

Literatuurstudie en ervaringen.

Deze tekst is een aanvulling op het onderzoeksverslag ‘Observaties aan het RekenWeb’.

Auteur: Joris Verrips

Mail : j.verrips@planet.nl

Datum laatste wijziging: 12 september 2007

Appendix 3

Literatuurstudie over testen van bruikbaarheid

Ik heb de laatste jaren veel literatuur bestudeerd op het gebied van ‘usability’, ‘user experience’, ‘informatie architectuur’ en van aanverwante vakgebieden in de hoop van mijn hobby mijn beroep te kunnen maken. In het verleden heb ik onderzoek gedaan naar sprekende software voor mensen met motorische beperkingen en me daarom ook met wat oudere literatuur beziggehouden. Deze paragraaf is losjes op al die literatuur gebaseerd en uiteraard mede geïnspireerd door ervaringen met kinderen. Niet op literatuur over e-learning, verwijzingen staan op <http://www.elearning.nl> en amper op literatuur over vakdidactiek voor wiskunde en rekenen, die ik als leraar oppervlakkig ken.

Wat is bruikbaarheid?

Met bruikbare software worden heel verschillende zaken aangeduid: leerbaar, efficiënt, interessant, foutarm, beleefd, aansluitend bij verwachting van klanten, van consumenten, begrijpelijk, en nog veel meer. Bruikbaarheid is niet gratis, alleen al omdat bij testen met gebruikers vaak veel suggesties voor wijzigingen ontstaan. In sommige situaties is bruikbaarheid onmisbaar: waar menselijke fouten grote gevolgen kunnen hebben, waar concurrentie dwingt om ‘gebruiksvriendelijkheid’ aan te bieden, op het internet waar onoverzichtelijke webpagina’s niet lang bezocht worden en dus weinig rendement hebben en bij het ontwerpen van nieuwe diensten en producten in een concurrerende markt. Noot *4.

Rol van experts

Experts in bruikbaarheid zijn nodig om observaties te interpreteren maar zijn in het algemeen, net als programmeurs, erg verschillend van de doelgroep. Zij kunnen daarom slecht beoordelen waar precies behoefte aan bestaat. Dit probleem speelt ook indien bij ontwikkelwerk met representatieve gebruikers wordt samengewerkt. Een half jaar later zijn die zo representatief niet meer. De literatuur over

bruikbaarheid stelt in talloos veel variaties dat met de doelgroep(en) geëvalueerd dient te worden. Kinderen die nog leren rekenen geven vast heel andere reacties dan volwassenen die dat al lang kunnen en geautomatiseerd hebben. In deze context is ook belangrijk wat en door wie van een bepaalde applet geleerd dient te worden en of dit theoretisch onderbouwd kan worden. In zekere zin lijkt het of de didactische interpretatie leidend moet zijn.

Testen op bruikbaarheid

In de praktijk wordt vaak door experts getest, of niet. Over evalueren volgens heuristieken zoals leerbaarheid, begrijpelijkheid, presentatie en efficiency bevat de literatuur over mens-machine interactie interessante gegevens. Volgens Nielsen en Möhlich kunnen vijf verschillende experts in enkele uren tijd 80% van de problemen met bruikbaarheid tevoorschijn halen. Deze problemen kunnen ook opgespoord en gedocumenteerd worden door leden van de doelgroep systematisch te testen, wat duurder is en meer voeten in aarde heeft. Een enkele expert haalt slechts circa 30%. Men doet daarom vaak een slordige test met twee of drie experts, afgewisseld met leden van de doelgroep die vaak moeilijker te vinden en te motiveren zijn. Uiteraard leveren ook 30% suggesties soms heel wat werk op en is daar niet altijd tijd voor. Dit rapport documenteert duidelijk minder dan 30% van wat in het hele en omvangrijke RekenWeb gevonden kan worden omdat een groot deel daarvan helemaal niet bekeken is. We stellen het optimistisch op 20%, en dat voorspelt dus nog veel werk.

Bruikbaarheid op het internet

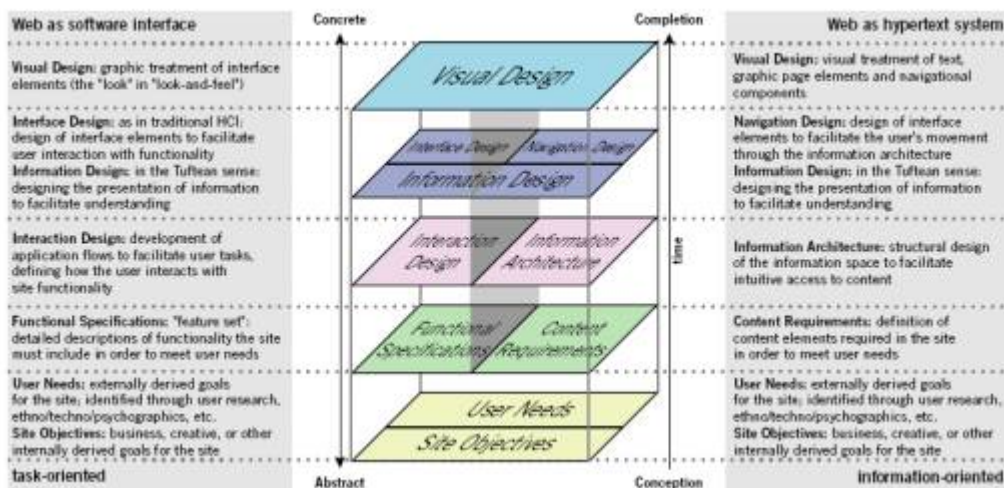
Bruikbaarheid toegepast op het internet onderscheidt zich doordat de gebruikers gemakkelijk weg kunnen gaan en dat ook vaak doen. Daardoor worden hogere eisen gesteld aan presentatie en ontwerp. In literatuur hierover is vaak sprake van ongeduldige gebruikers, van het scheppen van betrokkenheid (engagement), van toenemende ontsluiting (progressive disclosure), van klein en overzichtelijk en van beloning. Het vakgebied van bruikbaarheid is voortdurend in ontwikkeling, juist ook waar het het internet betreft, de laatste jaren wordt bijvoorbeeld steeds meer eye-tracking toegepast en al meer enquêtes via email. Toch is er wel degelijk een hoop kennis uitgekristalliseerd en zijn veel ontwerprichtlijnen de laatste tien jaar niet meer gewijzigd.

Dat het toch ook expertise en onderzoek vereist om deze toe te passen blijkt uit enerzijds <http://www.useit.com/alertbox/features.html> over het nádeel van veel 'features' (eigenschappen) en het belang van het scheppen van engagement en, anderzijds en van dezelfde auteur, <http://www.useit.com/alertbox/20050131.html> over het vóórdeel van 'features' als het om teenagers gaat, die in het algemeen snel verveeld zouden zijn en slecht zouden lezen en dus geboeid moeten worden. In het verleden ging het adagium van 'too little too late' wel op voor bruikbaarheidsonderzoek: te weinig effect en te laat in het ontwikkelproces. Tegenwoordig schijnt, doordat webgebruikers makkelijk weg kunnen gaan en het er –in potentie- ook heel veel kunnen zijn (meer dan een miljard mensen hebben tegenwoordig toegang tot het internet), de 'return on investment' nogal hoog te zijn. Wanneer verblijfstijd en percentage terugkerend bezoek gemeten worden is soms met eenvoudige ingrepen aanzienlijke winst te behalen. Zie voor meer achtergrond hierover bijvoorbeeld de columns (<http://www.useit.com/alertbox/intranet-usability.html> en <http://www.useit.com/alertbox/20030107.html>). De evidente successen van bruikbaarheidsonderzoek en van modern ontwerp nemen niet weg dat economische overwegingen bij vakken als software engineering, bruikbaarheid, interactie ontwerp en user experience erg belangrijk zijn. Lang niet alles wat kán hóeft ook.

Bij sommig gebruikersonderzoek worden vier tot zes gebruikers per dag geobserveerd met moderne techniek als doorkijkspiegels, video, eye-tracking en (elektronische) enquêtes. Gebruikers wordt gevraagd standaard opdrachten te doen, hardop te denken, na uitzetten van de monitor te beschrijven wat er op te zien was en uit te leggen wat het betekent, ‘teach back protocol’ te doen waar gebruikers elkaar uitleggen wat je ergens van leren kunt, samen te werken, en meer en verschilt dus nogal van een normale lessituatie omdat veel meer tijd beschikbaar is voor de precieze beleving van een individu. Het hoeft niet erg technisch te zijn, door gewoon kritisch en rustig te kijken, alles te controleren en veel te argumenteren kun je heel wat suggesties vinden.

Belang van ontwerp en van doelstellingen

Veel ontwerpers voor het internet kennen en waarderen het diagram van Figuur 110. Het diagram is niet makkelijk te interpreteren en wordt hier alleen opgenomen om duidelijk te maken dat ook voor een eenvoudige webzijde onderscheid gemaakt kan worden tussen verschillende lagen, zoals presentatie, informatie, navigatie en doelstellingen. Een organisatie in lagen komt vaak voor in de informatica.



Figuur 110. Uit: Garret, 2004.

Hoe verhelderend ook, het is niet duidelijk waar de leerprocessen van gebruikers in zoen diagram gesitueerd moeten worden, waar al uit blijkt dat ontwerpen van (reken)applets niet eenvoudig is. Dit geldt eens te meer indien wordt afgeweken van het beproefde concept van geprogrammeerde instructie waar verschillende stof telkens op dezelfde manier aangeboden en herhaald wordt. En waar al heel wat van ontwikkeld is. Er gelden denkelijk andere criteria dan voor webzijdets met als doelstelling om een wereldreis te boeken, om contactgegevens op te zoeken of om een webwinkel te bezoeken.

Doelstellingen van applets

Over wiskundesoftware voor het onderwijs is lezenswaard Boon en Drijvers, 2005. Zij verzuchten onder andere het gevoel te hebben gehad nieuw gebied te exploreren en merken op dat de rol van de docent belangrijk blijft en dat goede voorbereiding wel nodig is. Zij onderscheiden modelapplets, waarmee wordt bedoeld een (visuele) representatie als model voor wiskundige begrippen, van oefenapplets, waarmee bijvoorbeeld het oplossen van vergelijkingen kan worden geoefend. In hun onderzoek bleken de leerkrachten graag veel automatisch te laten administreren en registreren en dat zal voor rekenapplets wel niet anders zijn.

In 'Rekensoftware op het internet. Het RekenWeb gebruiken in de klas' (van Galen en Jonker, 2003) worden allerlei applets besproken en groepsgewijs verbonden met didactiek van het rekenen, met conform de uitgangspunten van realistisch rekenonderwijs oefenen wordt beschouwd als de laatste fase van het leerproces. 'In de meeste programma's van het RekenWeb staat het verwerven van inzicht voorop. Vaak biedt het programma de gelegenheid te experimenteren: wat gebeurt er als je het op de ene manier doet, wat als je het op een andere manier doet? Via dat experimenteren ontdekken de kinderen hoe dingen in elkaar zitten'. Ook elders in hun boek wordt 'onderzoekend leren' gepropageerd, en spreekt met van 'de overvloed aan oefenspellen voor rekenen'. Dit is niet de plaats en ik ben ook niet de persoon om daar nader op in te gaan.

In de vakdidactiek is veelal uitgangspunt dat een onderwijsleerproces een leerling van een beginsituatie naar een eindsituatie dient te leiden en zorgvuldig dient te worden ontworpen. Daartoe wordt nogal eens de voorkennis geëxpliciteerd en worden diverse 'voorstellingsvormen' onderscheiden zoals algebraïsch, grafisch of met concreet materiaal. En dan is er een (realistische) context, een rekenvoorbeeld, een formule in de (grafische) rekenmachine invoeren en plotten, oefenen, en spelletjes al dan niet als een applet. Het valt (mij) niet mee om omgekeerd, bij de beschikbare applets hun rol in diverse onderwijsleerprocessen te analyseren en te specificeren maar wel is duidelijk dat deze applets alléén daartoe niet voldoende zijn.

Visie op ontwerp en visie op effectonderzoek

De literatuur over bruikbaarheid bevat talloze juweeltjes als 'consistency is key to usability' (Nielsen, 2007a). Ze wekt ook hoge verwachtingen over inschakelen van ontwerpers, etnografen en vormgevers in het begin van het ontwikkelproces, met verwijzing naar het belang van een goed ontwerp voor een goed resultaat. Je zou er leerlingen, leraren, didactici en psychologen nog bij kunnen nemen. Hoe terecht die aanbevelingen ook zijn, bij ontwikkelen voor didactische doelen is het geen sinecure en geen garantie voor resultaat. Het kan uiteraard ook erg kostbaar zijn. Het maakt testen achteraf en met de doelgroep(en) absoluut niet overbodig, wat onvermijdelijk ook tot constructieve kritiek leidt, en leidt ook tot vragen over didactiek.

Achtergrond en eigen ervaring

Ik ben tweedegraads leraar wiskunde en heb een hobby in bruikbaarheid van spraaksoftware en van software voor het wiskundeonderwijs. Ik heb in de jaren 1990 en 1991 regelmatig met hele klassen in het computerlokaal toen beschikbare software gebruikt, en had in die tijd in mijn lokaal ook een beamer met Derive (waarmee een 4^e-klas proefwerk typisch 10%-20% sneller gemaakt kon worden dan met de hand) en een laptop waarop onder andere geprogrammeerde instructie van Wolters-Noordhoff aanwezig was, oefeningen met haakjes wegwerken en dergelijke, maar ook Vu-Grafiek, Ruimte, Parabool, Calis, software van de Wageningse methode en Derive.

Mijn doelstellingen waren toen zelf ervaring op te doen, de kinderen op voor de hand liggende manier te confronteren met de vele nieuwe en verbazingwekkende mogelijkheden van ict en om door aanbieden van andere voorstellingsvormen het onderwijs aantrekkelijker te maken. Deze doelstellingen werden ook behaald. Ik ben toen van school veranderd, heb tien jaar buiten het onderwijs gewerkt, en heb ook -onbetaald- onderzoek gedaan naar bruikbaarheid van sprekende software. Het bleek erg moeilijk en tijdrovend te zijn om met de resultaten van bruikbaarheidsonderzoek iets te doen, en helemaal in commercieel opzicht. Ook bleek het erg belangrijk of tussenpersonen (in dit geval hulpverleners maar leerkrachten zullen vast niet anders reageren) er belang bij hadden en er tijd mee konden besparen. Moeizame ervaringen met toepassen en distribueren zijn geen uitzondering in bruikbaarheidsonderzoek, zie Moggridge, 2007 voor treffende anekdotes door onderzoekers van naam.

Veel later (2005) heb ik diverse nieuwe programma's van de firma VuSoft getest, practicum geschreven voor Vu-Grafiek en met leerlingen op een huiswerkinstituut en op een school naar de bruikbaarheid van dit practicum gekeken. Weer viel op dat dat veel tijd kost. Een beetje interessant elektronisch werkblad wat in een uur of twee werd ontworpen en gemaakt vereiste typisch nog zes tot acht uur testen, aanpassen en verbeteren. En dan blééf de vraag of andere leraren er mee uit de voeten zouden kunnen. Leraren hebben geen tijd voor eindeloos schaven en testen wat pleit voor gespecialiseerde testers met een didactische achtergrond, waar er op het Freudenthal Instituut ook wel een aantal van zullen rondlopen. In België worden daarom werkbladen gemaakt en gedistribueerd aansluitend aan een bepaalde methode. De eerste ervaringen hiermee zijn gunstig. Voor mij is het een open vraag of methode onafhankelijke benadering haalbaar of zelfs beter is. En ook hoe ooit voldoende effectonderzoek te doen plaatsvinden, zelfs bij web-gebaseerde en dus voor veel mensen toegankelijke techniek.

Tijdens testen viel op dat veel kinderen waardeerden zelf werkbladen te maken. Ze zeiden vaak dat ze door grafische illustratie een begrip wat ze al kenden beter begrepen. Bijvoorbeeld dat de constante a in de functie $f: x \Rightarrow a x^2$ de vorm van de parabool bepaalt, en dat de grafiek van functie $g: x \Rightarrow a(x-b)^2$ dus dezelfde vorm heeft als de grafiek van f . Je kunt dus hopen dat door aanbieden van een andere voorstellingsvorm (namelijk families van functies) het begrip 'verdiept' wordt. Dat is iets anders dan dat dit begrip alleen door software zou kunnen worden aangebracht. De eenheidskring begrijpen en hanteren is wel iets anders dan er een plaatje van bekijken, al bevat dat een fraaie animatie van alle goniometrische functies. Ambitieuze doelstellingen voor het ontwerpen van eleren zijn sindsdien voor mij niet erg geloofwaardig meer. Tenslotte bleek distributie problematisch. Het spreekt niet vanzelf dat uitgever een getest practicum distribueren willen of dat scholen in een aan hun behoeften aangepast practicum geïnteresseerd zijn of daar voor betalen willen.

Zowel op mijn laatste als op mijn voorlaatste school heb ik af en toe kinderen kunnen observeren die met rekenapplets werkten van het RekenWeb. Op het Huygens College te Amsterdam gebeurt dit met hele groepen tijdens zogenaamde 'Pluslessen'. Dit loopt subiet mis als je de leerlingen allemaal verschillende applets laat doen en vereist instructie op maat. Je hebt in zo'n opzet weinig tijd om je met een kapotte muis of met een slecht geïnstalleerde Java-versie of Flash-player bezig te houden. Er worden soms hele voor de hand liggende vragen gesteld en de reacties van de leerlingen waren aanvankelijk nogal passief. Hoewel samenwerken voorkwam ging dat vooral over bedienbaarheid. Verder herinner ik me trotse leerlingen die bijvoorbeeld de rekenvierkanten allemaal af hebben en dat de leraar zelf wilden laten zien. Klassegesprekken over de software vonden (bijna) niet plaats.

De leerlingen op mijn laatste school, het Nova College in Amsterdam Slotervaart, zijn veelal kinderen van asielzoekers met een behoorlijke intellectuele bagage en een brede belangstelling. Ze hebben weinig kennis van de Nederlandse taal en een heel gevarieerde achtergrond. De teksten van alle gebruikte leerboeken en software zijn voor hen dan ook niet of slechts met grote moeite te volgen. Ik gaf hen les uit Getal en Ruimte VMBO 1A en 1B hoewel sommigen van hen eigenlijk een veel hoger niveau aankunnen. De CD uit het werkboek, waar ook enkele applets op staan, werd niet gebruikt. Uiteraard was de tekst van vele applets Het zou voor hen –en voor de scholen waar zij op zitten- prettig zijn om applets onmiddellijk in het Engels en het Arabisch te kunnen vertalen, en ook wel om snel te kunnen toetsen wat hun niveau is en welke applets (en welk onderwijs) ze nodig hebben.

Appendix 5

Inspirerende uitspraken

Your best friend, and worst enemy..... is the head of the independent test organisation.

(F.P. Brooks, 1975).

To me, error analysis is the sweet spot for improvement.

(D. Norman, 2005).

Organizations which design systems are constrained to produce designs which are copies of the communication structures of these organizations.

(M. Conway, 1968).

Noot *1 - * 3 staan in Observaties korteversie .

Noot

*4 Soms wordt investeren in bruikbaarheid niet gewenst. Als je er voor je inkomsten van afhankelijk bent onmisbaar te zijn is het niet prettig wanneer je computersysteem door elke collega en dus door elke potentiële concurrent gemakkelijk geleerd en gebruikt zou kunnen worden. Ben je expert in een bepaald vakgebied dan is een leerinspanning helemaal niet erg, en ruime functionaliteit belangrijker dan gemakkelijk terugvinden van zeldzaam gebruikte functie(s). Heel wat automatiseerders voelen zich expert in die zin. Als het herstellen van of het voorkomen van fouten belangrijk is stelt dat weer heel andere eisen aan leerbaarheid, presentatie, ‘mentaal model’ en programmatuur. Zo wordt usability weer wél gebruikt om het werk van piloten te analyseren, te vergemakkelijken en verbeteren.

Literatuur

Benyon, Turner & Turner, 2005.

Human Computer Interaction.

Boon, P Drijvers, P 2005.

<http://www.fi.uu.nl/wisweb/kloo/algebra-applets/Eindverslag.pdf>

Algebra en applets, leren en onderwijzen.

Brenda Laurel en Joy Mountford, 1990, Apple computer

The art of HCI design.

Brooks F.P. 1975, 1995.

The mythical Man-Month.

Tim Brown Strategy by Design.

[http://ideo.com/pdf/FastCo-StrategyByDesign\(TimBrown\).pdf](http://ideo.com/pdf/FastCo-StrategyByDesign(TimBrown).pdf)

Conway, M. 1968. Datamation. How Do Committees Invent?

<http://www.melconway.com/research/committees.html>

Cooper A., 2005.

The inmates are running the asylum.

Cooper A. & Reimann, 2003.

About Face 2.0.

De Souza, C.S. 2005.

Semiotic engineering of human computer interaction.

Van Galen, F Jonker, V. 2003.

Rekensoftware op het internet. Het RekenWeb gebruiken in de klas.

Garret, J. 2004.

The Elements of User Experience.

Jonker, 2005

<http://www.glsconference.org/2005/pop/jonker.htm>

Krug, A. 2005.

Do not make me think. A common sense approach to web usability. 2nd edition, paperback.

Moggridge, 2007.

Designing interactions.

Natan, 2000.

Adviseren als tweede beroep.

Nielsen en Mohlich, 1990.

Heuristic evaluation of user interfaces.

Proceedings of ACM CHI '90 Conference

Nielsen, 2005.

Newletter from NNG.

Nielsen, 2007.

Recommended reading. Books About Web Design and Usability

Op: www.nng.com

Nielsen en Loranger, 2005.

Prioritizing Web Usability.

Norman, 2004.

Things that make us smart.

Norman, 2005.

Human Centered Design considered harmful. Toelichting op http://www.jnd.org/dn.mss/hcd_harmful_a_clari.html

Saffer, 2005.

Designing for Interaction.

Stone, Jarret, Woodroffe, Minocha, 2004.

User Interface Design and Evaluation.

Veen, J. 2004.

Zie <http://www.adaptivepath.com/publications/essays/archives/000315.php>