**Online feedback op schriftelijk werk: betere feedback in minder tijd.**

Feedback is een van de krachtigste onderwijsstrategieën om het leren van studenten te bevorderen, mits de feedback voldoende aanknopingspunten biedt ter verbetering, tijdig wordt gegeven en studenten de mogelijkheid hebben om de feedback toe te passen. Daarom streven veel instellingen na dat studenten frequenter feedback op hun prestaties ontvangen. Dit beleid legt echter een zware tijdsdruk op docenten, zeker wanneer het gaat om het becommentariëren van geschreven werkstukken. Inmiddels zijn er digitale hulpmiddelen ontwikkeld die docenten kunnen helpen om de gewenste feedback op efficiënte wijze te geven en daarmee tijd te besparen.

In dit onderzoek is binnen vier instellingen voor hoger onderwijs onderzocht hoe twee digitale nakijktools (Turnitin en WinVision) in de onderwijspraktijk werden gebruikt en hoe docenten en studenten deze tools ervaren. Data zijn verzameld met behulp van online vragenlijsten en interviews. Een aanzienlijk deel van de docenten rapporteert een tijdsbesparing, deze ging niet te koste van de feedbackkwaliteit. De overgrote meerderheid van de docenten wil in de toekomst vaker met de digitale feedbacktool werken. Studenten blijken de digitale feedback positief te waarderen.

**Online feedback on writing: better feedback in less time**.

Feedback is a powerful teaching technic to raise students’ performance, provided that the feedback is informative on how to improve, is given in a timely manner and students have the opportunity to act upon it. Therefore, many institutions want their students to receive feedback on their performance more frequently. However, this policy seriously increases teachers’ workload, especially when written work has to be assessed. Recently, digital marking tools are developed that have the potential to help teachers to give good feedback in an efficient manner and thus to save time.

This study reports about teachers’ and students’ use and evaluation of two digital marking tools (Turnitin and WinVision) in normal educational practice. Data were gathered by online questionnaires and interviews. A substantial number of teachers reported that the use of the marking tool was time saving, without negatively impacting on the feedback quality. The vast majority of the teachers is willing to use the marking software in the future. Students were positive about the digital feedback tools as well.

Nederlandstalige trefwoorden: assessment of writing, online feedback, digital marking

Engelstalige trefwoorden: assessment of writing, online feedback, digital marking

**Online feedback op schriftelijk werk: betere feedback in minder tijd**.

Uit reviewstudies van Hattie en Timperley (2007), en Shute (2008), blijkt dat feedback een van de krachtigste strategieën is om het rendement van onderwijs te verhogen. Voorwaarde is dat de feedback de student voldoende aanknopingspunten biedt om de benodigde stappen te zetten van feitelijk naar gewenst prestatieniveau, dat de feedback tijdig wordt gegeven en er mogelijkheden zijn om de feedback toe te passen. Goede feedback helpt de student om te zien waar hij naar toe moet werken (’Where am I going?’), hoe ver hij nog van dat doel verwijderd is (‘How am I going?’) en wat de eerstvolgende stap moet zijn om daar te komen (‘Where to next?’), aldus Hattie & Timperley (2007).

De toetsmomenten in een cursus zijn veelal de enige momenten waarop de student feedback krijgt, maar goede feedback is daarmee niet gegarandeerd. Terecht krijgen toetsing en feedback de laatste jaren veel aandacht in het onderwijsbeleid van universiteiten en hogescholen. Zo wordt het belang benadrukt van een formatieve tussentoets en worden procedures ingesteld voor afstemming en transparantie in de beoordeling en feedback. De keerzijde hiervan is echter dat de werkdruk van docenten toeneemt. Verantwoorde toetsing en feedback vereisen nu eenmaal een stevige tijdsinvestering, vooral als het gaat om schrijfproducten.

Inmiddels zijn er diverse online systemen ontwikkeld voor het beoordelen en becommentariëren van geschreven werkstukken (Welch, 2012; Van Boxel, 2013). De vraag is of ze helpen om de kwaliteit van de toetsing te verbeteren en daarbij de tijdsinvestering van docenten beheersbaar te houden. In potentie hebben ze een aantal voordelen, zoals:

1. Een ingebouwde structuur van criteria en bijbehorende feedback, die het inzichtelijk maakt voor de studenten in welke opzichten hun prestatie aan de norm beantwoordt en op welke punten ze zich moeten verbeteren.
2. Studenten kunnen ook op afstand hun feedback inzien.
3. Het gebruik van één tool binnen een opleiding bevordert de consistentie van de beoordeling en feedback.
4. Door de mogelijkheden om feedback op te slaan en meermalen te gebruiken loont het de moeite om de feedback zorgvuldig te formuleren, wat de kwaliteit van de feedback ten goede komt.
5. Hergebruik van feedback bespaart tijd, doordat de feedback niet telkens hoeft te worden bedacht en uitgeschreven.
6. Vereenvoudiging van de administratie, hetgeen de logistieke rompslomp voor docenten beperkt.

Uiteraard brengt het gebruik van digitale nakijktools ook kosten met zich mee, zoals licentie- en beheerskosten, naast de tijdsinvestering voor docenten om zich de tool eigen te maken. Als die kosten aanvaardbaar zijn kunnen digitale beoordelings- en feedbacktools helpen om studenten frequent van feedback te voorzien met een aanvaardbare docenttijdsinvestering.

Om docenten en onderwijsmanagers onderbouwd te kunnen adviseren bij hun besluitvorming in deze zijn vier instellingen voor hoger onderwijs gezamenlijk met het SCALA-project gestart, met financiële ondersteuning vanuit SURF(www.surf.nl). In dit project zijn in totaal 28 pilots uitgevoerd, waarbij twee softwareapplicaties zijn gebruikt voor het beoordelen van en feedback geven op schriftelijk werk van studenten. Gewerkt is met de applicaties Turnitin (op de Vrije Universiteit, de Hogeschool van Amsterdam en de Universiteit Utrecht), en WinVision (Hogeschool Utrecht).

Turnitin (www.turnitin.com) bestaat uit drie delen: Originality check, GradeMark en PeerMark. De Originality check is een systeem voor plagiaatdetectie. GradeMark is het deel van Turnitin voor docenten waarin ze digitale feedback geven kunnen geven op schriftelijke producten en PeerMark faciliteert peer feedback. In het SCALA project is vooral gebruik gemaakt van GradeMark. GradeMark bevat drie tools waarmee al dan niet in combinatie feedback kan worden gegeven: een beoordelingsformulier (Rubric), (herbruikbare) feedback in de tekst (QuickMarks) en geschreven of gesproken algemeen commentaar (GeneralComments).

In WinVision Digitaal Portfolio leggen studenten een digitaal portfolio aan om hun competentieontwikkeling aan te tonen, te monitoren en te laten beoordelen. De docent kan zowel met gestandaardiseerde als open feedback werken. Gebruik is gemaakt van de tool beoordelingscriteria, waarmee de docent kan putten uit een bibliotheek van opgeslagen criteria alsook zelf criteria nieuw aanmaken.

**Methodologie**

*Pilots*

In het studiejaar 2012-2013 en in 2013-2014 zijn 28 pilots uitgevoerd en geëvalueerd, 22 met Turnitin en zes met WinVision. De keuze van de systemen lag per instelling vast: Turnitin voor de VU, HvA en UU; WinVision voor de HU.

Onder ‘pilot’ wordt hier verstaan: een reguliere cursus binnen een opleidingsprogramma, waarin de docent of het docententeam voor het eerst schrijfopdrachten nakijkt en/of van feedback voorziet via Turnitin of WinVision. De aard en omvang van de schrijfopdrachten liep per cursus uiteen. Vijf pilots zijn als ‘herhaalpilot’ uitgevoerd. Een herhaalpilot is een cursus die in hetzelfde of volgende studiejaar opnieuw is uitgevoerd in dezelfde opzet en met hetzelfde systeem, al dan niet met dezelfde docent(-en).

Om docenten voor de pilots te werven zijn binnen alle vier instellingen informatiebijeenkomsten gehouden over de mogelijkheden van Turnitin respectievelijk WinVision, alsook over de ondersteuning vanuit het SCALA-projectteam. Het projectteam werd gevormd door één tot twee onderwijsadviseurs per instelling. Deze adviseurs adviseerden en schoolden de docenten van de eigen instelling bij het gebruik van de online tools. Als projectteam waren zij verantwoordelijk voor de opzet en uitvoering van het onderzoek. Omdat de pilots plaatsvonden in het kader van ‘lopend onderwijs’, werden de docenten vrij gelaten in de manier waarop zij de digitale tools gebruikten. Er is tussen de pilots dan ook onderling een grote variatie, bijvoorbeeld in het aantal betrokken docenten per pilot; studentenaantal, opleidingsjaar, vakgebied, functie van de feedback (formatief of summatief) en het type schrijfproduct dat beoordeeld werd (uiteenlopende opdrachten, met als belangrijkste variabele de omvang).

De docenten die deelnamen aan de pilots kregen vooraf een hands-on training in het door hen te gebruiken systeem van twee uur. De studenten werden vervolgens door hun docenten geïnstrueerd.

*Onderzoeksvragen*

De pilots zijn geëvalueerd aan de hand van de volgende onderzoeksvragen:

1.In hoeverre heeft het gebruik van een digitale feedback- en beoordelingstool bijgedragen aan een verbetering van de feedback- en beoordelingskwaliteit?

2. In hoeverre heeft het gebruik van een digitale feedback- en beoordelingstool bijgedragen aan tijdwinst voor de docent bij het feedback geven en/of beoordelen?

*Participanten*

In totaal zijn data verkregen over 26 van de 28 pilots. Eén pilot werd op het laatste moment door de betreffende docent afgelast toen het niet mogelijk bleek om de zelf geformuleerde beoordelingscriteria –die waren opgemaakt met een van het systeem afwijkende systematiek- in te voeren. Bij een andere pilot is de evaluatie om roostertechnische redenen zowel bij de docenten als studenten achterwege gebleven.

Bij alle 26 pilots is de studentvragenlijst afgenomen. De docentvragenlijst is bij 24 pilots ingevuld. Bij twee pilots hebben de docenten de vragenlijst niet ingevuld, ook niet na een herhaald verzoek.

Van de 94 bij de pilots betrokken docenten hebben er 67 de vragenlijst geheel of voor een aanzienlijk deel ingevuld, dat betekent een responspercentage van 71,3 %. Van de 2406 studenten die aan een van de pilotcursussen hebben meegedaan hebben er 575 de vragenlijst geopend. De 23 studenten die na de eerste paar vragen over instelling en opleiding geen enkele vraag hebben ingevuld van het inhoudelijk deel zijn uit de dataset verwijderd, zodat de overall respons van de studenten 23,0% is (zie de tabellen 1 en 2 voor de respons per instelling). Verklaringen voor de lage respons bij de studenten kunnen zijn dat ze vlak voor de online vragenlijst ook de gebruikelijke cursusevaluatie van hun opleiding hadden moeten invullen. Een andere reden is dat niet alle docenten de link naar de vragenlijst snel genoeg doorstuurden naar hun studenten, waardoor het moment voor de studenten gepasseerd was waarop ze nog bereid waren terug te blikken.

Tabel 1. Respons docenten per instelling

|  |  |  |  |
| --- | --- | --- | --- |
| Instelling | Populatie | Aantal respondenten | Responspercentage |
| VU | 28 | 20 | 71,4 |
| HvA | 39 | 26 | 66,7 |
| UU | 20 | 14 | 70,0 |
| HU | 7 | 7 | 100,0 |
| Totaal | 94 | 67 |  |

Tabel 2. Respons studenten per instelling

|  |  |  |  |
| --- | --- | --- | --- |
| Instelling | Populatie | Aantal respondenten | Responspercentage |
| VU | 709 | 184 | 26,0 |
| HvA | 1085 | 111 | 10,2 |
| UU | 431 | 168 | 39,0 |
| HU | 181 | 89 | 49,2 |
| Totaal | 2406 | 552 |  |

Al met al is de respons van de docenten met 71% acceptabel. Dit is belangrijk omdat het SCALA-project primair op de docenten was gericht, het gaat immers uiteindelijk om de verbetering van de feedback- en beoordelingskwaliteit en de verlichting van hun werkdruk. De respons van de studenten mag met 23% aan de lage kant zijn, het is wel een gangbaar gemiddelde voor dit soort evaluaties.

*Instrumenten*

Gekozen is voor het verzamelen van kwantitatieve en kwalitatieve data. De kwantitatieve data zijn verzameld via twee online vragenlijsten, een voor de docenten en een voor de studenten. De link naar deze vragenlijsten is in alle pilots ongeveer twee weken na de laatste cursusbijeenkomst verstuurd. De kwalitatieve data zijn verzameld via open antwoordmogelijkheden in genoemde vragenlijsten en interviews. Voor de start van elke pilot is een interview met de docent gehouden, of in geval er meerdere docenten bij de uitvoering van dezelfde pilot betrokken waren, met de cursuscoördinator. In dit startinterview stonden de verwachtingen ten aanzien van het systeem centraal. Na afloop van de cursus is de docent of het docententeam opnieuw geinterviewd, waarin de ervaringen tijdens de pilot werden besproken.

De vragenlijsten voor de docenten en studenten zijn door het projectteam ontwikkeld. Dit gebeurde op basis van literatuurstudie, aangevuld met praktijkervaringen van de projectteamleden van de VU en de HvA, die al bekend waren met Turnitin.

Beide vragenlijsten zijn zoveel mogelijk analoog opgebouwd. Ze openen met enkele vragen die bedoeld zijn om respondentgegevens te verzamelen: over de instelling waar men doceert/studeert, de naam van de opleiding en die van de cursus. Aan de studenten is daarnaast gevraagd naar de naam van de docent.

Het inhoudelijke deel bestaat uit items over de thema’s Gebruiksgemak, Feedbackkwaliteit en Tijdbesparing (alleen in de docentvragenlijst), tabel 3 bevat enkele voorbeelden van de gestelde vragen. Daarnaast is zowel aan docenten als aan studenten om een algemeen eindoordeel gevraagd (in de zin van ‘ik zou vaker met dit systeem willen werken’, met de mogelijkheid om het antwoord toe te lichten). Per thema zijn vijf tot tien items opgenomen, meestal in de vorm van stellingen waarbij de respondent gevraagd wordt om op een Likertschaal van 1 (zeer oneens ) tot 5 (zeer eens) en de optie ‘niet van toepassing’ zijn oordeel te geven. In beide vragenlijsten is bij de items met een meer overkoepelende strekking steeds de mogelijkheid tot toelichting gegeven. Omdat er twee verschillende softwareapplicaties zijn gebruikt (Turnitin en WinVision) en er binnen Turnitin gekozen kon worden voor het gebruik van verschillende tools is een aantal vragen ook toolspecifiek gesteld. De totale docentvragenlijst omvatte 67 vragen, de studentvragenlijst omvatte 45 vragen, de respondenten zijn via routing door de vragenlijst geleid. De invulling van de studentvragenlijst was anoniem, in de docentvragenlijst werd de respondenten gevraagd om hun naam in te vullen, maar dit was niet verplicht.

De docent- en studentvragenlijsten, alsook de interviewleidraad zijn opvraagbaar bij de auteurs.

Tabel 3. Voorbeelden van vragen aan docenten resp. studenten over de thema’s

|  |  |  |
| --- | --- | --- |
| Thema | Docentvragenlijst | Studentvragenlijst |
| Gebruiksgemak | Ik was snel gewend aan het systeem (schaal 1 t/m 5, nvt)  Het systeem hielp me om overzicht te houden op de ingeleverde en nagekeken opdrachten (schaal 1 t/m 5, nvt) | Ik vond het gemakkelijk om de opdrachten digitaal in te leveren (schaal 1 t/m 5, nvt)  Ik kon de feedback op mijn werk gemakkelijk vinden (schaal 1 t/m 5, nvt) |
| Feedbackkwaliteit | De kwaliteit van mijn feedback c.q. beoordeling werd beter toen ik dit systeem ging gebruiken (schaal 1 t/m 5, nvt)  Door het gebruik van een rubric\* werd mijn feedback c.q. beoordeling duidelijker voor de student (schaal 1 t/m 5, nvt) | De feedback die ik op mijn werk heb gekregen beoordeel ik met het cijfer (1 t/m 10)….  De rubric\* maakte duidelijk op welke aspecten mijn werk beoordeeld is (schaal 1 t/m 5, nvt) |
| Tijdbesparing | In vergelijking met mijn gebruikelijke nakijktijd voor dit type opdrachten kostte het feedback geven c.q. beoordelen mij:   1. Meer tijd 2. Evenveel tijd 3. …% minder tijd |  |

\* Dezelfde vragen werden gesteld over de andere tools van Turnitin, te weten QuickMarks en GeneralComments, en voor WinVision: beoordelingscriteria.

*Uitvoering dataverzameling*

De vragenlijsten zijn omgezet in een online format. Direct na afloop van elke pilot kreeg de docent of cursuscoördinator twee e-mails van de onderzoekers, een met de link naar de docentvragenlijst de andere met de link naar de studentvragenlijst. De eerste e-mail bevatte het verzoek om de docentvragenlijst in te vullen en –in geval van een docententeam- de email met link door te sturen naar de andere docenten. De tweede e-mail was bedoeld om door te sturen naar alle studenten die aan de cursus hadden deelgenomen. De dataverzameling van beide enquêtes is afgesloten op 10 december 2013.

Het eerste interview met de docent of met het docententeam vond plaats vlak voor de start van elke pilot, het tweede interview is gemiddeld een maand na afloop van elke pilot afgenomen. Interviewers waren de leden van het projectteam, zij interviewden de docenten van hun eigen instelling, de gespreksduur was ongeveer een uur. Het tweede interview vond plaats nadat de docentvragenlijst en de studentvragenlijst waren ingevuld, de uitkomsten hiervan werden betrokken in het gesprek. De samenvattingen van de interviews zijn ter accordering en eventuele aanvulling aan de geinterviewden voorgelegd.

*Wijze van analyse*

De kwantitatieve data worden gepresenteerd met behulp van beschrijvende en inductieve statistiek, waarbij mogelijke verbanden en eventuele verschillen in gemiddelden zijn onderzocht met behulp van parametrische en non-parametrische toetsen. Getoetst is met als onafhankelijke variabelen: 1. systeem (Turnitin versus WinVision), en 2. herhaling (eerste pilots versus herhaalpilots). De interviews en de toelichtingen in de vragenlijst zijn gebruikt voor de interpretatie en illustratie van de kwantitatieve data.

**Resultaten**

De docenten zijn positief in hun eindoordeel over het digitaal feedback geven c.q. beoordelen via de door hen gebruikte systemen: 86,2% (n=56) geeft aan vaker met het door hen gebruikte systeem te willen werken. Van deze 56 docenten werkten er 51 met Turnitin en vijf met WinVision. Voor beide systemen geldt dat de docenten meerwaarde zien in de logistieke ondersteuning die het systeem biedt bij het inleveren en archiveren. Daarnaast zijn ze overtuigd geraakt van de inhoudelijke potentie die de specifieke tools hebben voor het transparanter en consistenter beoordelen en feedback geven.

De studenten die hun feedback c.q. beoordeling kregen via rubric en QuickMarks (beide tools van Turnitin, meestal in combinatie gebruikt) zijn eveneens positief: ruim 80% zou vaker op diezelfde manier hun feedback of beoordeling willen krijgen. Van de studenten die met WinVision werkten wil de helft dit vaker op deze manier.

De resultaten met betrekking tot de eerste onderzoeksvraag worden afzonderlijk gepresenteerd op docent- en studentniveau. Aangezien onderzoeksvraag 2 gesteld is vanuit het perspectief van de docent wordt hier alleen op docentniveau gerapporteerd. Hierbij worden tevens de correlaties gepresenteerd tussen tijdbesparing en feedback-c.q. beoordelingskwaliteit, alsook tussen tijdbesparing en het werken in een zgn. herhaalpilot.

*Onderzoeksvraag 1: In hoeverre heeft het gebruik van een digitale feedback- en beoordelingstool bijgedragen aan een verbetering van de feedback-en beoordelingskwaliteit?*

Voor de beantwoording van deze vraag is niet alleen de optiek van de docent van belang, maar ook die van de ontvanger, de student.

*Docenten*

Met betrekking tot de stelling ‘De kwaliteit van mijn feedback c.q. beoordeling werd beter toen ik dit systeem ging gebruiken’ zijn de docenten sterk verdeeld: 40,0% is het eens met de stelling, 36,7% antwoordt neutraal en 23,3% is het oneens. Eenzelfde antwoordpatroon doet zich voor bij de waardering voor de bijdrage van de specifieke tools binnen het systeem (zie Tabel 4).

Docenten die een kwaliteitsverbetering ervoeren lichten die toe met argumenten als bijvoorbeeld:’ omdat consequent bij elk verslag dezelfde feedback (QuickMarks) werd gebruikt. Hierdoor gebruik je consequent dezelfde criteria en eenzelfde formulering’. (Turnitin)‘.’Vaste beoordelingscriteria beschermen je tegen jezelf om alle fouten aan te strepen. Je beoordeelt nu alle producten op dezelfde wijze. Dat is kwalitatieve winst en tijdwinst’. (WinVision)

Docenten die geen kwaliteitsverbetering zagen wijten dit eerder aan tekortkomingen van het systeem: ‘Het is niet mogelijk aanpassingen in de tekst zelf te maken. Dit ervaar ik als een gemis bij het feedback geven’. (Turnitin)

‘Voor een van de opdrachten hebben we voor de beoordeling een rubric ontwikkeld, maar het was om technische redenen niet mogelijk om die in WinVision op te nemen’.

Het oordeel van docenten over een eventuele verbetering van de feedback- c.q. beoordelingskwaliteit correleert matig tot sterk positief (rs = 0,57; p< 0,01) met hun oordeel over het gebruiksgemak van het door hen gebruikte systeem. Oftewel: het gebruiksgemak van het systeem draagt aanzienlijk bij aan de mate waarin docenten een kwaliteitsverbetering ervaren op het gebied van feedback geven en/of beoordelen.

Ondanks genoemde problemen en/of minpunten geeft 86,2% van de docenten (56) vaker met het gebruikte systeem te willen werken. Van deze 56 docenten werkten er 51 met Turnitin en vijf met WinVision.

Tabel 4. Oordeel docenten m.b.t. de stelling ‘Door het gebruik van deze optie werd de kwaliteit van mijn feedback c.q.beoordeling beter’

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | (Zeer) eens (%) | Eens noch oneens (%) | (Zeer)oneens (%) | Tot. N |
| **Turnitin (overall)** | 38,9 (n=21) | 38,9% (n=21) | 22,2% (n=12) | 54 |
| Rubrics | 36,3 (n=16) | 40,9 (n=18) | 22,8 (n=10) | 54 |
| QuickMarks | 41,5 (n=22) | 35,8 (n=19) | 22,7 (n=12) | 53 |
| General Comments | 47,6 (n=20) | 26,2 (n=11) | 26,2 (n=11) | 42 |
| **WinVision** Beoordelingscriteria | n=4\* | n=2 | n=1 | 7 |

\*aangezien het hier om zeer kleine aantallen gaat worden de percentages niet vermeld.

*Studenten*

De studenten geven gemiddeld het rapportcijfer 7,1 (SD 1,38) voor de feedback op hun werk, 72,4% van hen vindt dat ze voldoende feedback op hun werk hadden gekregen. Een t-toets voor ongepaarde waarnemingen wijst uit dat er op beide punten geen significant verschil is tussen de Turnitin en de WinVisiongebruikers.

Van de 291 studenten die (in Turnitin) met behulp van een rubric zijn beoordeeld, vaak in combinatie met feedback in de tekst via QuickMarks, wil ruim 80% vaker op deze manier worden beoordeeld. Uitspraken die dit illustreren zijn bijvoorbeeld: ‘Door de rubrics is het duidelijk waar de focus ligt en hoe je dat aan kan passen en verbeteren. Erg handig’. ‘Een rubric is wat oppervlakkiger, feedback in de tekst geeft echt specifieke aanwijzingen’.

Van de 89 studenten die hun beoordeling in WinVision (via de tool Beoordelingscriteria) kregen wil ongeveer de helft (54,9%) vaker op deze manier worden beoordeeld. Ze vinden de feedback duidelijk en de digitale vorm handig: ‘Deze manier van beoordelen bevalt mij erg, omdat ik hierdoor een duidelijk beeld krijg van mijn sterke punten en punten waar ik nog aan moet werken. Hierdoor word ik gemotiveerd om mijzelf te ontwikkelen’. ’Het is erg makkelijk voor mij om mijn werk via WinVision in te leveren. Ik hoef namelijk niks meer uit te printen en/of met losse blaadjes rond te lopen, wat er voor zorgt dat ik mijn feedback formulier of juist mijn werk sneller kwijt raak’.

Sommige studenten die met WinVision werkten hadden behoefte aan meer instructie vooraf in het systeem: ‘Een digitale beoordeling is prima, maar het moet van tevoren duidelijk zijn hoe je de vraag naar een beoordeling in het digitale portfolio zet en hoe je de beoordeling weer terug krijgt. Op WinVision is het soms erg lastig terug te vinden. Een melding over goed ontvangst/verstuur bericht etc. zal helpen om dit te checken’.

*Onderzoeksvraag 2: In hoeverre heeft het gebruik van een digitale feedback- en beoordelingstool bijgedragen aan tijdwinst voor de docent bij het feedback geven en/of beoordelen?*

Deze vraag heeft alleen betrekking op de docent. Hierbij wordt tevens onderzocht of een eventuele tijdbesparing plaatsvindt zonder dat de docent hierbij een achteruitgang in feedback- c.q. beoordelingskwaliteit ervaart. Ook wordt nagegaan of er sprake is van een uitgesteld profijt: wellicht werkt het systeem de eerste keer nog niet tijdbesparend, maar wel daarna, als de investeringen eenmaal zijn gedaan.

Op de vraag ‘In vergelijking met mijn gebruikelijke nakijktijd voor dit type opdrachten kostte het feedback geven c.q. beoordelen mij 1. Meer tijd; 2. Evenveel tijd; 3. ..% Minder tijd’ antwoordt 38,5% van de docenten dat zij minder dan de voor hen gebruikelijke tijd hebben besteed bij het feedback geven c.q. beoordelen, deze tijdbesparing varieerde tussen 10 en 50%.

In Tabel 5 wordt de tijdbesparing vermeld, uitgesplitst per systeem.

Tabel 5. Ervaren tijdbesparing docenten bij gebruik Turnitin resp. WinVision

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Minder tijd (%) | Evenveel tijd % | Meer tijd % | Tot.N |
| Turnitin | 40,7 (n=24) | 45,7 (n=27) | 13,6 (n=8) | 59 |
| WinVision | n=1\* | n=3 | n=2 | 6 |
| Overall | 38,5 (n=25) | 46,2 (n=30) | 15,4 (n=10) | 65 |

\*aangezien het hier om zeer kleine aantallen gaat worden de percentages niet vermeld.

Docenten die tijd hadden bespaard lichten dit toe met uitspraken als bijvoorbeeld: ‘Handig om voorgeprogrammeerde feedback te slepen’. (Turnitin) ’Dit bespaart me veel e-mailverkeer’. (Turnitin) ‘Met de vaste set criteria kan ik een grotere hoeveelheid producten nakijken’. (WinVision)

Docenten die evenveel nakijktijd als gewoonlijk nodig hadden waarderen over het algemeen wel de logistieke voordelen van het systeem, dat geldt voor zowel de Turnitin- als de WinVisiongebruikers: ‘Voordeel van het systeem is dat het de mogelijkheid biedt om bestanden op een gestructureerde overzichtelijke manier te bewaren/te plaatsen. Deadlines kunnen makkelijker gehanteerd worden omdat studenten documenten niet meer kunnen plaatsen. Ook is het duidelijk voor hen wanneer er feedback verwacht kan worden’. (Turnitin)

‘De takenlijst die zichzelf opschoont was erg handig’.(WinVision)

Docenten die meer tijd besteedden geven vaak als reden aan dat men tijd nodig had om het systeem te leren kennen of ontwikkeltijd nodig had voor het ‘vullen’ van de tools.

Het oordeel over de tijdbesparing correleert niet met het oordeel over de feedback- c.q. beoordelingskwaliteit. Docenten die tijd bespaarden hebben dit dus niet moeten bekopen met een verminderde feedback- c.q. beoordelingskwaliteit.

Van de 65 docenten waren er twaalf die de cursus uitvoerden als herhaalpilot. In Tabel 6 wordt de tijdbesparing van de docenten die een pilot voor het eerst uitvoerden en die van de docenten in de herhaalpilots gepresenteerd.

Tabel 6. Tijdbesparing docenten herhaalpilots versus docenten eerste pilot

|  |  |  |  |
| --- | --- | --- | --- |
|  | N | M\* | SD |
| Eerste pilot | 53 | 2,13 | 0,68 |
| Herhaalpilot | 12 | 2,67 | 0,65 |
| Overall | 65 | 2,23 | 0,70 |

\* Gemiddelde tijdbesparing op een schaal van 1 (meer tijd besteed)-2 (evenveel tijd besteed)-3 minder tijd besteed

De Mann-Whitney U test wijst uit dat de docenten van de herhaalpilots meer tijd hebben bespaard dan de docenten voor wie dit de eerste keer was. De docenten in de herhaalpilots hadden minder nakijktijd nodig (gemiddelde rangordescore 44,46) dan de docenten in de eerste pilots (gem. rangordescore 30,41). Dit verschil is significant (Z=-2,535;p=0,011).

**Conclusie en discussie**

De resultaten ten aanzien van de twee onderzoeksvragen worden hieronder per onderzoeksvraag samengevat:

*Kwaliteit feedback-c.q. beoordeling*

Volgens 40,0 % van de docenten is de kwaliteit van hun feedback c.q. beoordeling verbeterd. De docenten beschouwen het als een verbetering dat er uniformer wordt beoordeeld omdat het hele docententeam gebruik maakt van dezelfde rubric, QuickMarks of beoordelingscriteria. Ze vinden het een verbetering dat voor de beoordeling van schrijfproducten consequent dezelfde criteria worden gebruikt en dat de beoordeling transparanter is geworden voor de studenten. De docenten die geen kwaliteitsverbetering zagen (23,3%) waren vaak niet tevreden over de mogelijkheden of de inhoudelijke invulling van de tools, of hadden technische problemen gehad met het systeem.

De studenten zijn over het algemeen tevreden over de feedback die ze hebben gekregen, dat geldt voor beide systemen. Ze waarderen de feedbackkwaliteit met gemiddeld het rapportcijfer 7,1 (SD 1,38). Bijna driekwart (72,4 %) vindt de hoeveelheid feedback voldoende. Opvallend hierbij is dat de studenten vooral de combinatie van rubric en QuickMarks als leerzaam ervaren. (Verder uitgewerkt in Van der Hulst, Van Boxel & Meeder, 2014).

*Tijdbesparing*

Van de docenten heeft 38,5% minder dan de gebruikelijke tijd besteed bij het feedback geven c.q. beoordelen. De opgegeven tijdbesparing loopt daarbij uiteen van 10 tot 50%. 46,2% heeft evenveel tijd als gewoonlijk nodig gehad en 15,4% heeft meer tijd moeten investeren. Dat er meer tijd besteed is komt vooral door de investering die nodig was om inhoudelijk vulling te geven aan de tools.

Docenten die tijd bespaard hebben bij het beoordelen c.q. feedback geven vinden niet dat de kwaliteit van hun feedback c.q. beoordeling erop achteruit is gegaan. Met andere woorden: ze hebben hun tijdwinst niet moeten ‘bekopen’ met een achteruitgang in feedbackkwaliteit.

Ervaring in het werken met het systeem lijkt te lonen: docenten die een cursus uitvoerden als een ‘herhaalpilot’ gaven vaker aan dat ze tijdswinst hadden geboekt dan docenten in cursussen waarin de tool voor het eerst werd ingezet. Ze konden profiteren van het eerder verrichte ontwikkelwerk.

**Discussie**

Deze resultaten maken duidelijk dat digitaal feedback geven en beoordelen –wanneer de benodigde investeringen eenmaal gedaan zijn- helpt om de feedback- c.q. beoordelingskwaliteit te verbeteren en daarbij de werkdruk beheersbaar te maken of te houden. Tot slot passen enkele kanttekeningen die van belang zijn bij de interpretatie van de gegevens.

In het onderzoek is geen rekening gehouden met de mate van innovativiteit die sommige docenten al aan de dag legden voordat ze met een van de hier gepresenteerde tools aan het werk gingen. Zo werkten sommigen zelf al langer met een bestand met veel voorkomende feedbackopmerkingen en was papierloos nakijken voor een aantal docenten vanzelfsprekend. Op de instellingen waar de docenten werkten was inleveren via de elektronische leeromgeving gebruikelijk, maar er waren ook docenten die er de voorkeur aan gaven om het na te kijken werk buiten de leeromgeving uit te wisselen, via mail of postvak. Kortom, de beginsituatie van de docenten liep in dit opzicht uiteen, ze zijn hierop ook bevraagd, maar het was met deze aantallen niet mogelijk om daarin substantieel verschillende groepen te onderscheiden. De oordelen van docenten over het effect van de tools op hun beoordelings- c.q. feedbackkwaliteit en op hun tijd zijn hiervoor dus niet gecorrigeerd.

De terminologie in de vragenlijst kan voor verwarring gezorgd hebben. Het komt voor dat studenten de vragen over de rubric hebben overgeslagen, terwijl ze in de desbetreffende pilot wel met een rubric werden beoordeeld. Ze hadden geen rubric herkend of ze hadden die misschien wel gezien maar waren vergeten wat het was of dat het ‘rubric’ heette.

Onze evaluatie richtte zich op de kwaliteit van feedback en beoordeling door docenten en studenten. In een enkele pilot was er sprake van een verschil tussen perceptie van de docenten (zijzelf vinden die verbeterd) en die van de studenten (die uiteraard niet kunnen vergelijken), die er een vrij laag rapportcijfer 6,2 aan toekenden. Of hier werkelijk een verbeteringsslag heeft plaatsgevonden is lastig te zeggen, daarvoor zouden we de studentevaluaties van diezelfde cursus in voorgaande jaren moeten raadplegen.

De pilots zijn uitgevoerd door docenten die voor het overgrote deel onbekend waren met Turnitin dan wel WinVision, in alledaagse onderwijscontexten met de onvermijdelijke malheur die daarbij komt kijken zoals tijdgebrek, last minute wijzigingen, uitval van de internetverbinding et cetera. Ook hadden sommige docenten bepaalde functies nog niet ontdekt en bleken er misvattingen te bestaan over wat het systeem wel en niet kan. Dit maakt dat de feedback- en nakijktools niet altijd op de meest doordachte of best mogelijke manieren zijn ingezet. De uitkomsten tonen daarom niet de volledige potentie van het nakijken en beoordelen door middel van deze systemen. Evenmin wordt exact duidelijk wat precies de merites zijn van de verschillende specifieke tools, afzonderlijk of in bepaalde combinaties gebruikt. Om daar meer inzicht in te krijgen hadden de pilots systematischer ontworpen moeten worden.

**Literatuur**

Boxel, P. (2013). Trends and developments in On-screen Marking and Grading of Essays [abstract]. *In:* *Book of abstracts of the 18TH International Conference on Technology Supported Learning and Teaching, November 28-30*, Berlin: Online Educa Berlin, 62.

Hattie, J., & Timperley, H. (2007). *The Power of Feedback. Review of Educational Research,77(1)*, 81–112.

Hulst, J. van der, Boxel, P. van, Meeder, S. (2014). Digitilizing feedback: Reducing teachers’ time investment while maintaining feedback quality*. In:* *Proceedings of the 13th European Conference on e-Learning – ECEL 2014, Aalborg University,Copenhagen, Denmark; 30-31 October 2014*, 243 – 250.

SCALA-project, zie: https://www.surf.nl/kennis-en-innovatie/innovatieprojecten/startdatum-2012/scala-scaffolding-assessment-for-learning.html

Shute, V.J. (2008). *Focus on formative feedback. Review of educational research, 78(1*), 153-189.

Weaver, M.R. (2006). *Do students value feedback? Student perceptions of tutors’ written responses. Assessment & Evaluation in Higher Education,31(3),* 379–394.

Welch, L. (2012). *The role of computer aided marking and feedback in the assessment process: a critical review.* The University of Edinburgh, Retrieved from http://online.education.ed.ac.uk/2012/02/lyndsey-welch-the-role-of-computer-aided-marking-and-feedback-in-the-assessment-process-a-critical-review/