

Gebruikers nudgen naar effectieve oefeningen

Joran DE BRAEKELEER

Promotor: Prof. dr. Katrien Verbert
Begeleider en assessor: Jeroen Ooge
Assessor: Nathan Cornille

Proefschrift ingediend tot het behalen van
de graad van Master of Science
in toegepaste informatica (multimedia)

Academiejaar 2022-2023

Voorwoord

Deze thesis was veruit de grootste uitdaging van mijn studieloopbaan, maar als afsluiting van mijn master vind ik dat wel terecht. Meermaals dacht ik mijn uitweg uit het bos te vinden, om daarna te botsen tegen de zoveelste boom. Daarom mag ik mezelf heel gelukkig prijzen met de mensen die me overeind hielpen, degenen wiens stemmen vanboven uit het bladerdek weergalmden en me leidden in de juiste richting.

Eerst en vooral wil ik mijn geweldige mentor Jeroen Ooge bedanken. Hij was er altijd om mij te steunen, door uitgebreide opbouwende kritiek, door mee na te denken over alternatieven, of gewoon door me aan te moedigen in tijden van nood. Hij steunde me omdat het zijn rol was als mentor, maar ook omdat hij echt geloofde in mij en het werk dat ik leverde.

Daarnaast wil ik Prof. dr. Katrien Verbert bedanken om dit interessante onderwerp aan te bieden en de hele Augment-groep voor hun nuttige feedback tijdens tussentijdse presentaties.

Last but not least wil ik mijn goede vriend Kobe Loppe bedanken voor zijn uitgebreide technische assistentie om dit onderzoek mogelijk te maken. Hij spendeerde talloze uren om mij bij te scholen in webapplicaties en databanken en wilde werkelijk dat het onderzoek zo goed mogelijk verliep. Zonder hem had ik lang zitten sukkelen.

Bedankt aan alle genoemden en alle andere vrienden en familieleden die me steunden tijdens de ontwikkeling van deze thesis, jullie bijdragen zal ik niet vergeten.

Samenvatting

Het is geen geheim dat wiskunde kampt met een imagoprobleem: het aantal studenten dat beschikt over universitaire wiskundekennis kent al jaren een dalende trend en dat is in strijd met het groeiende belang van het STEM-domein van *science, technology, engineering en mathematics*. Aan de grondslag van dat probleem ligt een tekort aan intrinsieke motivatie van leerlingen om wiskunde te leren.

Dit onderzoek focust op twee manieren om leerlingen te motiveren om wiskunde in te oefeningen. Enerzijds onderzoekt het hoe een onderwijssysteem gamificatie kan toepassen om leerlingen al spelend te laten leren. Personalisatie is daarbij belangrijk, aangezien gamificatie enkel werkt als spelelementen spreken tot de voorkeuren van leerlingen en geen twee leerlingen gelijk zijn. In een onderwijstoepassing is personalisatie evenzeer van belang, rekening houdend met het leerproces en de noden van individuele gebruikers. Oefeningen moeten uitdagend zijn, maar niet te moeilijk gegeven de vaardigheden van gebruikers. Door oefeningen te maken die aan de grens van hun kunnen liggen, leren gebruikers het meest effectief en efficiënt.

De centrale vraag van dit onderzoek is of gamificatie gebruikers kan motiveren om voor die doeltreffende oefeningen te kiezen, ondanks de hogere moeilijkheidsgraad. We bouwden speciaal voor deze studie de digitale onderwijsapplicatie Alchemath. De applicatie genereert oefeningen over vergelijkingen *on the fly* en maakt gebruik van vier gamificatie-elementen die we bewust kozen voor hun relevantie in de literatuur en de opzet van dit onderzoek.

Onze resultaten tonen aan dat hogere beloningen voor uitdagendere oefeningen ervoor zorgen dat leerlingen meer geneigd zijn om uitdagendere oefeningen te maken. Hoewel we een vermoeden hebben, kunnen we niet bevestigen dat uitdagendere oefeningen maken daarom ook de leerefficiëntie van gebruikers verhoogt.

Inhoudsopgave

1	Inleiding	1
2	Literatuurstudie	2
2.1	Games en gamificatie	2
2.1.1	Games en serious games	2
2.1.2	Gamificatie	3
2.1.3	Kritiek op gamificatie	5
2.2	Motivatie	6
2.2.1	Self-determination theory	6
2.2.2	Beloningen	6
2.3	Nudging	8
2.4	Gepersonaliseerde gamificatie	9
2.4.1	Big Five	9
2.4.2	BrainHex	9
2.4.3	Hexad	10
2.4.4	Kritiek op modellen	12
2.5	Gepersonaliseerd leren	13
2.5.1	Complexiteit en moeilijkheid	13
2.5.2	Optioneel leren	14
2.5.3	Elo-rating	14
2.5.4	Zone of proximal development	15
2.6	Open problemen in de literatuur	16
3	Methode	18
3.1	Onderzoeksvragen	18
3.2	Alchemath: gemotiveerd oefenen op maat	19
3.2.1	Automatisch gegenereerde oefeningen	19
3.2.2	Feedback bij oplossen	20
3.2.3	Gamificatie	21
3.3	Oefeningen aanbevelen	24
3.4	Deelnemers en verloop van de studie	25
3.4.1	Rekrutering	25
3.4.2	Gerandomiseerd onderzoek met controlegroep	26
3.4.3	Verloop	26
3.4.4	Loggegevens	27
3.4.5	Overige gegevens	28

3.5	Hypotheses en analyse	28
4	Ontwikkelingsproces	31
4.1	Papieren prototype	31
4.2	Digitale prototype	32
4.2.1	Eerste think-aloud	32
4.2.2	Tweede think-aloud	32
4.3	Technische implementatie	33
4.3.1	Oefeningen parsen	33
4.3.2	Verzamelen van data	34
4.3.3	Publicatie	34
5	Resultaten	36
5.1	Screening van testgroepen	36
5.2	Voorkeur voor spelelementen	37
5.3	Testgroepen en oefeningen kiezen	39
5.4	Testgroepen en leerefficiëntie	42
5.5	Testgroepen en zelf-gerapporteerde competentie	42
5.6	Feedback van leerkrachten	42
6	Discussie	45
6.1	Voorkeur voor spelelementen	45
6.2	Testgroepen en oefeningen kiezen	46
6.3	Testgroepen en leerefficiëntie	46
6.4	Testgroepen en zelf-gerapporteerde competentie	47
6.5	Beperkingen van het onderzoek en toekomstig werk	47
7	Conclusie	49
7.1	Resultaten voor de onderzoeksvragen	49
7.2	Bijdragen aan het mens-machine-interactiedomein	49
7.3	Bijdragen aan het onderwijsdomein	50
A	Vragenlijsten	51
B	Details proof of concept	53
C	Overige resultaten	57
D	Informatie aan betrokkenen en toestemmingsformulieren	58
	Bibliografie	69

Hoofdstuk 1

Inleiding

Motivatie is de drijvende factor in mensen die gedrag op gang brengt en aanstuurt. Mensen vinden motivatie in zichzelf voor talloze doeleinden, omdat ze persoonlijke interesse of plezier aanwakkeren. Anderzijds zijn er ook veel activiteiten waar mensen tegenop kijken zoals afgeraken van slechte gewoontes, terwijl ze er wel baat bij hebben. Zulke activiteiten vereisen aan externe vorm van motivatie. Gamificatie is één manier om externe motivatie te genereren en is de laatste jaren toegenomen in populariteit. Gamificatie neemt elementen uit spellen of games en plaatst ze in een niet-game context om een activiteit speelser en leuker te maken.

Deze thesis focust op gamificatie in onderwijs, specifiek in de wiskunde. Het is geen geheim dat wiskunde kampt met een imago probleem en gamificatie zou wel eens de oplossing kunnen zijn om een ander licht op het vak te laten schijnen. In een onderwijstoepassing is het bovendien van belang rekening te houden met het leerproces en de noden van individuele gebruikers. Oefeningen moeten uitdagend zijn, maar niet te moeilijk gegeven de vaardigheden van gebruikers. Door oefeningen te maken die aan de grens van hun kunnen liggen, leren gebruikers het meest effectief en efficiënt.

De centrale onderzoeksvraag van deze thesis is of gamificatie gebruikers kan motiveren om voor die doeltreffende oefeningen te kiezen, ondanks de hogere moeilijkheidsgraad. We bouwden speciaal voor deze studie de digitale onderwijsapplicatie Alchemath. De applicatie genereert oefeningen over vergelijkingen *on the fly* en maakt gebruik van vier gamificatie-elementen die we bewust kozen voor hun relevantie in de literatuur en de opzet van dit onderzoek.

Hoofdstuk 2

Literatuurstudie

Gamificatie komt steeds meer voor in talrijke toepassingen zoals onderwijs, gezondheid, marketing en bedrijfsproductiviteit. Naast gamificatie in alledaagse toepassingen, is ook de academische interesse gegroeid. Dit hoofdstuk verkent eerst de betekenis van gamificatie en het verwantschap met games. Vervolgens gaat het in op een typisch doel van gamificatie, namelijk motivatie. Mensen verschillen van aard en daarom bestaat er geen universele aanpak voor gamificatie. Om zo veel mogelijk gebruikers aan te spreken houdt het dus steek om de ervaring te personaliseren. Paragraaf 2.4 gaat in op verschillende modellen die personalisatie in games en gamificatiesystemen mogelijk maken. Ten slotte gaat Paragraaf 2.5 dieper in op de onderwijstoepassing van gamificatie en de aandachtspunten om ook het leerproces gepersonaliseerd en motiverend te maken.

2.1 Games en gamificatie

Games en gamificatie zijn sterk verwant, maar toch verschillend. Eerst bespreekt deze paragraaf wat games zijn en welke rol ze spelen in serieuze toepassingen zoals onderwijs, daarna gaat ze in op het begrip gamificatie en ten slotte kadert ze de kritiek op gamificatie en hoe die kritiek leidde tot de notie *meaningful gamification*.

2.1.1 Games en serious games

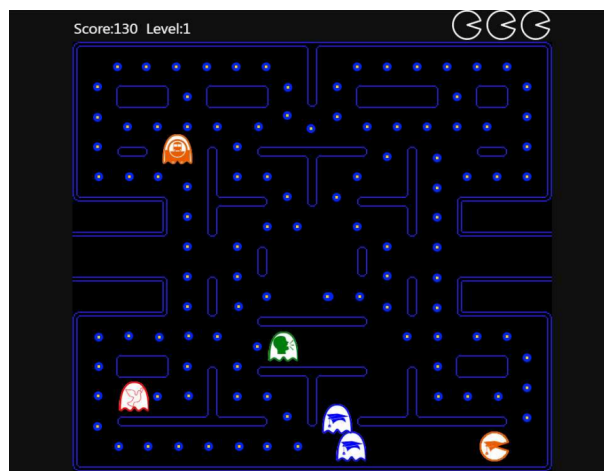
Een game is een elektronisch spel dat mensen spelen op een computer, console, mobiel toestel of ander apparaat [1]. Volgens die definitie vormen games een deelverzameling van spellen en beide hebben doorgaans als primair doel om te vermaken. Spellens waarvan het voornaamste doel daarvan afwijkt heten *serious games* [2] en hebben talloze toepassingen zoals wetenschappelijke instrumenten, trainings simulaties, ondersteuning bij medische behandelingen en onderwijs. Merk op dat *serious games* niet noodzakelijk games zijn; ze kunnen gebruikmaken andere media zoals bord- en kaartspellen [3].

Om de wetenschap te vorderen bestaan er bijvoorbeeld *human computation games* of HCGs. Dat zijn spellen die taken waar computers moeite mee hebben op een speelse manier aan mensen voorstellen. HCGs zijn door hun maatschappelijk nut een belangrijk aandachtspunt binnen de gamificatie [4]. Een bekend voorbeeld is het puzzelspel *Foldit*

[5] waarin spelers bepaalde proteïnes zo goed mogelijk proberen te plooien. Oplossingen met hoge scores kunnen toepassing hebben op proteïnen in de echte wereld.

Spellen gebruikt in onderwijs vormen een andere belangrijke toepassing van *serious games*, in de literatuur ook wel *serious educational games* of SEGs genoemd. Het onderzoeksdomein rond SEGs heet onder andere *game-based learning* of *edutainment*. SEGs kunnen langdurig de aandacht van spelers trekken en bieden een goedkoop, tijd- en ruimteonafhankelijke alternatief voor traditioneel lesgeven [6].

Een sleutelaspect om een spel zowel leuk als leerrijk te maken is intrinsiek leren, *intrinsic learning* of *intrinsic integration*. In een SEG die intrinsiek leren toepast maken spelers rechtstreeks gebruik van de leerstof om het speldoel te behalen. Cutting et al. [7] verduidelijkten de werking van intrinsiek leren in een game waarbij deelnemers een reeks afbeeldingen moesten onthouden, zie Figuur 2.1. Er waren twee versies van het spel, die verschilden in de manier waarmee spelers spoken konden opeten. In de versie met intrinsiek leren konden spelers enkel spoken opeten die dezelfde afbeeldingen hadden als zij. In de versie zonder intrinsiek leren moest slechts de kleur overeenkomen tussen de twee. Ze concludeerden dat aandacht de belangrijkste factor was voor de leerefficiëntie van intrinsiek leren.



Figuur 2.1: De *serious game* gebruikt in de studie van Cutting et al. [7] rond intrinsiek leren.

2.1.2 Gamificatie

In het algemeen verwijst gamificatie naar de technologische, economische, culturele en maatschappelijke ontwikkelingen die *gamefulness* vergroten. Die ontwikkelingen gaan gepaard met typische voordelen geassocieerd met spelen: vaardigheden opbouwen, motivatie verhogen, creativiteit, speelsheid, engagement, en over het algemeen positieve groei en geluk [8].

Gamificatie kan ontstaan op twee manieren: *emergent gamification* en *intentional gamification*. *Emergent gamification* is een gradueel opkomende ontwikkeling, mogelijk intentioneel, die ontstaat door de steeds groeiende rol van spellen in de maatschappij. Recente fenomenen zoals *e-sports*, YouTube en *game modding* tonen de integratie van werk en spel

en de culturele en sociale impact die dat heeft op het dagelijkse leven. *Intentional gamification* daarentegen is een weldoordachte ontwikkeling die de ervaring van een activiteit, systeem, dienst, etc. doet overeenkomen met een echt spel [8].

Ontwerpers die gamificatie bewust toepassen in een applicatie doen dus aan *intentional gamification* en daarvoor maken ze gebruik van elementen uit spellen in een niet-spelcontext zoals onderwijs. Er zijn talloze voorbeelden van zulke spelelementen; Figuur 2.2 toont een overzicht van enkele veelvoorkomende.



Figuur 2.2: Een overzicht van enkele veelvoorkomende gamificatie-elementen met korte omschrijvingen [9].

Net als spellen bevatten gamificatiesystemen dus spelelementen, maar dat maakt zo'n systeem nog geen *serious game*. Het onderscheid tussen beide komt tot uiting tijdens het ontwerpproces. Ontwerpers maken *serious games* met de niet-spelcontext in het achterhoofd; ze trachten beide te integreren. SEGs gebruiken bijvoorbeeld *intrinsic learning* om spel en leerstof samen te voegen en creëren een geheel dat beide delen overstijgt. Gamificatie daarentegen beschouwt de niet-spelcontext en de spelelementen als twee onafhankelijke lagen van de applicatie. In een gamificatiesysteem is het gemakkelijk om de spelelementen weg te denken zonder de werking van het systeem te beïnvloeden.

2.1.3 Kritiek op gamificatie

Vergeleken met games krijgt gamificatie vaak kritiek op haar tekort aan inhoudelijke waarde. Hoewel gamificatie spelelementen aanwendt, zijn deze elementen zelden geïntegreerd met de rest van het systeem zoals in een game. SEGs maken bijvoorbeeld gebruik van visualisaties die inzicht versterken of stappen uit oplossingen verduidelijken. Gameontwerper Radoff [10] haalt bijvoorbeeld narratief en *immersion* aan als elementen die zelden in gamificatiesystemen voorkomen. De term gamificatie doet sterk denken aan games en volgens sommigen is het die associatie niet waardig. Robertson [11] bijvoorbeeld verkiest de term *pointsification* omdat het meestal de rijkdom van een spel reduceert tot punten en badges.

“[Points and badges are] great tools for communicating progress and acknowledging effort, but neither points nor badges in any way constitute a game... They are the least important bit of a game, the bit that has the least to do with all of the rich cognitive, emotional and social drivers which gamifiers are intending to connect with” [11].

Die kritiek op de inhoud van gamificatie leidde tot nieuwe vormen die meer aanspreken tot de intrinsieke motivatie. Nicholson [12] beschrijft *meaningful gamification* als een humanistische aanpak om vaardigheden op verschillende manieren te versterken. Hij legt onder andere de nadruk op de link tussen spel en werkelijkheid; het spel is telkens een middel om vaardigheden in de echte wereld te versterken. Jonge kinderen vinden inspiratie in spellen die ze leuk vinden om buiten het spel hun kennis te verruimen over het onderwerp in boeken of websites en tekeningen of verhalen te creëren rond het thema. Gee [13] noemt het “education at its best”. Op basis van onder andere *self-determination theory* stelde Nicholson [12] zes aandachtspunten op om gamificatie *meaningful* te maken:

- *Play* biedt de kans om uit te proberen, te ontdekken en fouten te maken zonder consequenties. Met zo’n systeem kunnen gebruikers zelf bepalen wat ze leuk vinden en valt de nood voor externe beloningen weg.
- *Exposition* creëert een narratief dat de leerstof linkt aan de echte wereld. Als gebruikers dat verband begrijpen zullen ze beter inzien welke voordelen de beoogde vaardigheden voortbrengen. Een narratief in een fantastisch kader daarentegen zou gebruikers kunnen afleiden van dit praktisch nut.
- *Choice* geeft gebruikers de mogelijkheid om te bepalen wat ze willen leren en hoe ze het willen leren. Het is sterk gelinkt aan *autonomy* uit SDT.
- *Information* probeert gebruikers de motivering achter leerstof duidelijk te maken en welke toepassingen het heeft in de echte wereld.
- *Engagement* focust op sociale interactie. De motivatie van anderen kan inspireren en sociale interactie zoals samenwerking of competitie zijn drijvende factoren voor motivatie. Het is sterk gelinkt aan *relatedness* uit SDT.
- *Reflection* helpt gebruikers om hun ervaringen in het gamificatiesysteem te linken aan ervaringen uit hun leven. Het creëert een sterkere verbinding tussen de twee, waardoor het leerproces gemakkelijker vooruitgaat. Het is ook een kans om de leerstof te linken aan andere interesses en zo een mentaal netwerk rond de leerstof

op te bouwen.

Een ander probleem dat gamificatie soms heeft is een *one size fits all* aanpak. Uit recent onderzoek blijkt dat gamificatie veel doeltreffender is om de gebruikerservaring te verbeteren wanneer een systeem spelelementen voorziet op basis van de voorkeuren van individuele gebruikers. Paragraaf 2.4 gaat hier dieper op in.

2.2 Motivatie

Motivatie drijft mensen om met games of gamificatiesystemen te interageren. Er zijn echter verschillende vormen van motivatie die vaak gelinkt zijn aan menselijke behoeften. Deze paragraaf bespreekt eerst een gevalideerd psychologisch model rond motivatie en de verschillen tussen intrinsieke en extrinsieke motivatie. Daarna gaat ze in op beloningen en de verschillende factoren die deze bron van extrinsieke motivatie beïnvloeden.

2.2.1 Self-determination theory

Self-determination theory (SDT) is een psychologisch model over menselijke motivatie door Deci et al. [14]. Het maakt in de eerste plaats een onderscheid tussen extrinsieke motivatie en intrinsieke motivatie. Extrinsieke motivatie ontstaat door beloningen die niets te maken hebben met de taak. Die beloningen kunnen mensen motiveren die de taak niet uit zichzelf willen maken. Met intrinsieke motivatie daarentegen, halen personen voldoening uit de taak zonder externe overtuiging. Er zijn drie basisbehoeften die personen hebben om intrinsieke motivatie op te wekken [14]:

Autonomie: het verlangen om het leven in eigen hand te nemen en zelfstandig keuzes te maken. Zelfstandig betekent niet onafhankelijk van anderen, maar gaat eerder over psychologische vrijheid en vrije wil. Autonomie kan dus voldaan zijn zelfs in het geval een ander de taak opdraagt. Deci [15] heeft ook gevonden dat externe beloningen autonomie en dus intrinsieke motivatie onderdrukken.

Competentie: het verlangen de uitkomst van een taak te beïnvloeden en geassocieerde vaardigheden te beheersen. Positieve feedback is een manier om competentie te versterken doordat het de competentie bevestigt [15].

Verbondenheid: het verlangen om met anderen om te gaan, tot anderen te behoren en en voor anderen te zorgen. Als anderen in de directe omgeving van een persoon intrinsieke motivatie hebben voor een taak, dan kan die motivatie overdragen naar de persoon zelf.

2.2.2 Beloningen

Beloningen of *rewards* zijn essentiële elementen van bijna alle games en gamificatiesystemen. Ze kunnen een invloed hebben op het verloop van het spel, maar kunnen ook de vorm aannemen van een aanmoediging, een leuk waarneembaar effect of een certificaat van vooruitgang zoals punten of badges. De bedoeling van een beloning is telkens om positieve feedback te geven aan spelers en hun motivatie te versterken. Er bestaat geen exhaustieve lijst van beloningen, maar wel groeperingen zoals de taxonomie van Philips

et al. [16]. Tabel 2.1 toont de zes categorieën van beloningen volgens Philips et al. uit hun onderzoek dat 915 soorten beloningen vergeleek over 60 commerciële games.

Tabel 2.1: Zes categorieën van beloningen volgens Philips et al. [16]

Beloning	Beschrijving
<i>Rewards of Access</i>	Beloningen die toegang bieden tot exclusief spelmateriaal.
<i>Rewards of Facility</i>	Beloningen die spelers meer doeltreffend maken in de huidige speltoestand.
<i>Rewards of Sustenance</i>	Beloningen die de effecten van nadelen zoals beperkte levens, beperkte ammunitie, etc. verzachten.
<i>Rewards of Glory</i>	Beloningen die geen invloed hebben op de speltoestand, maar eerder dienen als certificaten van vooruitgang of competentie zoals badges of scoreborden.
<i>Rewards of Praise</i>	Aanmoedigingen in woorden.
<i>Sensory Feedback</i>	Beloningen met visuele of auditieve esthetische waarde die vaak gepaard gaan met andere beloningen.

Voordeel van variatie

Philips et al. [17] onderzochten hoe beloningen de ervaring van spelers beïnvloeden. Ze hebben geen resultaten kunnen vinden over individuele soorten beloningen, maar wel over de variatie van beloningen. De aanwezigheid van verschillende soorten beloningen heeft een positief effect op de spelervaring. Volgens hen is onder andere het verschil in soorten spelers een aspect dat een variatie aan beloningen vereist; bepaalde soorten spelers halen meer motivatie uit één beloning dan een andere.

Siu et al. [4] onderzochten of de keuze tussen beloningen een invloed heeft op de efficiëntie waarop gebruikers taken voltooien in een HCG. Het experiment gebruikte drie soorten beloningen: een plaats op een scorebord, een aanpasbare avatar, of een verhaaltje. Deelnemers die na elke taak de keuze kregen tussen de drie toonden een hogere correctheid van taken, hoger aantal voltooide taken en lagere gemiddelde antwoordtijd. Hieruit leidden Siu et al. af dat de deelnemers meer motivatie ervaarden.

Placebo-effect

Een klassieke soort beloning in games zijn *power-ups*. Wat *power-ups* uniek maken is hun optionele karakter. Spelers hebben altijd de keuze om een power-up al dan niet te verzamelen. Denisova et al. [18] onderzochten welk effect *power-ups* hebben op de spelervaring. Hun resultaten tonen aan dat *power-ups* de *immersion* en autonomie van spelers versterken. Ze vonden echter geen invloed op het gevoel van competentie. Naast de aanwezigheid van *power-ups* onderzochten Denisova et al. ook het verschil tussen *power-ups* met impact en zonder impact op de speltoestand. Interessant is dat de placebo *power-ups* niet significant afweken van de echte *power-ups* op het vlak van *immersion* en autonomie. Dat wilt zeggen dat een beloning met enkel een esthetisch effect nog steeds een sterke invloed heeft op de motivatie van spelers.

2.3 Nudging

Nudging is een begrip uit *nudge theory* dat vertaalt naar een duwtje geven in de goede richting. Volgens die theorie zijn *nudges* subtiele, overtuigende berichten om de keuze van mensen te beïnvloeden zonder de vrijheid van kiezen te ontnemen [19, 20]. De techniek heeft vaak als doel om het gedrag van mensen te verbeteren, zoals sterkere wachtwoorden kiezen [19] of meer lokale voeding kopen [21].

Beloningen zijn een vorm van nudging, specifiek *incentive nudging*. Er zijn echter ook vele andere manieren om gedrag te nudgen, zoals Tabel 2.2 samenvat.

Tabel 2.2: Verschillende soorten nudging, geordend volgens de mnemoniek MINDSPACE.

Nudging	Beschrijving
<i>Messenger</i>	De persoon die de informatie overbrengt heeft een belangrijke invloed.
<i>Incentives</i>	Psychologische irrationele gedragspatronen zoals verliesaversie en neiging naar de status quo beïnvloeden hoe mensen reageren op prikkels.
<i>Norms</i>	Het gedrag van anderen en van de meerderheid is een bepalende factor.
<i>Defaults</i>	Mensen volgen de keuzes die op voorhand bepaald zijn uit gemakszucht.
<i>Saliency</i>	De aandacht van mensen gaat naar wat nieuw en relevant lijkt.
<i>Priming</i>	Onbewuste signalen beïnvloeden ons gedrag.
<i>Affect</i>	Emotionele associaties hebben een sterke invloed op onze acties.
<i>Commitment</i>	Mensen willen zich houden aan publieke beloften.
<i>Ego</i>	Mensen handelen op een manier waarop ze zichzelf beter voelen.

Een voorbeeld vorm van nudging is het *watching-eye effect*. Illustraties van ogen, zoals in Figuur 2.3 zorgen ervoor dat mensen minder egoïstisch en minder antisociaal handelen, omdat mensen het gevoel krijgen dat ze onder toezicht staan. Dit kan leiden tot de daling en preventie van misdaad zoals diefstal [22].



Figuur 2.3: Een illustratie van ogen op een spandoek, een voorbeeld van *norm nudging* [23].

Nudging kent vele toepassingen en één daarvan is onderwijs. Nudging is gewenst wanneer bijsturing van gedrag kan leiden tot positieve effecten, zoals mensen helpen om hun

vaardigheden te versterken. Wambsganss et al. [24] onderzochten of ze studenten konden nudgen om beter onderbouwde teksten te schrijven door ze een vergelijking met andere studenten voor te schotelen. Vergeleken met de controlegroep zorgde die vorm van *norm nudging* voor meer overtuigende en beter geschreven teksten.

2.4 Gepersonaliseerde gamificatie

Er bestaat geen *one size fits all* aanpak voor gamificatie; de doeltreffendheid van een gamificatie-element of spelelement hangt af van de eindgebruiker. Om spelelementen op maat te maken volstaat het echter om te werken met een generalisatie van die eindgebruikers. Verschillende theoretische modellen proberen mensen in te delen volgens een reeks archetypes om zo uitspraken te doen over hun voorkeuren en gedrag patronen. Deze paragraaf gaat dieper in op drie zulke modellen: *Big Five* over persoonlijkheid, BrainHex over motivatie binnen games en Hexad over motivatie binnen gamificationsystemen.

2.4.1 Big Five

De Big Five-persoonlijkheidstreken of het Five Factor Model is een psychologisch model dat in ontwikkeling is sinds de jaren '80 [25]. Zoals de naam suggereert, onderscheidt het vijf eigenschappen van mensen:

- *Openness for experience* is een algemene appreciatie voor visuele, auditieve, etc. ervaringen. Het is ook verbonden aan nieuwsgierigheid, emotioneel bewustzijn en de wil om nieuwe dingen te proberen.
- *Conscientiousness* zorgt voor zelf-discipline en de wil om te presteren, zelfs tegen verwachtingen in.
- *Extraversion* is de neiging om met de externe wereld en anderen om te gaan. Extraverten krijgen energie door sociaal contact, terwijl introverten hierdoor energie verliezen.
- *Agreeableness* is de neiging om sociale harmonie te onderhouden en een optimistische blik op de menselijke aard te hebben. Het hangt samen met gulheid, vertrouwen, zorg en bereidheid anderen te helpen.
- *Neuroticism* is de neiging om negatieve emoties te ervaren. Het omgekeerde van neuroticisme is emotionele stabiliteit. Het is een kenmerk dat jaren is onderzocht voor het als deel van het FFM werd voorgesteld. [26].

2.4.2 BrainHex

BrainHex is een model gebaseerd op neurobiologie uit 2014, dat spelers van games classificeert in zeven archetypes op basis van motivatie. Door de focus op games komen de archetypes vaak overeen met gamegenres zoals *Survivors* en horrorgames of *Daredevils* en racegames. De classificatie is ook gelinkt aan oudere spelertypologiën zoals Myers-Briggs en Bartle [27]. De zeven archetypes die BrainHex onderscheidt zijn [28]:

- *Achievers* zijn doel-georiënteerd. Ze willen spellen voltooien en stellen doelen op lange termijn. RPG-spellen bouwen vaak rond die filosofie. Ze passen moeilijkheid

dynamisch aan en belonen spelers die de lange weg afleggen om een karakter tot het hoogste niveau te brengen. Het archetype is verwijst naar Bartle's achiever.

- *Conquerors* houden van uitdagingen, in het bijzonder uitdagingen die moeilijk te overkomen zijn. In zekere mate komen de archetypes *conqueror* en *achiever* overeen, maar ze verschillen van doelstelling. Conquerors spelen minder op lange termijn en zijn niet geïnteresseerd om slechts een lijst af te vinken.
- *Daredevils* houden van snelheid en risico's. Racing games zijn in het bijzonder geschikt voor dit archetype.
- *Seekers* zijn nieuwsgierig en willen de spelomgeving ontdekken en ervaren. Hun vorm van motivatie is gelinkt aan het deel van de hersenen dat waarneming registreert (nl. sensory cortices) en het associatief geheugen (nl. hippocampus).
- *Socialisers* interageren graag met anderen, helpen graag anderen en hechten veel waarde aan vertrouwen. Dat gedrag is gelinkt aan het sociaal centrum in de hersenen wat een bron is van een neurotransmitter geassocieerd met vertrouwen; oxytocine. Het archetype is verwijst naar Bartle's socialiser.
- *Survivors* genieten van de spanning van angst in een gecontroleerde omgeving. Epinephrine is de neurotransmitter dat achter die ervaring ligt en verstrekt de effecten van dopamine. Horror games zijn het typisch geschikt voor dit archetype.
- *Masterminds* houden van probleemoplossend denken en strategie. Het archetype hangt dicht samen met Bartle's explorer en vindt typisch voldoening in puzzle of strategy games.

Orji et al. [29] onderzochten gezond-gerelateerd gedrag door factoren uit het *health belief model* te koppelen aan BrainHex-archetypes. Zo was een belangrijke factor voor *achievers* de *perceived susceptibility*: hun inschatting van het risico op een gezondheidsprobleem. Voor *conquerors* daarentegen bleken *perceived benefits* belangrijker te zijn.

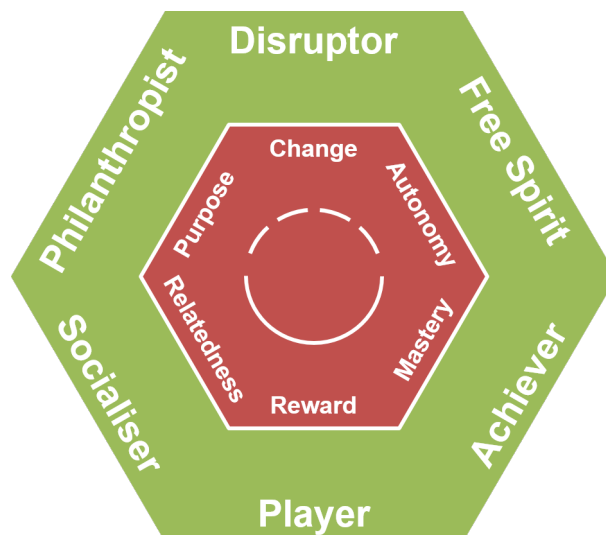
2.4.3 Hexad

Het Hexad-model dateert uit 2015 en classificeert gebruikers van gamificatiesystemen in zes archetypes [30]. Marczewski et al. introduceerden samen met het model een vragenlijst om de archetypes van mensen af te leiden, die Tondello et al. [31] empirisch valideerden. Een vergelijkende studie tussen FFM, BrainHex en Hexad concludeerde dat Hexad het meest geschikt is voor gepersonaliseerde gamificatie [32]. De zes archetypes die Hexad onderscheidt, zie ook Figuur 2.4, zijn [30]:

- *Achievers* (A) streven naar **mastery**. Ze willen obstakels overbruggen en hun vaardigheden verbeteren.
- *Disruptors* (D) streven naar **change**. Ze zoeken de grenzen op van een systeem, met positieve of negatieve gevolgen.
- *Free Spirits* (F) streven naar **autonomy**. Ze willen ontdekken en de vrijheid om eigen keuzes te maken.
- *Philanthropists* (Ph) streven naar **purpose**. Ze handelen sociaal en houden ervan

als anderen afhankelijk van hen zijn. Ze nemen graag verantwoordelijkheid en delen graag hun kennis om anderen te helpen.

- *Players* (P) streven naar **beloningen**. Ze willen vooral winnen en beloningen ontvangen en hebben weinig intrinsieke motivatie.
- *Socialisers* (S) streven naar **relatedness**. Ze handelen sociaal zoals filantropen, maar zoeken eerder naar sociale interactie zoals samenwerking en competitie.



Figuur 2.4: Marczewski's *User Type Hexad* [30]. De voorstelling toont de archetypes aan de buitenkant en achterliggende motivaties aan de binnenkant.

Tondello et al. [33] onderzochten de link tussen spelelementen en de archetypes van de Hexad- en FFM-modellen. Ze hebben eerst 59 *gameful design elements* vastgelegd en die elementen geclassificeerd in acht groepen. In hun studie bepaalden Tondello et al. eerst de archetypes van alle deelnemers volgens beide modellen [31]. Daarna moesten deelnemers de 59 spelelementen beoordelen op een 5-punts Likertschaal, op basis van hun ervaren plezier. De verbanden die ze vonden, zie Tabel 2.3, kunnen ontwerpers gebruiken om hun gamificationsystemen te personaliseren.

Tabel 2.3: Verbanden tussen soorten spelelementen en Hexad-gebruikertypes [33].

Element	A	D	F	Ph	P	S
<i>Socialization</i>	matig				matig	sterk
<i>Assistance</i>						zwak
<i>Immersion</i>	sterk	zwak	sterk	zwak		
<i>Risk/Reward</i>	matig	zwak			matig	
<i>Customization</i>						
<i>Progression</i>	zwak			zwak		
<i>Altruism</i>	zwak			sterk		matig
<i>Incentive</i>					sterk	

Om de de verbanden tussen Hexad-types en hun spelelementen bij voorkeur te onderzoeken, vergeleken Altmeyer et al. [34] de Hexad-profielen van deelnemers met hun affiniteit

met twaalf beschikbare *storyboards*, illustraties die telkens een bepaald spelelement afbeelden. De gebruikers gaven aan elke *storyboard* een waardering op een 7-punts Likertschaal en die vergeleken Altmeyer et al. met de verwachte voorkeuren voor spelelementen voor hun Hexad-types. Ze concludeerden dat deelnemers voor bijna alle spelelementen een affiniteit toonden die overeenkwam met hun Hexad-type.

Zoals eerder vermeld gebruiken studies vaak de standaard Hexad-24 vragenlijst [31] om het Hexad-type van gebruikers af te leiden. Hoewel die vragenlijst zeker geschikt is voor onderzoeksdoeleinden, kan hij in praktijk de *immersion* en algemene gebruikservaring verminderen [35, 36]. Altmeyer et al. [35] probeerden dit probleem aan te pakken door gamificatie toe te passen op de vragenlijst zelf, waardoor de vragenlijst ook minder haaks staat op de rest van het gamificatiesysteem. In het onderzoek van Altmeyer et al. vulden deelnemers de Hexad-24 vragenlijst in om hun Hexad-type te bepalen, alsook de 22-delige IMI-vragenlijst die achterhaalde hoe gebruikers de invulling van de Hexad-24 ervaarden. Daarna vulden de deelnemers de *gamified* variant van de Hexad-vragenlijst, *Cloud Clicker*, in. Door de resultaten van de twee lijsten te vergelijken, vonden Altmeyer et al. dat een *gamified* vragenlijst even effectief is als de standaard vragenlijst, maar dat gebruikers haar wel als leuker en meer *immersive* ervaarden.

Altmeyer et al. [35] stelden in diezelfde paper ook een alternatief voor om het gamificatietype af te leiden, zonder enige vorm van vragenlijst. Het idee is dat in de aanwezigheid van verschillende soorten spelelementen gebruikers geneigd zijn om meer met bepaalde elementen te interageren. Door dus te kijken naar het gedrag van gebruikers kan een digitaal systeem op een natuurlijkere manier personalisatie toepassen. Die aanpak vermijdt de kans dat mensen zich fout inschatten door een discrepantie in zelfkennis. Na de vermelde vragenlijsten in te vullen, speelden deelnemers het spel *Snowball Shooter*. Gebruikers bestuurden hierbij een sneeuwkanon waarmee hij of zij kan schieten op spelelementen naar keuze. Het spel bood ontgrendelbare items, badges en een virtueel karakter aan. Uit de resultaten bleek dat gebruikers meer interageerden met de spelelementen die overeenkwamen met hun Hexad-types en dat het gebruiksgedrag dus een goede indicatie is om het Hexad-type te bepalen.

2.4.4 Kritiek op modellen

Het is duidelijk dat er een hele waaier aan modellen bestaat rond psychologie en motivatie om gebruikers te verdelen onder archetypes. Welk model het meest geschikt is hangt af van de toepassing. Voor gamificatie is Hexad bijvoorbeeld het meest toepasselijk door de link met gamificatie.

Om de voorkeuren van gebruikers correct te modelleren, blijkt één archetype onvoldoende. Gebruikers uiten zich vaak als een gewogen gemiddelde van meerdere types en volgens Halifax et al. [32] hebben vooral *Socialiser*, *Disruptor*, *Achiever* en *Free spirit* een grote invloed. In de context van jongeren vonden Ooge et al. [37] dat Free Spirit een zwakke interne consistentie heeft en dat *Socialiser* en *Philanthropist* zo sterk aan elkaar gelinkt zijn dat ze wellicht beter onder één type vallen. Zij stelden in dit kader een vereenvoudiging van het Hexad-model voor.

Het is duidelijk dat de aannames van theoretische modellen zoals Hexad, Big Five en BrainHex niet altijd overeenkomen met het feitelijke gedrag van mensen. Krath et al.

[38] bevestigden dat suggesties van experts en theoretische modellen niet altijd te verenigen zijn met de werkelijkheid. Eén verklaring is dat de voorkeuren van mensen veranderen en dat het daarom niet voldoende is om het profiel van gebruikers eenmalig op te stellen; het systeem zou ze doorlopend moeten bijwerken [35]. Anderzijds kan het zijn dat gebruikers zichzelf slecht inschatten tijdens de pre-studie vragenlijst en dat hun echte voorkeuren past tot uiting komen in hun gedrag [38].

2.5 Gepersonaliseerd leren

Elke student leert op een andere manier, op een ander tempo en heeft andere noden om gemotiveerd te blijven tijdens het leren. Aanbevolen oefeningen moeten dus overeenkomen met het niveau van het individu. Deze paragraaf bespreekt eerst wat oefeningen moeilijk maakt en hoe die factoren kunnen worden gemeten. Daarna verkent ze optioneel leren, een manier om de autonomie van gebruikers te verhogen, een manier om de vaardigheid van gebruikers te modelleren met het Elo-systeem en ten slotte bespreekt ze hoe oefeningen in de *zone of proximal development* een leerling gemotiveerd houden.

2.5.1 Complexiteit en moeilijkheid

Complexiteit en moeilijkheid zijn twee nauw verwante, maar verschillende concepten. Volgens Pelánek et al. [39] is complexiteit een eigenschap van een probleem dat zijn structuur en aspecten beschrijft; het is objectief en onafhankelijk van wie het probleem oplost. Moeilijkheid daarentegen is een subjectieve eigenschap; het beschrijft de moeite die een individu heeft om een probleem op te lossen. Figuur 2.5 toont beide concepten op een schaal van subjectiviteit met bijhorende metrieken.

Er zijn verschillende manieren om complexiteit en moeilijkheid te meten. Voor verschillende domeinen zijn sommige metrieken beter geschikt dan andere. Pelánek et al. [39] destilleerden drie metrieken uit het algemeen overzicht van complexiteit van Liu et al. [40] die rechtstreeks toepassing hebben op onderwijs:

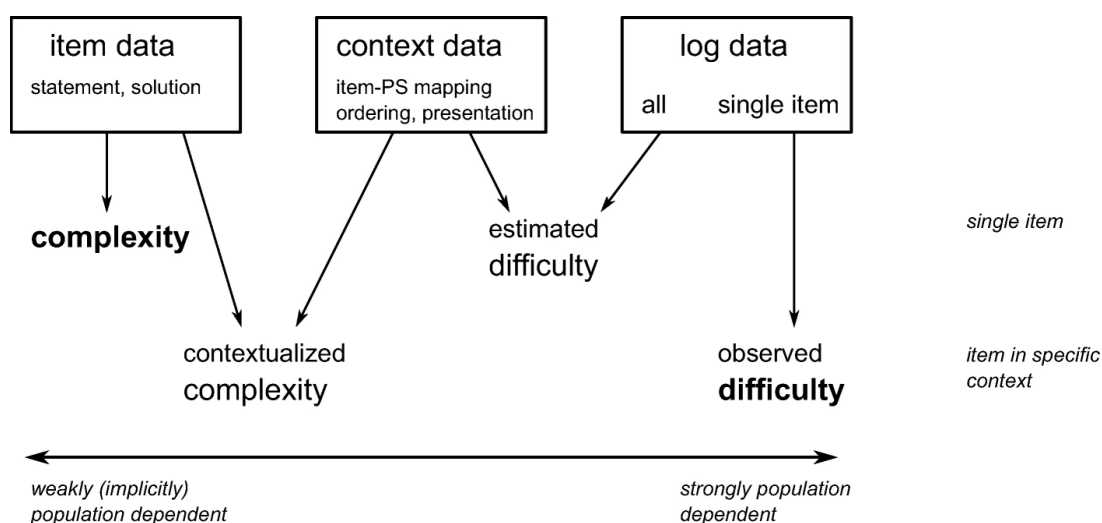
- *Grootte* kan gemeten worden a.d.h.v. de opgave of de oplossing van een probleem. Mogelijke lengte-eenheden zijn het aantal karakters, aantal woorden of aantal lijnen code.
- *Variatie* is het aantal verschillende concepten nodig om een probleem om te lossen. In een wiskundeprobleem kan een concept bijvoorbeeld optelling, aftrekking, vermenigvuldiging of deling zijn. Ook hier kan de metriek kijken naar de opgave, oplossing of oplossingsproces. De manier van concepten tellen is belangrijk, bijvoorbeeld optelling en aftrekking zouden één concept kunnen vormen. Mogelijk wegen bepaalde concepten meer door dan andere, waardoor een gewogen som van toepassing is.
- *Relaties* zijn afhankelijkheden tussen verschillende stappen van de oplossing. Voor een wiskundeprobleem kan de diepte van de betreffende *computation tree* dienen als een geschikte metriek.

Moeilijkheid vereist dan weer andere metrieken door haar subjectieve karakter. Om die metrieken te meten zijn per definitie proefpersonen nodig. Het is belangrijk om daarbij

de doelgroep van een toepassing in acht te nemen. Pelánek et al. haalden de volgende metrieken aan die voor bijna alle domeinen geschikt zijn:

- Mediane antwoordtijd generaliseert de benodigde tijd om een probleem op te lossen. De mediaan is een betere keuze dan het gemiddelde, omdat antwoordtijden vaak uitschieters hebben en de mediaan ze compenseert.
- Faalpercentage beschrijft de proportie mensen van een testgroep die een probleem niet konden oplossen. Faalpercentage is een natuurlijkere metriek dan slaagpercentage omdat het recht evenredig is met moeilijkheid.

Daarnaast zijn er nog probleem-afhankelijke metrieken zoals aantal pogingen, aantal gebruikte hints, kwaliteit van de oplossing en aantal bewerkingen om tot het antwoord te komen.



Figuur 2.5: Moeilijkheid verdeeld o.b.v. subjectiviteit. De rechthoeken stellen metrieken voor om moeilijkheid te meten [39].

2.5.2 Optioneel leren

Paragraaf 2.2.2 maakte duidelijk dat keuze en variatie een belangrijke rol spelen om beloningen motiverend te maken. Die resultaten linkt deze paragraaf aan gepersonaliseerd leren. Joshi [41] verkende de keuze van beloningen in SEGs en het positieve effect van optioneel leren. Hij stelde een SEG voor over programmeren, die de vrijheid en keuze van spelers vergroot zonder de educatieve inhoud te compromitteren. Het spel biedt spelers voordelen die ze kunnen verkrijgen door een optionele taak te maken rond de leerstof. Tabel 2.4 toont vier voorkomende spelmechanismen die hij aan die voordelen verbindt.

2.5.3 Elo-rating

Om oefeningen op maat te voorzien moet een systeem eerst weten hoe moeilijk oefeningen zijn en hoe competent iemand is om ze op te lossen. Eén van de systemen dat verschillende studies rond onderwijs [42, 43, 44] gebruikt hebben is het Elo-systeem, vernoemd naar zijn

Tabel 2.4: Vier voorkomende spelmechanismen die volgens Joshi [41] geschikt zijn om optioneel leren op toe te passen.

Spelmechanisme	Beschrijving
<i>Reverse failure</i>	Faalcondities vermijden om zo vooruitgang binnen het spel te behouden.
<i>Manipulate environment</i>	Tools om op exclusieve manieren met de spelomgeving om te gaan.
<i>Heighten success</i>	Bijkomende waardering voor de winconditie te behalen. Door duidelijk te maken dat zulke waardering ontbreekt zonder de taak te maken, spreekt het ook een neiging naar volledigheid aan.
<i>Unlock bonuses</i>	Toegang tot exclusieve delen van de spelomgeving, wat een gevoel van ontdekking en autonomie aanspreekt.

uitvinder Arpad Elo. Arpad ontwierp het systeem om de competentie van schaakspelers onderling te kunnen vergelijken en handhaven. Na elk spel bepaalt het systeem de nieuwe Elo-scores op basis van de huidige scores en de uitslag van het spel. Stel dat speler A tegen speler B heeft gespeeld en hun huidige scores respectievelijk R_A en R_B zijn. Gegeven hun relatieve scores is de kans dat A wint gelijk aan

$$E_A = \frac{1}{1 + 10^{(R_B - R_A)/400}}.$$

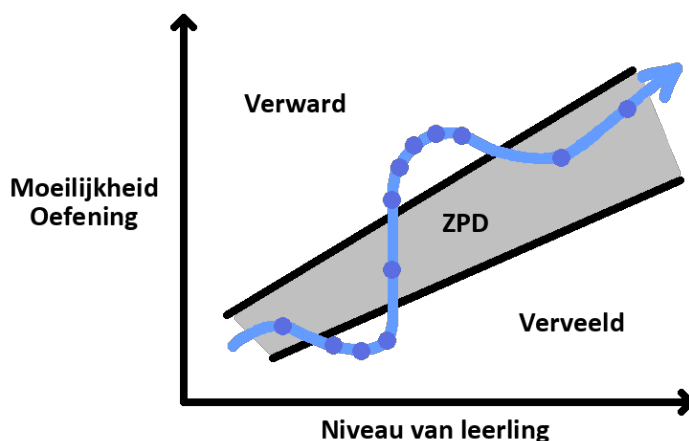
Vanuit het perspectief van speler A kan de uitslag S_A een waarde aannemen in $[0, 1]$, waarbij nul betekent dat A verloor en één dat A won. Sommige toepassingen laten tussenliggende waarden toe om bijvoorbeeld een gelijkstand aan te duiden. De nieuwe Elo-score van A volgt uit de formule

$$R'_A = R_A + K(S_A - E_A).$$

De waarde K , de K-factor, is een coëfficiënt die bepaalt hoe snel een speler stijgt of daalt. Ze varieert per toepassing, maar hangt meestal af van de spelers rating en het aantal partijen dat de speler voltooid heeft. De K-factor is essentieel om de initiële Elo-score van een speler te bepalen. Spelers starten meestal met een vaste Elo-score en door een grotere K-factor te kiezen voor de eerste partijen, zal de speler sneller convergeren naar zijn daadwerkelijke positie binnen het systeem [45].

2.5.4 Zone of proximal development

Onderwijssystemen streven twee doelen na; ze willen het leerproces zo effectief en zo efficiënt mogelijk maken. Met andere woorden ze willen dat mensen zo veel mogelijk leren is zo weinig mogelijk tijd. Opdrachten moeten uitdagend zijn, maar binnen de mogelijkheden van mensen blijven, gegeven hun vaardigheden. Onder die condities bevindt een opdracht zich in de *zone of proximal development* (ZPD) [46]. Oefeningen die voor een leerling binnen de ZPD vallen zijn positief frustrerend, waardoor de leerling gemotiveerd is om ze op te lossen. Figuur 2.6 toont hoe een traject van een leerling er mogelijk uit ziet in functie van de ZPD; de punten op het traject stellen oefeningen voor. Het systeem



Figuur 2.6: Voorbeeldtraject van een leerling in functie van de ZPD [46].

moet in staat zijn waar te nemen wanneer oefeningen buiten de ZPD vallen zodat het daarna de moeilijkheid van oefeningen kan bijstellen [46].

Hoewel gamificatie en SEGs spelelementen gebruiken om het leerproces te bevorderen, is er ook een natuurlijk leerproces aanwezig in games zelf. Games hebben tenslotte vaak complexe spelmechanismen die ze de spelers best stap voor stap aanleren. De volgende paragraaf is gebaseerd op werk van Gee [13] en onderzoekt welke factoren een spel leerrijk maken en wat onderwijstoepassingen daaruit kunnen leren.

Om complexere problemen op te lossen moet een student eerst de basis onder de knie hebben. Initiële oefeningen reiken telkens individuele concepten aan die een student kan veralgemenen en combineren. Zonder welbedachte volgorde of door een complexiteit die te snel stijgt, kunnen studenten zelf oplossingen vinden die achteraf niet veralgemeenbaar blijken. Een goed model om expertise te behalen binnen een domein heet de *cycle of expertise*. Eerst leren mensen één bepaalde vaardigheid door steeds dezelfde soort problemen op te lossen en pas nadat ze de vaardigheid beheersen geeft het model een nieuw soort probleem. Om dat nieuwe probleem aan te pakken komt de oude vaardigheid in een nieuw daglicht en moet de persoon beide vaardigheden verenigen. Daarna oefenen ze de nieuwe vaardigheid weer in en de cyclus herhaalt zich [47, 13].

2.6 Open problemen in de literatuur

Gamificatie in een onderwijscontext is geen ongekende toepassing in de literatuur [48], maar studies focussen vaak op volwassenen zoals universiteitsstudenten [49, 50, 51, 52]. Het aantal studenten dat beschikt over universitaire wiskundekennis is in delen van de ontwikkelde wereld sinds de jaren 80 langzaam afgenomen [53, 54]. Dat wekt zorgen over het aanbod van bekwaam personeel in het STEM-domein van *science, technology, engineering en mathematics* [55]. Gamificatie die zich toespits op middelbare scholieren lijkt dus relevanter dan ooit. De literatuur bevestigt daarbij het belang om gamificatie op basis van de eindgebruiker personaliseren [56, 57, 58]. Het blijkt echter dat het feitelijke gedrag van mensen sterk afwijkt van theoretische veronderstellingen en suggesties van experts [38]. Gepersonaliseerde gamificatie is dus nog lang geen opgelost probleem.

Misschien nog belangrijker in een onderwijscontext is de personalisatie van de leerstof [59]. Om oefeningen effectief te maken moet hun moeilijkheid geschikt zijn voor een individuele leerling. Verschillende studies gebruikten het Elo-systeem om die moeilijkheid van oefeningen te modelleren [42, 43, 44]. Er is echter weinig onderzoek rond het modelleren van individuele concepten van leerstof. Door Elo-scores op te stellen voor elk concept laat een systeem toe dat sommige concepten moeilijker zijn voor sommigen dan voor anderen. Oefeningen volgen zo niet langer één absolute rangschikking, maar vallen in een gepersonaliseerde rangschikking afhankelijk van het individu.

Zelfs al weet een systeem welke oefeningen het best overeenkomen met het niveau van een gebruiker en het effectiefst zijn, dan blijft er nog steeds de vraag hoe het die oefeningen moet aanbrenge. Door gebruikers te verplichten om telkens één welbepaalde oefening te maken, daalt hun vrijheid, wat een negatief effect kan hebben op hun intrinsieke motivatie [14]. Nudging daarentegen behoudt die vrijheid en kan toch gebruikers overtuigen om gewenst gedrag te vertonen, ook in onderwijstoepassingen [60, 61, 24]. Nudging in een onderwijscontext is echter een heel recente ontwikkeling in de literatuur en werk over nudging om effectieve oefeningen aan te bevelen ontbreekt nog.

Hoofdstuk 3

Methode

Dit hoofdstuk overloopt de opzet en het verloop van het onderzoek bestaande uit de onderzoeksvragen geformuleerd in Paragraaf 3.1. Het onderzoek verliep volledig online via de digitale onderwijsapplicatie Alchemath, die gamificatie toepast in de vorm van beloningen. Paragrafen 3.2 en 3.3 beschrijven de app en het bijhorende aanbevelingssysteem in meer detail, Paragraaf 3.4 gaat in op de rekrutering van deelnemers en het verloop van de studie en ten slotte vat Paragraaf 3.5 samen hoe de verzamelde gegevens een antwoord kunnen bieden op de gestelde onderzoeksvragen.

3.1 Onderzoeksvragen

Dit onderzoek tracht antwoorden te bieden op de open problemen in de literatuur uit Paragraaf 2.6. Specifiek stelt het twee ruime onderzoeksvragen (OV), met telkens een aantal specifiekere deelvragen.

OV 1. Hoe kan het gedrag van gebruikers aangewend worden om hun Hexad-type te bepalen?

OV 1.1 Is er een verband tussen het Hexad-type van gebruikers op basis van hun gebruiksgedrag en het type op basis van de pre-studie vragenlijst?

OV 1 gaat enerzijds na of de correlaties tussen Hexad-types en voorkeuren voor spelelementen die Altmeyer et al. [35] vonden herhaalbaar zijn in een wiskundecontext, waarbij deelnemers meer tijd krijgen. Gamificatiesystemen veronderstellen verbanden tussen Hexad-types en spelelementen, maar de literatuur is het niet altijd eens over die verbanden. Als er duidelijke correlaties zijn tussen Hexad-types en voorkeuren voor spelelementen, kunnen gamificatiesystemen die voorkeuren doorlopend monitoren om zo Hexad-profielen op te stellen zonder vervelende vragenlijsten.

OV 2. Hoe kan nudging toegepast worden om gebruikers te laten kiezen voor moeilijke oefeningen?

OV 2.1. Welke vorm van nudging, verhoogde beloningen of moeilijkheidslabels, zorgt voor een verschil in leerefficiëntie?

OV 2.2. Welke vorm van nudging, verhoogde beloningen of moeilijkheidslabels, zorgt voor een verschil in zelf-gerapporteerde competentie?

OV 2 gaat na of nudging een interessante techniek is om de doeltreffendheid van digitale onderwijssystemen te verbeteren. Het onderzoekt het principe van optioneel leren, wat de vrijheid en de autonomie van gebruikers verhoogt en daarmee de motivatie om wiskundeoefeningen te maken. De oefeningen die het systeem waardevol acht om te nudgen vallen binnen de *zone of proximal development* en zouden dus het meest efficiënt moeten zijn om de leerstof aan te brengen.

Daarnaast is het ook interessant om te kijken naar het de ervaring van de gebruikers zoals in OV 1.3. Dat functioneert ook als een vorm van validatie van de kwantitatieve resultaten.

3.2 Alchemath: gemotiveerd oefenen op maat

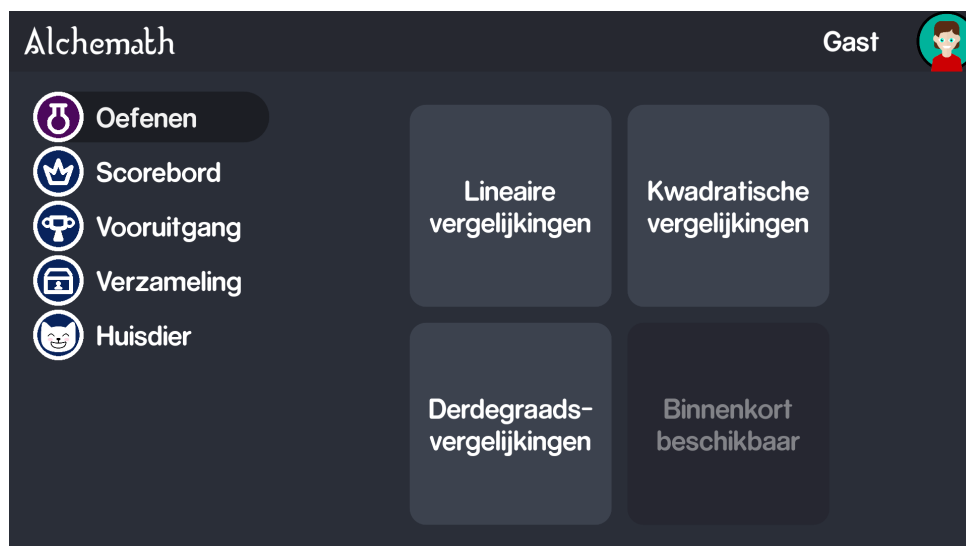
Het onderzoek gebeurde volledig via de digitale onderwijsapplicatie Alchemath (Figuur 3.1), dat speciaal voor deze studie gebouwd is. De applicatie biedt oefeningen aan in drie modules: lineaire, kwadratische en kubische vergelijkingen. Die leerstof komt aan bod in de eerste en tweede graad van het middelbaar onderwijs. De app genereert oefeningen *on the fly* op basis van een reeks templates en past ze aan het niveau van de individuele gebruikers aan. Het biedt ook ondersteuning om vergelijkingen stap voor stap op te lossen en geeft gepaste feedback op de invoer. Om motivatie tijdens het oefenen te verhogen biedt de app ook vier gamificatie-elementen aan in de vorm van beloningen met bijhorende interfaces, zoals te zien in Tabel 3.2.

3.2.1 Automatisch gegenereerde oefeningen

Om een bijna eindeloze hoeveelheid oefeningen te voorzien, genereert Alchemath ze *on the fly* op basis van de templates in Tabel 3.1. Er zijn twee groepen templates: algemene en specifieke modules. De algemene templates komen in elke module voor en vereisen simpele rekenkunde. De andere templates komen enkel voor in bepaalde modules en vereisen specifieke formules of regels. Om vergelijkingen van het type $X^2 + bX + c = 0$ op te lossen, moeten gebruikers de zogenaamde *wortelformule* kennen:

$$x = \frac{-b \pm \sqrt{D}}{2a},$$

waarbij de discriminant $D = b^2 - 4ac$. Willen ze vergelijkingen van het type $X^3 + aX^2 + bX + c = 0$ oplossen, dan moeten ze de regel van Horner kennen. Behalve een overzicht



Figuur 3.1: Titelscherm van Alchemath, de digitale onderwijsapplicatie ontwikkeld voor deze studie.

van de *wortelformule* bood Alchemath geen manier om die concepten te leren, aangezien de app focust op gekende leerstof inoefenen. Tabel 3.1 toont ook de concepten die nodig zijn om elk template op te lossen. Eén concept dat in het overzicht ontbreekt is *breuken*, wat zich niet manifesteert in de templates zelf, maar in de parameterwaarden a , b en c . Door vergelijkingen te categoriseren op basis van templates kan de app gemakkelijk oefeningen kiezen met de concepten waar een individuele gebruiker moeite mee heeft.

De app kiest de waarde van X , de oplossing van de vergelijking, willekeurig uit een uniforme verdeling waarbij

$$0 \leq X \leq M.$$

waarbij de maximale waarde M lager ligt bij vergelijkingen van hogere graden om het rekenwerk minimaal te houden. Een bijkomende vereiste is dat $X \neq 1$ in $aX = c$, omdat dat het antwoord triviaal zou maken.

Zoals eerder vermeld zijn de waardes a , b en c ofwel gehele getallen ofwel breuken. Het systeem kiest die getallen willekeurig uit de uniforme verdeling $[1, 20]$. Merk op dat het telkens één waarde moet bepalen die vastligt op basis van de gekozen waarden om de vergelijking te laten opgaan.

3.2.2 Feedback bij oplossen

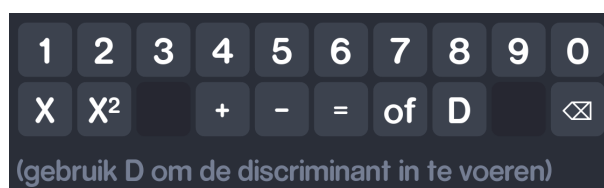
Om vergelijkingen op te lossen moeten gebruikers symbolen invoeren via het voorziene virtuele toetsenbord, zie Figuur 3.2. Door op controleren te drukken analyseert de app de invoer en reageert ze gepast. Onderstaande lijst toont de soorten invoer die de app onderscheidt en Figuur 3.3 toont hoe ze de feedback meedeelt aan de gebruikers.

1. **Getalwaarde:** de invoer bevat enkel een getalwaarde, die de app vergelijkt met de oplossing van de vergelijking. Gebruikers krijgen drie pogingen om het juiste antwoord in te voeren, daarna verschijnt het automatisch.

Tabel 3.1: De verschillende templates van vergelijkingen gebruikt om oefeningen te genereren in Alchemath. In de algemene templates kan de macht van X aangepast worden om te voldoen aan vergelijkingen van hogere graad. De onderste twee templates zijn uniek voor respectievelijk kwadratische en kubische vergelijkingen.

Template	Vereiste concepten
Algemeen	
$X + b = c$	optellen
$aX = c$	vermenigvuldigen
$aX + b = c$	optellen, vermenigvuldigen
$-(X + b) = c$	optellen, tekenverandering
$-(aX + b) = c$	optellen, vermenigvuldigen, tekenverandering
$a(X + b) = c$	optellen, vermenigvuldigen, distributiviteit
Specifieke modules	
$X^2 + bX + c = 0$	<i>wortelformule</i>
$X^3 + aX^2 + bX + c = 0$	regel van Horner

2. **Tussenstap:** als de invoer syntactisch correct is en een gelijkteken bevat, zal het systeem controleren of de ingevoerde vergelijking dezelfde oplossingen heeft als de opgave. Als dat het geval is verschijnt de invoer rechtsboven onder *tussenstappen*. Als de invoer bij tweedegraadsvergelijkingen een D en een gelijkteken bevat, interpreteert ze de invoer als de waarde van de discriminant die nodig is voor de *wortelformule*. We hebben bewust gekozen om de opgave niet steeds te updaten bij elke tussenstap om duidelijk het traject te tonen en toe te laten om verder te werken van voorgaande stappen. Wel was er een knop om alle stappen te wissen.
3. **Foute syntaxis:** de invoer mist een gelijkteken, X-teken of bevat een andere technische fout. De app meldt het probleem en beoordeelt de invoer als juist noch fout.



Figuur 3.2: Het virtuele toetsenbord om oplossingen in te voeren in Alchemath.

3.2.3 Gamificatie

In Alchemath krijgen gebruikers beloningen door oefeningen op te lossen. Die vorm van gamificatie probeert de motivatie om oefeningen te maken te verhogen. In totaal zijn er vier soorten beloningen die overeenkomen met de Hexad-gebruikertypes uit Tabel 3.2. Door beloningen te voorzien voor verschillende types is er een hogere kans dat er tenminste één spelelement tussenzit waar gebruikers van houden en dus motivatie uithalen. Hoewel



Figuur 3.3: Feedback op verschillende soorten invoer in Alchemath. Van boven naar beneden: foute syntaxis, fout antwoord, juist antwoord.

het scorebord bedoeld is voor *Socialisers*, heeft onderzoek ook al aangetoond dat het positief gecorreleerd is aan *Players*, *Achievers* en *Disruptors* [35].

Tabel 3.2: Overzicht van alle soorten beloningen in Alchemath, hun bijhorende interfaces en hun overeenkomstige Hexad-typen.

beloning element beschrijving	Sterren Scorebord Het scorebord rangschikt alle gebruikers tot dan toe op het aantal sterren dat ze verdiend hebben. Gebruikers kunnen enkel de vier personen boven en onder hen zien, om motivatie te bevorderen op kleine schaal. Ook verschijnt er een aanmoedigend zinnetje wanneer gebruikers maar drie sterren nodig hebben om een plaats te stijgen of de persoon onder hen drie sterren nodig heeft om hen in te halen.
bedoeld voor	<i>Socialiser</i> 
beloning element beschrijving	Munten Badges Er zijn vijf levels van badges met bijhorende titels en motiverende beschrijvingen, zoals te zien in Tabel B.1. Gebruikers moeten munten verzamelen om hun badgelevel te verhogen. De titels van de levels komen overeen met de getallenverzamelingen $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ en $\mathbb{C}$.
bedoeld voor	<i>Achiever</i> 
beloning element beschrijving	Schatkisten Verzameling Een schatkist genereert één of twee van de 24 mogelijke verzamelobjecten in Tabel B.2, afhankelijk van de grootte van de beloning. Die objecten komen voor met verschillende kansen: veelvoorkomend ($\frac{12}{17}$), zeldzaam ($\frac{4}{17}$) en legendarisch ($\frac{1}{17}$). Elk verzamelobject heeft ook een beschrijving die past bij zijn afbeelding met leuke tekstjes over wiskundige begrippen.
bedoeld voor	<i>Free Spirit</i>



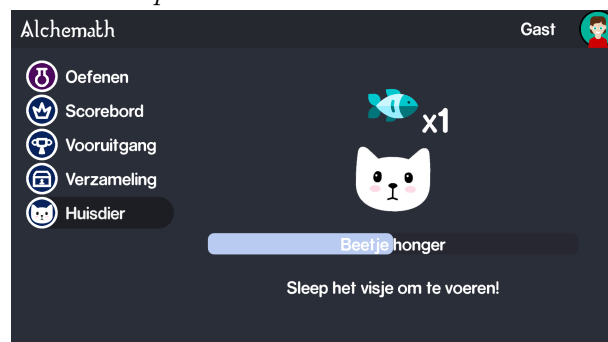
beloning
element
beschrijving

Visjes Huisdier

Om de groeiende honger van een virtuele kat tegen te gaan moeten gebruikers visjes verzamelen en aan haar voeren door te slepen. De gezichtsuitdrukking van de kat verandert zoals in Tabel B.3 te zien naarmate haar honger vermindert van wenend naar lachend, wat een gevoel van zorg opwekt. Ze heeft ook bijhorende animaties zoals op en neer springen om haar reacties zo echt mogelijk te maken. Haar honger kan variëren tussen $[0, 4]$, stijgt met 1 per visje en groeit automatisch met een willekeurige waarde tussen $[.5, 1]$ wanneer gebruikers een nieuwe oefening starten, om te verzekeren dat de visjes waardevol blijven.

bedoeld voor

Philanthropist



3.3 Oefeningen aanbevelen

Een belangrijke taak van de app is om het niveau van gebruikers te modelleren en op basis daarvan oefeningen aan te bieden. Oefeningen moeten uitdagend zijn, maar niet frustrerend; ze moeten liggen in de *zone of proximal development*. Door de toegespitste leerstofafbakening was het gemakkelijk om de vereiste concepten voor verschillende oefeningen te bepalen, zoals besproken in Paragraaf 3.2.1.

Het niveau van gebruikers is gerelateerd aan de moeilijkheidsgraad van de verschillende concepten, dus het systeem moet inschatten hoe goed gebruikers elk concept begrijpen. Daarvoor maakt het gebruik van Elo-rating. Gebruikers krijgen eerst een persoonlijke Elo-score van 1000 toegekend voor elk concept per module. Die scores zeggen hoe goed de gebruiker elk concept begrijpt. Om ze te interpreteren zal het systeem ze vergelijken met overkoepelende Elo-scores die alle gebruikers delen. Idealiter zijn die overkoe-

pelende Elo-scores een schatting van de absolute moeilijkheid van de concepten, maar daarvoor moeten genoeg gebruikers oefeningen rond die concepten maken zodat de scores convergeren. Bij aanvang van het onderzoek hebben we uit gebrek aan onderzoek over concept-gebaseerde Elo-scores manueel een inschatting gemaakt van de moeilijkheid met advies van een wiskundeleerkracht. Tabel C.1 toont zowel de initiële als de convergeerde waarden.

Als een gebruiker oefeningen wilt maken, genereert het systeem telkens vier opties met verschillende slaagkansen op basis van de persoonlijke en overkoepelende Elo-scores: 40%, 50%, 60%, en 70%. De oefeningen met kansen 40% en 50% vallen in de *zone of proximal development*: ze zijn moeilijk, maar haalbaar. Daarom zijn dat de oefeningen die het systeem aanraadt en probeert te *nudgen*, zoals Paragraaf 3.4.2 zal verduidelijken. Het systeem berekent de slaagkansen voor elk mogelijk template binnen de gekozen module, waarbij de Elo-score van een template overeenkomt met de hoogste Elo-score van zijn concepten. De kans S dat een gebruiker met persoonlijke Elo-score P een oefening juist oplost met overkoepelde Elo-score G is

$$S = \frac{1}{1 + 10^{(G-P)/400}}.$$

Uit alle templates kiest het systeem voor elke slaagkans het template dat er het dichtste bij ligt. Het systeem kan natuurlijk niet altijd templates kiezen met de exacte slaagkansen, omdat het aantal templates beperkt is.

3.4 Deelnemers en verloop van de studie

Deze paragraaf beschrijft de opzet van het experiment, die werd goedgekeurd door de Sociaal-Maatschappelijke Ethische Commissie (SMEC), referentienummer G-2023-6197.

3.4.1 Rekrutering

Aangezien de beschikbare leerstof op Alchemath aan bod komt in de eerste en tweede graad van het middelbaar onderwijs, is het vanzelfsprekend dat leerlingen uit die graden deel uitmaken van de doelgroep. We hebben echter ook deelnemers buiten die doelgroep toegelaten, omdat we dan verschillen in leeftijd konden bestuderen. Studenten of volwassenen die weinig of geen wiskunde nodig hebben voor hun opleiding of job verleren het vak bovendien snel en halen er voordeel uit om hun vaardigheden scherp te houden.

We rolden Alchemath enerzijds uit op Android via een *gesloten testtrack*. Om de app te downloaden moesten deelnemers eerst toegang krijgen door te lid worden van een testgroep op *Google Groups*. Anderzijds rolden we de app ook uit op het web via *Github Pages*. Om leerlingen te rekruteren stuurden we e-mails naar alle middelbare scholen in Vlaams-Brabant en Antwerpen met screenshots van de app. Naar leerkrachten die geïnteresseerd reageerden stuurden we links voor beide versies van de applicatie.

Op de ca. 300 verstuurd e-mails antwoordden ca. 30 geïnteresseerde leerkrachten, waarvan er uiteindelijk 10 bevestigden en effectief deelnamen aan het onderzoek. Er waren uiteindelijk 174 deelnemers, waarvan 141 leerlingen in klasverband.

3.4.2 Gerandomiseerd onderzoek met controlegroep

Deelnemers werden automatisch en stilzwijgend ingedeeld in één van vier testgroepen. Om de testgroepen ongeveer even groot te maken was de kans op indeling voor elke testgroep 25%. Ze verschilden van elkaar op twee vlakken:

1. **Verhoogde beloningen:** het systeem maakt de beloningen voor de oefeningen met slaagkansen van 40% en 50% twee keer groter dan die met slaagkansen 60% en 70%. Deze vorm van *incentive nudging* tracht gebruikers te doen kiezen voor de oefeningen met lagere slaagkansen die liggen in de *zone of proximal development*. In testgroepen zonder verhoogde beloningen is de grootte van de beloning voor alle oefeningen één.
2. **Moeilijkheidslabels:** het systeem informeert de gebruiker over de moeilijkheid van elke oefening aan de hand van een 4-punts schaal met gekleurde bolletjes. Het aantal gekleurde bolletjes komt overeen met de slaagkansen van de oefeningen: één voor 70%, twee voor 60%, drie voor 50% en vier voor 40%. Zonder moeilijkheidslabels kan de gebruiker enkel afgaan op de opgave om te moeilijkheid te bepalen.

De variatie van die twee variabelen komt overeen met vier mogelijke testgroepen, samengevat in Figuur 3.4.



Figuur 3.4: De verschillen tussen keuzeschermen op basis van verhoogde beloningen en moeilijkheidslabels die de vier testgroepen onderscheidden.

3.4.3 Verloop

Deelnemers downloadden of openden de app via de voorziene link op Android of in hun browser. Om aan het onderzoek deel te nemen en toegang te krijgen tot Alchemath moesten deelnemers eerst toestemming geven; dat konden ze doen op papier of digitaal

in de app zelf. Leerlingen die vijftien jaar of jonger waren hadden ook schriftelijke toestemming van hun ouders nodig. Leerkrachten in klasverband moesten een bijkomend formulier ondertekenen om te bevestigen dat ze hun leerlingen niet onder druk zouden zetten en oefeningen op papier zouden voorzien voor leerlingen die niet deelnamen aan het onderzoek.

Pre-studie vragenlijst

Deelnemers moesten één voor één de vragen van de Hexad-24 vragenlijst [31] beantwoorden om toegang te krijgen tot de inhoud van de applicatie. Die valideerde vragenlijst is een Nederlandse vertaling door Ooge [62] van 24 vragen op een 7-punts Likertschaal die aanleiding geeft tot de Hexad-types van de gebruikers. Tabel A.1 toont de volledige vragenlijst, inclusief een extra controlevraag met de instructie “Duid nu ‘Oneens’ aan om te bewijzen dat je nog wakker bent”. Deelnemers die deze vraag fout beantwoordden namen we niet in de analyse op omdat ze de vragenlijst hoogstwaarschijnlijk onaanachtig invulden.

Oefeningen maken

Oefeningen waren onderverdeeld in drie modules: lineaire, kwadratische en kubische vergelijkingen. Deelnemers waren vrij om een module te kiezen op basis van hun niveau en konden in de gekozen module zo veel oefeningen maken als ze wilden. Op het keuzescherf verschenen telkens vier mogelijke oefeningen van verschillende moeilijkheden zoals vermeld in Paragraaf 3.3. Om bias te vermijden wanneer gebruikers telkens kiezen voor oefeningen op dezelfde plaats, toonde de app op elke positie telkens een willekeurige moeilijkheid. Om bias te vermijden op basis het uiterlijk van de virtuele avatars, kregen de vier keuzes telkens dezelfde avatar.

Post-studie vragenlijst

De post-studie verscheen automatisch na vijftien opgeloste oefeningen en gebruikers konden verder oefeningen maken nadat ze de vragenlijst hadden ingevuld. Op die manier was de vragenlijst geïntegreerd met de rest van het onderzoek en was de kans groot dat deelnemers de lijst zouden zien en invullen. Zoals te zien in Tabel A.2 bestond die vragenlijst uit tien vragen. De eerste vier vragen polsten naar de gebruikers ervaring van de verschillende gamificatie-elementen en toonde screenshots van de bijhorende schermen zoals in Tabel 3.2. Daarna kregen de deelnemers vijf vragen uit de *perceived competence* subcluster van de *Intrinsic Motivation Inventory* [63], vertaald naar het Nederlands door Peeters [64]. De laatste vraag ging over waar gebruikers op letten wanneer ze oefeningen kozen, bijvoorbeeld de opgave, de beloning en de moeilijkheid. Merk op dat beloning en moeilijkheid enkel van toepassing waren op de testgroepen met verhoogde beloningen en moeilijkheidslabels respectievelijk. De laatste vraag voorzag ook open tekstvelden waarin gebruikers verder konden verduidelijken waar ze op letten en waarom.

3.4.4 Loggegevens

Naast de data verkregen uit de vragenlijsten logde de app ook gebruikerinteracties. Door deze kwantitatieve data te toetsen aan de menselijke feedback uit de post-studie vragen-

lijst, verhoogt de betrouwbaarheid van correlaties en patronen in de analyse.

- **Gamificatie:** de totale tijd dat elk van de vier gamificatieschermen te zien was tijdens het gebruik. Interactie met de schermen geeft aan welke beloningen gebruikers waardevol vonden.
- **Oefeningpogingen:** elke gestarte oefening, inclusief afgebroken oefeningen. Voor elke oefening logde de app de onderstaande gegevens. Het beloningstype kan vergeleken worden met een gebruikers Hexad-profiel en Elo-scores zijn nodig om leerefficiëntie te berekenen. Andere gegevens maken onder andere de analyse eenvoudiger.
 - De starttijd
 - De verstreken tijd
 - Het aantal pogingen
 - Of de oefening was opgelost
 - De plaats van de oefening op het keuzeschermb: linksboven, rechtsboven, links-onder of rechtsonder
 - De opgave
 - De relatieve moeilijkheidscore
 - Het beloningstype
 - De gebruikers Elo-scores

3.4.5 Overige gegevens

Naast de reeds vermelde data gaven een aantal deelgenomen leerkrachten ook feedback via e-mail na afronding van het onderzoek.

3.5 Hypotheses en analyse

Deze paragraaf beschrijft onze hypotheses over de onderzoeksvragen en de methodes waarmee we de verschillende data hebben geanalyseerd om antwoorden te bieden op die vragen. We bespreken elke onderzoeksvraag in een aparte paragraaf.

Om kwantitatieve resultaten te berekenen hebben we gebruikgemaakt van *RStudio* en voor statistische toetsen gebruikten we steeds significantieniveau $\alpha = 0.05$. Voor OV 2, OV 2.1 en OV 2.2 gebruikten we Welch t-toetsen. Deze variant van de t-toets gaat er niet vanuit dat de varianties tussen twee steekproeven gelijk zijn. Als de steekproeven niet normaal verdeeld zijn kunnen t-toetsen misleidende resultaten geven, maar niet in het geval de steekproeven groot genoeg zijn. Aangezien elk mogelijk paar testgroepen minstens 85 deelnemers bevat gaan we de normaliteit van de steekproeven dus niet na.

OV 1.1 Is er een verband tussen het Hexad-type van gebruikers op basis van hun gebruiksgedrag en het type op basis van de pre-studie vragenlijst?

We verwachtten de verbanden terug te zien uit Tabel 3.2 die Altmeyer et al. [35] ook vonden.

Om de Hexad-types van gebruikers op te stellen gebruikten we de standaardprocedure van Hexad-24 [31]. De voorkeuren bepaald door die profielen hebben we vergeleken met

de beloningen die gebruikers kozen, hoelang gebruikers naar de verschillende gamificatieschermen keken en hun zelf-gerapporteerde voorkeuren voor gamificatie-elementen uit de post-studie vragenlijst.

OV 2. Hoe kan nudging toegepast worden om gebruikers te laten kiezen voor moeilijke oefeningen?

We verwachtten dat verhoogde beloningen voor oefenen ervoor zou zorgen dat gebruikers meer geneigd zijn voor oefeningen, zelf als die oefeningen moeilijker zijn dan de oefeningen met normale beloningen.

Om OV 2 te beantwoorden hebben we de verhouding moeilijke oefeningen tot het totaal aantal gemaakte oefeningen vergeleken tussen testgroepen. We gebruikten om enerzijds eenzijdige Welch t-toetsen om testgroepen met en zonder verhoogde beloningen te vergelijken en anderzijds tweezijdige Welch t-toetsen voor testgroepen met en zonder moeilijkheidslabels. We kozen voor eenzijdige Welch t-toetsen omdat we sterk vermoedden dat verhoogde beloningen gebruikers overtuigt om te kiezen voor moeilijke oefeningen.

We hebben bijkomend gekeken naar factoren waarvan gebruikers in de post-studie vragenlijst aangaven dat ze hen beïnvloedden om oefeningen te kiezen.

OV 2.1. Welke vorm van nudging, verhoogde beloningen of moeilijkheidslabels, zorgt voor een verschil in leerefficiëntie?

We verwachtten dat in de testgroepen met verhoogde beloningen gebruikers meer gemotiveerd waren dan wanneer de beloningen voor alle oefeningen gelijk was. We verwachtten ook dat gebruikers meer hun best deden voor die grotere beloning, wat voor een hoger slaagpercentage zou zorgen en dus de gemiddelde Elo-score zou verhogen.

Leerefficiëntie (LE) hebben we gemeten aan de hand van de totale winst in persoonlijke Elo-scores, gesommeerd over alle modules en concepten:

$$LE = \sum_{\text{module}} \sum_{\text{concept}} (Elo_{\text{end}} - Elo_{\text{start}}).$$

Zoals in Paragraaf 3.3 vermeld is Elo_{start} voor alle gebruikers voor alle modules en concepten gelijk aan 1000; de uiteindelijke formule is dus:

$$LE = \sum_{\text{module}} \sum_{\text{concept}} (Elo_{\text{end}} - 1000).$$

Om OV 2.1 te beantwoorden hebben we de leerefficiëntie van deelnemers tussen de vier testgroepen vergeleken. We gebruikten tweezijdige Welch t-toetsen om enerzijds testgroepen met en zonder verhoogde beloningen te vergelijken en anderzijds testgroepen met en zonder moeilijkheidslabels.

OV 2.2. Welke vorm van nudging, verhoogde beloningen of moeilijkheidslabels, zorgt voor een verschil in zelf-gerapporteerde competentie?

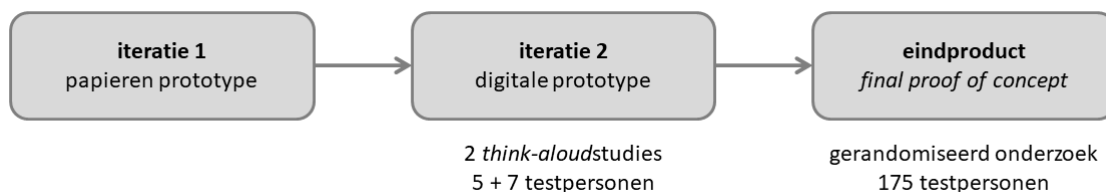
We verwachtten dat de testgroepen met moeilijkheidslabels hun eigen competentie hoger inschatten omdat ze bewust waren van het niveau van de gekozen oefeningen.

De zelf-gerapporteerde competentie van gebruikers bepaalden we aan de hand van hun antwoorden op de *perceived competence* subcluster van de IMI-vragenlijst die deelt uitmaakte van de post-studie vragenlijst. We gebruikten tweezijdige Welch t-toetsen om enerzijds testgroepen met en zonder verhoogde beloningen te vergelijken en anderzijds testgroepen met en zonder moeilijkheidslabels.

Hoofdstuk 4

Ontwikkelingsproces

Dit hoofdstuk beschrijft de verschillende stappen in de ontwikkeling van de digitale onderwijsapplicatie Alchemath, zie Figuur 4.1. Hoewel de app slechts een middel is om de onderzoeksvragen te beantwoorden, is het ook een volwaardig product dat kan blijven bestaan nadat deze thesis is afgerond. Aan de hand van twee gebruikersstudies probeerden we de bruikbaarheid van de app te verzekeren, zodat ze het onderzoek niet ondermijnde. Figuur 4.2 toont de ontwikkeling van papieren prototype naar finaal ontwerp aan de hand van een aantal schermen.



Figuur 4.1: De stappen in de ontwikkeling van de digitale onderwijsapplicatie Alchemath.

4.1 Papieren prototype

Het initiële ontwerp van Alchemath leek meer op een *serious educational game* dan op een *gamified* applicatie en is te zien in Figuur 4.1. Gebruikers zouden in die versie vergelijkingen oplossen door getallen te slepen en neer te zetten, ook bekend als *drag and drop*. Die manier van oplossen heeft als voordeel dat gebruikers elke oplossingsstap uitvoeren en daardoor een dieper begrip ontwikkelen van vergelijkingen. Het is echter enkel een voordeel voor leerlingen die voor het eerst kennismaken met vergelijkingen; voor de meeste gebruikers zijn verplichte tussenstappen die ze uit het hoofd kunnen doen eerder een belemmering. Nagashima et al. [65] beschreven in hun paper een software die het *drag and drop*-principe succesvol hanteert voor vergelijkingen.

4.2 Digitale prototype

Het digitale prototype is de vertaling van het papieren prototype naar een volledig functionerende applicatie en is te zien in Figuur 4.1. Om de bruikbaarheid van de app te verzekeren verbeterden we het prototype tijdens twee *think-aloud* studies, waarbij testpersonen een aantal vragen kregen en moesten navigeren door de interface van de app zonder context.

4.2.1 Eerste think-aloud

De eerste *think-aloud* studie is samengevat in Tabel 4.1 en telde vijf deelnemers: één middelbare leerling, drie universitaire studenten en een volwassene met een middelbaar diploma. Een belangrijk aandachtspunt was de flow van de app; gebruikers vonden dat ze te veel moesten klikken om na een opgeloste oefening weer een nieuwe te starten. Ook de symbolen van de beloningen waren niet allemaal even duidelijk.

Tabel 4.1: Feedback tijdens de eerste *think-aloud* van het digitale prototype.

		Aantal
	Hoe goed kunnen gebruikers de symbolen van de beloningen linken aan hun betekenissen?	
feedback	Symbolen voor munt, trofee, cadeau en verassing overlappen in betekenis.	2/5
actie	Symbolen herwerken om het onderscheid te verduidelijken.	
	Hoe gemakkelijk kunnen gebruikers oefeningen starten en oplossen?	
feedback	Het keuzeschermbetekenis bevat te veel informatie.	2/5
actie	Contrast verhogen van essentiële informatie zoals de opgave en de beloning.	
feedback	De flow van oefeningen kiezen en oplossen is te traag.	3/5
actie	Bevestigingsstap laten vallen bij het keuzeschermbetekenis en na de eerste klik meteen de gekozen oefening starten. Het beloningsschermbetekenis integreren met het oplossingsschermbetekenis in de vorm van een pop-up.	
feedback	Gebruikers verwachten dat ze na een oefening gekozen en gemaakt hebben, kunnen verdergaan met de andere oefeningen die op het keuzeschermbetekenis stonden, maar het scherm toont telkens vier nieuwe.	2/5
actie	De werking van het keuzeschermbetekenis verduidelijken met het zinnetje “Volgende keer krijg je 4 nieuwe oefeningen” bovenaan het scherm.	

4.2.2 Tweede think-aloud

De tweede *think-aloud* is samengevat in Tabel 4.2 en telde zeven deelnemers, waaronder twee middelbare leerlingen en vier universitaire studenten. In deze iteratie konden ze het soort beloning kiezen onafhankelijk van de gekozen oefening, konden ze tussenstappen invoeren om vergelijkingen op te lossen en verscheen automatisch de post-studie vragenlijst na een aantal oefeningen. Er waren ook een aantal interessante opmerkingen zoals de

voorziening van andere huisdieren dan een kat, die door tijdsgebrek buiten het kader van deze thesis vielen.

Tabel 4.2: Feedback tijdens de tweede *think-aloud* van het digitale prototype.

		Aantal
	Is de uitleg over beloningen duidelijk en begrijpen deelnemers hoe ze het soort beloning kunnen wijzigen?	
feedback	De aandacht gaat meteen naar de oefeningen, waardoor de knoppen om beloningen te wijzigen over het hoofd worden gezien.	6/7
actie	De knoppen mee tonen tijdens op het uitlegscherm en er visueel nadruk op leggen.	
	Zijn de gamificatieschermen duidelijk evenals hun verband met de beloningen?	
feedback	De knop van het huisdierscherm lijkt visueel niet te horen bij de andere knoppen. De knop om oefeningen te starten zou daarentegen juist verschillender moeten zijn.	3/7
actie	De knop van het huisdierscherm kaderen met een cirkel en de achtergrond van de knop voor oefeningen paars maken, zodat hij verschilt van de andere knoppen met een blauwe achtergrond.	
feedback	De werking van het badgescherm is niet meteen duidelijk, zelfs na de uitleg over de beloningen.	4/7
actie	De werking op het scherm zelf in woorden verduidelijken.	
feedback	Het zou leuk zijn mocht er ook een ander huisdier zijn dan een kat.	2/7
actie	Geen actie door lage prioriteit en gebrek aan tijd.	
	Begrijpen deelnemers dat ze tussenstappen kunnen invoeren om vergelijkingen op te lossen?	
feedback	De tussenstappen zouden duidelijker zijn onder de opgave dan rechtsboven.	2/7
actie	Geen actie want er is niet genoeg plaats onder de opgave.	
	Zijn de vragen van de post-studie vragenlijst duidelijk?	
feedback	De vragen zijn duidelijk.	7/7
actie	Geen actie vereist.	

4.3 Technische implementatie

We hebben de digitale applicatie Alchemath ontwikkeld in de gratis *open-source game engine* Godot 4.1 [66], in Godot's eigen programmeertaal *GDScript* dat sterk verwant is aan Python. Godot biedt gratis exportmogelijkheden naar verschillende platformen zoals HTML5 en Android die we gebruikt hebben voor deze studie.

4.3.1 Oefeningen parsen

Zoals in Paragraaf 3.2.1 vermeld, genereerde het systeem vergelijkingen op basis van een reeks templates. Om de invoer van gebruikers tijdens het oplossen te interpreteren,

maakte het gebruik van een soort *parser*. Eerst verwijderde het onbelangrijke tekens zoals spaties en splitste het de invoertekst op in antwoorden, op basis van het *of*-teken. Die ondersteuning was cruciaal voor vergelijkingen met meerdere oplossingen, zoals kwadratische vergelijkingen met een positieve discriminant. Dan liep het elk gesplitste deel af en bepaalde of het deel een getalwaarde of een tussenstap was. Getalwaarden kon het meteen vergelijken met de gezochte oplossingen, maar tussenstappen moest het eerst evalueren. Om dat te doen zette de *parser* de invoertekst om naar een logische voorstelling van een vergelijking. Als die omzetting mislukte, gaf het systeem een syntaxisfout met de boodschap “Hmm, dat lijkt geen vergelijking”. Op basis van de logische voorstelling kon het systeem de oplossing van de vergelijking berekenen en vergelijken met de invoer.

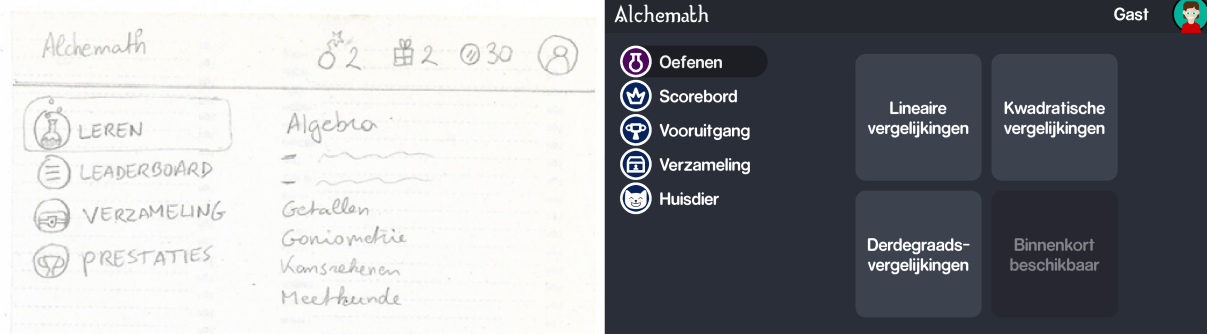
4.3.2 Verzamelen van data

De app stuurde verschillende data via HTML door naar een *RESTful* java-applicatie verbonden aan een *MYSQL*-databank, maar hield ook een lokaal *config*-bestand bij zodat de functionaliteit minimaal afhankelijk was van de *backend*. Enkel om het scorebord te tonen was een noodzakelijk om data van de backend te vragen. Zowel de java-applicatie als de databank werden gehost op de servers van KULeuven.

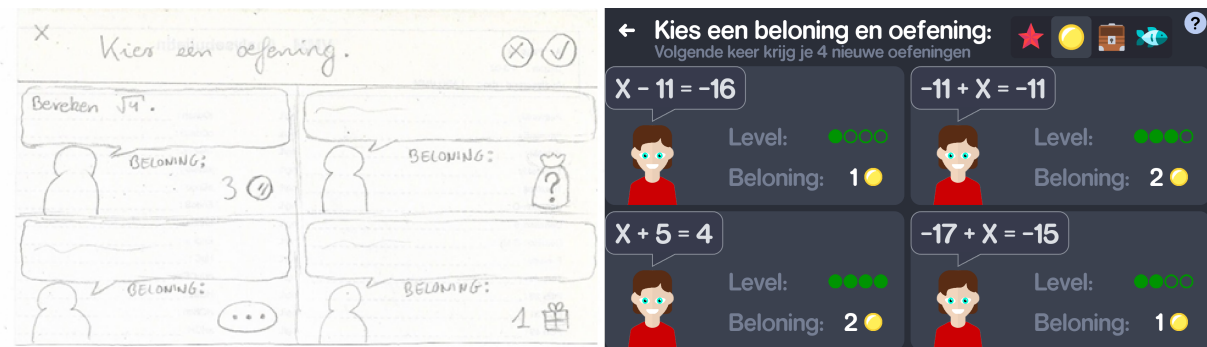
4.3.3 Publicatie

We hebben de mobiele versie van de app uitgerold voor Android op *Google Play*. *Google Play Console*, het platform dat ontwikkelaars in staat stelt om apps te publiceren en te monitoren, biedt een aantal mogelijkheden aan wat de reikwijdte van de uitrol betreft. De *productietrack* stelt een app publiek beschikbaar en laat iedereen toe om de app te vinden in *Google Play* zonder verwijzing. Voor de *gesloten testtrack* moeten gebruikers echter eerst toestemming krijgen om de app daarna via een unieke url te vinden. Ondanks die bijkomende vereiste, hebben wij gekozen om Alchemath uit te rollen op de *gesloten testtrack*. Voordat een app namelijk effectief verschijnt op de *productietrack* moeten werknemers van Google nagaan of de app conform is met de richtlijnen en overeenkomt met de opgegeven informatie rond reclame, leeftijdsclassificatie, enz. Dat proces kan dagen tot zelfs weken duren en we wilden het experiment starten zodra de finale implementatie klaar was. Apps op de *gesloten testtrack* ondergaan enkel een automatische technische controle die slechts een paar uur duurt. Ook in het geval dat we problemen met de app ondervonden na de uitrol wilden we er zeker van zijn updates vlot verliepen.

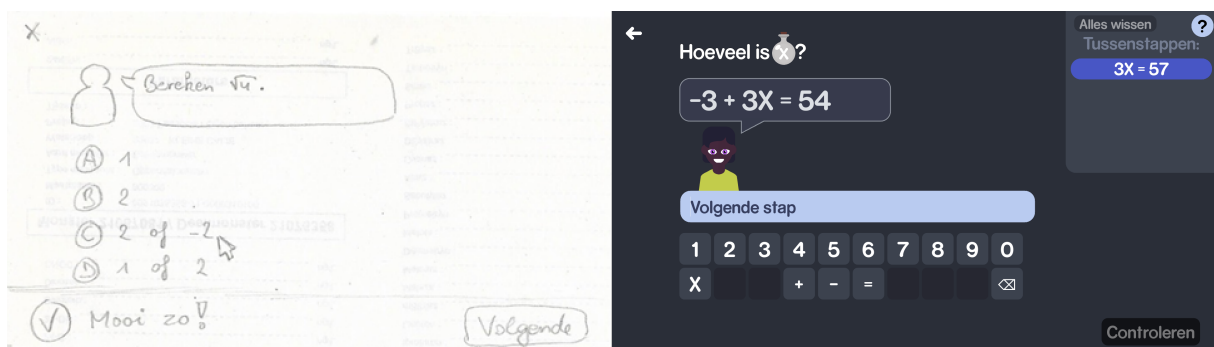
Godot vereiste dat geëxporteerde webapplicaties gehost werden op een beveiligde HTTPS-verbinding. Door technische problemen hebben we gekozen om de webversie van Alchemath te hosten via *Github Pages*.



(a) Startscreen



(b) Keuzescreen



(c) Oplossingsscreen

Figuur 4.2: Vergelijking tussen het papieren prototype en het finale ontwerp.

Hoofdstuk 5

Resultaten

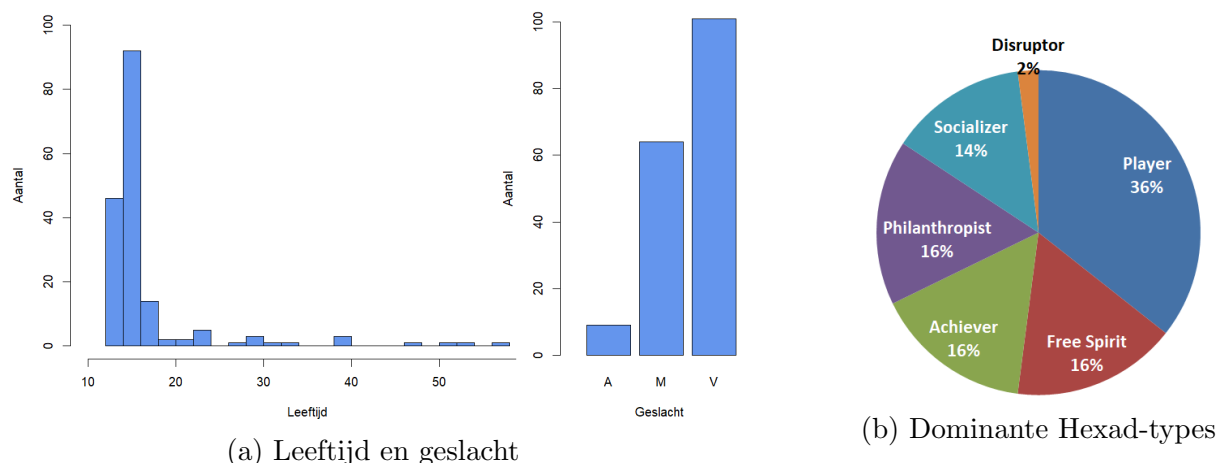
Dit hoofdstuk toont de resultaten van de 174 personen die deelnamen aan het onderzoek, waarvan 141 leerlingen in klasverband. Van het totaal aantal deelnemers hebben 146 de Hexad-24 vragenlijst geldig ingevuld, waarvan 91 ook de post-studie vragenlijst hebben ingevuld. Figuur 5.1 toont de verdeling van gebruikers op basis van leeftijd, geslacht en dominante Hexad-types; de types waar gebruikers het hoogst op scoorden. In de volgende paragrafen beperken we sommige analyses tot de 84 deelnemers die bovenop hun geldig Hexad-profiel ook minstens 15 oefeningen hebben gemaakt en minstens 30 seconden met de gamificatieschermen hebben geïnterageerd. Gebruikers zijn namelijk geneigd om bij het eerste gebruik van de applicatie elke soort beloning uit te proberen, onafhankelijk van hun Hexad-voorkeuren. Daarnaast verhoogt die voorwaarde de kans dat gebruikers de beloningen en hun verband met de gamificatieschermen begrepen; Tabel 5.3 bevestigt dat begrip.

Eerst bespreekt dit hoofdstuk enkele algemene verschillen tussen de vier testgroepen, dan wijdt het per onderzoeksvraag een paragraaf aan de resultaten die betrekking hebben op die vraag. Om kwantitatieve resultaten te berekenen hebben we gebruikgemaakt van *RStudio* en voor statistische toetsen gebruikten we steeds significantieniveau $\alpha = 0.05$.

5.1 Screening van testgroepen

Deze paragraaf behandelt enkele algemene verschillen tussen de vier testgroepen. Om de resultaten gemakkelijker te interpreteren zullen we vanaf deze paragraaf naar de testgroepen verwijzen met de codes B0/M0, B1/M0, B0/M1 en B1/M1, waarbij B en M staan voor verhoogde beloningen en moeilijkheidslabels en de waardes 0 en 1 betekenen dat die variabelen afwezig, respectievelijk aanwezig zijn. Tabel 5.1 toont het aantal deelnemers met geldige Hexad-profielen per testgroep.

Figuur 5.2 toont het totaal aantal gemaakte oefeningen per testgroep, per module. In totaal maakten deelnemers 8776 oefeningen over een totaal van 208,267 uur. Deelnemers maakten vooral oefeningen over lineaire vergelijkingen, waarschijnlijk omdat een groot deel van de deelnemers leerlingen uit de eerste graad waren en vaak de gemakkelijkste module kozen. Het valt ook op dat B0/M0 beduidend meer oefeningen maakte dan de



Figuur 5.1: Verdeling van 174 gebruikers hun leeftijd en geslacht en van 146 geldige gebruikers hun dominante Hexad-types.

Tabel 5.1: Verdeling van 174 deelnemers onder de vier testgroepen.

	Normale beloningen	Verhoogde beloningen
Geen moeilijkheidslabels	B0/M0: 43	B1/M0: 44
Moeilijkheidslabels	B0/M1: 42	B1/M1: 45

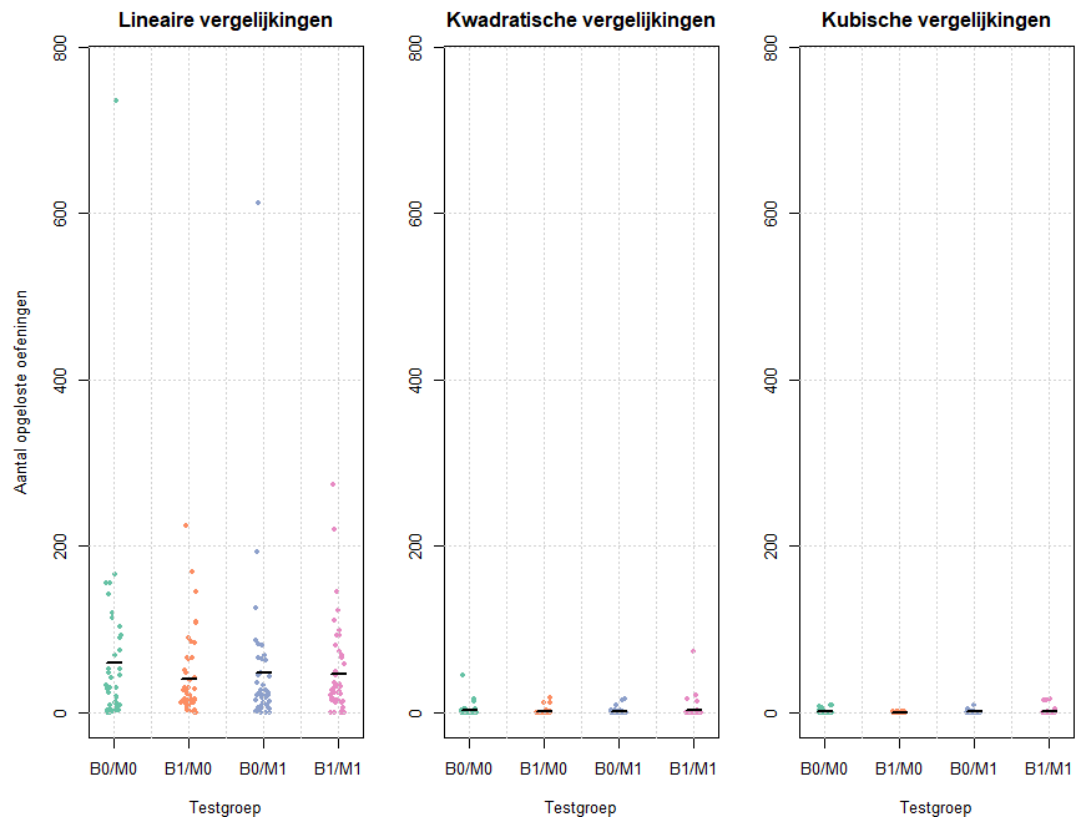
andere testgroepen. Twee gebruikers uit testgroepen B0/M0 en B0/M1 maakten respectievelijk 735 en 614 oefeningen, wat sterk boven de gebruiker op de derde plaats ligt met 274 oefeningen. Deze en andere uitschieters die meer verwijderden we uit de statistische analyse. Tweezijdige Welch t-toetsen vonden echter geen significant verschil tussen testgroepen op basis van het aantal gemaakte oefeningen.

Figuur 5.3 toont de totale tijd dat gebruikers gespendeerd hebben aan al die opgeloste oefeningen. Het valt op dat B0/M1 veel meer tijd nodig had om een gelijkaardig aantal oefeningen op te lossen dan de andere testgroepen. Ook op basis van deze data vonden tweezijdige Welch t-toetsen geen verschil tussen testgroepen.

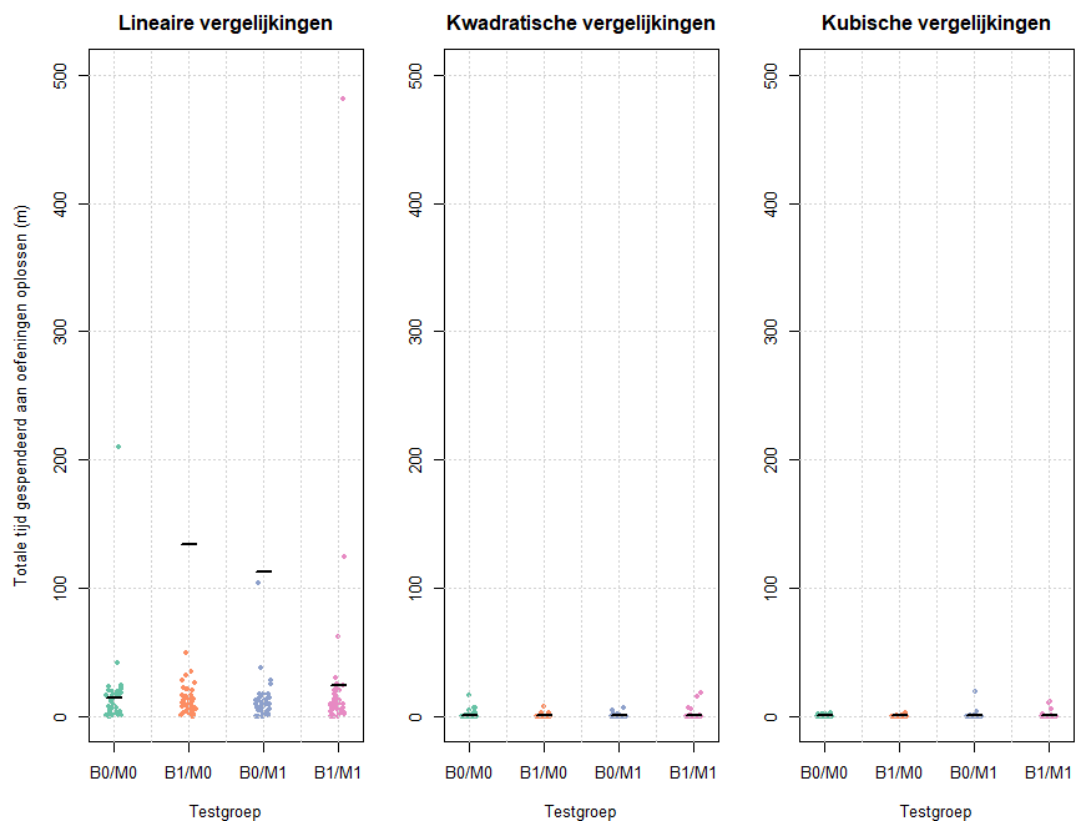
5.2 Voorkeur voor spelelementen

Zoals OV 1 stelt willen we weten of het gedrag van gebruikers een goede indicatie is om het Hexad-profiel uit af te leiden. Tabellen 5.2 tot 5.4 tonen relevante resultaten om die vraag te beantwoorden.

Allereerst valt op in Tabel 5.2 dat de zelf-gerapporteerde voorkeuren sterkere en significantere correlaties tonen dan de voorkeuren waargenomen door het systeem. Over *Socialiser* kunnen de resultaten geen enkele significante of sterke uitspraken doen. Qua beloningen is er maar één ietwat sterk positief verband: *Player*-sterren. Daarnaast zijn de sterkste negatieve verbanden *Disruptor*-munten, *Free Spirit*-kisten en *Philanthropist*-kisten. Qua gamificatieschermen is er slechts één significant ($p < 0,001$) verband: *Player*-scorebord.



Figuur 5.2: Aantal oefeningen die 174 gebruikers maakten, per testgroep per module.



Figuur 5.3: Totale tijd die 174 gebruikers gebruikten om oefeningen op te lossen, per testgroep per module.

Het op één na sterkste verband is tussen *Disruptor*-badgescherm. Qua zelf-gerapporteerde voorkeuren komen er duidelijkere verbanden aan het licht. De sterkste positieve verbanden zijn *Achiever*-verzameling, *Philanthropist*-huisdier en *Player*-scorebord. Daarnaast is *Disruptor* ook significant ($p < 0,01$) negatief gecorreleerd aan het badgescherm en de verzameling.

Tabel 5.3 toont de correlaties tussen beloningen en gamificatieschermen. De sterkste significante ($p < 0,001$) verbanden zijn ster-scorebord, munt-badgescherm, kist-verzameling en visje-huisdier.

Verder is het interessant om te kijken naar de correlaties tussen Hexad-types onderling, zoals te zien in Tabel 5.4. De sterkste positieve verbanden ($p < 0,001$) zijn *Achiever-Free Spirit*, *Achiever-Philanthropist*, *Free Spirit-Philanthropist* en *Philanthropist-Socialiser*.

Tabel 5.5 toont de correlaties tussen Hexad-types volgens Hexad-24 [31] en Hexad-12 [36]. De correlaties van alle Hexad-types met henzelf zijn significant ($p < 0,001$) en sterk positief.

5.3 Testgroepen en oefeningen kiezen

Zoals OV 2 stelt willen we weten hoe we gebruikers kunnen nudgen. Zoals Figuur 3.4 illustreerde toont de app aan elk van de vier testgroepen een ander keuzescherm. Figuur 5.4 vat relevante resultaten samen om de onderzoeksvraag te beantwoorden. Figuren 5.4a tot 5.4c tonen respectievelijk de verhoudingen van geslaagde oefeningen, gekozen moeilijke oefeningen en geslaagde moeilijke oefeningen. Het valt meteen op dat testgroepen met verhoogde beloningen gemiddeld meer moeilijke oefeningen kozen en gemiddeld op meer moeilijke oefeningen slaagden. Tweezijdige Welch t-toetsen vergeleken elk van die metrieken tussen de vier testgroepen. Tussen testgroepen met moeilijkheidslabels en zonder is er geen enkel significant verschil. Eenzijdige Welch t-toetsen vonden echter wel significante ($p < 0,001$) verschillen op basis van gekozen moeilijke oefeningen en geslaagde moeilijke oefeningen tussen testgroepen met verhoogde beloningen en zonder.

Naast de kwalitatieve resultaten tonen we ook de factoren waarvan gebruikers in de open tekstvelden van de post-studie vragenlijst aangaven dat ze de keuze tussen oefeningen beïnvloedden. Figuur 5.5 toont het aantal deelnemers per factor.

Beloningen waren duidelijk de meest invloedrijke factor. Gebruikers die voor beloningen kozen gaven aan dat ze “*een grote beloning*” wilden, omdat “*het motiveerde*”, al dan door het soort beloning te specificeren: “*ik wou mijn kat eten geven*”, “*scorebord nummer 1 halen*”, “*ik dingen wou verzamelen*”. Onder gebruikers die naar de opgave of het level keken wilden sommigen makkelijke oefeningen, sommigen juist uitdagingen en anderen vonden “*wiskunde leuk*” of wilden “*kijken of het lukte*”. Sommigen gaven ook aan dat ze de opgave gewoon als eerste zagen. Ten slotte waren er ook verschillende gebruikers die geen bewuste keuze maakten of bewust willekeurig kozen.

Tabel 5.2: Correlatiematrix van de Hexad-types van 84 gebruikers en hun voorkeuren voor beloningen en hun interactietijd met gamificatieschermen en de Hexad-types van 91 gebruikers en hun zelf-gerapporteerde (Z) voorkeuren voor gamificatie-elementen.

	A	D	F	Ph	P	S
Sterren	0.022	0.063	0.035	0.015	0.206 *	0.028
Munten	-0.002	-0.191 *	-0.155	0.007	-0.063	-0.021
Kisten	-0.083	-0.098	-0.197 *	-0.151	-0.095	-0.064
Visjes	-0.017	-0.034	-0.036	0.046	-0.004	0.009
Scorebord	0.077	0.036	0.018	0.048	0.292 ***	0.029
Badgescherm	0.004	-0.116	-0.099	0.062	0.122	-0.026
Verzameling	-0.020	0.013	-0.106	-0.062	0.146	0.023
Huisdier	-0.074	-0.033	-0.098	-0.038	0.099	-0.040
Scorebord (Z)	0.213 *	-0.108	0.160	-0.014	0.248 *	-0.010
Badgescherm (Z)	0.186	-0.292 **	0.160	0.219 *	0.101	0.034
Verzameling (Z)	0.308 **	-0.272 **	0.159	0.066	0.141	0.030
Huisdier (Z)	0.160	-0.150	0.151	0.405 ***	-0.004	0.177

*** $p < 0,001$, ** $p < 0,01$, * $p < 0,05$

Tabel 5.3: Correlatiematrix van beloningen en gamificatieschermen van 84 gebruikers.

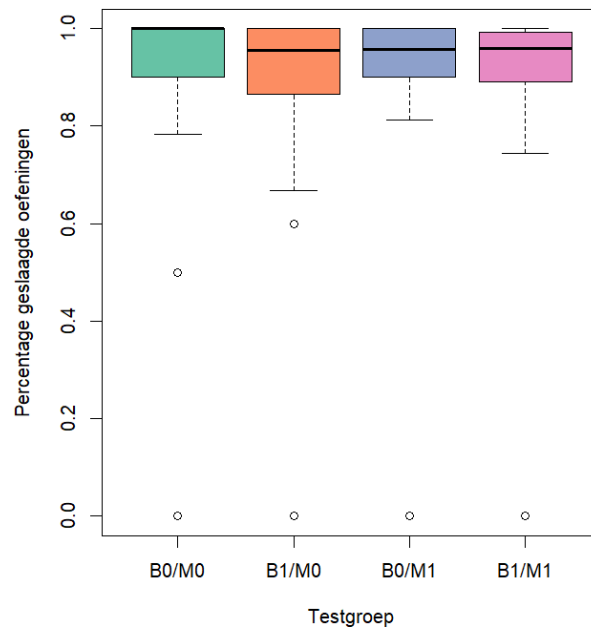
	Ster	Munt	Kist	Visje
Scorebord	0.646 ***	0.259 **	0.094	0.183 *
Badgescherm	0.285 ***	0.702 ***	0.296 ***	0.253 **
Verzameling	0.163	0.450 ***	0.661 ***	0.314 ***
Huisdier	0.175 *	0.335 ***	0.372 ***	0.753 ***

*** $p < 0,001$, ** $p < 0,01$, * $p < 0,05$

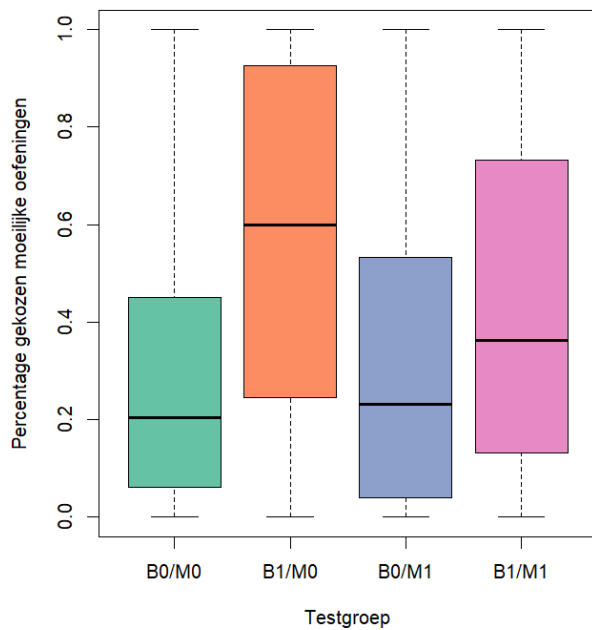
Tabel 5.4: Correlatiematrix van de Hexad-types van 146 gebruikers.

	A	D	F	Ph	P
Disruptor	0.041				
Free Spirit	0.377 ***	0.368 ***			
Philanthropist	0.442 ***	0.213 **	0.446 ***		
Player	0.212 *	0.260 **	0.281 ***	0.100	
Socialiser	0.228 **	0.308 ***	0.291 ***	0.506 ***	0.182

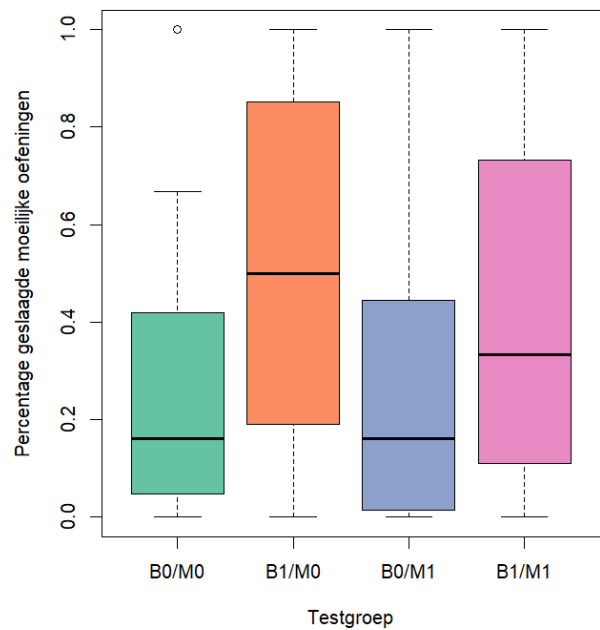
*** $p < 0,001$, ** $p < 0,01$, * $p < 0,05$



(a) Verhouding **geslaagde** oefeningen tot het totaal aantal geslaagde oefeningen

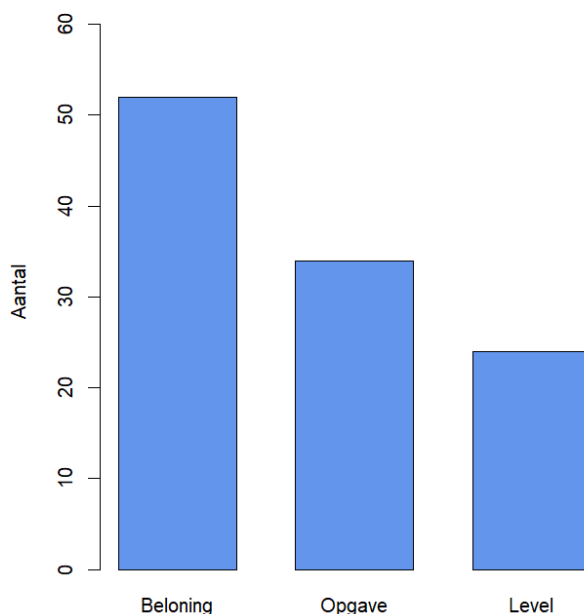


(b) Verhouding **gekozen moeilijke** oefeningen tot het totaal aantal gekozen oefeningen



(c) Verhouding **geslaagde moeilijke** oefeningen tot het totaal aantal gekozen oefeningen

Figuur 5.4: Verhoudingen van 174 gebruikers per testgroep. Een oefening starten en verlaten zonder pogingen telt niet als een gekozen oefening. Oefeningen zijn geslaagd als gebruikers ze oplossen in maximaal drie pogingen.



Figuur 5.5: Het aantal gebruikers uit de 91 dat aangaf welke factoren de keuze tussen oefeningen beïnvloedden.

5.4 Testgroepen en leerefficiëntie

Zoals OV 2.1 stelt willen we weten welke invloed moeilijkheidslabels en verhoogde beloningen hebben op de leerefficiëntie van gebruikers, gemeten aan de hand van hun Elo-scores. Figuur 5.6a toont de totale persoonlijke Elo-winst van 174 gebruikers, gesommeerd over alle modules en concepten, per testgroep. Het valt op dat testgroepen met verhoogde beloningen gemiddeld een hogere Elo-winst hebben dan testgroepen zonder. Tweezijdige Welch t-toetsen vonden echter geen significantie verschillen tussen de testgroepen op basis van verhoogde beloningen, noch moeilijkheidslabels.

5.5 Testgroepen en zelf-gerapporteerde competentie

Zoals OV 2.2 stelt willen we weten welke invloed moeilijkheidslabels en verhoogde beloningen hebben op de zelf-gerapporteerde competentie. Figuur 5.6b toont de gemiddelde zelf-gerapporteerde competentie per testgroep. Tweezijdige Welch t-toetsen vonden geen significantie verschillen tussen de groepen op basis van verhoogde beloningen, noch moeilijkheidslabels.

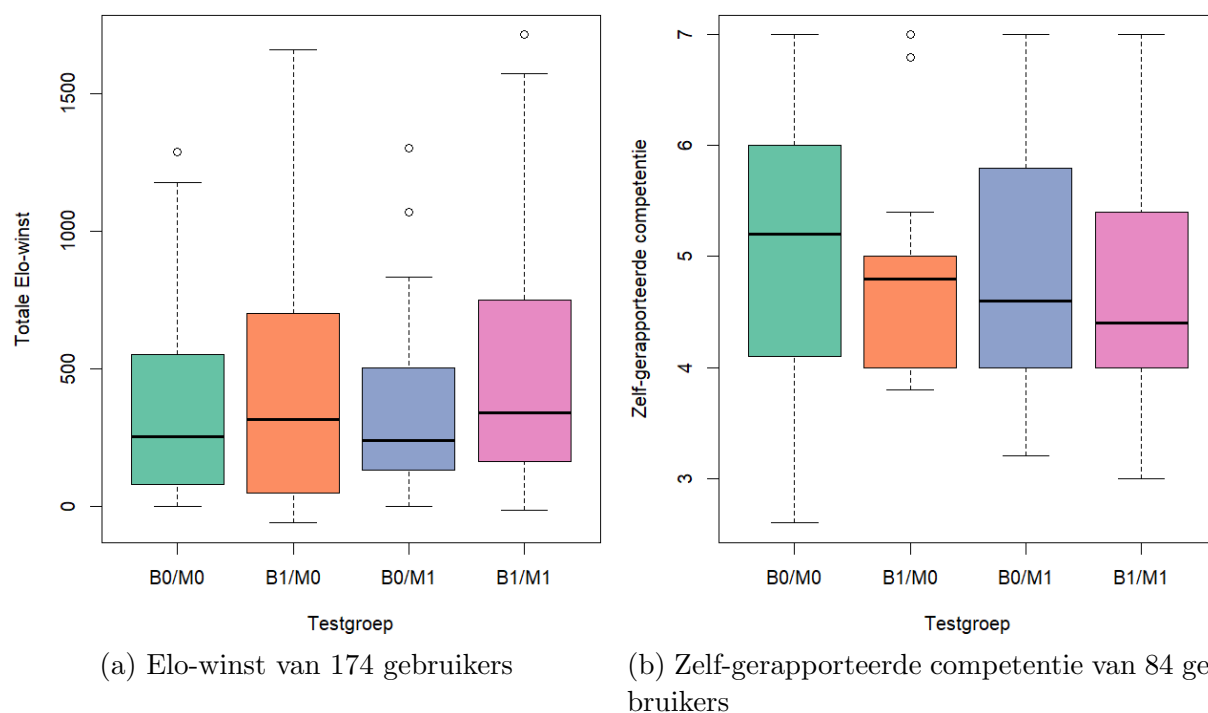
5.6 Feedback van leerkrachten

Meerdere leerkrachten reageerden positief over Alchemath na het onderzoek. Ze bevestigden dat leerlingen de app leuk vonden en dat ze hen kan helpen om wiskunde in te oefenen. Hun positieve feedback benadrukt dat het interessant is om Alchemath te laten bestaan na de afloop van deze thesis, of zoals één leerkracht vermeldde:

Tabel 5.5: Correlatiematrix van de Hexad-types van 146 gebruikers volgens Hexad-24 [31] langs de y-as en Hexad-12 [36] langs de x-as.

	A12	D12	F12	Ph12	P12	S12
A	0.817 ***	−0.001	0.297 ***	0.304 ***	0.208 *	0.169 *
D	0.201 *	0.859 ***	0.335 ***	0.095	0.304 ***	0.267 **
F	0.371 ***	0.342 ***	0.800 ***	0.330 ***	0.242 **	0.196 *
Ph	0.457 ***	0.131	0.357 ***	0.870 ***	0.099	0.334 ***
P	0.258 **	0.303 ***	0.329 ***	0.035	0.867 ***	0.165 *
S	0.267 **	0.212 *	0.298 ***	0.466 ***	0.148	0.894 ***

*** $p < 0,001$, ** $p < 0,01$, * $p < 0,05$



Figuur 5.6: Leerefficiëntie, de totale persoonlijke Elo-winst van gebruikers gesommeerd over alle modules en concepten, en gemiddelde zelf-gerapporteerde competentie per testgroep.

“ De leerlingen (..) vonden het echt leuk en gaven aan dat ze op deze manier beter kunnen inoefenen. Ik hoop dus dat we de app nogmaals volgend jaar kunnen gebruiken. ”

Ze gaven ook aan dat de app een aantal werkpuntjes had, zoals de trage verandering in Elo-scores:

“ Ik heb geen breuken gezien en voor mij duurde het gedeelte met enkel overbrengen van een term te lang. ”

Hoofdstuk 6

Discussie

In dit hoofdstuk interpretern we de resultaten uit Hoofdstuk 5 en vergelijken we ze met bevindingen in de literatuur. Op basis van die interpretatie kunnen we antwoorden bieden op de onderzoeksvragen uit Paragraaf 3.1.

6.1 Voorkeur voor spelelementen

OV 1.1 Is er een verband tussen het Hexad-type van gebruikers op basis van hun gebruiksgedrag en het type op basis van de pre-studie vragenlijst?

Vergeleken met Tondello et al. [31] toont Figuur 5.1b opmerkelijke verschillen qua dominante Hexad-types. *Player* komt veruit het meeste voor met 35%, terwijl dat type bij Tondello et al. maar 10% representeerde. *Philanthropist* en *Socialiser* komen ook minder voor in verhouding tot andere archetypes. Mogelijk heeft deze discrepantie te maken met de doelgroep van dit onderzoek en zou dus leeftijdsgebonden zijn. Eerder onderzoek suggereerde al dat het Hexad-model niet helemaal geschikt zou zijn om profielen van jongeren op te stellen [37].

De resultaten in Tabel 5.3 zijn een indicatie dat gebruikers het verband doorhadden tussen beloningen en gamificatieschermen; dat ze wisten waar de beloningen voor stonden en hoe ze het effect van de beloningen konden waarnemen. We verwachtten de verbanden uit Tabel 3.2 terug te vinden en dat is het geval.

Als we de correlaties in Tabel 5.2 vergelijken met die van Altmeyer et al. [35] dan zien we enkel het verband *Player*-scorebord terugkomen. Net zoals Altmeyer et al. vinden we geen sterk positief verband tussen *Free Spirit* en verzamelobjecten. De verbanden *Achiever*-scorebord en *Philanthropist*-huisdier komen wel overeen met Altmeyer et al. als we enkel de zelf-gerapporteerde voorkeuren in beschouwing nemen. Bovendien lijkt de verzameling volgens die voorkeuren meer overeen te komen met *Achievers*. Dit verband houdt steek aangezien *Achievers* geneigd zijn om reeksen te voltooien [31]. De discrepantie tussen de voorkeuren die het systeem waarnam en de zelf-gerapporteerde voorkeuren heeft mogelijk te maken met onderlinge interactie tussen leerlingen in klasverband. De sociale druk is hoger in klasverband, waardoor leerlingen mogelijk gemakkelijker geneigd zijn om

hun voorkeuren aan te passen aan de voorkeuren van klasgenoten.

De correlaties tussen Hexad-types onderling komen sterk overeen met Tondello et al. [31], behalve *Philanthropist-Player* en *Player-Socialiser* die bij ons ontbreken. Bovendien vonden wij verbanden die Tondello et al. niet vonden: *Disruptor-Player* en *Disruptor-Socialiser*. Het verband *Socialiser-Philanthropist* is veruit het sterkste, wat zowel overeenkomt met Tondello et al. [31] als Ooge et al. [37].

6.2 Testgroepen en oefeningen kiezen

OV 2. Hoe kan nudging toegepast worden om gebruikers te laten kiezen voor moeilijke oefeningen?

Figuur 5.4 en de bijhorende analyse in Paragraaf 5.3 geven een volwaardig antwoord op OV 2. Verhoogde beloningen als een vorm van *incentive nudging* overhaalt gebruikers om enerzijds meer te kiezen voor moeilijke oefeningen en zorgt er ook voor dat ze op meer moeilijke oefeningen slagen. Nudging lijkt dus een geschikte manier om een gedrag van leerlingen in een onderwijssysteem te beïnvloeden en ervoor te zorgen dat leerlingen meer uitdagende oefeningen kiezen. Komende paragrafen gaan in op welke effecten deze nudging kan hebben op het leerproces van leerlingen.

Gebruikers gaven ook zelf aan dat ze vooral naar beloningen keken, zie Figuur 5.5. De beloningen van oefeningen lijken dus gebruikers meer te beïnvloeden bij oefeningen kiezen dan de opgave of de moeilijkheid voorgesteld door moeilijkheidslabels. Beloningen lijken de andere factoren te overstemmen, wat overeenkomt met de kwantitatieve resultaten: als beloningen belangrijker zijn dan de waargenomen moeilijkheid aan de hand van de opgave of moeilijkheidslabels, dan zijn ze voldoende om gebruikers te overtuigen om toch voor moeilijker oefeningen te kiezen.

6.3 Testgroepen en leerefficiëntie

OV 2.1. Welke vorm van nudging, verhoogde beloningen of moeilijkheidslabels, zorgt voor een verschil in leerefficiëntie?

Zowel moeilijkheidslabels als verhoogde beloningen lijken geen effect te hebben op de leerefficiëntie van gebruikers. Hoewel Figuur 5.6a suggereert dat de leerefficiëntie hoger is voor testgroepen met verhoogde beloningen, komt dit verschil niet tot uiting in de resultaten van de statistische analyse. Aangezien verhoogde beloningen duidelijk gebruikers lijken te overtuigen om op meer moeilijke oefeningen te slagen en aangezien slagen op moeilijke oefeningen per definitie een grotere Elo-winst teweegbrengt, zou dat suggereren dat de leerefficiëntie mee moet groeien als gevolg. Gebruikers kunnen echter ook kiezen om steeds gemakkelijkere oefeningen te kiezen die een kleinere Elo-winst teweegbrengen, maar die ze sneller en met minder fouten kunnen oplossen. De Elo-score van gebruikers stijgt tenslotte enkel als ze oefeningen maken zonder fouten; vanaf één fout daalt hun Elo-score. Figuur 5.2 bevestigt dat gebruikers vooral lineaire vergelijkingen maakten

en fig. 5.4a toont aan dat B0/M0 het hoogste slaagpercentage heeft, wat zou kunnen suggereren dat veel gebruikers oefeningen kozen onder hun niveau.

6.4 Testgroepen en zelf-gerapporteerde competentie

OV 2.2. Welke vorm van nudging, verhoogde beloningen of moeilijkheidslabels, zorgt voor een verschil in zelf-gerapporteerde competentie?

Zowel moeilijkheidslabels als verhoogde beloningen lijken geen effect te hebben op de zelf-gerapporteerde competentie van gebruikers. Onze hypothese dat moeilijkheidslabels de zelf-gerapporteerde competentie verhogen gaat dus niet op. Figuur 5.6b suggereert zelf het tegengestelde: testgroepen met moeilijkheidslabels lijken zichzelf als minder competent in te schatten dan testgroepen zonder. De moeilijkheidslabels verhogen mogelijk de druk op gebruikers, waardoor ze meer twijfelen aan hun eigen kunnen. In conclusie, moeilijkheidslabels lijken geen geschikte vorm van nudging voor moeilijke oefeningen.

6.5 Beperkingen van het onderzoek en toekomstig werk

Deze paragraaf kadert de beperkingen van het onderzoek en hoe toekomstig werk kan verderbouwen op de aangesneden onderwerpen.

1. Het onderzoek verliep via de digitale onderwijsapplicatie Alchemath, waardoor we de onderzoeksvragen kaderden in de specifieke wiskundecontext van de app. Een nadeel van onderzoek in klasverband is dat leerlingen konden merken dat niet iedereen dezelfde grootte van beloningen kreeg. Toekomstig werk kan de effecten van nudging onderzoeken wanneer er geen interactie is tussen deelnemers.
2. Leerlingen in klasverband gebruikten Alchemath slechts één lesuur, volgens onze instructies om niet te veel wanorde te scheppen in het curriculum van leerkrachten. Hoewel niets hen weerhield om na dat lesuur verder gebruikt te maken van de app, haakten leerlingen snel af. Toekomstig werk kan de voorkeuren van gebruikers op lange termijn monitoren en nagaan in welke mate ze veranderen.
3. Om een breed testpubliek aan te spreken voorzag Alchemath drie modules met oefeningen over lineaire, kwadratische en kubische vergelijkingen. De app wist echter niet welk niveau de eindgebruikers hadden en gaf hen de vrijheid om zelf modules te kiezen. Uit de resultaten blijkt dat leerlingen vooral lineaire vergelijkingen maakten, hoewel dat normaal leerstof is voor de eerste graad. Voor leerlingen uit hogere graden waren deze oefeningen relatief eenvoudig en dat geeft mogelijk een vertekend beeld van de doeltreffendheid van verhoogde beloningen.
4. Het Elo-systeem om oefeningen te kiezen op niveau van de leerling ervaarde een *cold start* vanwege een tekort aan data over Elo-waarden van de concepten uit Tabel 3.1. De uiteindelijke Elo-scores die resulteerden uit het onderzoek zijn wellicht nog niet volledig geconvergeerd. Toekomstig onderzoek kan onze eindscores verder

blootstellen aan uitwisseling met gebruikers zodat ze een correcter beeld scheppen van de moeilijkheid van oefeningen.

Hoofdstuk 7

Conclusie

Dit hoofdstuk vat de resultaten van het onderzoek samen. Eerst overlopen we kort de resultaten voor de onderzoeksvragen en daarna tonen we een overzicht van de bijdragen die dit onderzoek levert aan het mens-machine-interactiedomein en het onderwijsdomein.

7.1 Resultaten voor de onderzoeksvragen

We kunnen slechts deels de verbanden tussen de Hexad-types van gebruikers en hun voorkeuren voor spelelementen bevestigen. Loggegevens zoals het aantal gekozen beloningen en de totale interactietijd met gamificatieschermen kunnen enkel de verbanden tussen *Players* en scoreborden bevestigen. Uit de voorkeuren die gebruikers zelf rapporteerden komt ook het verband tussen *Philanthropists* en het virtuele huisdier tot uiting.

Ons onderzoek bevestigt dat verhoogde beloningen als een vorm van *incentive nudging* ervoor zorgen dat gebruikers voor uitdagendere oefeningen kiezen en meer uitdagendere oefeningen oplossen. Hoewel we een onderbouwd vermoeden hebben, kunnen we niet bevestigen dat uitdagendere oefeningen maken daarom ook de leerefficiëntie van gebruikers verhoogt. We vonden ook geen significante verschillen tussen testgroepen op basis van zelf-gerapporteerde competentie.

7.2 Bijdragen aan het mens-machine-interactiedomein

De resultaten van dit onderzoek dragen op een aantal manieren bij tot het domein van de mens-machine-interactie:

1. De resultaten tonen concrete verbanden aan tussen de Hexad-types van gebruikers en hun voorkeuren voor spelelementen die deels overeenkomen met Altmeyer et al. [35] en daarmee hun bevindingen bekrachtigen.
2. De resultaten bevestigen dat *incentive nudging* met verhoogde beloningen een effectieve manier is om leerlingen te doen kiezen voor effectieve oefeningen. Hoewel het onderzoek geen statistische bevestiging kon geven, hebben wij een sterk vermoeden dat effectieve oefeningen leiden tot hogere leerefficiëntie.

3. De resultaten tonen aan dat Hexad-12 een goede benadering is voor Hexad-24 en bevestigen daarmee Krath et al. [36] in een onderwijscontext.

7.3 Bijdragen aan het onderwijsdomein

De resultaten van dit onderzoek dragen op een aantal manieren bij tot onderwijs in de praktijk en onderzoek rond gepersonaliseerde onderwijsplatformen:

1. De digitale onderwijsapplicatie Alchemath is een volwaardig product dat kan blijven bestaan nadat deze thesis is afgerond. Leerkrachten gaven na het onderzoek aan dat ze de app graag opnieuw zouden gebruiken in volgende schooljaren. Volgens hen ervaren leerlingen de app als leuk en leerzaam en vormde het een mooie aanvulling van de geziene leerstof.
2. Ons onderzoek spitste zich toe op wiskunde en vergelijkingen en gebruikte Elo-rating om de moeilijkheid van oefeningen te modelleren volgens moeilijkheid. De Elo-scores voor concepten van vergelijkingen zijn tamelijk geconvergeerd en bieden een goed startpunt voor toekomstig onderzoek.
3. Verder ontwierpen we in het kader van vergelijkingen een reeks templates met een overzicht van gebruikte concepten die alle lineaire, kwadratische en kubische vergelijkingen kunnen vertegenwoordigen. Deze templates vormen een indeling die toekomstige ontwerpers kunnen gebruiken om oefeningen *on the fly* te laten maken met mogelijkheden voor personalisatie.

Bijlage A

Vragenlijsten

Tabel A.1: De Hexad-24 [31] vragenlijst om een gebruikers Hexad-type te bepalen, vertaald naar het Nederlands door Ooge [62]. Gebruikers moeten elke uitspraak waarderen op een 7-punts Likertschaal: Helemaal oneens, Oneens, Eerder oneens, Neutraal, Eerder eens, Eens of Helemaal eens

Duid aan hoe goed de onderstaande uitspraken jou beschrijven.

Ik vind het leuk om de status quo in vraag te stellen.
Beloningen zijn een goede manier om me te motiveren.
Ik vind het leuk om obstakels te overwinnen.
Het maakt me gelukkig als ik anderen kan helpen.
Ik vind het leuk om nieuwe dingen uit te proberen.
Ik vind het moeilijk om een probleem los te laten voordat ik een oplossing heb gevonden.
Ik vind het leuk om te provoceren.
Ik geniet van groepsactiviteiten.
Ik zie mezelf als een rebel.
Ik hou ervan anderen te helpen met zich oriënteren in nieuwe situaties.
Ik hou van competities waarbij een prijs gewonnen kan worden.
Ik vind het belangrijk om taken altijd volledig uit te voeren.
Duid nu 'Oneens' aan om te bewijzen dat je nog wakker bent.
Ik vind onafhankelijk zijn belangrijk.
Ik vind het leuk om deel te zijn van een team.
Ik vind het belangrijk om mijn eigen weg te volgen.
Ik vind het belangrijk om me een deel van een gemeenschap te voelen.
Ik laat me vaak leiden door mijn nieuwsgierigheid.
Ik vind het belangrijk om iets terug te krijgen voor mijn inspanningen.
Ik vind het leuk om moeilijke taken de baas te kunnen.
Als de beloning voldoende is, dan zal ik de inspanningen leveren.
Ik hou ervan om mijn kennis te delen.
Ik hou niet van het volgen van regels.
Ik vind interactie met anderen belangrijk.
Ik vind het welzijn van anderen belangrijk.

Tabel A.2: De vragen in de post-studie vragenlijst van het experiment. Eerst vier vragen die polsen naar de gebruikerservaring van de verschillende gamificatie-elementen. Daarna de vijf vragen uit de *Perceived competence* subcluster van de Intrinsic Motivation Inventory [63], vertaald naar het Nederlands door Peeters [64]. Gebruikers moeten elke uitspraak van de eerste twee secties waarderen op een 7-punts Likertschaal: Helemaal oneens, Oneens, Eerder oneens, Neutraal, Eerder eens, Eens of Helemaal eens. Tenslotte een vraag over de invloed van verschillende factoren op het keuzescherf met betrekking tot nudging.

Waardering van gamificatie.

Ik vind het scorebord leuk.
 Ik vind het vooruitgangsscherf leuk.
 Ik vind het verzamelscherf leuk.
 Ik vind het huisdier leuk.

Zelf-gerapporteerde competentie.

Ik denk dat ik best wel goed ben in deze activiteit.
 Ik denk dat ik best wel goed ben in deze activiteit, vergeleken met andere leerlingen.
 Ik ben tevreden met hoe ik het heb gedaan bij deze activiteit.
 Ik voelde mij competent bij deze activiteit.
 Nadat ik enige tijd bezig was met deze activiteit, voelde ik mij best competent.

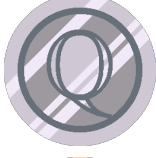
Manier van kiezen.

Als ik een oefening koos keek ik vooral naar:
 de beloning / de opgave / het level / andere, specificeer.

Bijlage B

Details proof of concept

Tabel B.1: Badges in Alchemath van level één tot level vijf.

Afbeelding	Titel	Beschrijving	Munten naar vol- gende level
	<i>Natuurlijke Noob</i>	Je bent natuurlijk nog een noob. Verzamel munten om een hogere badge te verdienen!	8
	<i>Gehele Gabber</i>	Je hebt de negatieve getallen omarmd, maar geheel tevreden ben je nog niet. Op naar het volgende niveau.	14
	<i>Rationele Rakker</i>	Door rationeel denken ben je rationaal geworden; in staat om zowel breuken als kommagetallen te manipuleren.	20
	<i>Reële Re- bus</i>	Je kennis is werkelijk reëel. Je beheerst alle getallen die echt zijn. Was er maar een manier om nieuwe getallen te imagineren.	30
	<i>Complexe Kastaar</i>	Je bent officieel heerser over alle getallen, zowel werkelijk als denkbeeldig. Tevreden kijk je terug naar je vooruitgang.	n.v.t.

Tabel B.2: Verzamelobjecten in Alchemath gegroepeerd per zeldzaamheid.

Afbeelding	Beschrijving
Veelvoorkomende verzamelobjecten ($p = \frac{12}{17}$)	
	Hoorn van een π -renese steenbok. Deze soort is vooral te zien tijdens het paarseizoen rond 14 maart.
	Het medaillon der kwadranten wordt gebruikt om te navigeren door het cartesisch coördinatenstelsel.
	Reservevingers. Voorwerp dat van pas komt wanneer je eigen vingers te vermoeid zijn van al dat schrijven.
	Door een breuk in deze parel kwamen teller en noemer tevoorschijn. Werkelijk een mooi exemplaar.
	Eerste gebruik van het integraalteken uit 1800 v.Chr, vervaardigd van een giftige bessentak genaamd rubus integralis.
	Een geodriehoek uit een ver verleden. In die tijd hadden ze nog geen driehoeken uitgevonden.
	Uit dit exemplaar zijn duidelijk scherpe scherven gehaald. Waarschijnlijk voor een puntspiegeling.
	Deze geluksbrenger wordt al generaties lang doorgegeven en dateert waarschijnlijk uit de tijd van Pythagoras.
Zeldzame verzamelobjecten ($p = \frac{4}{17}$)	
	Het dodenmasker van de aztekenkoning Uwatl. Volgens de legende zou hij het loodje hebben gelegd nadat hij 0 door 0 deelde.



Spiekniet, een magisch mineraal dat leerkrachten gebruiken om studenten te betrappen tijdens het spieken.



Dit naamloos kristal werd ooit gevonden tijdens een expeditie door het Gauss-gebergte. Het gloeit in de nabijheid van een normaalverdeling.



Het spiekend oog is een gegeerd voorwerp onder studenten, vooral tijdens de examenperiode. Leerkrachten gebruiken vaak spieknietstenen om zijn kracht te verminderen.



De ketting der permutaties verandert constant de volgorde van haar kralen; niemand snapt hoe dat mogelijk is. De faculteit wiskunde weet misschien meer.



Hoofd der brandende vragen. Neem altijd pauses na intense wiskundesessies, zo voorkom je een gelijkaardig lot.



De stok der natuurlijke getallen. Een symbolisch voorwerp vervaardigd uit telhout dat een geheel positieve werking heeft.



Het instrument dat Descartes in de 17de eeuw gebruikte om imaginaire getallen te vangen in zijn dromen. Zijn werking is zeer complex.

Legendarische verzamelobjecten ($p = \frac{1}{17}$)



Deze kaarten werden gebruikt in één van de grootste wiskundige tovertrucs ooit. Zelfs de functies die normaal altijd waakzaam zijn waren door de truc afgeleid.



Pure eureka, gevangen in glas. Er is geen probleem waar dit middel geen oplossing voor kan bedenken, maar soms wil je het antwoord liever niet weten.



De megahedron lijkt op een bol, maar is eigenlijk een regelmatig veelvlak met 1 miljoen vlakken. Het vuur ontstaat door hoeken van bijna 90° .



De kroon der getallen haalde haar kracht uit de wiskundige die haar opheeft. Vele wiskundigen hebben dit instrument gedragen.



Het bloed van een universitair wiskunde, een van de laatste van zijn soort.



De Bol der Antwoorden is de tegenhanger van het Enigma; het geeft op elke wiskundevraag een antwoord.



De Necrekenomicon is een goddeloos boek over de zogenoemde 'donkere wiskunde'. Het beschrijft onder andere hoe men deelt door nul.



Het Enigma, een bovennatuurlijk mechanisme dat onophoudelijk wiskundevragen genereert.

Tabel B.3: Gemoedstoestanden van het huisdier in Alchemath.

Afbeelding	Beschrijving
	Uitgehongerd
	Veel honger
	Beetje honger
	Weinig honger
	Geen honger

Bijlage C

Overige resultaten

Tabel C.1: De verschillen tussen initiële *hand-crafted* Elo-scores en de geconvergeerde Elo-scores per concept op het einde van het onderzoek.

Concept	Initiële Elo-score	Geconvergeerde Elo-score
Lineaire vergelijkingen		
Optellen	1050	894
Tekenverandering	1100	985
Vermenigvuldigen	1200	1065
Distributiviteit	1300	1232
Breuken	1350	1297
Kwadratische vergelijkingen		
Optellen	1050	999
Tekenverandering	1100	1061
Vermenigvuldigen	1200	1155
Distributiviteit	1300	1275
Breuken	1350	1326
<i>Wortel formule</i>	1400	1402
Kubische vergelijkingen		
Optellen	1050	1015
Tekenverandering	1100	1079
Vermenigvuldigen	1200	1178
Distributiviteit	1300	1286
Breuken	1350	1338
Regel van Horner	1400	1400

Bijlage D

Informatie aan betrokkenen en toestemmingsformulieren

TITEL

Beloningen personaliseren om de grens van competentie te verleggen.

BESCHRIJVING EN DOELSTELLINGEN

Gamificatie is een techniek die van alles een spel probeert te maken, met vaak de bedoeling om de motivatie van betrokken personen te verhogen. Gamificatie gebruikt hiervoor spelelementen zoals virtuele munten, badges en leaderboards. Uit eerder onderzoek bleek dat gamificatie het beste werkt wanneer het rekening houdt met voorkeuren van een persoon. Om te weten te komen welk soort spelelement een persoon leuk vindt, gebruiken onderzoekers meestal een vragenlijst. Dat werkt prima voor onderzoeksdoeleinden, maar wordt vaak negatief ervaren door gebruikers. Dit onderzoek probeert de voorkeuren van een persoon te bepalen op basis van wat een persoon doet binnen een applicatie en vergelijkt de resultaten met die van de vragenlijst. De applicatie in kwestie is een wiskundig leerplatform dat door de thesisstudent wordt gebouwd en gehost wordt op servers van KU Leuven.

ONDERZOEKSMETHODE**1. Vragenlijst**

Bij het eerste gebruik moeten deelnemers een korte vragenlijst over spelvoorkeuren invullen.

2. Oefeningen maken

Na registratie kunnen gebruikers onbeperkt oefeningen maken op de applicatie. Aan elke oefening hangt een beloning vast, die bepaald is op basis van wat gebruikers doen binnen de applicatie. Deze beloning is steeds bekend vóór gebruikers een oefening starten.

INFORMATIE OVER DE VERWERKING VAN UW PERSOONSGEGEVENS

Als gevolg van uw deelname aan het onderzoek “Beloningen personaliseren om de grens van competentie te verleggen” worden persoonsgegevens over u verzameld en verwerkt. Deze gegevens worden verwerkt in overeenstemming met de Algemene Verordening Gegevensbescherming (AVG). Met dit informatieblad willen wij u informeren over het gebruik en de opslag van uw gegevens.

Tijdens dit onderzoek worden met name de volgende persoonsgegevens over u verzameld: gebruiksgedrag in de applicatie en persoonlijke kenmerken: leeftijd, geslacht.

Gebruik van uw persoonsgegevens

Alleen de persoonsgegevens die nodig zijn voor de doeleinden van dit onderzoek zullen worden verzameld en verwerkt. Meer specifiek beoogt het onderzoek de motivatie om wiskunde te leren in een digitale leeromgeving te vergroten door beloningen te personaliseren. De verzamelde gegevens kunnen mogelijk worden hergebruikt in toekomstige studies.

De gegevens die voor dit onderzoek worden verzameld, worden gepseudonimiseerd. Dit betekent dat gegevens die u kunnen identificeren, zoals uw naam, worden gescheiden van de andere gegevens in het onderzoek en worden vervangen door een unieke willekeurige code. Op deze manier zijn de gegevens niet meer eenvoudig te herleiden tot een specifieke betrokkene. Alleen de onderzoeker kan door middel van de unieke code de gegevens koppelen aan een specifiek individu. Dit gebeurt echter alleen in uitzonderlijke gevallen, bijvoorbeeld als u gebruik wilt maken van uw recht op inzage, rectificatie of verwijdering van uw gegevens. Ook wordt u niet genoemd in publicaties die voortkomen uit het onderzoek.

De gegevens worden verwerkt op basis van algemeen belang. Dit betekent dat het onderzoek leidt tot kennisvoortgang en inzichten genereert die (direct of indirect) de samenleving ten goede komen.

Uw gegevens worden door de onderzoekers maximaal 10 jaar na het einde van de studie bewaard op een beveiligde opslagplaats aan de KU Leuven. Na deze periode worden uw persoonsgegevens definitief verwijderd indien deze niet langer nodig zijn voor de doeleinden van het onderzoek.

Uw rechten

U heeft het recht om meer informatie op te vragen over het gebruik van uw gegevens. Daarnaast heeft u het recht op inzage, rectificatie of verwijdering van uw gegevens, tenzij uitoefening van deze rechten het behalen van de onderzoeksdoelstellingen onmogelijk zou maken of ernstig zou belemmeren.

Indien u een van deze rechten wenst uit te oefenen, kunt u contact opnemen met de onderzoekers via de contactgegevens onderaan dit informatieblad.

Contactgegevens

Voor de doeleinden van dit onderzoek is de KU Leuven de verwerkingsverantwoordelijke. Meer bepaald zullen enkel de betrokken onderzoekers (Joran De Braekeleer, Katrien Verbert, Jeroen Ooge) toegang hebben tot uw persoonsgegevens. Mocht u specifieke vragen hebben over dit onderzoek, inclusief de verwerking van uw persoonsgegevens, neem dan gerust contact met hen op.

Joran De Braekeleer

E-mail: joran.debraekeleer@student.kuleuven.be

Voor alle verdere vragen en zorgen over de verwerking van uw persoonsgegevens kunt u contact opnemen met Toon Boon, de functionaris voor gegevensbescherming van de KU Leuven (dpo@kuleuven.be). Gelieve het betreffende onderzoek te specificeren door zowel de titel als de namen van de betrokken onderzoekers te vermelden.

Indien u na contact met de functionaris voor gegevensbescherming toch een klacht wenst in te dienen over het gebruik van uw persoonsgegevens, kan u zich wenden tot de Belgische Gegevensbeschermingsautoriteit (www.gegevensbeschermingsautoriteit.be).

Instemmingsverklaring leerkracht

Ik ga als wiskundeleerkracht akkoord met onderstaande uitspraken:

1. Ik ben betrokken bij het wetenschappelijk onderzoek voor de masterproef van Joran De Braekeleer, masterstudent toegepaste informatica aan KU Leuven. Ik begrijp hoe het onderzoek verloopt na het lezen van de informatiebrief.
2. Ik bied mijn leerlingen de kans om gebruik te maken van de wiskunde-app Alchemath tijdens een van mijn lessen wiskunde. Het gaat daarbij over een supplementaire oefeningen-les die aansluit bij mijn geziene leerstof.
3. Ik ben ervan op de hoogte dat leerlingen die Alchemath willen gebruiken een vragenlijst moeten invullen, wat 5-10 minuten kan duren.
4. Ik begrijp dat leerlingen die Alchemath gebruiken meewerken aan het onderzoek van Joran De Braekeleer. Daarom verplicht ik mijn leerlingen niet om deel te nemen en zet ik ze ook niet onder druk om dat te doen (bijvoorbeeld door hen te evalueren terwijl ze Alchemath gebruiken). Tijdens mijn les met Alchemath voorzie ik oefeningen op papier voor leerlingen die zich niet willen deelnemen of die daartoe geen toestemming krijgen van hun ouders.

Naam:

School:

Datum:

Handtekening voor akkoord:

Geïnformeerde toestemming

Titel van het onderzoek:

Beloningen personaliseren om de grens van competentie te verleggen.

Naam + contactgegevens promotor en onderzoeker(s):

Joran De Braekeleer

Departement Computerwetenschappen

E-mail: joran.debraekeleer@student.kuleuven.be

Prof. Dr. Katrien Verbert

Departement Computerwetenschappen

Mens-machine interactie, Leuven (Arenberg)

E-mail: katrien.verbert@kuleuven.be

Jeroen Ooge

Departement Computerwetenschappen

Mens-machine interactie, Leuven (Arenberg)

E-mail: jeroen.ooge@kuleuven.be

Doel en methodologie van het onderzoek:

Motivatie bijbrengen om wiskunde te leren in een digitale leeromgeving door personalisatie van beloningen.

Duur van het experiment:

Vragenlijst (5 min) + gebruik applicatie (45 min)

- Ik begrijp wat van mij verwacht wordt tijdens dit onderzoek.
- Ik weet dat ik zal deelnemen aan volgende proeven of testen:
Gebruik van een educatief leerplatform over wiskunde.
- Er zijn geen risico's of ongemakken verbonden aan mijn deelname.
- Ikzelf of anderen kunnen baat bij dit onderzoek hebben op volgende wijze:
Het gebruik van de applicatie kan wiskundevaardigheden versterken.
- Mijn deelname levert een bijdrage aan het wetenschappelijk onderzoek. Ik weet ik geen verdere beloning of compensatie voor mijn deelname zal ontvangen.
- Ik begrijp dat mijn deelname aan deze studie vrijwillig is. Ik heb het recht om mijn deelname op elk moment stop te zetten. Daarvoor hoef ik geen reden te geven en ik weet dat daaruit geen nadeel voor mij kan ontstaan.

Voor de verdere verwerking van de verzamelde gegevens geldt het algemeen belang als rechtsgrond volgens de AVG/GDPR. Stopzetting van deelname aan de studie houdt dus in dat de eerder verzamelde gegevens nog verder rechtsgeldig kunnen worden betrokken in de studie en niet moeten worden verwijderd door KU Leuven.

- Ik begrijp dat in de context van dit onderzoek gegevens worden verzameld die volgens de Algemene Verordening Gegevensbescherming als bijzonder gevoelig worden beschouwd. Ik geef dan ook expliciet mijn toestemming om deze gegevens in de context van dit onderzoek te verzamelen. Het gaat om de volgende gegevens:
Leeftijd, geslacht.
- Ik weet dat er in dit onderzoek opnames van mij gemaakt kunnen worden:
 - Audio-opnames voor transcriptie tijdens think-alouds over de applicatie. Deze opnames worden na transcriptie meteen verwijderd.
- De resultaten van dit onderzoek kunnen gebruikt worden voor wetenschappelijke doeleinden en mogen gepubliceerd worden. Mijn naam wordt daarbij niet gepubliceerd. Doorheen het onderzoek zullen mijn gegevens steeds vertrouwelijk behandeld worden. De onderzoekers zullen de volgende maatregelen nemen om mijn privacy te beschermen:
 - Alle data wordt gepseudonimiseerd (er wordt gewerkt met unieke codes i.p.v. namen)

Opgemaakt in tweevoud.

- De gegevens van de deelnemers worden enkel geraadpleegd via een beveiligde OneDrive gelinkt met de KU Leuven, deze gegevens worden nooit gedownload op de persoonlijke computer van de onderzoeker
- Ik wil graag op de hoogte gehouden worden van de resultaten van dit onderzoek. De onderzoeker mag mij hiervoor contacteren op het volgende e-mailadres:

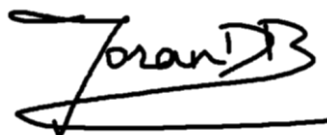
- Voor vragen evenals voor de uitoefening van mijn rechten (inzage gegevens, correctie ervan,...) weet ik dat ik na mijn deelname terecht kan bij:
Joran De Braekeleer (zie boven)
Meer informatie met betrekking tot privacy in onderzoek kan ik terugvinden op www.kuleuven.be/privacy. Verdere vragen over privacyaspecten kan ik richten tot de functionaris voor gegevensbescherming: dpo@kuleuven.be
- Deze studie werd beoordeeld en goedgekeurd door de Sociaal-Maatschappelijke Ethische Commissie (SMEC) van KU Leuven (G-2023-6197). Voor eventuele klachten of andere bezorgdheden omtrent ethische aspecten van deze studie kan ik contact opnemen met SMEC: smec@kuleuven.be

Ik heb bovenstaande informatie gelezen en begrepen en heb antwoord gekregen op al mijn vragen betreffende deze studie. Ik stem toe om deel te nemen.
Belangrijk: als je 15 jaar of jonger bent, dan moeten je ouders dit document verplicht ondertekenen.

Datum:

Handtekening proefpersoon

Handtekening onderzoeker:

A handwritten signature in black ink, appearing to read 'Joran DB', with a long horizontal stroke underneath.

Handtekening(en) ouder(s)
(verplicht voor kinderen 15 jaar en jonger)

Bibliografie

- [1] Computerspel. <https://nl.wikipedia.org/wiki/Computerspel>. Accessed: 2022-12-20.
- [2] Fedwa Laamarti, Mohamad Eid, and Abdulmotaleb El Saddik. An overview of serious games. *International Journal of Computer Games Technology*, 2014:11–11, 2014.
- [3] Justin D. Weisz, Maryam Ashoori, and Zahra Ashktorab. Entanglion: A board game for teaching the principles of quantum computing. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*, CHI PLAY '18, page 523–534, New York, NY, USA, 2018. Association for Computing Machinery.
- [4] Kristin Siu and Mark O. Riedl. Reward systems in human computation games. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play*, CHI PLAY '16, page 266–275, New York, NY, USA, 2016. Association for Computing Machinery.
- [5] Seth Cooper, Firas Khatib, Adrien Treuille, Janos Barbero, Jeehyung Lee, Michael Beenen, Andrew Leaver-Fay, David Baker, Zoran Popović, et al. Predicting protein structures with a multiplayer online game. *Nature*, 466(7307):756–760, 2010.
- [6] Ahmed M Abou Elfotouh, Eman S Nasr, and Mervat H Gheith. Towards a comprehensive serious educational games' ontology. In *Proceedings of the 3rd Africa and Middle East Conference on Software Engineering*, pages 25–30, 2017.
- [7] Joe Cutting and Ioanna Iacovides. Learning by doing: Intrinsic integration directs attention to increase learning in games. *Proc. ACM Hum.-Comput. Interact.*, 6(CHI PLAY), oct 2022.
- [8] Juho Hamari. *Gamification*, pages 1–3. John Wiley & Sons, Ltd, 2019.
- [9] Pandey. Gamification trends in 2019. <https://elearningindustry.com/gamification-trends-2019-tips-ideas-packed>. Accessed: 2022-12-20.
- [10] Radoff. Gamification. <https://web.archive.org/web/20110219171239/http://radoff.com/blog/2011/02/16/gamification/>. Accessed: 2022-12-20.
- [11] Robertson. Can't play, won't play. [web log comment]. <http://www.hideandseek.net/2010/10/06/cant-play-wont-play>. Accessed: 2022-12-20.
- [12] Scott Nicholson. A recipe for meaningful gamification. In *Gamification in education and business*, pages 1–20. Springer, 2015.

- [13] James Paul Gee. What video games have to teach us about learning and literacy. *Computers in entertainment (CIE)*, 1(1):20–20, 2003.
- [14] Edward L. Deci and Richard M. Ryan. The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality*, 19(2):109–134, 1985.
- [15] Edward L Deci. Effects of externally mediated rewards on intrinsic motivation. *Journal of personality and Social Psychology*, 18(1):105, 1971.
- [16] Cody Phillips, Daniel Johnson, Peta Wyeth, Leanne Hides, and Madison Klarkowski. Redefining videogame reward types. In *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction*, pages 83–91, 2015.
- [17] Cody Phillips, Daniel Johnson, Madison Klarkowski, Melanie Jade White, and Leanne Hides. The impact of rewards and trait reward responsiveness on player motivation. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play, CHI PLAY '18*, page 393–404, New York, NY, USA, 2018. Association for Computing Machinery.
- [18] Alena Denisova and Elliott Cook. Power-ups in digital games: The rewarding effect of phantom game elements on player experience. In *Proceedings of the Annual Symposium on computer-human interaction in Play*, pages 161–168, 2019.
- [19] Shipi Kankane, Carlina DiRusso, and Christen Buckley. Can we nudge users toward better password management? an initial study. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems, CHI EA '18*, page 1–6, New York, NY, USA, 2018. Association for Computing Machinery.
- [20] Peter King, H Richard, Cass R Thaler, et al. Nudge: Improving decisions about health, wealth and happiness. *Journal of Social Policy*, 38:726, 2009.
- [21] Vaiva Kalnikaite, Yvonne Rogers, Jon Bird, Nicolas Villar, Khaled Bachour, Stephen Payne, Peter M. Todd, Johannes Schöning, Antonio Krüger, and Stefan Kreitmayer. How to nudge in situ: Designing lambent devices to deliver salient information in supermarkets. In *Proceedings of the 13th International Conference on Ubiquitous Computing, UbiComp '11*, page 11–20, New York, NY, USA, 2011. Association for Computing Machinery.
- [22] Big brother’ eyes make us act more honestly. <https://www.newscientist.com/article/dn9424-big-brother-eyes-make-us-act-more-honestly/>. Accessed: 2022-12-20.
- [23] Putting eyeballs on billboards might help stop crime. <https://www.smithsonianmag.com/smart-news/putting-eyeballs-billboards-might-help-stop-crime-180956394/>. Accessed: 2022-12-20.
- [24] Thiemo Wambsganss, Andreas Janson, Tanja Käser, and Jan Marco Leimeister. Improving students argumentation learning with adaptive self-evaluation nudging. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2):1–31, 2022.

- [25] Arthur E Poropat. A meta-analysis of the five-factor model of personality and academic performance. *Psychological bulletin*, 135(2):322, 2009.
- [26] Jerome Kagan and Nancy C Snidman. *The long shadow of temperament*. Harvard University Press, 2009.
- [27] Richard Bartle. Hearts, clubs, diamonds, spades: Players who suit muds. *Journal of MUD research*, 1(1):19, 1996.
- [28] Lennart E. Nacke, Chris Bateman, and Regan L. Mandryk. Brainhex: A neurobiological gamer typology survey. *Entertainment Computing*, 5(1):55–62, 2014.
- [29] Rita Orji, Regan L Mandryk, Julita Vassileva, and Kathrin M Gerling. Tailoring persuasive health games to gamer type. In *Proceedings of the sigchi conference on human factors in computing systems*, pages 2467–2476, 2013.
- [30] Andrzej Marczewski. *User Types HEXAD*, pages 65–80. 10 2015.
- [31] Gustavo F. Tondello, Rina R. Wehbe, Lisa Diamond, Marc Busch, Andrzej Marczewski, and Lennart E. Nacke. The gamification user types hexad scale. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play, CHI PLAY '16*, page 229–243, New York, NY, USA, 2016. Association for Computing Machinery.
- [32] Stuart Hallifax, Audrey Serna, Jean-Charles Marty, Guillaume Lavoué, and Elise Lavoué. Factors to consider for tailored gamification. In *Proceedings of the annual symposium on computer-human interaction in play*, pages 559–572, 2019.
- [33] Gustavo F. Tondello, Alberto Mora, and Lennart E. Nacke. Elements of gameful design emerging from user preferences. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play, CHI PLAY '17*, page 129–142, New York, NY, USA, 2017. Association for Computing Machinery.
- [34] Maximilian Altmeyer, Marc Schubhan, Pascal Lessel, Linda Muller, and Antonio Krüger. Using hexad user types to select suitable gamification elements to encourage healthy eating. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–8, 2020.
- [35] Maximilian Altmeyer, Gustavo F Tondello, Antonio Krüger, and Lennart E Nacke. Hexarcade: Predicting hexad user types by using gameful applications. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, pages 219–230, 2020.
- [36] Jeanine Krath, Maximilian Altmeyer, Gustavo F Tondello, and Lennart E Nacke. Hexad-12: Developing and validating a short version of the gamification user types hexad scale. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–18, 2023.
- [37] Jeroen Ooge, Robin De Croon, Katrien Verbert, and Vero Vanden Abeele. Tailoring gamification for adolescents: a validation study of big five and hexad in dutch. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, pages 206–218, 2020.

- [38] Krath Jeanine, Carolina Ana, Klock Tomé, Morschheuser Benedikt, Legaki Nikoletta-Zampeta, Park Solip, Harald F. O. von Korfflesch, and Juho Hamari. Designing tailored gamification: A mixed-methods study on expert perspectives and user behavior in a gamified app for sustainability at work. 2023.
- [39] Radek Pelánek, Tomáš Effenberger, and Jaroslav Čechák. Complexity and difficulty of items in learning systems. *International Journal of Artificial Intelligence in Education*, 32(1):196–232, 2022.
- [40] Peng Liu and Zhizhong Li. Task complexity: A review and conceptualization framework. *International Journal of Industrial Ergonomics*, 42(6):553–568, 2012.
- [41] Amogh Joshi. Optional learning in games. In *Extended Abstracts of the 2020 Annual Symposium on Computer-Human Interaction in Play*, pages 24–26, 2020.
- [42] Andy Gray, Alma A.M Rahat, Tom Crick, Stephen Lindsay, and Darren Wallace. Using elo rating as a metric for comparative judgement in educational assessment. In *Proceedings of the 6th International Conference on Education and Multimedia Technology, ICEMT '22*, page 272–278, New York, NY, USA, 2022. Association for Computing Machinery.
- [43] Boban Vesin, Katerina Mangaroska, Kamil Akhuseyinoglu, and Michail Giannakos. Adaptive assessment and content recommendation in online programming courses: On the use of elo-rating. *ACM Trans. Comput. Educ.*, 22(3), jun 2022.
- [44] Katerina Mangaroska, Boban Vesin, and Michail Giannakos. Elo-rating method: Towards adaptive assessment in e-learning. In *2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT)*, volume 2161-377X, pages 380–382, 2019.
- [45] Elo rating system. https://en.wikipedia.org/wiki/Elo_rating_system#cite_ref-27. Accessed: 2022-12-20.
- [46] Tom Murray and Ivon Arroyo. Toward measuring and maintaining the zone of proximal development in adaptive instructional systems. In *Intelligent Tutoring Systems: 6th International Conference, ITS 2002 Biarritz, France and San Sebastian, Spain, June 2–7, 2002 Proceedings 6*, pages 749–758. Springer, 2002.
- [47] Marlene Scardamalia, Carl Bereiter, Robert S McLean, Jonathan Swallow, and Earl Woodruff. Computer-supported intentional learning environments. *Journal of educational computing research*, 5(1):51–68, 1989.
- [48] Simone de Sousa Borges, Vinicius HS Durelli, Helena Macedo Reis, and Seiji Isotani. A systematic mapping on gamification applied to education. In *Proceedings of the 29th annual ACM symposium on applied computing*, pages 216–222, 2014.
- [49] Gabriel Barata, Sandra Gama, Joaquim Jorge, and Daniel Gonçalves. Improving participation and learning with gamification. In *Proceedings of the First International Conference on gameful design, research, and applications*, pages 10–17, 2013.
- [50] Gabriel Barata, Sandra Gama, Joaquim AP Jorge, and Daniel JV Gonçalves. Relating gaming habits with student performance in a gamified learning experience.

- In *Proceedings of the first ACM SIGCHI annual symposium on Computer-human interaction in play*, pages 17–25, 2014.
- [51] Sepandar Sepehr and Milena Head. Competition as an element of gamification for learning: an exploratory longitudinal investigation. In *Proceedings of the First International Conference on Gameful Design, Research, and Applications*, pages 2–9, 2013.
- [52] Christopher Cheong, Justin Filippou, and France Cheong. Towards the gamification of learning: Investigating student perceptions of game elements. *Journal of Information Systems Education*, 25(3):233, 2014.
- [53] A Matthews and D Pepper. Evaluation of participation in a level mathematics: final report. *London: Qualifications and Curriculum Authority*, 2007.
- [54] Adrian Smith. Making mathematics count. 2004.
- [55] Andrew Noyes and Paula Sealey. Investigating participation in advanced level mathematics: a study of student drop-out. *Research Papers in Education*, 27(1):123–138, 2012.
- [56] Christian E Lopez and Conrad S Tucker. The effects of player type on performance: A gamification case study. *Computers in Human Behavior*, 91:333–345, 2019.
- [57] Alberto Mora, Gustavo F Tondello, Laura Calvet, Carina González, Joan Arnedo-Moreno, and Lennart E Nacke. The quest for a better tailoring of gameful design: An analysis of player type preferences. In *Proceedings of the XX international conference on human computer interaction*, pages 1–8, 2019.
- [58] Maurits Kaptein, Boris De Ruyter, Panos Markopoulos, and Emile Aarts. Adaptive persuasive systems: a study of tailored persuasive text messages to reduce snacking. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 2(2):1–25, 2012.
- [59] Hassan Khosravi, Shazia Sadiq, and Dragan Gasevic. Development and adoption of an adaptive learning system: Reflections and lessons learned. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education, SIGCSE '20*, page 58–64, New York, NY, USA, 2020. Association for Computing Machinery.
- [60] Lauren Eskreis-Winkler, Katherine L Milkman, Dena M Gromet, and Angela L Duckworth. A large-scale field experiment shows giving advice improves academic outcomes for the advisor. *Proceedings of the national academy of sciences*, 116(30):14808–14810, 2019.
- [61] Benjamin A Motz, Matthew G Mallon, and Joshua D Quick. Automated educational nudges to reduce missed assignments in college. *IEEE Transactions on Learning Technologies*, 14(2):189–200, 2021.
- [62] Jeroen Ooge. Het personaliseren van motivationele strategieën en gamificationstechnieken mbv recommendersystemen. *KU Leuven*, 2019.
- [63] Richard M. Ryan and Edward L. Deci. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25(1):54–67, 2000.

- [64] Wessel Peeters. Motivatie meten: 3 vragenlijsten. <https://www.vernieuwenderwijs.nl/motivatie-meten-2-vragenlijsten/>. Accessed: 2022-12-20.
- [65] Tomohiro Nagashima, John Britti, Xiran Wang, Bin Zheng, Violet Turri, Stephanie Tseng, and Vincent Aleven. Designing playful intelligent tutoring software to support engaging and effective algebra learning. page 258–271, Berlin, Heidelberg, 2022. Springer-Verlag.
- [66] Godot engine. <https://godotengine.org/>. Accessed: 2022-12-20.

Departement Computerwetenschappen

Celestijnenlaan 200 A bus 1

3000 LEUVEN, BE

tel. + 32 16 32 7444

fax + 32 16 32 7445

wms.cs.kuleuven

