

## MASTER

### Groepswerk bij TUE methode Zelf- en peerevaluatie versus docent beoordeling

Savelsberg, M.A.H.J.

*Award date:*  
2020

[Link to publication](#)

#### **Disclaimer**

This document contains a student thesis (bachelor's or master's), as authored by a student at Eindhoven University of Technology. Student theses are made available in the TU/e repository upon obtaining the required degree. The grade received is not published on the document as presented in the repository. The required complexity or quality of research of student theses may vary by program, and the required minimum study period may vary in duration.

#### **General rights**

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

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

Groepswerk bij TUE methode;  
Zelf- en peerevaluatie versus docent beoordeling.

Marc A.H.J. Savelsberg, 0187302,

Scheikunde

Onderzoek van Onderwijs (*10 EC*)

Begeleiding: M.C.G. Thurlings

## Verklaring inzake TU/e Gedragscode Wetenschapsbeoefening in het kader van de Masterscriptie

Ik heb kennis genomen van de TU/e Gedragscode Wetenschapsbeoefening<sup>1</sup>.

Hierbij verklaar ik dat mijn Masterscriptie conform de regels van de TU/e Gedragscode Wetenschapsbeoefening tot stand is gekomen.

Datum

16-08-2019

.....

Roepnaam M. Achternaam (student 1)

Marc A.H.J. Savelsberg. ....

.....

Handtekening (student 1).



.....

---

<sup>1</sup> Zie: <http://www.tue.nl/universiteit/over-de-universiteit/integriteit/wetenschappelijke-integriteit/>

Hier is ook de Nederlandse Gedragscode Wetenschapsbeoefening van de VSNU te vinden.  
Meer informatie over wetenschappelijke integriteit is te vinden op de websites van de TU/e en de VSNU.

**Inhoudsopgave**

1. Inleiding.....	5
2. Theoretisch kader en literatuur.....	6
Voordelen en nadelen van groepswerk .....	6
Beoordelen van groepswerk.....	7
3. Onderzoeksvragen .....	8
4. Methode.....	9
Participanten.....	9
Procedure en instrumenten.....	9
Analyse.....	10
5. Resultaten .....	12
Resultaten vragenlijst.....	14
6. Conclusie en discussie.....	16
Literatuur.....	19
Bijlagen .....	21

## Samenvatting

In dit onderzoek is gekeken naar het gebruik van zelf- en peerevaluatie in vergelijking met een beoordeling van de individuele bijdrage door de docent aan de hand van observatie van het groepswerk. De bijdrage aan het groepswerk wordt in dit onderzoek samengevat in de vorm van het taakgericht bezig zijn met de opdrachten.

De leerlingen maakten hierbij gebruik van een zelf- en peerevaluatie formulier wat aan het einde van de week werd ingevuld. De docent hield tijdens de lessen op een observatie formulier de activiteiten van de leerlingen bij. Door deze formulieren met elkaar te vergelijken werd een beeld gekregen van de overeenkomsten of verschillen tussen leerlingen en docent.

De resultaten van dit onderzoek lijken erop te wijzen dat leerlingen hun bijdrage aan het groepsproces (zowel kwantitatief als kwalitatief) hoger beoordelen dan uit de observaties van de docent blijkt. Uit een afgenomen vragenlijst kan worden geconcludeerd dat de meeste leerlingen een wekelijkse zelf- en peer-evaluatie wel veel vinden. In de praktijk zou de frequentie van het invullen van een zelf- en peer-evaluatie aangepast kunnen worden door deze online aan te bieden op het moment dat een leerling het huiswerk (online) gemaakt heeft. Er moet dan verder onderzocht worden hoe de resultaten van deze evaluaties kunnen worden verwerkt in het bestaande taakcijfer of dat het als een soort van formatieve beoordeling wordt meegenomen.

*Trefwoorden:* groepswerk, evaluatie

## 1. Inleiding

Op de onderzoeksschool<sup>2</sup> wordt bij het vak scheikunde al jaren gebruik gemaakt van de methode Theorie Uit Experimenten (TUE). Deze methode maakt gebruik van thematische werkboekjes, die de leerlingen in groepjes doorwerken. De werkboekjes zijn ingedeeld in practica waarin aan de hand van een context en meerdere vragen erover chemische concepten worden geïntroduceerd. Door het maken van samenvattingen en (digitaal) huiswerk komen de leerlingen vervolgens tot een inzicht in de chemische concepten en berekeningen die uiteindelijk in het centraal examen worden getoetst (www.theorieuitexperimenten.nl, Mellink, Kaper, Nool, & Ten Voorde, 1994).

Tijdens de contacturen werken de leerlingen in hun groep aan de opdrachten in het werkboekje. Na elke les worden de gemaakte opdrachten ingeleverd en geregistreerd. Dit registreren gebeurt elke les op basis van ingeleverd groepswerk uit de voorgaande les. Het werk wordt door de docent nagekeken en van commentaar voorzien. Hierbij controleert en registreert de docent ook of het betreffende groepje voldoende werk heeft uitgevoerd (voortgang) en wat de kwaliteit van het gemaakte werk is (inhoud). Één van de eisen die aan het inleveren van het groepswerk gesteld wordt is dat de leerlingen om beurten inleveren.

Omdat het huiswerk scheikunde op de onderzoeksschool al digitaal werd aangeboden en gecontroleerd of het gemaakt was, is dit digitale onderdeel ongeveer twee jaar geleden uitgebreid met een tool om ook het groepswerk (kwaliteit en voortgang) te registreren (taakcijfer). Het taakcijfer bestaat dus voor een deel uit de (automatische) registratie van het huiswerk en voor een deel uit de door de docent beoordeelde registratie van het groepswerk (*docent-evaluatie of DE*).

Om de leerlingen wat meer inzicht te geven in hun groepsproces kan gebruik gemaakt worden van *zelf-evaluatie* en *peer-evaluatie* (ZE en PE). Hiermee kunnen leerlingen zichzelf en hun groepsleden een score geven voor wat betreft het werken in de groepjes. Het inzicht wordt dan gevormd doordat de leerlingen zelf moeten nadenken over hun eigen bijdrage en de bijdrage van de groepsleden aan het werken in groepjes bij scheikunde.

In dit onderzoek worden ZE en PE vergeleken met een docent-beoordeling (DE) met als doel om inzicht te krijgen in de vergelijkbaarheid en uitvoerbaarheid van de methodes. Dit is van belang om te komen tot een methode om de taakcijfers een meer individueel karakter te geven met een minimum aan kosten in begeleidingstijd.

---

<sup>2</sup> Indien gewenst kunnen details over de onderzoeksschool bij de auteur worden nagevraagd.

## 2. Theoretisch kader en literatuur

De methode Theorie uit Experimenten (TUE) maakt gebruik van coöperatief leren, waarbij leerlingen in groepjes van drie of vier werken. Topping (Topping, 2005) geeft aan dat werken in groepjes een duidelijke structuur en begeleiding nodig heeft. Een docent die leerlingen in groepjes wil laten werken moet een gefundeerde keuze maken bij het samenstellen van groepjes wil het succesvol zijn. Bij het coöperatief leren wordt van de leerlingen verwacht dat ze elkaar aanvullen of verbeteren bij het maken van opdrachten.

Bij TUE maken de leerlingen kennis met scheikundige concepten door het uitvoeren van kleine experimenten. Het bestuderen van de experimenten en het verwerken van de resultaten wordt in deze methode door middel van gerichte vragen bewerkstelligd, waarbij de docent vooral als begeleider optreedt (Mellink et al., 1994). Belangrijk gegeven bij de TUE methode is dat er gebruik gemaakt wordt van een opdrachtenboekje, maar vooral dat leerlingen door middel van het maken van samenvattingen hun "eigen" scheikundeboek maken. Dit maken van samenvattingen wordt opgebouwd door in leerjaar 3 aan de leerlingen een samenvattingboekje te verstrekken waarin de kernbegrippen en belangrijkste inhoud van de verschillende paragrafen ingevuld kunnen worden. In leerjaar 4 wordt ook nog een boekje verstrekt, alleen daar wordt slechts beperkt informatie gegeven en is meer ruimte vrijgemaakt voor de eigen invulling van de leerling. In leerjaren 5 en 6 dienen de leerlingen dan zelf voor een schrift voor de samenvattingen te zorgen. Om ervoor te zorgen dat er geen misconcepties in de samenvattingen komen, wordt tijdens elke les het gemaakte werk van de les ervoor met de docent besproken, die dan ook aangeeft wat er thuis hoort in de samenvattingen van de leerlingen. Er wordt ook regelmatig een samenvattende les verzorgd, waarbij de leerlingen de gelegenheid hebben om hun samenvattingen te controleren op (juistheid van de) inhoud.

De leerresultaten van de leerlingen worden vervolgens (klassiek, summatief) getoetst, vaak met een startexperiment en een aantal vragen die in de groepjes gemaakt worden, gevolgd door een individueel deel waar de kennis van de leerlingen wordt getoetst. Tijdens deze toetsmomenten is het de leerlingen toegestaan om hun zelf gemaakte samenvattingen te raadplegen.

### **Voordelen en nadelen van groepswork**

Door het gebruik van groepswork worden leerlingen min of meer verplicht om met elkaar de concepten uit de leerstof te begrijpen in plaats van alleen maar te reproduceren; het zogenoemde diep-leren (Gillies, 2008; Kaldeway, 2006). Uit een onderzoek van Basili en Sanford (1991) is gebleken dat leerlingen die in groepjes hadden gewerkt beduidend minder misconcepties hadden dan leerlingen die niet in groepjes werkten. Indien nodig kijkt de docent naar de antwoorden om eventuele misconcepten aan te geven en te sturen op de juiste begripsvorming van de leerlingen (Nool, Kaper, & Voorde, 1995).

Daarnaast heeft het werken in groepjes ook als voordeel dat de leerlingen ook leren discussieren en luisteren naar elkaar zodat ze ook op sociaal gebied meer bedreven raken (Johnson & Johnson, 2004).

## BEOORDELEN GROEPSWERK TUE

Echter het toepassen van groepswork kan ook tot gevolg hebben dat leerlingen die minder goed zijn uiteindelijk niet genoeg kunnen leren. De docent moet dus uitermate alert zijn op de individuele leeropbrengsten van de leerlingen en moet indien nodig de groepsamenstelling wijzigen (Topping, 2005).

Een belangrijk punt van aandacht bij groepswork is de koppeling tussen de zogeheten meelifters en de sulletjes (Davies, 2009). Een meelifter (*free-rider*) wil met zo weinig mogelijk inspanning het werk voltooid hebben en een sulletje (*sucker*) is een (goede) leerling die in aanwezigheid van een vermoedelijke meelifter ook minder werk gaat doen. Dit kan leiden tot demotivatie, waardoor er een verband gemaakt zou kunnen worden tussen onvrede en motivatie bij het werken in groepjes (Burke, 2011).

### Beoordelen van groepswork

In het merendeel van de gevallen wordt groepswork beoordeeld met één groepspunt (“*klassieke beoordeling*”), waardoor onvrede kan ontstaan met name bij leerlingen die menen een groter aandeel hebben gehad (D. Thompson & McGregor, 2009; Towns, Kreke, & Fields, 2000). Uit deze literatuur blijkt echter ook dat dit soort beoordelingen vooral gericht zijn op de producten en niet zozeer op het proces tot het product.

Een mogelijkheid om deze onvrede te ondervangen is het gebruik van *peer-evaluation* (Fellenz, 2006), waarbij de leerlingen elkaars bijdrage beoordelen, onder andere op kwaliteit en kwantiteit van het werk. Bij zo'n *peer-evaluation* vullen groepsleden voor elk lid in wat de (procentuele) bijdrage is geweest op bijvoorbeeld inbreng van ideeën (kwaliteit) en uitvoeren van werkzaamheden (kwantiteit) (Paswan & Gollakota, 2004). Bij de onderzoekschool is het scheikundeonderwijs volledig gericht op groepswork en daarop zal dus een ZE en PE op afgestemd moeten zijn. Op deze manier kan de docent een beeld krijgen van de individuele bijdragen aan het proces van de verschillende groepsleden en aan de hand daarvan een beoordeling geven.

Naast een *peer-evaluation* wordt ook vaak gebruik gemaakt van een *self-evaluation* waarmee de leerling gevraagd wordt om zijn eigen bijdrage te beoordelen (Castelijns & Andersen, 2013; Ferrell & Barbera, 2015). In de meeste gevallen wordt echter gebruik gemaakt van de combinatie ZE en PE, zodat er maar één formulier ingevuld hoeft te worden (Race, 2001; Ryan, Marshall, Porter, & Jia, 2007; Sung, Chang, Chiou, & Hou, 2005)

Een andere mogelijkheid om deze onvrede bij leerlingen te ondervangen is dat de docent door observeren de activiteiten van de leerlingen bijhoudt, maar dit gaat dan wel ten koste van de begeleidingstijd (Cohen, 1994). In dit onderzoek wordt dit aangeduid met *docent-evaluatie (DE)*. Bij een docent-evaluatie observeert de docent de leerlingen en vult dan in welke activiteiten werden waargenomen.

Alle hierboven beschreven methoden zijn natuurlijk tot op zekere hoogte subjectief omdat leerlingen elkaar niet willen afvallen en een observatie in feite een momentopname is (Galton, 2010). Ook is er in de literatuur al vaker gekeken naar de verschillen in beoordeling door leerlingen zelf en door de docent (Chang, Tseng, & Lou, 2012; Falchikov & Goldfinch, 2000; Ryan, Marshall, Porter, & Jia, 2007). In veel gevallen werd er alleen gekeken naar een



beoordeling van een toets, en in hoeverre de cijfers die de leerlingen gaven overeenkwamen met de beoordeling door een docent. Voor een vergelijking van de ZE-PE versus DE zijn in de literatuur verschillende resultaten gevonden (Falchikov & Goldfinch, 2000; Ryan et al., 2007). De resultaten variëren van een groot verschil tussen leerlingen en docent tot een min of meer significante overeenkomst. In het algemeen kon worden aangetoond dat leerlingen zichzelf (en anderen) hoger inschalen dan de inschaling van de docent (Ryan et al., 2007; Sherrard & Raafat, 1994; Thompson, 2001).

### 3. Onderzoeksvragen

Om dus het taakcijfer een meer individueel karakter te geven worden moet er gebruik gemaakt worden van een beoordeling per leerling. Dit kan door een ZE-PE of door docent-evaluatie. Deze laatste heeft echter het nadeel dat dit ten koste gaat van de begeleidingstijd. Dit geeft dus aanleiding om te onderzoeken op welke manier de individuele bijdragen aan het groepswerk bij scheikunde het beste kan worden beoordeeld.

De hoofdvraag wordt dan als volgt geformuleerd: *“In welke mate zijn beoordelingen middels self-peer-evaluatie bruikbaar voor de beoordeling van de individuele inbreng bij groepswerk?”*.

Om na te kunnen gaan of een ZE-PE voldoende betrouwbaar is, moet natuurlijk gekeken worden of er verschillen zijn in beoordeling door leerlingen zelf en door de docent (DE). Het ligt namelijk voor de hand dat leerlingen hun eigen inbreng en die van hun groepsleden anders inschatten dan een docent (Falchikov & Goldfinch, 2000; Ryan et al., 2007). Vandaar dat de volgende vraag moet worden onderzocht voordat een antwoord op de hoofdvraag kan worden gegeven.

Subvraag 1 *“In hoeverre is er overeenstemming tussen DE en de resultaten uit een ZE-PE?”*.

De leerlingen zijn in de voorgaande jaren alleen op groepsniveau en alleen door de docent beoordeeld op basis van het ingeleverde werk per les. Het is dus ook van belang om te onderzoeken wat de leerlingen van deze nieuwe vorm van beoordeling vinden om te kunnen inschatten of een ZE-PE een instrument kan zijn voor het meten van de inbreng van individuele bijdragen aan het groepswerk.

Subvraag 2 *“Wat vinden leerlingen van beoordelingen middels ZE-PE-DE?”*

## 4. Methode

### Participanten

Het onderzoek is uitgevoerd bij twee lesgroepen scheikunde in het vierde leerjaar op havo en vwo, zie Tabel 1. Er is voor deze groep leerlingen gekozen, omdat die in het voorgaande leerjaar op de klassieke manier zijn beoordeeld en dus een vergelijking kunnen maken met de nieuwe beoordelingsmethode uit dit onderzoek en omdat de auteur in deze klassen de scheikundedocent is.

*Tabel 1, Participanten*

Niveau	Man	Vrouw	Totaal	Gemiddelde leeftijd
4 havo	5	13	18	15,8 ± 0,8
4 vwo	8	6	14	15,2 ± 0,4

### Procedure en instrumenten

In de onderzoeksklassen werd tijdens het werken aan één module van TUE gedurende drie weken door de leerlingen wekelijks een zelf- en peer-evaluatie ingevuld (bijlage 1). Het formulier is gedownload<sup>3</sup> en aangepast voor het gebruik in dit onderzoek. Met dit formulier vullen leerlingen voor vijf onderdelen hun eigen bijdrage en die van de groepsgenoten in. Ze moeten hierbij hun oordeel geven over hun bijdrage aan het groepswerk op de gebieden actief bezig zijn (A), bij het onderwerp blijven (B), meedenken (over de opdrachten) (C), hoeveelheid werk (aantal opdrachten gemaakt) (D) en de kwaliteit van het werk (juiste antwoorden en/of verwerken van opmerkingen van docent) (E). Op het formulieren geven leerlingen zichzelf en hun groepsgenoten een score van 0 tot en met 5. In grote lijnen zijn op dit formulier dan drie afzonderlijke onderdelen terug te vinden; taakgericht (onderdelen A, B en C), de hoeveelheid werk (onderdeel D) en de kwaliteit van het werk (onderdeel E).

Tijdens de onderzoeksperiode werd elke les (3 lessen/week) een aantal keren een formulier voor de DE ingevuld. Er is een formulier gebruikt, waarop van elke leerling werd bijgehouden welke activiteit bij een leerling geobserveerd werd (bijlage 2). Het formulier is een vinklijst waarop per leerling werd bijgehouden waar die leerling mee bezig was (Cohen, 1994). Omdat dit alleen via observatie gebeurde werd gescoord op de volgende onderdelen: praten/uitvoeren (1), lezen/schrijven (2), kijken/luisteren (3), niet bezig (4), wacht op docent (5). Hierbij moet worden opgemerkt dat met name de eerste drie onderdelen geen nauwkeurig onderscheid kunnen maken tussen wel of niet taakgerichtheid van waarneming. Daar waar via de observatie dit onderscheid wel gemaakt kon worden werd als categorie “niet bezig” ingevuld. Door het aantal vinkjes bij een bepaalde activiteit te delen door het totale aantal

<sup>3</sup> Beschikbaar via <https://www.dailyteachingtools.com/cooperative-learning-evaluate.html>.

## BEOORDELEN GROEPSWERK TUE

vinkjes werd zo een percentage gekregen dat aangeeft hoeveel tijd een leerling met die activiteit bezig is geweest.

De twee formulieren zijn niet hetzelfde, omdat de docent door observatie alleen kan bekijken wat een leerling op dat moment doet. De twee formulieren kunnen echter wel met elkaar vergeleken worden op het gebied van bijvoorbeeld het taakgericht of niet-taakgericht bezig zijn van een leerling. Om deze vergelijking tussen DE en ZE-PE duidelijk te maken werden de onderdelen A,B en C van de leerlingformulieren vergeleken met de onderdelen 1,2 en 3 van het docentenformulier.

De vergelijking tussen ZE-PE en DE voor wat betreft de hoeveelheid werk en de kwaliteit van het werk van de leerlingen kon worden gemaakt omdat na elke les het gemaakte werk van de groepjes werd ingeleverd. De docent registreerde daarvan het aantal gemaakte vragen en door het inhoudelijk nakijken van de antwoorden kon ook vastgesteld worden wat de kwaliteit van het gemaakte werk was.

Aan het einde van de onderzoeksperiode werd nog een vragenlijst afgenomen waarop de leerlingen konden aangeven wat ze vonden van het gebruik van de zelf- en peer-evaluaties en hoe ze omgingen met het groepswerk (bijlage 3). In deze vragenlijst waren vijftien schaalvragen (likert 5-puntschaal) en vier open vragen opgenomen.

### **Analyse**

De gegevens uit de ZE-PE-formulieren en de DE-formulieren werden verzameld en gestructureerd tot vergelijkbare waarden om onder andere via beschrijvende statistiek tot mogelijke antwoorden op de onderzoeksvragen te komen. Een heel belangrijk gegeven voor dit onderzoek is dat door het combineren van gegevens zoals in de vorige sectie aangegeven is toch een beeld gevormd kan worden over het groepswerk. De overige gegevens over hoeveelheid werk en kwaliteit van het werk komen niet uit de docentobservatie, maar uit het werk dat elke les door de leerlingen wordt ingeleverd. Dit geldt ook voor de kwaliteit, omdat het ingeleverde werk door de docent wordt nagekeken en van commentaar voorzien.

Als eerste is er natuurlijk de vergelijking tussen het “zelf” en de “peer” uit de evaluatieformulieren. Hier werden de door de leerling opgegeven scores voor zichzelf en voor de andere groepsleden vergeleken. Dit werd gedaan voor alle gescoorde categorieën. Per leerling kan de zelf-evaluatie vergeleken worden met maximaal drie peer-evaluaties omdat de groepsgrootte maximaal 4 leerlingen is.

In bijlage 4 is een deel van de ruwe data van de zelf-peer evaluatieformulieren van een klas weergegeven. Zo geeft leerling 1 zichzelf een 4 voor het actief meewerken (categorie A), een 3 voor het zorgen dat de groep bij het onderwerp blijft (categorie B), een 2 voor het bijdragen van bruikbare informatie (categorie C), een 5 voor de hoeveelheid werk dat is gedaan (categorie D) en 4 punten voor de kwaliteit van het gedane werk (categorie E). Uit de formulieren van de andere groepsleden worden de cijfers van deze vijf categorieën vervolgens in het werkblad verzameld en kan er per leerling een gemiddelde berekend worden hoe deze leerling door zijn groepsleden is beoordeeld.

## BEOORDELEN GROEPSWERK TUE

Op deze manier kan voor de elke klas een gemiddelde voor de zelf-evaluatie en de peer-evaluatie worden berekend. Er kan ook een vergelijking gemaakt worden tussen de leerlingen van havo en vwo op deze gebieden.

Ten tweede is er de vergelijking met de observaties van de docent: taakgericht of niet taakgericht in vergelijking met de categorieën A tot en met C uit de ZE-PE formulieren. Bij het observeren van de leerlingen is op het docentformulier een heel andere indeling gebruikt dan wat de leerlingen op het ZE-PE formulier invullen. Er is dus slechts een zeer algemene vergelijking te maken tussen de resultaten vanuit het docentformulier ten opzichte van de leerlingen formulieren. Op het leerlingformulier kan de leerling zichzelf en anderen scores geven die betrekking hebben op de deelname in het groepswork op basis van de eigen bijdrage en die van anderen. Deze categorieën A tot en met C zijn dan in zekere zin vergelijkbaar met de scores van een leerling bij “praten/uitvoeren”, “lezen/schrijven” of “luisteren/kijken” op het docentenformulier. Een leerling die volgens het ZE-PE formulier dus hoog scoort op categorieën A tot en met C zou dan op het docentenformulier overeenkomstig hoog moeten scoren. Tijdens de observaties werd zoveel mogelijk in de gesprekken van de groepjes meegelisterd om te kunnen beoordelen of de gesprekken over de leerstof gingen. Alleen als tijdens een observeringsmoment over de leerstof werd gesproken werd dit op het observatieformulier als praten genoteerd. Gesprekken die niet over de leerstof gingen tijdens een observatiemoment werden als niet bezig (niet taakgericht) genoteerd. Vanwege de andere manier van meten op het docentenformulier werden de percentages omgerekend naar een waarde tussen 1 en 5 door de percentuele verhouding met 5 te vermenigvuldigen en af te ronden.

En als laatste kan uit de geregistreeerde voortgang (hoeveel opdrachten zijn gemaakt) een vergelijking gemaakt worden met categorieën D en E van de zelf- en peer-evaluatie formulieren. Deze laatste categorieën laten de leerlingen hun interpretatie geven van de hoeveelheid werk dat gedaan is en de kwaliteit ervan. De voortgang werd geregistreeerd op basis van het gemiddelde aantal deelvragen uit een module dat per les gemaakt wordt. Dit gemiddelde is gebaseerd op een planning die jaarlijks gemaakt wordt en uit jarenlange ervaring altijd haalbaar moet zijn. Dit gegeven aantal deelvragen vormt dan het midden van de vijfpuntschaal en de afwijkingen naar boven (meer dan 15 deelvragen) en naar beneden (minder dan 15 deelvragen) vormen dan de basis om vast te stellen waar een groepje op de vijfpuntschaal uitkomt. Omdat elke les de gemaakte opdrachten worden ingeleverd kan de docent ook een score toekennen aan de kwaliteit van de gemaakte opdrachten. Het gaat dan niet persé over het goed of fout zijn van de antwoorden maar meer of de leerlingen blijk geven te hebben nagedacht over hun antwoorden.

De schaalvragen op de vragenlijst worden samengevat in gemiddelde en standaardafwijking. Waar mogelijk wordt ook gekeken naar de betrouwbaarheid van de resultaten. De open vragen van de vragenlijst konden gezien het aantal respondenten (n=29) relatief makkelijk geanalyseerd en samengevat worden.

## 5. Resultaten

Om de subvraag betreffende het overeenkomen van zelf-peer evaluaties en docent evaluatie te kunnen beantwoorden zijn in een tweetal klassen zoals vermeld in de voorgaande sectie een aantal zelf-peer evaluaties afgenomen. De ruwe data is verwerkt tot een gemiddelde per onderzoeksgroep. In bijlage 4 is een voorbeeld gegeven van deze ruwe data. Deze data werd vervolgens verwerkt tot een gemiddelde voor de zelf- en peerevaluatie. Op deze bijlage is een voorbeeld te zien hoe een gemiddelde tot stand komt. Elke leerling vult voor zichzelf en voor de groepsleden in hoe ze op de verschillende categorieën scoren. In tabel 2 en 3 worden de resultaten van de zelf-peer formulieren en de docent-observaties tijdens de lessen weergegeven. Vanwege de verschillen in wijze van self-peer evaluatie en observatie door docent zijn de categorieën A,B en C (deelname aan groepswerk) van de self-peer formulieren samen genomen en gemiddeld.

Als algemene opmerking moet meegenomen worden dat, vanwege de relatief kleine groepen, de resultaten vooral indicatief zijn en dat daarom een uitgebreide statistische toetsing weinig extra informatie zal geven.

*Tabel 2, Zelf, peer en docent observatie (week 36)*

Waarneming	4 havo			4 vwo		
	Taakgericht (SD)	Kwantiteit (SD)	Kwaliteit (SD)	Taakgericht (SD)	Kwantiteit (SD)	Kwaliteit (SD)
Zelf	3,7 (0,7)	3,8 (0,7)	3,8 (0,4)	3,7 (0,6)	3,3 (0,5)	3,8 (0,7)
Peer	3,4 (1,4)	3,6 (0,6)	3,9 (0,5)	3,6 (0,7)	3,5 (0,6)	3,8 (0,6)
Observatie	2,9 (1,0)	2,3 (1,9)	3,0 (1,0)	3,1 (0,7)	3,3 (1,2)	3,5 (0,8)

Vanwege diverse lesuitval zijn er tijdens de onderzoeksperiode in de klassen slechts resultaten van 2 lesweken beschikbaar voor een analyse, maar uit de tabellen 2 en 3 blijkt al dat er weinig overeenstemming is tussen de evaluaties van de leerlingen en de observaties van de docent. Over het algemeen vallen de zelf- en peerscores hoger uit dan de score die door de docent gezien wordt. Significantie voor waargenomen verschillen werd waar mogelijk vastgesteld met een tweezijdige t-toets voor de gemiddelden ( $\alpha=0,05$ ).

In Tabel 2 lijkt een duidelijk verschil te zijn tussen ZE en PE bij de havo leerlingen voor het groepswerk ( $p=0,009$ ). Het verschil in beoordeling van groepswerk tussen ZE-PE en observatie door docent is zeker significant ( $p\ll 0,05$ ). Dit betekent dat leerlingen zichzelf hoger inschatten op taakgerichtheid dan dat er geobserveerd werd.

De verschillen in kwantiteit tussen ZE en PE zijn niet significant ( $p=0,14$ ), echter het verschil tussen ZE-PE en observatie duidelijk wel ( $p\ll 0,05$ ).

Bij de vwo leerlingen zijn de verschillen tussen ZE en PE voor het taakgericht werken niet significant maar het verschil tussen ZE-PE en observatie wel ( $p\ll 0,05$ ) echter het is een

## BEOORDELEN GROEPSWERK TUE

kleiner verschil dan bij de havo leerlingen. De kwantiteit is bij de vwo leerlingen hetzelfde voor de ZE-PE en de observaties ( $p=0,013$ ).

Belangrijk om te vermelden is dat er een vrij groot verschil is in de geobserveerde kwantiteit tussen havo en vwo. Weliswaar hebben beide gemiddelden een relatief grote afwijking, maar lijkt wel significant te zijn ( $p=0,02$ ), echter in de gegevens van de havo zit een uitzonderlijk lage waarde van 0,2. Als deze waarde niet wordt meegenomen, dan stijgt het gemiddelde bij de havo naar 2,9 en wordt het verschil minder significant ( $p=0,05$ ).

Voor de kolom kwaliteit kan voorsnog alleen worden vastgesteld dat die bij observatie (door de docent nagekeken vragen) lager uitvallen dan door de leerlingen is ingeschat. Voor deze gemiddelden waren echter niet voldoende gegevens beschikbaar om hierop echt te toetsen.

*Tabel 3, Zelf, peer en docent observatie (week 38)*

Waarneming	4 havo			4 vwo		
	Taakgericht (SD)	Kwantiteit (SD)	Kwaliteit (SD)	Taakgericht (SD)	Kwantiteit (SD)	Kwaliteit (SD)
Zelf	3,7 (0,7)	3,8 (0,8)	3,8 (0,6)	3,6 (0,8)	3,5 (1,0)	3,7 (0,6)
Peer	3,8 (0,8)	3,9 (0,8)	3,9 (0,7)	3,5 (0,8)	3,5 (1,0)	3,8 (0,8)
Observatie	3,5 (1,1)	2,5 (0,9)	2,8 (0,8)	3,4 (0,7)	3,3 (0,7)	3,4 (0,9)

Tabel 3 laat in eerste instantie resultaten zien die vergelijkbaar zijn met de resultaten van tabel 2. Opvallend is wel dat met name de categorie taakgericht zowel bij havo als bij vwo significant is toegenomen. Het verschil in kwantiteit tussen havo en vwo is feitelijk hetzelfde gebleven waarbij de havo leerlingen in de geobserveerde hoeveelheid werk lager scoren dan dat ze zelf aangeven.

Bij de vwo leerlingen is in het taakgericht werken een lichte toename te zien bij het gebruik van de ZE-PE formulieren en dat de hoeveelheid werk eigenlijk hetzelfde blijft.

Als het gaat om het beoordelen van het taakgericht werken kan uit deze gegevens opgemaakt worden dat er weliswaar verschillen zijn tussen de resultaten van de ZE-PE formulieren en de docentobservaties, maar ook dat deze verschillen kleiner lijken te worden naarmate een ZEPE vaker gebruikt wordt.

De verschillen in de hoeveelheid werk tussen ZE-PE en docentobservatie lijkt bij de havo leerlingen groter te blijven dan bij de vwo leerlingen.

**Resultaten vragenlijst**

Om antwoord op subvraag 2 (wat vinden leerlingen van een zelf-peer beoordeling) te kunnen geven werd een vragenlijst (bijlage 3) afgenomen met daarop 15 schaalvragen en 4 open vragen. Met deze vragenlijst werd beoogd om te achterhalen wat de leerlingen vinden van een zelf- en peer-evaluatie als middel om taakcijfers een meer individueel karakter te geven.

Op basis van deze aanname is ervoor gekozen om zelf een indeling te maken op basis van de vraagstelling.

De items van de vragenlijst zijn ingedeeld in een drietal hoofdgroepen, waarvan de resultaten in tabellen hieronder zijn weergegeven en uitgesplitst in havo en vwo.

De eerste hoofdgroep bevat de vragen waarin naar de meningen over het groepswerk wordt gevraagd (vraag 1, 2, 3, 8 en 12). Er is voor deze groep gekeken naar Cronbach's alpha om te achterhalen of het als één concept ("Werken in groepjes") beschouwd mocht worden. Hierbij werden de resultaten van vraag 8 en 12 omgekeerd gebruikt, omdat bij een hoge score op vraag 8 (de groepsleden werken dus minder hard dan de respondent) naar verwachting in elk geval een lage score bij vraag 3 zou moeten horen. Eenzelfde redenering kan bij vraag 12 gebruikt worden.

Analyse van de resultaten levert echter een lage Cronbach alpha ( $\alpha_c = 0,54$ ) zodat dit algemene concept niet betrouwbaar gemeten kan worden.

*Tabel 4, Meningen over groepswerk*

Vraag		Havo (SD)	Vwo (SD)	Totaal (SD)
1	Het werken in groepjes vind ik prettig	3,87 (1,15)	4,29 (0,70)	4,07 (0,98)
2	Het werken in groepjes helpt me scheikunde beter begrijpen	3,47 (1,26)	3,79 (0,94)	3,62 (1,13)
3	Het groepje waarin ik zit werkt goed samen	3,73 (0,85)	3,36 (0,97)	3,55 (0,93)
8	Mijn groepsleden werken minder hard dan ikzelf	2,33 (1,07)	3,07 (0,88)	2,69 (1,05)
12	Als ik alleen werk, werk ik harder	3,07 (0,85)	3,71 (0,96)	3,38 (0,96)

Uit deze resultaten blijkt dat de meningen over het (uitvoeren van) groepswerk nogal verdeeld zijn. Echter de verschillen tussen havo en vwo zijn allemaal significant ( $p < 0,01$ ). Kijkend naar de vraagstelling zou een hoge score bij vraag 1 eigenlijk een lage score bij vraag 12 moeten geven. Het verschil is te zien, maar slechts marginaal.

Uit de tabel is ook af te leiden dat de leerlingen over het algemeen tevreden zijn over het gegeven dat bij de scheikunde lessen van groepswerk gebruik gemaakt wordt. Een mogelijk opvallend verschil is te zien bij de vergelijking van havo en vwo bij de laatste vraag. Het lijkt dat vwo'ers een lichte voorkeur hebben om alleen te werken.

## BEOORDELEN GROEPSWERK TUE

In Tabel 5 en 6 zijn de resultaten weergegeven voor de items die met het gebruik van de zelfpeer evaluatie te maken hebben. Het gaat hier om de vragen 5, 6, 9, 11, 13, 14 en 15. Ook hier is een analyse van de consistentie voor een bovenliggend construct uitgevoerd, maar ook dat leverde geen resultaat dat als betrouwbaar mag worden aangemerkt.

*Tabel 5, Meningen over zelf-peer evaluatie*

Vraag	Havo (SD)	Vwo (SD)	Totaal (SD)
5 Beoordelen van groepsgenoten vind ik moeilijk	2,80 (1,33)	2,71 (0,88)	2,76 (1,13)
6 Ik vind het fijn dat ik mezelf mag beoordelen	3,60 (1,25)	3,21 (0,94)	3,41 (1,13)
9 Door de zelf-peer evaluatie is mijn groepje beter gaan werken	2,07 (1,18)	1,79 (0,94)	1,93 (1,08)
11 Ik beoordeel liever mijn groepsleden dan mezelf	3,33 (1,19)	3,29 (0,96)	3,31 (1,09)

Uitgaande van het midden van de vijfpuntschaal (3) is te zien dat leerlingen het beoordelen van groepsgenoten niet heel erg moeilijk vinden. De leerlingen vinden echter wel dat de ZEPE vrijwel geen invloed heeft gehad op het groepswerk zelf, echter dit lijkt dan weer tegenstrijdig te zijn met de resultaten uit tabel 2 en 3.

*Tabel 6, Gemak van formulier*

Vraag	Havo (SD)	Vwo (SD)	Totaal (SD)
13 Ik vond het moeilijk om een zelf-peer formulier in te vullen	2,53 (0,88)	2,07 (0,80)	2,31 (0,88)
14 Bij het invullen heb ik gelet op wat de anderen invulden	2,40 (1,02)	2,36 (1,17)	2,38 (1,10)
15 De vragen op het zelf-peer formulier zijn voor mij duidelijk	4,27 (0,68)	4,21 (0,56)	4,24 (0,62)

Bij vraag 14 is opvallend dat de leerlingen wel een beetje hebben gekeken naar wat andere leerlingen (over hen) invulden en daardoor waarschijnlijk zelf ook vergelijkbare waarden op de ZE-PE formulieren hebben ingevuld.

Tabel 7 geeft de respons van de leerlingen op de items 4, 7 en 10. Deze items hebben te maken met de beoordeling door docent ten opzichte van de ZE-PE.

*Tabel 7, Mening over docentbeoordeling*

Vraag	Havo (SD)	Vwo (SD)	Totaal (SD)
4 Ik vind dat het taakcijfer goed weergeeft wat ik doe	3,20 (1,17)	2,93 (0,88)	3,07 (1,05)
7 Ik word liever door een docent beoordeeld dan door een andere leerling	2,80 (1,28)	3,00 (1,07)	2,90 (1,18)
10 Ik vind dat een docent mijn werkhouding niet kan beoordelen	2,47 (1,26)	2,57 (1,12)	2,52 (1,19)



## BEOORDELEN GROEPSWERK TUE

In deze tabel zijn met name de resultaten van vragen 7 en 10 interessant. Gezien de vraagstelling van 7 geeft een lage score aan dat leerlingen liever door elkaar beoordeeld worden en geeft een lage score bij vraag 10 aan dat ze liever door een docent beoordeeld worden. Uit deze tabel volgt dus dat ze in redelijke mate tevreden zijn met de mogelijkheid om elkaar te beoordelen, maar uit de relatief lage score bij vraag 10 volgt ook dat ze vinden dat een docent ze best wel kan beoordelen.

Als laatste worden de antwoorden op de open vragen nog samengevat.

Vraag 16, Wat vond je van het gebruik van de zelf-peer evaluatieformulieren om elkaar te beoordelen?

Deze vraag was op het formulier als verplicht aangeduid dus hier moest door elke leerling iets worden ingevuld. De onderstaande aantallen hebben steeds betrekking op alle respondenten (n=29).

De meeste leerlingen (n=18) hadden geen problemen met het gebruik van ZE-PE of vonden het zelfs nuttig en bruikbaar (n=5). Een kleiner aantal (n=8) vond het niet nodig of helemaal niet nuttig. Verder werden er door 3 leerlingen opmerkingen gemaakt over de eerlijkheid bij het invullen van de ZE-PE formulieren, ze menen dat door vriendschappen of alleen al door het werken in de groepjes de formulieren door alle groepsleden op dezelfde manier worden ingevuld.

Vraag 17, Op het zelf-peer evaluatieformulier beoordeelde je elkaar op een aantal punten.

Zijn er nog andere punten die volgens jou moeten worden meegenomen? Zo ja, welke?

Deze vraag was op het formulier niet verplicht en is dus niet door alle respondenten ingevuld.

Deze vraag werd door vrijwel alle leerlingen met nee beantwoord. Één uitzondering was de tip om ook een vraag toe te voegen over de docent.

Vraag 18, Wat zou je graag anders zien als een beoordeling zoals deze in de toekomst wordt gebruikt?

De meeste leerlingen vonden het prima zoals het nu gebruikt is, maar geven wel aan dat het niet te vaak moet worden gebruikt. Er was ook een voorstel om het zelf-peer formulier voortaan digitaal in te vullen en een vraag om meer ruimte voor opmerkingen van leerlingen (over elkaar).

Vraag 19, Heb je nog andere opmerkingen over het gebruik van de zelf-peer evaluatieformulieren?

Deze vraag werd met een volmondig nee beantwoord.

## 6. Conclusie en discussie

Op subvraag 1: “In hoeverre is er overeenstemming tussen een docent-beoordeling en de resultaten uit een zelf-peer evaluatie?” kan antwoord gegeven worden op basis van de

## BEOORDELEN GROEPSWERK TUE

gegevens uit tabel 2 en 3. Er lijkt een kleine discrepantie te zijn tussen de beoordelingen van de leerlingen en de docent-beoordeling, waarbij de leerlingen zichzelf gemiddeld hoger inschatten dan door de docent wordt waargenomen. Dit is een gegeven dat ook in andere onderzoeken voorkomt (Chang, Tseng, & Lou, 2012; Ryan et al., 2007). Een ander belangrijk resultaat uit dit onderzoek is met name dat het hier gebruikte docent formulier in opzet te veel afwijkt van de formulieren die de leerlingen gebruikten. Een betere optie zou zijn om te onderzoeken of het ZE-PE formulier ook door de docenten gebruikt kan worden door daarop dan aan te geven hoe een docent de deelname van de groepsleden zou inschatten.

Uit de resultaten kan vooralsnog geconcludeerd worden dat er over het algemeen bij de Vwogroep een redelijke overeenstemming is tussen de zelf-peer evaluaties en de docentobservaties, maar dat deze overeenstemming bij Havo leerlingen niet volledig is waargenomen (Tabel 2). In de tweede observatieperiode (Tabel 3) is met name bij de Havo leerlingen een toename waargenomen bij het groepswerk (categorie ABC), het lijkt er dus op dat de leerlingen beter met het groepswerk zijn omgegaan tijdens de onderzoeksperiode, die overigens voor het vak scheikunde toetsvrij was. Mogelijk kan dit verklaard worden doordat de leerlingen door het gebruik van de zelf-peer evaluaties meer op het groepswerk waren gefocust.

Op basis van de resultaten in de onderzoeksperiode (tabel 2 en 3) is er bij de vwo-leerlingen geen merkbaar verschil in het waargenomen groepswerk geweest.

Subvraag 2; “Wat vinden leerlingen van beoordelingen middels zelf-peer evaluatie?” werd beantwoord door het afnemen van een enquête waar met behulp van een aantal likert schaalvragen kon worden vastgesteld dat leerlingen zeker niet negatief zijn over deze vorm van beoordeling, maar ook nog altijd een vorm van docentbeoordeling willen behouden. In de huidige werkwijze bij de scheikundelessen wordt dit in zekere zin toegepast doordat het huidige taakcijfer voor een groot deel is samengesteld uit objectieve data (hoeveel vragen zijn in de les gemaakt en hoe zijn die vragen beantwoord). Een zelf-peer evaluatie zou dan als toevoeging kunnen dienen om de leerlingen zichzelf meer bewust te maken van hun aandeel in het groepswerk. Ook hierover zijn in de literatuur soortgelijke waarnemingen gedaan (Sengupta, 1998; Minkels, 2016; Gatfield, 1999).

Met behulp van dit onderzoek is getracht een antwoord te krijgen op de vraag over de bruikbaarheid van zelf-peer evaluaties in de beoordeling van een individuele inbreng bij groepswerk. Het evaluatie-instrument dat bij dit onderzoek gebruikt is blijkt voor de leerlingen voldoende duidelijk te zijn en is ook vrij snel in te vullen. Uit dit onderzoek kwam echter ook naar voren dat er een verschil is tussen de beoordeling door docent en leerling, waarbij de leerlingen over het algemeen hoger uitkomen dan de docent. Dit verschil is duidelijker aanwezig bij havo-leerlingen dan bij vwo-leerlingen. Als we echter in ogenschouw nemen dat de weging van de beoordeling van het groepswerk slechts een deel is van een totale eindbeoordeling dan hoeft dit geen probleem te zijn.

De resultaten van dit onderzoek zijn gebaseerd op een relatief kleine onderzoeksgroep. Om met name het gevonden verschil tussen zelf-peer beoordeling en docent-beoordeling beter in beeld te krijgen en of het daadwerkelijk een significant verschil is zou een grotere onderzoeksgroep gebruikt kunnen worden. Uiteindelijk kan dit dan leiden

## BEOORDELEN GROEPSWERK TUE

tot een uniforme methode om voor elke leerling een individuele beoordeling te geven voor de bijdrage aan het groepswerk. Als in de toekomst vaker gebruik gemaakt gaat worden van enige vorm van ZE-PE binnen het scheikunde onderwijs op de onderzoeksschool dan verdient het de voorkeur om deze ZE-PE digitaal en individueel af te nemen. Hiermee kan voorkomen worden dat leerlingen te veel naar hun 'vriendjes' kijken en wat die hebben ingevuld (Love, 1981; R. S. Thompson, 2001).

**Literatuur**

- Burke, A. (2011). Group Work : How to Use Groups Effectively. *Journal of Effective Teaching*, 11(2), 87–95.
- Castelijns, J., & Andersen, I. (2013). Beoordelen om te leren: Leerlingen als medebeoordelaars van hun eigen leerproces. In *Beoordelen om te leren*.
- Chang, C. C., Tseng, K. H., & Lou, S. J. (2012). A comparative analysis of the consistency and difference among teacher-assessment, student self-assessment and peer-assessment in a Web-based portfolio assessment environment for high school students. *Computers and Education*, 58(1), 303–320. <https://doi.org/10.1016/j.compedu.2011.08.005>
- Cohen, E. G. (1994). Designing Groupwork: Strategies for the Heterogeneous Classroom. In *Communities* (Vol. 2nd). Retrieved from <http://books.google.com/books?hl=en&lr=&id=lZ4Blydrh0YC&oi=fnd&pg=PR9&dq=DESIGNING+GROUPWORK++Strategies+for+the+Heterogeneous+Classroom+,&ots=HyIVAiZeQH&sig=gNHfdVg11AC8ErlhdGrBe0mSnQo>
- Davies, W. M. (2009). Groupwork as a form of assessment: Common problems and recommended solutions. *Higher Education*, 58(4), 563–584. <https://doi.org/10.1007/s10734-009-9216-y>
- Falchikov, N., & Goldfinch, J. (2000). Student Peer Assessment in Higher Education: A Meta-Analysis Comparing Peer and Teacher Marks. *Review of Educational Research*, 70(3), 287–322. <https://doi.org/10.3102/00346543070003287>
- Fellenz, M. R. (2006). Toward Fairness in Assessing Student Groupwork: A Protocol for Peer Evaluation of Individual Contributions. *Journal of Management Education*, 30(4), 570–591. <https://doi.org/10.1177/1052562906286713>
- Ferrell, B., & Barbera, J. (2015). Analysis of students' self-efficacy, interest, and effort beliefs in general chemistry. *Chem. Educ. Res. Pract. Chem. Educ. Res. Pract.*, 318(16). <https://doi.org/10.1039/c4rp00152d>
- Galton, M. (2010). Assessing group work. In *International Encyclopedia of Education* (pp. 342–347). <https://doi.org/10.1016/B978-0-08-044894-7.00356-0>
- Gatfield, T. (1999). Examining Student Satisfaction with Group Projects and Peer Assessment. *Assessment & Evaluation in Higher Education*, 24(4), 365–377. <https://doi.org/10.1080/0260293990240401>
- Gillies, R. M. (2008). The effects of cooperative learning on junior high school students' behaviours, discourse and learning during a science-based learning activity. *School Psychology International*, 29(3), 328–347. <https://doi.org/10.1177/0143034308093673>
- Johnson, D. W., & Johnson, R. T. (2004). The Power of Cooperative Groups. In *Assessing Students in Groups* (pp. 1–22). Minneapolis.

- Kaldewey, J. (2006). Diepte-en oppervlakteleren. *VELON Tijdschrift Voor Lerarenopleiders*, 27(1), 23–29.
- Love, K. G. (1981). Comparison of peer assessment methods: Reliability, validity, friendship bias, and user reaction. *Journal of Applied Psychology*, 66(4), 451–457. <https://doi.org/10.1037/0021-9010.66.4.451>
- Mellink, E. C., Kaper, W. H., Nool, B. P., & Ten Voorde, H. H. (1994). *Theorie uit experimenten. Een methodenvergelijking naar de effecten van groepswerk ten opzichte van klassikaal scheikunde-onderwijs*. Amsterdam: Stichting Kohnstamm Fonds voor Onderwijsresearch.
- Minkels, T. J. M. (Eindhoven S. of E. (2016). *Beoordelen van groepswerk*. Eindhoven University of Technology.
- Nool, B. P., Kaper, W. H., & Voorde, H. H. (1995). *De expliciete onderwijssituatie*. 13(2).
- Paswan, A. K., & Gollakota, K. (2004). Dimensions of Peer Evaluation, Overall Satisfaction, and Overall Evaluation: An Investigation in a Group Task Environment. *Journal of Education for Business*, 79(4), 225–231. <https://doi.org/10.3200/JOEB.79.4.225-231>
- Race, P. (2001). A briefing on Self, Peer & Group Assessment. *Assessment*, (9), 1–25. Retrieved from [http://www.heacademy.ac.uk/resources.asp?process=full\\_record&section=generic&id=9](http://www.heacademy.ac.uk/resources.asp?process=full_record&section=generic&id=9)
- Ryan, G. J., Marshall, L. L., Porter, K., & Jia, H. (2007). Peer, professor and self-evaluation of class participation. *Active Learning in Higher Education*, 8(1), 49–61. <https://doi.org/10.1177/1469787407074049>
- Sengupta, S. (1998). Peer evaluation : 1 am not the teacher. *ELT Journal*, 52(January), 19–28.
- Sung, Y. T., Chang, K. E., Chiou, S. K., & Hou, H. T. (2005). The design and application of a web-based self- and peer-assessment system. *Computers and Education*, 45(2), 187–202. <https://doi.org/10.1016/j.compedu.2004.07.002>
- Thompson, D., & McGregor, I. (2009). Online self- and peer assessment for groupwork. *Education + Training*, 51(5/6), 434–447. <https://doi.org/10.1108/00400910910987237>
- Thompson, R. S. (2001). *Reliability , Validity , and Bias in Peer Evaluations of Self-Directed Interdependent Work Teams*.
- Topping, K. J. (2005). Trends in peer learning. *Educational Psychology*, 25(6), 631–645. <https://doi.org/10.1080/01443410500345172>
- Towns, M. H., Kreke, K., & Fields, A. (2000). An Action Research Project: Student Perspectives on Small-Group Learning in Chemistry. *Journal of Chemical Education*, 77(1), 111. <https://doi.org/10.1021/ed077p111>

**Bijlagen**Bijlage 1, Formulier<sup>4</sup> gebruikt voor self-peer evaluatie

# Self-Peer Evaluatie Formulier

Naam \_\_\_\_\_ Klas \_\_\_\_\_ Datum \_\_\_\_\_

Noteer de namen van je groepsleden bij de nummers. Geef vervolgens jezelf een cijfer voor elk onderdeel en doe vervolgens hetzelfde voor je groepsleden. Als laatste zet je in de laatste rij te totalen per kolom.

Waarde: **5**=geweldig **4**=boven gemiddeld **3**=gemiddeld **2**=onder gemiddeld **1**=slecht

	<b>Onderdeel</b>	<b>Ikzelf</b>	<b>1</b>	<b>2</b>	<b>3</b>
<b>A</b>	Actieve deelname bij beantwoorden van vragen				
<b>B</b>	Zorgen dat de groep bij het onderwerp blijft				
<b>C</b>	Bijdrage van bruikbare ideeën				
<b>D</b>	Hoeveel werk is gedaan				
<b>E</b>	Kwaliteit van uitgevoerd werk				
	<b>Totaal</b>				

<sup>4</sup> Vertaald van: <http://www.lapresenter.com/CoopEvalPacket.pdf>

# BEOORDELEN GROEPSWERK TUE

## Bijlage 2, Observatieblad docent

### Observatieblad – Hele klas

Docent \_\_\_\_\_

Datum \_\_\_\_\_

Aantal leerlingen \_\_\_\_\_

Tijd \_\_\_\_\_

Klas \_\_\_\_\_

Module \_\_\_\_\_

In de onderstaande tabel score je wat je van elke leerling ziet op een aantal specifieke momenten tijdens een les. Het aantal streepjes is dan een maat voor het actief met de lesstof bezig zijn.

Groep/leerling	Praten/ uitvoeren (1)	Lezen/ schrijven (2)	Kijken/ luisteren (3)	Niet bezig (4)	Wacht op docent (5)	N rij
1						
2						
3						
4						
5						
6						
N kolom						

Totaal N in groepswerk \_\_\_\_\_

Niet in een groepje:

Taakgericht \_\_\_\_\_

Niet taakgericht \_\_\_\_\_

Total N, niet in groepswerk \_\_\_\_\_

Bijlage 3, Vragenlijst self-peer evaluatie

1. Het werken in groepjes vind ik prettig

Helemaal niet

Helemaal wel

1  2  3  4  5

2. Het werken in groepjes helpt me scheikunde beter te begrijpen.

3. Het groepje waarin ik zit werkt goed samen.

4. Ik vind dat het taakcijfer zonder self-peer evaluatie goed weergeeft wat ik gedaan heb.

5. Het beoordelen van groepsgenoten vind ik moeilijk.

6. Ik vind het fijn dat ik mezelf mag beoordelen.

7. Ik word liever door een docent beoordeeld dan door een andere leerling.

8. Mijn groepsleden werken minder hard dan ikzelf.

9. Door het gebruik van de self-peer evaluatie is mijn groepje beter gaan werken dan zonder.

10. Ik vind dat een docent mijn werkhouding niet kan beoordelen.

11. Ik beoordeel liever mijn groepsleden dan mezelf.

12. Als ik alleen werk, werk ik harder.

13. Ik vond het moeilijk om een self-peer formulier in te vullen.

14. Bij het invullen van het self-peer formulier heb ik gelet op wat de anderen invulden.

15. De vragen op het self-peer formulier zijn voor mij duidelijk.

Helemaal niet

Helemaal wel

1  2  3  4  5

16. Wat vond je van het gebruik van de self-peer evaluatieformulieren om elkaar te beoordelen?

17. Op het self-peer evaluatieformulier beoordeelde je elkaar op een aantal punten. Zijn er nog andere punten die volgens jou moeten worden meegenomen? Zo ja, welke?

18. Wat zou je graag anders zien als een beoordeling zoals deze in de toekomst wordt gebruikt?

19. Heb je nog andere opmerkingen over het gebruik van de self-peer evaluatieformulieren?