

SELF- & IPSATIVE ASSESSMENT EN DE RELATIE MET TESTANGST: EEN INTERVIEWSTUDIE BIJ LEERLINGEN EN HUN LEERKRACHTEN IN HET VLAAMS SECUNDAIR ONDERWIJS

Aantal woorden: 17 998

Nona Van De Maele

Studentennummer: 01804679

Promotor: Prof. dr. Bram De Wever

Masterproef voorgelegd voor het behalen van de graad master in de pedagogische wetenschappen, afstudeerrichting Pedagogiek en onderwijskunde

Academiejaar: 2022 – 2023

I. Woord vooraf

“Everybody is a genius. But if you judge a fish by its ability to climb a tree, it will live its whole life believing that it is stupid.”

Deze quote van Albert Einstein wakkerde begin 2018 het vuur in mij aan om mij te verdiepen in de onderwijskundige wereld. Het bleek een ontdekkingstocht met veel onvolmaaktheden, maar met des te meer kansen en lichtpunten. Het rechtvaardigheidsgevoel in mij wanneer onvoldoende gestaafde conclusies werden getrokken op basis van louter een minder resultaat, werd aanzienlijk gevoed tijdens de eerste lessen van deze opleiding. Ik herinner me mijn gevoel nog goed wanneer Prof. Dr. Martin Valcke voor het eerst het concept ‘*ipsative assessment*’ toelichtte. Nu, vijf jaar kennisverwerving en twee jaar schrijven later, kan ik deze mooie vicieuze cirkel afsluiten aan de hand van deze masterproef. ‘Thesissen’ is een heel karwei, maar dankzij de adequate ondersteuning en uiterst waardevolle feedback van mijn promotor Prof. Dr. Bram De Wever heb ik een tastbaar aandenken verworven aan deze leerrijke periode. Ook de gesprekken met de leerkrachten en de interacties met de leerlingen hebben mij doen inzien hoe belangrijk en boeiend het onderwijs is – een aantal interviews zijn namelijk getransformeerd van een sec bevraging naar een authentiek gesprek. Ik wil al mijn participanten dan ook bedanken voor het mee helpen nadenken, het willen meewerken en het bieden van fantastische inzichten. Bedankt aan iedereen rondom mij om mij te allen tijde gesteund te laten voelen.

Nona Van De Maele,

Mei, 2023

II. Abstract

Deze masterproef onderzoekt wat de percepties zijn van leerkrachten en leerlingen in het Vlaams secundair onderwijs met betrekking tot testangst en het gebruik van alternatieve toetsvormen. Dat gebruik is de laatste jaren namelijk enorm toegenomen alsook de rapportering van een dalend mentaal welzijn onder de schoolgaande jeugd. De relatie tussen zowel de testangst van leerlingen als formatieve toetsing in het algemeen als *self-* en *ipsative assessment* specifiek, wordt duidelijk door de stelling dat het gebruik van dergelijke toetsvormen de testangst van leerlingen zou opvangen. Om de ervaringen van de stakeholders in kaart te brengen en ze een stem te geven in het debat, worden 10 leerkrachten en 89 leerlingen verspreid over 2 secundaire scholen bevroegd aan de hand van semi-gestructureerde interviews en vragenlijsten met open vragen. Uit de thematische analyse van de resultaten blijkt dat vooral de jongere leerkrachten door tijdsgebrek een ethisch-pedagogisch dilemma ondervinden in het kiezen van de meest adequate toetsvorm alsook in het geven van feedback. Daarnaast wordt er een miscommunicatie gerapporteerd tussen leerlingen en leerkrachten met betrekking tot *self-assessment*, namelijk dat enkele leerlingen dit bewust foutief invullen omdat ze van mening zijn dat dit een effect zou hebben op hun resultaten. Wanneer de leerkrachten ten slotte aanhalen wat ondernomen wordt – op zowel micro- als mesoniveau – om tegemoet te komen aan de psychologische noden van de leerlingen, wordt onder andere het doorverwijzen naar de zorgcoördinatoren aangehaald. Echter blijkt dat deze niet steeds een vertrouwenspersoon zijn voor de leerlingen, wat het veilig leerklimaat kan belemmeren.

Toelichting van aanpak en eigen inbreng

De structuur van de inleiding alsook het eerste luik van het theoretisch kader werden vaak hervormd dankzij de waardevolle inzichten en feedback van mijn promotor, professor Bram De Wever. Op basis van hiaten in de literatuur zijn de hoofd- en deelonderzoeksvragen tot stand gekomen. Bij de formulering van specifiek de hoofdonderzoeksvraag én bij het opstellen van de interviewleidraad kon ik eveneens op veel input rekenen van mijn promotor. De dataverzamelmethode heb ik volledig zelf gekozen alsook de techniek waarop de data-analyse zou uitgevoerd worden. Ik heb de relevante thema's uit de transcripties namelijk gecodeerd en geïnterpreteerd via Nvivo. Aangezien ik in dit onderdeel relatief vrij gelaten ben, heb ik de hulp van een medestudent ingeroepen om na te gaan of die de codeboom op dezelfde manier zou opgesteld hebben. Zelf heb ik één dataset opnieuw gecodeerd. Ten slotte wens ik mijn eerste participant te bedanken, aangezien die zelf voorstelde een uur van haar les voor enkele klassen vrij te houden zodat ik dan gemakkelijk de vragenlijsten kon afnemen.

III. Inhoud

I.	Woord vooraf	3
II.	Abstract.....	4
III.	Inhoud	5
1.	Inleiding.....	7
2.	Theoretisch kader	9
2.1.	Veranderende visie omtrent onderwijs.....	9
2.2.	Vormen van toetsing	11
2.2.1.	Formatieve en summatieve toetsing	11
2.2.2.	Ipsative assessment.....	12
2.2.3.	Self-assessment.....	14
2.3.	Psychologisch en academisch welbevinden bij leerlingen.....	16
2.3.1.	Testangst	16
3.	Onderzoeksdesign	19
3.1.	Probleemstelling	19
3.2.	Onderzoeksofzet.....	19
3.3.	Respondenten.....	21
3.4.	Betrouwbaarheid en validiteit.....	21
4.	Resultaten.....	23
4.1.	Wat zijn de ervaringen van leerlingen met de door hun leerkrachten gegeven toetsen en feedback?.....	23
4.2.	Wat zijn de ervaringen van leerkrachten met de door hun gegeven toetsen en feedback?.....	26
4.3.	Wat zijn de ervaringen van leerlingen met testangst?	29
4.4.	Wat zijn de ervaringen van leerkrachten met testangst bij hun leerlingen?.....	31
4.5.	Welke concrete methodieken hanteren leerkrachten als het aankomt op het verminderen van testangst bij hun leerlingen?.....	33
5.	Discussie en conclusie.....	34
5.1.1.	Wat zijn de ervaringen van leerlingen met de door hun leerkrachten gegeven toetsen en feedback?	34
5.1.2.	Wat zijn de ervaringen van leerkrachten met de door hun gegeven toetsen en feedback?	34
5.1.3.	Wat zijn de ervaringen van leerlingen met testangst?	36
5.1.4.	Wat zijn de ervaringen van leerkrachten met testangst bij hun leerlingen?.....	37

5.1.5. Welke concrete methodieken hanteren leerkrachten als het aankomt op het verminderen van testangst bij hun leerlingen?	37
5.1.6. Overkoepelende discussie	37
5.2. Aanbevelingen	38
5.3. Beperkingen van dit onderzoek en suggesties voor vervolgonderzoek.....	39
Bibliografie	41
Bijlagen	48
Bijlage 1: interviewleidraad leerkrachten.....	48
Bijlage 2: vragenlijst leerlingen.....	49
Bijlage 3: informatieformulier leerkrachten.....	50
Bijlage 4: blanco informed consent voor leerkrachten	51
Bijlage 5: blanco informed consent voor leerlingen en hun ouders.....	52
Bijlage 6: codeboom.....	53

1. Inleiding

“*Investeer in ons voor het te laat is*”. Dit citaat komt van een jongere uit één van de vele krantenartikels die de afgelopen jaren verschenen zijn over de mentale gezondheid van leerlingen in het secundair onderwijs. Zo meldt onder andere het United Nations Children’s Fund (2021) in haar rapport ‘On my mind’ dat zelfdoding de op één na belangrijkste doodsoorzaak is bij Europese jongeren. Ook specifiek in Vlaanderen wordt het probleem aangekaart. Uit het SIGMA-onderzoek van KU Leuven blijkt namelijk dat één op de vijf jongeren tussen twaalf en zestien jaar kampt met matige tot ernstige psychische klachten (Hagemann et al., 2019). De recente studie van de Vlaamse Vereniging van Studenten (2021) bevestigt deze cijfers. Ook de jaarverslagen van de CLB’s en onderzoek van het Departement van Cultuur, Jeugd en Media melden een stijging van de mentale klachten van jongeren (KeKi vzw, 2022). Minister van Onderwijs Ben Weyts maakt bijgevolg 8,5 miljoen euro per jaar vrij voor de CLB’s om tegemoet te komen aan de toegenomen zorgnoden van leerlingen.

Daarenboven komt de maatschappelijke focus de laatste jaren echter almaar meer te liggen op de resultaten die Vlaamse leerlingen halen op grootschalige internationale onderzoeken. De zogeheten ‘*public value of learning*’ neemt toe gezien de bredere publieke en economische waarden die onderwijs dient (Jarvis, 2010). Door de rijzende (simplistische) focus op werk en leren, komt er minder tijd vrij voor ontspanning en zelfzorg (Amelinckx et al., 2023). Eén van de ondervraagde jongeren uit het rapport ‘On my mind’ pleit daarom voor het feit dat scholen zich niet enkel richten op kennisoverdracht, maar ook op het oprichten van ondersteunende programma’s voor leerlingen om beter te kunnen praten over en omgaan met stress en welzijn (United Nations Children’s Fund, 2021). Onderzoek van Demeester en Vanmaercke (2021) bevestigt dit door te stellen dat gevoelens van angst en depressie bij leerlingen voor een deel kunnen vermeden worden indien ze de juiste ondersteuning zouden krijgen op school. Leerlingen in het secundair onderwijs zijn kwetsbaar, maar toch is er net voor deze groep onvoldoende ondersteuning en aandacht – ook niet in het Vlaams onderwijs (Amelinckx et al., 2023). De school is echter een plaats waar jongeren veel tijd spenderen, idealiter worden eventuele steunmaatregelen dan ook daar geïmplementeerd.

De link tussen de nadruk op schoolprestaties en het ervaren van stress bij leerlingen in het secundair onderwijs kan aangetoond worden door middel van de toegenomen prevalentie van testangst binnen deze doelgroep (Rotenstein et al., 2016). Vanstone en Hicks (2019) stellen dat testangst in het secundair onderwijs dermate in de belangstelling staat omdat er zo’n hoog uitvalpercentage aan verbonden is. Dit kan verklaard worden door het feit dat testangst een belangrijke determinant is voor slechtere schoolresultaten, met het watervalstelsel als gevolg (Huntley et al., 2022; Van Praag et al., 2015). Voor het opvangen van dergelijke testangst raden talrijke studies het gebruik van formatieve toetsing aan (Holmes, 2014; Rotenstein et al., 2016; Yousefi et al., 2015; Yousefi, 2019). Een concrete vorm van dergelijke formatieve toetsing, met name *self-assessment*, wordt geacht de *self-efficacy* van lerenden te verhogen (Andrade et al., 2009). Aansluitende feedback hierop van de instructieverantwoordelijke zou de zelf toegewezen scores van de leerlingen nog verfijnen (Dochy et al., 1999). Ook *ipsative*

assessment krijgt een plaats binnen deze discussie. Eerder vermelde onbedoelde effecten van de huidige visie omtrent toetsen zouden kunnen worden tegengegaan indien leraren meer aandacht schenken aan *ipsative assessment* en de stem van de leerlingen betrekken in hun resultaten (Bourke et al., 2018; Hughes, 2015).

Hoewel reeds werd aangetoond dat de implementatie van *self-* en *ipsative assessment* bevorderlijk zou zijn voor het inperken van testangst bij leerlingen in het secundair onderwijs (Bourke et al., 2018; Holmes, 2014; Hughes, 2015; Rotenstein et al., 2016; Yousefi et al., 2015; Yousefi, 2019), zijn er weinig studies binnen het Vlaams secundair onderwijs die de ervaringen rapporteren van zowel leerkrachten als leerlingen omtrent het gebruik van dergelijke beoordelingsvormen en het ervaren van testangst. Deze studie wil tegemoet komen aan het hiaat dat bestaat op vlak van het geven van een stem aan zowel leerkrachten als leerlingen in het Vlaams secundair onderwijs. Meer bepaald worden deze stakeholders bevraagd omtrent hun ervaringen met alternatieve beoordelingsvormen en testangst. De hoofdonderzoeksvraag die richting gaf aan deze studie is: *Hoe wordt het gebruik van self- en ipsative assessment en het ervaren van testangst door leerkrachten en leerlingen gepercipieerd?*

In volgend theoretisch kader zullen eerst alle belangrijke begrippen toegelicht worden. Eerst wordt het toegenomen belang van leerprestaties gekaderd waarna formatieve en summatieve toetsing zullen worden besproken. Dan worden specifiek *self-* en *ipsative assessment* toegelicht. Vervolgens verschuift de focus naar het psychologisch welbevinden en vooral de toenemende ervaring van testangst bij leerlingen. Dit vormt de nodige wetenschappelijke basis om via kwalitatief onderzoek de eventuele relatie tussen beide luiken te onderzoeken en vooral de ervaringen ermee bloot te leggen. Daarna wordt de methodologie van dit onderzoek besproken waarna een omschrijving van de resultaten van deze interviewstudie gegeven wordt. Ten slotte worden de resultaten geïnterpreteerd waardoor nieuwe perspectieven worden geopend voor vervolgonderzoek.

2. Theoretisch kader

2.1. Veranderende visie omtrent onderwijs

België behoort tot de top 20 van de meest concurrerende kenniseconomieën ter wereld (Schwab, 2020). Om die positie te behouden, dient de kwaliteit van het onderwijs optimaal te zijn door voldoende te investeren in kennisontwikkeling. Een ‘zesjescultuur’ – waarbij inspanningen gericht zijn op een minimaal resultaat – volstaat niet langer en de schoolprestaties moeten omhoog (Dijkstra, 2015). De dalende prestaties van Vlaamse leerlingen in internationale onderzoeken als PISA, TIMMS en PIRLS leidden zo in 2020 tot de oprichting van de commissie Beter Onderwijs (Vlaams Ministerie van Onderwijs en Vorming, 2021). Hierin communiceren diverse experts adviezen die vertrekken vanuit het leerproces van een leerling, met als doel om leerlingen tot beter leren te brengen. De krachtlijnen van de adviezen zijn voornamelijk het teruggaan naar de kern van het onderwijzen, waarbij onderwijsinnovaties zonder wetenschappelijke onderbouwing moeten geweerd worden alsook dat de samenleving niet elke maatschappelijke taak mag doorschuiven naar het onderwijs (Vlaams Ministerie van Onderwijs en Vorming, 2021). Philip Brinckman, voorzitter van de commissie, stelt dat het curriculum vandaag te zwaar is door vele maatschappelijke topics, zoals gezonde voeding en verkeersveiligheid. *“Uiteindelijk moet de focus liggen op verbeteringen van het onderwijs in de klas, zoals meer leertijd, betere instructie, meer en betere feedback aan leerlingen, hoge eisen stellen en aandacht voor leerstrategieën”*, stelden Ledoux en collega’s (2009).

Ook internationaal gezien kent het meten van (louter) academische resultaten de laatste jaren een enorme toename in belang (Coates, 2014). Deze extra focus op prestaties en individuele verantwoordelijkheid van leerlingen brengt echter onbedoelde risico’s met zich mee. Met de komst van dit nieuwe denken in de OESO-landen verschijnen namelijk ook verantwoordingsverplichtingen en competitiviteit op het toneel. Denk hierbij bijvoorbeeld aan de recente ‘Onderwijzer’ van Het Nieuwsblad (2023) die alle data van alle secundaire scholen in Vlaanderen bundelt, om zo de schoolkeuze van ouders en hun kinderen te vergemakkelijken. Ook de Vlaamse Kwalificatiestructuur (VKS) geeft een internationaal vergelijkbare standaard aan ten aanzien van onder andere het optimaal kunnen functioneren in de maatschappij (AHOVOKS & Vlaams Ministerie van Onderwijs en Vorming, z.d.). Dit kan resulteren in het feit dat er dan vooral aandacht geschonken wordt aan resultaten, meer dan aan het initiële leerproces zelf (Bourke et al., 2018). Leerkrachten worden hierdoor sterk beïnvloed om de prestaties van hun leerlingen af te stemmen op de vooropgestelde nationale eindtermen en/of benchmarks. Dit brengt echter het natuurlijk pedagogisch inzicht in het gedrang dat de leraar in staat stelt diens lespraktijk adequaat af te stemmen op de noden van de lerenden (Ball, 2003). Stobart (2008) concludeert hieromtrent dat het initiële doel van toetsen, namelijk leerlingen ondersteunen in hun leerproces, in het gedrang komt.

Het fenomeen dat optreedt wanneer dergelijk beoordelingssysteem het onderwijs controleert en beïnvloedt in plaats van omgekeerd, heet *‘the tail waggles the dog’* (Valcke, 2007). Zo ontstaat het risico dat enkele belangrijke vaardigheden niet meer getoetst en uiteindelijk ook niet meer aangeleerd worden wanneer er een te sterke klemtoon gelegd wordt op gestandaardiseerde toetsen in plaats van op alternatieve toetsvormen (Moerkerke, 1996). Het wordt vooral zorgwekkend wanneer leraren te veel druk ervaren om hun instructie te oriënteren op de beoordeling die past binnen een bepaalde visie (Adams, 2017; Ball, 2003). Dat komt onder andere sterk terug wanneer een *high-stakes* vorm van toetsing gehanteerd wordt. Dit

systeem draagt grote gevolgen in zich voor lerenden, aangezien het dient om transparant te zijn naar zowel ouders (cfr. Onderwijzer) als beleidsmakers toe om de normen van het onderwijs hoog te houden en verantwoording af te leggen (cfr. Vlaamse Kwalificatiestructuur), al dan niet in ruil voor extra middelen (Heubert & Hauser, 1999; Klenowski & Wyatt-Smith, 2012). Dit vergroot de kans bij leerlingen om hun jaar te moeten overdoen of te moeten veranderen van school, maar ook bij leerkrachten wordt de kans vergroot op ethisch-pedagogische dilemma's over welke vorm van leren geïmplementeerd dient te worden (Alexander, 2011; Black & William, 1998, 2006). Dat heeft een effect op elke leerling en kan zorgen voor stress, lager zelfvertrouwen en testangst (Brandmo et al., 2019; Marchant, 2004; Reddell, 2010). Het primaire doel van een school met een *high stakes* beleid is namelijk eerder gericht op haar eigenbelang ten koste van de behoeften en het welzijn van haar leerlingen (Kaščák & Pupala, 2011).

Hoewel de OESO-landen meer en meer resultaatgericht denken op vlak van onderwijs, kennen het laatste decennium vooral Europa en de Verenigde Staten toch ook een beweging naar het denken over alternatieve toetsvormen (Birenbaum & Dochy, 1996). Waar namelijk voordien gesproken werd over toetsing an sich wordt nu eerder gesproken over 'beoordeling', waarbij toetsing en instructie niet langer strict gescheiden zijn van elkaar en waarbij de leerlingen – idealiter – niet langer passief worden onderworpen aan een test met een totaalscore die geen rekening houdt met hun context en ervaring (Birenbaum, 1996; Wolf et al., 1991). Dit resulteert in de behoefte om toetsingsinstrumenten te ontwikkelen die een eerlijke en valide beoordeling van leren mogelijk maken (Zlatkin-Troitschanskaia et al., 2015, 2016). Dat is nodig aangezien anders de betrouwbaarheid van de evaluatie in het gedrang komt omdat momentopnames er niet in slagen de totale beheersing van leerdoelen te meten en een louter leerstofgerichte interpretatie van resultaten te binair is (Valcke, 2007). Een score is namelijk geen precieze meting van intelligentie. Deze verschuiving in visie heeft als gevolg dat tests niet langer gezien worden als instrument om lerenden te voorzien van een diploma, maar ook als waardevol om hun vooruitgang te monitoren en eventueel extra ondersteuning te bieden. Met andere woorden kan gesteld worden dat beoordeling gebruikt dient te worden om te leren, en niet omgekeerd (Arter, 1997; Dochy & McDowell, 1997). De leerling wordt hierbij dan gezien als een actief persoon die verantwoordelijk is voor diens eigen resultaten (Dochy et al., 1999).

Aangezien toetsen onderdeel uitmaken van een bepaald onderwijsconcept, dienen ze hierin dan ook optimaal te passen. Dit wil zeggen dat afhankelijk van de visie die een school wenst na streven en van het doel dat ze wenst te bereiken, er gezocht wordt naar de meest adequate toetsvorm (Linn, 1990; Shepard, 2000). Om tegemoet te komen aan de huidige maatschappelijke vereisten, volstaan de enkel conventionele toetsvormen dus niet langer en dient er geïnvesteerd te worden in alternatieve toetsvormen (Joosten-ten Brinke, 2010). Beoordeling van louter kennisreproductie is namelijk niet langer voldoende (Dochy et al., 1999; Dochy & Moerkerke, 1997). Dergelijke alternatieve toetsvormen, zoals *self-* en *ipsative assessment* of een combinatie van verschillende beoordelingssystemen, zijn door deze evolutie in denken naast de traditionele toetsvormen op het toneel verschenen (Valcke, 2007). Een direct voordeel dat aan deze nieuwe vormen kan gekoppeld worden, is het feit dat leerlingen nu meer betrokken worden bij de toetsing omdat ze extra inzicht verwerven in hun eigen leerproces en dat ze bijgevolg een grotere verantwoordelijkheid en autonomie dragen in het tonen van hun kunnen (Boud & Falchikov, 1989; Fiddler & Marineau, 2008). Deze grotere betrokkenheid en verantwoordelijkheid bij toetsen brengt zeker voordelen met zich mee, maar

auteurs waarschuwen ook voor het feit dat dit systeem ze net meer zou controleren en disciplineren, in plaats van ze te bekrachtigen (Tan, 2010). Ook de mate van sturing door leraren verschilt per vorm, wat aanpassingsvermogen vereist (Joosten-ten Brinke, 2010).

2.2. Vormen van toetsing

Waar tests aangepast worden aan de onderwijsvisie, passen lerenden zowel hun leergedrag als perceptie op instructie aan aan de vorm van de tests (Beckwith, 1991; Torrance, 2007; Vaessen et al., 2017; Valcke, 2007). Die vormen zijn de laatste jaren in aantal toegenomen.

2.2.1. Formatieve en summatieve toetsing

Meetgestuurd onderwijs bestaat uit twee varianten, namelijk formatief en summatief – afhankelijk van de beslissing die de leerkracht wil nemen. *Summative assessment* kan herkend worden aan de sluitende bedoeling ervan, het is met name enkel gericht op het bepalen van het niveau van de lerende en of een eindterm bereikt is. Dit heeft echter niet enkel een effect op (het leergedrag van) leerlingen, maar ook op beleidsmakers gezien zij het onderwijs aanpassen op basis van externe beoordelingen, zoals bij *high stakes* testsystemen. Ook de effectiviteit van deze vorm is meer omstreden, omdat in de discussie erover het accent voornamelijk ligt op het subjectief handelen van leraren en scholen, verarming van het curriculum en onderstimulering van betere leerlingen (Ledoux et al., 2009).

De *Assessment Reform Group* (2002) definieert *formative assessment* als ‘een proces van het interpreteren van bewijs door zowel leerling als leerkracht om te beslissen over waar de leerling zich momenteel bevindt in het leerproces, waar die naartoe zou moeten en hoe die daar best geraakt’. Er zijn vier hoofdcomponenten te onderscheiden, namelijk: (1) het duidelijk communiceren van de leerdoelen en de beoordelingscriteria, (2) de kwaliteitsverbetering van dialoog, (3) eerlijke en kwalitatieve feedback en scoring en ten slotte (4) het gebruik van *self-assessment* (Black et al., 2003; Ozan & Kincal, 2018; William, 2011). Hierdoor worden leraren en leerlingen zich bewust van de individuele vooruitgang en behoeften waardoor de instructie van de leraar adequaat kan aangepast worden (Nishizuka, 2022).

Beoordeling is bij *formative assessment* eerder gericht op begeleiding met eventuele bijsturing en verbetering van het leerproces. Het dient met andere woorden als diagnostische tool en als bron van informatie over de ontwikkeling van het leren van leerlingen. Het wordt om die reden dan ook beoordeling *voor* het leren genoemd in plaats van het beoordelen *van* het leren (Ozan & Kincal, 2018). De kracht van formatieve toetsing ligt in het feit dat het feedback voor zowel leraar als leerling voorziet (Biggs & Tang, 1999; Brown & Knight, 1994; Ding, 1998). Samenvattend kan dus gesteld worden dat het summatieve eerder het leerresultaat quoteert terwijl de formatieve variant het leerproces van verbetering voorziet. Hoewel *formative assessment* misschien als de betere testvorm van beide overkomt, is het net belangrijk een en-en verhaal te creëren en beide vormen elkaar te laten aanvullen zodat alle sterktes ervan optimaal benut worden (Shavelson et al., 2008).

Hoewel zowel summatieve als formatieve toetsing voordelen kennen, is vooral de relatie tussen formatieve toetsing en bepaalde positieve factoren vaak onderzocht. Zo is uit onderzoek gebleken dat het kan bijdragen tot verbeterde leerresultaten, zeker wanneer de leerlingen inhoudelijke en taakgerichte feedback krijgen (Ledoux et al., 2009). Ook vonden Rakoczy en collega's (2019) dat bij deze beoordelingsvorm de feedback op een meer nuttige

manier wordt gebruikt en dat zowel de *self-efficacy* als de interesse van de leerlingen toenamen. Ook de academische motivatie en de metacognitieve strategieën groeien aan, aldus Afrashteh (2019). Een belangrijke studie van Leenknecht en collega's (2021) heeft daarenboven aangetoond dat frequenter gebruik van *formative assessment* geassocieerd kan worden met meer gevoelens van autonomie, competentie en autonome motivatie bij de leerlingen. Deze variabelen zouden leiden tot een groter gevoel van academische motivatie (Deci & Ryan, 2000; Lecluyse, 2018). Ook ligt dit in lijn met het vijfcomponentenmodel van Edmonds (1979) dat stelt dat effectieve scholen onder andere een veilig leerklimaat moeten garanderen, gericht moeten zijn op het bevorderen van leerprestaties en hoge verwachtingen moeten hebben ten aanzien van de leerprestaties van alle leerlingen (De Maeyer & Rymenans, 2004).

2.2.2. Ipsative assessment

Een toetsvorm die een degelijke aanvulling zou kunnen betekenen voor de toepassing van *formative assessment* is *ipsative assessment*. De belangrijkste reden hiervoor is dat deze toetsvorm feedback biedt die gekoppeld is aan individuele leerwinst in plaats van aan externe criteria en normen (Hughes, 2011). *Ipsative assessment* vergelijkt namelijk de prestaties van eenzelfde individu met diens prestaties in het verleden om gerichte feedback te geven over de persoonlijke vooruitgang, in plaats van dat het de prestaties van een individu vergelijkt met een norm of met diens peers (Hughes, 2011, 2014, 2015; Nishizuka, 2022; Seery et al., 2019). Dit komt zowel de taak zelf, het levenslang leren en de motivatie van leerlingen ten goede omdat het de ontwikkeling van autonomie ondersteunt. Rattray (2018) stelt vervolgens dat lerenden dankzij *ipsative assessment* gewapend zijn tegen de toekomst door middel van de ontwikkeling van veerkracht en optimisme. Bovendien is het langetermijngericht; het stelt lerenden namelijk in staat om haalbare en realistische doelen te stellen. Uit onderzoek is daarenboven gebleken dat leerlingen *ipsative assessment* als motiverend en nuttig beschouwen (Hughes et al., 2011). *Ipsative assessment* sluit dus naadloos aan bij de visie dat tests niet langer gezien worden als instrument om lerenden te voorzien van een diploma, maar ook als waardevol instrument om hun vooruitgang te monitoren en eventueel extra ondersteuning te bieden (Arter, 1997; Dochy & McDowell, 1997).

Het is belangrijk oog te hebben voor *ipsative assessment* omdat leerkrachten beïnvloed worden door het huidige beoordelingsbeleid dat doorgaans verantwoordingsverplichtingen omvat om de prestaties van leerlingen af te stemmen op nationale benchmarks. Vele OESO-landen kenden de laatste decennia namelijk een overmatige belangstelling voor resultaten in plaats van het initiële leren en instructie (Bourke et al., 2018). Dit beleid heeft vorm gekregen door onder andere wetenschappelijke opvattingen (Alexander, 2011; Bourke, 2016; Thrupp, 2018). Het werkt echter in op het natuurlijk pedagogisch inzicht van een leraar die diens leerlingen kent en bijgevolg weet wat hun noden zijn. De spanning die hierdoor ontstaat tussen weten wat het beste is voor een leerling en weten wat het beleid verwacht, wordt gedefinieerd als 'de potentiële spagaat tussen de eigen oordelen over *good practices* en de noden van leerlingen en de striktheid van de prestaties' (Ball, 2003). Het risico bestaat dan dat de initiële intentie van toetsing, namelijk leerlingen ondersteunen in het leren, in het gedrang komt (Stobart, 2008).

Volgens Bourke en collega's (2018) kunnen de onbedoelde effecten van dit beoordelings- en verantwoordingsbeleid in toom worden gehouden wanneer leraren aandacht schenken aan zowel *ipsative assessment* als aan de stem van de leerling in de beoordeling. Uit onderzoek van Bourke (2010) blijkt namelijk dat leerlingen *ipsative assessment* reeds uit zichzelf

toepassen in buitenschoolse authentieke contexten. Zo zouden ze reeds in staat zijn om in te schatten hoe ze zichzelf effectief kunnen motiveren en weten ze ook van zichzelf hoe ze persoonlijke topprestaties zouden kunnen bereiken.

Hoewel *ipsative assessment* slechts een korte onderzoeksgeschiedenis kent, worden toch reeds de eerste bevindingen genoteerd. Zo blijkt dat er na het gebruik van *ipsative assessment* een kwalitatieve verbetering in prestatiegericht leren is vastgesteld. Dat is mogelijk omdat deze beoordelingsvorm ook de ontwikkeling van cognitieve hogere-orde functies ondersteunt, zoals academisch schrijven en evalueren (Nishizuka, 2022). Deze zaken kunnen onmogelijk worden gemeten met de traditionele kennisreproductie-toetsen (Malecka & Boud, 2021; Malecka et al., 2021; Seery et al., 2019; Univio & Perez, 2019). Vervolgens biedt *ipsative assessment* een antwoord tegen de huidige neoliberale samenleving die de concurrentie tussen leerlingen en scholen bevordert. Zo creëert het een veilige leeromgeving voor lerenden die ieders identiteit respecteert en helpt om sociale emoties te ontwikkelen zoals veerkracht en nieuwsgierigheid (Lucas et al., 2013; Rattray, 2018).

Daarenboven zou *ipsative assessment* ook het geven van kwalitatieve feedback kunnen vergemakkelijken. Dat is belangrijk aangezien onderzoek vaststelt dat leerlingen de feedback van hun leraar vaak niet nuttig en/of bruikbaar vinden (Carless et al., 2011; Price et al., 2010; Scoles et al., 2013). Redenen hiervoor zijn dat het vaak te vaag geformuleerd is (Weaver, 2006) en dat het te laat wordt gegeven waardoor het zijn doel voorbijschiet (Price et al., 2010). Op die manier begrijpen leerlingen de feedback niet of verkeerd (Carless, 2006; Lea & Street, 1998). Katholiek Onderwijs Vlaanderen (2023) rapporteert daarenboven in haar jaarverslag van schooljaar 2021-2022 dat slechts 10% van de leerkrachten deelneemt aan het nascholingsproject 'Kunst en kunde van feedback'. *Ipsative assessment* zou aan dit hiaat kunnen tegemoetkomen aangezien het de prestaties in de tijd illustreert wat het eenvoudiger maakt voor leerkrachten om een individueel traject te kunnen opstarten met individuele doelen en verwachtingen. Ook laat het de leerling effectief handelen naar die feedback, aangezien het specifiek gericht is op diens individuele prestaties en de leerling zelf betreft in het proces waardoor het persoonlijk relevant wordt (Hughes, 2011).

Hughes (2015) erkent dat *ipsative assessment* niet nieuw is, maar dat de formele systemen die het toepassen toch schaars zijn. Dit komt omdat vele leraren deze beoordelingsvorm waarschijnlijk toepassen zonder het zelf te beseffen of zonder de term echt te kennen. Om *ipsative assessment* concreet toe te passen in de klaspraktijk kan gebruik gemaakt worden van alle beschikbare beoordelingsvormen om de beste scores van leerlingen op vast te leggen alsook om veranderingen in deze resultaten adequaat op te volgen. Dit laatste wordt vaak over het hoofd gezien maar is nochtans een cruciaal aspect aangezien het een potentieel motiverend instrument is. Het zorgt er namelijk voor dat zowel de leerdoelen van de leerlingen als de methodieken van de leerkracht op elkaar worden afgestemd (Bourke & Mentis, 2014). Ook kan een *cover sheet* met feedback voor opdrachten ontwikkeld worden om leerlingen te laten reflecteren over de feedback en leraren te laten nagaan hoe hun feedback wordt toegepast. Online alternatieven zoals e-portfolio's en zelftoetsen in virtuele leeromgevingen vormen zeker ook een waardevolle invulling (Malecka et al., 2021; Penn & Wells, 2018). Belangrijk om hierbij te onderkennen is dat *ipsative assessment* vooral een aanvulling dient te zijn op andere vormen van assessment in een duaal beoordelingssysteem (Hughes, 2015). De grootste uitdaging van *ipsative assessment* is dus de operationalisering ervan, aangezien

het inefficiënt en onpraktisch kan zijn om tijdige en relevante feedback te geven op maat van elke leerling (Hughes, 2011).

2.2.3. Self-assessment

Het gebruik van feedback is zowel binnen summatieve als formatieve toetsing essentieel. Belangrijk om hierbij te onderkennen is dat niet enkel de instructieverantwoordelijke, maar ook de leerlingen zélf kunnen instaan voor het verschaffen van deze feedback. Leerlingen worden op die manier actief betrokken bij het leer- en beoordelingsproces (Nicol, 2007). De verantwoordelijkheid die hierbij komt kijken om te begrijpen welke kwaliteiten en normen samenhangen met de gewenste prestaties, werkt een zelfregulerend vermogen in de hand dat essentieel is voor het leerproces (Zimmerman, 2000).

Concreet beslaat *self-assessment* 'het managen van zelf-gegenereerde gedachten, gevoelens en acties die aangepast zijn aan het bereiken van persoonlijke doelen' (Zimmerman, 2000). *Self-efficacy* wordt gedefinieerd als 'de overtuiging van iemands persoonlijke capaciteiten om een taak succesvol te kunnen verwezenlijken', wat zorgt voor een toegenomen gebruik van zelfregulatiestrategieën alsook het gebruik van *self-assessment* (Bandura, 1997; Pajares, 1996, 2012; Schunk & Ertmer, 1999). Deze overtuigingen van eigenwaarde zijn volgens Rabbani en Yousefi (2012) onmisbaar gezien ze de manier waarop leerlingen zichzelf motiveren vormgeven en dit ze bijgevolg minder kwetsbaar maakt voor testangst.

Over het algemeen zou *self-assessment* de *self-efficacy* van lerenden zelfs kunnen verhogen (Andrade et al., 2009). Deze stelling wordt bevestigd door de veronderstelling dat lerenden een dieper begrip van de taakvereisten zouden hebben, waarna ze als gevolg van de succesvolle prestaties gevoelens van zelfwaarde en bekwaamheid zouden ontwikkelen (Paris & Paris, 2001). Barana en Marchisio (2016) wijzen daarenboven op het feit dat er zelfs sprake is van een toegenomen betrokkenheid en een versterkt gevoel van eigenaarschap en motivatie en dat zo zelfs de prestaties van de lerenden verhoogd zouden kunnen worden.

Nu de relatie tussen *self-assessment* en *self-efficacy* duidelijk is, kan *self-assessment* ook concreter afgelijnd worden: 'het is de grote verscheidenheid aan mechanismen waarmee leerlingen de kwaliteiten van hun eigen leerprocessen en -producten beschrijven en er eventueel zelf een waarde aan toekennen' (Panadero et al., 2016). Het is met andere woorden de betrokkenheid van lerenden bij de beoordeling van hun eigen leren en prestaties (Boud & Falchikov, 1989). Bij deze definitie kan enerzijds echter opgemerkt worden dat het enkel gaat over het beoordelen van het eigen werk en niet over het signaleren van extra kennisverwerving aan de leerkracht, in tegenstelling tot *ipsative assessment*; en anderzijds dat er verschillende manieren bestaan om *self-assessment* in de klas toe te passen, zoals het toekennen van een cijfer aan het eigen werk of zoals een grondige analyse van werkpunten en sterke punten aan de hand van formatieve feedback (Andrade, 2010). De definitie van Boud (1995) is breder en houdt wel rekening met het volledige leerproces: 'het is een manier om de rol van studenten als actieve deelnemers aan hun eigen leren te vergroten'. Hierop voortbouwend wordt beweerd dat *self-assessment* pas effectief is wanneer er ook zelfreflectie aan gekoppeld wordt die verduidelijkt wat het resultaat betekent voor de verdere ontwikkeling van de leerling (Joosten-ten Brinke et al., z.d.).

Empirisch onderzoek ondersteunt de idee van *self-assessment* als essentie voor productief leren, gezien het feit dat indien een leerling blindelings de leerkracht volgt zonder het doel van de instructie en de evaluatie te begrijpen, die nooit zal leren diens eigen werk te controleren en te monitoren (Black & William, 1998). Verscheidene studies toonden het positieve effect aan van *self-assessment* op academische resultaten. Het zou namelijk leiden tot een betere effectiviteit en kwaliteit van leren (Topping, 2003) alsook tot zelfs betere prestaties op verschillende niveaus en vakgebieden (Brown & Harris, 2013) – al merken deze laatste auteurs wel op dat de basis van deze positieve effecten gelinkt kan worden aan de complexiteit van de *self-assessment*, eerder dan aan het type. Dankzij de toegenomen helderheid van de leerdoelen leren de leerlingen namelijk plannen, monitoren en evalueren (Andrade & Valtcheva, 2009; Brown & Harris, 2013; Jonsson, 2014; Nicol & McFarlane-Dick, 2006; Panadero & Alonso-Tapia, 2013). Hassmen en collega's (1996) stellen echter geen voorwaarden aan het gebruik van *self-assessment* en stellen dat lerenden die dit toepassen tijdens het leren hogere prestaties behaalden tijdens tests dan leerlingen die dit niet doen. De grootste reden waarom *self-assessment* zo bevorderlijk zou zijn voor het leren van leerlingen, is volgens Longhurst en Norton (1997) het feit dat het leerlingen in staat stelt zich bewust te worden over de metacognitieve aspecten van hun leren. Daarenboven zou het gebruik van zelfregulatiestrategieën en zelfreflectie de goede afloop van een taak ook bevorderen (Bandura, 1997; Pajares, 1996, 2012; Schunk & Ertmer, 1999).

Wanneer de scores vergeleken werden die toegekend zijn door de docent en door de leerling zelf, worden echter significante verschillen gerapporteerd: de score toegekend door de leerling zelf ligt namelijk hoger dan de score toegekend door diens docent (de Grez et al., 2012). Wel stelt onderzoek van Dochy en collega's (1999) dat enkel de zwakkere leerlingen de neiging hebben om zichzelf hoger in te schatten, en dat dit het tegengestelde is bij oudere en beter presterende leerlingen.

Williams (1992) besluit dat de grote meerderheid van de leerlingen de voordelen inziet van *self-assessment*, meer bepaald dankzij het feit dat ze de uitwisseling van informatie met de docent nuttig achten. Wanneer ze hun aan zichzelf toegekende cijfer kunnen vergelijken met het cijfer dat ze van hun docent hebben gekregen met bijhorende feedback, zijn in staat om realistisch te zijn over de kwaliteit van hun werk. Wel wensen de leerlingen dat er voorafgaand aan de *self-assessment* concrete richtlijnen worden meegegeven en dat er achteraf feedback op dient te volgen (Birenbaum & Dochy, 1996; Boud & Falchikov, 1989; Griffie, 1995; Williams, 1992). Ook de motivatie van de leerlingen speelt een cruciale rol in de ervaring met *self-assessment*, want hoe hoger hun motivatie, hoe nauwkeuriger de resultaten (Longhurst & Norton, 1997). Om deze motivatie te bevorderen, kan een leeromgeving gecreëerd worden waarin de leerling en diens leren centraal staat, en niet de instructie van de leerkracht. Het moet met andere woorden duidelijk zijn voor de lerenden dat *self-assessment* inherent deel uitmaakt van het leerproces (Zoller & Ben-Chaim, 1997).

Na de – veelal positieve – factoren besproken te hebben van *self-assessment*, kan gekeken worden naar hoe leraren dit nu concreet kunnen toepassen in hun lespraktijk. McNamara & Deane (1995) stellen voor dat de leerlingen een leerlogboek bijhouden en een portfolio opstellen. Dit zou leerlingen helpen met het identificeren van hun sterke en zwakke leerstrategieën alsook hun voortgang. Door zich bewust te zijn van de eigen leercontext, wordt het namelijk meer haalbaar om ook individuele doelen te stellen. Harrington (1995) voegt hier het gebruik van een lijst met bekwaamheden aan toe waarop leerlingen kunnen aanduiden

welke ze het best en het slechtst beheersen. Dit kan ook in de vorm van een Likert-schaal. Ook kan het helpen om bij elke bekwaamheid een voorbeeld te zetten van een toepassing van die bekwaamheid. Deze formulieren worden geacht goedkoper en minder tijdrovend te zijn dan traditionele manieren om leerlingen te beoordelen (Nevo, 1995).

2.3. Psychologisch en academisch welbevinden bij leerlingen

De prevalentie van psychologische problemen zoals depressie en testangst bij lerenden neemt toe in vergelijking met die van de algemene bevolking (Dyrbye et al., 2006). De onbedoelde risico's van beoordeling beïnvloeden namelijk het leerproces (Hattie & Timperley, 2007), testangst (Huxham et al., 2012), het gevoel van eigenwaarde (Betts et al., 2009) en leerpercepties- en gedrag (Reeves, 2006). Academisch welzijn kan beschouwd worden als de houding die leerlingen hanteren ten opzichte van onderwijs, hun academische betrokkenheid en de tevredenheid over hun academische prestaties (Nguyen & Deci, 2016). Academisch welzijn kan daarom overkoepelend als belangrijke indicator rechtstreeks in verband gebracht worden met academische prestaties, want leerlingen die zowel emotioneel als cognitief betrokken zijn bij hun studies zullen zich meer aanpassen aan hun behoeften en meer energie aan het leren besteden (Abbasi et al., 2016).

2.3.1. Testangst

Een hogere mate van psychologisch welbevinden bij lerenden leidt tot betere academische prestaties, al dan niet dankzij de implementatie van alternatieve assessmentvormen als *self-* en *ipsative assessment* (Abbasi et al., 2016). Ondanks die bevinding is echter ook naar voren gekomen dat dit psychologisch welzijn de laatste jaren aan het verzwakken is binnen de leerlingenpopulatie. Volgens een studie van Tian-Ci Quek en collega's (2019) zou namelijk één derde van de leerlingen testangst ervaren. Onderzoek in het Verenigd Koninkrijk rapporteert dat 15% van de secundaire scholieren een hoge mate van testangst ervaart, en een nog hoger percentage ervaart lichte tot matige testangst (Putwain & Daly, 2014).

Testangst staat in de volksmond beter bekend als 'faalangst'. Winn en collega's (2019) definiëren het als 'een onaangename emotionele toestand die bijzondere gedragsmatige en psychologische gevolgen heeft en wordt ervaren bij formele evaluatiesituaties'. In een volgende studie beweren Wadi en collega's (2022) dat het zelfs integraal deel uitmaakt van *assessment*. Dat komt omdat het kan variëren van een kleine bezorgdheid tot een slopende angst, die het cognitieve proces van leerlingen in de weg staat. Wel dient opgemerkt te worden dat een *lichte* mate van testangst net bevorderlijk kan zijn voor de academische prestaties van een leerling, gezien het de nodige stimulans geeft om te studeren en zich mentaal voor te bereiden op de test (Wadi et al., 2022). Een extreem niveau van testangst werkt echter averechts en kan zelfs de psychologische en/of fysiologische functies en taakrelevante processen van lerenden belemmeren (Wadi et al., 2022). Individuen die dit in hogere mate ondervinden, reageren met een overmatige bezorgdheid op de gevolgen van falen en onderschatten hun competenties (Zeidner, 1998). Deze bezorgdheid neemt mentale energie weg die anders goed bruikbaar is voor het maximaliseren van academische prestaties (Eysenck et al., 2007). Een mindere mate van concentratie en een minder functionerend werkgeheugen zijn namelijk nefast voor de studieresultaten (Cassady, 2004; Hembree, 1988; Zeidner, 2007). Op die manier kunnen de gevolgen van een hoge mate van testangst ook gelinkt worden aan een tekort aan metacognitieve (leer)vaardigheden. Zo kan er algemeen

geconcludeerd worden dat een hoge mate van testangst leidt tot verminderde academische prestaties (Dorenkamp et al., 2021; Pate et al., 2021; von der Embse & Witmer, 2014).

Niet alleen de uitvoering van en de prestaties op de test zelf, maar ook de leerstrategie verandert bij lerenden die veel testangst ervaren. Zij zouden namelijk opteren voor een meer oppervlakkige studiemethode (Spada & Moneta, 2012, 2014). Ook spieken wordt gerapporteerd als mogelijk gevolg, omdat dit het individu toch in staat stelt om – in zekere mate – ‘goed’ te presteren in testsituaties (Cassady & Johnson, 2002). Vele studies rapporteren daarenboven het feit dat leerlingen met testangst hun capaciteiten als minder goed inschatten, wat zich op zijn beurt ook effectief uit in lagere academische prestaties (Peleg, 2016; Putwain, 2008; Segool et al., 2013; Trifoni & Shahini, 2011).

Vooraf een uitgebreide hoeveelheid leerstof, onaangepaste studievaardigheden, moeilijke toetsvormen, negatieve gedachten en vrouwelijk geslacht kwamen als factoren uit de bus die het meest van invloed zouden zijn op een hoge mate van testangst (Guraya et al., 2018; Hashmat et al., 2008; Khoshhal et al., 2017; Tsegay et al., 2019). Bandura (1997) vond dat academische doelen die zowel specifiek, haalbaar als kortetermijngericht zijn een hoger motiverend effect genereren bij leerlingen. Wanneer deze doelen echter te ambitieus en complex zijn, kan dit inboeten aan motivatie en kan het vervolgens leiden tot angstgevoelens in prestatiesituaties (Drach- Zahavy & Erez, 2002).

Sociale en persoonlijke factoren, culturele waarden en het belang dat een individu hecht aan academisch succes spelen hierbij ook een grote rol. Leerlingen met een Aziatische etniciteit zouden bijvoorbeeld vatbaarder zijn voor examenstress dan leerlingen met een Kaukasische achtergrond (Parsons, 2008). Ook ouderlijke druk, gezinscontext en -klimaat en directe betrokkenheid van ouders zoals hun verwachtingen en steun worden opgenomen als determinanten van testangst (Zeidner, 1998, 2014). Schoolbetrokkenheid heeft dus een positieve invloed op academische prestaties, terwijl druk de schoolprestaties belemmert (Dotterer & Lowe, 2011; Fredricks et al., 2004; Hembree, 1988; Upadyaya & Salmela-Aro, 2013). Dit kan verklaard worden door het feit dat ouderlijke *steun* geassocieerd wordt met positieve attitudes ten aanzien van school en bijgevolg betere leerprestaties (Bowen et al., 2008; Woolley & Grogan-Kaylor, 2006), terwijl ouderlijke *druk* negatief correleert met leerprestaties en positief met testangst (Levpuscek et al., 2012; Putwain et al., 2010).

Om af te sluiten kan concreet gekeken worden naar wat een school zelf kan ondernemen ten voordele van het psychologisch welbevinden van haar leerlingen. Wadi en collega's (2022) pleiten voor een platform voor vroegtijdige interventie, zoals in de vorm van psychologische ondersteuning en adequate training van alle personeel om hiermee om te gaan (Huntley et al., 2022; Soares & Woods, 2020). Zo zou het ook quasi een gewoonte moeten worden leerlingen erop te wijzen dat er niet slechts één juist antwoord of methode is en dat fouten maken erbij hoort. Ze kunnen in deze fouten begeleid worden aan de hand van *modelling* (Greenwood, 1984; Valcke, 2007). Wanneer leerlingen blootgesteld worden aan een enkelvoudig modelantwoord kan dit de ontwikkeling van concepties van kwaliteitsvol werk bemoeilijken door het gebrek aan rijk vergelijkingsmateriaal (Reinholz, 2016). Wanneer hen de mogelijkheid geboden wordt om verschillende oplossingen te bekijken, zullen ze hun sterke en zwakke punten beter kunnen identificeren. Hierdoor ontwikkelen ze een dieper begrip van kwaliteit (Swan, 2006). Parsons (2008) raadt daarenboven aan om geheugenondersteuning zoals open-boekexamens toe te laten, leerlingen te bevragen naar hun mening omtrent de test en

de tijdsdruk te versoepelen. Het gebruik van alternatieve evaluatiemethoden sluit hierbij aan, al dan niet via permanente evaluatie zoals huiswerk, vragen tijdens de les en klassikale oefeningen (Miller & Mitchell, 1994).

Ten slotte spelen ook de zorgcoördinatoren en schoolpsychologen een belangrijke rol in het ondersteunen van leerlingen op school. Eerst en vooral dragen ze bij aan het positief leerklimaat door bijvoorbeeld samen te werken met leerkrachten om strategieën te ontwikkelen die de zelfevaluatie van leerlingen bevorderen. Ook kunnen ze leerlingen ondersteunen in het begrijpen en waarderen van *self-* en *ipsative assessment* als waardevolle instrumenten voor zelfreflectie (Skundberg-Kletthagen & Moen, 2017). Daarnaast zijn ze in staat om leerlingen met testangst te identificeren en te begeleiden (von der Embse et al., 2013). Deze begeleiding kan tot stand komen in overleg met leerkrachten en/of ouders om zo adequate interventies en aanpassingen te kunnen bieden, zoals extra tijd of een rustige omgeving tijdens het evaluatiemoment (Lyon et al., 2014). Kortom vormen zorgcoördinatoren de brug tussen leerlingen, leerkrachten en ouders. Een kritische noot bij de werking ervan is echter het belang van de therapeutische relatie tussen leerling en zorgcoördinator. Pas wanneer de leerlingen die namelijk percipiëren als vertrouwenspersoon, dan is de zorg effectief (Slaton et al., 2023).

3. Onderzoeksdesign

3.1. Probleemstelling

De prevalentie van testangst bij leerlingen in Vlaanderen stijgt de laatste jaren (Hageman et al., 2019; KeKi vzw, 2022; Rotenstein et al., 2016; United Nations Children's Fund, 2021; Vanstone & Hicks, 2019; Vlaamse Vereniging van Studenten, 2021). Dit zou mogelijks te wijten zijn aan het toegenomen belang van prestaties en een te verzadigd curriculum (Bourke et al., 2018; Coates, 2014; Jarvis, 2010; Vlaams Ministerie van Onderwijs en Vorming, 2021). Alternatieve vormen van assessment zoals *self-* en *ipsative assessment* hebben het potentieel om deze testangst in te perken (Barana & Marchisio, 2016; Bourke et al., 2018; Brown & Harris, 2013; Lucas et al., 2013; Nishizuka, 2022; Rabbani & Yousefi, 2012; Rattray, 2018; Topping, 2003). Wetenschappelijk onderzoek naar ervaringen omtrent zowel alternatieve vormen van assessment als van testangst van leerkrachten en leerlingen in Vlaanderen ontbreekt echter. Ook is onderzoek naar *ipsative assessment* op zich nog schaars en pril, zeker in de Vlaamse context. Concreet is het onderzoeksdoel van deze masterproef om een stem te geven aan de actoren om wie de discussie de facto draait en wordt zo getracht de gepercipieerde link te zoeken tussen het ervaren van testangst en het toepassen van alternatieve vormen van assessment. De hoofdonderzoeksvraag luidt: *Hoe wordt het gebruik van self- en ipsative assessment en het ervaren van testangst door leerkrachten en leerlingen gepercipieerd?* Om dit vraagstuk te beantwoorden en om de eventuele relatie tussen bovengenoemde concepten bloot te leggen, worden enkele deelvragen geformuleerd:

- **OV 1:** Wat zijn de ervaringen van leerlingen met de door hun leerkrachten gegeven toetsen en feedback?
- **OV 2:** Wat zijn de ervaringen van leerkrachten met de door hun gegeven toetsen en feedback?
- **OV 3:** Wat zijn de ervaringen van leerlingen met testangst?
- **OV 4:** Wat zijn de ervaringen van leerkrachten met testangst bij hun leerlingen?
- **OV 5:** Welke concrete methodieken hanteren leerkrachten als het aankomt op het verminderen van testangst bij hun leerlingen?
- **OV 6:** Hoe wordt de relatie tussen het gebruik van alternatieve toetsvormen en het ervaren van testangst door leerlingen en leerkrachten gepercipieerd?

3.2. Onderzoeksopzet

Om een antwoord te kunnen bieden op bovenstaande onderzoeksvragen is deze studie als volgt opgezet: eerst werd een literatuurstudie uitgevoerd om een zicht te krijgen op de bestaande kennis omtrent het ervaren van testangst en het gebruik van alternatieve assessmentvormen. Daarna werd een kwalitatief onderzoek uitgevoerd, waarbij twee categorieën van dataverzameling werden gehanteerd. Enerzijds werden namelijk semi-structureerde online of face-to-face interviews afgenomen bij de leerkrachten, anderzijds zijn er vragenlijsten met open vragen voorgelegd aan de leerlingen op hun school gedurende één lesuur. Elk interview met een leerkracht nam gemiddeld 30 minuten in beslag. De leerkrachten gaven steeds hun voorkeur voor online of face-to-face door. Deze keuzemogelijkheid wordt aanbevolen vanuit de bevindingen van McIntosh en Morse (2015) en Schober (2018) omdat beide mogelijkheden zowel voor- als nadelen kennen. De drijfveer voor de keuze was vooral het de leerkrachten zo gemakkelijk mogelijk maken.

Er werd specifiek voor interviews als dataverzamelmethode gekozen omwille van volgende redenen: een semi-gestructureerd interview is veelzijdig en flexibel (Irvine et al., 2013; Mason, 2004; Rubin & Rubin, 2005). Het doel van een interview is namelijk gegevens verzamelen over zaken die niet kunnen worden verkregen met louter kwantitatieve methoden. Onderwijs wordt namelijk beïnvloed door interacties tussen leerkrachten en leerlingen, die niet als wetmatigheden kunnen worden voorgesteld aangezien ze dynamisch zijn (Leeman & Wardekker, 2010). Het interviewen van mensen geeft zo inzicht in hun wereld, opvattingen en ervaringen waardoor ik beter in staat was hun situatie te schetsen (Foley et al., 2021; Hove & Anda, 2005). In de interviewleidraad werd vooraf de structuur van de interviews vastgelegd, maar de vragen werden niet volledig gestandaardiseerd waardoor er ruimte was voor vrije interactie met de participanten om rijkere informatie te ontlokken en door te vragen (Brown & Danaher, 2017; Hove & Anda, 2005; McGrath, 2018; Taylor, 2005). Het verloop van de interviews was bijgevolg flexibel en werd vooral gestuurd door de antwoorden van de participanten (Van Hove, 2014). Zowel de interviewleidraad voor de leerkrachten als de vragenlijst voor de leerlingen zijn te vinden onder Bijlage 1 en 2.

De data-analyse is de meest complexe stap van kwalitatief onderzoek omdat het zodanig dient te gebeuren dat de resultaten een zekere validiteit en betrouwbaarheid bezitten (Thompson, 2022; Varpio, 2020). In dit onderzoek wordt geopteerd voor een thematische analyse. De ervaringen van de ondervraagden worden namelijk gecontextualiseerd en samengevat in thema's die een fenomenologisch verhaal opleveren dat eerder overdraagbaar dan generaliseerbaar is. Op die manier gaat de analyse verder dan het kwantificeren van zinnen die opvallen door louter hun frequentie (Ozuem et al., 2022). De belangrijkste thema's uit het onderzoeksmateriaal worden bestudeerd en de antwoorden van de participanten worden vergeleken (Braun & Clarke, 2006; Van Hove, 2014). Aan de hand van een inductieve analyse wordt dan gezocht naar gemeenschappelijke antwoorden op de onderzoeksvragen (Braun & Clarke, 2006; Varpio et al., 2019). Dit kan overigens gedefinieerd worden als praktijkgericht onderwijsonderzoek aangezien deze masterproef tot doelstelling heeft om nieuwe kennis te genereren in de onderwijspraktijk met betrokkenheid van relevante actoren (Enthoven & Oostdam, 2012; NRO, 2021; Verschuren, 2009). Algemeen kan gesteld worden dat het gaat om een cross-sectioneel onderzoek aangezien het eenmalig in de tijd is zonder interventies. Dit stelt in staat om enkel correlatieve verbanden in kaart te brengen en geen causale.

In deze masterproef worden ten slotte de zes stappen van Braun en Clarke (2006) gebruikt om de thematische analyse op een systematische en transparante wijze uit te voeren: ten eerste dient de onderzoeker vertrouwd te geraken met de data, waarbij de interviews op een actieve manier afgenomen en getranscribeerd worden. Van Hove en Claes (2011) beschrijven dit als het stadium tussen de mondelinge verzameling van de data en de analyse. De tweede stap is het genereren van initiële, open codes waarbij de gegevens opgedeeld worden in compacte, betekenisvolle eenheden. Hierbij worden labels toegekend aan stukken tekst in de data. Vervolgens worden thema's gezocht in de initiële codes en worden de data gesorteerd (Van Hove & Claes, 2011). Tijdens de vierde stap worden deze thema's herzien en verfijnd (Corbin & Strauss, 2008). Er wordt een duidelijke en overzichtelijke thematische map gecreëerd, waarbij de essentie per thema wordt genoteerd. In de vijfde stap worden de thema's benoemd en concreet gedefinieerd. Ten slotte dient een verslag geschreven te worden waarin het resultaat van de analyses per thema wordt beschreven. Hier worden citaten ter illustratie bijgevoegd. Voor het coderen, namelijk stap twee tot en met vijf, wordt het programma Nvivo gebruikt. In Bijlage 6 is de codeboom opgenomen.

3.3. Respondenten

In dit onderzoek worden zowel leerkrachten als leerlingen in het Vlaams secundair onderwijs onderzocht. Om deze actoren te bevragen, wordt gebruik gemaakt van een gelegenheidssteekproef (*convenient sample*) aangezien er veel potentiële leerkrachten en leerlingen zijn die geïnterviewd zouden kunnen worden, maar wegens praktische overwegingen is het opportuun enkel de leraren te selecteren die het best en snelst bereikbaar zijn. Daarom worden leerkrachten uit twee secundaire scholen gecontacteerd – namelijk het Onze-Lieve-Vrouwcollege te Bevegem (O.L.V.C.) en het Oscar Romerocollege te Dendermonde (O.R.C.). Het enige gehanteerde inclusiecriteria ter afbakening van het steekproefuniversum is dat zowel de leerkrachten als de leerlingen zich in het secundair onderwijs dienen te bevinden (Robinson, 2013).

De leerkrachten worden allen via mail gecontacteerd en de leerlingen worden bereikt dankzij het feit dat één leerkracht die lesgeeft aan de gehele derde graad van het O.L.V.C. de vragenlijsten tijdens de lessen wil voorleggen aan 4 random geselecteerde klassen. Op die manier wordt het mogelijk om in één stap de antwoorden van een zeer grote steekproef te verzamelen. Dankzij beide gegevensbronnen kan het perspectief van de verschillende stakeholders op het onderzoeksthema gekaderd worden (Harding & Gantley, 1998).

Het onderzoek omvat semi-gestructureerde interviews met 10 leerkrachten alsook vragenlijsten met open vragen voor 89 leerlingen in het Vlaams secundair onderwijs. Dit waren oorspronkelijk 90 leerlingen maar door individuele technische problemen tijdens de afname is er uitval van één leerling. Er werden aanvankelijk 13 leerkrachten gecontacteerd maar 3 ervan bedankten. 9 leerkrachten zijn tewerkgesteld in het O.L.V.C. en 1 leerkracht in het O.R.C. De vragenlijsten worden voorgelegd aan 75 leerlingen in het O.L.V.C. en 14 leerlingen in het O.R.C. Ter rechtvaardiging van deze steekproefgrootte kan aangehaald worden dat voor fenomenologisch onderzoek ongeveer 6 tot 15 participanten wordt aangeraden (Marshall et al., 2013).

3.4. Betrouwbaarheid en validiteit

Aangezien deze studie een cross-sectioneel onderzoek is met een gelegenheidssteekproef, zijn de verkregen resultaten niet generaliseerbaar. Er wordt namelijk eerder een exploratief beeld ter aanzet van verder onderzoek gegeven. Zo is er weinig tot geen externe validiteit. Om de interne validiteit te maximaliseren, is er sprake van zowel triangulatie van gegevensbronnen (namelijk data van zowel leerlingen als leerkrachten) als van dataverzamelmethode (namelijk semi-gestructureerde interviews en vragenlijsten) (Carter et al., 2014). Ook wordt het onderzoeksopzet zo gedetailleerd mogelijk beschreven en worden bij het luik 'Resultaten' steeds de frequenties meegedeeld om te vermijden dat anekdotes geselecteerd worden die de conclusies ondersteunen, maar eigenlijk niet frequent voorkomen (Baarends & Simon, 2017). Dergelijk onderzoek zou volgens Anderson en Herr (2011) meer valide zijn naarmate onder andere de onderzochte partijen door het onderzoek een beter begrip krijgen van het onderliggende probleem (*catalic validity*). De gebruikte wetenschappelijke termen worden namelijk vooraf reeds uitgebreid toegelicht aan de participanten zodat er geen ruimte ontstaat voor eventuele verwarring.

Voorafgaand aan de interviews en vragenlijsten wordt door elke partij een schriftelijk informatie- en toestemmingsformulier ondertekend, met een extra exemplaar voor de ouders

van de leerlingen. Blanco versies hiervan zijn te vinden in Bijlage 3, 4 en 5. De betrouwbaarheid van een studie is daarenboven gelijk aan de mate waarin de onderzoeker in staat is om categorieën te formuleren en te definiëren zodat andere codeurs in staat worden gesteld om tot een eenduidige beslissing te kunnen komen over welke items tot een bepaalde categorie behoren. Dit wordt ook de inter-codeurbetrouwbaarheid genoemd. Om hieraan tegemoet te komen, wordt één van de medestudenten gevraagd de gedefinieerde codes na te kijken. Er wordt voor deze studie echter voornamelijk beroep gedaan op de intra-codeurbetrouwbaarheid. Dit geeft een indicatie van de mate van overeenkomst tussen de uitkomsten van dezelfde metingen door dezelfde onderzoeker. Enkel de thema's van de interviews met de leerkrachten zijn systematisch opnieuw gecodeerd. Een intra-codeurbetrouwbaarheid van 1 is perfect, een van 0,6 is normaal (McHugh, 2012). Om de intra-codeurbetrouwbaarheid van dit onderzoek te visualiseren, is volgende tabel opgemaakt (zie Tabel 1). De cijfers onder 'Codering 1' en 'Codering 2' geven aan aan hoeveel codes de verschillende thema's binnen één interview gekoppeld zijn. De gebruikte verhouding is het aantal overeenstemmende categorisaties tot het totale aantal categorisaties. Voor dit onderzoek wordt een intra-codeurbetrouwbaarheid van 0,5 gerapporteerd.

Tabel 1: intra-codeurbetrouwbaarheid van data 'Leerkrachten'

Respondent	Codering 1	Codering 2	Overeenkomst
R1	9	10	0
R2	11	12	0
R3	11	11	1
R4	11	11	1
R5	11	11	1
R6	7	7	1
R7	8	13	0
R8	10	10	1
R9	10	12	0
R10	7	9	0

4. Resultaten

4.1. Wat zijn de ervaringen van leerlingen met de door hun leerkrachten gegeven toetsen en feedback?

Eerst en vooral geven 49 van de 89 leerlingen aan dat ze voornamelijk summatieve toetsen dienen te maken met open vragen en zonder dat er achteraf iets met de antwoorden gedaan wordt. Wat hierbij vaak naar voren komt uit de antwoorden, is dat er vaak verwarring ontstaat bij de open vragen aangezien vele leerlingen moeite ondervinden met het achterhalen wat precies van hen verwacht en gevraagd wordt. De overige 40 leerlingen nuanceren dit door aan te geven dat de 'soort' toets vooral vakafhankelijk is, waarbij vakken als Filosofie, Godsdiens en Esthetica naar voren komen als vakken waarbij meer ruimte is voor bijvoorbeeld creatieve vragen en vragen waarbij linken dienen gelegd worden tussen andere leerstof, de actualiteit en de eigen mening van de leerlingen.

“Open vragen geven meer persoonlijke vrijheid: iedereen zal een ander antwoord hebben. Het geeft de kans om een eigen interpretatie of mening in je antwoord te steken wat ervoor zorgt dat leerlingen bij het studeren verdergaan dan enkel reproduceren, maar ook begrijpen. Bij meerkeuzevragen primeert de kennis en minder het begrip.” (LL15)

Wanneer de leerlingen gevraagd wordt of ze reeds vertrouwd zijn met *self-assessment*, geeft 95% van hen aan dit zowel te kennen als reeds ingevuld te hebben. Slechts één derde van de steekproef vindt echter dergelijke *self-assessments* nuttig. De meest voorkomende stellingen zijn dat de *self-assessments* voornamelijk bevroegd werden in de eerste graden van het secundair onderwijs – wat de meeste leerlingen overigens ook goed vinden aangezien ze beweren zichzelf nu wel goed te kunnen inschatten – en dat het vooral bevroegd wordt na examens en na spreekbeurten. Een volgende uitkomst van de bevraging is dat veel leerlingen denken dat wat ze invullen op de *self-assessment* van invloed is op hun resultaat, daarom vullen velen met opzet foutieve informatie in – bijvoorbeeld dat ze meer hebben gestudeerd of dat ze mindere punten verwachten.

“Meestal na een spreekbeurt of bij een examen moeten we een zelfevaluatie invullen, ik vind dit niet zo nuttig. Ik doe het ook niet graag, ik heb liever feedback van de leerkracht met de nodige tips. Bij een zelfevaluatie is het vaak ook moeilijk om kritisch te zijn over jezelf, want je wilt uiteraard goede punten. Soms moeten we op de zelfevaluatie ook vermelden hoe lang we ervoor geleerd hebben en dat vind ik wel goed. Hierdoor kan de leerkracht namelijk weten of je effectief moeite hebt gedaan voor het vak, want soms zijn je punten slecht ookal doe je enorm je best. Maar dit heeft ook een nadeel, ik weet dat er veel van mijn klasgenoten hierover liegen, waardoor dit een vertekend beeld kan geven.” (LL25)

Eén leerling haalt nog een belangrijke en terechte reden aan waarom veel leerlingen *self-assessments* niet nuttig vinden om te maken:

“Na heel veel examens moeten we een beoordeling doen van ons examen. Ik vind dit echter na een examen iets minder nuttig dan na een toets. Want na een toets kan je hier nog uit leren voor je examen. Wanneer je dit na een examen doet, is dit dus iets minder nuttig aangezien je deze leerstof niet meer opnieuw moet leren.” (LL24)

De leerlingen die expliciet het nut van *self-assessments* aanhalen, vinden het voordeel ervan vooral dat zo hun voorbereiding en context beter in kaart worden gebracht alsook dat ze kunnen zien hoe goed ze in staat zijn om zichzelf en hun werk te kunnen inschatten. Ook helpt het om concreet aan te wijzen welk deel van de leerstof nog bijschaving vereist en wordt een (zelf)kritische geest getraind waardoor je ook jezelf en je sterke en zwakke punten beter leert kennen.

Op vlak van *ipsative assessment* geven slechts 5 van de 89 leerlingen aan hun punten nog nooit (on)bewust vergeleken te hebben met peers of met hun eigen vorige resultaten. De voornaamste reden die deze leerlingen hiervoor aanduiden, is dat punten überhaupt niet kunnen vergeleken worden omdat enerzijds het niveau van de toetsen verschilt alsook het feit dat elke leerling over verschillende capaciteiten beschikt.

“Voor mij is het het belangrijkste dat ik de leerstof snap. Of ik nu een hoger of een lager punt heb gehaald maakt mij in principe niet zo veel uit. Het heeft volgens mij niet veel nut om te vergelijken met vorige toetsen aangezien het niveau kan variëren. Een mediaan is handig om je te situeren binnen de klasgroep, maar dit heeft volgens mij weinig waarde. Mochten er meer zwakke of sterke leerlingen bijkomen in de klasgroep zou je positie veranderen ten opzichte van de mediaan en het gemiddelde, maar dat wil niet zeggen dat je daarom beter of slechter gescoord hebt.” (LL1)

Hier tegenover staat een groep van 50 leerlingen die zichzelf systematisch vergelijkt met diens peers en/of het klasgemiddelde. De voornaamste redenen die hieraan gekoppeld worden zijn het feit dat dat het enige ‘objectieve’ referentiepunt is om de moeilijkheidsgraad van een bepaalde toets na te gaan. Zo zien de leerlingen namelijk of enkel zij het een moeilijke toets vonden en zien ze ook waar ze moeten bijsturen. Ze vinden het vooral belangrijk om te weten of een slecht punt enkel aan zichzelf te wijten is of aan de moeilijkheidsgraad van de vragen. Verder vinden sommigen het handig om hun positie in hun klas- of vriendengroep te bepalen en geeft het een leuk gevoel wanneer ze boven het klasgemiddelde scoren.

“Ik vergelijk mijn punten wel met de anderen in de klas omdat ik het wel belangrijk vind om te weten waar ik sta. Natuurlijk is dat iets dat enerzijds goed is en dat je moet doen om te weten waar je staat en wanneer het nodig is om meer te studeren. Anderzijds is dat iets wat heel frustrerend kan werken als je merkt dat je punten altijd minder zijn dan de anderen.” (LL15)

Daarnaast wordt de groep leerlingen besproken die hun resultaten enkel vergelijkt met hun eigen vorige resultaten. Dit is de groep die het meest aanleunt bij het principe van *ipsative assessment* aangezien hier enkel gekeken wordt naar de individuele leerwinst. Concreet gaat het om een totaal van slechts 6 leerlingen. Die rapporteren onder andere dat ze zelf kunnen inschatten in welke categorie ze het vaakst scoren en wat ze ongeveer aankunnen, dus dat ze het het belangrijkste vinden dat die punten steeds in dezelfde lijn liggen.

“Ik focus vooral op mezelf. Doe ik het beter of slechter, moet ik mijn studiemethode aanpassen...? Natuurlijk kijk ik ook wel waar ik mij bevind in de klasgroep, maar dat vind ik een deel van de zelfreflectie. Als ik een 7 heb gehaald voor een bepaald vak waar ik normaal maar een 6 haal, maar de rest van de klas die normaal ook een 6 haalt, haalt nu een 10; dan moet ik voor mezelf wel reflecteren waar het fout ging.” (LL5)

De overige 28 leerlingen vergelijken zichzelf zowel met hun peers als met hun eigen vorige resultaten. Hun verklaring hiervoor is dat het vooral een dubbel en contextafhankelijk verhaal is: ze kijken namelijk vooral naar de punten van hun klasgenoten wanneer ze gebuisd zijn – om te zien of iedereen het even moeilijk vond – of wanneer het gaat om een vak waar ze echt goed op willen scoren. Daarnaast geven velen ook aan dat het na verloop der jaren vooral een automatisme is geworden en dat beide informatiebronnen gecombineerd tot de beste inschatting van een prestatie leiden. Wanneer ze dan vooral kijken naar hun eigen punten, is dat vaak als het gaat om ofwel een vak waar ze standaard mindere of net betere punten voor halen.

“Ik hecht eigenlijk aan alle twee veel belang. Persoonlijke vooruitgang is natuurlijk belangrijk als ik zie dat mijn inzet een beetje mindert en mijn punten daardoor ook, dan weet ik dat ik mezelf nog eens moet herpakken. Helaas hecht ik wel het meeste belang aan het vergelijken met de rest van de klas waardoor ik daar wel een beetje down kan van zijn. Ik weet nochtans dat het belangrijker is om naar mezelf te kijken, maar ik vergelijk mezelf toch elke keer weer met anderen.” (LL16)

Vervolgens worden de leerlingen bevraagd over de feedback die ze ontvangen van hun leerkrachten, meer bepaald in welke vorm die het vaakst geïmplementeerd wordt alsook of ze het als nuttig percipiëren. Op vlak van vorm werden voor de bevraging drie brede categorieën onderscheiden: bestaat de feedback louter uit het doorstrepen en/of aanduiden van fouten, bestaat die uit het achteraf klassikaal overlopen en vragen beantwoorden of wordt er systematisch persoonlijke feedback voorzien per leerling. Deze laatste categorie wordt door slechts 8 leerlingen aangeduid als vaakst voorkomende vorm van feedback. Dit uit zich vooral in het opschrijven van het juiste antwoord of wanneer er een hele slechte score behaald wordt, schrijven de meeste leerkrachten ook tips op zoals ‘herhaal de theorie’. Vervolgens volgen het aanduiden van fouten (n=35) en het klassikaal overlopen en/of vragen beantwoorden (n=46), maar deze laatste twee worden wel vaak samen gezien. Hieruit komt naar voren dat er enkel iets zal worden bijgeschreven wanneer het volledig fout is. Volgende leerling vat het adequaat samen: “Er staat meestal wel wat er fout is maar niet waarom het fout is” (R19). Over het algemeen verwachten de leerkrachten dat de leerlingen zelf hun vragen en onduidelijkheden communiceren. Dit vinden de leerlingen echter geen groot probleem; ze zijn namelijk van mening dat de beste feedback toch mondeling wordt gegeven en dat elke leerkracht openstaat om vragen (klassikaal) te beantwoorden, hoewel dat soms een drempel kan zijn.

“Dat hangt ook af van leerkracht tot leerkracht. Sommige leerkrachten schrijven het juiste antwoord ernaast en anderen zullen de toets na het teruggeven overlopen en de juiste antwoorden bespreken in de klas. Dit is meestal bij de niet-wetenschapsvakken, bij de wetenschapsvakken gaan ze er van uit dat we het kunnen terugvinden in onze cursus aangezien het soortgelijke oefeningen zijn. Enkel als we erachter vragen bespreken we deze toetsen. Er staat zelden feedback op de toets, bij oudere leerkrachten zie je wel vaker feedback staan. Bij ons wordt er meestal feedback geschreven op het rapport zelf, maar zelfs dan is deze feedback redelijk algemeen en niet echt gericht naar de leerling zelf. Vaak hebben meerdere leerlingen in de klas hetzelfde tekstje op hun rapport.” (LL23)

Ten slotte wordt gekeken naar welke waarde de leerlingen hechten aan de feedback van hun leerkrachten. De resultaten spreken alleszins voor zich: slechts 16 leerlingen geven aan de

feedback bruikbaar en duidelijk te vinden en 31 leerlingen zeggen expliciet dat ze de feedback te vaag en dus onbruikbaar vinden. De overige leerlingen hebben zich hier niet over uitgesproken. De leerlingen die zich kunnen vinden in de voorziene feedback, zeggen dat het voor hen voldoende is. Wanneer ze toch niet tevreden zijn met de feedback van de leerkracht, is alles meestal opgelost met een vraag tijdens of na de les of met wat opzoekingswerk in de cursus. De leerlingen die momenteel eerder ontevreden zijn over de gegeven feedback, geven aan dat de leerkrachten te onduidelijk schrijven en dat alles te oppervlakkig en te algemeen geformuleerd is. Ook zeggen velen het een drempel te vinden om een vraag te stellen aan de leerkracht omdat ze niet als 'dom' willen bestempeld worden.

4.2. Wat zijn de ervaringen van leerkrachten met de door hun gegeven toetsen en feedback?

Qua brede vormen van toetsen geven drie van de tien leerkrachten aan voornamelijk voor summatief te opteren. Wat hierbij opvalt is dat niet voor deze toetsvorm gekozen wordt omwille van de persoonlijke pedagogische voorkeur, maar eerder omwille van praktische elementen zoals het aantal lessen, de grootte van de klasgroep en werk- of tijdsdruk. Over het algemeen ziet de meerderheid de waarde van de formatieve toetsing dus wel in en voelen ze dat ze moeten 'vernieuwen', maar valt ook op dat vooral de jongere leerkrachten moeite ondervinden met formatief toetsen aangezien het tijdsintensiever is en ze nog niet zo ingewerkt zijn als hun oudere collega's.

"(...) Dus heel veel formatief heb ik eigenlijk nog niet gedaan gezien het feit dat ik nog niet zo lang ben afgestudeerd en gewoon omdat er heel veel tijd in kruipt. Ik begrijp wel het principe van dat je formatieve toetsen geeft omdat je die punten dan meeneemt in een soort van evaluatie en bijles en zo. Maar gewoon ja, er kruipt zo veel tijd in en in het begin heb je andere dingen te doen. Dus als je naast toetsen die meetellen ook nog eens extra toetsen of oefenbundels of whatever moet opstellen, dat heb ik eigenlijk nog niet gedaan. Los daarvan kunnen de leerlingen altijd op bijles komen of uitleg vragen of zo. Ik weet dat formatief zo meer gezien wordt als goed omdat je dan meer de evolutie ziet en omdat je dan feedback en zo kan geven. Ik vind het gemakkelijker om dat te zeggen dan om dat eigenlijk effectief te doen." (LK9)

Twee leerkrachten rapporteerden dat ze resoluut kiezen voor formatieve toetsing. Opnieuw valt ook hierbij op dat het gaat om iets oudere leerkrachten die reeds vaak hebben kunnen experimenteren met verschillende toetsvormen bij verschillende leerlingen. Het verschil in de werking van deze leerkrachten en de leerkrachten die aangeven voor een mengvorm te kiezen (n=5), is dat het formatieve bij deze twee echt systematisch aanwezig is op elk toetsmoment en dat er ook consequent extra bijlessen worden ingelast en oefenbundels worden uitgedeeld.

"(...) maar dat formatieve zit dan in mijn zelfevaluaties, mijn feedback, mijn checklists... Daar zit dan het formatieve in om het dan mee te nemen naar een volgende toets, want ze moeten dan ook invullen wat er misliep tijdens de vorige toets en of ze daar nu dan rekening met hebben gehouden of niet." (LK6)

De overige helft van de bevroegde leerkrachten in de steekproef vermeldt dat het kiezen voor een toetsvorm vooral resulteert in een mengvorm. Dit is hoofdzakelijk afhankelijk van het vak of het specifieke stuk leerstof dat zich op dat moment dan het best tot een bepaalde toetsvorm

leent. Daarnaast worden beide toetsvormen door een deel van deze groep ook beschouwd als een aanvulling op elkaar. Samengevat kan tot dezelfde vaststelling gekomen worden als wat reeds betoogd wordt bij de leerkrachten die voor de summatieve vorm kiezen, namelijk dat ze van mening zijn dat het vooral een kwestie is van praktische elementen. Alle vijf de leerkrachten die gecodeerd worden onder 'mengvorm' zijn namelijk van mening dat het sterk afhangt van welk vak je geeft, en dat formatief eerder gezien wordt als een extra aanvulling waarbij het summatieve dan vooral het fundament vormt.

“Wat de bovenhand neemt, en dat is ook normaal denk ik in het onderwijs, is het summatieve. Maar dat wil natuurlijk niet zeggen dat wij met formatief niet bezig zijn. Ze moeten bijvoorbeeld een recensie schrijven voor Nederlands en die is ook het eindproduct logischerwijs, maar als tussenstap moeten ze al eens een paar alinea's schrijven van een middenstuk van een recensie. Dan geven ze dat af aan mij en dan schrijf ik daar feedback op, van 'dat is een goed idee, dat niet' en taalkundig ook. Het bijsturen na een toets en het geven van remediëringsoefeningen gebeurt ook uiteraard. (...) Dat is dan ook formatief zoals jij het ziet, denk ik.” (LK7)

Vervolgens heeft een duidelijke meerderheid van acht van de tien leerkrachten al eens *self-assessment* toegepast in diens klaspraktijk na een toetsmoment. Wel valt op dat er binnen deze groep een verschil bestaat tussen hoe dergelijke *self-assessment* effectief geïmplementeerd wordt. De helft van de leerkrachten integreert het namelijk systematisch schriftelijk na elke toets of examen, de andere helft overloopt mondeling eens welk gevoel de leerlingen hebben na de evaluatie. Wel komt bij iedereen terug dat dit sowieso schriftelijk aanwezig dient te zijn na elk examen en dat het daar bovendien uitgebreider is. Sommige leerkrachten voegen naast de standaard zelfevaluatievragen zelfs nog vragen toe met betrekking tot het niveau en de lengte van het examen.

Hoewel de grote meerderheid van de leerkrachten aangaf *self-assessment* te gebruiken, benoemt slechts één derde van de steekproef expliciet de voordelen ervan. De voordelen die het vaakst naar voren komen uit de interviews zijn vooral van diagnostische aard, namelijk om na te gaan waar enerzijds het probleem van de toets ligt en waar anderzijds het probleem van (de studiemethode van) de leerling ligt. Met behulp van deze inzichten kan er namelijk adequater ingespeeld worden op toetsmomenten in de toekomst.

“De voordelen zijn uiteraard dat het voor mij als leerkracht heel wat eenvoudiger wordt om te zien wanneer het misloopt en waar het mis is gegaan. Op die manier kan je achteraf veel gemakkelijker gaan remediëren. (...) Dus op basis van die zelfevaluatie kan je eigenlijk heel gericht gaan werken. (...) Het heeft alleszins heel veel voordelen voor ons om een inzicht te krijgen in wanneer het niet goed gaat, anderzijds denk ik ook dat het nuttig is voor de leerling om zichzelf te evalueren en na te denken over hoeveel die er heeft ingestopt en hoeveel diens prestatie waard is.” (LK2)

Waarom de overige zes leerkrachten vooral de nadelen benoemen van *self-assessment*, komt niet zozeer door de ideologie van het concept, maar eerder door – opnieuw – praktische barrières zoals tijdsgebrek en vooral de leeftijd van de leerlingen. Zo wordt meermaals aangegeven dat jongere leerlingen zichzelf en hun werk nog niet correct kunnen inschatten en soms zelfs niet eerlijk zijn in hun zelfbeoordeling. Vooral de sterkere leerlingen zullen zichzelf eerder onderschatten en de zwakkere leerlingen zullen zichzelf dan eerder overschatten.

Bovendien is er een algemene shift in deze vaardigheid waargenomen bij de ondervraagde leerkrachten na de coronaperiode, waarbij de meeste leerlingen zichzelf nu systematisch zwakker zouden inschatten – al dan niet wegens (toegenomen) perfectionisme. Wel geldt de consensus dat hoe ouder leerlingen zijn, hoe beter ze zichzelf kunnen beoordelen.

“Soms vond ik dat wel heel nuttig als leerlingen dat correct invulden en dat goed deden, maar je merkt wel dat degene bij wie dat dat niet goed is of die dat niet leren, die kruisen altijd hetzelfde aan. Dan zit je met de feedback die je dan baseert op die zelfbeoordeling, maar dan zit je altijd in hetzelfde straatje.” (LK9)

Als laatste toetsvorm zal *ipsative assessment* besproken worden, wat naar boven komt als element waarover de meeste consensus te vinden is. Er is namelijk slechts één leerkracht van de tien die het hele concept wat overroepen vindt en het net een meerwaarde vindt om als leerling te weten waar je je bevindt binnen de groep. De overige leerkrachten erkennen echter wel het belang van dergelijk denkbeeld aangezien elke leerling uniek is en zich middenin een groeiproces bevindt, dat niet vergeleken kan worden. Ook wordt *ipsative assessment* hier louter geïmplementeerd als rijke achtergrondinformatie die aangewend wordt tijdens bijvoorbeeld deliberaties en individuele feedback, eerder dan als een structureel instrument binnen de onderwijspraktijk. Zo worden er bijvoorbeeld geen extra punten of andere voordelen toegekend aan de leerling met de grootste individuele leerwinst. Verder gaven alle leerkrachten aan dat op de maandrapporten geen klasgemiddeldes meer staan, net om de vergelijking van leerlingen te vermijden. Ook hier kan de meerderheid zich in vinden.

“Ik probeer wel behalve summatief anders naar leerlingen te kijken. Ik vind dat een leerling een totaalpakket is, dat is niet alleen punten die je daarop kleeft. (...) Ik heb al leerlingen doorgelaten van wie ik weet dat het moeilijk gaat worden, maar waarvan ik denk ‘die hebben wel een hart voor die taal, ik ga dat toch proberen’. Ze groeien ook nog zo veel, dat is een groeiproces hè bij een kind. Je kan ze niet straffen omdat ze iets geprobeerd hebben. (...) Dat zeg ik al jaren, dat je in godsnaam nooit mag vergelijken met wat je vrienden hebben van punten. (...) Ik heb dat wel moeten leren, dat er nu leerlingen zijn bij wie het moeilijker gaat. Nu weet ik dat ook de achtergrond een rol speelt. In de eerste plaats vind ik altijd dat een leraar de leerlingen graag moet zien en dan pas zijn les mag geven, en als je uw leerlingen graag ziet, dan ga je niet ieder kind vergelijken. (...) Hoe zot is dat, dat ik die punten nodig heb om te weten wat een kind waard is. Dat is niet zo.” (LK3)

Hoewel 90% van de leerkrachten zich kan vinden in de ideologie van *ipsative assessment*, benoemt slechts 50% van de steekproef expliciet de voordelen ervan. De meest voorkomende baten van bijvoorbeeld het afschaffen van de klasgemiddelden zijn dat het zeer demotiverend is voor leerlingen die zich steeds onderaan de groep bevinden en dat de meeste ouders te hard focussen op dat cijfer en het daardoor uit hun context trekken. Op één school wordt alvast het goede voorbeeld gegeven door telkens onderaan de rapporten te vermelden dat de punten enkel dienen om een globaal zicht te krijgen op de prestaties van de leerling en dat het belangrijk blijft om te weten wat de capaciteiten van die leerling zijn. Als nadeel wordt vooral vernoemd dat het niet haalbaar en verdedigbaar is om dergelijke ideologie structureel te implementeren in het onderwijs alsook dat leerlingen moeten leren omgaan met het feit dat er van nature verschil zit op talent en capaciteit.

“Ik bedoel, je krijgt uw cijfer, maar cijfers moet je altijd interpreteren. Een cijfer op zich zegt niets. Dus wat ik wel vind, en daar ben ik wel ook een beetje mee bezig, dat is dat een cijfer alleen maar waarde krijgt op een rapport als daar feedback bij staat. En ook samen met het klasgemiddelde. Nu niet dat dat alles makend is, maar toch zodanig dat een leerling of een ouder dat cijfer kan interpreteren. Als je daar een 6/10 ziet staan en een klasgemiddelde van 9/10, maar er staat bij ‘zeer goed gewerkt, dit is een heel mooi resultaat’, dan weet jij als ouder van oké, mijn kind is niet sterk in dat vak, maar blijkbaar is dit wel een heel goed cijfer voor mijn kind.” (LK8)

Tot slot wordt besproken hoe de ondervraagde leerkrachten feedback implementeren in hun klaspraktijk. Ook hierbij is consensus te vinden, namelijk dat elke ondervraagde leerkracht diens leerlingen voorziet van feedback. Het enige verschil zit in de frequentie en de diepgang ervan, want twee leerkrachten geven namelijk aan slechts sporadisch feedback te geven. Meer bepaald gaan ze dan enkel in op de leerstof indien de hele klas slecht heeft gepresteerd of wanneer de leerlingen er zelf specifiek om vragen, anders niet. Indien er iets op een toets wordt geschreven, is het vaak zeer algemene feedback zoals ‘goed gedaan’. De overige leerkrachten die wel systematisch feedback integreren, doen dit op verschillende manieren: een verbeterleutel posten op SmartSchool, een kader bovenaan de toets waarin de leerkracht aandachtspunten kan afvinken en/of noteren, (online) vragenuurtjes na de lessen, klassikale verbetering, bijlessen en remediëringsoefeningen... Wel noteren veel leerkrachten nog steeds enkel iets op een toets wanneer het echt niet goed is, maar dan staat er wel gedetailleerd bij welk deel ze best nog eens opnieuw studeren en wat ze verder nog kunnen ondernemen om te slagen. Eén leerkracht benoemt wel steevast het goede zodat diens leerlingen ook beter kunnen inschatten welk niveau van hen verwacht wordt.

“Feedback is meer in de aandacht gekomen. Weet je, toen ik begon zo’n twintig jaar geleden, dan was dat eerlijk gezegd geen item. Dat is een aantal jaren geleden stilaan in de aandacht gekomen, en terecht hè. Vanuit de school wordt daar nu ook meer op gehamerd. Ik zit ook in aanvangsbegeleiding en dat is ook iets dat wij nu naar de beginnende leerkrachten ook sterk benadrukken dat ze rapportcommentaren moeten zetten. Eerlijk gezegd is dat wel nog niet bij alle collega’s doorgedrongen. (...) Ik heb jaren geleden al van een pedagogisch begeleider te horen gekregen dat je niet enkel commentaar mag zetten bij de slechte punten, dat mag overal. Maar sommige leerkrachten beginnen zelfs nog niet eens bij die slechte punten.” (LK7)

4.3. Wat zijn de ervaringen van leerlingen met testangst?

Om te beginnen geven 11 van de 89 leerlingen aan helemaal geen schoolgerelateerde stress te ervaren. Ze vertrouwen erop voldoende hun best gedaan te hebben en voldoende voorbereid te zijn, en relativiseren het belang van een toets. De leerlingen die naar eigen zeggen soms last ervaren van de gevolgen van testangst vormen de grootste groep met een totaal van 44 leerlingen. Zelf beschrijven ze dit echter als ‘gezonde stress’, namelijk wanneer ze naar eigen aanvoelen niet voldoende geleerd hebben of wanneer het een toets is voor een (hoofd)vak waarop ze al vaak gebuisd zijn. Ook geven velen aan vooral stress te ervaren voor vakken waarbij niet veel toetsen worden gegeven, tijdens examens of tijdens herkansingstoetsen. Daarenboven willen ze zichzelf, hun ouders of hun leerkrachten niet teleurstellen. Bij de 34 leerlingen die vaak testangst ervaren, wordt nog meer de nadruk gelegd op druk vanuit de omgeving en op het zelf hoog leggen van de lat. Velen hebben uit

perfectionisme een probleem met goed timemanagement waardoor ze nooit echt voldaan zijn na het studeren, wat de stress vlak voor een toets verhoogt.

“Ik ervaar bij een toets zeer veel druk. Dit komt door meerdere dingen. Ten eerste wanneer het gewoon moeilijke leerstof is, of wanneer de vorige toets slecht was en ik dus nu het gevoel heb dat het wel goed moet zijn. Maar ook door medeleerlingen ervaar ik stress, want als zij stressen voor iets dan heb ik het gevoel dat ik daar ook stress voor moet hebben. Maar ook door mijn ouders. Ze worden niet per se boos als ik een slecht punt heb, maar ze vinden het ook niet leuk. Mijn broer en zus zijn heel slim en hadden altijd goede punten, en ik minder. Hierdoor heb ik het gevoel dat ik ze niet mag teleurstellen.” (LL18)

Met betrekking tot hoe het ervaren van testangst zich concreet uit bij leerlingen, rapporteren 21 leerlingen hieraan een lager zelfbeeld overgehouden te hebben. Dit werkt echter in een vicieuze cirkel met het gegeven dat ervaringen met testangst al hebben geleid tot lagere schoolresultaten (n=38). Dat kan verklaard worden door het feit dat ze zich minder goed kunnen concentreren door de stress tijdens een toetsmoment, wat leidt tot het ervaren van een black-out en tijdsnood waardoor vermijdbare fouten sneller gemaakt worden, wat opnieuw resulteert in meer stress. Hierdoor voelen ze zichzelf automatisch ook meer onzeker over hun academische capaciteiten. Daarenboven wordt regelmatig aangegeven dat ze zich overrompeld voelen door de werklast waardoor ze minder tijd, energie en motivatie hebben om grondig te studeren. Ten slotte kampen 30 leerlingen met (milde) medische problemen zoals paniekaanvallen, hartkloppingen, migraine, tremor, huilbuien, vergetelheid, slapeloosheid, braakneigingen, prikkelbaarheid en derealisatie en dissociatie.

“Ik kan moeilijk slapen en herhaal alle informatie. Dan lig ik wakker omdat ik dingen niet meer kan herinneren. Op school is dit door prikkelbaarder te zijn en minder te kunnen genieten van gezelschap van vrienden. Je doorstaat de lessen op automatische piloot tot de toets en daarna neemt de druk wat af. Echter heb ik soms stress overdag over de hoeveelheid leerstof die ik in de avond – op beperkte tijd – moet verwerken. Het kost moeite om je over een slecht cijfer te zetten en motivatie te vinden om eraan te werken.” (LL15)

De oorzaak die de bevroegde leerlingen toeschrijven aan het ervaren van testangst, is voor 15 leerlingen wanneer ze naar eigen aanvoelen onvoldoende gestudeerd hebben. 23 leerlingen hebben vooral last wanneer het in het algemeen gaat om een zeer grote en/of belangrijke toets en hierop aansluitend geven 45 leerlingen aan testangst vooral meer specifiek te ervaren bij een buisvak. Slechts 5 van de 89 leerlingen rapporteren dat dit veroorzaakt zou worden door een te hoge werkdruk en nog een andere cluster van 5 leerlingen rapporteert vooral tijdsdruk. Vaak voorkomende verklaringen van deze specifieke oorzaken zijn dat de leerlingen soms slechts één toets per maand krijgen voor een bepaald vak en soms bestaan toetsen uit slechts drie vragen. Veel leerlingen geven aan niet om te kunnen met deze druk die gepaard gaat met hun toekomst in eigen handen te hebben. Daarom wordt dan ook de meeste stress gerapporteerd tijdens examenperiodes of herkansingstoetsen, aangezien dit de 'laatste' kansen zijn. Ten slotte ervaren 48 leerlingen prestatiedruk vanwege hun leerkrachten (n=9), hun ouders (n=19) en/of zichzelf (n=20). Ze willen vooral hun omgeving niet teleurstellen en zichzelf bewijzen. Ook leggen ze voor zichzelf de lat hoog om later uit een breder studieaanbod te kunnen kiezen.

“Ik ervaar vrij veel stress voor school, voornamelijk in weken waarin er veel toetsen ingepland staan. Dit is omdat ik zo goed mogelijk wil presteren voor elke toets, maar dat is vaak niet mogelijk omdat ik uitgeput geraak. Vooral moeilijke leerstof of een toets over een groot onderdeel stresst mij.” (LL42)

Dit stuk wordt afgerond met de methoden die de leerlingen hanteren om zichzelf te kalmeren alsook op welke (structurele) veranderingen hen echt zouden ondersteunen. Een aantal leerlingen rapporteren kalmeerpillen te nemen en naar de psycholoog te gaan vanwege hun testangst. De meerderheid van de leerlingen verzet hun gedachten met behulp van vrienden en/of hobby's, maar dit resulteert soms in tijdsdruk waardoor er opnieuw stress ervaren wordt. Wanneer hen gevraagd wordt wat de leerkrachten zouden kunnen doen om hen tegemoet te komen in het ervaren van testangst, komt bij een overgrote meerderheid naar voren dat een betere en meer begripvolle spreiding van toetsen en meerdere kleine toetsen al een enorme hulp zouden zijn. Eén leerling oppert ook om meer taken te geven in plaats van toetsen, waardoor de leerlingen nog steeds bezig zijn met de leerstof maar met minder druk. Ook de implementatie van alternatieve assessmentvormen, zoals een filmpje thuis opnemen in plaats van vooraan 'live' te presenteren, kan rekenen op veel steun. Verder wordt ook geopperd om vlak voor een toets een meer ontspannen sfeer te creëren, bijvoorbeeld door klassikaal even te praten of met behulp van ademhalingsoefeningen.

“Leerkrachten gaan er vaak van uit dat we dingen kennen en kunnen, wat ons extra prestatiedruk geeft. Vaak beseffen ze ook niet hoe veel werk we in hun vak of in een ander vak moeten steken en blijven ze maar werk geven waardoor veel leerlingen (waaronder ikzelf) slaap moeten opofferen om alles bolgewerkt te krijgen met werk en hobby's erbij. Het is niet gezond dat sommige leerlingen zo lang bezig zijn met hun schoolwerk. Er moet ook nog plaats zijn voor sociaal leven en ontspanning.” (LL80)

“School laat leerlingen te veel focussen op het behalen van goede punten, waardoor leerlingen niet meer naar school gaan door de menselijke leergierigheid.” (LL89)

4.4. Wat zijn de ervaringen van leerkrachten met testangst bij hun leerlingen?

Voor deze onderzoeksvraag wordt eerst gepeild naar een inschatting van de leerkrachten of ze de indruk hebben dat het aantal gevallen van testangst gestagneerd is dan wel toegenomen, om daarna te peilen naar welke oorzaken volgens hen aan de basis zouden kunnen liggen van het ervaren van testangst. Dat de eerste deelvraag een subjectief antwoordt inhoudt, wordt sterk weerspiegeld in de cijfers. Tijdens het coderen is namelijk gebleken dat vijf van de tien leerkrachten vinden dat het aantal leerlingen met testangst gestagneerd is en de overige vijf vinden dat dit aantal de laatste jaren is toegenomen. Niemand vermeldt dat het zou afgenomen zijn.

Volgens de vijf leerkrachten die een stagnatie rapporteren, zijn er nu misschien wel iets meer gevallen van leerlingen met een zorgbehoefte, maar zou vooral de intensiteit van de klachten iets minder ernstig zijn dan vroeger. Dit komt ook terug op de checklists van de *self-assessments*, waarbij zo goed als elke leerling aanduidt dat die stress ervaarde tijdens de toets alsook dat een leerling nu veel sneller in een zorgklas of op een individueel traject gezet wordt met maar zelden stress of testangst als aanleiding.

“Sporadisch vroeger eentje per klas, maar echt in de eerste graad is het nu denk ik drie vierden. Maar voor alles wordt zorg ingeroepen tegenwoordig hè, en de ouders moeten daar nu ook geen bewijs meer voor hebben van een dokter of een psycholoog. Als de ouders vinden dat hun kind zorg nodig heeft, dan krijgen die dat hè.” (LK10)

De vijf leerkrachten die aangeven dat het aantal leerlingen die testangst ervaren zou zijn toegenomen, geven als voornaamste reden dat leerlingen er vroeger minder mee naar buiten kwamen en dat de afstand tussen leerkracht en leerling toen ook veel groter was. Het feit dat de problematiek toen grotendeels onder de radar bleef, neemt echter niet weg dat ook enkele recente veranderingen in de maatschappij toch ook mee hebben gezorgd voor een toename aan klachten. Zo worden de coronalockdowns, prikkels en druk van sociale media en een te hoge werkdruk van taken en toetsen en genoemd als grootste katalysatoren van stress.

“Ja, veel meer dan vroeger. Zeker en vast, ja. Dat is echt iets wat ik enorm merk de laatste jaren. Ik vul heel veel middagen met gesprekken met leerlingen omdat ik weet dat ze zich niet goed voelen, omdat dat ze in een bepaalde problematiek zitten, omdat ze faalangstig zijn, omdat ze heel veel stress ervaren... Gisteren zijn ze nog eens naar mij gekomen als klas van kijk, wij willen echt een sein geven want het is te veel, het lukt niet meer, we zijn overladen door werk. Het is zo iets dat zo groeit de laatste jaren, dat is enorm. Ik vind het eigenlijk niet meer oké, op den duur zijn wij als leerkrachten ook psychologen en wij weten niet wat wij moeten doen want wij zijn daar niet voor opgeleid. Maar die psychologische problemen, die worden groter en groter. Dat merk ik zéker. Zeker en vast.” (LK5)

De leerkrachten die aangeven dat het aantal leerlingen met kenmerken van testangst zou toegenomen zijn, sommen uit zichzelf al een aantal belangrijke oorzaken op – zoals hierboven te lezen is. In deze sectie wordt desbetreffende bijvraag door elke leerkracht meer in de diepte beantwoord. Drie specifieke factoren worden door de helft van de leerkrachtengroep spontaan aangehaald als potentiële oorzaken voor toegenomen testangst, namelijk: druk vanwege de ouders van de leerlingen, een toegenomen werklast door veranderde leerplannen en het feit dat de band tussen leerling en leerkracht nauwer is geworden waardoor er nu veel gemakkelijker over dergelijke problemen kan gesproken worden. De factoren die vervolgens door vier van de tien leerkrachten worden vermeld, zijn: de lange isolatie tijdens de coronalockdowns alsook het feit dat het zorgaanbod van scholen de laatste jaren immens is uitgebreid. Ten slotte wordt ook – eerder sporadisch – verteld dat er meer druk op de schouders van leerlingen rust door sociale media specifiek en de (prestatiegerichte) maatschappij in het algemeen.

“De vraag die je kan stellen, is: is het aantal klachten toegenomen omdat er meer mogelijkheden zijn voor de leerlingen omdat wij meer zorg aanbieden in vergelijking met pakweg tien jaar geleden; of is het aantal klachten toegenomen omdat er andere onderliggende problemen zijn zoals grotere drukte door sociale media en druk van thuis, druk van de maatschappij... eventueel ook door milieuproblematiek en luchtvervuiling waardoor er meer ADD, ADHD, ASS, dyspraxie, dyscalculie... is bijvoorbeeld.” (LK4)

4.5. Welke concrete methodieken hanteren leerkrachten als het aankomt op het verminderen van testangst bij hun leerlingen?

Hoewel er enkele tegenstrijdigheden te vinden zijn bij de vraag of het aantal leerlingen met testangst nu toegenomen is of niet, is elke ondervraagde leerkracht het er wel over eens dat elke leerling recht heeft op zorg indien nodig. Elke leerkracht uit de bevroegde steekproef haalt namelijk het gebruik van faciliteiten aan. Dit is bijvoorbeeld het laten dragen van een hoofdtelefoon tijdens een toets voor hoogsensitieve leerlingen en de tekst van een examen op voorhand laten inkijken voor dyslectische leerlingen, maar ook leerlingen de antwoorden op een laptop laten invullen, hen de vragen mondeling laten toelichten, inhaaltoetsen voorzien, extra tijd geven, een presentatie enkel geven voor de leerkracht, taken en toetsen zo goed mogelijk trachten te spreiden... Dat laatste aspect blijkt in de praktijk echter vaak niet volledig haalbaar, aangezien niet elke leerkracht hieraan wenst mee te werken. Wat door bijna iedereen aangehaald wordt, is het bieden van steun tijdens een toets of examen. Wanneer leerkrachten namelijk zien dat leerlingen het moeilijk krijgen, grijpen ze direct in om de leerling eens apart te nemen en die gerust te stellen.

Bijna de helft van de steekproef maakt extra tijd vrij tijdens of na de lessen om klassikaal of individueel eens te polsen hoe het gaat met iedereen. Ook hier wordt de praktische kant ervan aangehaald, namelijk dat dit het best werkt bij kleinere klasgroepen en dat de meeste respons voornamelijk van meisjes komt. Drie leerkrachten implementeren dit standaard, één andere leerkracht zou dit wel willen doen maar is van mening dat hier te weinig tijd voor is en pleit dus voor meer marge tussen de lessen. Daarenboven wordt gesproken over het zorgbeleid en de zorgcoördinatoren op school, die een belangrijke rol spelen in het identificeren en begeleiden van leerlingen met testangst. Vier leerkrachten geven tijdens het interview expliciet aan dat ze leerlingen naar hen doorverwijzen indien de specifieke problematiek hen te complex lijkt. Een vorm van kritiek die hierop door twee leerkrachten wordt geuit, is dat leerlingen meer inspraak zouden moeten krijgen in het kiezen met wie ze willen praten aangezien opgemerkt wordt dat deze zorgcoördinatoren niet voor elke leerling een vertrouwenspersoon zijn. Van alle interviews zijn er ten slotte twee leerkrachten die expliciet het belang van een zorgbeleid vermelden en die stellen dat er standaard een brede basiszorg zou moeten zijn voor elke leerling, terwijl één leerkracht expliciet aanhaalt dat men moet opletten voor het geven van meer ondersteuning dan werkelijk nodig is:

“Ik vind dat er een verschil is tussen leerlingen ondersteunen en leerlingen steunen. Ik bedoel, je kan ze ook heel de tijd dragen waardoor op den duur hun spieren verslappen hè. Dat is iets anders dan eventjes een hand geven of eens vasthouden. Ik sta echt wel kritisch tegenover zo'n zorgbeleid. Je moet ze inderdaad helpen waar dat het nodig is, maar de bedoeling is dat je ze helpt waar het nodig is en dat je ze niet heel de tijd vasthoudt. Het is dus vandaar dat ik geen apart examen en zo doe, ik doe daar niets speciaal voor. Maar als je echt ziet dat er daar een probleem is daardoor, ja dan help je ze wel natuurlijk.” (LK8)

5. Discussie en conclusie

5.1.1. Wat zijn de ervaringen van leerlingen met de door hun leerkrachten gegeven toetsen en feedback?

Overkoepelend komt naar voren dat leerlingen het vaakst toetsen met open vragen dienen te maken, waarbij er echter vaak sprake is van verwarring omdat ze moeite ondervinden met het achterhalen van de verwachtingen van de leerkracht. Algemeen gezien kan hierbij een nood opgemerkt worden aan een duidelijk afgebakende context en helderheid, zonder ruimte voor diverse interpretaties. Dit komt overeen met de literatuur die stelt dat een toegenomen helderheid van leerdoelen leidt tot het beter kunnen plannen, monitoren en evalueren (Andrade & Valtcheva, 2009; Brown & Harris, 2013; Jonsson, 2014; Nicol & McFarlane-Dick, 2006; Panadero & Alonso-Tapia, 2013).

Zo goed als elke leerling uit de steekproef geeft aan *self-assessment* reeds toegepast te hebben. Slechts één derde vindt het echter nuttig, vooral gezien het het vaakst en meest uitgebreid gevraagd wordt na examens, waarna er niets meer met wordt gedaan. Joosten-ten Brinke en collega's (z.d.) beweren echter dat *self-assessment* pas effectief is wanneer er ook zelfreflectie aan gekoppeld wordt die duidelijk maakt wat het resultaat betekent voor de ontwikkeling van de leerling. Opvallend hierbij is dat enkele leerlingen vertellen dat ze hun *self-assessment* opzettelijk foutief invullen omdat ze denken dat dit van invloed is op hun resultaat. Hierdoor krijgt de leerkracht vaak te lezen dat diens leerlingen weinig hebben kunnen studeren en dat ze een slecht punt verwachten. Eerder onderzoek spreekt dit tegen, waaruit namelijk blijkt dat leerlingen zichzelf net beter zouden inschatten (de Grez, Valcke & Roozen, 2012). Ten slotte wordt hierbij aangegeven dat de leerlingen het nuttiger zouden vinden mocht de leerkracht geïndividualiseerde feedback geven met de nodige tips. Williams (1992) stelt inderdaad dat leerlingen de voordelen van *self-assessment* vooral inzien wanneer hun eigen toegekende punt vergeleken wordt met het punt én feedback van de leerkracht. Concreet is er nood aan duidelijke richtlijnen vooraf en bruikbare feedback achteraf (Birenbaum & Dochy, 1996; Boud & Falchikov, 1989; Griffie, 1995).

De grootste reden waarom leerlingen hun resultaten vergelijken met hun peers, is om de moeilijkheidsgraad van een toets beter en objectiever te kunnen inschatten. Op die manier maken ze uit of het aan hen ligt en zo ja, waar ze dienen bij te sturen. Ze geven aan dat vooral de combinatie van resultaten uit het verleden met resultaten van peers zou leiden tot de beste inschatting van een prestatie. Bourke (2010) stelt inderdaad dat leerlingen *ipsative assessment* reeds uit zichzelf toepassen en daarenboven in staat zijn om zichzelf te motiveren om persoonlijke topprestaties te halen. Ruim één derde van de ondervraagde leerlingen geeft vervolgens aan de feedback van hun leerkrachten onnuttig te vinden. Zo zouden hun leerkrachten te onduidelijk schrijven en zou de boodschap te oppervlakkig en te algemeen geformuleerd zijn. Er wordt dus vaak wel aangeduid *wat* er fout is, maar niet *waarom*. Onderzoek stelt inderdaad vast dat leerlingen de feedback van hun leerkrachten vaak niet nuttig vinden, onder andere door een te vage formulering en de te lange duur tussen de toetsafname en de feedback (Carless et al., 2011; Lea & Street, 1998; Price et al., 2010; Scoles et al., 2013; Weaver, 2006). De leerkrachten zouden er van uitgaan dat de leerlingen uit eigen initiatief hun vragen en onduidelijkheden communiceren, maar veel leerlingen vinden dit een drempel omdat ze bijvoorbeeld niet als 'dom' willen bestempeld worden.

5.1.2. Wat zijn de ervaringen van leerkrachten met de door hun gegeven toetsen en feedback?

Uit de resultaten is gebleken dat voor welke toetsvorm de ondervraagde leerkrachten ook kiezen, deze keuze vooral gebaseerd is op praktische elementen eerder dan pedagogische voorkeur. Zo zouden vooral de jongere leerkrachten opteren voor summatieve toetsing gezien dit de minst tijdsintensieve optie blijkt, hoewel ze de waarde van formatieve toetsing wel

erkennen. Eerder onderzoek stelde inderdaad dat dit ethisch-pedagogisch dilemma het natuurlijk pedagogisch inzicht van de leerkracht in het gedrang brengt waardoor die minder in staat is diens lespraktijk adequaat af te stemmen op de noden van de leerlingen (Alexander, 2011; Ball, 2003; Black & William, 1998, 2006). De overige leerkrachten, die variëren tussen beide toetsvormen, rapporteren dat de keuze afhangt van het vak of het specifieke stuk leerstof. Dit wijst op het aanpassingsvermogen dat verwacht wordt van leerkrachten in de mate van sturing die verschilt per toetsvorm (Joosten-ten Brinke, 2010). Daarom worden beide toetsvormen door deze groep dan ook beschouwd als een aanvulling op elkaar. Shavelson en collega's (2008) pleiten inderdaad voor een en-en verhaal tussen beide toetsvormen.

De grote meerderheid van de ondervraagde leerkrachten geeft aan *self-assessment* regelmatig toe te passen in de klaspraktijk, met een onderling verschil in implementatie ervan. De helft van deze meerderheid integreert het namelijk systematisch na elk evaluatiemoment, waarbij de andere helft mondeling het gevoel van de leerlingen bevroegt. Andrade (2010) stelt inderdaad dat *self-assessment* op verschillende manieren kan toegepast worden. Ook is er een overeenkomst te vinden tussen wat de leerlingen bij deelonderzoeksvraag 1 rapporteerden, namelijk dat de meest uitgebreide *self-assessments* pas plaatsvinden na elk examen. Echter is het doel van *self-assessment* om de vooruitgang van leerlingen te monitoren en ze extra ondersteuning te bieden tijdens het leerproces, en niet pas op het einde (Arter, 1997; Dochy & McDowell, 1997). De voordelen van *self-assessment* die de leerkrachten benoemen, zijn vooral van diagnostische aard zodat zowel het probleem bij de toets als bij de leerling kan opgespoord en aangepakt worden. Onderzoek stelt dat de kracht van *self-assessment* inderdaad ligt in het feit dat het feedback voor zowel leerkracht als leerling voorziet waardoor de instructie adequaat kan aangepast worden (Nishizuka, 2022). De nadelen die vermeld worden, komen exact overeen met wat zowel de leerlingen als de literatuur zeggen; namelijk dat sommige leerlingen niet eerlijk zijn in hun *self-assessment* (de Grez, Valcke & Roozen, 2012). Ook vinden de leerkrachten en de literatuur een consensus in het feit dat hoe ouder een leerling is, hoe beter die zichzelf zou kunnen inschatten (Dochy, Segers & Sluismans, 1999).

De meerderheid van de leerkrachten uit de steekproef staat achter het principe van *ipsative assessment*, aangezien elke leerling zich in een uniek groeiproces bevindt dat niet vergeleken kan worden met dat van diens peers (Hughes, 2011, 2014, 2015; Nishizuka, 2022; Seery et al., 2019). Het concept wordt door de leerkrachten wel eerder begrepen als rijke achtergrondinformatie die aangewend wordt tijdens bijvoorbeeld deliberaties en individuele feedback, eerder dan als een structureel instrument binnen de onderwijspraktijk. Hughes (2015) stelt inderdaad dat *ipsative assessment* vooral een aanvulling is op andere vormen van toetsing en dat veel leerkrachten het voornamelijk informeel toepassen.

Elke leerkracht uit de steekproef voorziet diens leerlingen van feedback, maar er zit een groot verschil op de diepgang van deze feedback. De leerkrachten die het sporadisch toepassen, doen dit vooral wanneer de resultaten significant lager zijn dan gewoonlijk of wanneer er veel onduidelijkheden zijn. De grootste uitdaging van feedback is namelijk de operationalisering ervan, aangezien het inefficiënt en onpraktisch kan zijn om tijdige en relevante feedback te geven op maat van elke leerling (Hughes, 2011). De leerkrachten die hun leerlingen wel systematisch van feedback voorzien, doen dit middels het noteren van aandachtspunten, een verbeterleutel, remediëringsoefeningen, extra vragenuurtjes of bijles... Wanneer leerlingen inhoudelijke en taakgerichte feedback ontvangen, verbeteren de leerresultaten (Ledoux et al.,

2009). Eén leerkracht heeft ook specifiek oog voor feedback op uitstekende antwoorden, zodat de leerlingen kunnen inschatten welk niveau van hen verwacht wordt en ze een dieper begrip van kwaliteit ontwikkelen (Swan, 2006).

5.1.3. Wat zijn de ervaringen van leerlingen met testangst?

Uit de resultaten blijkt dat de leerlingen die *soms* last ervaren van testangst, de grootste groep vormen. Zelf beschrijven ze dit als gezonde stress die een logisch gevolg is van onvoldoende te studeren of wanneer het gaat om een toets van een moeilijk vak of een hoofdvak. Dit sluit aan bij de bevinding van Wadi en collega's (2022) die stellen dat testangst integraal deel uitmaakt van assessment aangezien het kan variëren van een kleine bezorgdheid tot een slopende angst. Wel wordt ook aangegeven dat er vooral veel stress ervaren wordt voor vakken waarbij zeer weinig toetsmomenten zijn, voor examens of voor herkansingstoetsen. Dat komt omdat het werkt als een 'alles-of-niets' principe waarbij één evaluatiemoment zeer veel gewicht krijgt.

Wanneer de leerlingen gevraagd wordt naar hoe testangst zich uit bij hen, kan een vicieuze cirkel beschreven worden. Wanneer ze namelijk stress ervaren tijdens een toetsmoment, kunnen ze zich minder goed concentreren wat leidt tot een black-out en tijdsnood, waardoor vermijdbare fouten sneller worden gemaakt. Dit leidt dan weer tot meer stress. Onderzoek bevestigt dit door te stellen dat stress mentale energie wegneemt die nodig is voor het behalen van goede academische prestaties (Eysenck et al., 2007), aangezien een mindere mate van concentratie en een minder functionerend werkgeheugen nefast zijn voor de resultaten (Cassady, 2004; Hembree, 1988; Zeidner, 2007). Op die manier kan gesteld worden dat testangst kan leiden tot mindere resultaten (Dorenkamp et al., 2021; Huntley et al., 2022; Pate et al., 2021; Van Praag et al., 2015; von der Embse & Witmer, 2014). Door de lagere resultaten voelen de leerlingen uit de steekproef zich ook onzeker over hun academische capaciteiten (Peleg, 2016; Putwain, 2008; Segool et al., 2013; Trifoni & Shahini, 2011). Ook wordt aangegeven dat ze zich overrompeld voelen door de werklast waardoor ze naar eigen zeggen minder tijd, energie en motivatie vinden om grondig te studeren. Volgens Spada en Moneta (2012, 2014) zouden leerlingen inderdaad opteren voor een meer oppervlakkige studiemethode, aangezien zowel de resultaten alsook de leerstrategie verandert van leerlingen die veel testangst ervaren.

Een uitgebreide hoeveelheid leerstof is één van de factoren die van invloed zou zijn op een hoge mate van testangst (Guraye et al., 2018; Hashmat et al., 2008; Khoshhal et al., 2017; Tsegay et al., 2019). Wanneer de leerdoelen daarenboven te ambitieus en complex zijn, kan dit inboeten aan motivatie en zo leiden tot testangst (Drach-Zahavy & Erez, 2002). De leerlingen uit de steekproef geven inderdaad aan het meest testangst te ervaren wanneer het gaat om een zeer grote toets van een zeer moeilijk vak, bijvoorbeeld wanneer er slechts drie toetsen per schooljaar zijn met zeer weinig vragen. Door dit klein aantal toetsmomenten, wordt het pak leerstof groter wat een uitdaging vormt op vlak van timemanagement. Twee derde van de ondervraagde leerlingen geeft daarenboven aan prestatiedruk te ondervinden vanwege hun ouders, leerkrachten of zichzelf. Ouderlijke druk is inderdaad een determinant van testangst (Levpuscek et al., 2012; Putwain et al., 2010; Zeidner, 1998, 2014).

De meerderheid van de leerlingen zegt dat wanneer ze hun gedachten trachten te verzetten met behulp van vrienden en/of hobby's, dit resulteert in tijdsdruk waardoor er nog meer stress ervaren wordt. In het rapport van de United Nations Children's Fund (2021) staat ook te lezen

dat door de toenemende focus op werk en leren minder tijd vrijkomt voor ontspanning en zelfzorg. Zowel enkele leerlingen uit de steekproef als onderzoek pleiten voor meer begrip tijdens het inplannen van toetsen en een klassikaal meer ontspannen sfeer net voor de toets (Demeester & Vanmaercke, 2021; United Nations Children's Fund, 2021). Ten slotte wordt ook gepleit voor het gebruik van meer permanente evaluatie en dus meerdere, maar kleinere toetsen (Miller & Mitchell, 1994).

5.1.4. Wat zijn de ervaringen van leerkrachten met testangst bij hun leerlingen?

Wanneer de leerkrachten gevraagd wordt wat hun ervaringen zijn met de prevalentie van testangst bij hun leerlingen, rapporteert niemand een afname. De reden hiervoor die het vaakst genoemd wordt, is dat het thema nu minder taboe is en dat de afstand tussen leerkracht en leerling kleiner is, waardoor klachten automatisch zichtbaarder worden. Het feit dat het ervaren van testangst vandaag een grotere prevalentie kent, wordt door de leerkrachten echter niet louter toegeschreven aan de toegenomen mondigheid van de leerlingen. Ook recente maatschappelijke veranderingen zoals de opkomst van sociale media en de coronalockdowns zitten hier volgens hen zeker voor iets tussen. Door de recent veranderde leerplannen – onder andere ten gevolge van de prestatiegerichte maatschappij – wordt ook een hogere werklast gerapporteerd (Vlaams Ministerie van Onderwijs en Vorming, 2021). Ten slotte wordt ook vermeld dat het zorgaanbod (van scholen) de laatste jaren immens is uitgebreid, waardoor er bijna automatisch meer zorgvragen verschijnen.

5.1.5. Welke concrete methodieken hanteren leerkrachten als het aankomt op het verminderen van testangst bij hun leerlingen?

Elke leerkracht uit de steekproef zegt faciliteiten te implementeren om tegemoet te komen aan (zorg)leerlingen (Parsons, 2008). De ene gaat hier verder in dan de andere, maar iedereen biedt steun aan tijdens evaluatiemomenten en stelt leerlingen gerust (Reinholz, 2016). Elke leerkracht doet daarenboven beroep op de zorgcoördinatoren op school aangezien ze een waardevolle schakel zijn in het ondersteunen van leerlingen (met testangst). Ze dragen namelijk bij aan het welzijn van leerlingen dankzij gepaste interventies, al dan niet in samenwerking met de leerkrachten en/of ouders (Lyon et al., 2014; von der Embse et al., 2013). Een kritiek die enkele leerkrachten echter uiten op de zorgcoördinatoren, is dat deze personen niet voor elke leerling de vertrouwenspersoon zijn. Dit is volgens Slaton en collega's (2023) echter cruciaal aangezien de therapeutische relatie de effectiviteit van de zorg garandeert. Daarom wordt de wens geuit voor meer marge tussen de lessen zodat elke leerling even kan praten met een leerkracht naar wens. Wadi en collega's (2022) pleiten ook voor een platform voor vroegtijdige interventie, met adequate training van alle personeel om om te gaan met de zorgbehoeften en testangst van leerlingen.

5.1.6. Overkoepelende discussie

Aanvankelijk werd verwacht dat er opvallende conclusies zouden kunnen getrokken worden omtrent *ipsative assessment* en minder omtrent *self-assessment*. Dit blijkt echter net andersom te zijn. Dat kan verklaard worden door het feit dat *ipsative assessment* eerder dient als achtergrondinformatie in plaats van als 'tastbare' tool waardoor ook de vragen voor de leerlingen van andere aard waren dan die voor de leerkrachten. Voor *self-assessment* is gebleken dat dit toch op minder steun kan rekenen dan verwacht, mede door de miscommunicatie errond. Ook wordt in dit onderzoek enkel de focus gelegd op metacognitieve en procesgerichte *self-assessment*, in plaats van op taakgerichte. Het is opvallend dat zelfs

niemand van de participanten spontaan taakgerichte *self-assessment* vermeldde. Verder onderzoek moet uitwijzen of dat ligt aan de gestelde vragen tijdens het interview of aan een effectief gebrek aan implementatie ervan. Of *self-* en *ipsative assessment* ten slotte testangst kunnen minimaliseren, kan niet causaal bewezen worden gezien de methodologie van deze masterproef. Wel is er een positieve correlatie te vinden tussen leerkrachten die een ipsatieve houding hanteren ten opzichte van (de resultaten van) hun leerlingen en leerkrachten die extra handelingen stellen om de testangst bij hun leerlingen te minimaliseren.

5.2. Aanbevelingen

Allereerst is tijdens dit onderzoek gebleken dat vele leerlingen tijdens een toets verwarring ervaren bij het achterhalen van de verwachtingen van de leerkracht. Er is nood aan een duidelijke context zonder ruimte voor verschillende interpretaties. Daarenboven geven de jongere leerkrachten aan een ethisch-pedagogisch dilemma te ervaren als het aankomt op het kiezen van een toetsvorm. Ze kiezen het vaakst voor de minst tijdsintensieve vorm, waardoor de lespraktijk minder adequaat kan afgestemd worden op de noden van de leerlingen. Dat is nefast aangezien er niet zoals gewenst kan geanticipeerd worden op de verwarring van de leerlingen omtrent de leerdoelen en toetsvragen. Hieruit vloeit een dubbele aanbeveling voor de praktijk, namelijk dat leerkrachten voor het toetsmoment de leerdoelen zeer helder en duidelijk dienen te communiceren aan de leerlingen (Andrade & Valtcheva, 2009; Brown & Harris, 2013; Jonsson, 2014; Nicol & McFarlane-Dick, 2006; Panadero & Alonso-Tapia, 2013). Om de lespraktijk vervolgens adequaat af te kunnen stemmen op de noden van de leerlingen, is het aangeraden om in te gaan op de vragen van de leerlingen en achteraf tijd te investeren in taakgerichte feedback (Alexander, 2011; Ball, 2003; Black & William, 1998, 2006).

Joosten-ten Brinke en collega's (z.d.) beweren dat *self-assessment* pas effectief is wanneer er zelfreflectie aan gekoppeld wordt. De leerlingen uit de steekproef van dit onderzoek rapporteren echter dat er na het invullen van hun *self-assessment* niets meer met wordt gedaan en zowel de leerlingen als de leerkrachten geven aan dat de meest uitgebreide *self-assessments* pas plaatsvinden op het einde van een trimester. Onderzoek pleit echter voor het aanvullen van de *self-assessments* met individuele feedback van de leerkracht (Birenbaum & Dochy, 1996; Boud & Falchikov, 1989; Griffie, 1995; Williams, 1992) én dat het dient als diagnostische tool *tijdens* het leerproces (Arter, 1997; Dochy & McDowell, 1997). Daarenboven wordt er binnen de steekproef een miscommunicatie opgemerkt omtrent *self-assessments*, waarbij de leerkrachten aangeven dat vele leerlingen dit foutief invullen en dat de leerlingen aangeven dit bewust te doen. De leerlingen zijn namelijk van mening dat wat ze hierop invullen van invloed is op hun resultaat, terwijl een *self-assessment* net dient als hulpmiddel om ze zo adequaat mogelijk te kunnen helpen. De concrete praktijkaanbeveling luidt bijgevolg dat de uitgebreide *self-assessments* eerder verplaatst dienen te worden naar toetsen *tijdens* het schooljaar zodat het leerproces nog bijgeschaafd kan worden voor de examens, en dat er vooraf een heldere transparantie is over het doel ervan.

Vervolgens rapporteert Katholiek Onderwijs Vlaanderen (2023) dat slechts 10% van de leerkrachten deelneemt aan de nascholing 'Kunst en kunde van feedback', en zowel de bevraagde leerlingen als onderzoek rapporteren inderdaad dat de gegeven feedback als onnuttig beschouwd wordt omdat het te oppervlakkig geformuleerd is (Carless et al., 2011; Lea & Street, 1998; Price et al., 2010; Scoles et al., 2013; Weaver, 2006). De meerderheid van de leerkrachten uit de steekproef geven zelf ook aan enkel feedback te voorzien indien een resultaat significant lager is, onder andere vanwege tijdsdruk. Onderzoek bevestigt dat de

operationalisering van feedback een uitdaging is aangezien het geven van tijdige, relevante en individuele feedback inefficiënt en onpraktisch kan zijn (Hughes, 2011). De aanbeveling voor de praktijk luidt om niet enkel feedback te voorzien op minder goede toetsen en om die ook zo specifiek mogelijk te formuleren door ook te verklaren *waarom* iets fout is. De aanbeveling voor het beleid luidt om de nascholingsprojecten omtrent het geven van feedback meer te promoten en er ook meer in te investeren.

Ten slotte uiten een aantal leerkrachten uit de steekproef kritiek op het feit dat de zorgcoördinatoren niet voor elke leerling een vertrouwenspersoon zijn, waardoor de nood ontstaat om meer ruimte tussen de lesuren te creëren zodat elke leerkracht beschikbaar is voor een gesprek met een leerling die dat wenst. Dit zal bovendien het veilig leerklimaat op school bevorderen. Voor deze beleidsaanbeveling wordt het onderzoek van Wadi en collega's (2022) aangewend, namelijk dat zowel tijdens de lerarenopleiding als tijdens de nascholingsprojecten ingezet wordt op training van alle leerkrachten om in te kunnen spelen op het psychologisch welbevinden van leerlingen. Zorgcoördinatoren zijn nochtans een waardevolle sleutelfiguur in het begeleiden van leerlingen (met testangst), maar het is vooral de therapeutische relatie tussen leerling en hulpverlener die zorgt voor de effectiviteit van de zorg (Slaton et al., 2023; von der Embse et al., 2013).

5.3. Beperkingen van dit onderzoek en suggesties voor vervolgonderzoek

Een beperking met betrekking tot de dataverzameling van dit cross-sectioneel onderzoek is dat de steekproef van leerkrachten en leerlingen zich bevindt in slechts twee Vlaamse secundaire scholen, die zich bovendien beide situeren binnen het Algemeen Secundair Onderwijs (ASO). Dit in combinatie met het feit dat het een gelegenheidssteekproef is met niet-gestandaardiseerde interviewvragen, maakt dat de verkregen resultaten niet generaliseerbaar zijn naar alle leerkrachten en leerlingen in het Vlaams secundair onderwijs. Onderzoek bij andere participanten zal nieuwe en/of andere praktijken en ervaringen blootleggen. Dit onderzoek wordt dus eerder gezien als een inspiratie en aanzet voor verder onderzoek, en als eye-opener voor het brede publiek door de ervaringen van leerkrachten en leerlingen in kaart te brengen en ze een stem te geven. Ook namen de leerkrachten vrijwillig deel, waardoor de ervaringen van de overige gecontacteerde leerkrachten niet zijn opgenomen. Dat heeft mogelijks gevolgen voor de resultaten, aangezien nu enkel geïnteresseerden deelnamen. Daarenboven is na de dataverzameling gebleken dat de steekproef zich enkel situeert in de derde graad van het secundair onderwijs, aangezien de leerlingen hier doorgaans meer matuur en vertrouwd zijn met assessment. Een concrete suggestie voor vervolgonderzoek is bijgevolg dat er bij herhaling van dit onderzoek geopteerd moet worden voor gestratificeerde randomisatie, waarbij leerkrachten en hun leerlingen random gekozen worden binnen de drie strata van ASO, BSO en TSO in Vlaanderen over alle graden heen. Dit verhoogt de generaliseerbaarheid van de resultaten van dit onderzoek.

Deze masterproef beantwoordt de vooropgestelde onderzoeksvragen, maar roept daarbij ook extra vragen op. Zo is bijvoorbeeld uit de data-analyse gebleken dat de huidige leerplannen met de geactualiseerde eindtermen als te zwaar worden gepercipieerd waardoor zowel leerlingen als leerkrachten naar eigen zeggen een hogere mate van werklast rapporteren. De commissie Beter Onderwijs is opgericht in 2020 om ook hieraan tegemoet te komen (Vlaams Ministerie van Onderwijs en Vorming, 2021), maar het is nog te vroeg om daar nu reeds de resultaten van te rapporteren. Toekomstig onderzoek zou deze eerste bevindingen omtrent de commissie bijgevolg als context kunnen meenemen, naast het oorspronkelijke opzet van deze

studie. Ook is tijdens de interviews vaak spontaan het woord 'Corona' gevallen als mogelijke invloed op een verhoogde testangst bij leerlingen, zonder dat er specifiek naar gevraagd werd. Vervolgonderzoek zou een longitudinale studie kunnen opzetten met dezelfde leerkrachten, maar met telkens verschillende leerlingen, om te zien of de lockdowns echt een groot effect hadden of niet.

Ten slotte worden ook enkele negatieve consequenties opgemerkt ten gevolge van tijdsgebrek bij vooral jongere leerkrachten. Het is opportuun te onderzoeken of de aanvangsbegeleiding hieraan tegemoet zou kunnen komen, door bijvoorbeeld in het begin van de carrière meer administratieve ondersteuning te voorzien. Een volgende tekortkoming van dit onderzoek is dat gezien de methodologie en de gestelde vragen tijdens de interviews het niet mogelijk is een sluitend antwoord te bieden op deelonderzoeksvraag 6. Ten slotte toont Tabel 1 een intra-beoordelaarsbetrouwbaarheid van slechts 0,5 voor de codes onder 'Leerkrachten'¹. Het is bijgevolg aangewezen om de codes onder 'Leerlingen' ook te onderwerpen aan een controle én dit te laten voltooien door een peer, waardoor ook de inter-beoordelaarsbetrouwbaarheid kan gemeten worden.

¹ Tabel 1: intra-codeurbetrouwbaarheid van data 'Leerkrachten', p. 22

Bibliografie

- Abbasi, M., Ayadi, N., & Shafiee, H. (2016). Role of Social Well-Being and Academic Vitality in Predicting the Academic Motivation in Nursing Students. *Education Strategies in Medical Sciences*, 8(6), 49–54. <http://edcbmj.ir/article-1-899-en.html>
- AHOVOKS, & Ministerie van Onderwijs en Vorming. (n.d.). *Wat is VKS? | De Vlaamse Kwalificatiestructuur*. Vlaamse Overheid. <https://vlaamsekwalficatiestructuur.be/wat-is-vks/>
- Amelinckx, E., de Vos, F., de Peuter, F., Jacobs, K., & Dahri, N. (2023). “Hoeveel noodkreten over mentale gezondheid bij jongeren moeten er nog klinken?” *Knack*. <https://www.knack.be/nieuws/belgie/maatschappij/hoeveel-noodkreten-over-mentale-gezondheid-bij-jongeren-moeten-er-nog-klinken/>
- Andrade, H. L. (2010). Students as the definitive source of formative assessment: Academic self-assessment and the self-regulation of learning. *Handbook of formative assessment* (90-105). Routledge. <https://www.taylorfrancis.com/chapters/edit/10.4324/9780203874851-8/students-definitive-source-formative-assessment-heidi-andrade>
- Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory Into Practice*, 48(1), 12-19. <http://dx.doi.org/10.1080/00405840802577544>
- Andrade, H., Wang, X. L., Du, Y., & Akawi, R. L. (2009). Rubric-referenced self-assessment and self-efficacy for writing. *Journal of Educational Research*, 102(4), 287-301. <https://doi.org/10.3200/JOER.102.4.287-302>
- Assessment Reform Group (2002). *Assessment for learning: 10 Principles. Research-based principles to guide classroom practice*. University of Cambridge, School of Education.
- Baarends, E. & Simon, F. (2017). Over validering van praktijkgericht kwalitatief onderzoek. *KWALON*, 22(3). <https://doi.org/10.5117/2017.022.003.002>
- Bandura, A. (1997). Self-efficacy: The exercise of control. *New York*.
- Barana, A., & Marchisio, M. (2016). Ten good reasons to adopt an automated formative assessment model for learning and teaching Mathematics and scientific disciplines. *Procedia-Social and Behavioral Sciences*, 228, 608–13. <https://doi.org/10.1016/j.sbspro.2016.07.093>
- Biggs, J., & Tang, C. (1999). Teaching for quality learning at university Buckingham: SRHE and Open University Press. *Search in*.
- Black, P., Harrison, C., & Lee, C. (2003). *Assessment for learning: Putting it into practice*. McGraw- Hill Education (UK). <https://ap.lc/A8Clf>
- Bourke, R., O'Neill, J., & Loveridge, J. (2018). What starts to happen to assessment when teachers learn about their children's informal learning? *Australian Educational Researcher*, 45(1), 33–50. <https://doi.org/10.1007/S13384-018-0259-X>
- Brandmo, C., Bråten, I., & Schewe, O. (2019). Social and personal predictors of test anxiety among Norwegian secondary and postsecondary students. *Social Psychology of Education*, 22(1), 43–61. <https://doi.org/10.1007/S11218-018-9461-Y>

- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), 77-101. <https://doi.org/10.1191/1478088706qp063oa>
- Brown, S., & Knight, P. (1994). *Assessing Learners in Higher Education*. Psychology Press. <https://ap.lc/il62Y>
- Carter, N., Bryant-Lukosius, D., DiCenso, A., Blythe, J., & Neville, A. J. (2014). The Use of Triangulation in Qualitative Research. *Oncology Nursing Forum*, 41(5), 545–547. <https://doi.org/10.1188/14.onf.545-547>.
- Corbin, J., & Strauss, A. (2008). Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. Thousand Oaks, CA: Sage. *Organizational Research Methods*, 12(3), 614-617. <https://doi.org/10.1177/1094428108234514>
- de Grez, L., Valcke, M., & Roozen, I. (2012). How effective are self- and peer assessment of oral presentation skills compared with teachers' assessments? *Active Learning in Higher Education*, 13(2), 129–142. https://doi.org/10.1177/1469787412441284/ASSET/IMAGES/LARGE/10.1177_1469787412441284-FIG1.JPEG
- detra Maeyer, S., & Rymenans, R. (2004). *Onderzoek naar kenmerken van effectieve scholen*. <https://ap.lc/jZHxu>
- Demeester, B., & Vanmaercke, S. (2021). *Rapport Mentaal Welzijn - Vlaamse Vereniging van Studenten (VVS)*. <https://vvs.ac/dossier-mentaal-welzijn/>
- Ding, L. (1998). Revisiting assessment and learning implications of students' perspectives on assessment feedback. Paper presented to Scottish Educational Research Association Annual Conference, *University of Dundee*.
- Dijkstra, P. (2015). Effectiever leren met leerstrategieën Pieterneel Dijkstra BOOM. *Boom Test Uitgevers*. www.cedar.nl/pro
- Dochy, F., Segers, M., & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: A review - ProQuest. *Studies in Higher Education*. <https://www.proquest.com/docview/219534593?accountid=11077>
- Dorenkamp, M. A., Irrgang, M., & Vik, P. (2021). Assessment-related anxiety among older adults: associations with neuropsychological test performance. <https://doi.org/10.1080/13825585.2021.2016584>.
- Dyrbye, L.N., Thomas, M.R., & Shanafelt, T.D. (2006). Systematic review of depression, anxiety, and other indicators of psychological distress among US and Canadian medical students. *Academic medicine*, 81(4), 354-373. <https://ap.lc/IYvkm>
- Enthoven, M. & Oostdam, R. (2012). De functie en zin van praktijkgericht onderzoek door studenten van educatieve hbo-opleidingen. *Kenniscentrum onderwijs en opvoeding*. <https://www.hva.nl/kc-onderwijs-opvoeding/gedeelde-content/publicaties/publicaties-algemeen/2014-of-ouder/de-functie-en-zin-van-praktijkgericht-onderzoek---enthoven-en-oostdam.html>
- Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: attentional control theory. *Emotion*, 7(2), 336. <https://psycnet.apa.org/doi/10.1037/1528-3542.7.2.336>

- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74, 59–109. doi:10.3102/00346543074001059
- Guraya, S.Y., Guraya, S.S., Habib, F., Al Quiliti, K.W., & Khoshhal, K.I. (2018). Medical students' perception of test anxiety triggered by different assessment modalities. *Med Teach*. 40, S49- S55. <https://doi.org/10.1080/0142159X.2018.1465178>
- Hagemann, N., Akcaoglu, Z., Hiekkaranta, A., Achterhof, R., Bamps, E., Janssens, J., Hermans, K., Biesemans, T., Wampers, M., Apers, S., Kirtley, O., & Myin-Germeys, I. (2019). SIGMA: Hoe gaat het met jullie mentale welzijn? *Center for Contextual Psychiatry*. <https://ap.lc/SsSxr>
- Harding, G., & Gantley, M. (1998). Qualitative methods: beyond the cookbook. *Family Practice*, 15(1), 76–79. <https://doi.org/10.1093/fampra/15.1.76>
- Hashmat, S., Hashmat, M., Amanullah, F., & Aziz, S. (2008). Factors causing exam anxiety in medical students. *Journal-Pakistan Medical Association*, 58(4), 167.
- Hove, S. E., & Anda, B. (2005). Experiences from Conducting Semi-Structured Interviews in Empirical Software Engineering Research. *International Software Metrics Symposium*, 11, 10–23. <https://doi.org/10.1109/METRICS.2005.24>
- Van Hove, G., & Claes, L. (2011). Qualitative research and educational sciences: A reader about useful strategies and tools. *Hampshire: Ashford Colour Press*.
- Hughes, G. (2011). Towards a personal best: A case for introducing ipsative assessment in higher education. *Studies in Higher Education*, 36(3), 353–367. <https://doi.org/10.1080/03075079.2010.486859>
- Hughes, G. (2015). Ipsative assessment: motivation through marking progress. <http://Dx.Doi.Org/10.1080/00071005.2015.1035906>, 63(2), 246–248.
- Huntley, C., Young, B., Tudur Smith, C., Jha, V., & Fisher, P. (2022). Testing times: the association of intolerance of uncertainty and metacognitive beliefs to test anxiety in college students. *BMC Psychology*, 10(1). <https://doi.org/10.1186/S40359-021-00710-7>
- Jarvis, P. (2010). Assessing and evaluating. *Adult education & Lifelong learning. Theory and Practice (4th ed., pp.210-226)*. Routledge. <https://ap.lc/hc3mX>
- Jonsson, A. (2014). Rubrics as a way of providing transparency in assessment. *Assessment & Evaluation in Higher Education*, 1-13. <http://dx.doi.org/10.1080/02602938.2013.875117>
- Joosten-ten Brinke, D. (2010). *Eigentijds toetsen en beoordelen*. <https://ap.lc/AriPQ>
- Kašćák, O., & Pupala, B. (2011). *Školy v prúde reformem* (pp. 137-194). Renesans. https://is.muni.cz/publication/1074291/Kascak_pupala_recenze.pdf
- Katholiek Onderwijs Vlaanderen. (2023). *Jaarverslag 2021-2022*. <https://pro.katholiekonderwijs.vlaanderen/jaarverslag-2021-2022/cijfers>
- Kenniscentrum Kinderrechten vzw. (2022). De impact van de coronamaatregelen op kinderen, jongeren en jongvolwassenen: een systematische literatuurstudie. *Jeugdonderzoeksplatform*. www.gezondleven.be/kwaliteitsvolle-gezondheidsbevordering/

- Khoshhal, K.I., Khairy, G.A., Guraya, S.Y., & Guraya, S.S. (2017). Exam anxiety in the undergraduate medical students of Taibah University. *Medical teacher*, 39, S22-S26. <https://doi.org/10.1080/0142159X.2016.1254749>
- Klenowski, V., & Wyatt-Smith, C. (2012). The impact of high stakes testing: The Australian story. *Assessment in Education: Principles, Policy and Practice*, 19(1), 65–79. <https://doi.org/10.1080/0969594X.2011.592972>
- Lecluyse, L. (2018). *Wat doen (de)motiverende leerkrachten? Inzichten vanuit de Zelf-Determinatie Theorie en het model voor Interpersoonlijk Leraarsgedrag*. https://libstore.ugent.be/fulltxt/RUG01/002/481/857/RUG01-002481857_2018_0001_AC.pdf
- Ledoux, G., Blok, H., Boogaard, M., & Krüger, M. (2009). Opbrengstgericht werken: over de waarde van meetgestuurd onderwijs. *SCO-Kohnstamm Instituut*. <http://www.sco-kohnstammstituut.uva.nl/pdf/sco812.pdf>
- Leenknecht, M., Wijnia, L., Köhler, M., Fryer, L., Rikers, R., & Loyens, S. (2021). Formative assessment as practice: the role of students' motivation. *Assessment & Evaluation in Higher Education*, 46(2), 236–255. <https://doi.org/10.1080/02602938.2020.1765228>
- Lyon, A. R., Ludwig, K., Romano, E., Koltracht, J., vander Stoep, A., & McCauley, E. (2014). Using Modular Psychotherapy in School Mental Health: Provider Perspectives on Intervention-Setting Fit. <https://doi.org/10.1080/15374416.2013.843460>, 43(6), 890–901.
- Marshall, B., Cardon, P., Poddar, A., & Fontenot, R. (2013). Does Sample Size Matter in Qualitative Research? A Review of Qualitative Interviews in is Research. *Journal of Computer Information Systems*, 54(1), 11–22. <https://doi.org/10.1080/08874417.2013.11645667>
- McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, 22(3), 276–282. <https://doi.org/10.11613/BM.2012.031/FULLARTICLE>
- McIntosh, M. J., & Morse, J. M. (2015). Situating and Constructing Diversity in Semi-Structured Interviews. *Global Qualitative Nursing Research*, 2. <https://doi.org/10.1177/2333393615597674>
- Nguyen, T.T., & Deci, E.L. (2016). Can it be good to set the bar high? The role of motivational regulation in moderating the link from high standards to academic well-being. *Learning and individual differences*, 45, 245–251. <https://doi.org/10.1016/j.lindif.2015.12.020>
- Nicol, D., & McFarlane-Dick, D. (2006). Formative assessment and self-regulated learning, a model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199–218. <http://dx.doi.org/10.1080/03075070600572090>
- Nishizuka, K. (2022). Significance and Challenges of Formative Ipsative Assessment in Inquiry Learning: A Case Study of Writing Activities in a “Contemporary Society” Course in a Japanese High School. *SAGE Open*, 12(2). https://doi.org/10.1177/21582440221094599/ASSET/IMAGES/LARGE/10.1177_21582440221094599-FIG2.JPEG
- Ozan, C., & Kincal, R. (2018). The effects of formative assessment on academic achievement, attitudes toward the lesson, and self-regulation skills. *Educational Sciences: Theory and Practice*. 18(1): 85–118. <http://doi.org/10.12738/estp.2018.1.0216>

- Ozuem, W., Willis, M. & Howell, K. (2022). Thematic analysis without paradox: sensemaking and context. *Qualitative Market Research: An International Journal*, 25(1), 143–157. <https://doi.org/10.1108/qmr-07-2021-0092>
- Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research*, 66(4), 543e578. <http://dx.doi.org/10.3102/00346543066004543>
- Pajares, F. (2012). Motivational role of self-efficacy beliefs in self-regulated learning. *Motivation and self-regulated learning*. (111-168). Routledge. <https://ap.lc/s6Cvo>
- Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, 22, 74–98. <https://doi.org/10.1016/J.EDUREV.2017.08.004>
- Parsons, D. (2008). Is there an alternative to exams? Examination stress in engineering courses. *International Journal of Engineering Education*, 24(6), 1111–1118. http://www.ijee.dit.ie/latestissues/Vol24-6/s10_ijee1993.pdf
- Pate, A. N., Neely, S., Malcom, D.R., Daugherty, K.K., Zagar, M., & Medina, M. S. (2021). Multisite study assessing the effect of cognitive test anxiety on academic and standardized test performance. *American Journal of Pharmaceutical Education*, 85(1), 43–54. <https://doi.org/10.5688/AJPE8041>
- Peleg, O., Deutch, C., & Dan, O. (2016). Test anxiety among female college students and its relation to perceived parental academic expectations and differentiation of self. *Learning and Individual Differences*, 49, 428–436. <https://doi.org/10.1016/J.LINDIF.2016.06.010>
- Tian-Ci Quek, T., Wai-San Tam, W., X. Tran, B., Zhang, M., Zhang, Z., Su-Hui Ho, C., & Chun-Man Ho, R. (2019). The global prevalence of anxiety among medical students: a meta-analysis. *International journal of environmental research and public health*, 16(15), 2735. <https://doi.org/10.3390/ijerph16152735>
- Rabbani, Z., & Yousefi, F. (2012). Examine the mediator role of the goal orientation in the relationship between self-efficacy and task value with cognitive strategies. *JISL*. 4(2): 49–80
- Rakoczy, K., Pinger, P., Hochweber, J., Klieme, E., Schütze, B., & Besser, M. (2019). Formative assessment in mathematics: Mediated by feedback's perceived usefulness and students' self-efficacy. *Learning and Instruction*, 60, 154-165. <https://doi.org/10.1016/j.learninstruc.2018.01.004>
- Rattray, J. (2018). Affect and ipsative approaches as a counter to pedagogic frailty: the guardian of traditional models of student success. 37(7), 1489–1500. <https://doi.org/10.1080/07294360.2018.1494141>
- Redactie. (2023). ONDERWIJZER. Bekijk hier hoe de scholen in jouw buurt scoren: hoeveel C-attesten reiken ze uit en hoeveel leerlingen halen later hoger diploma? *Het Nieuwsblad*. https://www.nieuwsblad.be/cnt/dmf20230308_92996215
- Robinson, O. C. (2013). Sampling in Interview-Based Qualitative Research: A Theoretical and Practical Guide. *Qualitative Research in Psychology*, 11(1), 25–41. <https://doi.org/10.1080/14780887.2013.801543>
- Rotenstein, L.S., Ramos, M.A., Torre, M., Bradley Segal, J., Peluso, M.J., Guille, C., Sen, S., ... & Mata, D.A. (2016). Prevalence of depression, depressive symptoms, and suicidal ideation among medical students a systematic review and meta-analysis. *Jama*, 316(21), 2214-2236. <https://doi.org/10.1001/jama.2016.17324>

- Schober, M. F. (2018). The future of face-to-face interviewing. *Quality Assurance in Education*, 26(2), 290–302. <https://doi.org/10.1108/qaе-06-2017-0033>
- Schunk, D.H., & Ertmer, P.A. (1999). Self-regulatory processes during computer skill acquisition: Goal and self-evaluative influences. *Journal of Educational Psychology*, 91(2), 251-260. <http://dx.doi.org/10.1037/0022-0663.91.2.251>
- Schwab, K. (2020). The Global Competitiveness Report. *World Economic Forum*. <https://www.weforum.org/reports/the-global-competitiveness-report-2020/in-full>
- Seery, N., Buckley, J., Delahunty, T., & Canty, D. (2019). Integrating learners into the assessment process using adaptive comparative judgement with an ipsative approach to identifying competence-based gains relative to student ability levels. *International Journal of Technology and Design Education*, 29(4), 701–715. <https://doi.org/10.1007/S10798-018-9468-X/TABLES/2>
- Skundberg-Kletthagen, H., & Moen, Ø. L. (2017). Mental health work in school health services and school nurses' involvement and attitudes, in a Norwegian context. *Journal of Clinical Nursing*, 26(23–24), 5044–5051. <https://doi.org/10.1111/JOCN.14004>
- Slaton, C. R., Lammers, W., & Park, A. (2023). How school belongingness in diverse students moderates student perceptions of teachers' cultural humility in predicting student–teacher working alliance. *Psychology in the Schools*. <https://doi.org/10.1002/PITS.22862>
- Soares, D., & Woods, K. (2020). An international systematic literature review of test anxiety interventions 2011–2018. *Pastoral Care in Education*, 38(4), 311–334. <https://doi.org/10.1080/02643944.2020.1725909>
- Systematic methodological review: developing a framework for a qualitative semi-structured interview guide. *Journal of Advanced Nursing*, 72(12), 2954–2965. <https://doi.org/10.1111/jan.1303>
- Tan, K. H. K. (2010). Does student self-assessment empower or discipline students? 29(6), 651–662. <https://doi.org/10.1080/0260293042000227209>
- Thompson, J. (2022). A Guide to Abductive Thematic Analysis. *The Qualitative Report*. <https://doi.org/10.46743/2160-3715/2022.5340>
- United Nations Children's Fund. (2021). The state of the world's children in 2021: On My Mind - promoting, protecting and caring for children's mental health. *UNICEF*. <https://www.unicef.org/media/119751/file>
- Valcke, M. (2007). *Onderwijskunde als ontwerpwetenschap: een inleiding voor ontwikkelaars van instructie en voor toekomstige leerkrachten*. Academia Press. <https://ap.lc/xSYkX>
- Van Hove, G. (2014). *Qualitative Research for Educational Sciences*. Hampshire: Pearson.
- van Praag, L., Boone, S., Stevens, P., & van Houtte, M. (2015). De paradox van het watervalstelsel: wanneer het groeperen van studenten in homogene groepen tot meer heterogeniteit leidt in het beroepsonderwijs. *Sociologos*, 36(2), 82–101

- Vanstone, D.M., & Hicks, R.E. (2019). Transitioning to university: Coping styles as mediators between adaptive-maladaptive perfectionism and test anxiety. *Personality and Individual Differences*, 141, 68–75. <https://doi.org/10.1016/J.PAID.2018.12.026>
- Varpio, L. (2020). Thematic analysis of qualitative data: AMEE Guide No. 131. *Medical Teacher*, 42(8), 846–854. <https://doi.org/10.1080/0142159x.2020.1755030>
- Vlaams Ministerie van Onderwijs en Vorming. (2021). *Commissie Beter Onderwijs stelt basiskennis centraal in 58 adviezen en 10 speerpunten*. Vlaams Ministerie van Onderwijs En Vorming. <https://ap.lc/sBVmE>
- von der Embse, N., Barterian, J., & Segool, N. (2013). Test Anxiety Interventions for Children and Adolescents: A Systematic Review of Treatment Studies from 2000–2010. *Psychology in the Schools*, 50(1), 57–71. <https://doi.org/10.1002/PITS.21660>
- von der Embse, N.P., & Witmer, S.E. (2014). High-Stakes Accountability: Student Anxiety and Large- Scale Testing. *Journal of Applied School Psychology*, 30(2), 132–156. <https://doi.org/10.1080/15377903.2014.888529>
- Wadi, M., Yusoff, M.S.B., Abdul Rahim, A.F., & Lah, N.A.Z.N. (2022). Factors affecting test anxiety: a qualitative analysis of medical students' views. *BMC Psychology*, 10(1). <http://doi.org/10.1186/S40359-021-00715-2>
- Leeman, Y., & Wardekker, W. (2010). Verbeter onderzoek het onderwijs? *VELON Tijdschrift voor lerarenopleiders*, 31(1), 19-22. <https://ap.lc/TasyP>
- Wiliam, D. (2011). Embedded formative assessment. *Solution Tree Press*. <https://ap.lc/4IXuV>
- Williams, E. (1992). Student Attitudes Towards Approaches to Learning and Assessment, *Scopus*. <https://ap.lc/k8Zjq>
- Winn, A.S., DelSignore, L., Marcus, C., Chiel, L., Freiman, E., Stafford, D., & Newman, L. (2019). Applying cognitive learning strategies to enhance learning and retention in clinical teaching settings. *MedEdPORTAL*, 15, 10850. https://doi.org/10.15766/mep_2374-8265.10850
- Woolley, M.E., & Grogan-Kaylor, A. (2006). Protective family factors in the context of neighborhood: Promoting positive school outcomes. *Family Relations*, 55, 93–104. doi:10.1111/j.1741-3729.2006.00359.x
- Yousefi Afrashteh, M., & Rezaei, S. (2022). The mediating role of motivated strategies in the relationship between formative classroom assessment and academic well-being in medical students: a path analysis. *BMC Medical Education*, 22(1). <https://doi.org/10.1186/S12909-022-03118-Y>
- Yousefi, A.M. (2019). The Relationship Between formative assessment with Academic engagement and Using Metacognitive Strategies in Medical Students. *Educ Strategy Med Sci*. 11(6), 147–52
- Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology*, 25, 82–91

Bijlagen

Bijlage 1: interviewleidraad leerkrachten

Topic	Vragen
Toelichting onderzoek & Informed Consent	<ul style="list-style-type: none">• Korte voorstelling van mezelf en van het thema en doel van de masterproef• Ondertekenen Informed Consent• Toestemming audio-opname gesprek
Algemene informatie	<ul style="list-style-type: none">• Toelichting van alle termen
Soorten assessment	<ul style="list-style-type: none">• Voor welke brede evaluatievorm opteert u het liefst; summatief of formatief? En waarom?• Welk gevoel heeft u bij de door u gebruikte evaluatievormen? Wat lijkt u ideaal? Waarom?• Heeft u al eens <i>self-</i> of <i>ipsative assessment</i> gebruikt? Kent u die laatste? Is dit haalbaar om toe te passen in de praktijk? Waarom wel/niet? Welke voor- en nadelen zie je?• Hoe waakt u over het feit dat u uw leerlingen voldoende voorziet van (constructieve) feedback?
Psychologisch welbevinden en testangst	<ul style="list-style-type: none">• Denkt u dat veel leerlingen momenteel testangst ervaren door stress voor toetsen? Wat lijkt u de oorzaak van een eventuele toename? Wat lijkt u een mogelijke oplossing?• Hoe tracht u de testangst bij uw leerlingen te verminderen? Heeft u ooit al eens een toets aangepast qua vorm? Zijn er verdere concrete methodieken die u toepast?

Bijlage 2: vragenlijst leerlingen

1. Hoe zien de meeste van jouw toetsen eruit? Is dat gewoon vraag en antwoord, een vraagstuk oplossen...? Of mag je af en toe ook eens je eigen mening schrijven, origineel en creatief zijn...?
2. Wanneer je je toets terugkrijgt, is de score voor jou dan duidelijk? Staat er feedback bij juiste en foute antwoorden? Weet je m.a.w. waarom je iets juist of fout hebt gedaan en legt de leraar dan ook uit hoe het wel moet? Of wordt er niets met de antwoorden gedaan?
3. Welk soort toets maak je het liefst (meerkeuze, open vragen, toepassen...)?
Waarom?
4. Heb je jezelf al eens moeten beoordelen (bv. na een spreekbeurt of na een toets/examen)? Vind je dat leuk en/of nuttig om te doen?
5. Vergelijk je jouw punten met je klasgenoten of eerder met je eigen vorige toetsen? Hecht je m.a.w. veel belang aan het klasgemiddelde of focus je je vooral op je eigen, persoonlijke vooruitgang?
6. Wanneer je een toets hebt, ervaar je dan veel stress? Zo ja, door wat komt dat denk je? Bv. onduidelijke vragen, het is een heel belangrijke/grote toets, je hebt niet voldoende kunnen studeren, je ervaart druk van ouders/leraren/vrienden...
7. Indien je op de vorige vraag 'ja' hebt geantwoord; hoe uit deze stress zich en hoe ga je daarmee om? Heeft dat soms al eens geleid tot een lager zelfbeeld of mindere punten?
8. Indien je op de vorige 2 vragen 'ja' hebt geantwoord; als je veel stress voor school ervaart, hoe tracht je dat tegen te gaan en jezelf te kalmeren? Is er iets wat de leraar of school zou kunnen doen om je écht te helpen?
9. Is er nog iets extra dat je wenst te vertellen? Iets wat nog niet aan bod is gekomen?

Bijlage 3: informatieformulier leerkrachten

Informatie over de studie

U wordt uitgenodigd om deel te nemen aan een studie van de Universiteit Gent. Neem voldoende tijd om deze informatiebrief aandachtig te lezen voor u beslist deel te nemen aan deze studie. Aarzel niet om vragen te stellen aan de onderzoeker als er onduidelijkheden zijn of indien u bijkomende informatie wenst. Zorg ervoor dat u alles begrijpt. Eens u beslist heeft om deel te nemen aan de studie zal men u vragen om het toestemmingsformulier achteraan deze bundel te ondertekenen.

Wat is het doel van het onderzoek?

In mijn thesis ga ik op zoek of er een relatie is tussen het (al dan niet regelmatige) gebruik van *self-* en *ipsative assessment* in de klaspraktijk en testangst bij leerlingen in het Vlaams secundair onderwijs. Hiervoor tracht ik zowel leerkrachten als leerlingen te bevragen om een zo breed mogelijk beeld te kunnen verkrijgen.

Informatie m.b.t. deelname

Wat houdt deelnemen aan dit onderzoek in?

De deelname aan deze studie is volledig vrijwillig en er kan op geen enkele manier sprake zijn van dwang. U kunt weigeren om deel te nemen aan de studie en u kunt zich op elk ogenblik terugtrekken uit de studie zonder dat u hiervoor een reden moet opgeven. Indien u weigert om deel te nemen, of wanneer u beslist zich terug te trekken uit een lopende studie, zal dit op geen enkele manier een invloed hebben op uw verdere relatie met de onderzoeker, uw evaluatie en/of studiebegeleiding (indien u student bent) of uw behandeling (indien u een therapeutische relatie heeft met de onderzoeker).

Als u dit wenst, kan u een samenvatting van de onderzoeksbevindingen krijgen nadat de studie is afgerond en de resultaten bekend zijn. Om een samenvatting te krijgen, kan u dit aanvragen bij de onderzoeker met wie u contact heeft.

Informatie m.b.t. Privacy en Persoonsgegevens

Het wettelijk kader voor de verwerking van persoonsgegevens en vertrouwelijke informatie in het kader van dit onderzoek wordt bepaald door:

- De Europese Algemene Verordening Gegevensbescherming 2016/679 van 27 april 2016, die van kracht is sinds 25 mei 2018 (dit is de AVG of GDPR);
- De Belgische Wet betreffende de bescherming van natuurlijke personen met betrekking tot de verwerking van persoonsgegevens van 30 juli 2018

De onderzoekers dienen zich te houden aan de generieke gedragscode voor de verwerking van persoonsgegevens van de UGent.

Welke persoonsgegevens worden verzameld?

De volgende persoonsgegevens zullen worden verwerkt:

- Beroep

Bijlage 4: blanco informed consent voor leerkrachten

Toestemming m.b.t. deelname aan de studie

Gelieve het juiste bolletje aan te kruisen	Ja	Nee
Ik neem vrijwillig deel aan deze wetenschappelijke studie.		
Ik weet dat ik me op elk ogenblik uit de studie mag terugtrekken zonder een reden voor deze beslissing op te geven en zonder dat dit op enige wijze een invloed zal hebben op mijn verdere relatie met de onderzoeker.	<input type="radio"/>	<input type="radio"/>
Ik heb het informatieformulier gelezen en heb voldoende uitleg gekregen over de aard, het doel, de duur, en de voorziene effecten van de studie. Ik kreeg de gelegenheid om vragen te stellen en ik heb op al mijn vragen een bevredigend antwoord gekregen.		

Toestemming m.b.t. de verwerking van persoonsgegevens

Gelieve het juiste vakje aan te kruisen	Ja	Nee
Ik weet dat ik rechten heb om mijn privacy te vrijwaren (o.a. inzage, verbetering, verwijdering) en tot wie ik me moet richten om deze rechten uit te oefenen.	<input type="radio"/>	<input type="radio"/>
Ik geef toestemming aan de onderzoekers om voor de doelstellingen van deze studie (persoons)gegevens van mij te verzamelen, verwerken, bewaren, analyseren en erover te rapporteren.		

Naam deelnemer	Naam onderzoeker
Datum:	Datum:
Handtekening	Handtekening

Bijlage 5: blanco informed consent voor leerlingen en hun ouders

Toestemming m.b.t. deelname aan de studie

Gelieve het juiste vakje aan te kruisen	Ja	Nee
Ik neem vrijwillig deel aan deze wetenschappelijke studie en geef toestemming aan de onderzoekers om mijn gegevens te verwerken, ze te bewaren, te analyseren en erover te rapporteren.		
Ik weet dat ik me op elk ogenblik uit de studie mag terugtrekken zonder een reden voor deze beslissing op te geven en zonder dat dit op enige wijze een invloed zal hebben op mijn verdere relatie met de onderzoeker.	<input type="radio"/>	<input type="radio"/>
Ik heb het informatieformulier gelezen en heb voldoende uitleg gekregen over de aard, het doel, de duur, en de voorziene effecten van de studie. Ik kreeg de gelegenheid om vragen te stellen en ik heb op al mijn vragen een bevredigend antwoord gekregen.		

Naam deelnemer	Naam onderzoeker
Datum:	Datum:
Handtekening	Handtekening

Toestemming m.b.t. deelname aan de studie

Gelieve het juiste vakje aan te kruisen	ja	nee
Mijn zoon/dochter neemt vrijwillig deel aan deze wetenschappelijke studie.		
Ik weet dat mijn kind zich op elk ogenblik uit de studie mag terugtrekken zonder een reden voor deze beslissing op te geven en zonder dat dit op enige wijze een invloed zal hebben op de verdere relatie met de onderzoeker of leerkracht. Het vroegtijdig stopzetten van de deelname heeft geen negatieve invloed op de schoolresultaten van mijn kind.	<input type="radio"/>	<input type="radio"/>
Ik heb het informatieformulier gelezen en heb voldoende uitleg gekregen over de aard, het doel, de duur, en de voorziene effecten van de studie. Ik kreeg de gelegenheid om vragen te stellen en ik heb op al mijn vragen een bevredigend antwoord gekregen.		

Naam leerling
Naam ouder/voogd
Datum
Handtekening

Bijlage 6: codeboom

