

1. Inleiding

UniC is een school waarbinnen op grote schaal gebruik gemaakt wordt van computers bij het geven van onderwijs. Inzet van computers is slechts een voorbeeld van wijze waarop het onderwijs gebruik kan maken van de nieuwe media. Leerlingen zijn vaak de eerste gebruikers van deze nieuwe media en doen dat spelenderwijs. Voor onze school is het een vraag en ook een uitdaging om te onderzoeken op welke wijze wij in ons onderwijs gebruik kunnen maken van deze nieuwe media. Dit met als doel ons onderwijs beter te laten aansluiten op de leefwereld van onze leerlingen en op de ontwikkelingen in veel Aziatische landen. In die landen is het gebruik van nieuwe media in toenemende mate een integraal onderdeel van het onderwijsproces. Aansluiten op die ontwikkeling is een noodzakelijke voorwaarde om geen grote achterstand op te lopen.

Binnen UniC bestaan al langere tijd plannen om games in het onderwijs te gebruiken, maar daartoe ontbrak tot dusver de benodigde expertise. Docenten van UniC staan zeer open voor innovaties, maar hebben nog geen ervaring opgedaan met games in het onderwijs. Dit onderzoeksproject biedt een uitgelezen mogelijkheid om samen met de onderzoekers van het Academieteam en Game On de expertise van docenten op dit punt te vergroten. Docenten ervaren hiermee wat de mogelijkheden van games in het onderwijs zijn en hoe ze games efficiënt en effectief in kunnen zetten.

Het gaat in dit project echter niet alleen om de inzet van games op zich,

ook inhoudelijk is dit project van belang voor de instelling. Net als overal elders in het middelbaar onderwijs zijn ook voor UniC natuurkundige misconcepties een herkenbaar lastig didactisch probleem. Daarbij komt dat docenten eerder oog zullen hebben voor de mogelijkheden van games in hun specifieke vakgebied wanneer zij zien dat games zelfs (!) voor een lastig vak als natuurkunde een goed leermiddel vormen.

De koppeling van een school en een lerarenopleiding in dit project draagt er toe bij dat ook toekomstige leraren weten hoe zij games effectief in het onderwijs in kunnen zetten, omdat de uitvoering en de resultaten van het onderzoek een waardevolle bijdrage kunnen leveren aan het curriculum van de lerarenopleiding.

1.1 De meerwaarde van games

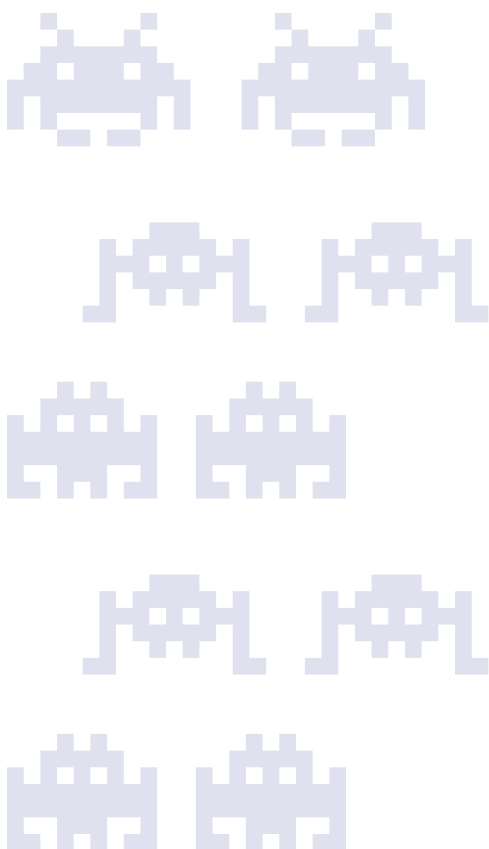
Games vormen een uitstekend leermiddel: ze bieden de mogelijkheid om zaken aan te leren die moeilijk (of zelfs niet) via andere leermiddelen aangeleerd kunnen worden, bijvoorbeeld vanwege veiligheidsrisico's of te hoge kosten. In een game is het bijvoorbeeld niet erg maar wel heel leerzaam wanneer iemand een vinger verliest terwijl hij leert zagen; in het echte leven zou een opleiding dat niet graag zien gebeuren. De meerwaarde van games ligt, met andere woorden, dicht bij de meerwaarde die simulaties bieden: een actie leidt direct tot resultaat. Wat games onderscheidt van simulaties zijn de game elementen die designers inbouwen om ervoor te zorgen dat een gamer

‘op het puntje van zijn stoel zit’: de flow factor. Goede games zijn verslavend, in de positieve zin van het woord. Deze eigenschap van games is wat games ook direct interessant maakt voor het onderwijs: welke docent zou nu niet willen dat zijn leerlingen op het puntje van hun stoel zitten tijdens zijn lessen!

Er zijn in diverse onderzoeken leereffecten van games aangetoond. In het theoretisch kader (1.2) worden daarvan enige voorbeelden gegeven. Op dit moment is Huizenga¹ aan de UvA bezig met een review van het onderzoek naar de effecten van games in het onderwijs. Wat bij haar inventarisatie duidelijk naar voren komt is dat er de laatste jaren weliswaar steeds meer onderzoek gedaan wordt naar de effecten van games, maar dat veel van dat onderzoek methodologisch gezien te wensen overlaat, of dat de rapportage niet helder is¹. Met dit onderzoek willen we dan ook een bijdrage leveren aan de ‘evidence’, die is gebaseerd op kwalitatief goed onderzoek. Bij goed onderzoek in de praktijk hoort een quasi-experimenteel design, met voor- en natests, en een controlegroep.

1.2 Games in het onderwijs

Ondanks het feit dat games aantoonbare leereffecten hebben, worden ze nog maar weinig in het onderwijs ingezet, zeker binnen Nederland. In andere landen, zoals de Angelsaksische en Scandinavische landen, zijn games in het onderwijs veel meer ingeburgerd. Wat docenten ervan weerhoudt om games in te zetten ligt vooral op het gebied van mediawijsheid rond dit onderwerp:



<http://www.surfspace.nl/nl/themas/Games%20en%20Virtuele%20Werelden/Verslagen/Pages/PresentatievanJantinaHuizengaop14april.aspx>

de meeste docenten weten niet welke games ze zouden kunnen gebruiken, hoe ze die in zouden moeten zetten en vooral ook, welke meerwaarde games hebben. Onderzoek waarbij de effecten van games op een deugdelijke wijze wordt aangetoond zou zeker helpen docenten 'over de streep te trekken'. Dat gaat echter niet vanzelf; probleem in de Nederlandse onderwijswereld is ook dat er niet echt sprake is van een platform waarbij docenten kunnen leren hoe ze innovaties effectief in hun onderwijs in kunnen zetten. Er zijn veel beuzen op het gebied van onderwijs en ict, maar daar gaat het vooral om de technische snufjes zelf en niet om de didactische inzet ervan. En dat terwijl toch allang bekend is dat de implementatie van innovaties niet 'technology driven' verloopt. Om die reden is in dit project dan ook nadrukkelijk ruimte ingebouwd voor de disseminatie van de resultaten.

1.3 Games en natuurkunde

In het project worden de effecten van een game onderzocht op een gebied waarbij de veronderstelling dat games hier meerwaarde kunnen hebben zeer plausibel is. Natuurkundige misconcepties zijn voor het VO wat breuken voor het PO zijn: een didactisch struikelblok. Leerlingen kunnen feilloos de juiste natuurkundige principes benoemen terwijl ze tegelijk ideeën over natuurkundige principes in hun hoofd hebben die daar niet mee kloppen. Gebrek aan werkelijk inzicht speelt hier de hoofdrol. In zo'n situatie past een levensechte

game die natuurkundig gezien klopt en waarin de beoogde misconcepties een rol spelen in de problemen die de leerling moet oplossen.

De misconceptie waar het in dit onderzoek om gaat betreft het idee dat objecten, wanneer ze eenmaal in beweging zijn, vanzelf een keer tot stilstand komen, met andere woorden, dat er sprake is van een soort ingebouwde rem. Leerlingen missen het inzicht dat beweging en stilstand met weerstand te maken hebben, ook al leren ze dit wel formeel. Om het beoogde inzicht bij te brengen biedt een game type waarbij het luchtledige een rol speelt, zoals bijvoorbeeld een space game, grote mogelijkheden. Deze game willen wij vergelijken met een serious game waarin dezelfde principes aan de orde komen.

Deze vergelijking maken we in dit onderzoek om twee redenen. Een daarvan is een maatschappelijke. Het ontwikkelen van educatieve games is vaak duur; de meeste scholen zullen daar geen budget voor hebben. Dat is naar onze overtuiging ook helemaal niet nodig, en zelfs ook niet wenselijk. Het inzetten van bestaande, commerciële games is niet alleen veel goedkoper, maar ook veel efficiënter.

Het zou de inzet van games in het onderwijs zeker ten goede komen wanneer scholen voor deze aanzienlijk goedkopere oplossing zouden kiezen. Het onderzoek is daarom ook gericht op fundering van de claim dat het niet alleen goedkoper, maar zelfs ook beter is voor een bepaald

leerdoel simpelweg een aantal exemplaren van een bestaande game aan te schaffen.

De claim dat het ook beter is commerciële games te gebruiken hangt direct samen met onze aanname over de flow factor, die hierboven al even werd aangestipt. Naar onze verwachting vertonen commerciële games deze factor in hogere mate dan educatieve games. Deze verwachting is o.a. gebaseerd op het onderzoek van Verheul (2009), een van de onderzoekers in dit project.

1.4 De flow factor als verklaring voor de meerwaarde van games

Behalve de beoogde leereffecten vond Verheul (2009) in haar onderzoek ook interessante resultaten t.a.v. de rol van motivatie bij het leren via games. De leerlingen die via het spelen van de game aan de leerdoelen werkten waren minder gemotiveerd dan de leerlingen die hetzelfde op een andere manier leerden. Bovendien gaven de leerlingen in de controlegroep aan meer geleerd te hebben dan de leerlingen in de experimentele groep. Toch bleek het omgekeerde het geval. Deze resultaten waren voor de onderzoekster reden om haar eerdere veronderstelling dat de fun factor in games een sleutelrol vervult te herzien. Leerlingen behalen geen betere resultaten omdat ze het leuker vinden games te spelen. De flow factor, datgene in games dat ervoor zorgt dat je wilt doorgaan nadat je eenmaal begonnen bent, lijkt een betere kandidaat om de meerwaarde van games te verklaren.

Serious games missen deze factor doorgaans. Dat heeft vooral te maken met het feit dat designers van serious games meestal gefocused zijn op inhoud en niet op het spelgedrag waarmee de beoogde leerdoelen kunnen worden behaald.

Een andere factor die hierbij een rol zou kunnen spelen is de mate van levensechtheid van een game. Zoals gezegd ligt de meerwaarde van een game dichtbij de meerwaarde van een simulatie: een handeling leidt direct tot een resultaat. Onze aanname is dat hoe levensechter dat resultaat is, hoe meer flow er ervaren wordt, wat vervolgens weer bijdraagt aan de leereffecten. Toch is de mate van flow niet het hele verhaal, als het om leereffecten gaat.

Om via games te kunnen leren moeten leerlingen vrij in de game wereld kunnen rondlopen, zelf ervaringen opdoen, dingen kunnen uitproberen, maar ook pas op de plaats maken om eens goed na te denken wanneer ze niet verder kunnen. Met andere woorden: zowel ervaren: de experiential mode, als nadenken: de reflective mode zijn nodig om de beoogde leerdoelen te kunnen behalen. De flow factor speelt vooral een rol in de experiential mode. Bij het inzetten van een game moet altijd voldoende tijd en gelegenheid voor reflectie in de les worden ingeruimd.

Doelstellingen van het project

De ambities bij dit project liggen op zowel maatschappelijk als op wetenschappelijk vlak. Deze ambities staan niet los van elkaar, maar grijpen op elkaar in.

- Door aan te tonen dat niet alleen generieke competenties als samenwerken en probleemoplossen, maar ook inhoudelijke en didactisch lastige concepten uit de natuurkunde met een game beter kunnen worden aangeleerd dan met een traditioneel leermiddel, kan de bereidheid van docenten om games in hun onderwijs in te zetten positief worden beïnvloed, zowel op de school zelf als op andere scholen.
- De koppeling van een school en een lerarenopleiding in dit project draagt er toe bij dat ook toekomstige leraren weten hoe zij games effectief in het onderwijs in kunnen zetten, omdat de uitvoering en de resultaten van het onderzoek een waardevolle bijdrage kunnen leveren aan het curriculum van de lerarenopleiding.
- In het project wordt nadrukkelijk op de disseminatie van de resultaten ingezet: games kunnen zo'n waardevol leermiddel vormen, dat deze innovatie alle mogelijke inspanningen verdient om te zorgen dat dit leermiddel ook daadwerkelijk wordt benut.
- Door de vergelijking van een educatieve en een commerciële game wordt een bijdrage geleverd aan de fundering van de claim dat de inzet van commerciële games niet alleen op financiële gronden – dat staat immers buiten kijf – maar ook op inhoudelijke gronden de voorkeur verdient boven het ontwikkelen van educatieve games.
- Het onderzoek levert een bijdrage aan het wetenschappelijk debat over de rol die de flow factor speelt bij de effectiviteit van games.

Onderzoeksvragen

A

Verwerven leerlingen die een game spelen het inzicht dat beweging met weerstand te maken heeft beter dan leerlingen die dit inzicht via traditionele natuurkundelessen krijgen aangeboden?

Verwerven leerlingen dit inzicht beter via een voor dit doel geselecteerde commerciële game (COTS) dan via een op dit doel gerichte serious game?

B

C

Is er een verschil in flow bij het spelen van de COTS game en van de serious game?

Zoja:

Is de mate van flow die leerlingen ervaren bij het spelen van beide type games een verklarende factor voor verschillen in effecten?

D

2. Theoretisch kader

2.1 Wat is een game?

Net als bij simulaties gaat het in games om interactiviteit: een actie leidt direct tot een resultaat. Wat een game onderscheidt van een simulatie zijn de elementen die game designers gebruiken om games de flow factor mee te geven. Een goede game bevat de volgende elementen:

- UITDAGEND DOEL
- ONDERLIGGENDE REGELS
- HINDERNISSSEN EN BELONINGEN
- VORM VAN COMPETITIE
- ENIGE MATE VAN ONZEKERHEID
- SPELER CONTROLE
- CONTEXT:
 - AANTREKKELIJKE GRAPHICS
 - VERHAALLIJN

Vooraf competitie en beloning zijn van belang. Er zit veel waars in de stelling van Thomas Hobbes van vierhonderd jaar geleden, dat de mens zich het best voelt wanneer hij zich beter voelt dan een ander. Voor beloningen geldt dat deze beter werken wanneer ze niet altijd op handelingen volgen. Dit berust op een principe uit de klassieke conditioneringstheorie: gedrag dat zo nu en dan beloond wordt, dooft veel minder snel uit dan gedrag dat voortdurend beloond wordt. Enige mate van onzekerheid is daarom een belangrijk element als het gaat om de flow factor.

Commerciële games kunnen worden ingedeeld in de volgende genres:

- BORD- EN KAARTSPELLEN EN PUZZELS
- SIMULATIE EN STRATEGIE.
- ACTIESPELLEN
- AVONTURENSPELLEN
- ROLLENSPELLEN/MMORPG'S

In bord- en kaartspellen en puzzels is de speler zelf niet in beeld. In simulaties en strategiespellen bestuurt de speler een groter geheel: een wereld, een veldslag, een bedrijf. In een rollenspel kiest een speler een karakter (avatar), dat zich door het spel heen ontwikkelt. Hierbij maakt een speler zijn eigen keuzes.

De scheidslijn tussen de genres is niet scherp en vervaagt ook steeds meer: een adventure bestaat uit een verhaallijn die doorlopen wordt door puzzels op te lossen. De meeste adventures hebben tegenwoordig een behoorlijk grote actiecomponent. MMORPG's zijn multiplayer online rollenspellen, met veel actie, een verhaallijn en vaak ook puzzlelementen.

De educatieve game die in dit project gebruikt wordt, Space Challenge, is een simulatiegame. De speler moet een ruimteschip naar bepaalde doelen leiden door de 'straalmotoren' op de juiste manier te bedienen. In de game zijn levels met en zonder weerstand aangebracht.

De commerciële game die voor dit project geselecteerd is, is Unreal Tournament, een actiespel van het type First Person shooter. In dit type game kijkt de speler door de ogen

van zijn karakter (first person) en manoeuvreert zich door een virtuele wereld, waar hij doorgaans van alle kanten belaagd wordt door vijanden. Het level dat voor dit project geselecteerd is, is echter geweldloos.

2.2 Leereffecten van games

Verheul, een van de onderzoekers in dit project, heeft een onderzoek gedaan naar de effecten van een commerciële game (Verheul en van Dijk, 2009). Zij vond effecten op het gebied van samenwerken, oriëntatie en Engels: leerlingen in de experimentele groep bereikten hierop hogere resultaten dan leerlingen in de controlegroep.

Uit een review onderzoek van Randel e.a. ('92) bleek dat in 32% van de 68 studies gaming een groter effect op het verwerven en onthouden van kennis had dan conventioneel onderwijs; het omgekeerde gold voor 5% van de studies. Egenfeldt-Nielsen (2006) deed een vergelijkbaar onderzoek. De uitkomst was overwegend positief wat betreft de effectiviteit van games, maar er waren wel diverse bias. Controlegroepen bleken weinig gebruikt te worden.

De beste resultaten zijn behaald in onderzoek op het gebied van wiskunde en management, zoals blijkt uit het literatuuroverzicht in het proefschrift van Leemkuil (2005). In 7 van de 8 onderzochte studies op het gebied van wiskunde werden er met games betere resultaten gehaald. In alle studies op het gebied van management in een overzicht van Wolfe ('97) bleken games tot een sig-

nificante vooruitgang te leiden.

In een onderzoek van Huizenga e.a. (2007) speelden leerlingen uit het voortgezet onderwijs een stadspel, gericht op de geschiedenis van Amsterdam in de middeleeuwen. Cityteams werden door hun medeleerlingen in het headquarter met mobieltjes op diverse missies de stad rond gestuurd en van informatie voorzien. De leerlingen wisselden halverwege het spel van rol. Verwacht werd dat deze opzet vooral voor een toename in verhalende kennis zou zorgen; dit bleek echter niet het geval. Wel was de feitenkennis van de experimentele groep significant hoger dan van de controlegroep, die dezelfde stof in een regulier pakket aangeboden had gekregen.

2.3 Games als leermiddel bij natuurkundige misconcepties: de rol van de flow factor

In het ervaringsleren is het belangrijk dat een leerling in de gelegenheid is om ervaring op te doen met een verschijnsel: in dit geval beweging zonder wrijving. Deze ervaring kan heel wel opgedaan worden binnen een game, waarin de speler "in het fenomeen wordt ondergedompeld".

Een aantal onderzoeken die zich specifiek richten op begripsverwerving in het domein van de Newtonse mechanica geven aan dat het ontwerpen van een reflectieslag in een dergelijke game essentieel is. De speler is niet in staat om tijdens het spelen (in de flow) voldoende effectief te reflecteren op het eigen handelen en daaruit een begripkader te ontwikkelen.

White (1984) beschrijft een ontwerp van een game dat Newtons wetten aanleert. Ze beschrijft onder andere hoe middels model-progressie het leren in een game kan worden ondersteund. White doet aanbevelingen ten aanzien van het ontwerp van een educatieve natuurkundegame omtrent de wrijvingsloze beweging en geeft aan dat een dergelijk game een goede aanvulling kan zijn binnen het school curriculum.

In een artikel van Hul e.a. (1989) wordt een spel besproken dat in de Nederlandse schoolpraktijk is getest. Ook hieruit komen aanbevelingen, onder andere ten aanzien van het ontwerpen van een reflectiefase, naast het spel.

Rieber (2008) onderzoekt het leerrendement van een simulatiegame over versnelling en kracht en vindt dat een game weliswaar de vaardigheden in het domein kan bevorderen (spelers worden veel handiger in het besturen van de simulatie na het spelen van het spel), maar dat dit ook ten koste kan gaan van het rendement op begripvragen (spelers beantwoorden vragen in het domein niet beter na het spelen van het spel dan daarvoor, soms zelfs minder goed).

De vraag die uit deze onderzoeken rijst is in hoeverre de flow-state bijdraagt aan het conceptuele leereffect. Kan er teveel flow zijn waardoor de speler / leerling de mechanismen niet meer herkent? Dat zou betekenen dat een (eenvoudige, saaie) educatieve game meer leeropbrengst zou genereren dan een commerciële game die volledig op de flow is gericht. Als dit niet bezwaarlijk blijkt en de flow

wel volledig bijdraagt aan de leeropbrengst is een commerciële game juist een heel geschikt alternatief (goed passend in de belevingswereld van de gamende leerling) voor de speciaal ontwikkelde, dure educatieve games.

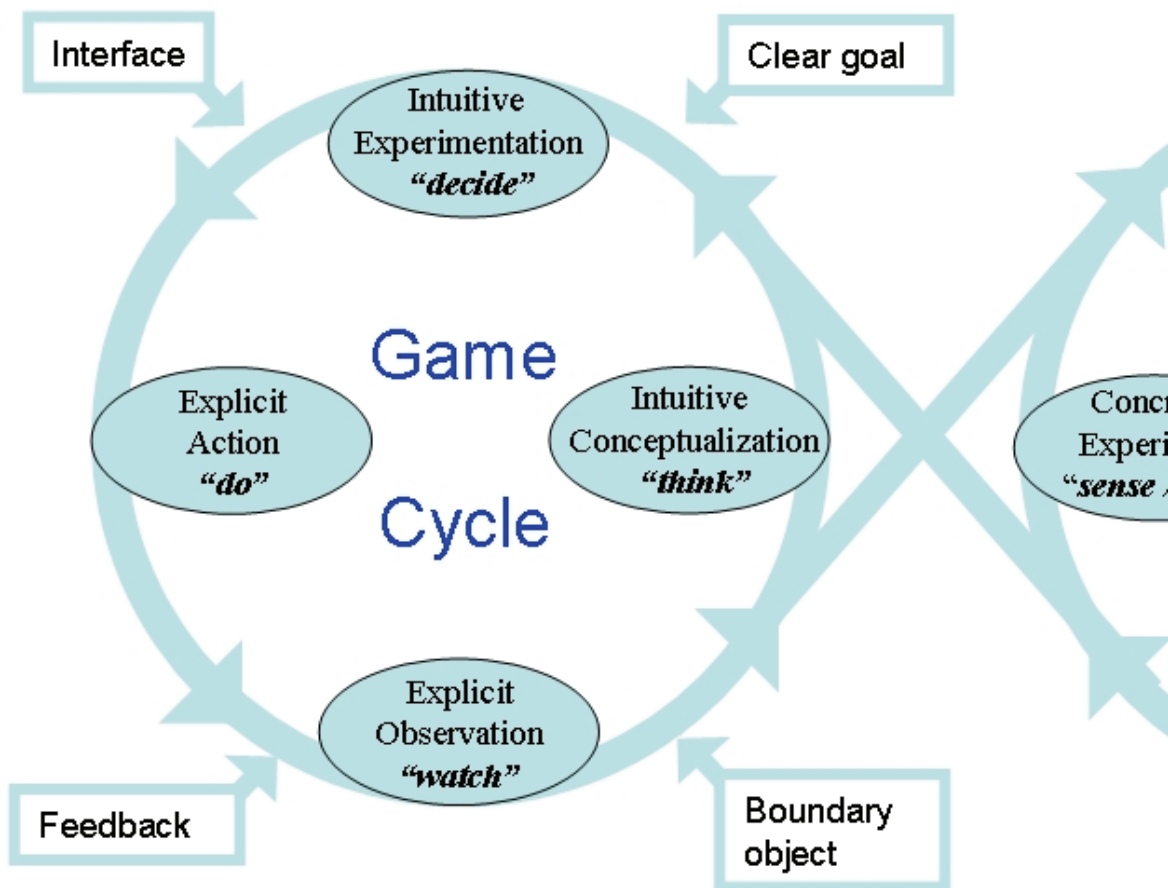
2.4 Een model voor het leren bij games: het *Serious Gaming Lemniscaat model*

Het Serious gaming lemniscaat model (SGLM) is door Martijn Koops, een van de onderzoekers in dit project van het Academieteam, ontwikkeld op basis van het Kennis Management model van Medini & Medini (2005), waarbij de taak-actie cyclus wordt geïntegreerd in de leercyclus. In het model van Koops bestaan de gamecyclus en de leercyclus naast elkaar, in plaats dat ze in elkaar opgaan of elkaar opvolgen. De concrete ervaring, het 'gevoel' staat centraal van het ervaringsleren. De cirkels vertegenwoordigen aan de ene kant de abstracte conceptualisering, het denkproces in de leercyclus en aan de andere kant de actie, het experimenteren, in de gamecyclus.

De leercyclus komt overeen met de cyclus van Kolb (Kolb, 1984), waarbij de leerder Concrete Ervaring (CE), Reflectieve Observatie (RO), Abstracte Conceptualisatie (AC) and Actieve Experimentatie (AE) gebruikt. De game cyclus is een vergelijkbare cyclus die de intuïtieve staat van de gamer weergeeft en de vorming van spontane concepten beschrijft. In de game cyclus is er sprake van een duidelijk doel (Kiili, 2005), dat de speler aanzet tot

Intuïtieve Experimentatie (IE). Om dit mogelijk te maken moet de game een intuïtieve interface hebben, zodat de speler daadwerkelijk handelingen uit kan gaan voeren (Expliciete Actie, EA). De handelingen van de speler leiden tot feedback van de game, het essentiële aspect van simulatie (White, 1984), waar de speler zijn aandacht op kan richten (Expliciete Observatie: EO). Wanneer de resultaten van de handeling duidelijk zijn, kan de speler een Intuïtief Concept (IC) creëren om de feedback van het systeem op de handeling te verklaren.

Op de scheidslijn van de twee cirkels kan de speler besluiten de game cyclus te verlaten en de leercyclus in te gaan, of om door te gaan met de game, zonder te reflecteren op wat hij aan het doen is. Op dit punt kunnen we proberen de speler in de leercyclus te krijgen. Dit punt vormt de essentie van het model. Wil een onderwijs opzet waarbij het gaat om effectieve leerervaringen gebaseerd op een (simulatie) game succes hebben, dan moet die gericht zijn op juist dit punt: het gaat er om de speler op een natuurlijke manier steeds van de

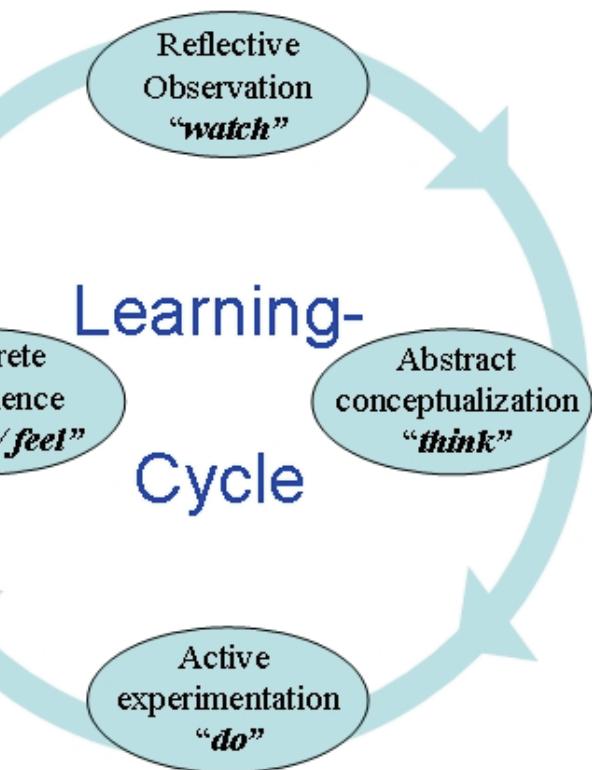


figuur 2.1: Het Serious Gaming Leminiscaat model van Koops

ene cyclus naar de andere te krijgen. In het model worden de verschillende leerfasen in een op simulatie gebaseerde leerervaring weergegeven en de volgorde daarvan in de tijd. Het model geeft aan welke interacties in de game cyclus moeten plaatsvinden, wanneer er naar de leercyclus geswitcht moet worden voor diepere reflectie en de vorming van wetenschappelijke concepten en wanneer de gamer weer terug moet naar de game cyclus om nieuwe ervaringen op te doen om op te kunnen reflecteren. Het is bijvoorbeeld zinloos

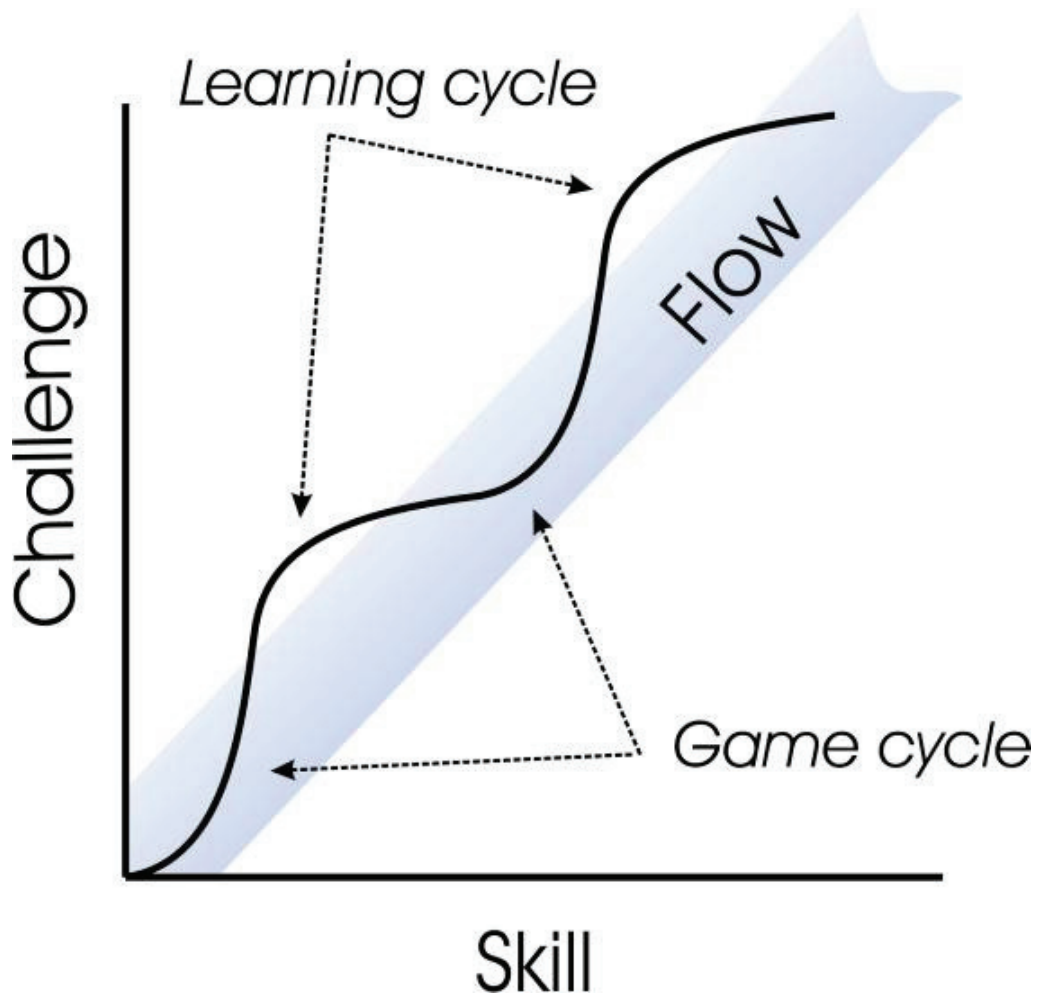
om conceptuele informatie te geven wanneer een speler zich in de game cyclus bevindt (Leutner, 1993).

Frustratie is een machtig instrument als het erom gaat een speler de switch naar de leercyclus te laten maken. Als een game zo is ontworpen dat de uitdaging fors toeneemt, wanneer de speler onoverkomelijke hindernissen tegenkomt of wanneer er nieuwe concepten worden toegevoegd aan het onderliggende regelsysteem, raakt een speler gefrustreerd en valt 'uit de flow', dus uit de game cyclus. De speler gaat op zoek naar hulp om weer terug te kunnen naar de game cyclus. Wij bieden de speler die hulp door de leercyclus te faciliteren (zie figuur 2.2).



Voor een goede verbinding tussen de game cyclus en de leer cyclus moet de representatie in de game gelinkt worden aan de wetenschappelijke context waar het om gaat. Het is bijvoorbeeld goed om het begrip 'kracht' in een game weer te geven met een pijl, omdat dat intuïtief goed te begrijpen is maar ook staat voor het wetenschappelijke begrip 'vector' (White, 1984). Zo'n object wordt in dit model een grensobject genoemd, een object dat in twee verschillende contexten een natuurlijke relevantie heeft, in dit geval zowel in de game als de leercyclus. (zie bijv. Trompette and Vinck, 2009).

In de leercyclus gaat het erom reflectie te faciliteren en te ondersteunen, bijvoorbeeld d.m.v. observatie en monitoring van het leerproces door de docent met behulp van bestaande tools (De Jong en Van Joolingen,



figuur 2.2:

Door een speler 'uit de flow' te laten raken kunnen we 'contact met hem maken' om hem door de leercyclus te begeleiden, zodat hij zich kan voorbereiden op de volgende game ervaring.

2001), of door gestructureerde hulp te bieden. We kunnen de speler helpen de nodige abstracte concepten te vormen door dingen uit te leggen en achtergrondinformatie te geven. Wanneer de speler zover is dat hij een nieuwe hypothese uit kan gaan testen in de game kunnen we hints geven als 'varieer een ding tegelijk' of 'pro-

beer het eens met extreme waarden' (De Jong en Van Joolingen, 2001).

3. Het onderzoek

Opzet van het onderzoek

In het onderzoek wordt gebruik gemaakt van een quasi-experimenteel design, met de volgende opzet:

- **pre- en posttests, waarbij de volgende variabelen worden gemeten:**
 - inzicht en toepassing van het beoogde natuurkundige concept
 - motivatie t.a.v. het vak natuurkunde. Deze variabele kan van invloed zijn op de leerresultaten en wordt daarom als covariaat meegenomen in de analyses
- **In de voortest worden tevens een aantal vragen gesteld over het gamegedrag van de leerlingen. Dit betekent dat de voortest bestaat uit het meten van:**
 - inzicht en toepassing van het beoogde natuurkundige concept
 - motivatie t.a.v. het vak natuurkunde
 - basisgegevens en gamegedrag
- **In de natest worden, naast de vragen over leerprestaties en motivatie, ook enige evaluatieve vragen over de lessen gesteld en wordt de flow factor gemeten. De natest bestaat daarmee uit het meten van:**
 - Inzicht en toepassing van het beoogde natuurkundige concept
 - motivatie t.a.v. het vak natuurkunde
 - evaluatie van de lessen de ervaren flow (voor de beide experimentele groepen)
- **twee experimentele groepen, met game type (commerciële versus serious game) als onafhankelijke factor**
- **een controlegroep: leerlingen die via de reguliere methode les krijgen; deze methode wordt wel zodanig aangepast dat het onderwerp samenvalt met de lessen in de experimentele condities**
- **verzameling van procesgegevens d.m.v.:**
 - observaties van minimaal twee lessen per conditie

4. Methode

4.1 Doelgroep

Op UniC zitten momenteel zo'n 480 leerlingen. In de onderbouw (t/m klas 3) zitten alle leerlingen in een Havo/VWO brugklas. Per bouw (OB-MB-BB) is er een vast docenten-mentorenteam verantwoordelijk voor de groep met leerlingen. Dit zijn tevens de vakexperts die de resultaten van het onderzoek gaan integreren in hun lessen.

Het onderzoek wordt uitgevoerd in de derde klassen. De klassen zijn in drie groepen verdeeld, op basis van een musical waar de leerlingen aan werkten: de acteurs, de 'bouwers' en de overige leerlingen. Deze groepsindeling is aangehouden bij de verdeling van leerlingen over de condities. Op voorhand is gecontroleerd of er verschillen waren tussen deze groepen ten aanzien van hun natuurkunde cijfer. Dit bleek niet het geval.

Het projectteam bestaat uit:

Ruben Koeweiden, de coördinator van het project voor UniC

Cees-Willem de Boer
docent-onderzoeker van UniC

Martijn Koops en **Rinus Tiesma**,
docenten van het Academieteam van het Instituut Archimedes, de lerarenopleiding Natuurkunde van de faculteit Educatie van de Hogeschool Utrecht.

Ineke Verheul,
externonderzoeker van Game On

4.2 Organisatie en projectteam

Het projectteam komt eens in de maand als geheel bij elkaar voor voortgangsoverleg. Daarnaast komen leden van het team zeer regelmatig – ongeveer elke twee weken – bij elkaar voor werkoverleg. Het team maakt gebruik van Surfgroepen voor informatie en communicatie management.

Unic

De docent-onderzoeker is niet alleen betrokken bij de uitvoering van het onderzoek, maar speelt in alle fasen van het onderzoek een actieve rol: bij de opzet van het onderzoek, de constructie van les- en onderzoeksmateriaal, de analyse van de resultaten en de rapportage. Wat de onderzoeksmatige kant betreft is dit project vooral een professionaliseringstraject voor deze docent. De kennis die hij in dit project opdoet over de opzet en uitvoering van onderwijs kan hij vervolgens weer inzetten in de school, bij toekomstige nieuwe projecten. UniC heeft de ambitie ambassadeur te worden voor games in het onderwijs.

In het project houdt de docent-onderzoeker nauw contact met de teamleider Natuurkunde voor de onderbouw, en met de ICT dienst van de school. Bij de constructie van het lesmateriaal zijn niet alleen leden van het projectteam, maar ook stagiaires op UniC betrokken.

De coördinator bewaakt de voortgang van het project, samen met de extern onderzoeker van Game On,

zorgt waar nodig voor de contacten binnen de school en houdt de administratie van het project bij.

Academieteam

De docenten van het Academieteam zijn vooral betrokken bij de inhoudelijke en didactische kant van het project: de games, het lesprogramma en het lesmateriaal. Het Academieteam houdt zich daarnaast ook bezig met de onderzoekskant van het project. Voor het Instituut Archimedes is het belangrijk om te weten wat er met serious gaming mogelijk is, zodat studenten kunnen worden opgeleid in het gebruik ervan. Daarom richt het Academieteam Serious Games zich onder andere op het beantwoorden van vragen omtrent de implementatie van serious gaming in het voortgezet onderwijs. Het team doet onderzoek naar de manier waarop serious games optimaal kunnen worden ingezet in een lessituatie. Dit is van belang voor:

- **het eigen onderwijs:**
 - begeleiding van studenten die in de masteropleiding hun praktijkgericht onderzoek doen op het gebied van serious gaming;
 - opdoen van inzichten over de wijze waarop de nieuwe didactiek van serious gaming kan worden opgenomen in het curriculum;
- **de studenten: welke strategieën hebben studenten nodig om serious gaming optimaal**

in te zetten in hun lessen en hun werkzaamheden op de scholen?

- **de partnerscholen. Het team begeleidt studenten die op partnerscholen schoolprojecten uitvoeren op het serious gaming vlak. Hiervan profiteren ook de partnerscholen.**

Game On

De extern onderzoeker houdt zich vooral bezig met de onderzoeksmatige kant van het project. Daarnaast is zij vooral vanuit de feedback rol betrokken bij de inhoudelijke en didactische aspecten, en bewaakt zij samen met de coördinator van UniC de procesvoortgang. Door haar opleiding als onderzoekspsycholoog en haar ervaring in diverse onderzoeksprojecten beschikt de extern onderzoeker over een gedegen expertise als onderzoekster. Daarnaast is ze een groot deel van haar loopbaan werkzaam geweest als opleidingsadviseur. Met haar bedrijf Game On onderwijs Onderzoek (Game On) geeft zij advies en voorlichting over en doet zij onderzoek naar de inzet van games in het onderwijs.

4.3 De educatieve game

Als educatieve game wordt in dit project Space Challenge gebruikt, een game die door een student van Archimedes, Martijn Hoevenaar (2010) is ontwikkeld en die specifiek gericht is op misconcepten rond de eerste en tweede wet van Newton,

het onderwerp waar het in dit project om gaat.

De game is uitgetest bij leerlingen uit de vierde klas van het VWO. In het onderzoek van Hoevenaar bleken leerlingen die via de game ervaring hadden opgedaan met de betreffende Newtoniaanse concepten meer te hebben geleerd dan leerlingen die deze concepten via een traditionele methode kregen aangeboden. De leerlingen in de experimentele groep zaten in een ‘moeilijke’ en een ‘makkelijke’ conditie. De makkelijke conditie was zo opgezet dat leerlingen in principe de hele game individueel konden spelen zonder ‘uit de flow’ te raken. De verbinding tussen de ervaringen die de leerlingen in de game hadden opgedaan en de abstracte concepten waar het om ging werd in de lessen erna gelegd.

De opzet van de moeilijke conditie was om er d.m.v. een grotere complexiteit van de game voor te zorgen dat leerlingen tijdens het spelen van de game ‘uit de flow’ raakten, zodat ze gedwongen pas op de plaats zouden moeten maken en er reflectie plaats kon vinden, al dan niet met hulp van de docent. Deze conditie bleek voor een nogal chaotische gang van zaken in de klas te zorgen; leerlingen raakten snel gefrustreerd en zeker voor één docent was het vrijwel ondoenlijk hulp te bieden waar nodig. Niettemin leerden ook de leerlingen in deze conditie méér dan de leerlingen in de controle conditie.

De ervaringen die met Space Challenge in het onderzoek van Hoevenaar werden opgedaan waren goed bruikbaar voor dit project. Zo leidden de

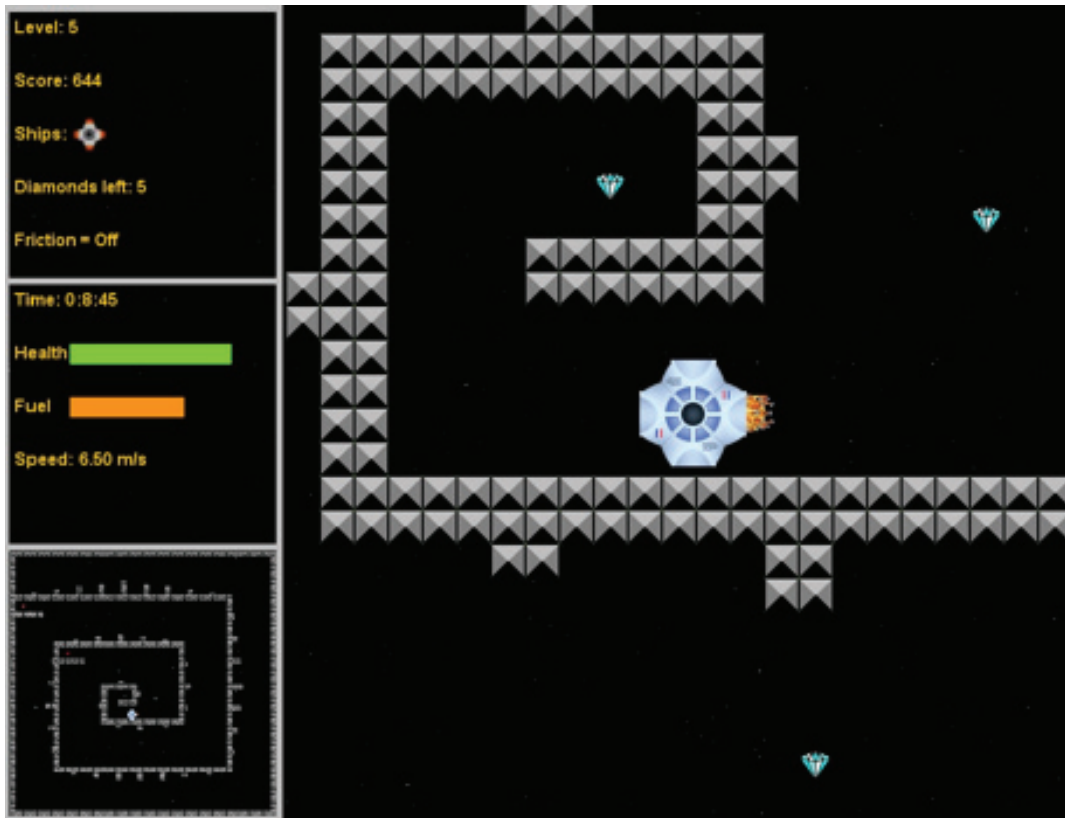
ervaringen met de moeilijke conditie ertoe voor het huidige project in principe uit te gaan van een opzet waarbij leerlingen niet van buitenaf uit de flow zouden worden gebracht. Daarnaast konden de resultaten op de voor- en natest worden meegenomen bij de selectie van een instrument voor het meten van de conceptuele kennis.

Space Challenge

Space Challenge is in essentie een puzzelgame, een ‘maze’. Met een ruimteschip moet een bepaald traject worden afgelegd om een diamant op te pakken. Hierbij kan zowel de wrijving als de zwaartekracht zijn aan- of uitgeschakeld. Om een zo hoog mogelijke score te halen mag niet tegen de muren van de maze worden gebotst en moet een level zo snel mogelijk worden uitgespeeld met zo min mogelijk verbruikte brandstof. Levels nemen toe in complexiteit door:

- het uitschakelen van de wrijving
- een grotere complexiteit van het af te leggen traject
- een toename van het aantal doelen, d.w.z. op te pakken diamanten
- hindernissen als mijnen en meteorieten
- het gebruik van meer dan één dimensie (vector).

In het eerste level is er geen wrijving, en gaat het om een kracht langs een rechte lijn, dus in één dimensie. In dit level speelt de mogelijke misconceptie dat er continu een kracht



figuur 4.1 Screenshot Space Challenge

nodig is om een voorwerp met constante snelheid te laten bewegen. Spelers hanteren hierbij de naïeve strategie om kracht te blijven uitoefenen om zo snel mogelijk bij het voorwerp te komen; hierdoor kan niet op tijd geremd worden en crasht het ruimteschip tegen de muur.

In het tweede level moet hetzelfde traject worden afgelegd, maar dan met wrijving zodat de speler kan ervaren dat de strategie die in de wrijvingsloze situatie niet tot het gewenste resultaat leidde, nu wel succesvol is.

In de volgende twee levels spelen krachten in twee dimensies een rol; hier gaat het om de mogelijke misconceptie dat een voorwerp altijd in de richting van de kracht beweegt die

op het voorwerp wordt uitgeoefend. Ook deze situatie wordt zowel met als zonder wrijving aangeboden.

In de levels erna worden de vaardigheden die geleerd zijn in de eerste levels verfijnd en wordt de complexiteit opgevoerd door een toename van hindernissen en meer ingewikkelde trajecten.

Om de implicaties van de wetten van Newton zichtbaar te maken zijn naast het speelveld een aantal kleine veldjes gecreëerd waarin extra informatie zoals snelheid, score, schade, brandstof en tijd zichtbaar wordt gemaakt.

De besturing van het ruimteschip verloopt simpel: door het indrukken van een van de vier mogelijke cursor-toetsen ontsteekt een stuwkraket in de tegenovergestelde richting. De tijdsduur van het indrukken van de toets is een maat voor de tijdsduur van het uitoefenen van stuwkracht. Die stuwkracht wordt gevisualiseerd d.m.v. 'vlammen' en ook auditief weergegeven met het geluid van een brander.

Aanpassingen van de educatieve game

Ondanks het feit dat Space Challenge in deze vorm reeds succesvol is gebleken waren de onderzoekers van mening dat een aantal aanpassingen het leren via de game zouden kunnen verbeteren. Besloten is tot de volgende aanpassingen:

- feedback over aantal ruimte schepen, tijd, resterende brandstof is grafisch gemaakt.
- doolhof is duidelijk getekend.
- feedback schermen zijn verhelderd
- interface is verduidelijkt: andere-toetsen voor Quit en Start)
- informatie uit het Engels vertaald in NL.
- Het verschil in complexiteit tussen het vierde en vijfde level was aanzienlijk groter dan tussen de voorgaande levels. Niet alleen het aantal hindernissen, maar ook

het aantal op te pikken objecten neemt daarbij aanzienlijk toe. Daarmee wordt het principe van een continue toename van de moeilijkheidsgraad doorbroken.

De aanpassingen zijn uitgevoerd door studenten van instituut Archimedes, onder begeleiding van het Acadamieteam.

4.4 De commerciële game

Unreal Tournament is een First Person Shooter (FPS): de speler baant zich gewapend en wel een weg door een veelal vijandige wereld en ziet die wereld door de 'ogen' van zijn avatar. De school heeft er bezwaar tegen dat er in de game daadwerkelijk geweld wordt gebruikt. Rond games zijn meestal spelers communities actief op fora e.d. en met de creatie van nieuwe levels die andere spelers vervolgens weer kunnen gebruiken. Voor dit project wordt een geweldloos level gebruikt dat door leden van de Unreal Community is ontworpen.

Unreal Tournament

Voor het onderzoek was een commerciële game nodig die inzicht biedt en ervaring oplevert op het gebied van de newtonse mechanica, en dan met name de wrijvingsloze beweging.

Hiervoor hebben we een grote verscheidenheid aan commerciële ruimtevaart games bekeken. Veel van deze games blijken zich weliswaar wel in de ruimte af te spelen maar niet volgens de wetten van de fysica te bewegen. Andere werkelijk natuurlijke games waren weer erg gedetailleerd en si-

muleerden bijvoorbeeld de koppeling van de spaceshuttle aan het ruimtestation. Deze games waren heel lastig te besturen en vergden een grote inwerkperiode van de speler.

Uiteindelijk viel ons oog op Shattered Horizon, een 3D FPS waarbij je rond een ruimtestation zweeft met een raket op je rug. De bewegingen van deze game waren precies wat we zochten. Helaas bleek dit spel erg gewelddadig en bedoeld voor 16+.

Een maal op dit spoor gezet hebben we gekeken naar Unreal Tournament. Dit is een iets oudere FPS game met een level editor waarin we zelf de nodige wijzigingen konden aanbrengen. Via de gebruikersgemeenschap op internet leerden we hoe de game kon worden aangepast zodat er geen wapens in voorkomen. Omdat het spelelement van een shooter zonder schieten ver te zoeken is maakten we ons hierover enige zorgen tot we het Greed level ontdekten. In dit level van Unreal moet de speler goudkisten zoeken en de inhoud in zogenaamde drop-points inleveren. Hiermee is de gameplay weer gegarandeerd. Het spel is te spelen met meerdere spelers in een virtuele wereld en het wedstrijdelement is volop aanwezig.

De fysica van de wrijvingsloze beweging hebben we gerealiseerd door in bestaande (mooi vormgegeven) 3D werelden de wrijving uit te schakelen. Standaard hebben we alle spelers met een jet-pack uitgerust zodat ook vliegen tot mogelijkheden behoorde. Unreal is ook rated 16+ maar door de wapens uit het spel te halen en sterven onmogelijk te maken worden de spelers de moorddadige scenes

bespaard.

Om het spel enige didactische opbouw te geven hebben we 5 verschillende werelden aangepast, te beginnen met een wereld waarin de wrijving wel bestaat. Hierin kunnen de speler wennen aan de spelbesturing. Vervolgens komen de spelers in steeds bergachtiger levels waarin geen wrijving bestaat. Omdat de speler ziet hoe de beweging in deze levels afwijkt van de “gewone” beweging in het eerste level ervaart de speler het effect van het ontbreken van wrijving helemaal.

De game is uitgetprobeerd bij studenten (deeltijd en voltijd) van de lerarenopleiding. Daarbij ging het er vooral om de game zelf uit te proberen en om enige conceptinstrumenten uit te testen. Wat vonden studenten van de game, verliep de game zoals verwacht? De game was niet ingebed in een lesprogramma, zoals in het onderzoek zelf. Effecten van de game op leerprestaties waren daarom niet aan de orde.

Studenten waren enthousiast over de game en de technische uitvoering verliep prima. Wel bleek dat voltijdstudenten, die over het algemeen jonger zijn dan deeltijdstudenten, beter uit de voeten konden met de game dan deeltijdstudenten. De navigatie in een FPS kan desoriënterend werken voor wie dit soort games niet gewend is, zeker wanneer de wrijving geen rol speelt. Een aantal deeltijdstudenten gaf aan enigszins misselijk te worden door het spelen van de game; de voltijdstudenten hadden daar geen last van.



figuur 4.1 Screenshot Unreal Tournament

4.5 Lesprogramma en lesmateriaal

Het onderzoek is in twee weken uitgevoerd.

De opzet van de eerste week was voor alle drie de condities hetzelfde.

Week 1:

- plenaire introductie voor alle drie de groepen van het thema en van het onderzoek: hierbij werd aangegeven dat er tijdens de eerste twee weken van het thema een onderzoek zou worden uitgevoerd en dat leerlingen daarom vragenlijsten zouden worden voorgelegd. Er werd echter niet inhoudelijk op het onderzoek ingegaan.
- Direct daarna volgde de afname van de pretests: de test over natuurkundige concepten (FCI, zie hieronder bij Onderzoeksmateriaal), de algemene vragen vooraf en de test over motivatie voor natuurkunde. Bij deze afname waren observanten aanwezig.

- Gedurende de rest van de eerste week werkten leerlingen aan opdrachten rond de onderwerpen waar het in dit thema om ging: snelheid, bewegen, versnelling, tijddiagrammen. In UniC verwerven leerlingen de nodige kennis en vaardigheden voornamelijk door het zelfstandig maken van opdrachten. De leerlingen zijn in kleine groepen opgedeeld, die met elkaar samenwerken en elkaar helpen. Plenaire behandeling van stof past niet in het systeem van UniC en komt daarom niet of nauwelijks voor.
- Aan het eind van week 1 vonden de gamesessies en het practicum plaats, voor alle drie de groepen in hetzelfde tijdblok van 2 uur. Deze planning op hetzelfde tijdstip had in de eerste plaats onderzoekstechnische redenen, om de plaats in het curriculum voor alle drie de groepen zoveel mogelijk constant te houden en om te voorkomen dat leerlingen die nog geen sessie hadden gehad iets over de sessie van andere groepen zouden horen.

Deze opzet had als consequentie, dat de UniC docent niet alle drie de groepen kon begeleiden. De leerlingen die de educatieve game speelden (Space groep) werden begeleid door de ontwikkelaar van deze game, die eerder een onderzoek met Space Challenge heeft uitgevoerd (Martijn Hoevenaar). De leerlingen die de COTS game speelden, werden begeleid door een lid van het Akademieteam. In alle drie de groepen waren tevens 'ordebewakers' van UniC aanwezig, en observanten.

De begeleiders kregen de instructie geen inhoudelijke informatie te geven of inhoudelijke hulp te bieden, om de docentvariantie tot een minimum te beperken.

De leerlingen in de controle conditie voerden tijdens het practicum een aantal proeven uit, die gericht waren op dezelfde principes die in de games een rol spelen: krachten, richtingen van krachten, snelheid en versnelling. Hierbij werd een carrousel opstelling gebruikt, waarbij groepjes leerlingen met een verschillend startpunt dezelfde reeks opdrachten uitvoerden.

De leerlingen die de educatieve game speelden (Space groep) kregen een korte introductie en een schriftelijke instructie voor het spelen van de game. Daarna konden ze in eigen tempo de verschillende levels van de game uitspelen.

De leerlingen die de COTS game speelden (Unreal groep) kregen eveneens een korte introductie en een schriftelijke spelinstructie. Daarna speelden ze de game eerst individueel (off line) en vervolgens tegen elkaar (on line), toen aangenomen mocht worden dat de leerlingen voldoende oefening hadden gehad in het manoeuvreren binnen de game.

De opdrachten voor het practicum en de instructies voor beide games zijn te vinden in bijlage 1.

Week 2:

- De leerlingen werkten verder aan de opdrachten van het thema, tot elk van de groepen aan de onderwerpen gekomen was waar het in het practicum en de game sessies om ging. Op dit punt kregen de leerlingen verwerkingsopdrachten, die teruggrepen op hetgeen de leerlingen tijdens het practicum en de respectievelijke gamesessies gedaan hadden (zie bijlage 2). Om roostertechnische redenen was het niet mogelijk de games sessies en het practicum zo te plannen dat deze direct voorafgingen aan de terugkoppeling door middel van de verwerkingsopdrachten. De groepen kregen de verwerkingsopdrachten wel op dezelfde dag, maar niet op hetzelfde tijdstip aangeboden. Tijdens de uitvoering van deze opdrachten waren observanten aanwezig.
- Aan het eind van de week werden de pretests afgenomen: de FCI, de motivatietest en voor de experimentele groepen, ook een aantal evaluatievragen over het spelen van de game tijdens de les.

4.6 Onderzoeksmateriaal

4.6.1 FCI

Voor het testen van de leerresultaten wordt de Force Concept Inventory (FCI, Hestenes e.a. 1992)) gebruikt. Met dit instrument worden natuurkundige concepten over kracht en beweging gemeten. De keus van afleiders geeft informatie over de misconcepties van de leerling. Het instrument is eerder gebruikt in het (voor)onderzoek van Hoevenaar (2010).

In het onderzoek van Hoevenaar werd het totale instrument gebruikt. De FCI bestaat uit 30 5-keuze vragen en meet ook andere concepten dan de concepten die de eerste en tweede wet van Newton betreffen. Hoevenaar vond leerwinst voor de leerlingen die de beoogde concepten via games geleerd hadden ten opzichte van leerlingen die deze concepten via traditioneel onderwijs kregen aange-

boden. Deze leerwinst speelde zich echter af in een beperkt gebied van de FCI: de gemiddelde score van de leerlingen met de grootste leerwinst lag rond de 14 vragen. Het onderzoek van Hoevenaar werd uitgevoerd bij leerlingen uit de vierde klas van het VWO. De leerlingen in ons onderzoek zitten in de derde klas en hebben nog maar weinig natuurkunde gehad. Het is daarom te verwachten dat hun gemiddelde score nog lager zal liggen dan bij de leerlingen uit het onderzoek van Hoevenaar.

De FCI is niet ontworpen voor de doelgroep in dit onderzoek. Het instrument is vertaald uit het Engels, maar niet aangepast in 'leerlingentaal'.

Leerlingen een test voorleggen die in hoge mate vragen bevat die leerlingen (nog) niet kunnen beantwoorden kan voor een behoorlijke bias zorgen. Leerlingen kunnen gefrustreerd raken wanneer ze steeds het antwoord

schuldig blijven en daardoor steeds minder gemotiveerd worden de test serieus in te vullen. Bovendien bevat de test, uit het oogpunt van ons onderzoek, teveel ballast: kennis die in het experimentele lesprogramma geen rol speelt.

Om al deze redenen leek het niet verstandig de FCI in zijn geheel voor te leggen aan de leerlingen in het onderzoek. Besloten werd een selectie van bruikbare vragen te maken en deze voor zover nodig ‘leerlingvriendelijk’ te maken. Daarbij was wel geboden de vragen qua formulering zo dicht mogelijk bij de oorspronkelijke vragen te houden. De FCI is in eerder onderzoek grondig uitgetest en een valide en betrouwbaar instrument gebleken.

Selectie van vragen

In eerste instantie is een selectie op zuiver inhoudelijke gronden gemaakt: alleen die vragen die direct betrekking hadden op de concepten waarvan zou mogen worden aangenomen dat die worden aangeleerd in het onderzoek. Dit leverde een selectie van 12 vragen op.

Voor het bepalen van de selectie konden de gegevens van het onderzoek van Hoevenaar gebruikt worden. Daarnaast is de selectie van 12 vragen voor gelegd aan de voltijd- en deeltijdstudenten in de pilot waarbij de commerciële game werd uitgetest.

De betrouwbaarheid van de totale FCI voor de data van Hoevenaar bedroeg .59 voor de pretest en .77 voor de posttest. Dat de alpha voor de posttest hoger ligt dan voor de

pretest kan verklaard worden uit een toename van de variantie: de leerlingen in de experimentele condities leerden meer dan de leerlingen uit de controle condities, zodat de scores meer uit elkaar kwamen te liggen in de posttest. De betrouwbaarheid van de FCI in de pretest is relatief laag, wat een indicatie is voor onze aanname dat de test als geheel eigenlijk te moeilijk is.

De betrouwbaarheden van de selectie van 12 vragen voor de data van Hoevenaar en voor de pilot bij deeltijd- en voltijdstudenten van de lerarenopleiding staan in tabel 4.1.

TABEL 4.1 BETROUWBAARHEID VAN EEN SELECTIE VAN 12 VRAGEN UIT DE FCI

	ALPHA PRETEST	ALPHA POSTTEST
DATA MARTIJN HOEVENAAR	.41	.61
DEELTIJD STUDENTEN PILOT	.49	.54
VOLTIJD STUDENTEN PILOT	.60	.68

Bij de gegevens in tabel 1 moet in aanmerking genomen worden dat de leerlingen in het onderzoek van Hoevenaar de hele FCI kregen aangeboden, wat consequenties kan hebben voor de betrouwbaarheid van een selectie eruit. Dat de betrouwbaarheid van de selectie voor de studenten in pre- en posttest dichter bij elkaar ligt zal te maken hebben met het feit dat

de game niet in hun lesprogramma was ingebed, zoals in het onderzoek van Hoevenaar. De studenten kregen na een introductie van de concepten waar het om ging de pretest van 12 vragen uit de FCI, speelden de game en kregen de posttest. Een eventueel verschil in scores tussen pre- en posttest zou dan aan het spelen van de game op zich moeten worden toegeschreven, iets dat niet in de lijn der verwachtingen ligt. Ons uitgangspunt is dat games in het onderwijs geïntegreerd moeten worden, wil er op doelmatige wijze reflectie kunnen plaatsvinden.

De gevonden betrouwbaarheden zijn niet hoog. Betrouwbaarheid van tests wordt, behalve door de variantie, ook bepaald door de lengte van een vragenlijst. Hoe meer vragen, hoe hoger de te verwachten betrouwbaarheid. Het zou de betrouwbaarheid van de selectie dus ten goede kunnen komen wanneer er meer vragen zouden kunnen worden meegenomen.

Bij een nadere beschouwing van de oorspronkelijke FCI leken er op inhoudelijke gronden toch meer items voor selectie in aanmerking te komen. Aanvankelijk waren bijvoorbeeld vragen waarin de zwaartekracht een rol speelde niet meegenomen in de selectie, omdat zwaartekracht geen rol speelt in de games. Bij een veruiming van de inhoudelijke criteria kwamen 20 van de 30 vragen voor selectie in aanmerking.

Voor elk item in de oorspronkelijke FCI werd het volgende in kaart gebracht:

- het onderwerp;
- (indien van toepassing) de betrouwbaarheid van het item (rir) in een selectie van 20 vragen, gebaseerd op de gegevens van Hoevenaar;
- (indien van toepassing) de betrouwbaarheid van het item (rir) in de selectie van 12 vragen, gebaseerd op de pilot bij voltijd- en deeltijd studenten van de lerarenopleiding;
- de gemiddelde score op de pre- en posttest van het item in het onderzoek van Hoevenaar, zowel voor het totaal als uitgesplitst naar de drie condities van Hoevenaar (makkelijk, moeilijk en controle)
- significante leerwinst (afhankelijke t-test): is de score op het item significant toegenomen op de posttest?

Een overzicht van deze gegevens is te vinden in bijlage 3. Op grond van al deze gegevens werd een uiteindelijke selectie van 19 vragen vastgesteld (zie bijlage 4). De betrouwbaarheid van deze selectie bedraagt .54 voor de pretest en .71 voor de posttest.

4.6.2 Flowinstrument (BIJLAGE 5)

Voor het vaststellen van de hoeveelheid flow die leerlingen ervaren bij het spelen van de game wordt de Flow Scale for Games van Kiili (2006) gebruikt. Dit instrument is in het onderzoek van Kiili

uit het Fins vertaald naar het Engels. Voor ons onderzoek is de lijst vervolgens weer in het Nederlands vertaald. Bij deze vertaling is er zoveel mogelijk rekening mee gehouden dat de vragen in leerlingentaal werden geformuleerd. De vertaalde vragenlijst is vervolgens voorgelegd aan een docente Engels van de lerarenopleiding.

De Flow schaal meet, behalve ervaren flow, ook gepercipieerde 'flow antecedenten': elementen van de game die er voor kunnen zorgen dat de speler flow ervaart. De schaal bevat de volgende subschalen:

Antecedenten:

- A1. UITDAGING
- A2. DOEL
- A3. FEEDBACK
- A4. CONTROLE
- A5. SPEELBAARHEID

Dimensies van flow ervaring:

- D1. CONCENTRATIE
- D2. TIJDVERVORMING
- D3. AUTOTELISCHE ERVARING
- D4. VERLIES VAN ZELFBEWUSTZIJN

Het begrip autotelisch heeft betrekking op de intrinsieke motivatie. De vragenlijst bestaat uit 23 geslo-

ten vragen met een 5-puntschaal, die loopt van 'eens' tot 'oneens'. Daarnaast zijn er nog twee open vragen:

- als je bij het spelen een flow ervaring had, wat zorgde er in de game dan voor dat je die ervaring kreeg?
- als je geen flow ervaring had, wat zorgde er in de game dan voor dat je die ervaring niet kreeg?

Om de betrouwbaarheid van de vertaalde flow schaal te bepalen is het instrument afgenomen in de pilot onder studenten van de lerarenopleiding, na het spelen van de commerciële game. De betrouwbaarheid voor de beide groepen studenten, deeltijd en voltijd, over de schaal als geheel en over de onderdelen 'antecedenten' en 'dimensies' staan in tabel 4.2.

TABEL 4.2 BETROUWBAARHEID FLOW SCHAAL EN SUBSCHALEN VOOR STUDENTEN VAN DE LERARENOPLEIDING

	VOLTijd	DEELTijd
FLOW SCHAAL ALS GEHEEL	.90	.96
ANTECEDENTEN	.82	.77
DIMENSIES VAN FLOW ERVARING	.85	.93

De betrouwbaarheid van zowel de schaal in zijn geheel als van de onderdelen blijkt hoog. De schaal kan dus zo in het onderzoek worden ingezet. Dit instrument wordt alleen in de posttest afgenomen.

Omdat bij de pilot gebleken was dat voltijdstudenten beter uit de voeten konden met de game dan de wat oudere deeltijdstudenten, is nagegaan of de flowervaring voor deze studenten ook anders was. Inderdaad bleken de voltijd studenten meer flow te hebben ervaren dan de deeltijdstudenten ($F=3,89$, $p=.055$). Dit verschil op de totale schaal bleek alleen terug te vinden op het onderdeel ‘dimensies’ ($F=6,52$, $p=.02$). Op het onderdeel ‘antecedenten’ verschillen de beide groepen studenten niet. Voltijd- en deeltijdstudenten schatten de elementen in de game

die voor flow zouden kunnen zorgen dus niet anders in. Voorbeelden van die elementen zijn:

- het gemak van de interface
- de mate uitdaging, dwz de balans tussen de moeilijkheid van de games en de eigen vaardigheden.

4.6.3 Motivatie voor natuurkunde (BIJLAGE 6)

Voor het meten van de motivatie van leerlingen voor het vak natuurkunde is gebruik gemaakt van een vragenlijst bestaande uit 17 gesloten vragen met stellingen op een 5-puntschaal die loopt van ‘Helemaal mee eens’ tot ‘Helemaal niet mee eens’.

Deze vragen zijn verdeeld over drie subschalen:

TABEL 4.3 SUBSCHALEN MOTIVATIE VRAGENLIJST

SUBSCHAAL	AANTAL VRAGEN
ANGST/GEMAK	6
PLEZIER	5
RELEVANTIE/NUT	6

De vragenlijst is samengesteld op basis van drie vragenlijsten gebruikt in onderzoek naar de houding van leerlingen tegen over natuurkunde(Alting, 2003), houding tegen over wiskunde (wiskunde belevingsschaal, Meijer e.a. 2008) en houding tegen over het vak natuur, leven en technologie (de Graaf e.a. 2010). Waarbij de vragenlijst in het

laatst genoemde onderzoek gebaseerd is op de vragenlijst gebruikt in het onderzoek van Alting. De drie vragenlijsten laten naast het gebruik van dezelfde vragen ook verschillen zien in formulering en gebruikte subschalen. Als eerste stap zijn die vragen geselecteerd die in de drie vragenlijsten voorkomen, vervolgens zijn de vragen aan gevuld met vragen

waarbij er kleine verschillen zitten in formulering. Bijvoorbeeld de stelling:

- Wiskunde is van belang om later een baan te krijgen,
- Natuurkunde is van belang om later een baan te krijgen,
- Met NLT in je pakket heb je later een betere kans op een baan.

is geworden:

- Met natuurkunde in je pakket heb je later een betere kans op een baan.

Waarbij die formulering is gekozen die het meest aansloot bij de context en doelgroep.

Er zijn drie subschalen gekozen die in de drie onderzoeken min of meer met elkaar overeenstemmen qua benaming of juist het tegenovergestelde begrip aanduiden. Deze zijn ook vergeleken met de Intrinsic Motivation Inventory (IMI) met subschalen: ervaren competentie, interesse, ervaren keuze, druk en inzet (Vos, 2009). Dit leidde tot de drie bovengenoemde subschalen.

De vragenlijst is uitgetest bij vierdeklassers van UniC (20 leerlingen). De betrouwbaarheid voor de vragenlijst bedroeg .82. Voor de subschalen bedroeg deze resp. .77 voor Angst/Gemak (6 items), .79 voor Plezier (5 items) en wat lager, .45, voor relevantie/Nut (6 items).

4.6.4 Vragen voor- en achteraf

De meer algemene vragen die vooraf en achteraf aan de leerlingen worden gesteld (zie bijlage 7) zijn ontleend aan het onderzoek van Verheul (2009).

Bij leerlingen in alle condities wordt vooraf gevraagd naar de datum, de naam, de leeftijd, de klas en het geslacht. Daarnaast wordt gevraagd of de leerlingen thuis games spelen, hoeveel uur per week ze dat doen en wat voor soort games ze spelen.

De leerlingen in de experimentele condities wordt bij de pretest gevraagd:

- hoe ze het vonden om tijdens de lessen een game te spelen
- of ze door het spelen van de games iets geleerd hebben, en wat dat dan is
- wat ze vinden van games in het onderwijs
- of ze verder nog iets kunnen zeggen over wat ze van het spelen van de game in de les vonden.

4.6.5 Observatie instrument

Voor de observaties tijdens de uitvoering van het practicum en de gamesessies en de afname van de verwerkingsopdrachten zijn formulieren geconstrueerd met de punten waarop de observanten moesten letten. Deze punten waren 'open', dus niet geprecodeerd geformuleerd. De formulieren zijn te vinden in bijlage 8.

4.6.6 Interview vragenlijst

Er zijn interview vragenlijsten opgesteld voor de interviews met:

- LEERLINGEN VOORAF
- LEERLINGEN ACHTERAF
- DOCENTEN VOORAF
- DOCENTEN ACHTERAF

De vragen hadden o.a. betrekking op de mening van leerlingen en docenten over het gebruik van games in de les, de gang van zaken tijdens de uitvoering en de gepercipieerde resultaten. De verschillende vragenlijsten zijn te vinden in bijlage 9.

5. Resultaten

5.1 Procesgegevens

Aan het onderzoek namen in totaal 93 leerlingen deel: 33 in de practicumgroep, en 30 in de beide gamegroepen. Niet alle leerlingen waren aanwezig bij alle momenten die van belang waren voor het onderzoek: de pre- en posttests, het practicum en de gamesessies en de verwerkingsopdrachten, zodat het aantal leerlingen waarover de analyses konden worden uitgevoerd lager ligt.

5.1.1 Observaties

Pre- en posttest

Bij het innemen van de tests is er steeds op gelet dat leerlingen hun naam invulden. De tests konden niet anoniem worden afgenomen, omdat er per leerling meerdere meetmomenten waren die aan elkaar moesten worden gekoppeld.

Helaas is er wat dat betreft toch iets fout gegaan bij de afname van de flowtests in de experimentele condities. Zes leerlingen bleken hun naam doorgekrast te hebben en het was niet meer te reconstrueren welke leerlingen dat waren. Voor de resultaten op de flow tests bleek dit gelukkig weinig verschil te maken.

Practicum

De leerlingen zijn een uur bezig geweest met de practicumopdrachten. Elke opdracht ronde duurde 10 minuten. Daarna hebben de leerlingen nog tien minuten een demonstratie gekregen over iets anders van de TOA.

De beschikbare tijd per ronde bleek soms wat te ruim. Leerlingen in een groep waren bijvoorbeeld al na vijf minuten klaar met de proef met de bak en het flesje. De amanuensis maakte zich duidelijk zorgen over het wat minder geordende karakter van het geheel, dat hierdoor met name aan het eind ontstond. Van hem kwam de suggestie het practicum aan te vullen met activiteiten voor leerlingen die al klaar waren.

De leerlingen waren zo te zien gefocused bezig. Het was niet erg rustig, maar dat hoorde bij de aard van de activiteiten. Leerlingen werkten samen, in groepen.

Er traden geen technische problemen op. Hulpvragen kwamen niet vaak voor. De practicumgroep werd begeleid door de docent die alle derde klassers les geeft. De docent liep rond en hield zelf de gang van zaken in de gaten. De ordebewaakster schoof wat meer aan bij groepjes. Vragen van (individuele) leerlingen gingen vooral over woorden en begrippen en soms stelden zij vragen als 'wat gebeurt er nu eigenlijk?' De docent ging daar niet inhoudelijk op in, maar gaf hooguit hints. Bovendien gaf de docent waar nodig individuele uitleg over de bonuspunten.

Relevantie duidelijk: CW geeft individueel uitleg over het bonuspunt.

De uitwerkingen van de opdrachten werden niet ingenomen. Achteraf gezien was het beter geweest dat wel te doen, gezien het feit dat er toch een soort van beoordeling van inzet zou zijn.

Twee leerlingen van deze groep kwamen bij de Space groep binnen

en hoorden toen dus dat die leerlingen aan het gamen waren. Ze waren wat minder te spreken over wat ze zelf moesten doen.

Space sessie

Alle derdeklassers startten in de ruimte waar de Space groep haar sessie zou hebben. Het duurde daardoor even voordat alle andere leerlingen hun bestemming hadden gevonden en de Space groep kon beginnen. Daarna wilde de conciërge enkele leerlingen meenemen voor corvee, maar deze werd er snel van overtuigd dat de onderzoekssessie dat niet toeliet. Het is wel een mooie demonstratie van hoe de praktijk ‘zuiver’ wetenschappelijk onderzoek in de weg kan zitten.

Na een korte instructie gingen de leerlingen aan de gang. Space was op het netwerk van de school geïnstalleerd. Voor een leerling met een Mac (waarop Space niet werkt) was al een andere laptop geregeld. Helaas bleek het inloggen daar niet te lukken. Ook andere leerlingen bleken inlogproblemen te hebben. Dit soort problemen werd opgelost door leerlingen samen achter één computer te laten werken.

De leerlingen maakten een behoorlijk gemotiveerde indruk. Sommige leerlingen vroegen naar elkaars scores en waren er duidelijk op uit het minstens even goed te doen als de buurman/vrouw. Ook speelden sommige leerlingen die al vroeg klaar waren de game nog een keer, om hun score te verbeteren (hoewel ze daar niet altijd zelf op kwamen). Zo werkte

ook hier het competitie element, ook al wordt Space niet online gespeeld.

Zeker het eerste half uur was het redelijk rustig; de docent die aanwezig was als ordebewaker noemde dat tijdens het interview zelfs ‘verbazend rustig’. Deze groep stond als niet een van de makkelijkste te boek. Gedurende het tweede half uur ontstond er wat meer onrust omdat er al leerlingen klaar waren. De ordebewaker wil er toen al een punt achter zetten, maar de onderzoekers gaven aan dat dat nog niet kon. Enige meisjes raakten nogal gedemotiveerd, maar lieten zich met succes inspireren om verder te gaan door de begeleider.

Er kwamen wat leerlingen binnen die niet in deze groep horen en in deze ruimte wilden werken. Ze gingen er tussen zitten.

Op hulpvragen werd door de begeleider individueel ingegaan; zo nu en dan leidde dat tot een boodschap voor de hele groep. Vragen kwamen niet veel voor en hadden vooral betrekking op de redenen waarom ze dit moesten doen. Leerlingen accepteerden het antwoord dat ze dat later nog wel zouden horen. Sommige leerlingen waren zo verbaasd over het feit dat ze nu ‘legaal’ konden gamen dat ze dachten dat het een 1 april grap was.

De begeleider had de indruk dat er minder vragen over de game waren dan toen hij zijn eerste onderzoek met Space deed (in een vierde klas) en schreef dat toe aan het feit dat leerlingen nu een schriftelijke instructie hadden. Nog een verschil met zijn eerdere ervaringen was dat deze leerlingen voorzichtig waren

met het geluid van de game; wellicht heeft dat te maken met het feit dat deze leerlingen gewend zijn binnen een grote ruimte in kleine groepjes samen te werken.

Opmerkelijk tenslotte was dat sommige leerlingen zelf met begrippen als wrijving op de proppen kwamen.

Unreal sessie

Na een korte instructie van de docent en het uitdelen van de schriftelijke instructie speelden de leerlingen Unreal gedurende ongeveer een half uur individueel. De game was vooraf op de vaste pc's in een computerlokaal geïnstalleerd. Vervolgens speelden de leerlingen twintig minuten tegen elkaar in de online, multiplayer modus. Het invullen van de flowtest tenslotte nam zo'n tien minuten in beslag.

De introductie van het spel leidde tot een sterke betrokkenheid en opwinding. Tijdens het spelen waren leerlingen zeer geconcentreerd bezig. Dat de betrokkenheid groot was, moge blijken uit deze verzuchting van een leerling: 'jammer dat we moeten stoppen!' en het feit dat leerlingen erna nog druk hun prestaties in het spel aan het bespreken waren.

Vier meisjes werden tijdens het eerste half uur helaas misselijk; het karakter van het spel (first person shooter) kan daartoe aanleiding hebben gegeven, zeker in combinatie met de 'onvaste' besturing wanneer er sprake was van een wrijvingsloze situatie. Gezien in het licht van de pilot met Unreal, waarbij een aantal

deeltijdstudenten dezelfde verschijnselen vertoonden als deze meisjes, kan onbekendheid met gaming an sich ook zeker een factor zijn. Deeltijdstudenten zijn over het algemeen wat ouder en zullen minder gewend zijn games te spelen; voor drie van de vier meisjes in de Unreal groep gold dat ze thuis geen games speelden. Ze vonden de game in het begin overigens wel leuk. Bij de analyses met herhaalde metingen zijn de scores van deze vier meisjes niet meegenomen, om de vergelijking zo zuiver mogelijk te houden.

Het was zeker niet rustig te noemen tijdens het spelen van de game, wat bij Space wel het geval was. Zeker in het begin was er een opgewonden sfeer, met veel kreten en luide verkondigingen van de resultaten. Na een kwartier spelen werd het wat rustiger. De opwinding was weer in volle hevigheid terug tijdens het online spelen, toen leerlingen elkaar in de speelruimte ontdekten.

Technische problemen zijn er niet gesignaleerd. Tijdens het single player gedeelte kwamen er vragen over de score, of over de besturing en tijdens het multiplayer deel nam het aantal vragen toe. Leerlingen werkten in het begin soms samen. De begeleider liep in het begin alle leerlingen langs voor een korte uitleg en wees plenair op het feit dat er werelden met en zonder wrijving waren.

Een leerling gaf aan het spel niet leuk te vinden: ze zag niet de relatie met natuurkunde. Andere leerlingen maakten duidelijk dat ze het in beide opzichten niet met haar eens waren.

5.1.2 Interviews

Interviews met leerlingen: vooraf

Per groep is één leerling geïnterviewd, net na de afname van de pretests. De leerling uit de Space groep was een jongen, de beide andere leerlingen waren meisjes.

Opvallend is dat zowel deze leerlingen als de leerlingen die na afloop van het onderzoek zijn geïnterviewd, behoorlijk positief zijn over het vak natuurkunde. Ze zijn daar door de docent niet op geselecteerd; hun communicatievaardigheden waren een veel belangrijker factor. Een van de leerlingen vindt het interessant, maar wel moeilijk. De leerlingen vinden het een leuk vak en interessant omdat je ook zelf dingen moet uitzoeken, het is niet alleen maar lezen: 'nadenken over hoe dingen gaan en hoe dat komt'. Een van de leerlingen wil er ook in de toekomst iets mee doen (astrofysica).

Specifieke moeilijke onderwerpen noemen ze niet, maar deze leerlingen hebben ook nog maar weinig natuurkunde gehad.

De leerlingen kregen de vraag voorgelegd wat zij zouden doen als ze docent waren om ervoor te zorgen dat hun leerlingen alles goed snapten. De leerling die natuurkunde moeilijk vindt zou dan (meer?) demonstraties geven, maar de beide andere leerlingen zijn wel tevreden over hoe het vak nu gegeven wordt. Omdat natuurkunde een moeilijk vak is, krijgen leerlingen meer dan

bij andere vakken kant en klare bronnen aangeboden en hoeven ze dus niet alles zelf op te zoeken.

De leerlingen vinden het geen slecht idee om games op school te gebruiken, vooral vanwege de motiverende werking, maar opvallend genoeg vinden ze dat vooral wel goed voor anderen. Als aandachtspunten noemen ze wel het risico dat leerlingen het alleen als games zien, en dat niet iedereen gamen even leuk vindt.

Afstemming op de doelgroep is wel belangrijk, merkt een leerling op op. Games die voor kleine kinderen lijken bedoeld zijn uit den boze. Over een game voor natuurkunde:

'Nou ik heb wel eens op internet gelezen dat er zo iets bestaat. Waar je dan met een soort verhaal volgt waarin je dan oplossingen moet zoeken voor problemen die zich dan in dat verhaal voordoen. Ja, soort raadsel oplossen schat vinden'

Op Unic worden nog geen games gebruikt volgens deze leerlingen, en daarin zijn ze het eens met de geïnterviewde docenten, net als over het feit dat er wel mogelijkheden zijn omdat iedereen een laptop heeft. Eén leerling noemt een voorbeeld van een quiz over alcohol, met kennisvragen en levels, waar simulaties doorheen waren vervlochten. Ze vond dat op zich een leuke leer-methode.

Van deze drie leerlingen is er één een regelmatige gamer (kleine 4 uur per week), speelt de tweede leerling een enkele keer en de derde nooit.

Interviews met leerlingen: achteraf

Per groep is één leerling geïnterviewd, net na de afname van de posttests. De leerling uit de Space conditie was een meisje, de beide andere leerlingen waren jongens.

De leerlingen vonden alle drie dat ze de afgelopen tijd veel geleerd hadden over natuurkunde, over de onderwerpen waar het om ging in deze twee weken. Alle drie de leerlingen zijn erg enthousiast over het vak. De lessen hebben daar geen invloed op gehad.

‘hoe je dingen moet berekenen (vergeet het ook niet meer)’

‘nu realiseer ik mij steeds meer als je op straat bent hoeveel dingen te maken hebben met natuurkunde’

Over de vraag hoe ze het gevonden hadden in deze lessen een game te spelen, werd Unreal wat positiever beoordeeld dan Space leerling. De leerling die Unreal had gespeeld vond het leuk om een game te spelen, in tegenstelling tot de keer dat hij eerder voor een onderzoek een game had moeten spelen en die nogal kinderachtig had gevonden. De leerling uit de Space groep vond het leuk ter afwisseling, maar vindt dat ze meer leert van werkbladen en uitleg. Over de game zelf waren ze positief, vooral de Unreal leerling.

‘Ik vind zo ie zo online spelen heel erg leuk dat je bezig bent met andere leerlingen want ja dat is toch leuker dan dat je tegen de computer speelt.’

Dus dan kan je laten zien hoe goed je bent en het spel was wel leuk in elkaar gezet ... schatkistjes zoeken en zo ... Dus zag er wel leuk uit. Zag er goed uit’

Ook de leerling die practicum had gedaan was wel te spreken over de inhoud ervan. Vooral de tijdtikker vond hij leuk, ‘omdat je het nu echt doet, ipv uit een boek’.

Beide leerlingen in de game conditie vonden het niet zo makkelijk om de game te spelen. De leerling in de Unrealgroep Unreal gaf aan dat het lastig was om te manoeuvreren; elk level was anders, dus omschakelen was lastig en hij gleed alle kanten op. De leerling in de Space groep vond het begin lastig, daarna ging het beter. Het besturen bleek heel rustig te moeten, vooral bij smalle doorgangen. Als de interviewer naar de factor wrijving vraagt, blijkt de leerling daar niet echt op te hebben gelet.

De leerling in de practicumgroep ondervond geen problemen bij het uitvoeren van de opdrachten. Hij had ook voldoende tijd, in tegenstelling tot de beide andere leerlingen. De leerling in de Space conditie had het laatste level nog niet gehaald. De leerling in de Unreal groep vond het jammer dat hij nog niet alle levels had kunnen oefenen voor ze online gingen; vooral over dat online gedeelte was hij erg enthousiast.

De leerlingen hadden alleen wat praktische hulp van de docent nodig, zoals bij het opstarten van Unreal.

De leerlingen hebben niet het ge-

voel dat ze veel geleerd hebben door de game/het practicum en ze leggen de link met de verwerkingsopdrachten als feedback niet direct zelf.

De leerling uit de Unreal groep heeft misschien een beetje gevoel gekregen van hoe de bewegingen gaan. De bedoeling van de game was niet duidelijk. Als hij echt iets moet leren, kan hij zich daarop concentreren.

‘als ik iets wil leren moet ik mij ook ergens op kunnen concentreren en dat werd niet gegeven dus dat was wel ... Dus dan weet je niet zo goed wat je zou kunnen leren van de game’

Als de interviewer de link met de verwerkingsopdrachten legt, geeft hij aan dat de game daar wel bij hielp; had wel enige vragen goed.

Ook de Space leerling komt zelf niet op het verband met de verwerkingsopdrachten. Als de interviewer haar daar op wijst zegt ze dat ze wel wat moest terugdenken aan de game, maar niet vaak. Game zou moeten worden uitgebreid met andere dingen.

‘Nou als wij bijvoorbeeld een opdracht moesten doen dat het zeg maar de richting van waar ergens iets naar toe ging dan herken je dat wel als je dat moest besturen zeg maar.’

De leerling die practicum heeft gedaan zei over de muntenproef dat hij wel geleerd heeft dat het zo is, niet waarom. Ook hij legde de link met de verwerkingsopdrachten kennelijk niet:

‘maar misschien als ik in het vervolg een opdracht wat hier op aansluit krijg misschien zou ik die opdracht beter kunnen maken door het practicum.’

Bij de vragen over de FCI had de leerling uit de Unreal groep niet direct door welke test nu eigenlijk bedoeld werd. Evenals het feit dat leerlingen niet direct zelf de link met de verwerkingsopdrachten legden, bevestigt dit het beeld dat de docent-onderzoeker tijdens zijn interview schetst, dat dit onderzoek voor docenten en onderzoekers veel meer leeft dan voor leerlingen – voor leerlingen hoort het bij de ‘schoolingen’.

De leerling uit de Unrealgroep deed het de tweede keer naar eigen zeggen beter op de test, maar hij had nog niet veel bijgeleerd over richtingen.

‘Beter wel niet helemaal goed want ik had nog niet veel geleerd over richtingen waarin ze glijden Maar van die krachtenwist ik het wel een beetje met constante snelheid en zo’

De leerling uit de Space groep vond de vragen wel duidelijk, maar haar kennis was niet zo groot en is ook niet toegenomen. De leerling uit de practicumgroep vond vragen duidelijk en interessant, en niet heel moeilijk.

De leerlingen in de game condities hebben het niet inhoudelijk met andere leerlingen over de games gehad. Volgens de leerling in de Unreal groep waren zijn medeleerlingen best ja-

loers. Ouders waren verbaasd, maar vonden het wèl leuk. Hij wist dat er drie groepen waren en had begrepen dat het andere spel wat saaier was.

De leerling in de practicumgroep heeft het er met zijn vriend over gehad, maar weet niet zeker of die in de game groep zat. Ze hebben het vooral over de lesstof gehad

Beide leerlingen in de game condities vinden het een goed idee om games in de lessen te gebruiken, vooral vanwege de motiverende werking. Leerlingen die normaal niet zoveel deden, deden nu ook gewoon mee. De leerling uit de Unreal groep legt de link met de opdrachten nu wel:

‘Maar als het zo wordt gebracht is het wel erg leuk ja Ik vond het ook een erg goed idee Nu was ik erg in verdiept normaal zit een beetje aan een thema te werken nu was ik erg bezig met dat spel Bij die vragen ben je weer bezig met dat spel’

Space mag wel wat moeilijker, zodat je er meer van leert.

De leerling in de practicum groep is aanzienlijk minder positief. Hij vindt het geen aanrader, maar weet niet precies hoe de games in z'n werk gingen, dus kan er niet echt over oordelen. Deze leerling is geen gamer en vindt een practicum ook leuk. Games zouden geen boeken mogen vervangen.

Gevraagd naar de vakken waar games bij gebruikt zouden kunnen worden doen de leerlingen suggesties als maatschappijleer, wiskunde en Frans.

Interviews met docenten: vooraf

De interviews vooraf zijn gehouden met twee docenten. Een van hen zou tijdens de uitvoering als ‘ordebewaker’ optreden. Hij geeft wiskunde aan deze groepen. De andere docent is de teamleider van de onderbouw en sportdocent.

Beide docenten kennen de opzet van het onderzoek maar weten er inhoudelijk weinig tot niets over. Hun verwachtingen liggen vooral op het vlak van motivatie en van de voordelen die simulaties met zich mee brengen. De wiskunde docent ziet vooral voordelen in de mogelijkheid de spanningsboog te verlengen, maar verwacht geen leerrendement, omdat leerlingen energie moeten steken in een nieuwe leervorm. Voor de wiskunde docent is duidelijk het controle aspect in een klas belangrijk. Net als de teamleider voorziet hij, gevraagd naar mogelijke problemen, deze op het technische vlak, maar ook vraagt hij zich af of de docent wel goed kan controleren of de goede game gespeeld wordt. De docent moet de game zelf ook kunnen spelen.

Beide docenten verwachten dat leerlingen het leuk zullen vinden, maar vinden daarbij beiden sturing belangrijk. Leerlingen hebben niet vanzelf de nodige reflectie, schat de wiskunde docent in. Leerlingen hebben het idee dat gamen is iets anders is dan leren. Ze hebben geen helicopterview, en kunnen dus niet beoordelen of ze iets geleerd hebben. Hij verwacht dat ze niet denken dat ze iets geleerd hebben, dat ze

liever anderhalf uur een verhaal van de docent horen.

Naar de mening van de teamleider is de koppeling met het leerdoel belangrijk. Havo/vwo leerlingen willen weten wat ze eraan hebben. Wanneer je leerlingen vooraf bewust maakt van dat doel kijken leerlingen anders. Hij verwacht dat leerlingen zullen vragen: waarom moeten we dat doen? Of ze vinden het zo leuk dat ze gewoon gaan spelen.

Beide docenten zien goede mogelijkheden voor games in het onderwijs. Daarbij is een goede organisatie en aangetoond rendement belangrijk. Het past goed bij het karakter van de school en bij de technologie die op school gebruikt wordt. Alle leerlingen hebben laptops. Uitdagend onderwijs past bij het UniC en de school zou op dit gebied een voorloper moeten zijn.

Zij hebben zelf geen ervaring met games in het onderwijs. De wiskunde docent zou ze vooral in de lagere klassen gebruiken, om de concentratie te verhogen. Leerlingen in de hogere klassen hebben dat hulpje niet nodig. Leerlingen zijn naar zijn idee zeer gevoelig voor vorm en moeten zo goed mogelijk voorbereid worden op het examen (niet met games dan dus).

De teamleider ziet geen mogelijkheden voor games in zijn vakgebied (sport); daar zijn geen digitale middelen voor nodig.

Gevraagd naar de reden waarom er nog zo weinig gebruik gemaakt wordt van games in het onderwijs, noemen beide docenten het examen: het moeten voldoen aan de eisen ervan kan een struikelblok zijn. De regels zijn streng, dus vernieuwing is riskant.

Het eindexamen is al 50 jaar hetzelfde.

De wiskunde docent geeft aan dat docenten behoudend zijn; games zijn dan erg out of the box. Nieuwe digitale middelen worden gebruikt als vervolg op wat er al was (digitale stift); bij games zijn er geen voorlopers.

‘De beamer is een vervanger van het whiteboard wat we al kennen, en dat is met smartboards ook, het percentage docenten dat een smartboard gebruikt is relatief groot. Maar het percentage daarvan dat het smartboard niet gebruikt als smartboard maar als veredeld projectiescherm, dat percentage is nog steeds hoog dus je kunt je afvragen of een smartboard echt een onderwijsvernieuwing is, of dat we eigenlijk nog steeds met een krijtje op een bord staan te kalken. Alleen is dat krijtje veranderd in een digitale stift. Snap je, dus wat is daar in vorm nou echt aan veranderd, nou dan kun je je afvragen of, en met een game is het probleem dat er geen voorloper is... en.. en daarom is het veel meer out of the box.’

Docenten zijn conventionele beroepsgroep;

Ook de teamleider ziet verschillen tussen een innovatie als het smartboard en games. Het smartboard is een hulpmiddel; een game is iets anders. De docent geeft daarmee dingen uit handen en het heeft meer gevolgen voor je inhoudelijke onderwijs. Games hebben daarbij negatieve connotaties als verslaving etc. Een voordeel games is dat je data kunt opslaan.

Interviews met docenten: achteraf

De interviews na afloop (na de posttests) werden gehouden met de docent-onderzoeker en met een docent die tijdens de uitvoering als ordebewaker was opgetreden (een andere docent dan de docent die vooraf was geïnterviewd. Deze docent geeft management en economie.

De docent-onderzoeker vond de twee weken van het onderzoek een interessant leerproces, met veel organisatie en beetje stress. Het onderzoek is een deel van het geheel; hij moet ook gewoon lesgeven.

Er waren tijdens de uitvoering (games en practicum) niet echt problemen. De economie docent vond het verrassend soepel verlopen. De leerlingen gingen snel aan de slag (sneller dan anders). Leerlingen in de derde klas zijn op een moeilijke leeftijd: ze puberen volop en hebben niet alle remmingen, wel de emoties.

Beide docenten hadden zeker het idee dat de leerlingen gemotiveerd meededen. De docent-onderzoeker gaf wel aan, dat naar zijn idee zelf niet de link met de test leggen. Bij het napraten met sommige leerlingen kreeg hij de vraag: waarom doe ik dit eigenlijk? Hij merkte op dat leerlingen hier veel minder mee bezig zijn dan docenten/onderzoekers: het is gewoon school.

Bij doorvragen over noodzaak leggen link: uitleg beïnvloedt testresultaten.

Hij vindt de link tussen theorie en

praktijk vooral bij Unreal lastig, als je niet goed in de stof zit. Space is meer to the point, dus de koppeling is beter te leggen.

‘Op het moment dat jij een omgeving aanbiedt die niet helemaal naadloos aansluit met de steriele theorie om het zo maar te zeggen, wat de Space Challenge wat meer heeft omdat het voor de kaders van het materiaal is gemaakt, zou je dus II moeten helpen in het proces. Ik denk dat een game als Space veel makkelijker te implementeren is omdat je er minder verwerkingsopdrachten in hoeft te stoppen’

Ook het Unic concept maakt het leggen van die koppeling lastig. Achteraf gezien is er, vergeleken met de andere voorbereidingen, toch te weinig aandacht besteed aan die koppeling. Dat is dus een leerpunt voor een herhaling.

Het in het vooruitzicht stellen van bonuspunten heeft naar zijn idee zeker geholpen.

De games werden op 1 april gespeeld en sommige leerlingen dachten dat het een grap was, vertelt de economie docent.

‘Ja zeker als ze in een keer te horen krijgen nadat ze jaren te horen dat ze niet mogen gamen dat ze in een keer moeten gamen. Nou dat motiveert wel. Daarom vond ik het leuk om te zeggen: nou jongens verplicht gamen’

Bij het afnemen van de tests waren er geen problemen. Wel con-

stateerde de docent-onderzoeker vooral bij de pretests ontbrekende voorkennis. Ook blijkt de woordenschat van leerlingen soms minder groot dan je zou verwachten.

De docent-onderzoeker schat in dat leerlingen weinig met elkaar over het onderzoek gepraat hebben: als je er zelf geen big deal van maakt, doen leerlingen dat ook niet.

Beide docenten zag hun rol tijdens de uitvoering (resp. practicum en Space) vooral als organisatorisch. De docent-onderzoeker gaf tijdens het practicum waar nodig wat verheldering van de opdrachten. De economie docent hield zich vooral bezig met klassenmanagement en motiveerde leerlingen waar nodig een hogere score te halen.

De docent-onderzoeker ziet wel enige mogelijke verbeterpunten. De: FCI is wel erg moeilijk, ook vanwege het taalgebruik. Uitvoering van het onderzoek in de vierde klas stuit echter op praktische bezwaren (te weinig leerlingen). De verwerkingsopdrachten verdienen meer aandacht en zouden direct na practicum/game moeten worden afgenomen, met een docent die daarbij coacht. De klassieke vraag bij practica was altijd al: doe je dat voor of na de theorie?

Een ander probleem bij de opzet was dat leerlingen een game/practicum kregen toen ze nog niet echt bij de stof waren waar het om ging; dat was bij de verwerkingsopdrachten wel zo. Organisatorisch is dit moeilijk te regelen.

De groepssamenstelling was gebaseerd op de rol die leerlingen in

een musical hadden: A – acteurs, B= bouwers en C= overigen. De leerlingen in de Space groep (C) waren mogelijk de wat minder gemotiveerde leerlingen.

Beide docenten zijn positief over het idee om games in het onderwijs te gebruiken. De docent-onderzoeker merkt daarbij op dat het rendement wel belangrijk is en ook, dat games niet top-down worden ingevoerd.

‘Als je iemand hebt die zijn hele leven met boeken is opgegroeid moet je niet zo’n game geven. Daar wordt hij helemaal niet gelukkig van. Toen het studiehuis werd ingevoerd zat ik in Havo 4 zaten er docenten die hun hele leven klassikaal hadden lesgegeven waren daar hartstikke goed in. Werd in eens gezegd nu moet je op deze manier lesgeven. Dat werkt gewoon niet.’

De economie docent geeft aan dat het wel praktisch haalbaar moet zijn, in 1 à 2 lesuren.

Beide docenten zouden wel games willen gebruiken. De docent-onderzoeker vindt het dan wel belangrijk dat dat geleidelijk gebeurt. Wat werkt? Waar voel ik me comfortabel bij? Er zijn uren en ruimte nodig, en je moet als docent verantwoording afleggen. De eerste opdracht voor een docent is, het curriculum te halen, met voldoende rendement. Daarnaast is er ruimte voor innovatie.

De economie docent ziet wel wat in games waarbij de bij variabele kosten bij onderneming inzichtelijk worden gemaakt, bijvoorbeeld door

kleurgebruik.

Beide docenten spelen zelf games. In hun opinie past het gebruik van games goed in het schoolklimaat. Er is volop ruimte en stimulatie om te innoveren.

‘Dat heeft er mee te maken hoe het team is samengesteld. Het zijn of jonge ambitieuze docenten die hier toekomst in zien of oudere docenten die wouden innoveren maar op hun oude school niet konden doen en op Unic terecht zijn gekomen.’

Naar de inschatting van de docenten is een groot deel van hun collega's niet op de hoogte van dit onderzoek. De docent-onderzoeker schat dat aantal wat lager in (50%) dan de economie docent (80%). De docent-onderzoeker heeft het er wel over met andere docenten; iedereen in het team is naar zijn idee op de hoogte. Volgens de economie docent weten alleen de staf en de direct betrokken ervan; hijzelf is op de hoogte omdat hij gevraagd is op te treden als ordebewaker.

5.2 Resultaten: tests en vragen

5.2.1 Analyses

Om de betrouwbaarheid vast te stellen van de tests die in het onderzoek gebruikt zijn, zijn de Apha waarden ervan bepaald. Bij de berekening van de alpha worden tests waarin missing values voorkomen niet meegenomen.

De betrouwbaarheid van een test wordt vooral bepaald door het aan-

tal items en door de variantie tussen de scores. Als de 'treatment' effect heeft, wordt de variantie tussen de groepen leerlingen groter en zal er dus bij de posttests een hogere betrouwbaarheid optreden. De meest zuivere meting is daarom de betrouwbaarheid op de pretest, omdat er dan nog geen verschillen tussen leerlingen zijn 'aangebracht'. Wij zullen de betrouwbaarheid op de posttests wel vermelden, maar de betrouwbaarheidsanalyses op de subschalen beperken tot de pretests.

Vervolgens is nagegaan of de leerlingen in de drie condities op de pre-test al verschilden op de afgenomen tests. De analyse naar de effecten van de treatment: Space en Unreal, bestond uit een herhaalde metingen analyse. Dit is de aangewezen analyse in een quasi experimenteel pretest-posttest design, zoals dat in dit onderzoek is gehanteerd. In zo'n analyse worden de scores van iedere leerling op de pre-test vergeleken met de scores op de posttest en wordt nagegaan wat daarbij het effect is van de conditie: Space, Unreal of controle. Omdat gebleken was dat er op alle tests en op de vragen vooraf van de posttest aanzienlijke schoolverschillen waren, werd in de herhaalde metingen analyse ook de school als factor meegenomen.

Waar we via een herhaalde metingen analyse naar op zoek zijn, zijn 'within' effecten: verschillen tussen de scores op de pre- en posttests. Wil de experimentele conditie het hierbij beter doen dan de controle

conditie, dan moet dat blijken uit ‘within interactie effecten’. Hierbij ging het zowel om de scores op de FCI als om de scores op de natuurkundemotivatietest.

De herhaalde metingen analyses zijn uitgevoerd over de gegevens van leerlingen die zowel een pretest als een posttest hebben ingevuld, en die daadwerkelijk de game hebben gespeeld/het practicum hebben uitgevoerd. De gegevens van de vier leerlingen die tijdens het spelen van Unreal misselijk zijn geworden en niet verder hebben meegedaan, zijn in deze analyses niet meegenomen. Eén van de leerlingen bleek bij geen van de drie gelegenheden (pretest, posttest en uitvoering) aanwezig te zijn geweest, zodat bij de analyses de data van maximaal 92 leerlingen beschikbaar waren.

Over de gegevens van de flow tests en de evaluatievragen achteraf zijn variantie analyses uitgevoerd.

Daarnaast is gekeken naar factoren die in eerder onderzoek van invloed zijn gebleken: geslacht en de vraag of leerlingen thuis ook games spelen, en naar de invloed van het natuurkundefcijfer op een reguliere toets. De groepen verschilden vooraf niet op dit cijfer.

5.2.2

Betrouwbaarheid van de tests

FCI

De betrouwbaarheid van de FCI, die uiteindelijk uit 19 van de 30 oorspronkelijke items bestond, was niet hoog: .37 in de pretest ($n=92$) en .51 in de posttest ($n=67$). Bij de berekening van de betrouwbaarheid in de posttest zijn de leerlingen die bij de uitvoering niet aanwezig waren en de leerlingen die tijdens het spelen van Unreal zijn weggegaan, niet meegenomen. Dat is de zuiverste betrouwbaarheidsbepaling, omdat voor hen niet de (volledige) ‘treatment’ geldt. Wanneer de gegevens van deze leerlingen wel worden meegenomen, ligt de betrouwbaarheid in de pretest iets hoger: .57 ($n=82$).

De betrouwbaarheid is aanzienlijk lager dan de betrouwbaarheid over deze 19 items voor de data uit het onderzoek van Martijn Hoevenaars, die ook al niet hoog lag (.54). Wat bij de nog lagere betrouwbaarheid in dit onderzoek een rol zal spelen is het feit dat de variantie tussen de scores niet erg groot kan zijn, omdat de scores erg laag liggen. Het gemiddelde op de pretest is 4,86 ($sd=2,26$), met een range van 1-13. Tachtig procent van de scores ligt tussen de 1 en 6.

De gemiddelde scores van de vierdeklassers in het onderzoek van Martijn Hoevenaars was 6,59 ($sd=2,74$), met een range van 3 tot 16. Hier ligt 90% van de scores van 3 tot 9.

De gemiddelden, standaarddeviaties en rirwaarden van de items staan vermeld in tabel 5.1.

TABEL 5.1 GEMIDDELDEN, STANDAARDDEVIATIES EN RIR-WAARDEN VAN DE FCI ITEMS IN DE PRETESTS

	GEM.	SD	RIR	N
FCIPRE1	,12	,326	,137	92
FCIPRE2	,07	,248	-,074	92
FCIPRE3	,59	,495	,086	92
FCIPRE4	,45	,500	,210	92
FCIPRE5	,24	,429	,147	92
FCIPRE6	,27	,447	,049	92
FCIPRE7	,14	,350	,416	92
FCIPRE8	,36	,482	-,165	92
FCIPRE9	,13	,339	,183	92
FCIPRE10	,17	,381	-,089	92
FCIPRE11	,49	,503	,279	92
FCIPRE12	,25	,435	,214	92
FCIPRE13	,34	,475	,105	92
FCIPRE14	,22	,415	,188	92
FCIPRE15	,13	,339	,246	92
FCIPRE16	,16	,371	,130	92
FCIPRE17	,12	,326	-,017	92
FCIPRE18	,26	,442	-,178	92
FCIPRE19	,36	,482	,211	92

Binnen de FCI kunnen drie subsets worden onderscheiden:

Newton (item 1,2,3,4,7,8,9,10,16,17,18,19)

Vector (item 5,6,13,14,15)

Kinematica (item 11 en 12)

De betrouwbaarheden voor deze subsets zijn weergegeven in tabel 5.2.

TABEL 5.2 BETROUWBAARHEID VAN DE SUBSETS VAN DE FCI

SUBSET	ALPHA	GEMIDDELDE	SD	AANTAL ITEMS
NEWTON	.02	2,92	1,41	12
VECTOR	.27	1,20	1,15	5
KINEMATICA	.56	.74	.78	2

De items waar het in dit onderzoek om gaat (Newton) vertonen dus geen samenhang, itt tot de twee items die over Kinematica gaan. Deze beide items gaan over twee apparaatjes die langs een meetlint naar rechts bewegen en iedere 0,20 seconde een blokje laten vallen.

Natuurkunde motivatie

De betrouwbaarheid van de vragenlijst Natuurkunde Motivatie is hoog, zowel op de pretest (alpha =.94, n=86) als op de posttest (alpha = .94, n=80). Wanneer bij de posttest alleen de data worden meegenomen van de leerlingen die (volledig) bij de uitvoering waren, blijft de betrouwbaarheid even hoog (n=65). De vragenlijst bestaat uit 17 items.

De betrouwbaarheden voor de subschalen van de pretest zijn weergegeven in tabel 5.3.

TABEL 5.3 BETROUWBAARHEID VAN DE SUBSCHALEN VAN DE VRAGENLIJST NATUURKUNDE MOTIVATIE				
SUBSCHAAL	ALPHA	GEMIDDELDE	SD	AANTAL ITEMS
ANGST/GEMAK	.88	20,53	5,54	6
NUT/RELEVANTIE	.86	20,66	5,25	6
PLEZIER	.91	16,68	5,42	5

Flow test

Ook de betrouwbaarheid van de flow test is hoog: alpha = .91 (n=34). Ook wanneer de scores van de vier leerlingen die dev uitvoering slechts gedeelte-lijk hebben meegemaakt wel worden meegenomen blijft de betrouwbaarheid hoog (.91, n=38). De Flow test bestaat uit 23 items.

De betrouwbaarheden voor de subschalen van de pretest zijn weergegeven in tabel 5.4.

TABEL 5.4 BETROUWBAARHEID VAN DE SUBSCHALEN VAN DE FLOW TEST				
SUBSCHAAL	ALPHA	GEMIDDELDE	SD	AANTAL ITEMS
ANTECEDENTEN	.80	34,54	8,48	10
ERVARING	.88	41,39	11,31	13

-

5.2.3 Verschillen tussen de groepen

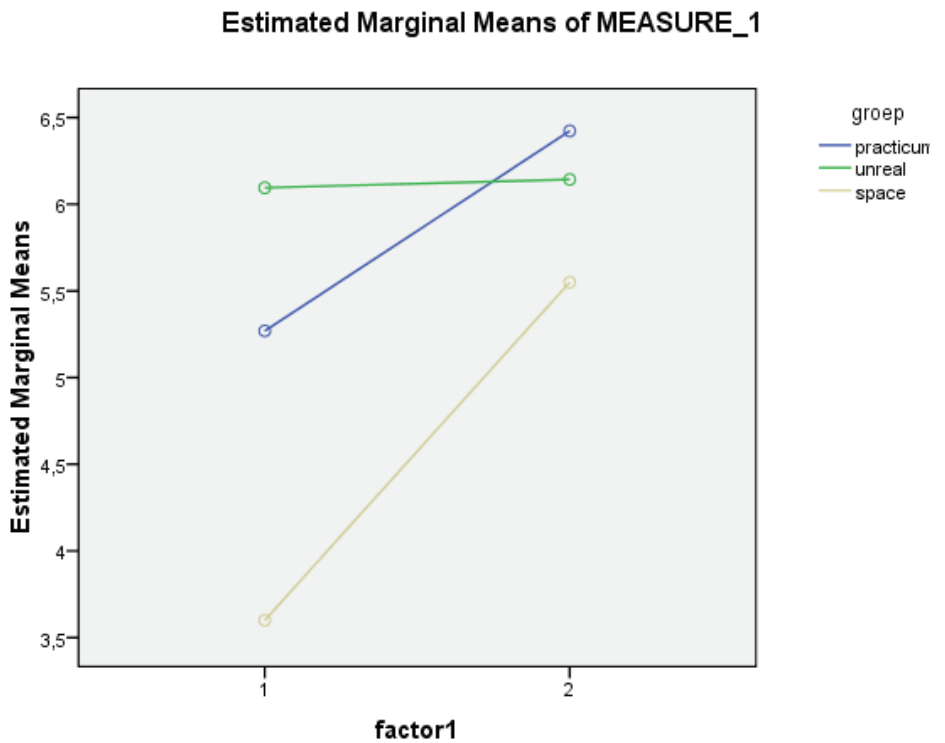
FCI

De score op de FCI pretests bleek significant samen te hangen met de score op de pretest van de vragenlijst Natuurkunde Motivatie ($r = .32$, $p = .003$). Op de voortest waren er geen verschillen tussen de drie groepen in hun motivatie voor natuurkunde, evenmin als in hun cijfer. Er was dus geen reden om het cijfer en/of de motivatie voor natuurkunde als covariaat mee te nemen in de herhaalde metingen analyse over de FCI scores. De resultaten van deze analyse zijn weergegeven in tabel 5.5.

TABEL 5.5 RESULTATEN HERHAALDE METINGEN ANALYSE OP DE SCORES VAN DE FCI

EFFECT	F	P
OVERALL WITHIN EFFECT	15,5	.000
INTERACTIE WITHIN EF- FECT/GROEP	3,991	.02

Het overall within effect geeft aan dat er overall sprake is van een sterk significante leerwinst: leerlingen scoorden op de posttest hoger dan op de pretest. De groepen verschillen echter ook in de mate van leerwinst; dat is de interactie waar we op uit waren. In figuur 5.1 en in tabel 5.6 staat weergegeven hoe de leerwinst er voor de drie groepen uitziet:



Figuur 5.1 Leerwinst op de FCI in de drie groepen

TABEL 5.6 GEMIDDELDEN EN STANDAARDDEVIATIES OP DE FCI VOOR DE DRIE GROEPEN (19 ITEMS)

	GROEP	GEMIDDELDE	SD	N
FCIPRETEST	practicum	34,54	8,48	10
	unreal	6.10	2.682	21
	space	3.60	1.667	20
FCIPOSTTEST	practicum	6.42	2.580	26
	unreal	6.14	3.582	21
	space	5.55	1.731	20

De verschillen blijven bestaan wanneer de vier leerlingen die tijdens het spelen van Unreal zijn weggegaan wegens misselijkheid wèl worden meege-
nomen. De F waarde voor het interactie effect de overall herhaalde metingen
analyse is dan $F=4,78$, $p= .01$. Ook het verschil tussen de Unreal groep en
de Space groep is dan nog steeds significant ($F=10,63$, $p= .002$).

Subsets van de FCI

Wanneer voor de drie verschillen-
de subsets van de FCI een herhaalde
metingen analyse wordt uitgevoerd,
blijken de verschillen die voor de FCI
in zijn geheel golden alleen op te
treden bij de vragen die direct met
de wetten van Newton te maken heb-
ben. Op deze vragen treedt overall
een significante leerwinst op en een
interactie effect.

TABEL 5.7 RESULTATEN HERHAALDE
METINGEN ANALYSE OP
DE SCORES VAN DE
NEWTON SUBSET

EFFECT	F	P
OVERALL WITHIN EFFECT	7,19	.009
INTERACTIE WITHIN EFFECT/ GROEP	3,34	.04

Net als bij de FCI in zijn geheel
blijkt ook bij deze subset het interac-
tie effect op conto te komen van het
verschil tussem de Space en de Un-
real groep. De F waarde voor het inter-
actie effect bij de vergelijking tussen
deze twee groepen is 6,79 ($p=.01$).
Bij de overige vergelijkingen treden
geen significante verschillen op.

Dat de verschillen die voor de hele
FCI gevonden worden alleen terug te
vinden zijn bij de Newton subset kan
met de inhoud te maken hebben,
maar ook – en dat is waarschijnlijker
– met het feit dat dit veruit de groot-
ste subset is: 12 vragen, tegenover
5 vragen in de Vector subset en 2
vragen die over Kinematica gaan.

TABEL 5.8 GEMIDDELDEN EN STANDAARDDEVIATIES OP DE SCORES VAN DE NEWTON SUBSET (12 ITEMS_)

	GROEP	GEMIDDELDE	SD	N
NEWTONPRE	practicum	3.1538	1.04661	26
	unreal	3.7619	1.64027	21
	space	2.1500	1.18210	20
NEWTONPOST	practicum	3.8846	1.58308	26
	unreal	3.5714	2.39940	21
	space	3.3000	1.52523	20

Ook op item niveau is nagegaan waar de leerwinst precies zit. Daartoe zijn voor de hele populatie en per groep afhankelijke t-tests uitgevoerd. De resultaten daarvan zijn weergegeven in tabel 5.9.

TABEL 5.9 VERGELIJKING TUSSEN DE PRE- EN POSTTEST VOOR DE FCI OP ITEMNIVEAU

		GM TOTAAL	SD TOTAAL	GEM PRACT	SD PRACT	GEM UNREAL	SD UNREAL	GM SPACE	SD SPACE
PAIR 1	FCIPRE1	.13	.341	.04A	.192	.33	.483	.05	.224
	FCIPOST1	.16	.371	.04A	.192	.38	.498	.10	.308
PAIR 2	FCIPRE2	.07	.265	.08	.272	.05	.218	.10	.308
	FCIPOST2	.13	.344	.04	.196	.29	.463	.10	.308
PAIR 3	FCIPRE3	.57	.499	.69	.471	.57	.507	.40	.503
	FCIPOST3	.63	.487	.81	.402	.62	.498	.40	.503
PAIR 4	FCIPRE4	.42	.497	.54	.508	.52	.512	.15	.366
	FCIPOST4	.46	.502	.58	.504	.24	.436	.55	.510
PAIR 5	FCIPRE5	.30	.461	.19	.402	.48	.512	.25	.444
	FCIPOST5	.37	.487	.38	.496	.38	.498	.35	.489
PAIR 6	FCIPRE6	.25	.438	.31	.471	.24	.436	.20	.410
	FCIPOST6	.40	.494	.46	.508	.29	.463	.45	.510
PAIR 7	FCIPRE7	.15	.359	.15	.368	.24	.436	.05	.224
	FCIPOST7	.28	.454	.27	.452	.33	.483	.25	.444
PAIR 8	FCIPRE8	.42	.497	.38	.496	.33	.483	.55	.510
	FCIPOST8	.36	.483	.35	.485	.33	.483	.40	.503
PAIR 9	FCIPRE9	.15	.359	.12	.326	.24	.436	.10	.308
	FCIPOST9	.21	.410	.27	.452	.19	.402	.15	.366
PAIR 10	FCIPRE10	.21	.410	.27	.452	.19	.402	.15	.366
	FCIPOST10	.25	.438	.35	.485	.19	.402	.20	.410

TABEL 5.9 VERGELIJKING TUSSEN DE PRE- EN POSTTEST VOOR DE FCI OP ITEMNIVEAU

		GM TOTAAL	SD TOTAAL	GEM PRACT	SD PRACT	GEM UNREAL	SD UNREAL	GM SPACE	SD SPACE
PAIR 11	FCIPRE11	.51	.504	.54	.508	.57	.507	.40	.503
	FCIPOST11	.58	.497	.65	.485	.48	.512	.60	.503
PAIR 12	FCIPRE12	.27	.447	.35	.485	.33	.483	.10	.308
	FCIPOST12	.28	.454	.42	.504	.33	.483	.05	.224
PAIR 13	FCIPRE13	.30	.461	.46	.508	.14	.359	.25	.444
	FCIPOST13	.37	.487	.27	.452	.38	.498	.50	.513
PAIR 14	FCIPRE14	.21	.410	.23	.430	.24	.436	.15	.366
	FCIPOST14	.28	.454	.19	.402	.48	.512	.20	.410
PAIR 15	FCIPRE15	.15	.359	.04	.196	.33	.483	.10	.308
	FCIPOST15	.16	.373	.15	.368	.24	.436	.10	.308
PAIR 16	FCIPRE16	.15	.359	.19	.402	.24	.436	.00	.000
	FCIPOST16	.13	.344	.12	.326	.14	.359	.15	.366
PAIR 17	FCIPRE17	.12	.327	.04	.196	.19	.402	.15	.366
	FCIPOST17	.21	.410	.19	.402	.19	.402	.25	.444
PAIR 18	FCIPRE18	.25	.438	.12	.326	.33	.483	.35	.489
	FCIPOST18	.27	.447	.27	.452	.19	.402	.35	.489
PAIR 19	FCIPRE19	.40	.494	.54	.508	.52	.512	.10	.308
	FCIPOST19	.51	.504	.62	.496	.48	.512	.40	.503

Legenda:

grijs: significante leerwinst

blauw: significante afname in scores

rood: trendmatig verschil ($.05 < p < .10$)

Het globale beeld is dat de Spacechallenge groep het meest stabiele beeld m.b.t. leerwinst op aangereikte concepten vertoont.

De FCI meet meer concepten dan er in SpaceChallenge worden “behandeld”. De concepten die worden behandeld komen uit de verf. Waar we winst verwachten vinden we die ook, zij het soms niet significant.

De Unrealgroep heeft vreemde uitschieters. Het leren bij de Unreal groep staat ogenschijnlijk los van de game. Leerwinst is soms conform de verwachtingen en dan weer geheel tegengesteld en soms zelfs significant negatief. Misschien dat op andere plaatsen in de lessenreeks deze concepten zijn aangeleerd?

Het overall within effect geeft aan dat er overall sprake is van een sterk significante leerwinst: leerlingen scoren op de posttest hoger dan op de pretest. De groepen verschillen echter ook in de mate van leerwinst; dat is de interactie waar we op uit waren.

In figuur 5.1 en in tabel 5.6 staat weergegeven hoe de leerwinst er voor de drie groepen uitziet:

SpaceChallenge is ontwikkeld met de volgende leerdoelen voor ogen:

- GEEN KRACHT -> RECHTE BAAN (FCI VRAAG 4)
- ZONDER KRACHT BLIJFT SNELHEID CONSTANT (FCI VRAAG 7)
- REMMEND EFFECT VAN WRIJVING OP DE BEWEGING (FCI VRAAG 19)

Op de vragen in de FCI waar deze concepten getest worden scoren de

SpaceChallenge spelers beter, zij het niet altijd significant. De Unreal spelers scoren soms beter, soms slechter. Het stabiele beeld van begripsverbetering voor deze vragen dat we bij de SpaceChallenge groep zien ontbreekt hier.

Vraag 4 kwam bijna letterlijk in het practicum voor, maar deze groep scoorde niet beter op deze vraag na het practicum. Dat is conform het bekende verschijnsel dat practicum nauwelijks kan bijdragen aan cognitieve begripsverandering.

Op vraag 11 wordt door de Spacegroep beter gescoord in de posttest. Dat kan komen door de representatie in deze vraag die erg doet denken aan het raketje uit het spel. Vraag 12 kent weliswaar dezelfde representatie, maar vertoont geen leerwinst. Vermoedelijk komt dat doordat in deze vraag naar versnelling gevraagd wordt; een te lastig concept voor de doelgroep. Vraag 16 heeft de context van de raket. Spacechallenge spelers vertonen een trendmatige toename.

Het lijkt er dus op dat het spelen van Space Challenge een betere voorbereiding biedt op de vragen van de FCI dan Unreal. Een verklaring zou kunnen zijn dat de leerlingen die Unreal speelden niet genoeg overzicht hadden. De game was heel immersive, maar bood weinig ruimte voor het systematisch beoordelen van bewegingen. Omdat het een First person shooter is, draait alles

continue, terwijl de speler geen helder beeld van zichzelf (het be-

wegende) object in de ruimte) krijgt. Vraag 4 gaat bijvoorbeeld over een cirkelbeweging, dat is heel lastig beoordelen vanuit 1st person perspectief.

Natuurkunde Motivatie

Op de vragenlijst Natuurkunde Motivatie wordt geen hogere score behaald op de posttest; wat er tussen de pretest en de posttest heeft plaatsgevonden heeft dus geen effect, gunstig of ongunstig, op de motivatie die leerlingen voor het vak Natuurkunde hebben. Ook wanneer de analyse wordt uitgevoerd over alle leerlingen, dus inclusief de leerlingen die afwezig waren bij de uitvoering van het practicum/de games, is er geen verandering te constateren in de motivatie voor het vak.

Die verandering is er echter wèl op een van de subschalen van de vragenlijst: Angst/gemak. Overall neemt de score op deze subschaal toe ($F=9,43$, $p= .003$). er is daarbij echter geen verschil tussen de groepen (zie tabel 5.10).

De schaal Angst/Gemak gaat over de inschatting door leerlingen van hun competentie t.a.v. het vak natuurkunde. Geconcludeerd kan dus worden dat leerlingen het gevoel hebben dat ze beter zijn geworden in het vak in de twee weken dat het onderzoek – en dus het onderwijs – hebben geduurd. Deze verbetering kan echter niet worden toegeschreven aan (een van de) interventies, omdat er geen interactie effect optreedt op deze schaal.

Motivatie voor natuurkunde hangt sterk samen met het cijfer voor natuurkunde ($r = .48$) en met de score op de fci op de pretest ($r=.32$). Leerlingen die beter zijn in natuurkunde scoren zijn dus meer gemotiveerd voor het vak. Dit geldt voor zowel de hele vragenlijst als voor de scores op de subschalen (zie tabel 5.11)

TABEL 5.10 GEMIDDELDEN EN STANDAARDDEVIATIES OP DE SUBSCHAAL ANGST/GEMAK VAN DE VRAGENLIJST NATUURKUNDE MOTIVATIE

	GROEP	GEMIDDELDE	SD	AANTAL ITEMS
ANGSTGEMAKPRE	practicum	34,54	.75620	26
	unreal	3.3500	1.21142	20
	space	3.3148	.83213	18
ANGSTGEMAK-POST	practicum	3.8462	.64278	26
	unreal	3.5583	1.11512	20
	space	3.3889	.85749	18

De schaal Angst/Gemak gaat over de inschatting door leerlingen van hun competentie t.a.v. het vak natuurkunde. Geconcludeerd kan dus worden dat leerlingen het gevoel hebben dat ze beter zijn geworden in het vak in de twee weken dat het onderzoek – en dus het onderwijs – hebben geduurd. Deze verbetering kan echter niet worden toegeschreven aan (een van de) interventies, omdat er geen interactie effect optreedt op deze schaal.

Motivatie voor natuurkunde hangt sterk samen met het cijfer voor natuurkunde ($r = .48$) en met de score op de fci op de pretest ($r=.32$). Leerlingen die beter zijn in natuurkunde scoren zijn dus meer gemotiveerd voor het vak. Dit geldt voor zowel de hele vragenlijst als voor de scores op de subschalen (zie tabel 5.11)

Flowtest

Zes leerlingen in de Space groep hadden hun naam doorgekrast. Voor de variantie analyse op de scores op de flow test maakt dat op zich niet uit, omdat de scores daarbij niet aan andere data hoeven te worden gekoppeld.

Er bleek geen verschil te zijn in de flowervaring van leerlingen die Space of Unreal hadden gespeeld. Dit gold zowel voor de gehele test, als voor de subschalen flow antecedenten en flow ervaring. Ook wanneer de vier leerlingen die voortijdig zijn weggegaan buiten de vergelijking worden gehouden zijn er geen verschillen.

TABEL 5.11 CORRELATIES TUSSEN DE SUBSCHALEN VAN DE VRAGENLIJST NATUURKUNDE MOTIVATIE EN DE FCI (PRETESTS)

		FCIPRETEST
ANGSTGEMAKPRE	R	,303**
	P	,004
	N	89
PLEZIERMOTPRE	R	,303**
	P	,004
	N	91
NUTMOTIVATIEPRETEST	R	,258*
	P	,014
	N	90
	P	
	N	92

TABEL 5.12 GEMIDDELDEN EN STANDAARDDEVIATIES VOOR DE FLOWTEST EN SUB-SCHALEN, VOOR BEIDE GAME

		N	GEMIDDELDE	SD
FLOW	unreal	21	3.1056	.76447
	space	19	3.2220	1.01758
FLOW ANTECEDENTEN	unreal	24	3.1708	.81105
	space	25	3.2920	1.25297
FLOWERVERING	unreal	22	3.0979	.82000
	space	21	3.0549	.93142
	Total	43	3.0769	.86590
FLOW23	unreal	23	2.52	1.201
	space	27	2.11	1.311

Evaluatievragen achteraf

Op twee evaluatievragen die achteraf gesteld werden was een trendmatig verschil te vinden tussen de beide game condities. De Unreal groep vond het spelen van de game wat interessanter ($F=3,59$, $p= .07$) en makkelijker ($F= 3,72$, $p= .06$) dan de Space groep. In deze analyse zijn

de vier leerlingen die voortijdig weggingen niet meegenomen.

Op de vraag of het spelen van de game leuk was, nuttig was en of ze er iets van hadden geleerd was er geen verschil tussen de groepen.

TABEL 5.13 GEMIDDELDEN EN STANDAARDDEVIATIES OP DE EVALUATIEVRAGEN ACHTERAF, VOOR DE BEIDE GAME CONDITIES

		N	GEMIDDELDE	SD
LEUK	UNREAL	21	3.48	1.250
	SPACE	19	2.95	1.393
INTERESSANT	UNREAL	21	3.05	1.244
	SPACE	18	2.33	1.085
NUTTIG	UNREAL	21	2.33	.913
	SPACE	18	2.22	1.114
MAKKELIJK	UNREAL	21	3.19	1.078
	SPACE	18	2.44	1.338
IETSGELEERD	UNREAL	19	2.11	1.049
	SPACE	19	2.11	1.197

Verschillen tussen jongens en meisjes

De verdeling van jongens en meisjes over de drie groepen zag er scheef uit, vooral in de gamecondities:

- PRACTICUMGROEP: 16 JONGENS, 17 MEISJES
- UNREALGROEP: 12 JONGENS, 18 MEISJES
- SPACEGROEP: 9 JONGENS, 21 MEISJES.

Deze verdeling bleek bij een X2 analyse echter niet significant. Toch zouden bijvoorbeeld verschillen tussen jongens en meisjes voor ruis in de resultaten kunnen zorgen.

Op de voortests bleken inderdaad geslachtsverschillen op te treden: jongens scoorden (trendmatig) hoger op de FCI dan meisjes ($F=3,53$, $p=.06$) en waren meer gemotiveerd voor het vak natuurkunde ($F=7,71$, $p=.007$).

TABEL 5.14 **VERSCHILLEN TUSSEN JONGENS EN MEISJES OP DE FCI EN MOTIVATIE PRETESTS**

	MV	N	GEMIDDELDE	SD
FCIPRETEST	man	34	5,50	2,799
	vrouw	42	4,52	1,685
MOTIVATIE PRETEST	unreal	31	3,7400	,73640
	space	39	3,2112	,83263

Nagegaan is of er in een herhaalde metingen analyse ook verschillen tussen jongens en meisjes optraden op de FCI en op de motivatie vragenlijst. Dit bleek niet het geval. Jongens scoren weliswaar gemiddeld hoger ($F=5,58$, $p=.02$) dan meisjes op de FCI, maar ze behalen niet meer leerwinst. Ook zijn jongens gemiddeld gemotiveerder ($F=8,72$, $p=.05$), maar ze raken niet meer gemotiveerd. De gemiddelden voor jongens en meisjes op de pre- en posttests voor de FCI en de motivatie vragenlijst staan in tabel

Wanneer geslacht als tweede factor wordt meegenomen in een herhaalde metingen analyse over de FCI treden er geen significante interacties op.

TABEL 5.15 **GEMIDDELDEN EN STANDAARDDEVIATIES OP DE FCI EN MOTIVATIE PRE- EN POSTTESTS VOOR JONGENS EN MEISJES**

		N	GEMIDDELDE	SD
FCIPRETEST	man	30	5,60	2,799
	vrouw	37	4,57	1,708
FCIPOSTTESTW	man	30	6,90	3,144
	vrouw	37	5,41	2,140

TABEL 5.15 GEMIDDELDEN EN STANDAARDDEVIATIES OP DE FCI EN MOTIVATIE PRE- EN POSTTESTS VOOR JONGENS EN MEISJES

		N	GEMIDDELDE	SD
MOTIVATIE PRETEST	man	26	3,7828	,79545
	vrouw	33	3,1872	,89058
MOTIVATIEPOST	man	26	3,8620	,65035
	vrouw	33	3,2228	,87060

Jongens en meisjes verschillen niet alleen op de FCI en de motivatie, maar ook op de flowtest. Jongens scoren daarop hoger dan meisjes, zowel op de totale test ($F = 4,17$, $p = .05$), als op de flow ervaring subschaal ($F = 7,75$, $p = .008$).

TABEL 5.16 VERSCHILLEN TUSSEN JONGENS EN MEISJES OP DE FLOW TEST

		N	GEMIDDELDE	SD
FLOWTOTAAL	man	16	3,5516	,65009
	vrouw	22	3,0198	,88044
FLOW ANTECEDENTEN	man	20	3,6400	,69312
	vrouw	24	3,2458	1,02447
FLOWERVARING	man	16	3,5673	,73668
	vrouw	24	2,8590	,81996

Ook op de evaluatievragen zijn verschillen tussen jongens en meisjes te vinden. Jongens vonden het spelen van een game interessanter ($F = 6,24$, $p = .02$), nuttiger ($F = 6,17$, $p = .02$) en ze hadden wat meer geleerd (trend, $F = 3,76$, $p = .06$) dan meisjes.

Wanneer in een variantie analyse over de evaluatievragen zowel de factor groep als de factor geslacht wordt meegenomen, zijn er echter geen interactie effecten te vinden.

TABEL 5.17 VERSCHILLEN TUSSEN JONGENS EN MEISJES OP DE EVALUATIEVRAGEN ACHTERAF

		N	GEMIDDELDE	SD
LEUK	MAN	18	3,56	1,423
	VROUW	22	2,95	1,214
INTERESSANT	MAN	17	3,24	1,300
	VROUW	22	2,32	,995

TABEL 5.17 VERSCHILLEN TUSSEN JONGENS EN MEISJES OP DE EVALUATIEVRAGEN ACHTERAF

		N	GEMIDDELDE	SD
NUTTIG	MAN	17	2,71	1,047
	VROUW	22	1,95	,844
MAKKELIJK	MAN	17	2,53	1,231
	VROUW	22	3,09	1,231
IETSGELEERD	MAN	1	2,50	1,211
	VROUW	22	1,82	,958

Effecten van thuis gamen

Jongens zijn vaker dan meisjes ook de leerlingen die thuis regelmatig games spelen ($C = .52, p .000$), maar de factor geslacht en thuisgamen overlappen niet totaal. Ook de verdeling wel/niet thuis gamen is scheef over de groepen; net als bij de factor geslacht is die scheve verdeling niet significant.

- practicum: 21 gamers versus 12 niet-gamers
- unreal: 18 gamers versus 11 niet gamers
- space: 12 gamers versus 18 niet gamers

Wanneer alleen naar de game groepen wordt gekeken is er wel een trendmatig verschil ($X^2 = 2,87, p = .09$).

Zoals te verwachten is, gezien het verband met geslacht, zijn vergelijkbare verschillen op de andere variabelen te vinden:

- op de FCI pretest ($F = 4,91, p = .03$)
- op de motivatie pretest ($F= 5,46, p = .02$)
- op de evaluatievragen: gamers vonden het interessanter ($F= 11,49, p = .002$), nuttiger ($F= 12,45, p=.001$) en hebben ook meer geleerd ($F=6,46, p = .02$) dan leerlingen die thuis niet gamen
- flow: gamers scoorden hoger op de flowtest in totaal ($F = 4,45, p = .04$) en op de subschaal flowervaring ($F= 7,05, p = .01$) dan leerlingen die thuis niet gamen

Wanneer in een variantie analyse over de evaluatievragen zowel de factor groep als de factor thuisgamen wordt meegenomen, blijkt er op ‘leuk’ een significant interactie effect ($F=6.42, p = .05$). Bij de Space groep hebben de gamers het duidelijk leuker gevonden dan de niet-gamers; bij de Unreal groep gaat het geringe verschil de andere kant op.

TABEL 5.18 INTERACTIE EFFECT OP DE EVALUATIEVRAAG 'LEUK' MET GROEP EN THUISGAMEN ALS FACTOREN.

LEUK	unreal	JA	3,300	,386
		NEE	3,444	,406
	space	JA	4,000	,461
		NEE	2,455	,368

TABEL 5.19 GEMIDDELDEN EN STANDAARDDEVIATIES VOOR LEERLINGEN DIE THUIS GAMEN EN LEERLINGEN DIE THUIS NIET GAMEN OP DE PRETESTS, DE FLOW TESTS EN DE EVALUATIEVRAGEN ACHTERAF

		N	GEMIDDELDE	SD
LEUK	JA	20	3,55	1,276
	NEE	20	2,90	1,334
INTERESSANT	JA	19	3,32	1,157
	NEE	20	2,15	,988
NUTTIG	JA	19	2,79	,918
	NEE	20	1,80	,834
MAKKELIJK	JA	19	2,63	1,212
	NEE	20	3,05	1,276
IETSGELEERD	JA	18	2,56	1,199
	NEE	20	1,70	,865
FCIPRETEST	JA	39	5,51	2,635
	NEE	37	4,38	1,705
MOTIVATIE PRETEST	JA	36	3,6634	,71557
	NEE	34	3,2145	,88761
FLOWTOTAAL	JA	22	3,4743	,72318
	NEE	16	2,9266	,87473
FLOW ERVARING	JA	22	3,4441	,80044
	NEE	18	2,7735	,78645
FLOWANTE- CEDENTEN	JA	25	3,6000	,77728
	NEE	19	3,1947	1,01897

Net als bij de verschillen voor geslacht halen thuisgamers een hogere gemiddelde score op de FCI ($F= 4,57$, $p = .04$) en op de motivatie vragenlijst ($F= 6,63$, $p = .01$), maar ze halen geen grotere leerwinst. Ook hier levert het geen interacties op wanneer de factor thuisgamen als tweede factor wordt meegenomen.

TABEL 5.20 **GEMIDDELDEN EN STANDAARDDEVIATIES OP DE FCI EN DE MOTIVATIE VRAGENLIJST VOOR LEERLINGEN DIE THUISGAMEN EN LEERLINGEN DIE DAT NIET DOEN.**

	THUIS- GAMES	N	GEMIDDELDE	SD
FCIPRETEST	JA	35	5,60	2,614
	NEE	32	4,41	1,739
FCIPOSTTEST	JA	35	6,60	3,051
	NEE	32	5,50	2,214
MOTIVATIE PRETEST	JA	30	3,6941	,77661
	NEE	39	3,1968	,94906
MOTIVATIE- POST	JA	30	3,7941	,65078
	NEE	29	3,2049	,91354

6. Discussie

Onderzoeksvragen

A

Verwerven leerlingen die een game spelen het inzicht dat beweging met weerstand te maken heeft beter dan leerlingen die dit inzicht via traditionele natuurkundelessen krijgen aangeboden?

Verwerven leerlingen dit inzicht beter via een voor dit doel geselecteerde commerciële game (COTS) dan via een op dit doel gerichte serious game?

B

C

Is er een verschil in flow bij het spelen van de COTS game en van de serious game?

Zoja:

Is de mate van flow die leerlingen ervaren bij het spelen van beide type games een verklarende factor voor verschillen in effecten?

D

6.1 Resultaten: **onderzoeksvragen**

Onderzoeksvraag 1:

Deze onderzoeksvraag is in dit onderzoek niet bevestigd. Er is geen verschil in leerwinst tussen enerzijds de practicumgroep, anderzijds de beide game groepen samen genomen. Ook is er geen verschil tussen de practicumgroep en de beide game condities afzonderlijk. Leerlingen verwerven het inzicht dat beweging met weerstand te maken dus net zo goed met een goed gekozen practicum als door het spelen van een game waarin dit inzicht een belangrijke rol speelt.

Onderzoeksvraag 2:

Deze onderzoeksvraag is in dit onderzoek verworpen: leerlingen verwerven dit inzicht niet beter via een COTS game dan via een op dit doel gerichte serious game. Integendeel: de leerlingen die de educatieve game spelen (Space Challenge) behalen een grotere leerwinst dan de leerlingen die de COTS game spelen (Unreal). De Space groep heeft bij de pretest een aanzienlijk lagere score dan de andere twee groepen, maar haalt dit verschil bij de posttest grotendeels in.

De Unreal groep start relatief hoog en blijft op dat niveau – relatief, want deze leerlingen behalen overall op de posttest gemiddeld niet hoger dan een derde van de mogelijke score.

De verschillen tussen de beide groepen zijn terug te vinden op de subset vragen die direct met de wet-

ten van Newton te maken hebben, en niet met de andere twee subsets, Vectoren en Kinematica.

Onderzoeksvragen 3&4:

Er is geen verschil in ervaren flow gemeten tussen leerlingen die de COTS game speelden en leerlingen die de educatieve game speelden. De mate van flow die leerlingen ervaren tijdens het spelen kan dus de verschillen in leerwinst tussen de beide typen games niet verklaren.

6.2 Discussie

Games in natuurkunde onderwijs?

In dit onderzoek is dus niet gebleken dat games een betere manier zijn om het inzicht dat beweging met weerstand te maken heeft te verwerven dan een traditionele lesmethode als een practicum. Kan hieruit worden geconcludeerd dat het dus geen zin heeft games in natuurkunde onderwijs te gebruiken?

Die conclusie lijkt niet terecht. Een onderzoek is een momentopname en voor onderzoek binnen het onderwijs geldt dat er veel ruis optreedt, hoezeer onderzoekers ook hun best doen de meting zo zuiver mogelijk te houden. Bovendien zijn er beperkingen voor de zuiverheid van het design die vooral te maken hebben met het feit dat een school onderwijs moet geven en dat een opzet waarbij leerlingen de kans lopen iets niét te leren door de methode – terecht – niet mogelijk is. Ook het practicum is zo ingericht

dat de principes die een rol spelen maximaal aan bod komen.

Een andere reden waarom deze conclusie niet terecht lijkt, is dat een practicum als dit, dat specifiek gericht is op de concepten waar het in de games ook om gaat, steeds weer opnieuw zou moeten worden ingericht, terwijl een eenmaal ontworpen game zo kan worden hergebruikt. Om te reflecteren op de geleerde concepten zou bovendien ook makkelijk weer kunnen worden teruggegrepen op de game, terwijl dat met een practicum praktisch moeilijk te realiseren is. Gezien de vergelijkbaarheid van leeropbrengst heeft de inzet van een game hier dus ook zeker organisatorische voordelen.

Bij het onderzoek dat de ontwikkeling van de educatieve game, Martijn Hoevenaar, eerder uitvoerde bleken er wel degelijk verschillen op te treden ten gunste van de leerlingen die Space Challenge speelden. Het onderzoek van Hoevenaar verschilde in een aantal opzichten van ons onderzoek.

In de eerste plaats ging het om vierde klassers, terwijl het hier om derdeklassers ging, die nog maar weinig natuurkunde kennis hebben opgedaan. Die keus had vooral te maken met het idee dat natuurkundige misconcepties beter bij de aanvang van het natuurkunde onderwijs 'uit de weg geruimd' kunnen worden, en niet pas wanneer leerlingen daar mogelijk al een tijd door gehinderd worden.

Een ander verschil betrof de setting. UniC is een school waarbij het 'traditionele onderwijs' al niet zo

traditioneel is. Leerlingen werken zelfstandig en in groepen aan hun opdrachten en daarin past de traditionele setting van een docent die voor de klas iets uitlegt niet. Dat was wel het geval in de school waar het onderzoek van Hoevenaar werd uitgevoerd. Het is goed mogelijk dat het gebruik van games in het onderwijs beter tot zijn recht komt in een meer traditionele setting.

Het belang van nabesprekingen bij het gebruik van games in het onderwijs komt overal in de literatuur terug. In het onderzoek van Hoevenaar was die nabespreking goed mogelijk, op een school als UniC is dat veel lastiger, zeker gezien de opzet waar we voor gekozen hadden. Deze opzet had in de eerste plaats te maken met een derde belangrijk verschil met het onderzoek van Hoevenaar: de uitvoerende docent. In het onderzoek van Hoevenaar was de uitvoerende docent in de experimentele condities (hijzelf) een andere dan de docent in de controle conditie. In ons onderzoek was de uitvoerend docent – de docent/onderzoeker – in alle drie de condities dezelfde. Dit had in de eerste plaats praktische redenen: deze docent verzorgde het onderwijs voor alle derde klassen gedurende het gehele thema. Als verschillende docenten de drie condities hadden moeten begeleiden, dan hadden er dus voor minimaal twee weken (het thema zelf duurde langer) twee extra docenten moeten worden ingezet. Dat is binnen een schoolrooster natuurlijk niet te realiseren. Bovendien is het de vraag of het wel wenselijk zou zijn. Wanneer

één en dezelfde docent wordt ingezet bij alle drie de condities, is de factor 'docent' – waarvan altijd en overal blijkt dat die voor de grootste variatie zorgt bij onderzoek naar effecten van leermiddelen – tenminste zoveel mogelijk constant. De verwachtingen die de docent heeft van de uitkomsten kunnen dan natuurlijk nog wel degelijk voor ruis zorgen.

De verwachtingen van de docent in ons onderzoek bleken bij het interview achteraf (gehouden voordat de resultaten bekend waren) redelijk neutraal. De docent is zeker overtuigd van het nut van games in het onderwijs, maar niet per se omdat daarvan een grotere leerwinst te verwachten zou zijn. Hij zou al tevreden zijn als de games het even goed zouden doen als het practicum, gaf hij aan. Variatie in leervormen vindt hij belangrijk, omdat daarmee kan worden ingespeeld op verschillen tussen leerlingen.

Zowel om praktische als methodische redenen was de docent dus in alle drie de condities dezelfde. Zoals al eerder aangegeven, vond de 'treatment' voor alle drie de groepen op hetzelfde moment plaats. Dit had vooral methodische redenen. Op deze manier was de plaats in het curriculum vergelijkbaar voor de drie groepen en was er geen kans dat leerlingen zouden vertellen wat ze gedaan hadden tegen leerlingen die nog geen game/practicum hadden gedaan.

Er was dus sprake van een dilemma: wat zou zwaarder wegen, de ruis die zou kunnen ontstaan wanneer leerlingen op verschillende tijdstip-

pen het practicum zouden uitvoeren/de games zouden spelen? Of het feit dat voor dat onderdeel van het onderzoek de docent niét dezelfde kon zijn, maar voor de rest van het onderwijs in de twee weken wel? We kozen voor het eerste: de groepen werden tijdens het praktische gedeelte dus begeleid door drie verschillende begeleiders: de docent/onderzoeker, een onderzoeker van het Akademieteam en Martijn Hoevenaar.

Om te voorkomen dat er docentverschillen op zouden treden kregen de begeleiders de instructie om géén inhoudelijke informatie/hulp te geven. Ook werd er geen nabespreking gehouden in de drie groepen, direct na het practicum/de games; ook daarmee zouden anders docentverschillen kunnen ontstaan.

Om toch de koppeling naar de praktijk te kunnen maken, werden er verwerkingsopdrachten gebruikt. Deze opdrachten – voor elke conditie een eigen versie – werden enige dagen na het practicum/de game zelfstandig door leerlingen gemaakt. Wanneer de leerlingen de opdrachten hadden afgerond, konden ze de goede antwoorden (op internet) vergelijken met hun eigen antwoorden. De opdrachten werden door de hele groep tegelijk in één ruimte gemaakt, onder begeleiding van de docent/onderzoeker. Gezien het grote aantal leerlingen (90) hield de docent zich vooral bezig met organisatie en orde. Er waren twee observanten aanwezig die qua expertise inhoudelijke hulp hadden kunnen bieden waar nodig, maar die waren geïnstrueerd dat niet te doen – ook weer om docentver-

schillen te voorkomen.

De opzet was er dus op gericht het onderzoek zo zuiver mogelijk te houden, maar had als consequentie – zeker binnen de setting van een school met een onderwijsvisie als deze – dat de nodige reflectie achteraf erdoor bemoeilijkt werd.

Op basis van de interviews realiseerden wij ons dat ten minste één van de factoren die wij hiermee onder controle wilden houden, waarschijnlijk veel minder belangrijk was dan wij dachten. Leerlingen zouden waarschijnlijk helemaal niet zo geneigd geweest zijn (voortijdig) met elkaar over het onderzoek, en in het bijzonder de games, te praten. Zoals de docentonderzoeker waarschijnlijk zeer terecht opmerkte is de school voor leerlingen de school, vol dingen die je doet omdat dat nu eenmaal moet en daarmee iets volstrekt anders dan wat je in je vrije tijd doet. Zo'n onderzoek hoort daar in de perceptie van leerlingen gewoon bij. Leerlingen vinden een onderzoek dus veel minder belangrijk dan de onderzoekers die het uitvoeren. Wat de leerlingen bij de interviews achteraf zeiden, klopt wel met dit beeld.

De verschillen tussen dit onderzoek en het onderzoek van Hoevenaar die de revue gepasseerd zijn zouden stuk voor stuk hebben kunnen zorgen voor de verschillen in uitkomsten. Nader onderzoek zou moeten uitwijzen welk(e) verschil(len) hierbij de doorslag gaf/gaven. Naar ons idee is de grootste kanshebber daarbij de gelegenheid voor reflectie. Het zou heel goed kunnen dat games beter

passen in een onderwijs systeem waarin reflectie met de hele groep na het spelen van de game makkelijk kan plaatsvinden. Om de games 'een faire kans' te geven in de vergelijking met het practicum zou het daarom aanbeveling verdienen om het onderzoek nogmaals uit te voeren en daarbij enige temporale ongelijkheid en de mogelijkheid dat leerlingen voortijdig iets horen over de andere condities voor lief te nemen. De docent - één docent – zou dan de drie condities tijdens het practicum/de games kunnen begeleiden, daarbij direct in kunnen spelen op inhoudelijke vragen en direct na de games/het practicum klassikaal de terugkoppeling van de theorie naar de praktijk kunnen maken. In een dergelijke opzet is de mogelijkheid tot reflectie vergelijkbaar met het onderzoek van Hoevenaar; de verschillen zitten dan in de mogelijke invloed van de docent en het niveau van de leerlingen (derde versus vierde). Het laatste verschil is op het UniC moeilijk op te lossen, omdat er veel minder vierde dan derdeklassers zijn die natuurkunde hebben. De groepen waartussen vergeleken moet worden, worden dan veel te klein.

Een educatieve game of een COTs game?

In dit onderzoek deed de educatieve game het duidelijk beter dan de COTS game: leerlingen in de Spacegroep behaalden een significant grotere leerwinst. Daaraan ligt geen verschil in ervaren flow ten grondslag: de game groepen verschilden niet

op hun scores op deze vragenlijst. Waar is het verschil dan wel door te verklaren?

Een factor die niet in de onderzoeksvragen is opgenomen, maar wel is gemeten is de motivatie voor het vak natuurkunde. Deze factor is zowel vooraf als achteraf gemeten, om na te gaan of de motivatie voor natuurkunde toeneemt door/na de verschillende treatments, maar ook als mogelijke verklarende variabele. Wanneer leerlingen die een grotere leerwinst halen ook beter worden in natuurkunde, er meer plezier in hebben en/of de relevantie van het vak meer inzien, zou dat die grotere leerwinst (mede) kunnen verklaren. Omgekeerd zou ook die grotere motivatie kunnen voortkomen uit het feit dat zij er meer van snappen. Er is inderdaad een sterke samenhang tussen de score op de inhoudelijke (pre)test en de motivatie score, net als tussen de motivatiescore en het natuurkunde cijfer. Leerlingen die beter zijn in natuurkunde zijn dus meer gemotiveerd. Ook blijkt de motivatie aan het eind van week 2 voor de hele groep leerlingen toegenomen op de subschaal Angst/Gemak: leerlingen hebben het gevoel dat ze beter zijn geworden in natuurkunde in de twee weken dat het onderzoek heeft geduurd. Het onderwijs heeft dus effect gehad, maar welke rol spelen het gamen/practicum hierin?

Er is geen verschil tussen de groepen qua toename in ervaren competentie in het vak. De grotere leerwinst in de Space groep heeft dus niet te maken met een grotere motivatie. Waarmee dan wel?

Mede gezien het feit dat er al eerder effecten van de educatieve game gevonden zijn, zal minstens een deel van de grotere winst voor de Space groep met de game zelf te maken hebben. De game ligt dicht bij waar het in deze twee onderwijsweken om gaat: snelheid, krachten, beweging en het is ook duidelijk waarneembaar dat dat zo is. De game is er op gemaakt. De winst is ook te vinden op de vragen die direct met de eerste twee wetten van Newton te maken hebben.

Als onze veronderstelling klopt dat te weinig expliciete reflectiemogelijkheid in de opzet er de oorzaak van is dat gamen niet beter werkt dan meer traditioneel onderwijs, dan zal vooral Unreal daarvan te lijden hebben gehad. De relatie naar de theorie wijst zich in het geval van Space Challenge veel meer vanzelf; bij Unreal zal die relatie veel meer aangestuurd moeten worden in een nabespreking.

Dat wil niet zeggen dat we op grond van deze resultaten kunnen concluderen dat de Unreal groep het bij een replicatie met meer aangestuurde reflectiemogelijkheid het vanzelf dus beter zal doen. Eén van de hypothesen in dit onderzoek was ook dat leerlingen die een COTS game spelen, daarbij meer flow ervaren. Flow zou tot betere leereffecten kunnen leiden, of juist tot lagere, omdat flow reflectie in de weg zit. De beide game condities verschilden echter niet van elkaar in de ervaren flow. Wel bleken er enige lichte verschillen te zijn tussen de Space en Unreal groep bij de evaluatievragen achteraf. De verschillen

waren zeker niet groot ($.05 < p < .10$), maar ze waren in het voordeel van de Unreal groep. Deze leerlingen hadden het spelen van de game interessanter en makkelijker gevonden dan de leerlingen die de educatieve game speelden. Dit verschil is niet toe te schrijven aan een verschil tussen jongens en meisjes, of tussen gamers en niet gamers. Wanneer deze factoren mee worden genomen in een variantie analyse blijkt er alleen een interactie effect tussen thuis gamen en 'een game tijdens de les leuk vinden', in het voordeel van de space groep. Op 'interessant' en 'makkelijk' zijn geen interactie effecten te vinden.

Behalve met de game zelf, kan het verschil in leerwinst ook te maken hebben met de samenstelling van de groepen. De groepen waren niet speciaal samengesteld met het oog op dit thema, maar met het oog op heel iets anders: een musical. De space groep – de groep die zo'n lage begin-score had op de FCI – bestond uit leerlingen die hierbij geen specifieke keuze voor of acteren, of bouwen hadden gemaakt. Voorzichtig gezegd zou geconcludeerd kunnen worden dat deze groep bestond uit de overall wat minder gemotiveerde leerlingen. Dat verklaart op zich echter niet waarom deze leerlingen dan zo'n relatief grote leerwinst zouden behalen tov de leerlingen in de Unreal groep. Vooraf is wel gecontroleerd of er geen verschil in niveau qua natuurkunde bestond; er bleek geen verschil tussen de groepen in het cijfer voor natuurkunde. Wel bleek achteraf sprake van een scheve verdeling tussen jongens

en meisjes, zij het niet significant scheef. In de practicumgroep waren vrijwel evenveel jongens als meisjes, in de beide game groepen waren de meisjes in de meerderheid, bij de space groep zelfs veruit in de meerderheid. Jongens bleken wat hoger te scoren op de FCI pretest, duidelijk meer gemotiveerd te zijn voor het vak natuurkunde, significant meer flow te ervaren en het spelen van een game in de les interessanter, nuttiger en leuker te hebben gevonden en naar hun mening iets meer te hebben geleerd dan meisjes. Jongens behalen echter geen grotere leerwinst dan meisjes en wanneer het geslacht als tweede factor wordt meegenomen in een herhaalde metingen analyse treden er geen interactie effecten op. Dit verschil tussen jongens en meisjes kan dus geen verklaring zijn voor de grotere leerwinst van de Space groep, maar mogelijk wel voor de lagere uitgangspositie.

Hetzelfde geldt voor de factor thuisgamen, die deels overlapt met geslacht. Jongens gamen zoals bekend thuis vaker dan meisjes. In de space groep zitten meer niet-gamers dan gamers, in de Unreal groep is dat omgekeerd. De verschillen tussen gamers en niet gamers zijn vergelijkbaar met de verschillen tussen jongens en meisjes, maar iets meer uitgesproken. Ook hier treden geen interactie effecten op wanneer deze factor wordt meegenomen in de herhaalde metingen analyse. Dat er in de Space groep zoveel meisjes en dus (grotendeels) niet-gamers zaten, kan wel hun lage uitgangspositie verklaren, maar niet hun grotere leerwinst.

Dit gegeven, en het gegeven dat de ervaren flow, de motivatie voor natuurkunde en de evaluatievragen achteraf evenmin een verklaring kunnen bieden maakt de veronderstelling dat de aard van de game hier vooral een rol heeft gespeeld des te waarschijnlijker.

Een laatste alternatieve mogelijkheid is dat we hier te maken hebben met een plafondeffect: dat 6 van de 19 vragen goed het maximaal haalbare is voor de derdeklassers in dit onderzoek. Unreal spelers 'zitten dan gelijk al tegen hun plafond', terwijl er voor de andere groepen nog winst te behalen is. Ook wat dit betreft zou een replicatie onderzoek dat meer is toegespitst op reflectiemogelijkheden waarmee de scores mogelijk positief kunnen worden beïnvloed, uitsluitel kunnen geven.

Literatuur

- ALTING, A. (2003). NUT, VERTROUWEN, TOEGANKELIJKHEID. WAT DOCENTEN KUNNEN DOEN OPDAT MEER MEISJES NATUURKUNDE GAAN KIEZEN
- B.WHITE (1984), DESIGNING COMPUTERGAMES TO HELP PHYSICS STUDENTS UNDERSTAND NEWTONS LAWS OF MOTION ,COGNITION AND INSTRUCTION, 1984, 1(1) 69-108
- EGENFELDT-NIELSEN, S. (2006) OVERVIEW OF RESEARCH ON THE EDUCATIONAL USE OF VIDEO GAMES. DIGITAL KOMPETANSE, 3-2006, 1, 184 – 213.
- GRAAF, C. (2010). OPBRENGSTEN VAN DIGITAAL LEERMATERIAAL: NLT
- HESTENES, D., WELLS, M. & SWACKHAMER, G. (1992). FORCE CONCEPT INVENTORY. THE PHYSICS TEACHER. VOL. 30, 1992, 141 – 158.
- HUIZENGA, J., ADMIRAAL, W. & AKKERMAN, S. (2007.). ONDERZOEKSRAPPORT FQ 1550 x 10. UNIVERSITEIT VAN AMSTERDAM & UNIVERSITEIT UTRECHT.
- HUL, F VAN 'T, W. VAN JOOLINGEN AND P. LIJNSE (1989) DE TRAGE INBREKER, TIJDSCHRIFT VOOR DIDACTIEK DER -WETENSCHAPPEN 3:172-191
- RIEBER, L. P & NOAH, P (2008). GAMES, SIMULATIONS, AND VISUAL METAPHORS IN EDUCATION: ANTAGONISM BETWEEN ENJOYMENT AND LEARNING EDUCATIONAL MEDIA INTERNATIONAL, 45, 77-92.
- JONG, T DE, W.R. VAN JOOLINGEN ,(1998) SCIENTIFIC DISCOVERY LEARNING WITH COMPUTER SIMULATIONS OF CONCEPTUAL DOMAINS. REVIEW OF EDUCATIONAL RESEARCH 68:22, 179-201
- KIILI, K. (2006) EVALUATIONS OF AN EXPERIENTIAL GAMING MODEL. HUMAN TECHNOLOGY, VOLUME 2 (2), OCTOBER 2006, 187–201
- KIILI, K. (2005). DIGITAL GAME-BASED LEARNING: TOWARDS AN EXPERIENTIAL GAMING MODEL. THE INTERNET AND HIGHER EDUCATION, 8 (2005), 13–24.
- KOLB, DAVID A. 1984. EXPERIENTIAL LEARNING: EXPERIENCE AS THE SOURCE OF LEARNING AND DEVELOPMENT. PRENTICE-HALL, INC., ENGLEWOOD CLIFFS, N.J.
- KOOPS, M. C. AND HOEVENAAR, M. (2010A). CONCEPTUAL CHANGE BY SIMULATION GAMES. MANUSCRIPT SUBMITTED FOR PUBLICATION
- KOOPS, M.C. (2010). THE SERIOUS GAMING LEMNISCATE MODEL FOR
- ACQUIRING KNOWLEDGE THROUGH SIMULATION GAMES. PROCEEDINGS OF THE 41TH ANNUAL CONFERENCE FOR THE INTERNATIONAL SIMULATION AND GAMING ASSOCIATION, JULY 2010.
- LEEMKUIL, H. & DE JONG, T. (IN PRESS).. CAN COMPUTER GAMES BE USED FOR INSTRUCTION? INSTRUCTIONAL SUPPORT IN GAMES. IN S. TOBIAS & D. FLETCHER (EDS) ROUTLEDGE PRESS
- LEEMKUIL, H. (2005) IS IT ALL IN THE GAME? ENSCHEDE: UNIVERSITY OF TWENTE.
- LEUTNER, D. (1993). GUIDED DISCOVERY LEARNING WITH COMPUTER-BASED SIMULATION GAMES: EFFECTS OF ADAPTIVE AND NON-ADAPTIVE INSTRUCTIONAL SUPPORT. LEARNING AND INSTRUCTION, 3, 113-132.
- MEDENI, T. & MEDENI, . T. AN EXPERIENCE-BASED TACIT-EXPLICIT KNOWLEDGE INTERACTION MODEL OF ACTION & LEARNING, IN NITE 2004, OCTOBER, TURKEY
- MEIJER, J. , E.A. (2008). LEREN MET MEER EFFECT; RAPPORTAGE VAN HET ONDERZOEK
- RANDEL, J.M., MORRIS, B.A., WETZEL, C.D. & WHITEHILL, B.V. (1992) THE EFFECTIVENESS OF GAMES FOR EDUCATIONAL PURPOSES: A REVIEW OF THE RESEARCH. SIMULATION AND GAMING, 25, 261 – 276.
- RIEBER, L. P & NOAH, P (2008). GAMES, SIMULATIONS, AND VISUAL METAPHORS IN EDUCATION: ANTAGONISM BETWEEN ENJOYMENT AND LEARNING EDUCATIONAL MEDIA INTERNATIONAL, 45, 77-92.
- VERHEUL, I. & DIJK, W. VAN (2009). EFFECTIVITEIT VAN EEN COTS GAME IN HET MBO: OBLIVION. [HTTP://ONDERZOEK.KENNISNET.NL/ONDERZOEKEN/RENDEMENT/COTS](http://onderzoek.kennisnet.nl/onderzoeken/rendement/cots)
- Vos. N. W. (2009). GAMENDERWIJS. ONDERZOEK NAAR DE EDUCATIEVE WAARDE VAN GAMES OVER SPREEKWOORDEN.
- WHITE, B. (1984). DESIGNING COMPUTER GAMES TO HELP PHYSICS STUDENTS UNDERSTAND NEWTON'S LAWS OF MOTION. COGNITION AND INSTRUCTION, 1(1), 69,108
- WOLFE, J. (1997) THE EFFECTIVENESS OF BUSINESS GAMES IN STRATEGIC MANAGEMENT COURSE WORK. SIMULATION AND GAMING, 28, 4, 360.