijfer af). Een student kan dan kiezen tussen langzamer en zorgvuldiger werken, maar weet dan ook dat dat effect heeft op cijfer. Het is dan een keuze geworden in plaats van een druk. Voor digitale toetsing binnen het bèta-domein zijn momenteel nog maar beperkte mogelijkheden.

Nog een stap verder in het gebruik van digitale ondersteuning is het inzetten van Eye-tracking om het leerproces in de probleemoplossing te sturen. Fysisch probleem oplossen is een complexe samengestelde hogere orde vaardigheid. Als bekend is op welke items in een vraagstuk de aandacht van een student gevestigd is, dan kan deze student feedback op maat krijgen. Feedback op maat is gunstig voor het individuele leerproces. Om dit te realiseren is meer ervaring met eye-tracking onderzoek noodzakelijk in domein van fysisch probleem oplossen.

In dit onderzoek blijkt de schaal toetsvaardigheden niet consistent. Binnen deze onderzoeksgroep was de factor toetsresultaatgericht werken significant aanwezig. De vragen zaten in de oorspronkelijke vragenlijst, maar in de factoranalyse bleken ze nu omgekeerd gecodeerd te correleren. Mogelijke oorzaken binnen de opleiding ligt een sterke focus op studenten om voldoende te oefenen met de

oefenopgaven en tevens kent de opleiding een hoog percentage mannen. Wellicht speelt resultaatgericht werken daardoor een belangrijkere rol. Verder onderzoek is hier aan te bevelen. De huidige onderzoeksgroep had een te beperkte omvang voor eenduidige conclusies.

Het experiment van de toetsanalyse laat zien dat het wegvallen van de toetstijd een grote impact kan hebben. Het is aan te bevelen om dit onderzoek op meerdere vakgebieden te laten plaatsvinden. Dan kan vastgesteld worden of de resultaten algemeen gelden of dat ze domein specifiek zijn. Bij het verruimen van de toetstijd gaf de strategie van twee penkleuren veel input voor de analyses. Het experiment kan ook goed andersom uitgevoerd worden. Door te gaan toetsen met een scherpere tijdslimiet. Dit kan wel in het kader van low-stake toetsen, maar is niet wenselijk in een toetssituatie waar voor de student veel van afhangt.

De algemene aanbeveling die op basis van dit onderzoek gesteld kan worden is dat het binnen toetsing belangrijk is te zorgen voor een goede balans tussen enerzijds voldoende tijd om problemen op te lossen en anderzijds een werktempo dat passend is bij het gewenste expertiseniveau. Verder onderzoek met grotere aantallen studenten kan inzicht geven in hoeverre verschillen tussen studenten op het gebied van toetsangst en mate van ervaren tijdsdruk mogelijk tot ongelijkheid in de toetsing leiden.

## Referenties

- Biggs, J., Kember, D., & Leung, D. Y. P. (2001). The revised two-factor Study Process Questionnaire: R-SPQ-2F. *British Journal of Educational Psychology*, 71(1), 133–149.  
<https://doi.org/10.1348/000709901158433>
- Butler, A. C., Karpicke, J. D., & Roediger, H. L. (2007). The effect of type and timing of feedback on learning from multiple-choice tests. *Journal of Experimental Psychology: Applied*, 13(4), 273–281. <https://doi.org/10.1037/1076-898X.13.4.273>
- Byun, T., & Lee, G. (2014). Why students still can't solve physics problems after solving over 2000 problems. *Physics Education Research American Journal of Physics*, 82(9), 906.  
<https://doi.org/10.1119/1.4881606>
- Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and Representation of Physics Problems by Experts and Novices. *Cognitive Science*, 5(2), 121–152.  
[https://doi.org/10.1207/s15516709cog0502\\_2](https://doi.org/10.1207/s15516709cog0502_2)
- De Paola, M., & Gioia, F. (2016). Who performs better under time pressure? Results from a field experiment. *Journal of Economic Psychology*, 53, 37–53.  
<https://doi.org/10.1016/j.joep.2015.12.002>
- Devine, A., Fawcett, K., Szűcs, D., & Dowker, A. (2012). Gender differences in mathematics anxiety and the relation to mathematics performance while controlling for test anxiety. *Behavioral and*



- Brain Functions*, 8(1), 33. <https://doi.org/10.1186/1744-9081-8-33>
- Docktor, J. L. (2009). *Development and validation of a physics problem solving assessment rubric*. University of Minnesota. Retrieved from [http://groups.physics.umn.edu/physed/People/Docktor\\_dissertation\\_submitted\\_final.pdf](http://groups.physics.umn.edu/physed/People/Docktor_dissertation_submitted_final.pdf)
- Docktor, J. L., Strand, N. E., Mestre, J. P., & Ross, B. H. (2015). Conceptual problem solving in high school physics. *Physical Review Special Topics - Physics Education Research*, 11(2), 020106. <https://doi.org/10.1103/PhysRevSTPER.11.020106>
- Dodeen, H. M. (2008). Assessing test-taking strategies of university students: developing a scale and estimating its psychometric indices. *Assessment & Evaluation in Higher Education*, 33(4), 409–419. <https://doi.org/10.1080/02602930701562874>
- Dodeen, H. M., Abdelfattah, F., & Alshumrani, S. (2014). Test-taking skills of secondary students: the relationship with motivation, attitudes, anxiety and attitudes towards tests. *South African Journal of Education*, 34(2), 1–18. <https://doi.org/10.15700/201412071153>
- Driscoll, M. P. (2005). Meaningful Learning and Schema Theory. In *Psychology of Learning for Instruction* (3rd ed., pp. 111–152). Florida: Pearson Education.
- Duggan, G. B., & Payne, S. J. (2009). Text skimming: The process and effectiveness of foraging through text under time pressure. *Journal of Experimental Psychology: Applied*, 15(3), 228–242. <https://doi.org/10.1037/a0016995>
- Education, P. (2018). Mastering | Physics - Reach every student. Retrieved October 15, 2018, from <https://www.pearsonmylabandmastering.com/northamerica/masteringphysics/>
- Eklöf, H., & Nyroos, M. (2013). Pupil perceptions of national tests in science: perceived importance, invested effort, and test anxiety. *European Journal of Psychology of Education*. Springer. <https://doi.org/10.2307/23421906>
- Fakcharoenphol, W., Morphey, J. W., & Mestre, J. P. (2015). Judgments of physics problem difficulty among experts and novices. *Physical Review Special Topics - Physics Education Research*, 11(2), 020128. <https://doi.org/10.1103/PhysRevSTPER.11.020128>
- Feil, A., & Mestre, J. P. (2010). Change Blindness as a Means of Studying Expertise in Physics. *Journal of the Learning Sciences*, 19(4), 480–505. <https://doi.org/10.1080/10508406.2010.505139>
- Field, A. (2013). *Discovering Statistics using IBM SPSS Statistics* (3rd ed.). London: SAGE Publications/Sage Ltd, UK.
- Gegenfurtner, A., Lehtinen, E., Jarodzka, H., & Säljö, R. (2017). Effects of eye movement modeling examples on adaptive expertise in medical image diagnosis. *Computers & Education*, 113, 212–225. <https://doi.org/10.1016/j.compedu.2017.06.001>
- Giancoli, D. C. (2014). *Physics for Scientists & Engineers with Modern Physics* (4th ed.). Harlow,

UK: Pearson Education.

- Hu, Y., Wu, B., & Gu, X. (2017). An Eye Tracking Study of High- and Low-Performing Students in Solving Interactive and Analytical Problems. *Journal of Educational Technology & Society*, 20(4), 300–311. <https://doi.org/10.2307/26229225>
- Huang, P.-S., & Chen, H.-C. (2016). Gender Differences in Eye Movements in Solving Text-and-Diagram Science Problems. *International Journal of Science and Mathematics Education*, 14(S2), 327–346. <https://doi.org/10.1007/s10763-015-9644-3>
- Hull, M. M., Kuo, E., Gupta, A., & Elby, A. (2013). Problem-solving rubrics revisited: Attending to the blending of informal conceptual and formal mathematical reasoning. *Physical Review Special Topics - Physics Education Research*, 9(1), 010105. <https://doi.org/10.1103/PhysRevSTPER.9.010105>
- Hunt, T. E., & Sandhu, K. K. (2017). Endogenous and exogenous time pressure: Interactions with mathematics anxiety in explaining arithmetic performance. *International Journal of Educational Research*, 82, 91–98. <https://doi.org/10.1016/j.ijer.2017.01.005>
- Jaeger, M., & Adair, D. (2017). Time pressure in scenario-based online construction safety quizzes and its effect on students' performance. *European Journal of Engineering Education*, 42(3), 241–251. <https://doi.org/10.1080/03043797.2016.1153042>
- Kastens, K. A., Shipley, T. F., Boone, A. P., & Straccia, F. (2016). What Geoscience Experts And Novices Look At, And What They See, When Viewing Data Visualizations. *Journal of Astronomy & Earth Sciences Education (JAESE)*, 3(1), 27–58. <https://doi.org/10.19030/jaese.v3i1.9689>
- Kuo, E., Hallinen, N. R., & Conlin, L. D. (2017). When procedures discourage insight: epistemological consequences of prompting novice physics students to construct force diagrams. *International Journal of Science Education*, 39(7), 814–839. <https://doi.org/10.1080/09500693.2017.1308037>
- Landelijk Opleidingsprofiel Hbo-bachelor Technische Natuurkunde*. (2016). Retrieved from <https://appliedscience.nl/app/uploads/2018/01/Landelijk-Opleidingsprofiel-TN-december-2016.pdf>
- Lee, H. S., & Ahn, D. (2018). Testing prepares students to learn better: The forward effect of testing in category learning. *Journal of Educational Psychology*, 110(2), 203–217. <https://doi.org/10.1037/edu0000211>
- Lee, Y.-H., & Haberman, S. J. (2016). Investigating Test-Taking Behaviors Using Timing and Process Data. *International Journal of Testing*, 16(3), 240–267. <https://doi.org/10.1080/15305058.2015.1085385>
- Lu, Y., & Sireci, S. G. (2007). Validity Issues in Test Speededness. *Educational Measurement: Issues*

- and Practice*, 26(4), 29–37. <https://doi.org/10.1111/j.1745-3992.2007.00106.x>
- Madsen, A. M., Larson, A. M., Loschky, L. C., & Rebello, N. S. (2012). Differences in visual attention between those who correctly and incorrectly answer physics problems. *Physical Review Special Topics - Physics Education Research*, 8(1), 010122. <https://doi.org/10.1103/PhysRevSTPER.8.010122>
- Mehdinezhad, V., & Bamari, Z. (2015). The Relationship between Test Anxiety, Epistemological Beliefs and Problem Solving among Students. *Journal of New Approaches in Educational Research*, 4(1), 2–8. <https://doi.org/10.7821/naer.2015.1.97>
- Morphew, J. W., Mestre, J. P., Ross, B. H., & Strand, N. E. (2015). Do experts and novices direct attention differently in examining physics diagrams? A study of change detection using the flicker technique. *Physical Review Special Topics - Physics Education Research*, 11(2), 020104. <https://doi.org/10.1103/PhysRevSTPER.11.020104>
- Pieters, R., & Warlop, L. (1999). Visual attention during brand choice: The impact of time pressure and task motivation. *International Journal of Research in Marketing*, 16(1), 1–16. [https://doi.org/10.1016/S0167-8116\(98\)00022-6](https://doi.org/10.1016/S0167-8116(98)00022-6)
- Roach, V. A., Fraser, G. M., Kryklywy, J. H., Mitchell, D. G. V., & Wilson, T. D. (2017). Time limits in testing: An analysis of eye movements and visual attention in spatial problem solving. *Anatomical Sciences Education*, 10(6), 528–537. <https://doi.org/10.1002/ase.1695>
- Roediger, H. L., & Karpicke, J. D. (2006a). Test-Enhanced Learning. *Psychological Science*, 17(3), 249–255. <https://doi.org/10.1111/j.1467-9280.2006.01693.x>
- Roediger, H. L., & Karpicke, J. D. (2006b). The Power of Testing Memory: Basic Research and Implications for Educational Practice. *Perspectives on Psychological Science*, 1(3), 181–210. <https://doi.org/10.1111/j.1745-6916.2006.00012.x>
- Rowland, C. A. (2014). The effect of testing versus restudy on retention: a meta-analytic review of the testing effect. *Psychological Bulletin*, 140(6), 1432–1463. <https://doi.org/10.1037/a0037559>
- Sireci, S. G., Scarpati, S. E., & Li, S. (2005). Test Accommodations for Students With Disabilities: An Analysis of the Interaction Hypothesis. *Review of Educational Research*, 75(4), 457–490. <https://doi.org/10.3102/00346543075004457>
- Snyder, J. L. (2000). An investigation of the knowledge structures of experts, intermediates and novices in physics. *International Journal of Science Education*, 22(9), 979–992. <https://doi.org/10.1080/095006900416866>
- Sung, Y.-T., Chao, T.-Y., & Tseng, F.-L. (2016). Reexamining the relationship between test anxiety and learning achievement: An individual-differences perspective. *Contemporary Educational Psychology*, 46, 241–252. <https://doi.org/10.1016/j.cedpsych.2016.07.001>
- Susac, A., Bubic, A., Kaponja, J., Planinic, M., & Palmovic, M. (2014). Eye Movements Reveal

- Students' Strategies in Simple Equation Solving. *International Journal of Science and Mathematics Education*, 12(3), 555–577. <https://doi.org/10.1007/s10763-014-9514-4>
- Tai, R. H., Loehr, J. F., & Brigham, F. J. (2006). An exploration of the use of eye-gaze tracking to study problem-solving on standardized science assessments. *International Journal of Research & Method in Education*, 29(2), 185–208. <https://doi.org/10.1080/17437270600891614>
- Tang, H., & Pienta, N. (2012). Eye-Tracking Study of Complexity in Gas Law Problems. *Journal of Chemical Education*, 89(8), 988–994. <https://doi.org/10.1021/ed200644k>
- Van der Linden, W. J. (2011). Setting time limits on tests. *Applied Psychological Measurement*, 35(3), 183–199. <https://doi.org/10.1177/0146621610391648>
- Van Meeuwen, L. W., Jarodzka, H., Brand-Gruwel, S., Kirschner, P. A., de Bock, J. J. P. R., & van Merriënboer, J. J. G. (2014). Identification of effective visual problem solving strategies in a complex visual domain. *Learning and Instruction*, 32, 10–21. <https://doi.org/10.1016/j.learninstruc.2014.01.004>
- Yeo, D. J., & Fazio, L. K. (2018). The Optimal Learning Strategy Depends on Learning Goals and Processes: Retrieval Practice Versus Worked Examples. *Journal of Educational Psychology*. <https://doi.org/10.1037/edu0000268>
- Young, H. ., & Freedman, R. A. (2012). *University Physics with modern physics* (13th ed.). San Francisco, United States: Pearson Education.

## Appendix

### A. Vragenlijst Surveyonderzoek

#### Algemene gegevens

2.  jaar  
Mijn leeftijd:
3.  Man  
 Vrouw  
Geslacht:
4.  2017  
 2016  
 2015  
 2014  
 2013 of eerder  
Cohort (jaar van start studie):
5.  P-fase  
 K-fase  
 A-fase  
Studiefase:
6.  ja  
 nee  
Is er sprake van een studiebeperking (denk aan dyslexie, ASS, ADHD, etc.) ?  
Indien nee, ga verder met vraag 11.
7.   
Indien ja, wil je dan aangeven wat jouw beperking is ?
8.  ja  
 nee  
Vervolg vraag: heb je hiervoor extra voorzieningen tijdens toetsen verkregen via de examencommissie (bijv. extra tijd, werken met een A3, etc.) ?
9.  Extra toetstijd  
 Opgaven op groot (A3-) formaat  
 Prikkelarme toetsruimte  
 Gebruik van programma Kurzweil  
 Andere voorziening  
Zo ja, welke voorzieningen zijn toegekend ?  
(meerdere antwoorden mogelijk)
10.   
Alleen indien gekozen voor "andere voorziening", welke voorziening was dat dan?
11.  ja  
 nee  
In dit onderzoek zou ik graag gebruik maken van gegevens uit Progress. Dit wordt enkel in anonieme vorm verwerkt. De verzamelde informatie uit dit onderzoek wordt anoniem opgeslagen en kan niet worden teruggeleid naar specifieke personen. Geef je hiervoor toestemming?
12.   
Indien ja: mijn naam en studentnummer zijn:

### Studeren

Dit eerste deel van de vragenlijst gaat over je houding tot de studie en jouw manier van studeren in zijn algemeenheid. Er is geen **goede** manier van studeren. Het hangt maar net af van jouw leerstijl en welk onderwerp je aan het bestuderen bent. Het is belangrijk dat je zo oprecht mogelijk invult. Als je denkt dat het antwoord afhangt van het onderwerp of vak, geef dan het antwoord dat geldt voor de meeste vakken binnen de opleiding Technische Natuurkunde. Mogelijke antwoorden:

- Dit is **nooit of zelden** waar voor mij
- Dit is **soms** waar voor mij
- Dit is **regelmatig** waar voor mij
- Dit is **(bijna) altijd** waar voor mij

Kruis aan welk antwoord het best past bij iedere stelling. Denk niet te lang na, maar geef het antwoord wat het eerst in je opkomt. Alle antwoorden worden vertrouwelijk behandeld en tot een anoniem resultaat verwerkt.

	nooit of zelden	soms	regelmatig	(bijna) altijd
13. Ik werk hard voor mijn studie technische natuurkunde, want de vakken zijn zo interessant	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14. Ik ervaar dat ieder vak interessant is, als ik er eenmaal induik	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
15. Als ik een vak niet interessant vind, probeer ik het met minimale inspanning te halen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
16. Ik vind dat docenten niet moeten verwachten dat studenten tijd besteden aan studiematerialen die niet worden getoetst op de toets	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
17. Ik vind dat ik net zo lang moet studeren tot ik mijn eigen conclusies kan vormen over de kernbegrippen binnen het vakgebied.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
18. Je verder verdiepen in een vak is verwarrend en tijdverspilling. Je moet dicht bij de rode draad van het vak blijven.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
19. Ik kom naar de meeste lessen met vragen die ik graag beantwoord wil hebben	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
20. Ik besteed een deel van mijn vrije tijd aan het vinden van meer informatie over interessante onderwerpen die zijn bediscussieerd in de lessen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
21. Ik bestudeer enkel de onderwerpen, die in de les behandeld worden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
22. Ik vind het nutteloos materiaal te bestuderen dat niet in het tentamen komt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
23. Studeren geeft me soms een diep gevoel van persoonlijk voldoening.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
24. Ik vind de meeste onderwerpen interessant en besteed vaak veel extra tijd aan het verkrijgen van meer informatie.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
25. In het algemeen beperk ik mijn studie tot wat specifiek benoemd is, want ik denk dat het onnodig is iets extra's te doen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

26.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ik vind dat het bestuderen van natuurkundige onderwerpen soms net zo interessant als een goed boek of een goede film				
27.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mijn doel is met zo min mogelijk werk een voldoende te halen				
28.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ik zorg dat ik meestal ook alle extra informatie die wordt aangeboden via de vakwijzers bekijk				
29.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ik ben bang dat ik toetsen niet haal				

### Vorbereiden op toetsen

Dit tweede deel van de vragenlijst gaat over jouw voorbereiding op toetsen. Er is geen **goede** manier van studeren. Het is belangrijk dat je zo oprecht mogelijk invult. Als je denkt dat antwoord afhangt van het onderwerp, geef dan het antwoord dat geldt voor de meeste vakken.

	nooit of zelden	soms	regelmatig	(bijna) altijd
30.				
Ik test mezelf op belangrijke onderwerpen totdat ik ze volledig begrijp	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
31.				
Ik ga <b>niet</b> naar de lessen in de laatste lesweek	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
32.				
De avond en nacht voor de toets ben ik hard aan het studeren	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
33.				
Het is vaak makkelijker een tentamen te halen door zaken uit het hoofd te leren dan door het te proberen te begrijpen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
34.				
Ik ga door met het bestuderen en doorlezen van de stof tot de laatste minuut.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
35.				
Soms leer ik dingen uit mijn hoofd door ze telkens te herhalen totdat ik ze erin gestampt heb, zelfs als ik het niet begrijp.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
36.				
Voor een toets begint, spreek ik met medestudenten over de toets	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
37.				
Ik vind de beste manier om toetsen te behalen, het onthouden van antwoorden op vragen met een grote kans om op de toets te komen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
38.				
Ik lees de instructies over de toets op de vakwijzer zorgvuldig door	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
39.				
Voorafgaand aan toetsen ben ik bezorgd over hoe moeilijk de toets zal zijn.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
40.				
Ik drink veel koffie, frisdrank of energiedrank voor een toets	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
41.				
Ik neem alle noodzakelijke middelen mee naar de toetsen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
42.				
Ik ontbijt <b>niet</b> als ik een toets heb	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	nooit of zelden	soms	regelmatig	(bijna) altijd
43. Ik gebruik de volledige toetstijd voor een toets	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
44. Als andere studenten het toetslokaal verlaten, heb ik het gevoel dat ik ook moet vertrekken	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
45. Ik maak een schatting hoeveel tijd ik heb voor iedere vraag.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
46. Als ik de toets afgerond heb, kijk ik mijn antwoorden nog na	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
47. Als ik te weinig tijd heb, dan geef ik bij de resterende vragen een indicatie van hoe ik de opgave zou oplossen/aanpakken	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
48. Ik hou me zelf aan de tijd die ik voor iedere vraag heb.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
49. Ik ben kalm en zelfverzekerd als ik toetsen maak	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
50. Ik markeer opgaven die ik niet weet	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
51. Ik lees iedere opgave zorgvuldig voor ik hem probeer te beantwoorden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
52. Ik onderstreep belangrijke informatie in een opgave / Maak een overzicht van de belangrijkste informatie	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
53. Als ik een vraag niet begrijp, schrijf ik zoveel mogelijk informatie op	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
54. Als ik een antwoord niet weet dan laat ik de vraag open.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
55. Ik lees niet de hele vraag als de vraag er bekend uitziet	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
56. Ik bedenk/plan eerst het antwoord in mijn hoofd voor ik het opschrijf	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
57. Ik ben vaak zo nerveus dat ik dingen vergeet die ik eigenlijk wel weet.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
58. Ik denk meer aan het te behalen cijfer dan aan de toets zelf	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
59. Als ik het antwoord niet weet dan doe ik een gok.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
60. Ik controleer mijn werk ook als het er goed uitziet	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
61. Als iets niet helder geformuleerd is , dan vraag ik voor een verduidelijking.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
62. Ik kijk naar het aantal punten bij iedere vraag voor ik de vraag probeer te beantwoorden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
63. Na de toets tel ik de punten van de toetsonderdelen op en vergelijk het met de toetsscore.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



64.
- Na afloop van de toets check ik de oplossingsstrategie van iedere vraag.
65.
- Ik bepaal wat de oorzaak is van mijn fouten
66.
- Op basis van mijn toetsresultaten, pas ik mijn studeer- en toetsstrategieën aan
67.
- Ik ga naar iedere toetsinzage en luister aandachtig naar de aandachtspunten

Graag willen we weten hoe jullie de toetstijd van 50 minuten in de propedeuse ervaren.

- ja nee
68.
- Heb je wel eens tijddruk ervaren tijdens de toetsen ?

69. Zo ja, bij welke vakken?

Als je dit studiejaar geen toetsen in de propedeuse hebt gedaan, mag je de volgende acht vragen **overslaan**

- |  | nooit of zelden       | soms                  | regelmatig            | (bijna) altijd        |
|--|-----------------------|-----------------------|-----------------------|-----------------------|
| 70. Ik kan alle vragen beantwoorden binnen de geplande toetstijd                       | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 71. Tijdens de toets maak ik me zorgen of ik voldoende tijd heb om alles af te krijgen | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 72. Door de tijdslijmet maak ik vaak rekenfouten.                                      | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 73. Vaak moet ik opgaven sneller maken dan ik eigenlijk zou willen.                    | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 74. Ik laat vragen onbeantwoord, omdat daar onvoldoende tijd voor is.                  | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 75. Door de tijdslijmet heb ik onvoldoende tijd de vragen zorgvuldig te lezen          | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 76. Ik heb tijd om mijn antwoorden na te kijken.                                       | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 77. Als de toetsen geen tijdslijmet hadden zou ik betere resultaten halen              | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

78.

Heb je nog aanvullingen met betrekking tot de toetstijd van 50 minuten, dan kun je die hier delen:

## B. Survey

### B.1 Analyse onderzoeksschaal Toetsvaardigheid

Een nadere analyse van alle toetsitems van de factor toetsvaardigheid laat zien dat er in de hoofdschaal en de subschalen een aantal items zijn met een lage item-correlatie ( $<0.100$  of negatief) of met een negatieve gemiddelde covariantie. Deze items zijn beoordeeld op hun relevantie voor toetsvaardigheden en de relevantie in relatie tot de specifieke toetssituatie bij de opleiding Technische natuurkunde (Tabel 15). Op basis van deze analyse zijn vijf items beoordeeld als een lage relevantie voor de schaal toetsvaardigheden. In de itemanalyse zijn vier items geïdentificeerd als tegenstrijdige indicatoren. Deze items zijn niet effectief voor lange termijn leereffecten, maar de strategie heeft mogelijk wel een positief effect op het korte termijn resultaat (Roediger & Karpicke, 2006a).

Tabel 15

*Nadere beschouwing van de Items met een negatieve of lage item-correlatie ( $<0,100$ ) in de hoofdschaal toetsvaardigheid (TV. Deze items) zijn beoordeeld op relevantie voor de schaal toetsvaardigheden of relevantie op basis van de specifieke toetssituatie bij de opleiding.*

Items	Relevantie	Motivatie en beslissing
TV2 : blokken in de avond voor de toets TV6 : door werken tot de laatste minuut TV18 : bij antwoord niet weten zoveel mogelijk opschrijven TV22 : Cijfer belangrijker dan toets	Tweeledig	Deze items kunnen een positief effect hebben op het toetsresultaat, maar zijn niet positief voor het langetermijn leereffect (Roediger & Karpicke, 2006a) → onderbrengen in aparte factor : controleren met factoranalyse
TV3 : gebruiken van cafeïne houdende dranken TV5 : ontbijten	Laag	Dit zijn goede leefgewoonten, maar niet specifieke toetsvaardigheden → items verwijderen
TV10 : als anderen het toetslokaal verlaten heb ik het gevoel dat ik ook moet vertrekken TV15 : ik markeer opgaven die ik niet weet	Laag	Toetsen duren 50 minuten. De korte tijd zorgt ervoor dat weinig studenten voortijdig lokaal verlaten en toetsen weinig verschillende opgaven hebben → items verwijderen
TV25 : als iets niet helder geformuleerd is vraag ik voor verduidelijking	Laag	Opleiding werkt met externe surveillanten → item verwijderen

De schaal toetsvaardigheden is teruggebracht tot 26 items door de items met een lage relevantie voor de toetsvaardigheid te verwijderen. Het verwijderen van deze items heeft geen positief effect op de betrouwbaarheid van de hoofdschaal toetsvaardigheid en de bijbehorende subschalen (Dodeen, 2008), zie Tabel 15. Een oblique geroteerde factoranalyse is uitgevoerd op de 26 items resterende items van de schaal toetsvaardigheid om te bepalen of de schaal opgesplitst kan worden in verschillende subschalen. Op basis van de screeplot lijken drie of zes factoren aannemelijk. Gezien het lage aantal respondenten ( $n=60$ ) in dit onderzoek is besloten drie factoren te bepalen. Deze factoren hebben alle een eigenwaarde van boven de 2 en verklaren 32,9% van de variantie in de schaal toetsvaardigheden. De drie factoren (Tabel 16) kunnen worden beschreven als “zorgvuldig werken”, “timemanagement” en “toetsresultaatgericht werken”. Ze hebben alle drie een relatief lage betrouwbaarheid, maar beter dan de hoofdschaal en de oorspronkelijke subschalen. De factor toetsresultaatgericht werken bevat de items die ongewenst zijn vanuit leerrendement, maar een positief effect kunnen hebben op het korte termijn resultaat (Tabel 15).

Tabel 16

Resultaten van de factoranalyse uitgevoerd op schaal toetsvaardigheid (N=60). Factorlading van 0,3 of lager zijn onderdrukt. F1=Zorgvuldig en nauwgezet werken, F2=timemanagement keuzes tijdens de toets; F3=toetsresultaatgericht. TV24 meegenomen in de factor zorgvuldig en nauwgezet werken (op basis van factorthema); TV23 en TV27 niet meegenomen, omdat ze niet passen bij de factorthema's.

Item	F1	F2	F3
TV12 Als ik de toets afgerond heb, kijk ik mijn antwoorden nog na	0.585	0.348	
TV31 Ik ga naar iedere toetsinzage en luister aandachtig naar de aandachtspunten	0.573		
TV16 Ik lees iedere opgave zorgvuldig voor ik hem probeer te beantwoorden	0.525		
TV20 Ik lees niet de hele vraag als de vraag er bekend uitziet	-0.497		
TV17 Ik onderstreep belangrijke informatie in een opgave / Maak een overzicht van de belangrijkste informatie	0.485		
TV8 Ik lees de instructies over de toets op de vakwijzer zorgvuldig door	0.404	0.353	
TV30 Op basis van mijn toetsresultaten, pas ik mijn studeer- en toetsstrategie aan	0.332		
TV31 Ik bepaal wat de oorzaak is van mijn fouten	0.309		
TV24 Ik controleer mijn werk ook als het er goed uitziet	0.443	0.449	
TV11 Ik maak een schatting hoeveel tijd ik heb voor iedere vraag		0.641	
TV14 Ik hou mezelf aan de tijd die ik voor iedere vraag heb		0.496	
TV26 Ik kijk naar het aantal punten voor iedere vraag voor ik de vraag probeer te beantwoorden		0.485	
TV13 Als ik weinig tijd heb, dan geef ik bij de resterende vragen een indicatie van hoe ik de opgave zou aanpakken		0.444	
TV23 Als ik een antwoord niet weet doe ik een gok		0.385	
TV6 Ik ga door met het bestuderen en doorlezen van de stof tot de laatste minuut	0.432		0.686
TV2 De avond en nacht voor de toets ben ik hard aan het studeren	0.372		0.604
TV7 Voor een toets begint, spreek ik met medestudenten over de toets			0.572
TV22 Ik denk meer aan het te behalen cijfer dan aan de toets zelf			0.440
TV27 Na de toets tel ik de punten van de toetsonderdelen op en vergelijk het met de toetsscore			0.310
Eigenwaarde	3.809	2.549	2.196
% van de variantie	14.65	9.80	8.45
$\alpha$	0.70	0.61	0.63

Op basis van de factoranalyse is besloten om in dit onderzoek de schaal toetsvaardigheden op te splitsen in een schaal “Toetsvaardigheid (nieuw)” en een schaal “Toetsresultaatgericht werken”.

Toetsvaardigheid (nieuw) omvat de items die te maken hebben met zorgvuldig

Tabel 17

*Betrouwbaarheid van de verschillende varianten in de schaal toetsvaardigheid. Toetsvaardigheid (-5 items) en de subschalen volgen de indeling van (Dodeen, 2008) zonder de niet relevante items. Toetsvaardigheid nieuw en Toetsresultaatgericht werken zijn afkomstig uit de factoranalyse.*

	items	$\alpha$	Opmerkingen
<b>Toetsvaardigheid (31-5 items)</b>	26	0.51	Op basis van itemanalyse 5 items verwijderd
Voor de toets	6	-0.085	Negatieve gemiddelde covariantie
Time management	5	0.54	
Tijdens de toets	10	0.25	
Na de toets	5	0.53	
<b>Toetsvaardigheid (nieuw)</b>	13	0.66	Op basis van factoranalyse F1 + F2
Zorgvuldig werken	9	0.70	
Timemanagement	4	0.61	
<b>Toetsresultaatgericht werken</b>	4	0.63	Op basis van factoranalyse F3

## B.2 Vergelijking van groepen binnen de onderzoeksschalen

De variabelen studiebenadering (diep en oppervlakkig), toetsvaardigheid nieuw, Toetsresultaatgericht werken, toetsangst en ervaren tijdsdruk zijn omgezet naar ordinale variabelen (Tabel 18). Met behulp van een ANOVA vergelijking met Post Hoc tests is er voor iedere schaal bepaald of er verschillen zijn tussen de groepen met een verschillende mate van die onderzoeksvariabele.

Tabel 18

*Overzichtstabel van de groepen binnen de onderzoeksschalen*

	Lage mate van		Gemid. mate van		Hoge mate van	
	score	n	score	n	score	n
<b>Diepe studiebenadering</b>	2.06 ± 0.27	17	2.61 ± 0.15	25	3.11 ± 0.28	18
<b>Opp. studiebenadering</b>	1.57 ± 0.14	15	1.96 ± 0.13	28	2.59 ± 0.32	17
<b>Toetsvaardigheid nieuw</b>	2.12 ± 0.19	15	2.58 ± 0.14	30	3.10 ± 0.20	15
<b>Toetsresultaatgericht werken</b>	1.60 ± 0.19	18	2.51 ± 0.34	30	3.48 ± 0.17	12
<b>Toetsangst</b>	1.44 ± 0.11	13	2.20 ± 0.32	32	3.43 ± 0.43	15
<b>Ervaren tijdsdruk</b>	1.50 ± 0.21	10	2.14 ± 0.20	27	3.16 ± 0.46	14

**Studiebenadering:** Hier zijn enkel significante verschillen aanwezig voor de twee subschalen onderling (Diep → oppervlakkige studiebenadering  $F(2,57)=4.434, p=.016$ ).

**Toetsvaardigheid (nieuw):** geen significante verschillen aanwezig voor de mate van toetsvaardigheid.

**Toetsresultaatgericht werken:** De mate van toetsresultaatgericht werken laat verschillen zien voor de mate van toetsangst ( $F(57,2)=4.724, p=.013$ ) en de mate van ervaren tijdsdruk ( $F(2,48)=4.221, p=.020$ ). Studenten met een hoge mate van toetsresultaatgericht werken scoren 0.82 hoger op de schaal van toetsangst ( $SE=0.27, p=.011, CI\ 95\%: 0.15 - 1.49$ ) en 0.65 op de schaal van ervaren tijdsdruk ( $SE=0.24, p=.024, CI\ 95\%: 0.07- 1.24$ ) dan studenten met een lage mate van toetsgericht werken. Voor het aantal behaalde vakken in de propedeuse is de variantie verschillend. Op basis van Welch F zijn er enkel niet-significant verschillen tussen de groepen met verschillende maten van toetsresultaatgericht werken.

**Toetsangst:** Significante verschillen voor toetsresultaatgericht werken ( $F(2,57)=5.781, p=.005$ ) en ervaren tijdsdruk ( $F(2,48)=3.668, p=.000$ ). Studenten met een hoge mate van toetsangst werken meer resultaatgericht ( $verschil=0.85, SE=0.25, p=.004$ ) en ervaren meer tijdsdruk ( $verschil=0.79, SE=0.22, p=.002$ ) dan studenten met een lage mate van toetsangst. Voor het toetsresultaat aantal behaalde vakken in de propedeuse ( $P_n$ ) is de variantie verschillend. Op basis van Welch F zijn er significant verschillen tussen de groepen met verschillende mates van angst voor het aantal behaalde vakken ( $F(2,13.4)=5.197, p=.021$ ). Studenten met een lage toetsangst hebben meer vakken behaald dan studenten met een gemiddelde toetsangst ( $verschil=2.6, SE=0.8, p=.011$ ) en hoge toetsangst ( $verschil=1.54, SE=1.3, p=.502$ ). Het effect is enkele significant tussen de groepen lage en gemiddelde toetsangst. De groep studenten met een lage toetsangst scoort voor alle indicatoren van toetsresultaat beter dan studenten met een gemiddelde of hoge toetsangst.

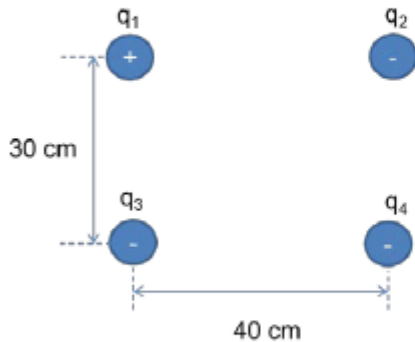
**Ervaren tijdsdruk:** Voor toetsangst (@TA) en perceptie van de student in hoeverre toetstijd van resultaat nadelig beïnvloedt (TD8) is de variantie verschillend. Op basis van Welch F zijn er significant verschillen tussen de groepen met verschillende mates van ervaren tijdsdruk voor toetsangst ( $F(2,26.7)=15.360, p=.000$ ) en de perceptie ( $F(2,26.7)=61.671, p=.000$ ). Op het gebied van toetsresultaat is ervoor het gemiddeld cijfer in de K-fase is er ook een significant verschil ( $F(2,10)=5.062, p=.030$ ). Studenten die veel tijdsdruk ervaren scoren op alle indicatoren voor toetsresultaat het laagst ( $P_{gem}; P_n; K_{gem}$ ). Het aantal deelnemers per indicator is echter klein.

## C. Toets & Toetsanalyse

### C.1 Toets

#### Opgave 1: 4 ladingen (2+4)

Gegeven zijn 4 puntladingen met lading  $q_1 = 5 \mu\text{C}$ ,  $q_2 = -8 \mu\text{C}$ ,  $q_3 = -15 \mu\text{C}$ ,  $q_4 = -16 \mu\text{C}$ . De ladingen hebben relatieve posities ten opzichte van elkaar zoals aangegeven in Figuur 1.



Figuur 1 Relatieve posities 4 ladingen

- Teken** het vrije lichaamsdiagram van lading  $q_4$ . De bijdrage van de zwaartekracht mag worden verwaarloosd.
- Bereken** de resulterende kracht (grootte en richting) op lading  $q_4$

#### Opgave 2: Wet van Gauss (2+3+2)

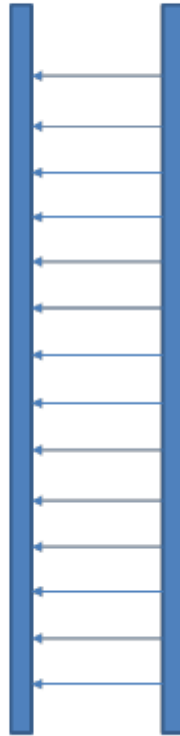
Een holle bolvormige geleider heeft een netto lading  $+Q$ , de binnenstraal is  $R_1$  en de buitenstraal is  $R_2$ . De bol is geïsoleerd van zijn omgeving.

In het middelpunt van deze bol bevindt zich een puntlading met lading  $-Q/2$ .

- Teken de situatie**. Geef hierin duidelijk aan waar er lading zit, of deze lading positief of negatief is en hoe de elektrische veldlijnen lopen.
- Bereken**  $E(r)$  voor  $0 < r < \infty$
- Schets** de grafiek van  $E(r)$  voor  $0 < r < \infty$

### Opgave 3: Potentiaal en potentiële energie (2+2)

Gegeven is een oneindig lange plaatcondensator met bijbehorende elektrische veldlijnen, zie Figuur 2. De sterkte van het elektrisch veld tussen de condensatorplaten is  $E$ .



Figuur 2 Schets van een plaatcondensator met bijbehorend veldlijnenpatroon

Stel de potentiaal op de linker condensatorplaat gelijk aan 0. De afstand tussen de condensatorplaten is gelijk aan  $d$ . Neem als assenstelsel de  $x$ -as positief naar rechts met de oorsprong op de linkercondensatorplaat.

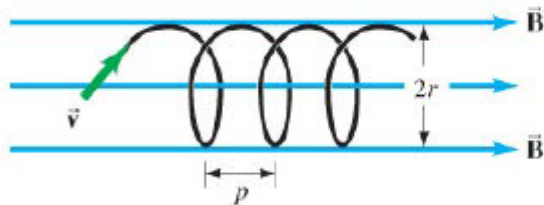
Neem als proeflading een elektron.

- Bereken en schets**  $V(x)$  voor dit elektron voor  $0 < x < d$
- Bereken en schets** de potentiële energie  $U(x)$  voor dit elektron voor  $0 < x < d$



### Opgave 4 Elektron in een homogeen magnetveld (1+2+2)

Een elektron komt met een snelheid  $3.0 \cdot 10^6$  m/s een homogeen magnetveld binnen. De hoek tussen het magnetveld ( $B = 0.50$  T) en de snelheidsvector is  $40^\circ$ . De beweging die het elektron maakt is 3 dimensionaal en noemen we een helix, zie Figuur 3.



Figuur 3 Beweging elektron dat onder een hoek in een homogeen magnetveld beweegt

- Leg uit hoe deze beweging ontstaat
- Bereken de straal ( $r$ ) van deze beweging.
- Bereken de pitch ( $p$ ) van deze beweging.

### C.2 Vragen uit de evaluatie van de verlengde toets

Deze korte vragenlijst gaat over jouw gedrag tijdens de toets. Het is belangrijk dat je zo oprecht mogelijk invult. Mogelijke antwoorden:

- Volledig oneens → dit geldt absoluut niet voor mij
- beetje oneens → dit is niet voor mij geldig
- beetje eens → dit is wel voor mij geldig
- Volledig mee eens → dit is absoluut voor mij geldig

Kruis aan welk antwoord het best past bij iedere stelling. Denk niet te lang na maar geef het antwoord wat het eerst in je opkomt. Alle antwoorden worden vertrouwelijk behandeld en tot anonieme resultaten verwerkt:

	Stelling	Volledig oneens	Beetje oneens	Beetje eens	Volledig mee
1	Ik heb mezelf goed voorbereid op deze toets.				
2	Het ontbreken van de tijdslimiet op deze toets heeft volgens mij geleid tot een beter resultaat				
3	Ik heb gebruik gemaakt van de verlengde toetstijd.				

4	Ik heb de vragen aandachtiger gelezen dan bij toetsen waar de tijdslimiet 50 minuten is.				
5	Ik heb netter/overzichtelijker gewerkt dan bij toetsen waar de tijdslimiet 50 minuten is.				
6	Ik was kalm en zelf verzekerd tijdens het maken van dit tentamen.				
7	Voor aanvang van het tentamen was ik bezorgd over dit tentamen.				
8	Ik was bang om een onvoldoende te halen.				
9	Ik was zo nerveus dat ik zaken vergeten ben die ik meestal wel weet.				
10	Ik heb mijn werk nagekeken.				

- Ik heb recht op extra toetstijd (toegekend door examencommissie) Ja / Nee
- Wat is volgens jou de ideale toetstijd voor de vakken in de propedeuse  
..... minuten

### C.3 Procedure van de analyse op basis van de Physics Problem Solving Rubric

#### Opgave 1

Met opgave 1A is reeds groot deel effectieve beschrijving toegekend. Deze opgave vraagt het vrije lichaamsdiagram te tekenen. Een student die deze vraag beantwoordt heeft zit minimaal op 3p in de categorie Effectieve beschrijving. Specifieke toepassing en wiskundige methodes liggen hier dicht bij elkaar. Logische voortgang ontstaat deels door de opbouw van deelvraag 1A en 1B.

#### *Effectieve beschrijving (EB1):*

- 5p – opgave bevat een **duidelijke tekening** met daarin de **krachten** die de 3 overige deeltjes op het 4<sup>e</sup> deeltje uitvoeren. Er is een **x-y richting** gedefinieerd. De krachten hebben een **naam**. Er is een **overzicht van de bekende / onbekende gegevens**.
- 4p – Er ontbreekt iets uit bovenstaande factoren / bekende factoren foutief overgenomen
- 3p – Er ontbreken 2 van bovenstaande factoren
- 2p – Grootste deel niet bruikbaar / bevat foutieve informatie
- 1p – Beschrijving is niet bruikbaar
- 0p – beschrijving is niet aanwezig

#### *Natuurkundige aanpak (NA1)*

- 5p – De natuurkundige aanpak is helder gedefinieerd. De formules zijn te herleiden die de student gebruikt om het probleem op te lossen (in symbolen): (Wet van Coulomb, Vectorieel optellen, Grootte en richting bepalen). In de keuzes is de richting van de krachten in de figuur van 1A correct (afstotend/aantrekkend)
- 4p – natuurkundige aanpak is correct, formules direct ingevuld of klein deel aanpak ontbreekt/bevat foutje
- 3p – Sommige delen van de aanpak ontbreken / bevat fouten (bijv. grootte en richting bepalen verder compleet)
- 2p – grootste deel ontbreekt bevat fouten
- 1p – Aanpak is niet van toepassing / incorrect

- 0p – geen aanpak

*Specifieke toepassing natuurkunde (STN1)*

- 5p – De specifieke toepassing wordt correct geïmplementeerd. De student rekent de verschillende krachten correct uit en kiest juiste richtingen bij de verschillende coulombkrachten. De vectoriële optelling is correct uitgevoerd. De grootte en de richting van de resulterende zijn correct berekend. Eenheden worden correct omgerekend. De uitwerking is lijn met het zelf gekozen X-y-stelsel.
- 4p – Op een van bovenstaande factoren ontbreekt de specifieke uitwerking is niet correct
- 3p – Sommige delen ontbreken / bevatten fouten
- 2p – grootste deel ontbreekt bevat fouten
- 1p – Aanpak is niet van toepassing / incorrect
- 0p – geen specifieke uitwerking

*Wiskundige methodes*

- 5p – wiskundige methodes zijn volledig en passend
- 4p – wiskundige methodes niet volledig of bevatten kleine foutjes
- 3p – Sommige delen van de wiskundige methodes ontbreken / bevatten fouten
- 2p – grootste deel ontbreekt bevat fouten
- 1p – Aanpak is niet van toepassing / incorrect
- 0p – geen specifieke uitwerking

*Logische Voortgang:*

- 5p – Werk van de student is gestructureerd en het is duidelijk dat student doel voor ogen heeft. In een aantal stappen komt student bij de Resulterende kracht op deeltje 4.
- 4p – Werk is (redelijk) goed gestructureerd maar student maakt enkele slordigheidsfouten
- 3p – Werk is niet gestructureerd / bevat meerdere slordigheidsfouten
- 2p – Werk is niet samenhangend (bijv. uitwerkingen in 1B zijn in strijd met tekening 1A / doelvraag grootte richting niet beantwoord.)
- 1p – Volledige oplossing is niet gefocust / inconsistent
- 0p – Geen logische voortgang

**Opgave 2**

In opg 2A zit de basis van de effectieve beschrijving. Als in deze tekening verkeerde keuzes worden gemaakt is dit beoordeeld onder natuurkundige aanpak. De uitwerkingen bij 2C moeten in lijn zijn met 2A en 2B (logische voortgang). De opgestelde formules voor de drie relevante deelgebieden zijn beoordeeld onder de specifieke toepassing. Wiskundige methodes spelen geen rol er hoeft niets uitgerekend te worden. Dus in principe niet beoordeeld. Wel score ingevuld als wiskundige bewerkingen zorgden voor fouten

*Effectieve beschrijving (EB2):*

- 5p – opgave bevat een **duidelijke tekening** van de doorsnede van de holle bol met daarin de **ladingsverdeling op de bol en – Lading in oorsprong**. De tekening bevat veldlijnen. R1 en R2 herkenbaar aangegeven.
- 4p – Er ontbreekt iets uit bovenstaande factoren / bekende factoren foutief overgenomen
- 3p – Er ontbreken 2 van bovenstaande factoren

- 2p – Grootste deel niet bruikbaar / bevat foutieve informatie
- 1p – Beschrijving is niet bruikbaar
- 0p – beschrijving is niet aanwezig

#### *Natuurkundige aanpak (NA2)*

- 5p – De natuurkundige aanpak is helder gedefinieerd. Het elektrisch veld is gedefinieerd in (1) de holle ruimte ( $r < R_1$ ), (2) bolmateriaal ( $R_1 < r < R_2$ ) en (3) er buiten ( $r > R_2$ ). In de natuurkundige aanpak houdt de student rekening met de volgende basis principes
  - De getekende veldlijnen (4) in 2A zijn in overeenstemming met de theorie van + naar – en loodrecht op het oppervlak binnen en buiten de bol
  - De lading Q op de bol is correct aangebracht. Rekening houdend met het feit dat de bol een geleider is (lading op oppervlakte) en dat de puntlading in de oorsprong gelijk is aan  $-1/2Q$
  - Elektrisch veld op basis van bolsymmetrie  $E = \frac{Q_{omsl}}{4\pi\epsilon_0 r^2}$
- 4p – natuurkundige aanpak is grotendeels correct. Klein deel aanpak ontbreekt/bevat foutje
- 3p – Sommige delen van de aanpak ontbreken / bevat fouten (bijv. grootte en richting bepalen verder compleet)
- 2p – grootste deel ontbreekt bevat fouten
- 1p – Aanpak is niet van toepassing / incorrect
- 0p – geen aanpak

#### *Specifieke toepassing natuurkunde (STN2)*

- 5p – De specifieke toepassing wordt correct geïmplementeerd. De student rekent de verschillende Elektrische velden correct uit en kiest de juiste omsloten ladingen. De grensvlakken R1 en R2 laten een sprong zien in elektrisch veld. De uitwerking is in lijn met de getekende situatie in 2A.
- 4p – Op een van bovenstaande factoren ontbreekt de specifieke uitwerking is niet correct
- 3p – Sommige delen ontbreken / bevatten fouten
- 2p – grootste deel ontbreekt bevat fouten
- 1p – Aanpak is niet van toepassing / incorrect
- 0p – geen specifieke uitwerking

#### *Wiskundige methodes (WM2)*

- 5p wiskundige methodes correct gebruikt
- 3p indien formule boloppervlak incorrect, verkeerde handelingen in integraal, breukomschrijven
- 1p meer fouten tegen wiskundige handelingen
- 0p geen wiskundige methodes gebruikt

#### *Logische Voortgang (LV2):*

- 5p – Werk van de student is gestructureerd en het is duidelijk dat student doel voor ogen heeft. In een aantal stappen komt student bij de schets van het verloop van E(r).
- 4p – Werk is (redelijk) goed gestructureerd maar student maakt enkele slordigheidsfouten
- 3p – Werk is niet gestructureerd / bevat meerdere slordigheidsfouten, doelvraag 2A niet beantwoord Veldlijnen of Lading ontbreken intekening)
- 2p – Werk is niet samenhangend (bijv. uitwerkingen in 2C zijn in strijd met tekening 2A /
- 1p – Volledige oplossing is niet gefocust / inconsistent

- 0p – Geen logische voortgang

### Opgave 3

Kwalitatieve vraag over potentiaalverschil en potentiële energie. Er wordt enkel gevraagd naar het inzicht in de situaties. Wiskundige methodes spelen geen rol er hoeft niets uitgerekend te worden.

Logische voortgang ligt grotendeels vast in de deelvraag a en b.

#### *Effectieve beschrijving (EB3):*

- 5p – opgave bevat een **duidelijke tekening** of overzicht van de gegevens. **potentiaal  $V(0)=0$** , **V, plaatafstand d, proeflading  $q=-e$** . Het plaatje bij de opgave is voor vele reeds een effectieve beschrijving (dus veel mensen geen EB toegevoegd). Hierin is echter nog niet aangegeven dat links de potentiaal 0V is. Dat x-as van links naar rechts gaat en dat daardoor het E-veld in de richting x gelijk aan  $-E$  is.
- 4p – Er ontbreekt iets uit bovenstaande factoren / bekende factoren foutief overgenomen
- 3p – Er ontbreken 2 van bovenstaande factoren
- 2p – Grootste deel niet bruikbaar / bevat foutieve informatie
- 1p – Beschrijving is niet bruikbaar
- 0p – beschrijving is niet aanwezig

#### *Natuurkundige aanpak (NA2)*

- 5p – De natuurkundige aanpak is helder gedefinieerd. Relatie  $V_x - V_0 = - \int_0^x E dx$  en  $U_x - U_0 = - \int_0^x F dx = - \int_0^x qE dx \rightarrow$  lineaire relaties voor parallelle platen
- 4p – natuurkundige aanpak is grotendeels correct. Klein deel aanpak ontbreekt/bevat foutje
- 3p – Sommige delen van de aanpak ontbreken / bevat fouten (bijv. grootte en richting bepalen verder compleet)
- 2p – grootste deel ontbreekt bevat fouten
- 1p – Aanpak is niet van toepassing / incorrect
- 0p – geen aanpak

#### *Specifieke toepassing natuurkunde (STN2)*

- 5p – De specifieke toepassing wordt correct geïmplementeerd. De student stelt vast wat potentiaal op  $x=0$  en  $x=d$  vast. Student stelt potentiële energie correct vast rekening houdend met de negatieve lading van het elektron en de negatieve richting van E-veld.
- 4p – Op een van bovenstaande factoren ontbreekt de specifieke uitwerking is niet correct
- 3p – Sommige delen ontbreken / bevatten fouten
- 2p – grootste deel ontbreekt bevat fouten
- 1p – Aanpak is niet van toepassing / incorrect
- 0p – geen specifieke uitwerking

#### *Wiskundige methodes (WM2)*

- 5p wiskundige methodes correct gebruikt
- 3p indien formule boloppervlak incorrect, verkeerde handelingen in integraal, breukomschrijven
- 1p meer fouten tegen wiskundige handelingen
- 0p geen wiskundige methodes gebruikt

#### *Logische Voortgang (LV2):*

- 5p – Werk van de student is gestructureerd en het is duidelijk dat student doel voor ogen heeft. In een aantal stappen komt student bij de schets van het verloop van  $E(r)$ .
- 4p – Werk is (redelijk) goed gestructureerd maar student maakt enkele slordigheidsfouten
- 3p – Werk is niet gestructureerd / bevat meerdere slordigheidsfouten
- 2p – Werk is niet samenhangend
- 1p – Volledige oplossing is niet gefocust / inconsistent
- 0p – Geen logische voortgang

#### Opgave 4

Analytische vraag over deeltje in magneetveld. Er wordt gevraagd naar inzicht in de beweging.

Wiskundige methodes worden de beschrijvende parameters uitgerekend.

*Effectieve beschrijving (EB3):*

- 5p – opgave bevat een **duidelijke tekening** of overzicht van de gegevens:  $\mathbf{v}_\perp$  en  $\mathbf{v}_\parallel$ ,  $m_e$ ,  $q_e$ ,  $\mathbf{B}$  en  $\mathbf{o}$ .
- 4p – Er ontbreekt iets uit bovenstaande factoren / bekende factoren foutief overgenomen
- 3p – Er ontbreken 2 van bovenstaande factoren
- 2p – Grootste deel niet bruikbaar / bevat foutieve informatie
- 1p – Beschrijving is niet bruikbaar
- 0p – beschrijving is niet aanwezig

*Natuurkundige aanpak (NA2)*

- 5p – De natuurkundige aanpak is beschreven in opgave 4A
- 4p – natuurkundige aanpak is grotendeels correct. Klein deel aanpak ontbreekt/bevat foutje
- 3p – Sommige delen van de aanpak ontbreken / bevat fouten (bijv. grootte en richting bepalen verder compleet)
- 2p – grootste deel ontbreekt bevat fouten
- 1p – Aanpak is niet van toepassing / incorrect
- 0p – geen aanpak

*Specifieke toepassing natuurkunde (STN2)*

- 5p – De specifieke toepassing wordt correct geïmplementeerd. Formule voor berekenen straal volgt uit  $F = ma = \frac{mv_\perp^2}{r} = qv_\perp B$ . Omlooptijd bepalen  $T = \frac{2\pi r}{v_\perp}$  en dit levert de pitch  $p = Tv_\parallel$
- 4p – Op een van bovenstaande factoren ontbreekt de specifieke uitwerking is niet correct
- 3p – Sommige delen ontbreken / bevatten fouten
- 2p – grootste deel ontbreekt bevat fouten
- 1p – Aanpak is niet van toepassing / incorrect
- 0p – geen specifieke uitwerking

*Wiskundige methodes (WM2)*

- 5p wiskundige methodes correct gebruikt
- 3p rekenfout / in correct gebruik  $\sin$  /  $\cos$
- 1p meer fouten tegen wiskundige handelingen
- 0p geen wiskundige methodes gebruikt

*Logische Voortgang (LV2):*

- 5p – Werk van de student is gestructureerd en het is duidelijk dat student doel voor ogen heeft. In een aantal stappen komt student bij de schets van het verloop van E(r).
- 4p – Werk is (redelijk) goed gestructureerd maar student maakt enkele slordigheidsfouten
- 3p - Werk is niet gestructureerd / bevat meerdere slordigheidsfouten
- 2p – Werk is niet samenhangend
- 1p – Volledige oplossing is niet gefocust / inconsistent
- 0p – Geen logische voortgang

#### C.4 Resultaten van de toetsanalyse op basis van eigenschappen van de populatie

Tabel 19

Op basis van een erkend studiehandicap (voorziening extra tijd toegekend)

	Reguliere toets		Verlengde toets	
	Regulier	ET	Regulier	ET
	RR	RET	VR	VET
Besteedde toetstijd [min]	50	60	78 ± 21	97 ± 23
Aantal pagina's	3.4 ± 0.8	3.6 ± 1.4	4.1 ± 1.2	4.2 ± 1.0
Aantal deelnemers	42	14	37	19
Toegekende punten [22p]	8.3 ± 3.8	7.6 ± 4.6	12.5 ± 4.5	12.2 ± 4.4
Punten op basis van werk in extra tijd			3.4 ± 2.6	2.8 ± 2.5
Gemid Eff. Beschrijving	2.7 ± 0.9	3.0 ± 1.0	2.6 ± 0.9	2.6 ± 1.0
Gemid Nat. Aanpak	2.7 ± 0.9	2.5 ± 1.2	3.5 ± 1.0	3.3 ± 0.9
Gemid. Spec. Toepassing	2.2 ± 1.0	1.9 ± 1.0	3.1 ± 1.0	3.0 ± 1.0
Gemid. Wisk. Methodes	2.6 ± 1.2	2.4 ± 1.6	2.8 ± 1.1	2.5 ± 1.2
Gemid. Logische voortgang	2.4 ± 1.0	2.4 ± 1.2	3.3 ± 1.0	3.1 ± 1.1
Toetsangst			2.55 ± 0.57	2.30 ± 0.58

Significante verschillen tussen de groepen voor:

- Toetstijd  $F(3,107)=49.38; p=0.000$ :
  - Studenten in de verlengde toetssituatie gebruiken significant meer tijd dan studenten in de reguliere toetstijd
  - Studenten met extra tijd gebruiken significant meer tijd (18,7 min) dan de reguliere studenten in de verlengde toetstijd situatie.
- Aantal toegekende punten (totaal en voor iedere opgave)
  - In verlengde toetssituatie worden meer punten gescoord dan in de reguliere toetssituatie
  - Geen significante verschillen zichtbaar tussen reguliere student en student met extra tijd

- Studenten met extra tijd lijken gemiddeld een beetje lager te scoren dan reguliere studenten. Deze effecten zijn klein als voor de verschillende opgaven cohen's d wordt berekend ( $-0.6 < d < +0.2$ )
- Gemiddelde effectieve beschrijving, natuurkundige aanpak en logische voortgang
  - Alle verschillen zitten enkel tussen de reguliere toetstijd en de verlengde toetstijd

Tabel 20

*Vergelijking toetsanalyse op basis van gender*

	Reguliere toets		Verlengde toets	
	man	vrouw	man	vrouw
Besteedde toetstijd [min]	53 ± 5	53 ± 5	76 ± 19	86 ± 31
Aantal pagina's	3.6 ± 1.0	3.9 ± 0.8	4.2 ± 1.2	4.1 ± 1.2
Aantal deelnemers	14	9	29	8
Toegekende punten [22p]	9.9 ± 3.2	8.1 ± 2.4	13.0 ± 4.3	11.0 ± 3.7
Punten op basis van werk in extra tijd			2.9 ± 2.1	2.4 ± 1.7
Gemid Eff. Beschrijving	2.9 ± 0.8	3.3 ± 1.0	2.5 ± 1.0	2.7 ± 0.9
Gemid Nat. Aanpak	3.1 ± 0.7	3.0 ± 0.8	3.5 ± 1.0	3.2 ± 0.9
Gemid. Spec. Toepassing	2.4 ± 0.8	2.1 ± 0.7	3.1 ± 1.1	2.7 ± 0.8
Gemid. Wisk. Methodes	3.0 ± 1.1	3.0 ± 0.9	2.9 ± 1.0	2.5 ± 1.1
Gemid. Logische voortgang	2.7 ± 0.9	2.6 ± 0.8	3.4 ± 1.0	2.8 ± 1.1
Toetsangst				

Geen significante verschillen indien de toetsresultaten van de mannen en vrouwen met elkaar vergeleken worden. Het aantal deelnemers is hier laag aangezien enkel van de deelnemers van het survey onderzoek de gender bekend is. De dataset was reeds geanonimiseerd. Vrouwen lijken in alle gevallen lager te scoren dan de mannen daarom zijn er ook contrastvergelijkingen uitgevoerd. De resultaten van de contrasten zijn niet significant (reguliere toets:  $t(56)=1.13$ ,  $p=.264$ ; verlengde toets  $t(56)=1.32$ ,  $p=0.193$ ; overall  $t(56)=1.72$ ,  $p=0.090$ ).



Tabel 21

Toetanalyse op basis van mate van toetsangst

	Verlengde toets		
	Lage toetsangst	Gemid toetsangst	Hoge toetsangst
Besteedde toetstijd [min]	79 ± 20	90 ± 23	77 ± 24
Aantal pagina's	4.2 ± 1.1	4.2 ± 0.9	3.5 ± 1.3
Aantal deelnemers	11	15	9
Toegekende punten [22p]	14.1 ± 4.0	13.5 ± 4.6	9.8 ± 3.8
Punten op basis van werk in extra tijd	4.2 ± 2.1	2.7 ± 2.5	2.6 ± 3.0
Gemid Eff. Beschrijving	3.0 ± 1.0	2.5 ± 1.0	2.3 ± 0.6
Gemid Nat. Aanpak	3.7 ± 0.9	3.6 ± 1.1	2.9 ± 0.8
Gemid. Spec. Toepassing	3.4 ± 1.0	3.4 ± 1.1	2.4 ± 0.9
Gemid. Wisk. Methodes	2.6 ± 1.3	2.8 ± 1.2	2.4 ± 0.8
Gemid. Logische voortgang	3.5 ± 1.3	3.4 ± 1.0	2.6 ± 1.0
Toetsangst	1.8 ± 0.2	2.5 ± 0.2	3.2 ± 0.4

Verschillen tussen de groepen voor:

- Aantal toegekende punten op de gehele toets  $F(2,32)=2.95$ ,  $p=0.067$  geen significant verschil
  - De groep studenten met een lage toetsangst heeft meer punten gescoord dan de groep studenten met een hoge toetsangst (*verschil*: 4.3,  $p=0.078$ ). Het verschil is niet significant.
  - De groep studenten met een gemiddeld toetsangst scoort beter dan de groep studenten met een hoge toetsangst (*verschil* 3.7;  $p=0.114$ ). Dit verschil is echter niet significant
  - In iedere losse opgave is het effect zichtbaar dat studenten met een lage mate van toetsangst beter scoren dan studenten met een hoge toetsangst. In geen enkele opgave is dit verschil significant
  - In iedere losse opgave is het effect zichtbaar dat studenten met een gemiddelde toetsangst beter scoren dan studenten met een hoge toetsangst. In geen enkele opgave is dit verschil significant.
- Voor alle beoordelingsaspecten uit de Physics Problem Solving Rubric (Docktor, 2009) is de score lager voor de studenten met een hoge mate van toetsangst. Voor geen enkel aspect is het verschil significant.

Op basis van geplande contrasten kan geconstateerd worden dat studenten met een hoge mate van toetsangst significant minder punten scoren dan studenten met lage toetsangst ( $t(32)=2.248, p=0.32, r=0.37$ ) en studenten met een gemiddelde toetsangst ( $t(32)=2.063, p=0.047, r=0.34$ ). Het is een medium effect. Voor de gemiddelde specifieke toepassing van de Natuurkunde is een medium effect aanwezig als de vergelijking wordt gemaakt tussen studenten met een hoge toetsangst en studenten met een lage toetsangst ( $t(32)=2.207, p=0.035, r=0.36$ ) en studenten met hoge toetsangst en een gemiddelde toetsangst ( $t(32)=2.319, p=0.027, r=0.38$ ).

## D Aannames in het quasi-experiment

### D.1 Samenstelling toetsrooster samenstelling

Tabel 22

Toetsrooster voor periode 4 in de afgelopen 3 jaar

	2016			2017			2018		
	20-6	ma		14-6	wo		11-jun	ma	
	22-6	wo		16-6	vr		13-6	wo	
	24-6	vr		20-6	di		15-6	vr	
				21-6	wo		19-6	di	
	gemid	stdev	n	gemid	stdev	n	gemid	stdev	n
Vak 1	4,30	1,70	67	5,32	2,10	62	5,52	2,29	63
Vak 2	6,60	2,10	62	6,37	1,88	61	6,73	1,77	60
Vak 3	5,70	2,30	64	6,15	1,84	52	5,76	2,28	58
Vak 4	5,40	1,90	62	6,24	1,64	57	6,06	1,89	60
Vak 5	5,60	2,20	63	3,78	1,96	58	6,26	1,50	58
Electromagnetisme	5,90	1,80	59	4,90	1,49	65	5,94	1,81	68

### D.2 Vergelijking resultaten 2018 en 2017 met het lange termijn gemiddelde

Tabel 23

Vergelijking van de resultaten van alle vakken in periode 4 ten op zichte van het vijfjarig-gemiddelde. Uitgevoerd voor de toetsen in juni 2018 en juni 2017

	verschil 2018 tov langdurig gemiddelde					verschil 2017 tov langdurig gemiddelde				
	verschil	stdev	t	df	p	verschil	stdev	t	df	p
Vak 1	0,37	0,29	1,28	391	0,176	0,17	0,89	0,19	390	0,391
Vak 2	0,54	0,23	2,33	383	0,027	0,18	0,73	0,24	384	0,387
Vak 3	0,44	0,30	1,46	374	0,138	0,83	0,77	1,08	368	0,222
Vak 4	0,46	0,25	1,85	389	0,072	0,64	0,67	0,95	386	0,253
Vak 5	0,96	0,20	4,82	387	0,000	-1,52	0,78	-1,94	387	0,061
Elektromagnetisme	1,05	0,22	4,75	413	0,000	0,01	0,61	0,02	410	0,399