

GESPIEKT?!

SPIEKBRIEFJES ALS HULPMIDDEL BIJ INFORMATICATOETSEN OP HET VO

Rick de Vries (s1843141)

Faculteit Behavioural, Management and Social
Sciences
Onderzoek van Onderwijs (10EC)

Begeleiders:
dr. L.E.I. Breymann
dr. T.C. Visser

Tijdspad:
juni - november 2022

Samenvatting

Het gebruik van naslagwerken bij toetsen is, zeker bij universiteiten, heel normaal. Naast de bekende "open-boektoets" bestaan echter ook andere varianten, zoals het toestaan van meegebrachte spiekbriefjes. Uit onderzoeken bij universiteiten is gebleken dat het toestaan van zelfgemaakte spiekbriefjes o.a. helpt bij het verminderen van stress bij studenten. In dit onderzoek is een analyse uitgevoerd op de spiekbriefjes die H4-, V4- en V5-leerlingen hebben gebruikt bij hun informaticatoetsen. Met deze analyse en de antwoorden op een vragenlijst zijn de effecten van de spiekbriefjes op de stresservaring, toetsvoorbereiding en vraagscores van de leerlingen in kaart gebracht.

In de analyse zijn de spiekbriefjes onderzocht op inhoud die mogelijk relevant is voor de beantwoording van bepaalde toetsvragen. Deze informatie is vervolgens gekoppeld aan de behaalde vraagscores om zo een mogelijke relatie te onderzoeken. Door dit te doen op zowel toetsvraag- als leerlingniveau is het mogelijk om zowel algemene trends als de effecten op specifieke leerlingen (met bijvoorbeeld een hoog of juist laag cijfer) te onderzoeken.

Uit de analyse van de vragenlijst is gebleken dat leerlingen minder stress ervaren als een spiekbriefje mag worden gebruikt. Ook bleek dat de leerlingen meer motivatie en structuur bij hun toetsvoorbereiding ervaren als ze een spiekbriefje tijdens de toets mogen gebruiken. Na analyse van de spiekbriefjes en vraagscores bleek dat er hooguit een zeer lichte correlatie is tussen de inhoud van een spiekbriefje en de behaalde vraagscores. Verder kwalitatief én kwantitatief onderzoek moet uitwijzen hoe het komt dat sommige leerlingen meer lijken te profiteren van hun spiekbriefje en hoe deze relatie verandert met de aard van de toets zelf.

Inhoudsopgave

1	Inleiding en probleemstelling	3
2	Definities en theoretische inbedding	4
2.1	Variatie in spiekbriefjes	4
2.2	Effecten gebruik spiekbriefjes	4
2.3	Invloed vraagniveau's	6
2.4	Evaluatie van toetsonderdelen	8
2.5	Onderzoeksvragen	9
3	Methode	10
3.1	Context	10
3.2	Procedure	10
3.3	Respondenten	11
3.4	Instrumenten	12
3.5	Analyse	13
3.6	Betrouwbaarheid, validiteit en overdraagbaarheid	18
4	Resultaten	20
4.1	Invloed op stresservaring	20
4.2	Invloed op toetsvoorbereiding	21
4.3	Correlatie spiekbriefjes en vraagcores	23
5	Conclusies en discussie	28
5.1	Conclusie	28
5.2	Beantwoording deelvragen	28
5.3	Discussie	30
	Bibliografie	38
A	Vragenlijst	41
B	Toetsinformatie	42
C	Oefentoets	46

1 Inleiding en probleemstelling

Binnen de studie Technical Computer Science (TCS) op de Universiteit Twente zijn bij een aanzienlijk deel van de toetsen meerdere hulpmiddelen toegestaan. Veel scholieren zijn al bekend met “standaard” hulpmiddelen als een rekenmachine of een woordenboek. Op universiteiten zijn, afhankelijk van de studie, soms ook open-boektoetsen toegestaan. Enkele studies, zoals TCS in Twente, pakken het echter anders aan. Zij bieden naslagwerk in andere vormen aan: door presentaties van colleges toegankelijk te maken tijdens de toets, door de toets open-boek te maken of door zelfgemaakte spiekbriefjes binnen bepaalde kaders toe te staan. Ook bij programmeertoetsen van andere instituten wordt gebruikgemaakt van spiekbriefjes, vaak met de achterliggende gedachte dat het studenten helpt in hun voorbereiding (Daly & Waldron, 2004; de Raadt, 2012).

Geïnspireerd door dit voorbeeld sta ik bij mijn eigen lessen als docent op het VO ook toe dat spiekbriefjes gebruikt worden, op grotendeels dezelfde manier als bij technische universitaire opleidingen. Er is vanuit de literatuur echter weinig bekend over de effecten van zelfgemaakte spiekbriefjes op toetsing bij middelbare scholieren. Uit een meta-analyse van Larwin, Gorman en Larwin (2013) is gebleken dat universitaire studenten hogere scores halen en minder stress ervaren bij het gebruik van spiekbriefjes, maar bij middelbare scholieren ontbreekt het bewijs hiervoor tot dusver.

Met dit onderzoek wil ik vaststellen of bij middelbare scholieren dezelfde trends zichtbaar zijn als al vastgesteld bij (bachelor)studenten. Deze trends gaan verder dan de behaalde scores zelf. Naast de eerdergenoemde stressvermindering (Erbe, 2007) blijkt bijvoorbeeld ook dat studenten meer tijd in hun toetsvoorbereiding stoppen dankzij de spiekbriefjes (Wachsman, 2002).

Er zijn ook interessante trends waargenomen wat betreft de inhoud van de spiekbriefjes. Allereerst is vastgesteld dat het opnemen van stof over een specifieke vraag de score op die vraag kan verhogen (Hamouda & Shaffer, 2016). Mogelijk komt dit doordat de leerling de stof verwerkt om deze op het spiekbriefje te zetten. De manier van verwerking is echter wel van belang. Zo blijkt dat studenten die louter codevoorbeelden of oefentoetsantwoorden overnemen op hun spiekbriefje gemiddeld slechter presteren (Hamouda & Shaffer, 2016). Door hierin te coachen, zouden toetsresultaten mogelijk verbeterd kunnen worden.

Op mijn eigen school, Stad & Esch in Meppel, heb ik spiekbriefjes bij alle klassen ingevoerd, vanaf havo-4 tot vwo-6. Dit heb ik gedaan omdat ik geloof in de voordelen betreffende gestructureerde stressreductie, toetsvoorbereiding en vraagscores. Deze invoering is echter gedaan onder de aanname dat de literatuurresultaten zich goed laten vertalen naar informatica op het VO. Dit verslag beschrijft mijn onderzoek naar de meerwaarde van spiekbriefjes op het VO. De hoofdvraag luidt: *Welke relatie is er tussen het toestaan van het gebruik van spiekbriefjes bij informaticatoetsen in het VO en stresservaring, toetsvoorbereiding en vraagscores?*

2 Definities en theoretische inbedding

Er zijn al diverse studies verricht naar het gebruik van spiekbriefjes bij universitaire toetsen. Deze zijn vaak verricht in een academische context, met studenten die een spiekbriefje mochten maken voor een universitair vak. Deze onderzoeken bieden daarom waardevolle inzichten in zowel de effecten van spiekbriefjes (bij studenten) als mogelijke opzetten voor onderzoek in dit gebied. In deze sectie wordt verder ingegaan op de bestaande literatuur op het gebied van spiekbriefjes en (waar gerelateerd) toetsopzetten.

2.1 Variatie in spiekbriefjes

Erbe (2007) definieert een spiekbriefje als *“a piece of paper on which one has answers or notes for a test, used to cheat on or prepare for a test”*. Het vereiste formaat en (de beperkingen op) de invulling van de spiekbriefjes zijn echter niet gegeven in deze definitie. Hierin zijn door verschillende onderzoeken dan ook verschillende keuzes gemaakt.

Formaat Allereerst hanteerden verschillende studies verschillende formaten voor de spiekbriefjes. De meeste studies stonden papier van grofweg dubbelzijdig A4-formaat toe (Erbe, 2007; Daly & Waldron, 2004; Hamouda & Shaffer, 2016). Er waren echter ook studies die alleen kleinere spiekbriefjes, zoals een enkelzijdig en gehalveerd A4, toestonden (Funk & Dickson, 2011). De gedachte achter het beperken van het formaat is dat studenten dan beter moeten samenvatten en dus mogelijk beter de stof verwerken (Funk & Dickson, 2011).

Invulling Een ander gebied van variatie was de wijze van het vullen van het spiekbriefje. Sommige onderzoeken stonden hierbij iedere manier van tekst of afbeeldingen aanbrenge toe. Te denken valt aan typen, schrijven, tekenen en alle andere varianten, mogelijk gekopieerd of overgeschreven van klasgenoten (Erbe, 2007; Hamouda & Shaffer, 2016). Andere onderzoeken stonden alleen handgeschreven spiekbriefjes toe (de Raadt, 2012). De motivatie hierachter is dat iets beter onthouden wordt wanneer dit handgeschreven in plaats van getypt is (Mangen, Anda, Oxborough & Brännick, 2015; Shibata & Omura, 2018). Ook ontmoedigt dit het kopiëren: het is nog wel mogelijk, maar door de eisen omtrent handschrift is dit meer werk dan bij getypte spiekbriefjes (de Raadt, 2012).

2.2 Effecten gebruik spiekbriefjes

De onderzoeken in academische setting hebben diverse inzichten gebracht over (onder anderen) veelvoorkomende inhoud van spiekbriefjes en de effecten van de spiekbriefjes op de stresservaring bij studenten tijdens de toets. Hoewel dit weliswaar niet de doelgroep is van dit onderzoek, zijn de conclusies wel interessant om te benoemen.

Stresservaring Deelnemende studenten in de onderzoeken van Larwin e.a. (2013), Hamouda en Shaffer (2016) en Trigwell (1987) rapporteerden minder stress voor aanvang of tijdens de toets wanneer een spiekbrieftje gebruikt mocht worden. De deelnemers gaven dit aan in de vragenlijsten of interviews die na afloop van de toets werden afgenomen. Ook Erbe (2007) geeft aan dat studenten minder stress ervaren als een spiekbrieftje gebruikt mag worden, en juist door deze stressvermindering beter presteren.

Toetsvoorbereiding Er is minder onderzoek over de effecten van spiekbrieftjes op het verloop van de toetsvoorbereiding. Bij de economietoets in de studie van Wachsmann (2002) besteedden de studenten die een spiekbrieftje gebruikten meer tijd aan hun toetsvoorbereiding dan studenten die geen spiekbrieftje meenamen naar de toets. In het onderzoek van Visco, Swaminathan, Zagumny en Anthony (2007) gaven studenten via interviews aan dat de spiekbrieftjes hielpen om “op te frissen”, te “organiseren” en te “herhalen”. Tenslotte stelde de Raadt (2012) vast dat studenten hoger scoorden wanneer de inhoud van het spiekbrieftje de volgorde van het vak volgde. Mogelijk duidt dit er volgens hem op dat sommige studenten profiteren van een meer gestructureerde “begin-tot-eind”-aanpak van de toetsvoorbereiding.

Inhoud en toetsscores De inhoud van spiekbrieftjes zijn door verschillende studies geanalyseerd op diverse criteria, sommigen afhankelijk van de vakinhoud. In de studie van de Raadt (2012) bleek dat studenten die een spiekbrieftje meebrachten een hoger cijfer haalden dan studenten die dit niet deden. Studenten die op hun spiekbrieftje letterlijk antwoorden van oefeningen hadden overgeschreven, scoorden gemiddeld lager dan studenten die dit niet deden. Dit was andersom wanneer studenten hun spiekbrieftje vulden met abstracte en generieke concepten: zij haalden juist hoger dan studenten die dit niet deden.

Hamouda en Shaffer (2016) onderzochten niet wat er op het spiekbrieftje stond, maar of voor een bepaalde vraag info op het spiekbrieftje aanwezig was. Hieruit bleek dat studenten met een gemiddelde toetsscore hun vragen beter maakten wanneer voor de betreffende vraag informatie op het spiekbrieftje stond. Studenten met een boven- of ondergemiddelde toetsscore haalden juist overwegend dezelfde scores, ongeacht of er voor een specifieke vraag informatie op hun spiekbrieftje stond.

Vorbereiding en gebruik spiekbrieftje De effecten op toetsscores kunnen zowel veroorzaakt zijn door het *voorbereiden* van een spiekbrieftje, als door het *gebruiken* van een spiekbrieftje. Om hier onderscheid in te maken is in het onderzoek van Dickson en Bauer (2008) dezelfde toets twee keer achter elkaar afgenomen: eerst zonder spiekbrieftje, en vervolgens met spiekbrieftje. Hieruit bleek dat leerlingen slechter scoorden wanneer ze hun spiekbrieftje niet konden gebruiken, waaruit Dickson en Bauer concludeerden dat de studenten te sterk leunden op hun spiekbrieftjes. Dit staat haaks op het werk van Wachsmann (2002), die concludeerde dat zowel de constructie als het gebruik van spiekbrieftjes de toetsscores verbeterden, maar dat van over-afhankelijkheid geen sprake was.

Niveau	Voorbeeldvraag
R	De snelheid van bubble sort wordt niet gemeten in seconden, maar in <i>stappen</i> . Wat wordt met een “stap” bedoeld? Een vergelijking, een berekening of een iteratie?
T1	Sorteer de lijst [4, 2, 6, 4, 7, 1] met bubble sort. Schrijf de lijst na iedere verwisseling opnieuw op.
T2	Wat is het meest geschikte algoritme om de lijst [2, 3, 5, 7, 8, 1] zo snel mogelijk te sorteren? Leg je antwoord uit.
I	Geef een algoritme waarmee de sokken uit een rommelige wasmand gesorteerd kunnen worden in paren. Geef je algoritme als een genummerd stappenplan.

Tabel 1: Voorbeeldvragen over sorteeralgoritmen bij ieder RTTI-niveau.

2.3 Invloed vraagniveau's

Het is belangrijk om te noemen dat de effectiviteit van spiekbriefjes mogelijk samenhangt met de opzet van de toets. Om deze reden is het goed om kort in te gaan op de verschillende facetten van het opzetten en evalueren van een toets. Een belangrijk onderdeel hiervan zijn de verschillende vraagniveau's, zoals gedefinieerd in een taxonomie.

2.3.1 Taxonomieën

Door de jaren heen zijn verschillende manieren bedacht om toetsvragen te categoriseren. Internationaal geniet de taxonomie van Bloom (Bloom, Engelhart, Furst, Hill & Krathwohl, 1956) veel bekendheid, die begint bij “onthouden” en in stappen omhoog werkt naar creëren. In Nederland worden tevens RTTI (Drost & Verra, 2012) en OBIT (Mesie, 2013) veel gebruikt.

RTTI Bij RTTI wordt onderscheid gemaakt in vier niveau's: Reproductie (R), Toepassing 1 (T1), Toepassing 2 (T2), en Inzicht (I). Het grootste verschil in deze niveau's is te vinden in de context van de vraag. Bij T1 wordt toegepast binnen een bekende of geoefende situatie. Bij T2 wordt de stof toegepast in een nieuwe situatie of combinatie. Delen van de geleerde stappenplannen kunnen hiervoor echter wel worden gebruikt. Bij inzicht moet de de methode helemaal zelf geconstrueerd worden. Een probleem moet worden geanalyseerd om tot een (creatieve) oplossing te komen. Een voorbeeld van een vraag over sorteren op ieder van de vier niveau's is te vinden in tabel 1.

OBIT Waar bij RTTI het hoogste niveau “inzicht” is en de middelste niveau's “toepassing”, eindigt de taxonomie van OBIT bij “toepassing”. OBIT staat namelijk voor “Onthouden, Begrijpen, Integreren, Toepassen”. Hier wordt met toepassen echter al om cre-

ativiteit vanuit de leerling gevraagd (Mesie, 2013). Zo zijn er meer subtiele verschillen tussen RTTI en OBIT, maar hebben ze wel grotendeels hetzelfde doel. Dankzij de taxonomieën kan inzichtelijk gemaakt worden welke kennis en vaardigheden bij leerlingen missen, waardoor gerichter kan worden gewerkt (Mesie, 2013).

Omzetting Bovenstaand voorbeeld over de verschillende opvattingen bij RTTI en OBIT over het niveau “toepassing” illustreert een vaker voorkomend probleem bij taxonomieën. Als niet overal dezelfde taxonomie wordt gebruikt, kan communiceren over vraagniveau’s snel lastig worden. Er is namelijk een omzetting nodig. Zo moet voor “integreren”-vragen bij OBIT eerst opgezocht worden wat OBIT hier precies mee bedoelt, om vervolgens tot de conclusie te komen dat het om een T2-vraag gaat in de RTTI-taxonomie. In het werk van Kwakernaak (2013) zijn de definities van RTTI, OBIT en de niveau’s van Bloom op een rijtje gezet en omzettingstabellen tussen deze niveau’s opgesteld. Binnen dit onderzoek is de keuze gemaakt om de RTTI-taxonomie te hanteren. Indien deze niveau’s moeten worden omgezet naar andere taxonomieën, is het werk van Kwakernaak een goed hulpmiddel.

2.3.2 Vraagniveau en geschiktheid spiekbrieftjes

Diverse onderzoekers hebben opmerkingen gemaakt over de invloed van de compositie van de toets op de geschiktheid van spiekbrieftjes. Zo stelt Erbe (2007) dat spiekbrieftjes geschikter zijn bij toetsen die een vorm van toepassing, analyse of inferentie vereisen. Zij geeft als reden hiervoor dat leerlingen bij zulke vragen informatie niet rechtstreeks kunnen overschrijven vanaf hun spiekbrieftje. Wachsmann (2002) stelt dat leerlingen minder afhankelijk worden van spiekbrieftjes (en deze dus effectiever zijn) als de vragen “analytisch” van aard zijn, waarbij voor een oplossing moet worden “geredeneerd, en niet alleen onthouden”. De onderzoekers stellen dus dat spiekbrieftjes minder gebruikt zouden moeten worden bij toetsen waarin veel reproductievragen voorkomen.

Weinig publicaties hebben een link tussen vraagniveau’s van een toetstaxonomie en de effectiviteit van spiekbrieftjes daadwerkelijk aangetoond. Bovenstaande uitspraken zijn door de onderzoekers namelijk gedaan in de context van andere onderzoeken, waarin (bijvoorbeeld) geen informatie over de vraagniveau’s is gegeven. De enige uitzondering hierop was Hamouda en Shaffer (2016), die een (zwak) positief effect waarnam bij vragen op lage niveau’s (zoals reproductievragen) als de bijbehorende stof op het spiekbrieftje stond. Bij alle andere vraagtypes werd geen effect waargenomen. Al met al suggereren de uitspraken van de onderzoekers en de (weinig) resultaten dus dat een spiekbrieftje een verhogend effect kan hebben op de scores bij reproductievragen, maar verder onderzoek in dit gebied is nodig voor definitieve conclusies.

Het bovenstaande in acht nemend zouden spiekbriefjes bij informatica op het VO goed geschikt moeten zijn voor (bijna) alle examendomeinen, mits tijdens het maken van de toets de keuze is gemaakt relatief weinig reproductievragen op te nemen. Belangrijk is dat leerlingen hun antwoorden niet rechtstreeks over kunnen nemen vanaf hun spiekbriefje. Bij praktische onderwerpen, zoals programmeren, is dit vrij gemakkelijk. Bij meer theoretische onderwerpen, zoals architectuur of grondslagen, vergt dit wellicht meer creativiteit of gebruik van contexten. Ook bij deze onderwerpen zijn echter prima toepassings- of inzichtsvragen te formuleren, zoals o.a. gedaan door de SLO (Bergervoet e.a., 2021).

2.4 Evaluatie van toetsonderdelen

Taxonomieën (zie sectie 2.3) zijn praktisch bij het opstellen van de toets, maar geven geen uitsluitel over de kwaliteit van een toetsvraag. Deze kan alleen achteraf worden bepaald, wanneer de behaalde scores worden gebruikt voor een toetsevaluatie. Het Cito hanteert twee statistieken voor het beoordelen van de kwaliteit van items: de moeilijkheidsgraad, en de discriminatie-index (Moelands, 2017).

Moeilijkheidsgraad De moeilijkheidsgraad, ook wel bekend als de p-waarde (gesloten vraag) of de p'-waarde (open vraag), geeft de gemiddelde score van leerlingen op een vraag als deel van de maximale score. Zo heeft een open vraag waarop gemiddeld 1.5 van de 2 punten is behaald, een p'-waarde van 0.75. In de literatuur is er discussie over de exacte p'-waarde waar naar gestreefd zou moeten worden. Wel is men het eens dat extreem lage of extreem hoge p'-waardes vermeden zouden moeten worden, omdat deze geen onderscheid maken tussen goed en minder goede leerlingen (Moelands, 2017). Omdat in dit onderzoek veel gebruik is gemaakt van kansrekening, wordt niet de term "p'-waarde" maar de volledige naam "moeilijkheidsgraad" gebruikt in de rest van dit onderzoek. Dit is gedaan om verdere verwarring te voorkomen.

Discriminatie-index Zelfs bij een gemiddelde moeilijkheidsgraad is het discriminerend vermogen van een toetsvraag niet gegarandeerd. Discriminerend vermogen is gedefinieerd als de mate waarin de toetsvraag onderscheid maakt tussen leerlingen met hoge of juist met lage toetsscores (Moelands, 2017). Dit wordt uitgedrukt in de discriminatie-index, Rit, met een waarde tussen -1 en +1. Hoe hoger de waarde, hoe meer een hoge score op de vraag correleert met een hoge score op de toets. In feite geeft de Rit dus aan hoe representatief een vraag is voor de toets als geheel. Een negatieve correlatie resulteert in een negatieve Rit.

Er bestaan geen absolute normen voor Rit-waarden. Mogelijk komt dit omdat de waarde sterk afhankelijk is van het aantal onderdelen van een toets. Over het algemeen wordt een waarde boven de 0.3 als goed gezien (Moelands, 2017; Goldebeld, 1992).

2.5 Onderzoeksvragen

De hoofdvraag is al geformuleerd aan het einde van de inleiding, en luidde: *Welke relatie is er tussen het toestaan van het gebruik van spiekbriefjes bij informaticoetsen in het VO en stresservaring, toetsvoorbereiding en vraagscores?* Zoals hierboven genoemd zijn deze effecten al (deels) onderzocht bij toetsen op universiteiten. In dit onderzoek wordt, gebruikmakend van vergelijkbare technieken, het effect van het toestaan van spiekbriefjes bij VO-leerlingen onderzocht. Concreet wordt een antwoord gezocht op de volgende deelvragen.

1. In hoeverre helpen spiekbriefjes bij het verminderen van de ervaren stress bij informaticoetsen op het VO?
2. Welke invloed hebben spiekbriefjes op de motivatie en voorbereiding bij informaticoetsen op het VO?
3. Welke correlatie is zichtbaar tussen de inhoud van een spiekbriefje en de behaalde score op toetsvragen bij informaticoetsen op het VO?

3 Methode

Het onderzoek was gericht op drie facetten waarmee de spiekbriefjes mogelijk een relatie hebben: de stresservaring, de toetsvoorbereiding en de behaalde vraagscores. Leerlingen in drie klassen mochten bij hun informaticoetsen een spiekbriefje gebruiken. Na afloop van de drie (verschillende) toetsen zijn vragenlijsten afgenomen en de gebruikte spiekbriefjes verzameld. De beantwoording van de vragenlijsten en de inhoud van de spiekbriefjes zijn vervolgens gebruikt voor de beantwoording van de deelvragen van het onderzoek.

De vragenlijsten bevatten vragen over o.a. de ervaren stress en de (structuur van de) toetsvoorbereiding. Bij de analyse van de spiekbriefjes is bijgehouden of een leerling op zijn/haar spiekbriefje informatie voor bepaalde vragen had staan. Deze gegevens zijn vervolgens gekoppeld aan de behaalde scores op de betreffende vragen en de toets als geheel, om zo de relatie tussen spiekbriefjesinhoud en vraagscores te onderzoeken.

3.1 Context

Het onderzoek heeft plaatsgevonden op Stad & Esch in Meppel. Binnen alle klassen informatica op deze school werd al voor aanvang van het onderzoek gebruikgemaakt van spiekbriefjes. Dit maakt het een zeer geschikte plaats voor het onderzoek.

Stad & Esch is een openbare middelbare school die alle niveau's vanaf het praktijkonderwijs tot en met de bovenbouw van het VWO aanbiedt. De school biedt als CultuurProfiel-School onder anderen alle kunstvakken aan. Hiernaast worden ook technische vakken aangeboden, zoals informatica voor leerlingen van havo en vwo in de profielen EM, NG en NT. Tijdens het onderzoek was er 1 klas per leerjaar, voor een totaal van 5 klassen (H4, H5, V4, V5, V6). Iedere informaticaklas bevatte tussen de 10 en 20 leerlingen.

Het onderzoek heeft plaatsgevonden tijdens de laatste reguliere toetsweek van het schooljaar. Hierin had de H4-klas een toets over domein B (grondslagen), de V4-klas over domein E (architectuur) en de V5-klas over de domeinen L en N (netwerken en beveiliging). De V5-toets was een schoolexamen. De H4- en V4-toetsen telden alleen voor de overgang. De examenklassen waren al klaar met hun toetsing konden daarom niet deelnemen aan het onderzoek.

3.2 Procedure

De spiekbriefjes werden door de leerlingen meegebracht, waarna ze verder als normaal hun toets konden maken. Het spiekbriefje leverden ze na afloop van de toets in met hun antwoorden. De afname van de vragenlijst vond plaats direct na de afronding van de toets. De leerlingen werden vooraf geïnformeerd over het onderzoek middels een informatiebrief en een (verbale) samenvatting.

Doordat de toetsen werden afgenomen in een toetsweek werd de surveillance waargenomen door docenten of onderwijsassistenten die niet inhoudelijk bekend waren met de stof. Zij kregen instructies bij de toets over het toelaten van spiekbriefjes en het afnemen van de vragenlijsten. Deze docenten surveilleerden gedurende de 50 (V4) of 100 (H4 en V5) minuten zoals bij iedere andere toets. Zodra de leerlingen klaar waren, konden ze hun gemaakte werk op de hoek van hun tafel leggen.

Dit werk werd vervolgens opgehaald door de surveillant. Bij dit ophalen kreeg de leerling een korte vragenlijst over de creatie en het gebruik van het spiekbriefje terug. Het invullen van de vragenlijst ging dus niet af van de toetstijd van de leerling¹. Omdat in de rest van het lokaal nog leerlingen bezig waren, werden de vragenlijsten in stilte ingevuld. Na afronding van de vragenlijst werd ook deze, voorzien van naam, opgehaald door de surveillant.

De leerlingen werden al voor aanvang van de toetsweek ingelicht over het plaatsvinden van het onderzoek. Aangezien niet alle leerlingen 16 jaar of ouder waren, kregen zij de informatie op papier uitgereikt om met hun ouders te delen. Vooraf was aan leerlingen verteld dat zij de vragenlijst ook weer blanco mochten inleveren. Hiermee konden zij aangeven geen deel te willen aan het onderzoek. Dit gaf leerlingen de keuze om aan het onderzoek deel te nemen, ondanks dat de toets zelf wel verplicht was.

Nadat de data uit de verschillende bronnen aan elkaar gekoppeld was, is de data geanonimiseerd en zijn alle referenties naar persoonlijke gegevens verwijderd. Dit onderzoek is goedgekeurd door de ethische commissie van de BMS-faculteit van de Universiteit Twente onder aanvraagnummer 220869.

3.3 Respondenten

Alle respondenten zijn leerlingen die informatica volgen op Stad & Esch. Hierbij zijn wel enige verschillen in de achtergrond van de leerlingen.

Zoals eerder vermeld is het onderzoek verricht binnen drie klassen: H4 (14 leerlingen²), V4 (15 leerlingen) en V5 (12 leerlingen). In V5 had iedere leerling een NG- of NT-profiel. In V4 had 1 leerling een EM-profiel en in H4 had ongeveer de helft van de leerlingen dit profiel. In de H4-klas bevonden zich ook enkele doublerende, opgestroomde en afgestroomde leerlingen. In V5 waren twee leerlingen doublerend. De leeftijd van de leerlingen varieerde van 15 tot 17 jaar. De overgrote meerderheid van de leerlingen waren jongens.

Alle leerlingen hadden evenveel ervaring met het gebruik van spiekbriefjes. Alle klassen mochten al het hele schooljaar bij hun toetsen spiekbriefjes gebruiken. Informatica is

¹Op Stad & Esch is het de regel dat de toetszaal alleen na 50 en 100 minuten verlaten mag worden. De leerlingen hadden dus weinig anders te doen dan het invullen van de vragenlijst. Gelijk na inleveren weggaan kon immers niet.

²Deze klas was oorspronkelijk groter, maar een aantal leerlingen stopte met actief deelnemen voor het einde van het jaar. Deze leerlingen zijn al uit dit totaal en andere statistieken gehaald.

	Geen deelname	Wel deelname
Geen spiekbriefje	0	6
Wel spiekbriefje	1	34

Tabel 2: Deelnamestatistieken van het onderzoek. Leerlingen die geen vragenlijst hebben ingevuld zijn niet meegenomen in de analyses op de volgende pagina's.

Deelvraag	Vragenlijst	Spiekbriefjes	Toetsvragen	Vraagcores
Stresservaring (1)	X			
Toetsvoorbereiding (2)	X			
Relatie vraagcores (3)	X	X	X	X

Tabel 3: Gebruikte instrumenten voor de beantwoording van de verschillende deelvragen.

het enige vak op Stad & Esch dat het meenemen van spiekbriefjes toestaat. Dit is van belang voor de vragenlijst, aangezien leerlingen daarin is gevraagd om vergelijkingen te maken met een situatie zonder spiekbriefjes. Er is aangenomen dat leerlingen hier toe in staat zijn, omdat ze bij alle andere vakken geen spiekbriefjes hebben.

Een enkele leerling zag af van deelname aan het onderzoek. Zes andere leerlingen namen wel deel, maar hadden geen spiekbriefje meegenomen naar de toets. In totaal namen 40 leerlingen daadwerkelijk deel aan het onderzoek. De exacte verdeling van de leerlingen over deelname en spiekbriefjesgebruik is te vinden in tabel 2.

3.4 Instrumenten

Binnen het onderzoek zijn veel verschillende instrumenten gebruikt om data te verzamelen. Een overzicht van de instrumenten en hun gebruik voor de beantwoording van de deelvragen is te vinden in tabel 3. In deze sectie worden de verschillende instrumenten individueel besproken.

Vragenlijst De vragenlijst is opgesteld naar het voorbeeld van verschillende universitaire studies, zoals die van Song, Guo en Thuente (2016). De lijst begint met vier stellingen over stress en spanning voor de toets. De vier vragen daarna gaan in op de toetsvoorbereiding. De lijst eindigt met twee vragen over inzet tijdens de hele periode en drie vragen over de (geschatte) invloed op de toetsscore. De volledige vragenlijst is te vinden in appendix A.

Leerlingen antwoordden op de stellingen middels een vijfpunts-Likertschaal. Deze schaal begint bij "helemaal niet mee eens" en eindigt met "helemaal mee eens". De neutrale optie is verwoord als "niet mee eens en niet mee oneens".

Als extra data bevat de lijst ook vragen over de (geschatte) invloed op het cijfer en de tijd die is besteed aan de toetsvoorbereiding. De tijdsinschattingen zijn uiteindelijk niet verder gebruikt in dit onderzoek.

Spiekbriefjes Aan de leerlingen waren de eisen gesteld dat het spiekbriefje in het eigen handschrift geschreven diende te zijn op een enkelzijdig A4. Deze grootte is een halvering ten opzichte van de universitaire onderzoeken, die overwegend een dubbelzijdig A4 toestonden. Dit is gedaan omdat de stof van VO-toetsen relatief minder is, waardoor door de halvering van de grootte ervoor zorgt dat de hoeveelheid informatieverwerking vergelijkbaar is. De handschrift-eis is overgenomen uit het onderzoek van de Raadt (2012). Hiermee wordt het kopiëren van spiekbriefjes door leerlingen ontmoedigd.

Na het afnemen van de toets zijn alle spiekbriefjes eerst van naam voorzien. Ook is toen gecontroleerd of aan alle voorwaarden (enkelzijdig en handgeschreven) is voldaan door de leerling. Tijdens de beoordeling van de toets zijn de spiekbriefjes verder niet gebruikt. Pas ver na de publicatie van de toetscijfers zijn de vragenlijsten en spiekbriefjes verder verwerkt voor dit onderzoek. De exacte details van deze verwerking zijn toegelicht in sectie 3.5.

Toetsvragen De toetsvragen zijn bij het opstellen van de toets gecategoriseerd op RTTI-niveau en op onderwerp(en). De toetsen bevatten weinig reproductievragen, in lijn met de discussie in sectie 2.3. De vragen zijn grotendeels gegeven als een nieuwe context waarin de opgedane kennis toegepast dient te worden. Een voorbeeldtoets die de vragenstijl illustreert voor domein E (architectuur) is opgenomen in appendix C.

Niet alle afgenomen toetsvragen zijn opgenomen in dit onderzoek. Dit is deels gedaan uit tijdsoverwegingen, maar ook door de wisselende kwaliteit van toetsvragen. In lijn met de aanbevelingen van het Cito (sectie 2.4) hebben alle geanalyseerde toetsvragen een moeilijkheidsgraad tussen de 0.30 en 0.85 en een Rit-waarde van minstens 0.3. Een overzicht van de geselecteerde vragen is te vinden in tabel 4. Dezelfde data voor *alle* afgenomen toetsvragen is opgenomen in appendix B.

Vraagcores Alle toetsen zijn door dezelfde docent nagekeken om inconsistenties bij nakijken te voorkomen. Tijdens het nakijken waren het spiekbriefje en de vragenlijst (indien ingevuld) niet bereikbaar voor de corrector om eventuele invloed hiervan op de correctie te voorkomen.

3.5 Analyse

De antwoorden op de vragenlijsten zijn verwerkt door statistische toetsen toe te passen op een numerieke interpretatie van de antwoordmogelijkheden. Daarnaast is binnen elk spiekbriefje gezocht naar inhoud die helpt bij de beantwoording van specifieke toetsvragen. Vervolgens is de correlatie tussen de aanwezigheid van informatie op het spiekbriefje en de score op een toetsvraag onderzocht. Deze vormen van analyse zijn hieronder verder toegelicht.

Klas	Vraag	Niveau	Moelijkheidsgraad	Rit-waarde
H4	1b	T1	0.71	0.43
	4	I	0.50	0.48
	5b	T1	0.50	0.66
	5c	I	0.57	0.61
	6	T2	0.35	0.74
	7c	T1	0.64	0.85
	V4	1j	R	0.70
1k		I	0.50	0.56
2c		T2	0.67	0.48
2d		I	0.60	0.54
3a		T2	0.84	0.69
3c		T1	0.33	0.81
V5		1c	R	0.71
	3c	T1	0.61	0.54
	3e	I	0.50	0.55
	3h	I	0.58	0.40
	4b	T2	0.58	0.40
	5c	T1	0.71	0.57

Tabel 4: Overzicht van de geanalyseerde toetsvragen.

3.5.1 Verwerking Likertschalen vragenlijst

Hoewel een snelle blik op een histogram een goede indruk geeft over resultaten van een vragenlijst, geeft dit weinig houvast om de juiste conclusies te kunnen trekken. Voor zulke doeleinden is een statistische toets beter geschikt.

Wilcoxon-toets Ieder antwoord van de vijfpunts-Likertschaal is omgezet naar een getal van 1 tot 5, waarbij 1 overeenkomt met “helemaal niet mee eens” en 5 met “helemaal mee eens”. Het neutrale antwoord, tussen deze uitersten in, komt dan overeen met de waarde 3. Vervolgens is een Wilcoxon-toets gebruikt om te bepalen of de antwoorden significant ($\alpha = 5\%$) afwijken van de neutrale waarde 3.

Er is overwogen om de bekendere *t*-toets te gebruiken om significantie te bepalen. Deze neemt echter aan dat de data een normaalverdeling volgt en dat data “continu” (met oneindig veel numerieke antwoorden) is, wat overduidelijk niet het geval is bij slechts vijf antwoordmogelijkheden. De analyse is daarom in lijn met de aanbevelingen uit Schrum, Johnson, Ghuy en Gombolay (2020), die een niet-parametrische toets (zoals Wilcoxon) voorschrijven.

Uitvoering toets Alle ingevulde vragenlijsten zijn eerst handmatig gedigitaliseerd en vervolgens ingeladen in de statistiektaal R. De ingebouwde Wilcoxon-toets is gebruikt

om de p-waardes te berekenen. De diverse figuren zijn gegenereerd via het *likert*-package³.

3.5.2 Verwerking spiekbrieftjes

Om een mogelijke relatie tussen de inhoud van de spiekbrieftjes en de scores op de toets te onderzoeken, zijn alle spiekbrieftjes doorzocht op inhoud die gerelateerd is aan de geselecteerde toetsvragen.

Selectie toetsvragen Het was niet mogelijk om voor alle toetsvragen van alle toetsen een spiekbrieftjesanalyse te doen. Dit had o.a. te maken met de beschikbare hoeveelheid tijd en de wisselende kwaliteit van de toetsvragen (zie sectie 3.4). Om deze redenen zijn van iedere toets 6 vragen geselecteerd waarop de analyse is uitgevoerd. Deze selectie is zo gemaakt dat verschillende vraagniveau's (reproductie, toepassing en inzicht) waar mogelijk aanwezig zijn.

Categorisering spiekbrieftjes Ieder spiekbrieftje is gecontroleerd op aanwezigheid van begrippen, definities, tekeningen of voorbeelden die mogelijk zouden kunnen helpen bij het beantwoorden van de desbetreffende toetsvraag. Deze informatie is gekoppeld aan de vraagscores, zodat bij iedere vraagscore bekend is of deze is verkregen met of zonder mogelijk behulpzame informatie op het spiekbrieftje. Gebruik van meer gedetailleerde info (zoals "geen info", "definitie" en "voorbeeld") was niet mogelijk omdat dan het aantal vraagscores per categorie te klein zou worden voor enige statistische significantie.

3.5.3 Statistische analyse data

De data uit de spiekbrieftjes zijn op twee manieren verwerkt in statistische toetsen. De eerste manier was op vraagniveau, waarbij iedere *toetsvraag* apart geanalyseerd is. De tweede manier was op leerlingniveau, waarbij iedere *leerling* apart geanalyseerd is.

Het verschil tussen de twee analyses is geïllustreerd in figuur 1. In deze tabel zijn de scores uit het nakijken en de data uit de spiekbrieftjescategorisering weergegeven. Voor ieder spiekbrieftje is gekeken of voor een toetsvraag wel of niet relevante informatie op het spiekbrieftje staat. Dit is gecodeerd als "wel" of "niet" bij iedere toetsvraag. De analyse per toetsvraag is de verticale manier van kijken: scoren op een vraag leerlingen met info op het spiekbrieftje beter (t.o.v. de klas)? De analyse per leerling is juist horizontaal afgebeeld: scoort een specifieke leerling hoger (t.o.v. zichzelf) als er info op het spiekbrieftje staat voor die vraag? Op beide analyses wordt hieronder verder ingegaan.

Analyse per toetsvraag Om te beoordelen of er een significant verschil was tussen de scoreverdeling van de leerlingen met en leerlingen zonder info op het spiekbrieftje, is gebruikgemaakt van een Wilcoxon-toets. Voor deze toets wordt een nulhypothese gesteld

³<https://jbryer.github.io/likert/>

	Vraag 1	Vraag 2	Vraag 3
Leerling A	0,5 / Niet	0 / Niet	1 / Wel
Leerling B	2 / Wel	0 / Niet	1 / Niet
Leerling C	2 / Niet	0 / Niet	0 / Niet
Leerling D	1 / Niet	1 / Wel	1 / Wel

Figuur 1: Visualisatie van de verschillende analyses. Verticaal is de analyse per toetsvraag. Horizontaal is de analyse per leerling. "1/Niet" geeft aan dat deze leerling op deze vraag 1 punt heeft gehaald en geen relevante informatie op het spiekbriefje had staan.

dat de twee groepen metingen (scores met en scores zonder info op spiekbriefje) uit dezelfde populatie afkomstig zijn. De Wilcoxon-toets geeft over deze nulhypothese een p-waarde. Bij een lage p-waarde ($\alpha = 5\%$) wordt de nulhypothese verworpen. Dit betekent dat de twee groepen metingen (scores met en zonder info op spiekbriefje) significant van elkaar verschillen. Deze significantietoets is uitgevoerd voor iedere geanalyseerde toetsvraag.

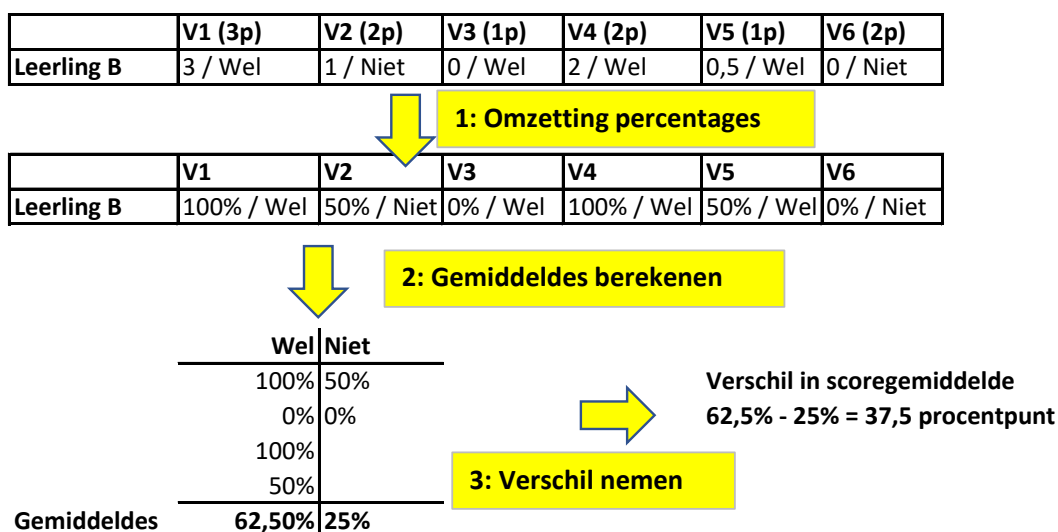
Analyse op leerlingniveau De significantietoets zoals hierboven beschreven is slechts één manier om de data te duiden. De tweede analyse, op leerlingniveau, werkt op de omgekeerde manier. Hierbij is gekeken of een specifieke *leerling* gemiddeld hoger scoort (t.o.v. zichzelf) bij vragen waar relevante info op het spiekbriefje stond.

Iedere score voor een toetsvraag is eerst uitgedrukt in een percentage. Zo levert 1 punt op een vraag van 2 punten een score op van 50%. Voor iedere leerling zijn twee gemiddeldes berekend: de gemiddelde score bij vragen waar geen info aanwezig was op het spiekbriefje, en een gemiddelde score waar deze info wel aanwezig was. Voor de analyse is gekeken naar het *verschil* tussen deze scoregemiddeldes. De hele berekening is uitgebeeld voor één leerling in figuur 2.

Het verschil is niet altijd te berekenen, bijvoorbeeld bij leerlingen zonder spiekbriefje. Ook kan het zijn dat een scoregemiddelde gebaseerd wordt op slechts 1 toetsvraag, bijvoorbeeld wanneer het spiekbriefje slechts informatie bevatte over 1 onderwerp. Om deze redenen zijn geen exacte p-waarden berekend, maar is alleen gezocht naar een trend in de data.

3.5.4 Beantwoording onderzoeksvragen

Met de informatie uit de vragenlijsten, vraagscores, spiekbriefjes en bovenstaande analyses is voldoende data verzameld voor de beantwoording van de drie deelvragen. In deze laatste sectie wordt besproken hoe alle informatie is gecombineerd om de drie deelvragen van het onderzoek te beantwoorden.



Figuur 2: Berekening van het verschil in scoregemiddelde voor één leerling. De “wel” of “niet” in de bovenste tabel geeft aan of deze leerling voor de betreffende vraag relevante info op het spiekbriefje had staan. Het getal aan de linkerkant hiervan bevat de score die deze leerling op deze vraag behaald had.

Stresservaring In de vragenlijst zijn vier vragen opgenomen over de hoeveelheid stress en spanning die een leerling ervaart. Met de uitgevoerde Wilcoxon-toets is onderzocht of de antwoorden significant afwijken van een neutrale waarde. Is dit consistent het geval, dan geeft een blik op een histogram duidelijkheid over de richting van deze afwijking. Met andere woorden, zijn leerlingen het vaak eens of juist vaak oneens met de stelling? Zijn de leerlingen het vaak significant eens, dan hebben de spiekbriefjes een significant verminderend effect op de stresservaring.

Motivatie en voorbereiding Een vergelijkbare werkwijze is gebruikt voor de beantwoording van de tweede deelvraag. In de vragenlijst zijn vier stellingen opgenomen over verschillende aspecten van toetsvoorbereiding en de motivatie hiervoor. De statistische toetsen geven opnieuw aan of de antwoorden significant verschillen van de neutrale waarde. Een blik op een histogram geeft de richting van deze afwijking. Hieruit kan worden afgeleid of de spiekbriefjes een positief of negatief effect hebben gehad op de motivatie of de (structuur van de) toetsvoorbereiding.

Correlatie vraagcores Er zijn twee analyses uitgevoerd op de informatie uit spiekbriefjes en vraagcores: de analyse per toetsvraag en de analyse per leerling. Voor de analyse per toetsvraag zijn 18 significantie-toetsen uitgevoerd, eentje voor iedere geanalyseerde toetsvraag. Hierdoor zijn er ook 18 p-waardes. Tegelijkertijd is er van de meerderheid van de leerlingen een verschil in scoregemiddelden bepaald.

Als de p-waardes uit de analyse per toetsvraag vaak significante verschillen aangeven, dan is er sprake van een correlatie tussen de inhoud van het spiekbriefje en de scores op een toetsvraag. Een blik op de data wijst dan uit of dit een positieve of negatieve correlatie is. Mogelijk is de correlatie niet overal zichtbaar, maar bijvoorbeeld alleen bij reproductievragen. Een gedetailleerde blik op de vragen waar wel of geen significante waarden zijn is dus vereist.

De analyse op leerlingniveau geeft aanvullende informatie hierop. Zijn de verschillen in scoregemiddelden overwegend positief, dan scoren leerlingen gemiddeld hoger bij relevante info op hun spiekbriefje. Verdere nuances hierop, bijvoorbeeld een onderscheid in sterke en zwakke leerlingen, zijn te maken door de data te combineren met andere variabelen (zoals het toetscijfer) in een grafiek.

3.6 Betrouwbaarheid, validiteit en overdraagbaarheid

In dit onderzoek is gestreefd om betrouwbare, valide en overdraagbare resultaten te produceren. Een onderzoek is betrouwbaar als een herhaling van het onderzoek of een uitvoering door een andere onderzoeker dezelfde resultaten oplevert. Validiteit betekent dat wordt onderzocht wat men daadwerkelijk wil onderzoeken. Overdraagbaarheid gaat om de vraag in hoeverre de opbrengsten van het onderzoek vertaald kunnen worden naar de beroepscontext van anderen (Van der Donk & Van Lanen, 2020).

Betrouwbaarheid De betrouwbaarheid wordt vergroot door toevallige verstoringen zoveel mogelijk te voorkomen (Van der Donk & Van Lanen, 2020). De afname van de vragenlijst heeft plaatsgevonden in de toetsruimte, terwijl de leerlingen volgens toetsweekprotocol nog niet weg mochten. Hierdoor konden de leerlingen in alle stilte en zonder tijdsdruk de vragenlijst invullen. De geanalyseerde vragen op de toetsen zijn gekozen op basis van criteria opgesteld door Cito (zie secties 2.4 en 3.4). Hierdoor is de invloed van "slechte" toetsvragen geminimaliseerd. Alle toetsen zijn door dezelfde docent opgesteld en nagekeken om consistentie bij deze activiteiten te waarborgen. Tevens is gezorgd voor een zo groot mogelijke leerlingpopulatie en is de analyse uitgevoerd met statistieksoftware R, wat een correcte uitvoering van statistische toetsen mogelijk maakt (Chambers, 2010, p. 10).

Validiteit De beantwoording op de vragenlijst is verlopen via een standaard vijfpunts-Likertschaal. De vragenlijst is geordend in blokken met een korte toelichting, waardoor de lijst overzichtelijker is en de leerling weet wat er wordt verwacht. Tevens bevatte de vragenlijst bij veel aspecten (zoals stresservaring) controlevragen die gebruikmaken van vergelijkbare begrippen. Bij de analyses is gebruikgemaakt van statistische toetsen om de juiste conclusies te trekken uit de resultaten bij o.a. uiteenlopende beantwoording en variërende klasgroottes. Tenslotte is bij het beantwoorden van de deelvragen rekening gehouden met de beperkingen van de instrumenten. Zo is, bij gebrek aan meetapparatuur, niet onderzoek gedaan naar de (biologische) stress, maar naar de stresservaring.

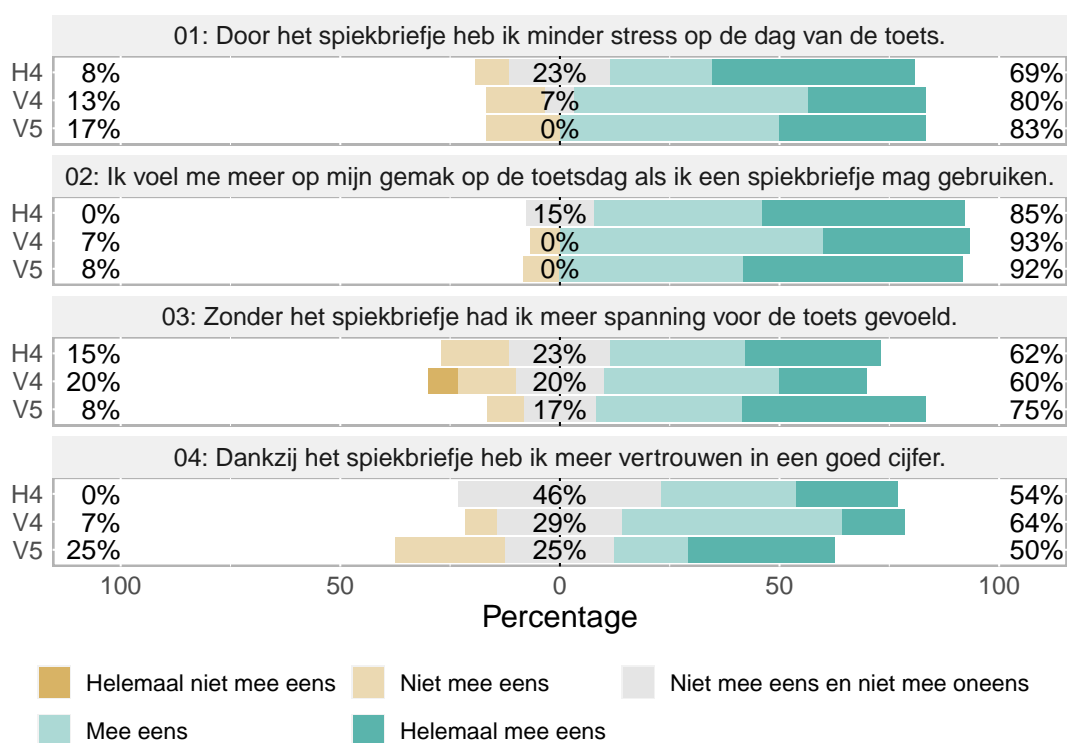
Overdraagbaarheid De overdraagbaarheid kan vergroot worden door de context te specificeren en doelgroepgericht de resultaten te rapporteren (Van der Donk & Van Lanen, 2020). De context is al eerder gespecificeerd in sectie 3.1. Hierin zijn o.a. de school, de leerlingenpopulatie en de toetsonderwerpen nader toegelicht. De toetswijze is tevens zichtbaar gemaakt middels een oefentoets, opgenomen in appendix C. Alle resultaten zijn per klas weergegeven en toegelicht, waardoor docenten op verschillende jaarlagen en niveau's zich kunnen focussen op de resultaten die voor hun leerlingenpopulatie relevant zijn.

4 Resultaten

De resultaten van de vragenlijst en de analyse van de spiekbrieftjes zijn in de volgende pagina's toegelicht. De resultaten zijn per deelvraag besproken. Waar relevant zijn de p-waarden van de Wilcoxon-toetsen toegevoegd in kleine tabellen bij iedere sectie.

4.1 Invloed op stresservaring

Alle informatie over de stresservaring is afkomstig uit de beantwoording van de eerste vier vragen van de vragenlijst. De resultaten zijn visueel gepresenteerd en gerangschikt per klas gepresenteerd. De beantwoording van de vragen over stresservaring zijn te vinden in figuur 3. De percentages links, rechts en in het midden van iedere horizontale balk staan voor het aantal leerlingen die het respectievelijk (helemaal) niet eens, (helemaal) mee eens en neutraal waren over de stelling.



Figuur 3: Antwoorden op vragen over de stresservaring, gerangschikt per klas.

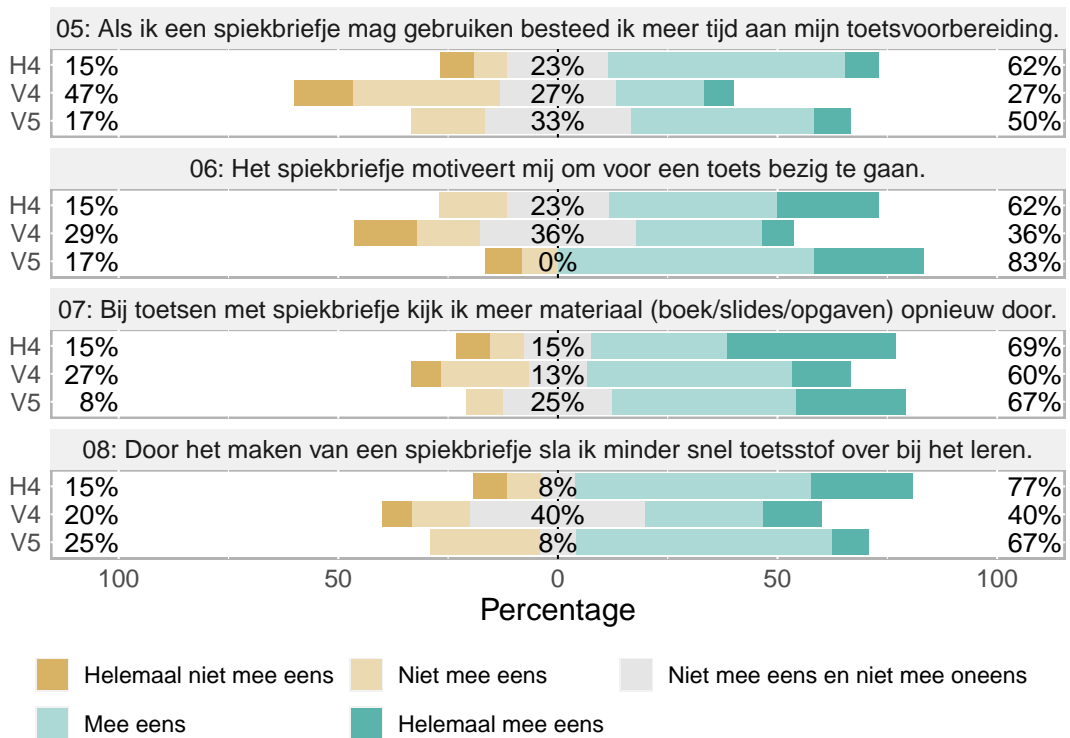
Uit de statistische toetsen blijkt dat bij alle stellingen een significant verschil vanaf de neutrale waarde bestaat. Figuur 3 laat de richting van deze afwijking zien. Leerlingen geven aan dat er bij gebruik van een spiekbriefje minder stress wordt ervaren. Hetzelfde beeld ontstaat bij de vervolgvragen met andere verwoordingen, zoals "spanning" en (minder) "op je gemak zijn". De complete lijst met p-waardes is te vinden in tabel 5.

Vraag	p-waarde
Minder stress (1)	< 0.0001
Meer op gemak (2)	< 0.0001
Minder spanning (3)	0.0003
Meer vertrouwen (4)	0.0002

Tabel 5: p-waardes bij stellingen over stresservaring.

4.2 Invloed op toetsvoorbereiding

De gegevens over de toetsvoorbereiding zijn ook afkomstig uit de vragenlijst. Direct na de vragen over de stresservaring bevatte de lijst vier vragen (nummer 5 t/m 8) over o.a. de besteedde tijd en de (ervaren) motivatie en structuur van de toetsvoorbereiding. De verdeling van de antwoorden is te vinden in figuur 4.



Figuur 4: Antwoorden op de vragen over de toetsvoorbereiding, gerangschikt per klas.

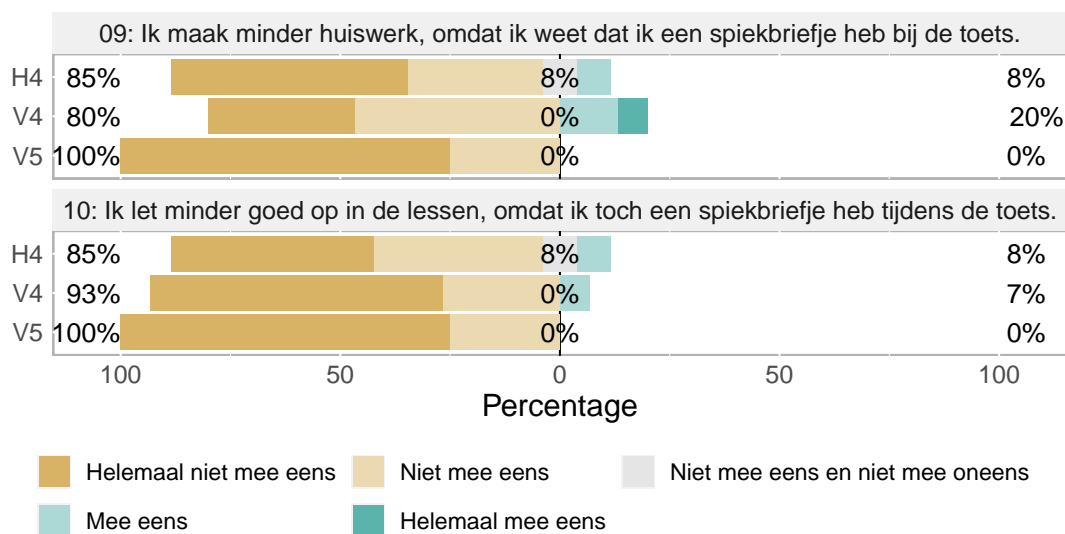
Uit de statistische toetsen blijkt dat bij drie van de vier stellingen een significante afwijking van de neutrale waarde bestaat. Alleen bij stelling 5 is de nulhypothese niet verworpen. Bij alle andere stellingen laten de p-waarden in tabel 6 en de waarden in figuur 4 zien dat een significante afwijking bestaat, namelijk dat er meer motivatie en structuur in de toetsvoorbereiding ervaren wordt.

Vraag	p-waarde
Meer bestede tijd (5)	0.339
Extra motivatie (6)	0.0208
Meer doorkijken (7)	0.00161
Minder overslaan (8)	0.00938

Tabel 6: p-waarden bij stellingen over stresservaring.

Leerlingen denken soms bij toetsvoorbereiding alleen aan het stampen en samenvatten voor een toets. Bij de vragen hierboven werden leerlingen bevraagd over deze "smalle" interpretatie. Toetsvoorbereiding in "brede" zin gaat echter verder dan de enkele activiteiten direct voor de toets. Leerlingen werken gedurende een hele periode zowel op school als thuis naar een toets toe. De lesactiviteiten en het huiswerk gedurende de periode kunnen dus, bij een "brede" interpretatie, ook gezien worden als toetsvoorbereiding.

In de vragenlijst zijn twee stellingen (nummer 9 en 10) opgenomen over de leshouding en het huiswerk gedurende de hele periode (de "brede" toetsvoorbereiding). De beantwoording hierop is te vinden in figuur 5.



Figuur 5: Antwoorden op vragen over de houding in de les, gerangschikt per klas.

Bij beide stellingen vond de statistische toets een p-waarde ver onder de 0.0001. Er was dus een significant verschil vanaf de neutrale waarde. Figuur 5 laat de richting van de afwijking zien. Leerlingen waren het oneens met de stellingen, die suggereerden dat door spiekbriefjes minder huiswerk gemaakt wordt en minder aandacht in de les is gedurende de hele periode.

4.3 Correlatie spiekbriefjes en vraagscores

De impact op de vraagscores is onderzocht met twee analyses: de analyse per toetsvraag en de analyse per leerling. Bij de analyse per toetsvraag zijn voor iedere toetsvraag p-waardes berekend middels een Wilcoxon-toets. Bij de analyse per leerling zijn verschillen in scoregemiddelden berekend en uitgetekend in grafieken. Beide analyses volgen hieronder, te beginnen met de analyse per toetsvraag.

4.3.1 Analyse per toetsvraag

Bij de analyse per toetsvraag is binnen een klas onderzocht of leerlingen met relevante informatie op hun spiekbriefje significant hoger scoorden op een toetsvraag dan leerlingen zonder zulke informatie op hun spiekbriefje. Voor iedere klas is dit verband voor 6 toetsvragen onderzocht.

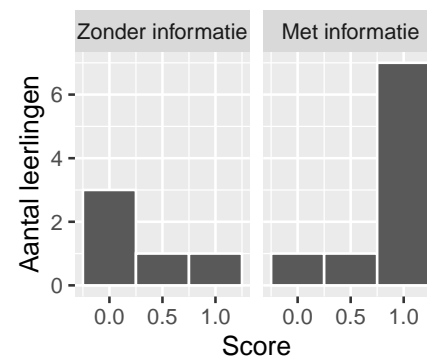
H4-toets Bij de H4-toets was er bij de helft van de vragen een significant verschil verschil in score tussen de groepen met en zonder informatie op hun spiekbriefje. Bij drie van de zes geanalyseerde vragen was de p-waarde dus kleiner dan 0.05. Nadere bestudering van de histogrammen toonde aan dat bij deze vragen leerlingen met info op het spiekbriefje significant hoger scoorden. Al met al scoorden dus bij drie van de zes vragen de leerlingen met info op het spiekbriefje significant hoger dan de leerlingen zonder. De gemiddelde scores en p-waardes per toetsvraag zijn weergegeven in tabel 7.

De puntenverdeling van toetsvraag 7c is weergegeven in figuur 6. Deze vraag gaf het "minst significante" verschil. Desondanks is in het histogram duidelijk te zien dat er een groot verschil was in de puntenverdeling. Bij de andere toetsvragen met een lagere p-waarde was dit verschil nog groter.

Het is tenslotte interessant om te benoemen dat alle vragen met een significant verschil toepassingsvragen waren. Bij geen van de inzichtsvragen was er een noemenswaardig verschil. De toets bevatte geen reproductievragen.

Vraag	Type	Gem. score	Gem. score	p-waarde
		zonder informatie	met informatie	
1b	T1	0.63	1.75	0.0120
4	I	1.1	1.4	0.644
5b	T1	0.22	0.30	0.645
5c	I	0.56	0.60	0.938
6	T2	0.14	1.3	0.0102
7c	T1	0.30	0.83	0.0430

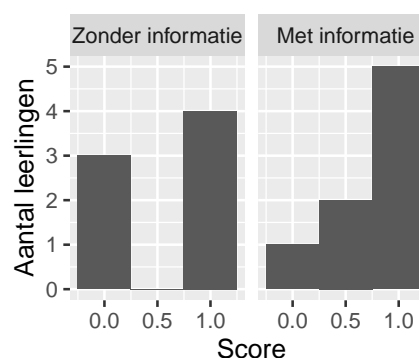
Tabel 7: Geanalyseerde vragen, gemiddelde scores en p-waardes bij H4-toets.



Figuur 6: Puntenverdeling vraag 7c bij H4.

V4-klas Bij de V4-klas was bij geen van de geanalyseerde vragen een significant verschil in score tussen de groepen met en zonder informatie op hun spiekbriefje te zien. De statistieken per onderzochte toetsvraag zijn te vinden in tabel 8. Om een idee te krijgen van de verschillen tussen de scores, geeft het histogram in figuur 7 de scoreverdeling weer bij vraag 2c. Bij beide groepen zijn een aantal leerlingen met volle en met halve of geen punten te vinden. Ditzelfde patroon herhaalde zich bij veel andere vragen. Een lichte “uitzondering” was vraag 1j. Bij deze reproductievraag had iedereen met informatie op het spiekbriefje de vraag goed. Dit gaf echter geen significante p-waarde, omdat een aantal leerlingen zonder spiekbriefje-info de vraag ook goed hadden.

Vraag	Type	Gem. score	Gem. score	p-waarde
		zonder informatie	met informatie	
1j	R	0.59	1.00	0.135
1k	I	0.55	0.40	0.632
2c	T2	0.57	0.75	0.597
2d	I	0.55	0.75	0.469
3a	T2	2.75	2.45	0.506
3c	T1	0.25	0.36	0.750



Tabel 8: Geanalyseerde vragen, gemiddelde scores en p-waardes bij V4-toets.

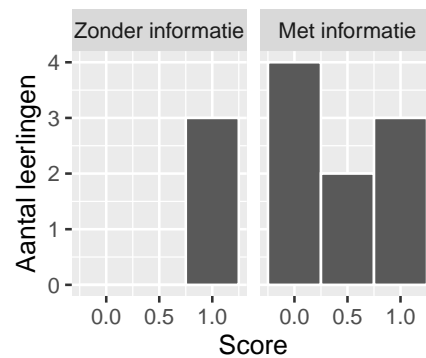
Figuur 7: Puntenverdeling vraag 2c bij V4.

V5-klas Bij de V5-klas ontstond een beeld dat vergelijkbaar was met de V4-klas. Opnieuw was er bij geen van de (deel)vragen een significant verschil tussen de groepen met en zonder relevante info op hun spiekbriefje. De complete statistieken zijn te vinden in tabel 9. Twee waarden zijn de moeite waard om in het bijzonder toe te lichten. De enige reproductievraag in de lijst, vraag 1c, gaf een p-waarde van 1. Bij deze vraag hadden de spiekbriefjes dus geen enkel geobserveerd effect. De enige vraag die een p-waarde had die in de buurt kwam van significantie, was vraag 4b. De Wilcoxon-toets kijkt echter alleen of er een verschil tussen de twee groepen is. Het histogram in figuur 8 toont dat leerlingen *zonder* informatie op het spiekbriefje juist gemiddeld *hoger* scoorden.

Tenslotte gaven twee andere vragen (3c en 5c) histogrammen die, bij een hoger leerlingaantal, mogelijk wel een significante waarde in hadden opgeleverd. Deze histogrammen lieten een duidelijke trend zien dat leerlingen met info op hun spiekbriefje hoger scoorden. Met slechts 12 leerlingen in deze klas was dit echter niet genoeg voor een significante p-waarde.

Vraag	Type	Gem. score zonder informatie	Gem. score met informatie	p-waarde
1c	R	0.70	0.71	1.00
3c	T1	0.60	1.14	0.229
3e	I	0.50	0.50	1.00
3h	I	1.00	1.25	0.648
4b	T2	1.00	0.44	0.0862
5c	T1	0.58	0.83	0.336

Tabel 9: Geanalyseerde vragen, gemiddelde scores en p-waardes bij V5-toets.



Figuur 8: Puntenverdeling vraag 4b bij V5.

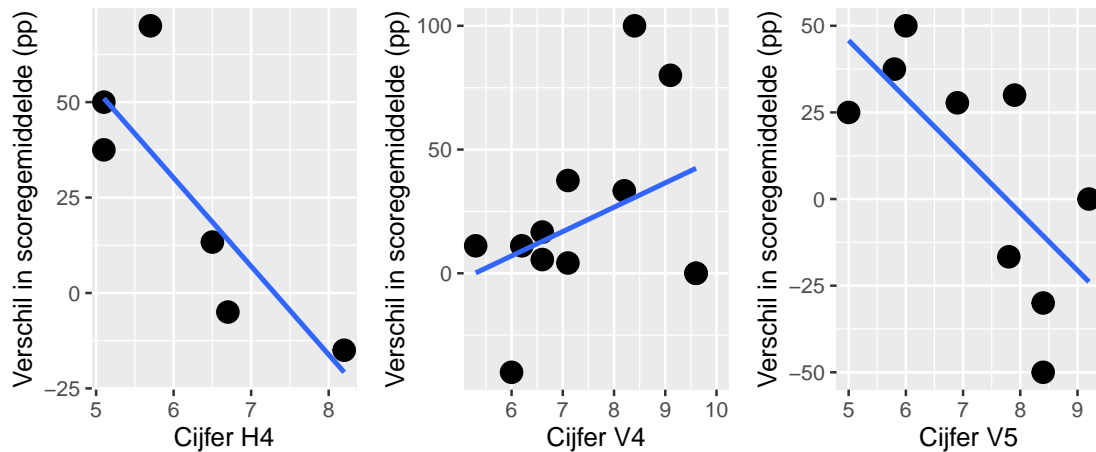
4.3.2 Analyse per leerling

In de voorgaande analyse is op een klassenniveau gekeken naar de vraagcores. Gegeven een bepaalde vraag, scoren leerlingen in een klas met info op het spiekbriefje significant anders dan andere leerlingen in dezelfde klas, zonder info over deze vraag op hun spiekbriefje?

Bij de tweede analyse, op leerlingniveau, is niet gekeken naar het gemiddelde over een klas, maar naar de scores van een individuele leerling. Hiermee is te onderzoeken of een specifieke leerling profiteert van aanwezige informatie op het spiekbriefje. Hierdoor kan de impact onderzocht worden buiten de context van de klas als geheel, waardoor het verband bij sterke of zwakke leerlingen los van elkaar te onderzoeken is.

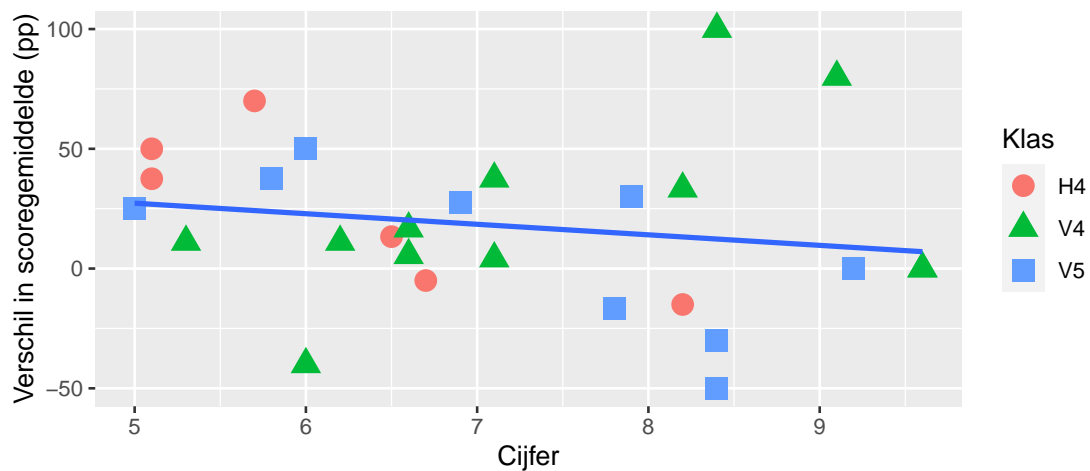
Uitgesloten leerlingen Een belangrijke voorwaarde voor deze methode is dat een leerling bij sommige vragen niet, en bij andere vragen juist wel informatie op het spiekbriefje heeft. Als een leerling bijvoorbeeld bij geen enkele vraag iets relevants op het spiekbriefje heeft, is er geen verschil tussen scores, omdat alle vragen in dezelfde groep vallen. Hierdoor kon dit onderzoek op leerlingniveau bij 14 leerlingen (een derde van het totaal) niet uitgevoerd worden, omdat ze geen (gevuld) spiekbriefje hadden meegenomen of omdat hun spiekbriefje perfect compleet was. Beide mogelijkheden kwamen ongeveer even vaak voor.

Vershil in scoregemiddelde Zoals beschreven op pagina 16 is gekeken naar het verschil tussen scoregemiddelden. In de statistieken hieronder is dit berekend als "(gemiddelde score met info) - (gemiddelde score zonder info)". Een positieve waarde betekent dus dat een leerling gemiddeld hoger scoorde bij vragen waar wel informatie voor op het spiekbriefje stond. Bij een negatieve waarde haalde de leerling juist een lagere score. Omdat hier een verschil tussen percentages is genomen, is de eenheid in procentpunten (pp). Haalt een leerling gemiddeld 40% score bij vragen met info en 50% score bij vragen zonder info op het spiekbriefje, dan is er een verschil in scoregemiddelde van -10pp.



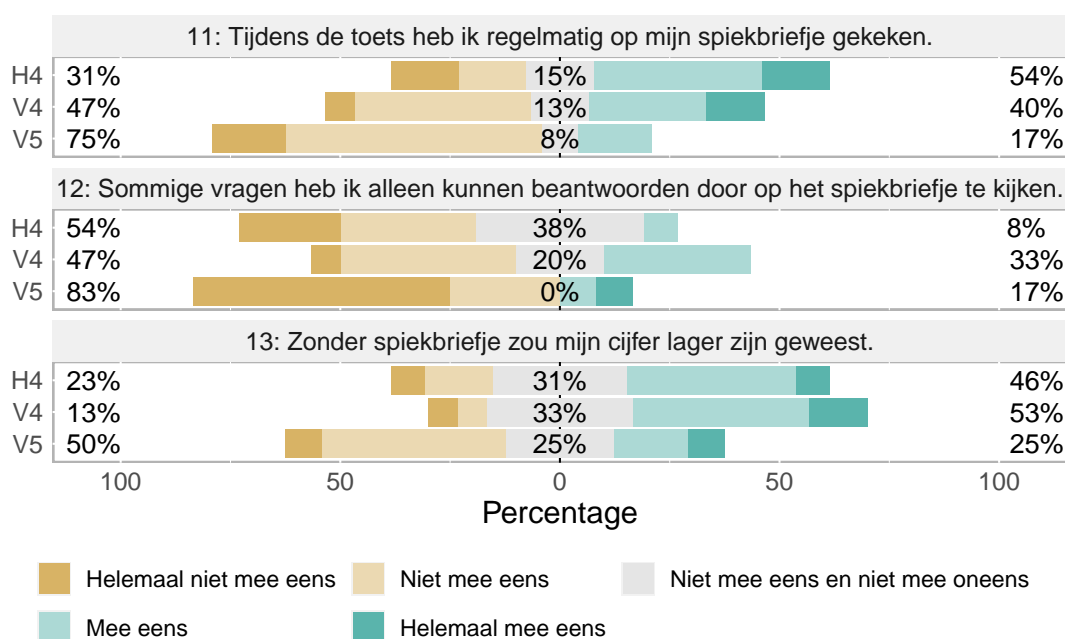
Figuur 9: Verschil in scoregemiddelde van leerlingen in H4, V4 en V5, afgezet tegen hun cijfers. Ieder punt in de grafiek is het verschil in scoregemiddelde van 1 leerling. Let op dat de verticale assen op verschillende hoogten nulpunten hebben.

Verschillen tussen klassen Om te onderzoeken of leerlingen met lagere cijfers mogelijk meer of minder profiteerden van hun spiekbrieftje dan leerlingen met hogere cijfers, is het verschil in scoregemiddelde afgezet tegen het toetscijfer. Figuur 9 laat dit zien voor iedere klas apart. Te zien is dat bij twee klassen leerlingen meer profiteerden bij een laag cijfer, en bij een klas leerlingen juist (ietwat) minder.



Figuur 10: Verschil in scoregemiddelde van alle leerlingen, afgezet tegen hun cijfer. Voor iedere leerling is dit berekend als "gemiddelde score met info op spiekbrieftje - gemiddelde score zonder info op spiekbrieftje". Ieder punt in de grafiek is de uitkomst van deze som voor één leerling. De berekening is eerder uitgebeeld in figuur 2.

Gecombineerde trend De drie grafieken uit figuur 9 zijn gecombineerd in één grafiek te vinden in figuur 10. Wanneer alle klassen samen worden genomen, is er een licht dalende trend in de data te zien. Dit betekent dat het verschil in scoregemiddelde meestal iets hoger is bij leerlingen met een lager cijfer. Tegelijkertijd blijft de trendlijn boven de 0. Over het algemeen heeft een leerling dus een hogere score bij vragen waar wel info over op het spiekbriefje stond, dan vragen waar dit niet zo is. Dit verschil wordt kleiner naarmate het cijfer van de leerling hoger wordt. Tenslotte is het goed om te benadrukken dat de spreiding in deze grafiek gigantisch is. Het gevonden verband is daardoor ook vrij zwak.



Figuur 11: Antwoorden op de vragen over het nut van het spiekbriefje tijdens de toets, gerangschikt per klas.

Vragenlijst De leerlingen hebben in de vragenlijst zelf ook een inschatting gemaakt van de impact van de spiekbriefjes op hun toetsscores. Hun antwoorden zijn te vinden in figuur 11. Alleen de middelste stelling, over de noodzaak van het spiekbriefje voor sommige vragen (nummer 12), leverde een significante ontcrachting op. De eerste vraag (nummer 11) ging over het gebruik van spiekbriefjes *tijdens* de toets. De antwoorden van de leerlingen bevestigen deze stelling niet, maar wijzen deze ook niet af. Ook de laatste vraag (nummer 13), over de invloed op het cijfer, werd niet bevestigd noch ontcracht. De leerlingen zijn het dus niet met deze stelling eens, maar er ook niet mee oneens. Er is simpelweg geen eenduidige mening over.

5 Conclusies en discussie

De voorgaande hoofdstukken hebben het doel, de uitvoering en de resultaten van het onderzoek beschreven. In dit hoofdstuk is eerst de eindconclusie van het onderzoek gegeven. Na deze eindconclusie volgt de onderbouwing op basis van de drie deelvragen. Het hoofdstuk wordt afgesloten met de discussie, waarin ook aanbevelingen voor toekomstig onderzoek zijn opgenomen.

5.1 Conclusie

In dit onderzoek is middels afname van een vragenlijst en analyse van ingeleverde spiekbriefjes onderzocht wat de effecten zijn van spiekbriefjes op informaticaleerlingen in het VO. Er is voor 40 leerlingen in H4, V4 en V5 uitgezocht wat de effecten zijn op de stresservaring en toetsvoorbereiding. Ook is de correlatie tussen de inhoud van het spiekbriefje en de scores op de toets onderzocht.

Uit de vragenlijst is gebleken dat leerlingen significant minder stress ervaren in aanloop naar en tijdens de toets. Ook geven leerlingen aan dat ze meer motivatie en structuur in hun toetsvoorbereiding ervaren. Leerlingen besteedden echter niet significant meer of minder tijd aan hun toetsvoorbereiding.

Uit de analyse van de spiekbriefjes is gebleken dat er weinig tot geen correlatie bestaat tussen de inhoud van het spiekbriefje en de score op een toetsvraag. Slechts drie van de 18 onderzochte toetsvragen lieten een significant verschil zien in score zien. De data suggereert wel dat zwakkere leerlingen een hoger scoreverschil (meer "winst") hebben dan sterke leerlingen, wanneer info voor een toetsvraag op het spiekbriefje staat. De inhoud en opzet van de spiekbriefjes verschilden sterk, ook tussen leerlingen met dezelfde toetsscores.

Meer onderzoek is nodig om de exacte impact van de spiekbriefjes op de vraagcores vast te stellen. Mogelijk verschilt deze impact sterk per leerling, onderwerp of toetsopzet. Zowel replicatie-onderzoek als alternatieve onderzoeksopzetten, met meer focus op kwalitatieve data, zijn daarom aan te raden voor verder onderzoek naar de impact van spiekbriefjes op de vraagcores van leerlingen op het VO.

5.2 Beantwoording deelvragen

Bovenstaande conclusie geeft al kort antwoord op de deelvragen die zijn opgesteld aan het einde van hoofdstuk 2. Voor de volledigheid zijn deze hieronder kort herhaald.

1. In hoeverre helpen spiekbriefjes bij het verminderen van de ervaren stress bij informaticoetsen op het VO?
2. Welke invloed hebben spiekbriefjes op de motivatie en voorbereiding bij informaticoetsen op het VO?

3. Welke correlatie is zichtbaar tussen de inhoud van een spiekbriefje en de behaalde score op toetsvragen bij informaticoetsen op het VO?

In deze sectie wordt verder ingegaan op de beantwoording van deze deelvragen en de totstandkoming van de conclusie.

5.2.1 Impact op stresservaring

Voor de beantwoording van deze deelvraag is gebruikgemaakt van de vragenlijst. De antwoorden op de bijbehorende vragen zijn eerder gepresenteerd in sectie 4.1. Bij alle vragen die met stress te maken hadden, gaf een statistisch significant deel van de leerlingen aan dat het spiekbriefje hielp om de stress te verminderen. Leerlingen voelden zich bijvoorbeeld meer op hun gemak, ervaarden minder spanning en hadden meer vertrouwen in een goed cijfer.

Al met al kan worden geconcludeerd dat spiekbriefjes een zeer significante bijdrage leveren aan het verminderen van de stresservaring van leerlingen. Aangezien leerlingen dit hebben beantwoord op basis van hun eigen indrukken, gaat het hier expliciet om de *stresservaring* en niet de (biologische) *stress*. De discussie (sectie 5.3.2) gaat verder op dit verschil in.

5.2.2 Impact op toetsvoorbereiding

In de vragenlijst waren vragen opgenomen die ingingen op de verschillende aspecten van de toetsvoorbereiding. Hierbij valt te denken aan de motivatie, de bestede tijd en het gestructureerd werken. De antwoorden van de leerlingen zijn eerder gepresenteerd in sectie 4.2, en varieerden sterk bij de verschillende aspecten.

Op de vraag of leerlingen meer tijd besteden aan hun toetsvoorbereiding, kwam geen significante afwijking van het neutrale antwoord. Er kan dus niet geconcludeerd worden dat leerlingen meer tijd besteden aan hun toetsvoorbereiding als ze een spiekbriefje mogen gebruiken. De trend in de resultaten (figuur 4 op pagina 21) geeft aan dat leerlingen in ieder geval niet *minder* tijd in hun toetsvoorbereiding investeren.

Dat leerlingen niet meer of minder tijd besteden, betekende echter niet dat de toetsvoorbereiding verliep zoals bij andere toetsen. Zo rapporteerden leerlingen wel een significant hogere motivatie, keken leerlingen meer materiaal opnieuw door en hadden leerlingen het idee dat ze minder snel toetsstof oversloegen. Hieruit kan geconcludeerd worden dat leerlingen weliswaar niet meer tijd aan hun toetsvoorbereiding besteden, maar deze wel (naar hun idee) met meer motivatie en structuur aanpakken. Bovendien hebben de spiekbriefjes geen negatief effect op de lessen en het huiswerk voorafgaand aan de toets, zoals is gebleken uit sectie 4.2.

5.2.3 Correlatie inhoud en vraagscores

De uitgebreide analyses in sectie 4.3 zijn uitgevoerd om de laatste deelvraag te beantwoorden. Deze deelvraag zocht naar de correlatie “tussen de inhoud van een spiekbriefje en de behaalde score op toetsvragen bij informaticoetsen op het VO”. In sectie 4.3 zijn de resultaten van twee analyses beschreven: per toetsvraag, en per leerling. Om een antwoord op deze vraag te formuleren zijn de inzichten uit de twee analyses gecombineerd. De antwoorden uit de vragenlijst zijn hierbij slechts ter bevestiging gebruikt.

Analyse per toetsvraag Uit de analyse per toetsvraag bleek dat er bij de overgrote meerderheid van de toetsvragen geen duidelijk verband was. In totaal zijn 18 toetsvragen onderzocht. Bij drie vragen presteerden leerlingen met info op hun spiekbriefje significant beter. Bij enkele andere vragen was dezelfde trend zichtbaar, al was er bij één vraag ook een sterk omgekeerde trend, waarbij leerlingen met info slechter scoorden. Al met al is er hooguit sprake van een lichte correlatie, als daar al van gesproken kan worden.

Analyse per leerling Het onderzoek op leerlingniveau schetst een vergelijkbaar beeld. De gecombineerde data in figuur 10 laat een trendlijn zien die (nipt) boven het nulniveau blijft. Dit betekent dat leerlingen licht hogere scores als relevante info op het spiekbriefje staat, vergeleken met vragen zonder zulke info op het spiekbriefje. Tegelijkertijd is de spreiding hierin zeer groot. De trendlijn daalt licht, wat de indruk geeft dat bij een zwakkere leerling de correlatie sterker is dan bij een leerling die hogere cijfers haalt. De daling van deze lijn is echter zo zwak, dat men voorzichtig moet zijn over effecten van dit verband, als het verband überhaupt al bestaat. De enorme verschillen tussen de klassen (zie figuur 9) maken dat goed duidelijk.

Beantwoording vragenlijst Bovenstaand beeld werd al gelijk na het maken van de toets bevestigd in de vragenlijst die de leerlingen hebben ingevuld. Uit de antwoorden over de invloed van de spiekbriefjes op de behaalde scores kwam geen eenduidig beeld. De stelling over gebruik van het spiekbriefje tijdens de toets kreeg een neutraal antwoord, evenals de stelling over de invloed op het cijfer. De leerlingen bevestigen hiermee het beeld dat de analyses ook schetsen: er is hooguit een zeer lichte relatie tussen de inhoud van het spiekbriefje en de behaalde vraagscores.

5.3 Discussie

Hoewel de methode van het onderzoek de lijn volgde van andere onderzoeken naar spiekbriefjes op o.a. universitair niveau, moeten enkele beperkingen in acht genomen worden. Bovendien speelden enkele praktische problemen mee bij de uitvoering van het onderzoek. Voordat hiernaar gekeken kan worden, is het echter eerst goed om de resultaten in perspectief te plaatsen.

5.3.1 Verklaring van de resultaten

Op het eerste oog lijken de resultaten misschien vreemd. Waarom zou een spiekbriefje amper helpen om een hoger toetscijfer te behalen? Om dit te begrijpen, is het goed om een blik te werpen op een paar voorbeelden van spiekbriefjes.

Figuur 12 bevat vier voorbeelden van spiekbriefjes. Alle spiekbriefjes zijn afkomstig uit de V5-klas, waar deze toets als schoolexamen meetelde. De inhoud en opbouw van de spiekbriefjes verschillen sterk, maar ook de door leerlingen behaalde toetscijfers. Zo bevat het spiekbriefje van de leerling met het hoogste cijfer (figuur 12a) niet noodzakelijk de meeste informatie, de netste figuren, de best abstract gemaakte begrippen of een optimale indeling. Het meest treffende voorbeeld hiervan was het spiekbriefje in figuur 12d, waar alleen een paar begrippen waren opgeschreven die door de leerling (vermoedelijk) soms door elkaar werden gehaald ⁴.

De variatie in spiekbriefjes kan een verklaring zijn voor de sterk wisselende resultaten bij de vraagscore-analyses. Een spiekbriefje waar amper iets op staat is waarschijnlijk niet erg behulpzaam tijdens de toets. Dit roept echter weer andere vragen op. Waarom wisselden de spiekbriefjes zo sterk, als ze over hetzelfde onderwerp gingen? En wat is de invloed van de toets zelf op de effectiviteit van het spiekbriefje? In het vervolg van deze sectie wordt op deze en andere vragen ingegaan om zo een verklaring te zoeken voor de sterk wisselende spiekbriefjes en de zeer lichte correlatie tussen inhoud en vraagscores.

Vaardigheden maken spiekbriefje Het zou kunnen zijn dat sommige leerlingen geen beter spiekbriefje *kunnen* maken. Het maken van een spiekbriefje vereist immers vaardigheden in samenvatten en overzicht creëren, en dat is iets wat niet iedere leerling goed beheerst. Anderzijds zou je dan verwachten dat leerlingen tenminste *iets* opschrijven, wat dan misschien niet zou helpen. De lege ruimte op de spiekbriefjes in figuren 12c en 12d laat echter zien dat leerlingen niet heel selectief hoefden te zijn. Bovendien was dit al de vierde toets waarbij leerlingen een spiekbriefje mochten gebruiken. Leerlingen hadden dus al ervaringen uit vorige toetsen, en hadden dus al een goed idee wat wel en niet helpt tijdens een informaticatoets. Mogelijk speelden gebrekkige vaardigheden dus een rol, maar het verklaart waarschijnlijk niet de gigantische verschillen tussen leerlingen zoals te zien in figuur 10.

Persoonlijke voorkeuren Een andere verklaring voor de variatie in spiekbriefjes is dat leerlingen hun spiekbriefje simpelweg op hun eigen manier *willen*. De ene leerling begrijpt de stof beter dan de ander, en voelt daardoor wellicht geen noodzaak om het briefje goed te vullen. Het maken van een spiekbriefje kost immers veel tijd, en die tijd kan ook gebruikt worden voor andere toetsen in de toetsweek. Leerlingen stellen in een toetsweek prioriteiten, en die liggen niet altijd bij informatica.

⁴Enkele leerlingen, waaronder deze, hadden ook het alfabet op hun spiekbriefjes geschreven. Dit had te maken met een Caesar Cipher-vraag op de toets, waarvoor leerlingen een geschreven alfabet fijn vinden werken. Aangezien ze toch ruimte over hadden op hun spiekbriefje, scheelde dit weer tijd op de toets zelf.

Handwritten notes on a grid background discussing network protocols and security. Includes a list of letters and numbers at the top, followed by sections on ARP, DNS, DHCP, AP, and HTTPS. A diagram shows the flow of data between a client and a server through a proxy, involving public and private keys and certificates. The notes conclude with the phrase "Positive Energy It's all GREEN".

(a) Toetscijfer: 9.2

Handwritten notes on a grid background titled "Symmetrische Encriptie". It discusses symmetric encryption methods like XOR, Caesar, and Vigenere, and compares them to asymmetric encryption. It also covers digital signatures, hybrid encryption, and protocols like TLS and SSH. A list of letters and numbers is at the top. The notes end with a list of protocols: SSL/TLS, IPsec, and S/MIME.

(b) Toetscijfer: 8.4

Handwritten notes on a grid background titled "Spreekbriefje informatie Pa". It covers various network and security topics such as DNS, DHCP, ARP, NAT, and firewalls. Includes diagrams for NAT and DHCP processes. The notes are organized into several columns with sub-headers.

(c) Toetscijfer: 6.0

Handwritten notes on a grid background discussing network security and protocols. It includes a list of letters and numbers at the top, followed by text about "Beveiliging" and "Integriteit". A list of protocols is at the bottom: SSL/TLS, IPsec, and S/MIME. The notes also mention "Beveiliging" and "Integriteit" in a separate box.

(d) Toetscijfer: 8.7

Het hoeft echter geen kwestie van tijd te zijn: de leerlingen die informatica makkelijker vinden willen er misschien wel tijd in steken, maar weten misschien al dat ze die informatie toch al paraat hebben. Een voorbeeldantwoord van de oefentoets is misschien fijn voor sommigen, maar ook als de vraag toen toch al gelijk lukte? Deze gedachte wordt versterkt door de dalende trendlijn in figuur 10, die er (licht) op hintte dat sterke leerlingen minder profiteren van informatie op hun spiekbriefje.

Resultaten literatuur De (beperkte) literatuur over de invloed van spiekbriefjes op vraag-scores bij universiteiten schetst een beeld dat vergelijkbaar is met de resultaten uit dit onderzoek. De enige studie die onderzoek naar deed naar invloed op vraag-scores, Hamouda en Shaffer (2016), vond een zwak positief effect op de scores bij een klein aantal toetsvragen als relevante info op het spiekbriefje stond. Zij gaven echter geen verklaring voor de zwakke correlatie. Op basis van de discussie hierboven is het vermoeden dat de meeste leerlingen simpelweg weinig profijt kunnen halen uit een spiekbriefje, maar verder onderzoek is hiervoor nodig.

Invloed vraagniveau's Het "klein aantal toetsvragen" waar Hamouda en Shaffer wel een effect vonden, waren op het "comprehension"-niveau van de taxonomie van Bloom. Dit is het tweede niveau, net boven de pure reproductie. Volgens Kwakernaak (2013) valt dit niveau tussen de RTTI-niveau's R en T1 in. Bij geen van de R-vragen en bij slechts een klein deel van de T1-vragen zijn echter significante verschillen in scores gezien bij de analyse. Op dit punt verschillen de uitkomsten van dit onderzoek dus wél van de bestaande literatuur. De reden hiervoor is niet helemaal duidelijk. Mogelijk heeft dit bijvoorbeeld te maken met de verschillen tussen het VO en universitair onderwijs. In sectie 5.3.4 wordt verder ingegaan op toekomstig onderzoek om dit verschil te verklaren.

Stress en motivatie Met de informatie uit de vragenlijst zijn ook conclusies getrokken over o.a. de stress en motivatie rondom de toetsvoorbereiding. Deze conclusies komen wel overeen met de observaties uit de literatuur, zoals o.a. Hamouda en Shaffer (2016), Erbe (2007) en Larwin e.a. (2013).

5.3.2 Bedreigingen validiteit

Alhoewel de resultaten grotendeels overeenkomen met de literatuur, moeten wel enkele kanttekeningen gemaakt worden bij de opzet en uitvoering van het onderzoek. Hieronder worden de kanttekeningen over de opzet besproken, die mogelijk een effect hebben op de validiteit van het onderzoek.

Zelfrapportage stress In de vragenlijsten werd aan leerlingen gevraagd om hun indruk te geven over de impact op hun stresservaring. Dit was ook precies de stelling in de onderzoeksvraag, omdat dit in de praktijk als enige gemeten kon worden. Hierbij rijst echter wel de vraag in hoeverre de stresservaring overeenkomt met de daadwerkelijk

meetbare stress.

Uit onderzoek blijkt dat zelfrapportage van stress een redelijke correlatie (0.39) heeft met de stress zoals gemeten aan de hand van o.a. hartslag (Masood, Ahmed, Choi & Gutierrez-Osuna, 2012). Er is de leerlingen echter wel gevraagd om de stresservaring te vergelijken met andere toetsen zonder spiekbriefje. Deze andere toetsen waren minstens enkele uren eerder en bovendien van andere vakken. Stressrapportage over gebeurtenissen in het verleden is minder accuraat dan voor recente gebeurtenissen (Masood e.a., 2012). Ook zou het kunnen dat leerlingen slecht de impact van het spiekbriefje los kunnen zien van de stress voor het vak informatica, aangezien ze altijd het spiekbriefje hebben mogen gebruiken.

Al met al kan de zelfrapportage over stress zeer behulpzaam zijn om het verschil in stress te meten. Idealiter zou hierbij echter een meer objectieve meting gebruikt moeten worden dan een zelfrapportage.

Zelfrapportage toetsvoorbereiding Vergelijkbare zorgen spelen bij de zelfrapportage over de toetsvoorbereiding. In hoeverre kunnen leerlingen accuraat inschatten of ze minder toetsstof hebben overgeslagen? Het kan immers zo zijn dat leerlingen niet weten wat ze hebben gemist, en dit bewustzijn kan per vak verschillen. Het is lastig hier exacte cijfers voor te geven, zeker omdat leerlingen vergelijken tussen verschillende vakken en typen toetsen.

Het is de inschatting dat leerlingen beter zijn in het rapporteren over de hoeveelheid materiaal dat ze wél hebben doorgekeken, dan wat ze hebben overgeslagen. Op deze vraag kwamen vanuit de leerlingen grotendeels dezelfde antwoorden als op de overslaan-vraag. De conclusie over een meer gestructureerde toetsvoorbereiding wordt verantwoord geacht doordat de antwoorden op beide vragen hetzelfde beeld schetsen. Het zou echter goed zijn om hier meer objectieve informatie over te verzamelen, om zo de zelfrapportage in de vragenlijst te verifiëren.

Aspecten toetsvoorbereiding Bij het onderzoek naar de toetsvoorbereiding zijn vooraf enkele aspecten geselecteerd (zoals de motivatie), waarop leerlingen konden aangeven in hoeverre ze het hiermee eens waren. Deze aspecten zijn geselecteerd uit de resultaten van andere onderzoeken bij universiteiten. Het kan echter zijn dat spiekbriefjes invloed hebben op de toetsvoorbereiding op manieren die niet in die onderzoeken, en daarom ook dit onderzoek, naar voren zijn gekomen. De conclusies over de toetsvoorbereiding zijn dus zeker niet onjuist, maar mogelijk wel incompleet.

5.3.3 Bedreigingen betrouwbaarheid

Verwerking spiekbriefjes Bij de analyse is herhaaldelijk een inschatting gemaakt of de inhoud op het spiekbriefje mogelijk relevant is voor de desbetreffende vraag. Hierbij komt een zekere mate van subjectiviteit kijken. Wat de een inschat als waardeloos, kan immers voor de ander juist als relevant worden gezien.

Bij dit onderzoek zijn alle spiekbriefjes door dezelfde persoon verwerkt. De criteria voor

“bevat wel info” waren hier ruim genomen. Zo gold ook het noemen van een begrip, zonder een echte definitie, al als relevante info. Het kan zijn dat een andere onderzoeker, met een strengere definitie, hierdoor andere resultaten vindt.

Bij wijze van proef is voor enkele vragen de analyse opnieuw uitgevoerd, waarbij twijfelgevallen (zoals een begrip zonder definitie) zijn gemarkeerd als “bevat geen info”. Dit bracht bij deze steekproef geen significant verschillende resultaten. Hierom is het niet de verwachting dat resultaten sterk zouden afwijken bij andere onderzoekers of beoordelaars, maar de mogelijkheid kan niet worden uitgesloten.

Berekening scoreverschillen Om het verschil in scoregemiddelde te berekenen zijn eerst de gemiddelde scores berekend bij vragen mét info op het spiekbriefje, en zonder info. Deze beide gemiddelden werden echter berekend over een totaal van 6 vragen. In sommige gevallen is hierdoor een “gemiddelde” ontstaan uit één enkele vraag, als het spiekbriefje bijvoorbeeld zeer weinig info bevatte. Mogelijk zou een nieuw onderzoek bij een andere toets hierdoor bij dezelfde leerling een totaal ander beeld geven.

De analyse per leerling is uitgevoerd om twee redenen. Enerzijds konden hiermee de conclusies uit het onderzoek per toetsvraag in een ander licht worden gebracht. Anderzijds konden hiermee de conclusies uit Hamouda en Shaffer (2016), die het verband met zwakkere studenten eerder onderzocht, worden onderzocht voor middelbare scholieren. Bij het onderzoek van Hamouda en Shaffer waren echter veel meer studenten, waardoor daar niet rekenkundig vragen gecombineerd hoefden te worden. De aanpak van Hamouda en Shaffer is ongetwijfeld nauwkeuriger dan de benadering met scoregemiddelden. Deze aanpak geniet dan ook de voorkeur wanneer de hoeveelheid leerlingen deze vorm van analyse mogelijk maakt.

Houding leerlingen De H4-klas was de enige klas waarbij de analyse op klasniveau significante scoreverschillen liet zien tussen leerlingen met en leerlingen zonder relevante info op het spiekbriefje. Dit kan met vele factoren te maken hebben, waaronder de toetsstof. Een mogelijke factor die niet onbesproken moet blijven, is de houding van een groot aantal leerlingen. Tegen het einde van de periode wist een kwart van de leerlingen al dat ze het volgende jaar op een andere school verder zouden gaan. Ook was het gemiddelde van sommige leerlingen al dusdanig hoog, dat hun overgang niet in gevaar kwam. Aan de toetsen was te zien dat iedereen weliswaar hun best had gedaan in de examenruimte, maar dat sommigen geen enkele tijd in hun voorbereiding gestopt hadden. Zo hadden bepaalde leerlingen exclusief punten behaald op vragen die zij verplicht geoefend hadden in de les, en geen punten elders. Mogelijk heeft dit de vraagscores, en daarmee de resultaten van de analyse, vertekend. Bij andere klassen speelde dit probleem niet of nauwelijks.

Opzet toetsen Informatica op een specifieke school kan flink verschillen van andere scholen in Nederland. Waar de ene school bijvoorbeeld robotica centraal stelt, kan een andere school juist kiezen voor een theoretische en algoritmische aanpak op papier.

Ook de kerndomeinen, die op iedere school behandeld moeten worden, kunnen op verschillende manieren getoetst worden. Op de school van dit onderzoek is meer gekozen voor een theoretische benadering op basis van de methode Fundament. Een voorbeeldtoets, vergelijkbaar met de V4-toets van dit onderzoek, is opgenomen in appendix C. In deze oefentoets is ook te zien dat sommige vragen ook leesvaardigheid van leerlingen vereisen, wat niet iedere leerling even goed beheerst en zo invloed kan hebben op de observaties over de spiekbriefjes. Wie de vraag immers niet goed begrijpt, heeft weinig aan relevante info op het spiekbriefje. Het is dus mogelijk dat, vanwege de wijze van behandeling en toetsing, spiekbriefjes een andere impact hebben op andere scholen. Replicatie-onderzoek op andere scholen is nodig om dit nader te bepalen.

5.3.4 Toekomstige onderzoeksrichtingen

De resultaten en bovenstaande discussiepunten bieden een hoop openingen voor toekomstig onderzoek. Enkelen zijn hierboven al kort aangestipt. Een aantal richtingen zijn hieronder kort toegelicht.

Toetsopzetten en klassen Het is interessant om te zien of de resultaten van dit onderzoek zich herhalen bij andere lesstijlen, contexten voor onderwerpen en toetsopzetten. Wat is bijvoorbeeld de impact bij een schriftelijke toets voor domein F, Interactie? Of bij een programmeertoets in een robotica-context? Er is veel ruimte voor variatie in toetsing en contexten, en mogelijk leiden deze variaties ook tot andere uitkomsten. Hetzelfde geldt voor het type vragen dat gesteld wordt op een toets. Wat is de precieze impact bij scholieren als er meer reproductievragen worden opgenomen? Is er dan een sterkere correlatie tussen spiekbriefjesinhoud en vraagscore?

Interviews toetsvoorbereidingen Tijdens dit onderzoek is voornamelijk kwantitatieve data verzameld. Het is echter interessant om ook meer kwalitatieve data te verzamelen over de keuzes achter het spiekbriefje. Mogelijk kan hiermee ook een verklaring worden gevonden voor de wisselende effectiviteit van het spiekbriefje, zoals besproken in sectie 5.3.1. Waarom is bepaalde informatie wel of niet opgenomen? Hoe kan het dat sommige leerlingen gemiddeld lager scoren als er juist wel info over een vraag op hun spiekbriefje staat? Een interview met een aantal leerlingen zou bijvoorbeeld een goede manier zijn om de keuzes achter en de impact van spiekbriefjes verder te onderzoeken.

Impact op toetsscores In dit onderzoek is alleen de correlatie tussen spiekbriefjesinhoud en de score op toetsvragen onderzocht. Er bleek amper sprake van correlatie, maar dit hoeft niet te betekenen dat de spiekbriefjes geen impact hebben op de gehele toetsscores. Zo is het goed mogelijk dat de meer gestructureerde voorbereiding een algemeen verhogend effect heeft op de toetsscores.

In de praktijk is het lastig om de precieze impact van de aanwezigheid van spiekbriefjes vast te stellen. Onderzoeken die deze impact bestudeerd hebben, zoals de Raadt (2012),

trokken vaak hun conclusies op basis van studenten die vrijwillig geen spiekbriefje meebrachten. De waargenomen verschillen in scores bij deze onderzoeken kunnen echter ook veroorzaakt zijn door andere factoren die correleren met het meebrengen van een spiekbriefje, zoals de motivatie van een student. Een slechte score kan dan wellicht veroorzaakt zijn door het gebrek aan een spiekbriefje, maar ook door gebrekkige inzet bij (werk)colleges.

In het werk van o.a. Dickson en Bauer (2008) is getracht de impact van de *aanwezigheid* van het spiekbriefje los te onderzoeken van het effect op de *voorbereiding*. Dit is toen gedaan middels een onverwachtse extra toets, waar het voorbereide spiekbriefje niet gebruikt mocht worden. Een mogelijk beperkende factor van dit onderzoek was dat de extra toets niet meetelde in de berekening van het eindcijfer voor het vak, wat mogelijk de motivatie verminderde. Tevens kan met deze opzet nog steeds niet de impact van de veranderde voorbereiding gemeten worden.

Het meten van de exacte impact op vraag- en toetsscores blijft een uitdaging. Kern van het probleem blijft dat leerlingen in dezelfde jaarlaag op een school onder gelijke omstandigheden hun toetsing dienen te maken, wat een A/B-test lastiger maakt.

Onderzoekopzet impact op scores Desondanks zijn er waarschijnlijk wel mogelijkheden om dit goed te onderzoeken. Mogelijk kan hierbij gebruikgemaakt worden van "stapeldomeinen". Zo bouwt keuzedomein G (algoritmiek) voort op domein B (grondslagen). Voor het onderzoek zou in het ene jaar of op de ene school wel een spiekbriefje gebruikt mogen worden bij de toets van domein G, maar in een ander jaar of een andere klas niet.

Om de impact van het spiekbriefje te bepalen, zou vervolgens gekeken kunnen worden naar de verschillen ten opzichte van de "basistoets" (domein B), die wel onder gelijke omstandigheden in beide groepen is afgenomen. Wanneer leerlingen bijvoorbeeld gemiddeld 1 punt lager scoren op de tweede toets zonder spiekbriefje en "slechts" 0.5 punt lager wanneer er wel een spiekbriefje gebruikt mag worden, kan het verschil alleen toegeschreven worden aan het spiekbriefje. De vaardigheid van de leerlingen in het desbetreffende onderwerp zijn hiermee immers genormaliseerd. Dit is ook waarom in deze opzet voor een stapeldomein is gekozen in plaats van een vergelijking met het gemiddelde cijfer. Het gemiddelde kan immers beïnvloed zijn door aanleg voor andere onderwerpen of andere vormen van toetsing, zoals PO's.

Referenties

- Bergervoet, P., Bruggink, M., Gijssen, A., Gnodde, J., Hendriks, C., Hoekstra, J. & Weeda, R. (2021). *Toetsing domein B: Grondslagen*. SLO. Amersfoort.
- Bloom, B. S., Engelhart, M. B., Furst, E. J., Hill, W. H. & Krathwohl, D. R. (1956). *Taxonomy of educational objectives. The classification of educational goals*. New York: Longmans Green.
- Chambers, J. (2010). *Software for data analysis*. doi:[10.1007/978-0-387-75936-4](https://doi.org/10.1007/978-0-387-75936-4)
- Daly, C. & Waldron, J. (2004). Assessing the Assessment of Programming Ability. In *Proceedings of the 35th SIGCSE Technical Symposium on Computer Science Education* (pp. 210–213). doi:[10.1145/971300.971375](https://doi.org/10.1145/971300.971375)
- de Raadt, M. (2012). Student Created Cheat-Sheets in Examinations: Impact on Student Outcomes. In *Proceedings of the Fourteenth Australasian Computing Education Conference - Volume 123* (pp. 71–76). Melbourne, Australia: Australian Computer Society, Inc.
- Dickson, K. L. & Bauer, J. J. (2008). Do Students Learn Course Material During Crib Sheet Construction? *Teaching of Psychology*, 35(2), 117–120. doi:[10.1080/00986280801978343](https://doi.org/10.1080/00986280801978343)
- Drost, M. & Verra, P. (2012). *Sturen van leerprocessen met RTTI - Handboek RTTI*.
- Erbe, B. (2007). Reducing Test Anxiety While Increasing Learning: The Cheat Sheet. *College Teaching*, 55(3), 96–98. doi:[10.3200/CTCH.55.3.96-98](https://doi.org/10.3200/CTCH.55.3.96-98)
- Funk, S. C. & Dickson, K. L. (2011). Crib Card Use During Tests: Helpful or a Crutch? *Teaching of Psychology*, 38(2), 114–117. doi:[10.1177/0098628311401584](https://doi.org/10.1177/0098628311401584)
- Goldebeld, P. (1992). *TOETS- EN ITEMANALYSE MET TIA: Toelichting bij het lezen en interpreteren van toetsen itemanalyses voor gesloten en/of open vragen*. Cito. Arnhem.
- Hamouda, S. & Shaffer, C. A. (2016). Crib sheets and exam performance in a data structures course. *Computer Science Education*, 26(1), 1–26. doi:[10.1080/08993408.2016.1140427](https://doi.org/10.1080/08993408.2016.1140427)
- Kwakernaak, E. (2013). RTTI, OBIT, Bloom en het vreemdetalenonderwijs. *Levende talen magazine*, 100(5), 10–16.
- Larwin, K. H., Gorman, J. & Larwin, D. A. (2013). Assessing the Impact of Testing Aids on Post-Secondary Student Performance: A Meta-Analytic Investigation. *Educational Psychology Review*, 25(3), 429–443.
- Mangen, A., Anda, L. G., Oxborough, G. H. & Brñnick, K. (2015). Handwriting versus keyboard writing: effect on word recall. *Journal of writing research*, 7(2), 227–247.
- Masood, K., Ahmed, B., Choi, J. & Gutierrez-Osuna, R. (2012). Consistency and validity of self-reporting scores in stress measurement surveys. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 4895–4898). IEEE.

- Mesie, M. (2013). Is dit een vmbo-, een havo- of een vwo-leerling? OBIT: instrument om voorspellende waarde van toetsen te verhogen. *Levende Talen Magazine*, 100(3), 10–14.
- Moelands, H. (2017). *Toetsconstructie in 8 stappen: Een handleiding voor toetsconstructeurs met informatie over de acht stappen van het toetsconstructieproces*. Cito. Arnhem.
- Schrum, M. L., Johnson, M., Ghuy, M. & Gombolay, M. C. (2020). Four Years in Review: Statistical Practices of Likert Scales in Human-Robot Interaction Studies. In *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 43–52). doi:[10.1145/3371382.3380739](https://doi.org/10.1145/3371382.3380739)
- Shibata, H. & Omura, K. (2018). Reconsideration of the Effects of Handwriting. *ITE Transactions on Media Technology and Applications*, 6(4), 255–261. doi:[10.3169/mta.6.255](https://doi.org/10.3169/mta.6.255)
- Song, Y., Guo, Y. & Thuente, D. (2016). A quantitative case study on students' strategy for using authorized cheat-sheets. In *2016 IEEE Frontiers in Education Conference (FIE)* (pp. 1–9). doi:[10.1109/FIE.2016.7757656](https://doi.org/10.1109/FIE.2016.7757656)
- Trigwell, K. (1987). THE CRIB CARD EXAMINATION SYSTEM. *Assessment & Evaluation in Higher Education*, 12(1), 56–65. doi:[10.1080/0260293870120106](https://doi.org/10.1080/0260293870120106)
- Van der Donk, C. & Van Lanen, B. (2020). *Praktijkonderzoek in de school* (Vierde herziene druk). Bussum: Uitgeverij Coutinho,
- Visco, D., Swaminathan, S., Zagumny, L. & Anthony, H. (2007). Interpreting Student Constructed Study Guides. In *2007 Annual Conference & Exposition* (pp. 12.955.1–12.955.10). doi:[10.18260/1-2--1818](https://doi.org/10.18260/1-2--1818)
- Wachsman, Y. (2002). Should Cheat Sheets be Used as Study Aids in Economics Tests? *Economics Bulletin*, 1(1), 1–11.

A Vragenlijst

Naam: _____

	Helemaal niet mee eens	Niet mee eens	Niet mee eens en niet mee oneens	Mee eens	Helemaal mee eens
Door het spiekbriefje heb ik minder stress op de dag van de toets.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ik voel me meer op mijn gemak op de toetsdag als ik een spiekbriefje mag gebruiken.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Zonder het spiekbriefje had ik meer spanning voor de toets gevoeld.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Dankzij het spiekbriefje heb ik meer vertrouwen in een goed cijfer.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

De vragen hieronder gaan over je voorbereiding op de toets.

“Toetsvoorbereiding” betekent hier al het werk dat je specifiek voor de toets hebt gedaan, zoals het maken van de oefentoets, het teruglezen van slides en het maken van je spiekbriefje.

Vergelijk je toetsvoorbereiding met een situatie waarin geen spiekbriefje zou zijn voor hetzelfde vak.

Als ik een spiekbriefje mag gebruiken besteed ik meer tijd aan mijn toetsvoorbereiding.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Het spiekbriefje motiveert mij om voor een toets bezig te gaan.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Bij toetsen met spiekbriefje kijk ik meer materiaal (boek/slides/opgaven) opnieuw door.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Door het maken van een spiekbriefje sla ik minder snel toetsstof over bij het leren.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

De volgende vragen gaan over je gedrag tijdens *alle* lessen in de periode.

Ik maak minder huiswerk, omdat ik weet dat ik een spiekbriefje heb bij de toets.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ik let minder goed op in de lessen, omdat ik toch een spiekbriefje heb tijdens de toets.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

De laatste vragen gaan over het gebruik tijdens de toets zelf.

Tijdens de toets heb ik regelmatig op mijn spiekbriefje gekeken.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sommige vragen heb ik alleen kunnen beantwoorden door op het spiekbriefje te kijken.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Zonder spiekbriefje zou mijn cijfer lager zijn geweest.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Hoeveel uur heb je ongeveer besteed aan het maken van je spiekbriefje? _____

Hoeveel uur heb je ongeveer besteed aan je totale toetsvoorbereiding? _____

Bedankt voor het invullen!

B Toetsinformatie

In de tabellen op deze en de volgende pagina's zijn de complete statistieken over de afgenomen toetsen te vinden. De vragen waarbij de spiekbriefjes-analyse zijn uitgevoerd, zijn dikgedrukt weergegeven.

Bij sommige vragen is de moeilijkheidsgraad 1 en is de Rit-waarde "n.v.t." (niet van toepassing). Bij deze vragen hadden alle leerlingen het juiste antwoord. De Rit-waarde is dan niet te berekenen. Het omgekeerde, waarbij alle leerlingen een vraag fout hebben, is niet voorgekomen.

Vraag	Niveau	Moeilijkheidsgraad	Rit-waarde
1a	T1	0.84	0.38
1b	T1	0.71	0.43
1c	T1	0.82	0.80
1d	T2	0.21	0.31
2a	T1	0.29	0.37
2b	I	0.50	0.37
2c	T2	0.79	0.39
3a	I	0.38	0.34
3b	I	0.25	0.24
4	I	0.50	0.48
5a	T1	1.0	n.v.t.
5b	T1	0.50	0.66
5c	I	0.57	0.61
6	T2	0.35	0.74
7a	T1	0.50	0.60
7b	T1	0.79	0.67
7c	T1	0.64	0.85
7d	T1	0.50	0.70

Tabel 10: Statistieken over de afgenomen H4-toets. De geanalyseerde vragen zijn dikgedrukt.

Vraag	Niveau	Moeilijkheidsgraad	Rit-waarde
1a	T1	1.0	n.v.t.
1b	T1	0.80	0.40
1c	T1	0.97	0.28
1d	T1	0.70	0.29
1e	T2	0.67	0.78
1f	T2	0.53	0.73
1g	R	0.73	0.63
1h	T1	0.87	-0.02
1i	T2	0.83	0.45
1j	R	0.70	0.58
1k	I	0.50	0.56
2a	T1	0.87	0.24
2b	T2	0.80	0.03
2c	T2	0.67	0.48
2d	I	0.60	0.54
3a	T2	0.84	0.69
3b	R	1.0	n.v.t.
3c	T1	0.33	0.81

Tabel 11: Statistieken over de afgenomen V4-toets. De geanalyseerde vragen zijn dikgedrukt.

Vraag	Niveau	Moeilijkheidsgraad	Rit-waarde
1a	T1	0.67	0.15
1b	T1	0.81	0.58
1c	R	0.71	0.73
2a	T1	0.83	0.74
2b	T2	1.0	n.v.t.
2c	T2	0.96	-0.09
2d	I	0.75	0.40
3a	I	0.83	0.24
3b	I	0.75	0.30
3c	T1	0.61	0.54
3d	T1	0.79	0.76
3e	I	0.50	0.55
3f	R	0.58	-0.14
3g	T2	1.0	n.v.t.
3h	I	0.58	0.40
4a	T1	1.0	n.v.t.
4b	T2	0.58	0.40
4c	T2	0.58	0.02
4d	R	0.88	-0.08
4e	T2	0.56	0.20
4f	T1	0.96	0.37
5a	T2	0.83	0.72
5b	I	0.54	-0.01
5c	T1	0.71	0.57
5d	I	0.71	0.82

Tabel 12: Statistieken over de afgenomen V5-toets. De geanalyseerde vragen zijn dikgedrukt.

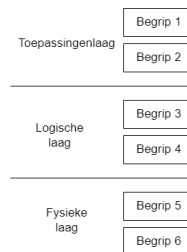
C Oefentoets

Oefentoets architectuur

juni 2022 (V4)

- Deze toets bestaat uit 5 vragen, verdeeld over 3 pagina's. In totaal zijn er 21 punten te behalen.
- Toegestane hulpmiddelen: normale (niet-grafische) rekenmachine en een zelf meegebracht handgeschreven enkelzijdig A4 met notities.
- De puntenaantallen op de echte toets kunnen nog verschillen van de aantallen op deze oefentoets. De soort vragen en moeilijkheidsgraad zijn wel zeer vergelijkbaar.

Vraag 1 Op de wifipunten van het merk Ubiquity draait een speciaal besturingssysteem. Het zogenaamde UnifiOS gebruikt intern Linux. Op UnifiOS draait een webserver om data uit het apparaat te halen. Dankzij deze server kan de gebruiker de status en belasting van het wifipunt zien.



Het wifipunt moet informatie bijhouden over alle verbonden apparaten. Als een wifipunt te veel apparaten helpt, kan dit ten koste gaan van de latency.

- 3p (a) Verdeel de volgende begrippen over de drie architectuurlagen: *webserver*, *UnifiOS*, *wifi-antenne*, *Linux*, *statuspagina*.
Geef je antwoord zoals in de figuur rechtsboven, zodat ook zichtbaar is waar iedere laag "staat" ten opzichte van de andere lagen. Let op dat het aantal begrippen per laag kan verschillen.
- 1p (b) Zoals bijna ieder systeem maken wifipunten gebruik van *drivers*. Leg uit binnen welke van de vier begrippen uit (a) drivers bevatten.
- 1p (c) Veel wifi-apparaten zijn wel verbonden, maar sturen amper data. Leg uit of deze "slapende" apparaten eisen stellen aan het RAM, de processor of de grafische chip van het wifipunt.
- 1p (d) Sommige leerlingen spelen spelletjes als Tetris tijdens de les. Leg uit of voor het spelen van Tetris (een singleplayer-game voor het plaatsen van blokken) een lage latency nodig is.
- Binnen school worden aan apparaten IP-adressen toegekend uit de zogeheten "private range".
- 1p (e) Leg uit of de kans groot is dat een ander apparaat op de wereld hetzelfde IP-adres heeft als jouw telefoon wanneer je op school bent.
- 1p (f) Leg uit waarom andere websites kunnen zien dat je vanaf de regio Meppel contact legt met hun site.

Pagina 1/3 Vervolg op volgende pagina. . .

Toets juni 2022 (V4)

Stad & Esch Lyceum Meppel

Vraag 2 In december 2015 werd een lek gevonden in de officiële fansite voor Hello Kitty, genaamd sanriotown.com. Hierdoor kwamen de gegevens van miljoenen fans op straat, samen met hun wachtwoorden. Deze wachtwoorden waren opgeslagen met moderne hashingtechnieken.

- 1p (a) Leg uit of hier sprake is van een inbreuk op de integriteit, de vertrouwelijkheid of de beschikbaarheid.

De webhostingssoftware maakt intern gebruik van een database met gegevens van gebruikers. Doordat deze onderliggende database niet goed afgeschermd was, bleek de data toegankelijk van buitenaf.

- 1p (b) Leg uit of dit een zwakheid in de architectuur, in de communicatie of bij de gebruikers betreft.

Normaal gesproken moet de webserver inloggen op de database met een gebruikersnaam en wachtwoord via een beveiligde verbinding. Een database kan meerdere gebruikers bevatten: sommigen mogen alleen data uitlezen, anderen kunnen data bewerken en toevoegen, weer anderen mogen data verwijderen. Ook moeten meerdere gebruikers tegelijkertijd bewerkingen kunnen uitvoeren op de database.

- $\frac{1}{2}$ p (c) Geef uit de tekst een voorbeeld van authenticatie rond het inlogproces op een database. Motiveer je antwoord.

- $\frac{1}{2}$ p (d) Geef uit de tekst een voorbeeld van autorisatie rond het inlogproces op een database. Motiveer je antwoord.

Vraag 3 Techwebsite Tweakers mailde in maart 2018 hun gebruikers omdat hun wachtwoord niet veilig zou zijn. Dit was echter niet door een datalek bij Tweakers zelf: de mensen achter deze site kwamen erachter dat de wachtwoorden van hun gebruikers op andere sites waren uitgelekt, en dat dezelfde wachtwoorden op Tweakers gebruikt werden.

- 1p (a) Leg uit of deze waarschuwing voortkomt uit een zwakheid in de architectuur, in de communicatie of bij gebruikers.

- 1p (b) Is de kans groot dat de websites waar de wachtwoorden uitlekten hashing gebruikten? Leg je antwoord uit.

Veel wachtwoorden lekken uit doordat ontwikkelaars hun database versleutelen via symmetrische encryptie. Hiermee claimen websitebouwers dat hun wachtwoorden veilig zijn opgeslagen. Hiermee wordt bijvoorbeeld voorkomen dat bij een inbraak in de serverruimte data uitlekt. Het kan echter niet alle aanvallen voorkomen.

- 1p (c) Klaarblijkelijk wordt bij het stelen van schijven een zwakheid in de architectuur misbruikt. Leg uit of deze zwakheid ligt bij de fysieke laag, de logische laag of de toepassingslaag.

- 1p (d) Leg kort uit hoe het gebruik van symmetrische encryptie misbruik van de zwakheid uit (c) voorkomt.

Een van de manieren om je account toch te beveiligen, is gebruik van authenticatie via meerdere *factoren*.

Pagina 2/3 Vervolg op volgende pagina. . .

Toets juni 2022 (V4)

Stad & Esch Lyceum Meppel

- 1p (e) Het wachtwoord is een manier om het account te beveiligen met iets wat je *weet*. Noem een manier om je account extra veilig te maken met iets wat je *hebt*.
- 1p (f) Leg uit of gebruik van 2-factorauthenticatie een vorm is van *preventie*, *detectie*, *repressie* of *correctie*.
- 2p **Vraag 4** Sommige bedrijven draaien “verdachte” software in een virtuele machine. Het idee hiervan is dat de schade dan beperkt is als de software inderdaad kwaadaardig was.
Leg kort uit wat een virtuele machine is, en waarom dit de schade zou beperken.
- Vraag 5** Geef voor elk van de volgende stellingen aan of deze waar of niet waar zijn. Leg je antwoord uit!
- 1p (a) Om een bericht te sturen naar Eve met asymmetrische encryptie gebruiken Charlie en Dave hun eigen public keys.
- 1p (b) Bij communicatie tussen een laptop en een webserver, is vanaf de webserver het MAC-adres van de bestemming het adres van de laptop.