

**Evaluatie Comproved  
Rapport Educate-IT**

Stephanie Kruiper  
Januari 2021

## **Samenvatting**

### **Context en Interventie**

In de cursus Inleiding Onderwijswetenschappen (blok 1, Sociale Wetenschappen, 2021-2022) is gebruik gemaakt van comparatief beoordelen voor het beoordelen van de langlopende opdracht van de cursus. Deze vorm van beoordelen gaat om het paarsgewijs vergelijken van opdrachten voor een beoordeling, in plaats van het afzonderlijk beoordelen van iedere opdracht. Het programma Comproved is hierbij ingezet, waarbij zowel tijdens het formatieve als summatieve beoordelingsmoment gebruik is gemaakt van de functionaliteiten van het programma door zowel studenten als docenten.

### **(verondersteld) Mechanisme**

Het werken met comparatief beoordelen in Comproved had in deze cursus meerdere doelen:

- Betrouwbaarheid en kwaliteit van beoordelen verhogen
- Feedbackgeletterdheid bij studenten bevorderen
- Geen verhoging, liefst verlaging van de nakijkdruk bij docenten

### **Uitkomsten**

Een schriftelijke evaluatie (N = 51) en evaluatiegesprek (N = 2) bij studenten en een evaluatiegesprek met docenten (N = 7) laten zien dat comparatief beoordelen door middel van Comproved als positief wordt ervaren. Comproved maakt statistisch inzichtelijk wat de betrouwbaarheid van de beoordeling is. Studenten en docenten vinden het daarbij erg prettig dat opdrachten door meerdere ogen worden bekeken en gezamenlijk een beoordeling tot stand komt. Comproved kan een goede rol spelen in het bevorderen van feedbackgeletterdheid bij studenten. Studenten bekijken in het programma de opdrachten van hun medestudenten en de feedback die daarop is gegeven, en zien de meerwaarde daarvan in het verbeteren van hun eigen werk. Om de kwaliteit van de feedback te verbeteren en feedbackgeletterdheid verder te bevorderen zijn meerdere feedbackrondes volgens de studenten nodig. De docenten zijn aan het nakijken van de werken niet meer tijd kwijt dan voorheen, al is wel inwerktijd nodig. Ze zien de meerwaarde van het werken met Comproved voor studenten en willen het programma blijven gebruiken.

### **Opvallendheden**

- Studenten geven aan dat de kwaliteit van de feedback van medestudenten soms te wensen overliet, en na de eindbeoordeling waren sommige studenten minder positief over comparatief beoordelen. De keuzes die gemaakt worden in het onderwijs(ontwerp) zijn erg belangrijk om hier verbeteringen in aan te brengen.
- Het werken met comparatief beoordelen vraagt een grotere aanpassing in de gehele cursus, waarbij werken met Comproved een onderdeel van die aanpassing is.

### **Belangrijkste kanttekeningen**

- Deze evaluatie is uitgevoerd door de docent die dit project is gestart en als enige uit de cursus gewerkt heeft met de 'achterkant' van Comproved. Hierdoor kan een mogelijke bias zijn ontstaan, al is geprobeerd de uitkomsten van de evaluaties zo evenwichtig mogelijk weer te geven.
- Het aantal deelnemers aan de evaluatie is beperkt, waardoor wellicht geen volledig beeld is ontstaan van de ervaringen met Comproved bij de studenten.

## Introductie en doelen

De cursus Inleiding Onderwijswetenschappen wordt aangeboden in blok 1 van het studiejaar voor alle eerstejaars bachelorstudenten Onderwijswetenschappen en de Academische Lerarenopleiding (ALPO) aan de faculteit Sociale Wetenschappen. Daarnaast is de cursus voor studenten buiten deze studies als keuzevak te volgen, bijvoorbeeld als onderdeel van een minor.

Bij de cursus waren in september 2021 131 studenten en 7 docenten betrokken. In de cursus is gewerkt met comparatief beoordelen om de langlopende opdracht van de cursus (het schrijven van een inleiding van een wetenschappelijk paper) van feedback te voorzien en te beoordelen. Deze vorm van beoordelen gaat om het paarsgewijs vergelijken van opdrachten voor een beoordeling, in plaats van het afzonderlijk beoordelen van iedere opdracht met behulp van een beoordelingsformulier of rubric. Comparatief beoordelen gaat ervan uit dat het makkelijker is om prestaties te vergelijken dan ze los van elkaar te moeten beoordelen. Het wordt gezien als een betrouwbare methode, vooral voor beoordelen met meerdere docenten (Lesterhuis et al., 2017; van Daal et al., 2019). Daarnaast kan comparatief beoordelen ook formatief worden ingezet, waarbij het voordeel is dat studenten meerdere opdrachten van medestudenten zien en ook de feedback op opdrachten van medestudenten kunnen inzien om hun eigen opdracht verder te verbeteren.

Om deze manier van beoordelen in te zetten is gebruik gemaakt van het programma Comproved. Het werken met comparatief beoordelen en Comproved had meerdere doelen:

- Betrouwbaarheid en kwaliteit van beoordelen verhogen
- Feedbackgeletterdheid bij studenten bevorderen
- Geen verhoging, liefst verlaging van de nakijkdruk bij docenten

In de cursus is een aantal aanpassingen gedaan om deze doelen te bereiken, waarbij Comproved een rol speelde:

- We ontwikkelden de feedbackgeletterdheid van de studenten met de toevoeging van verschillende opdrachten in de werkgroepen. Hierbij is ook Comproved gebruikt, waarbij met behulp van zogenoemde 'ankerteksten' een beeld is geschetst van de kwaliteit van verschillende voorbeeldopdrachten. Studenten vergeleken de vooraf ingevoerde anker teksten (uit eerdere jaren) met elkaar en gaven aan wat het ene verslag beter maakt dan het andere. Hieruit ontstond een gesprek over de beoordelingscriteria van de opdracht.
- Studenten vergeleken anoniem elkaars werken en gaven feedback op elkaars teksten via Comproved. Docenten deden mee in deze feedbackronde. Studenten zagen na deze ronde anoniem alle teksten van medestudenten op een rangorde, waarbij hun eigen plaats in die rangorde stond aangegeven. Alle andere teksten en de feedback daarop waren in te zien door de studenten.
- De summatieve beoordeling van de opdracht is via Comproved uitgevoerd. Docenten gebruikten het programma om vergelijkingen te maken. De opdrachten zijn naar aanleiding van deze vergelijkingen in Comproved gerangschikt op een rangorde. Op basis hiervan zijn de cijfers bepaald.

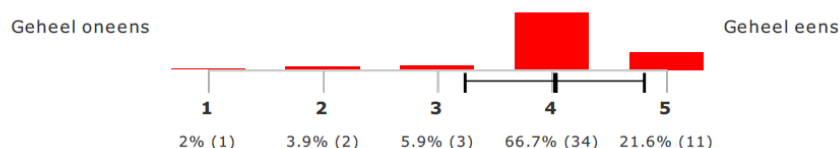
## Evaluatieonderzoek

Om de bruikbaarheid van Comproved te beoordelen is zowel bij de studenten als docenten informatie verzameld. Bij de studenten is dit gedaan door middel van een panelgesprek en een vragenlijst. Bij de docenten heeft de evaluatie plaatsgevonden gelijk met het evaluatiegesprek over de cursus. De evaluatie is afgenomen en geanalyseerd door Stephanie Kruiper (co-coördinator en mede-docent in de cursus).

### Studenten – vragenlijst (N = 51)

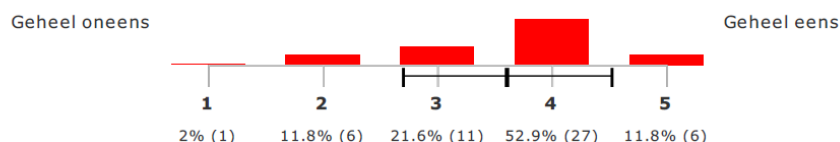
De studenten hebben in het algemene evaluatieformulier van de cursus extra vragen beantwoord over comparatief beoordelen en Comproved. Het relevante onderdeel van het evaluatieformulier is hieronder weergegeven. Daarnaast is de studenten gevraagd een toelichting te geven op de laatste 2 weergegeven vragen. In bijlage 1 zijn alle antwoorden op deze open vragen weergegeven.

#### De toetsing als geheel sloot goed aan bij de inhoud van de cursus



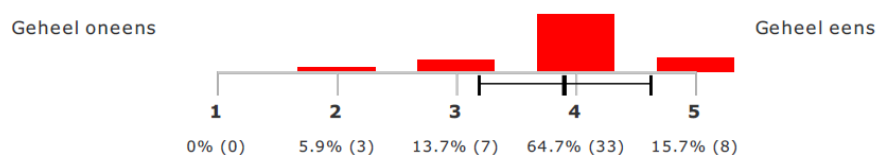
N = 51
Avg. = 4.0
Std. dev. = 0.8

#### De toetsing als geheel sloot goed aan bij het niveau van de cursus



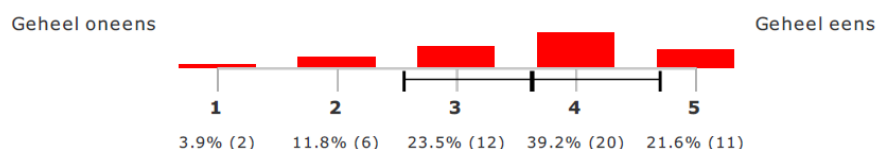
N = 51
Avg. = 3.6
Std. dev. = 0.9

#### De toetsing als geheel sloot goed aan bij het niveau van de opleiding



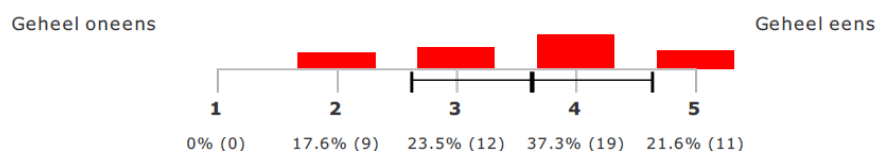
N = 51
Avg. = 3.9
Std. dev. = 0.7

#### De criteria waaraan de opdrachten moesten voldoen waren voor mij duidelijk.



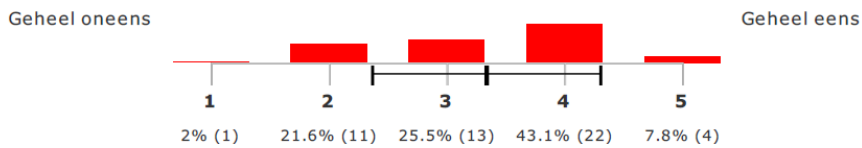
N = 51
Avg. = 3.6
Std. dev. = 1.1

#### De feedback die ik tijdens de cursus heb ontvangen (op opdrachten, conceptversies, tentamen) was bruikbaar



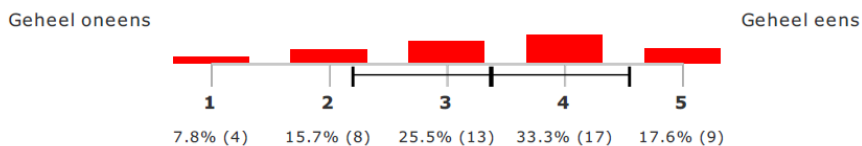
N = 51
Avg. = 3.6
Std. dev. = 1.0

**Ik heb tijdens deze cursus een goed beeld gekregen van mijn sterke en verbeterpunten in academisch schrijven.**



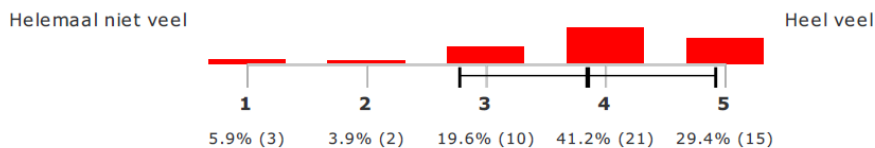
N = 51  
Avg. = 3.3  
Std. dev. = 1.0

**Het werken met comparatief beoordelen heeft mij geholpen om mijn academische schrijfvaardigheden te verbeteren.**



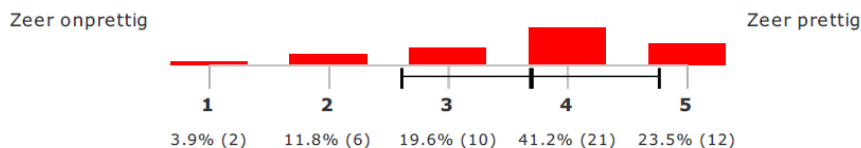
N = 51  
Avg. = 3.4  
Std. dev. = 1.2

**In hoeverre heb je gebruik gemaakt van het bekijken van andermans werk en feedback in Comproved?**



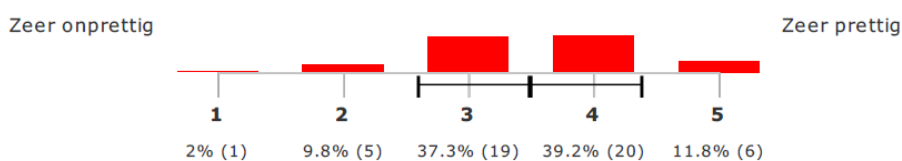
N = 51  
Avg. = 3.8  
Std. dev. = 1.1

**Ik vond het werken met comparatief beoordelen voor de feedback op de langlopende opdracht:**



N = 51  
Avg. = 3.7  
Std. dev. = 1.1

**Ik vond het werken met comparatief beoordelen voor de eindbeoordeling van de langlopende opdracht:**



N = 51  
Avg. = 3.5  
Std. dev. = 0.9

Over het algemeen waren studenten tevreden over de toetsing, waarbij het grootste deel van hen het eens is met de stellingen dat de toetsing goed aansluit bij de inhoud en het niveau van de cursus, en het niveau van de opleiding. Hierbij moet genoemd worden dat de toetsing bestond uit zowel een langlopende opdracht (waar Comproved bij is ingezet) en een tentamen. Ten aanzien van de langlopende opdracht vonden de studenten het redelijk duidelijk waaraan de opdracht moest voldoen ( $M = 3.6$ ,  $SD = 1.1$ ). De ontvangen feedback op de opdracht werd als bruikbaar ervaren ( $M = 3.6$ ,  $SD = 1.0$ ). Studenten hebben enigszins een beeld kunnen vormen van hun sterke- en verbeterpunten in academisch schrijven ( $M = 3.3$ ,  $SD = 1.0$ ). Het werken met comparatief beoordelen hielp de studenten enigszins om hun schrijfvaardigheden te verbeteren ( $M = 3.4$ ,  $SD = 1.2$ ). Voor Comproved als programma blijkt dat studenten veel gebruik hebben gemaakt van het bekijken van andermans werk en

feedback ( $M = 3.8$ ,  $SD = 1.1$ ). Het werken met Comproved werd over het algemeen als prettig ervaren, zowel voor de feedbackronde ( $M = 3.7$ ,  $SD = 1.1$ ) als voor de eindbeoordeling ( $M = 3.5$ ,  $SD = 0.9$ ).

#### *Feedbackronde*

In de toelichting op deze laatste twee vragen was de overgrote meerderheid van de studenten positief, waarbij onder andere werd genoemd dat het zien van alle andere (anonieme) papers en de feedback die daarop was gegeven een grote toegevoegde waarde had. Daarnaast waardeerden studenten het feit dat ze van meer dan 1 persoon feedback ontvingen. Niet alleen kregen ze door comparatief beoordelen een idee van de verschillen in opdrachten, maar ze konden de voorbeelden en feedback uit hoger geplaatste opdrachten gebruiken voor het verbeteren van hun eigen werk: 'Ik vond het fijn dat je op deze manier veel feedback kon krijgen en daardoor kon je je langlopende opdracht gericht verbeteren. Ook kreeg je zo een beetje een idee hoe anderen het schrijven hadden aangepakt. Hier kon je zelf ook weer ideeën of verbeterpunten uithalen.' Eén student gaf aan: 'Ik merkte dat je juist ook veel leert van andermans feedback'. Daar tegenover stond wel dat niet iedere student het geven van feedback serieus had genomen, waardoor niet iedereen van kwalitatief goede feedback was voorzien: 'Ik vond het best wel prettig dat ik van meerdere mensen feedback kreeg. Het principe is goed maar niet iedereen deed er serieus aan mee waardoor je geen betrouwbaar resultaat hebt, dat is erg jammer!'. 18 van de 51 studenten gaven aan dat hierdoor problemen ontstonden, waardoor 5 van deze 18 studenten ook minder tevreden waren over het werken met comparatief beoordelen.

#### *Eindbeoordeling*

In de toelichting op de vraag of studenten het comparatief beoordelen een fijne werkwijze vonden voor de eindbeoordeling van hun opdracht, gaf het grootste deel aan dat zij comparatief beoordelen als een eerlijke manier van beoordelen zagen: 'Ik denk dat er een eerlijke en goede ranking uit zal komen. Soms is het gewoon erg lastig om aan bepaalde eisen/elementen een cijfertje te hangen, zoals '0,3 van de 0,7 op bron gebruik'. Dit lijkt me een top manier om te beoordelen, omdat het heel relatief is. En omdat er heel duidelijk aangegeven moet worden waarom de een dan beter is dan de ander, heb ik het idee dat er goed in de papers gedoken wordt en dat dit een serieuze en geschikte manier van beoordelen is.' Er werd ook genoemd dat het een fijn idee is dat meerdere docenten naar het ingeleverde werk kijken in plaats van één. Er waren nog wel wat zorgen, die vooral leken neer te komen op het feit dat studenten nog niet altijd goed begrepen hadden hoe de beoordeling werkte: 'Ik heb liever dat mijn werk door een expert wordt nagekeken, dan door mijn medeleerlingen. Ik weet namelijk niet in hoeverre iedereen serieus zijn beoordelingen uitvoert. Bij sommige studenten heb ik het idee dat ze zomaar wat aanklikken.' Dit was een onterechte zorg, omdat docenten de eindversies hebben nagekeken.

Belangrijk punt bij de antwoorden die hier gegeven zijn is dat deze evaluatie door het gros van de studenten is ingevuld voordat zij hun beoordeling ontvingen, waardoor zij het idee hadden dat ze deze vraag nog niet (goed) konden beantwoorden.

#### *Comproved als programma*

Over het functioneren van Comproved als programma voor comparatief beoordelen hebben de studenten geen grote opmerkingen. Eén student gaf aan technische problemen te hebben met het programma. Twee studenten gaven verdere suggesties voor de

functionaliteiten van het programma. Zij zien graag dat er commentaar geplaatst kan worden bij het geüploade werk van medestudenten, en dat er een functie komt om tijdens het vergelijken tussendoor je commentaar op te kunnen slaan.

### ***Studenten - evaluatiegesprek (N = 2)***

Twee studenten hebben deelgenomen aan een evaluatiegesprek over comparatief beoordelen en de rol van Comproved. Hierin kwam grotendeels hetzelfde beeld naar voren als uit de vragenlijst. 'Je forceert studenten [met comparatief beoordelen] om actief na te denken' werd er gezegd. Het belangrijkste aanvullende punt uit dit gesprek was dat het aantal feedbackmomenten volgens deze studenten omhoog zou moeten om daadwerkelijk de feedbackgeletterdheid van studenten te vergroten. Eén feedbackronde leek te weinig om goed door te krijgen hoe je feedback geeft en ontvangt: 'Er werd wel iets gezegd over hoe er door alle studenten werd beoordeeld in de feedbackronde, maar je kon er vervolgens niet meer iets mee doen in de cursus. Je hebt oefening nodig om dit goed te doen.' Het toevoegen van meerdere feedbackrondes zou volgens de studenten ook de kwaliteit van de peerfeedback verhogen. Ondanks het feit dat docenten en studenten anoniem in dezelfde ronde feedback gaven, was voor de studenten makkelijk te zien welke feedback door docenten was gegeven. Hierdoor werd de feedback van medestudenten sneller aan de kant gezet.

De studenten is ook gevraagd wat zij vonden van Comproved als programma. Hier hadden zij geen commentaar of opmerkingen over, ze gaven aan tevreden te zijn met Comproved als tool voor comparatief beoordelen.

### ***Docenten - evaluatiegesprek (N = 7)***

Alle docenten stonden voor de start van de cursus positief tegenover het werken met comparatief beoordelen. Achteraf zijn zij allemaal nog steeds positief, vooral omdat zij zien dat het een meerwaarde heeft voor de studenten.

Het werken met comparatief beoordelen was wel wennen. Vooraf en tijdens de cursus hebben veel gesprekken plaatsgevonden over de manier van werken, waarbij het belangrijk was om alle docenten mee te nemen in de filosofie. De omslag in denkwijze van analytisch naar holistisch toetsen heeft tijd gekost. Twee docenten gaven aan dat het vergelijken voor hen persoonlijk niet fijn werkte: 'Ik zag op een gegeven moment zo veel papers achter elkaar dat ik niet meer wist wat ik aan het vergelijken was.' Daarnaast kwamen dezelfde papers nog wel eens terug, waardoor er wat herhaling was. Hier werd door anderen bij genoemd dat zij vaak een limiet van 1 à 2 uur per dag stelden aan de tijd die zij bezig waren met vergelijken. Een docent gaf aan dat het vergelijken juist makkelijker tussendoor te doen was vergeleken met hele opdrachten met een beoordelingsformulier nakijken.

Om tot een betrouwbare beoordeling te komen bij comparatief beoordelen is het nodig een minimum aantal vergelijkingen te maken. De algemene rekensom hiervoor is het aantal werken dat nagekeken moet worden te vermenigvuldigen met 7. Vervolgens wordt dit totale aantal vergelijkingen verdeeld over het aantal docenten. Hoe meer docenten, hoe lager de nakijklast. In de specifieke set-up van deze cursus en met 7 docenten is de rekensom als volgt:

	Oude situatie	Nieuwe situatie
Feedbackronde (58 opdrachten)	30 min per opdracht = 29 uur /7 = 4,1 uur per docent  1 uur per opdracht = 58 uur /7 = 8,3 uur per docent	5 vergelijkingen per docent Op basis van tijd in Comproved (hier werd feedback gegeven)  4 tot 5 uur per docent
Eindversie (59 opdrachten)	30 min per opdracht = 29,5 uur /7 = 4,2 uur per docent  1 uur per opdracht = 59 uur /7 = 8,4 uur per docent	63 vergelijkingen per docent Op basis van tijd in Comproved (hier werd alleen vergeleken)  +/-5 uur per docent

Het evaluatiegesprek van de docenten vond plaats na de eindbeoordeling van de opdrachten van studenten, waardoor ook de ervaringen rondom de eindbeoordeling meegenomen konden worden. Na de eindbeoordeling was er een aantal studenten dat minder tevreden was geworden over comparatief beoordelen. Bij een deel van hen kwam dit omdat zij in eerste instantie bij de feedbackronde hoog in de rangorde stonden en in de eindronde ver gezakt zijn waardoor zij een onvoldoende haalden. Dit ondanks het feit dat de docenten in de cursus hebben aangegeven dat de rangorde niks zegt over wáár die rangorde zich dan bevindt ten opzichte van cijfers (de rangorde kan van een 1 tot een 10 lopen maar ook van een 5 tot een 6). Het vermoeden bij de docenten bestond dat studenten hoog in de rangorde wellicht het gevoel hadden 'er al te zijn' en minder kritisch hebben gekeken naar waar hun werk verbeterd kon. De opleidingscommissie gaf tijdens het panelgesprek met de cursuscoördinator aan dat studenten het op prijs hadden gesteld om ook na de eindversie nog feedback te ontvangen, omdat zij nu niet wisten wat er precies goed en minder goed was aan hun opdracht. In gesprekken met studenten die een onvoldoende hadden, bleek dat de studenten nog steeds moeite hadden om te zien wat een goede opdracht onderscheidt van een minder goede opdracht. Er heerste bij deze groep een gevoel van hulpeloosheid.

Vooraf is door het docententeam besloten geen feedback te geven op de eindbeoordeling omdat dit een significante verhoging van de nakijklast zou betekenen. Het grote aantal vergelijkingen dat gemaakt moest worden gaf geen ruimte om ook bij iedere vergelijking feedback te geven. In plaats daarvan gingen de studenten met een onvoldoende in gesprek met een docent om tot een verbeterplan te komen. De ervaring rondom deze manier van werken gaf aanleiding tot een discussie over de geschiktheid van comparatief beoordelen in een eerstejaarscursus. Er werd door studenten toch nog veel afhankelijkheid getoond van de feedback van docenten, ondanks de insteek feedbackgeletterdheid te bevorderen. Sommige docenten noemden hierbij dat dit juist vraagt om een grotere focus op de waarde van peerfeedback, waarbij docenten hun focus misschien moeten verplaatsen van feedback op het werk zelf naar feedback op de peerfeedback. Een ander deel van de docenten was het hier niet mee eens en vroeg zich af of het mogelijk is al zo veel verantwoordelijkheid bij de studenten te leggen in het eerste jaar van hun studie. Daarnaast was een vraag of 10 weken genoeg is om al veel resultaat te zien. De vraag werd gesteld: 'Wat betekent dit voor andere cursussen?' Voor het ontwerp voor de cursus volgend jaar wordt er verder nagedacht over de rol van feedback in de eindbeoordeling.

Op technisch vlak heeft het werken met Comproved een behoorlijke tijdsinvestering gevraagd van de docent die alle assessments heeft klaargezet in het programma. Het kost, ondanks de instructie vooraf, toch tijd om goed uit te zoeken hoe het programma achter de



schermen precies werkt en wat studenten wel en niet kunnen zien. Verder werkt het programma wel intuïtief en is het gebruiksvriendelijk. De docenten gaven net als de studenten aan dat het fijn zou zijn om tussendoor te kunnen opslaan. Op één moment heeft het programma niet goed gefunctioneerd, dit was meteen nadat een update was geïnstalleerd waarbij bepaalde bugs ervoor zorgden dat studenten het programma niet konden gebruiken. Echter was het erg eenvoudig om contact op te nemen met de helpdesk en werd het probleem snel verholpen. Bij problemen of vragen werd altijd binnen een dag gereageerd.

Alle docenten gaven unaniem aan dat ze volgend jaar opnieuw met Comproved zouden willen werken, omdat zij zien dat het werkt voor de studenten: 'Wij waren ook nog aan het leren. We kunnen er meer uit halen, het zou zonde zijn om dit niet nog eens te doen.' De beginnende docenten vonden het fijn om op deze manier samen na te kijken, in plaats van alleen de verantwoordelijkheid te dragen. Het was volgens de docenten een fijn idee dat meerdere docenten naar de ingeleverde opdrachten keken, dit voelde eerlijker en dat kon ook worden aangetoond in de betrouwbaarheid die in Comproved wordt weergegeven. Een docent gaf ook aan dat de holistische manier van beoordelen beter aansloot bij zijn eigen werkwijze.

## Conclusie en aanbevelingen

Doelen in het werken met comparatief beoordelen in deze cursus (Comproved) waren het verhogen van de betrouwbaarheid en kwaliteit van de beoordeling en het bevorderen van feedbackgeletterdheid bij studenten, met daarbij het vermijden van een verhoging van de nakijkdruk bij docenten.

Op basis van de evaluaties bij zowel docenten als studenten blijkt dat de betrouwbaarheid van comparatief beoordelen hoog is als genoeg vergelijkingen worden gemaakt. Of daadwerkelijk een verhoging van de betrouwbaarheid is gerealiseerd is lastig te zeggen omdat er geen betrouwbaarheidscijfers bestaan over de eerdere methode van beoordelen. Studenten en docenten ervaren het beoordelen door meerdere docenten als eerlijker. Comproved als programma geeft aan wat de betrouwbaarheid is van de rangorde en laat het mogelijke misfits in de beoordeling zien, waardoor betrouwbaarheid ook daadwerkelijk is te meten. De kwaliteit van de beoordeling is voor de formatieve functie van comparatief beoordelen afhankelijk van de kwaliteit van de feedback die medestudenten en docenten geven. Hierin werd het als een voordeel gezien dat meerdere studenten/docenten feedback geven, dit is een verbetering ten opzichte van de oorspronkelijke manier van werken. Het is wel nodig om de kwaliteit van peerfeedback goed in de gaten te houden, vooral als het belang van peerfeedback in de opdracht groter wordt. In de eindbeoordeling is het als docententeam belangrijk duidelijk te communiceren naar studenten over de interpretatie van de rangorde. Ook is voor docententeams van belang goed na te denken over de manier waarop wordt gewerkt in zowel de feedbackronde als de eindbeoordeling, aangezien hier veel keuzes in te maken zijn afhankelijk van de specifieke situatie. Denk aan zaken als: wie geeft er feedback, hoe geven we feedback, werken we per werkgroep of over de hele cursus heen, geven we wel of geen feedback aan het eind, (hoe) worden cijfers bepaald op basis van de rangorde, etc.

De feedbackgeletterdheid van studenten bevorderen was één van de doelen van deze specifieke cursus waarbij comparatief beoordelen als middel is ingezet. Op basis van de evaluatie lijken hier stappen in gezet, waarbij Comproved een nuttig hulpmiddel bleek. Dit is vooral te zien aan het feit dat studenten veel gebruik hebben gemaakt van Comproved om het werk en de feedback van andere studenten te bekijken. Studenten gaven zelf ook aan dit fijn te vinden en de meerwaarde te zien voor het verbeteren van hun eigen werk. Uit de evaluaties blijkt ook dat één feedbackmoment eigenlijk te weinig is om goed te leren hoe feedback te geven en ontvangen, het leerrendement is hoger bij meerdere feedbackmomenten.

Ten aanzien van het vermijden van een verhoging van de werkdruk is in de eerder weergegeven tabel te zien dat er weinig verschil is tussen beide methoden. Het verschilt soms wel per docent hoe veel tijd het nakijken kost. Vooral in het begin was het voor sommige docenten lastig niet te veel tijd te besteden aan één vergelijking. Hierin moesten zij leren zich te focussen op de vraag 'welk paper is beter' in plaats van het beoordelen van de kwaliteit van ieder afzonderlijk (onderdeel van een) paper. Uiteindelijk konden alle docenten binnen 5 minuten één vergelijking maken. In het daadwerkelijke nakijken kost comparatief beoordelen dus niet per se meer tijd. Het implementeren van deze werkwijze vraagt daarentegen natuurlijk wel een investering.

Technische verbeterpunten voor Comproved:

- Tijdens een vergelijking kunnen opslaan
- Toevoegen van opmerkingen bij teksten (bij docenten: onzichtbaar voor studenten)

- Wordenaantal van teksten weergeven
- Student preview voor docenten
- Opnieuw laten inleveren van werk makkelijker maken
- Docentfeedback evenredig verdelen over alle papers, zodat alle studenten zowel peer- als docentfeedback ontvangen

## Literatuur

- Lesterhuis, M., Verhavert, S., Coertjens, L., Donche, V., & de Maeyer, S. (2017). Comparative judgement as a promising alternative to score competences. In E. Cano & G. Ion (Eds.), *Innovative practices for higher education assessment and measurement*. doi: 10.4018/978-1-5225-0531-0.ch007
- Van Daal, T., Lesterhuis, M., Coertjens, L., Donche, V., & De Maeyer, S. (2019). Validity of comparative judgement to assess academic writing: examining implications of its holistic character and building on a shared consensus. *Assessment in Education: Principles, Policy & Practice*, 26(1), 59-74, <https://doi.org/10.1080/0969594X.2016.1253542>.

## Bijlage 1: antwoorden studenten

### *Werken met Comproved voor feedback:*

Ik vond het goed dat we leerden hoe comperatief beoordelen werkt en wat goede feedback is. Echter kreeg ik ook feedback terug waarmee ik niks kon en dit is jammer. Ook vond ik het minder dat we nu niet wisten welke feedback van een docent komt. Bij feedback van een docent 'let je toch altijd wat beter op'. Diegene heeft er namelijk (denk ik) meer verstand van dan beginnende studenten.

Ik vond het best wel prettig dat ik van meerdere mensen feedback kreeg. Het principe is goed maar niet iedereen deed er serieus aan mee waardoor je geen betrouwbaar resultaat hebt, dat is erg jammer!

Ik vond het fijn om feedback te krijgen maar vond het lastig om zelf feedback te geven.

Goede ontwikkeling

Het is voor mij nog moeilijk om te kunnen kiezen welk paper beter is. Het vergelijken vind ik ook minder nuttig. De feedback die je aan een groepje geeft vond ik wel heel nuttig.

Het is fijn om te weten waar je op dat moment staat.

Je leert veel van feedback geven en krijgen en je leert kritisch te kijken naar je eigen werk door naar andermans werk te kijken.

Het gaf me een goed beeld van andere papers en daarnaast leer je ook hoe je moet lezen op een manier dat je de juiste feedback kan geven.

Het weten hoe je ervoor staat met betrekking tot anderen is erg fijn, ook al moet dit wel gerelativeerd worden. De feedback is erg handig en goed te verwerken in je eindproduct zolang je medestudenten er echt serieus naar hebben gekeken. Het punten van het echt serieus kijken naar andermans werk is wel een dingetje, omdat niet iedereen dit gedaan heeft. Dat heeft er voor gezorgd dat niet iedereen even goede feedback heeft kunnen krijgen.

Ik vond het erg fijn om inspiratie op te kunnen doen bij de mensen boven jou in de ranking, om je eigen paper tot een hoger niveau te tillen. Ook was het prettig om zoveel feedback te ontvangen (ook al was niet iedereen even serieus hiermee). Nuttig om te zien waar je staat en hoe je kunt verbeteren.

Het is fijn om meerdere feedback gevers te horen over het paper en daar tips en tops over krijgen. Zo weet je waar je staat met je opdracht.

Wij stonden helemaal onderaan en ik werd daar ontzettend onzeker door en ben bang het niet te halen( mijn gedachtes). Wel was het fijn als er goede feedback werd gegeven (een aantal mensen gaven feedback die te beknopt en onzinnig waren).

Ik vond het fijn om van veel mensen feedback te krijgen en te kunnen zien hoe anderen (die beter beoordeeld waren) het hadden gedaan.

Ik zou het fijn vinden om wat meer te weten waar je staat. Doordat veel mensen de feedback heel snel hadden afgeraffeld, had ik niet het idee dat de beoordeling de realiteit weerspiegelde. Wel vond ik het prettig dat je op deze manier veel feedback ontving en niet superveel verslagen moest lezen

Het was fijn dat je veel feedback kreeg, soms vond ik de feedback die anderen gaven wel niet duidelijk genoeg waardoor je er soms niks mee kon.

Comparatief beoordelen was iets wat ik nog nooit had ervaren, de feedback was erg behulpzaam, of het ook voor de beoordeling positief zal uitpakken is nog afwachten.

Doordat de feedback in Comproved gegeven werd door studenten, had ik niet heel erg het gevoel dat de feedback 'betrouwbaar' was. Daarom was de rangorde voor mijn gevoel niet heel erg 'waar'. Wel was het fijn om andermans essays te kunnen bekijken, al wist ik dus niet goed of andermans versies wel 'goed' waren en of ik daar dus van kon leren.

Je kon enigzins zien hoe je ervoor staat tov. de andere studenten.

Het is fijn om uitgebreide feedback vanuit meerdere personen te krijgen. Ook hielp het mij om papers van andere studenten te kunnen lezen, om daar de sterke en zwakke punten uit te halen en die mee te nemen voor onze eigen paper.

Het zorgde voor meer feedback, omdat je paper werd gelezen door meerdere mensen ipv 1 studiegenoot. Ik kon er dus meer mee.

Het is fijn om je werk te kunnen vergelijken met dat van anderen, maar het beoordelen welk werk beter was, vond ik erg lastig. De feedback was niet altijd even nuttig, en omdat het van andere studenten komt twijfelde ik of ik er echt wat mee moest doen.

De feedback is niet altijd even nuttig omdat niet iedereen serieus de teksten heeft gelezen. Ook weet ik niet of ik de juiste feedback heb gegeven. Ik vond mijn feedback prettig, maar dat betekent niet dat het klopt volgens de criteria van de docenten. Ik vraag me af of het eerlijk is om met Comproved te werken in deze cursus vanwege de grote verschillen in opleidingsniveau. Eerstejaars studenten worden nu direct beoordeeld met master-studenten die dit als bijvak volgen. Bij het gebruik van een rubric vindt directe vergelijking plaats met de vereisten van de opdracht in plaats van medestudenten, wat me in dit geval passender lijkt.

Ik vond het fijn om tussendoor te weten waar je stond, echter was er een te lange tussenperiode, waardoor we al verder waren en niets aan de feedback hadden.

Helaas werkte Comproved tot drie keer toe niet naar behoren. Van het beoordelen van het verslag van een medestudent leer je zelf wel veel.

Deze cursus voelde niet als de juiste cursus om Comproved mee uit te testen. Vooral omdat het echt zoeken was hoe alles werkte in dat programma, maar ook omdat het later bleek dat de grading helemaal niet zo betrouwbaar was. Dan vraag ik me af waarom het op de eerste plaats al gebruikt wordt. Daarbij was het ook vaag om zowel een rubric te hebben en te horen krijgen dat de langlopende opdracht nagekeken wordt via Comproved. Dat was een beetje tegenstrijdig.

Ik vond het fijn om ook de feedback op andermans werk te kunnen bekijken en daarmee af te leiden hoe ik mijn eigen werk kon verbeteren.

- ik vond de werkwijze zelf wat minder prettig omdat sommige opdrachten op hun eigen manier goed zijn en het dan lastig is om ze te vergelijken en dus een te kiezen die beter zou zijn. - ik vond het wel weer fijn dat je van ongeveer 5 mensen feedback had gekregen ipv maar 1 iemand zodat je kunt kijken wat volgens alle feedbackgevers punten zijn die echt aangepast moeten worden of dat er maar 1 vd 5 iets minder goed vindt en de rest wel goed. Ik vond het fijn dat je op deze manier veel feedback kon krijgen en daardoor kon je je langlopende opdracht gericht verbeteren. Ook kreeg je zo een beetje een idee hoe anderen het schrijven hadden aangepakt. Hier kon je zelf ook weer ideeën of verbeterpunten uithalen.

Het idee van anoniem de langlopende opdrachten vergelijken en beoordelen vind ik erg goed. Er was wel een groot verschil in hoe uitgebreid iemand de feedback gaf en of die feedback gerelateerd was aan de rubric. De feedback sprak elkaar ook regelmatig tegen, dus dan is het lastig om iets met de feedback te doen. Ik had wel graag willen weten welke feedback van de docent was geweest en welke van studenten, aan de feedback van de docent zou ik dan meer gewicht hangen dan aan feedback van een andere student (helemaal omdat dit een eerstejaars vak is en zij nog weinig ervaring hebben met wetenschappelijk schrijven).

het was wel handig

Je leert om kritisch te kijken naar andermans werk.

Ik merkte dat je juist ook veel leerde van andermans feedback.

Comparatief beoordelen geeft je de mogelijkheid een betere vergelijking te maken met je medestudenten. Dit geeft je de kans om te zien wat je zelf beter kan doen en wat de ander beter kan doen

Omdat je het anoniem is en je van iedereen de feedback kan zien. Hierdoor krijg je een goed beeld van wat beter kan.

Ik denk dat comparatief beoordelen best wel een betrouwbaar beeld geeft van de beoordeling. De beoordeling is anoniem, waardoor mensen ook meer durven op te schrijven, wat ze er van vonden. Verder krijg je ook van meerdere mensen feedback. Ik had zelf ook veel aan de feedback.

Je zag meteen wat er verschilde met andere, betere papers en daaraan kon je je dan een beetje optrekken.

De toetsing vond ik erg ingewikkeld, ik begreep ook vaak niet hoe een vraag op de juiste manier te beantwoorden. misschien had dit meer geoefend kunnen worden of concreet geven hoe je een academisch vraag beantwoord.

Van Comproved vond ik het erg vervelend dat je niet tussendoor je antwoorden kon opslaan. Ik heb namelijk een keer mijn tekst weggeklikt waardoor ik de gehele opdracht opnieuw moest uitvoeren en dit kostte best wat tijd omdat ik zo goed als klaar was.

Er waren grote niveaoverschillen tussen de studenten waardoor het vergelijken van papers soms erg lastig was en de feedback die je ontving niet altijd even nuttig was

Ik vond het heel fijn dat we een beetje konden kijken naar hoe anderen iets verwoord hadden en dat we konden kijken naar wat voor feedback er over het algemeen gegeven was. Overigens denk ik niet dat ik er persoonlijk heel veel aan heb gehad om zelf die feedback te geven aan andere studenten.

Het waren de studenten die de feedback mochten geven en deze feedback was vaak niet heel uitgebreid, omdat er toch niet gecontroleerd kon worden welke student welke feedback had gegeven. Onze feedback was vooral: dit ontbreekt, dit kan aangevuld worden, taalgebruik mag iets academischer en zorg voor een duidelijkere opbouw en dat was het. Deze punten hebben we alsnog meegenomen in het verbeteren van de opdracht, maar uiteindelijk was de score toch heel laag op het eind. De feedback had dus weinig zin. Heel misschien dat de feedback van de docent wel had geholpen, maar het was niet duidelijk wie dit was en ook deze feedback was misleidend, want iemand die een hele lange tekst als feedback had gegeven (hoogstwaarschijnlijk de docent) zei dat we goed op weg waren, maar nog niet ver genoeg. Uiteindelijk viel het eindpunt heel erg tegen in vergelijking met de feedback.

Ik heb nog geen feedback ontvangen op mijn tentamen en mijn langlopende opdracht. Alleen tussentijds van medestudenten en niet van docenten. Dit was daarom niet voldoende. Ik kan dus niet zeggen dat ik door comparatief beoordelen een beeld heb gekregen van mijn schrijven en dat mijn schrijven hierdoor verbeterd is. Zoals al eerder vermeld, is het belangrijk dat ook tussentijds in ieder geval ook 1 docent feedback geeft. Op basis van mede-studenten is er te weinig feedback geweest om constructief aan het artikel te werken.

Ik vond het wel interessant om te zien waar mijn paper stond in de ranglijst en de feedback was best heel nuttig.

*Werken met Comproved voor eindversie:*

Zelfde als hierboven. Je weet niet hoe een docent tegen je werk aan kijkt. Zij zijn toch de uiteindelijke beoordelers.

Geldt hetzelfde voor als hierboven

Je wist zo wat er nog verbeterd moest worden, al had je niet aan alle feedback wat.

Ik heb liever dat mijn werk door een expert wordt nagekeken, dan door mijn medeleerlingen. Ik weet namelijk niet in hoeverre iedereen serieus zijn beoordelingen uitvoert. Bij sommige studenten heb ik het idee dat ze zomaar wat aanklikken. Duidelijke feedback en afsluiting.

Het is een eerlijk systeem

De eindbeoordeling heb ik nog niks van gehoord dus kan ik denk ook weinig over zeggen  
NVT

Ik denk dat er een eerlijke en goede ranking uit zal komen. Soms is het gewoon erg lastig om aan bepaalde eisen/elementen een cijfertje te hangen, zoals '0,3 van de 0,7 op bron gebruik'. Dit lijkt me een top manier om te beoordelen, omdat het heel relatief is. En omdat er heel duidelijk aangegeven moet worden waarom de een dan beter is dan de ander, heb ik het idee dat er goed in de papers gedoken wordt en dat dit een serieuze en geschikte manier van beoordelen is. De eindbeoordeling is nog niet binnen

/

Ik denk dat het een hele eerlijke manier van nakijken is.

Ik denk dat het comparatief beoordelen een goede manier is om papers te beoordelen omdat hierdoor een veel eerlijker beeld naar voren komt dan bij het los beoordelen door meerdere docenten.

Ik weet het einderesultaat nog niet. Nog geen beoordeling gehad

Ik vind het een fijn idee dat docenten samen nakijken, en dat mijn werk vergeleken wordt met dat van anderen.

Ik vond het jammer geen comments in Comproved te kunnen plaatsen. Daardoor moet je al je gedachten samenvoegen en blijven onthouden terwijl je verder leest, waardoor ik persoonlijk minder goed feedback heb kunnen geven.

Ik ben bang dat er op deze manier te vluchtig beoordeeld wordt.

Nog geen idee, want we hebben nog geen beoordeling/punt. Het inleveren ging wel gewoon gemakkelijk.

We hebben ons cijfer nog niet, dus we weten nog niet echt hoe het zal gaan. Ik denk dat het wel fijn zal zijn op deze manier, aangezien het werk door meerdere mensen wordt beoordeeld en er zo een objectiever resultaat zal komen.

.

Ik weet niet zo goed wat ik ervan vind. Het lijkt me aan de ene kant een goede manier, omdat iedereen zo op waarde kan worden geschat en er daardoor geen ongelijkheid is tussen de opdrachten. Daarnaast voelt het ook wel een beetje raar dat je werkstuk vergeleken wordt met dat van een ander en dat het niet 'individueel' beoordeeld wordt. Wel erg onduidelijk welke bijlagen erbij moesten en hoe deze toegevoegd moesten worden. het was handig

Doordat jouw eigen werk vergeleken wordt met dat van andere studenten en je daarna kan zien op welke 'plek' je geëindigd bent, weet je of jouw niveau van schrijven een beetje overeenkomst met dat van de andere studenten.

-

Ik denk dat het een goed is dat het vanuit meerdere mensen wordt beoordeeld

Ik vind comparatief beoordelen voor de eindbeoordeling van de langlopende opdracht goed, omdat het anoniem is en er meer mensen naar kijken.



Je zag meteen wat er verschilde met andere, betere papers en daaraan kon je je dan een beetje optrekken. comparative werkte prima een normaal programma die net als alle anderen een tijdje kost om uit te vogelen hoe deze werkt. Op dit moment heb ik de eindbeoordeling nog niet terug, dus ik kan hier nog niks over zeggen.

We hebben nog geen eindbeoordeling gehad, dus ik kan niet zo goed zeggen wat ik daarvan vind.

Zoals ik al heb gezegd vind ik het raar dat er zo veel mensen laag gescoord hebben op de langlopende opdracht, terwijl de tussentijdse feedback niet eens zo slecht was. Dit was de allereerste keer voor veel studenten dat ze ooit een inleiding moesten schrijven en dan wordt dit voor veel zo slecht beloond. Ik denk dat het systeem niet geheel eerlijk is, want als veel versies gelijkwaardig waren, moeten die versies allemaal hetzelfde punt verdienen en moeten die versies niet alsnog gerangschikt worden. Ook wilde ik graag de feedback op de eindversie zien, zodat we voor de herkansing wel een voldoende kunnen halen, maar die feedback is niet zichtbaar.

Ik zie alleen het resultaat en er is geen feedback beschikbaar. Ik kan daarom niet zeggen of dit prettig is of niet.

Vond het interessant om te zien waar mijn paper eindigde, al kan ik me voorstellen dat het niet zo prettig is als je erg laag eindigt.