

LEREN VAN COMPLICATIES IN DE ANESTHESIE

Bachelorproject

Roel E. Stamhuis, s1819674, r.e.stamhuis@student.rug.nl

Samenvatting: De detectie van complicaties op een anesthesische bewakingsmonitor kan verbeterd worden door gebruik van grafische displays. De gebruiker dient wel adequaat getraind te zijn voor herkenning. In onze replicatie van een experiment van Doesburg (2011) en Moes (2012) moesten onervaren proefpersonen vijf complicaties leren herkennen. Van de proefpersonen moest de ene groep de complicaties expliciet leren vanaf een symptomenlijst. De proefpersonen uit de andere groep moesten de complicaties impliciet leren: Ze kregen een screenshot van een complicatie te zien, moesten een diagnose stellen en kregen vervolgens feedback – dit herhaalde zich totdat ze alle complicaties driemaal achtereenvolgend correct herkend hadden. Het experiment bestond uit drie delen: Een diagnosefase, een leerfase en weer een diagnosefase om de leerprestaties te testen. De groep die screenshots en feedback had gekregen presteerde significant beter dan de andere groep. Deze vondst heeft mogelijk consequenties voor medische opleidingen. Vervolgonderzoek naar de generaliseerbaarheid van dit effect, en naar de factoren die meespeelden, is echter nodig.

1. Introductie

Verpleegkundigen op de Intensive Care maken gebruik van bewakingsmonitoren om de fysieke conditie van patiënten in de gaten te houden. Hartslag, bloeddruk en een aantal andere vitale functies worden weergegeven op digitale displays. Ook anesthesisten in de operatiekamer maken gebruik van deze monitoren. Het is echter gebleken dat zowel anesthesisten als IC-verpleegkundigen suboptimaal presteren in het interpreteren van deze displays: complicaties worden regelmatig onderschat of over het hoofd gezien (Cooper, Newbower, & Kitz, 1984; Eichhorn, 1989). Daarom zijn er vele grafische alternatieven ontwikkeld en getest, al dan niet met succes (Drews & Westenskow, 2006). De potentiële prestatiewinst van de gebruikers is groot (Cook, Potter, Woods, & McDonald, 1991; Michels, Gravenstein, & Westenskow, 1997; Van Amsterdam, Cnossen, Ballast, & Struys, submitted). Het ontwerp van de monitordisplay is echter niet de enige factor die het succes bepaalt, kennis van de gebruiker speelt natuurlijk een belangrijke rol. Men zou verwachten dat ervaren verpleegkundigen beter presteren dan leken, ongeacht het uiterlijk van de display. Dit blijkt echter niet altijd het geval (Effken, Kim, & Shaw, 1994).

Uit eerder onderzoek blijkt dat informatie vaak beter onthouden wordt als het (tussentijds) uit het geheugen opgehaald moet worden (Kang, McDaniel, & Pashler, 2011). Dit fenomeen staat

bekend als het *testing effect* en is aangetoond bij zowel verbale als visuele informatie (Kang, 2010). Er is een onderscheid te maken tussen expliciet en impliciet leren. Met expliciet leren bedoelen we bestudering van bijv. een stuk tekst met de uitdrukkelijke intentie het stuk te leren; met impliciet leren bedoelen we leren als ‘bijproduct’ van een andere activiteit – dus zonder de expliciete intentie om te leren. Het testing effect valt onder deze laatste categorie. Als blijkt dat het detecteren en herkennen van complicaties beter of sneller gaat na een impliciet leertraject, heeft dat wellicht consequenties voor medische opleidingen.

Ons onderzoek was een replicatie van een experiment van Doesburg (2011) en Moes (2012). We hergebruikten hun zeshoekdisplay, oorspronkelijk van Gurushanthaiah, Weinger en Englund (1995). Deze display is een implementatie van het *emergent feature* concept (Wickens & Hollands, 1999), een holistische beschouwing van de display is voldoende om een indruk te krijgen of de patiënt in orde is.

Ons experiment was verdeeld in drie fasen. In het eerste deel (de gewinningsfase) kreeg de proefpersoon negen van deze bewakingsmonitoren te zien, waarop fictieve patiënten bewaakt werden. De proefpersoon moest aangeven wanneer hij dacht dat het misging met een patiënt. In het tweede deel (de leerfase) moest de proefpersoon vijf complicaties uit het hoofd leren. De proefpersonen waren ingedeeld in twee groepen. De ene groep kreeg in de leerfase een tabel met de symptomen per

complicatie aangeboden, en de (expliciete) instructie om deze te leren. De andere groep leerde de complicaties met een interactief leerprogramma waarin screenshots van de complicaties gepresenteerd werden. De proefpersoon werd dan gevraagd welke complicatie er te zien was, en kreeg vervolgens feedback op dit antwoord. In het derde deel tenslotte (de testfase) moest de proefpersoon weer hetzelfde doen als in de gewenningsfase, maar diende nu ook aan te geven welke complicatie hij dacht dat er optrad, als op een knop gedrukt was.

Elk type complicatie veroorzaakt een bepaalde unieke vervorming van de zeshoek. We gingen er vanuit dat de interactieve leer methode het leren van deze vervormingen faciliteerde. Deze methode was bovendien een implementatie van het testing effect. Onze hypothese luidde daarom dat het interactieve programma zou resulteren in betere herkenning van de complicaties tijdens de testfase.

2. Methoden

2.1. Proefpersonen

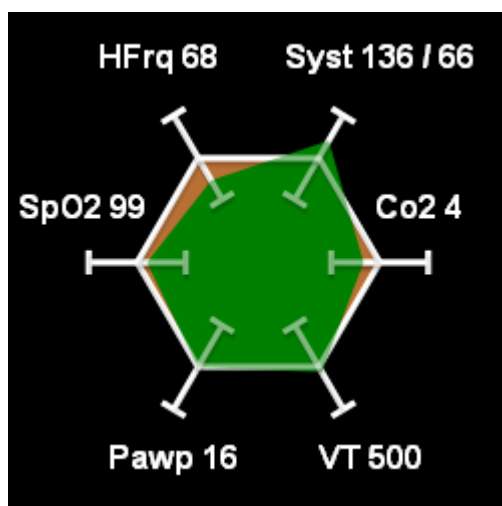
Er hebben 21 studenten van allerlei opleidingen op HBO- en WO-niveau meegedaan aan het onderzoek. De proefpersonen kregen een

beloning van 5 euro en een gratis kop koffie of thee tijdens het experiment. De gemiddelde leeftijd bedroeg 21,9 jaar. Van de proefpersonen deed 57% een bètawetenschappelijke opleiding, niemand deed een medische opleiding.

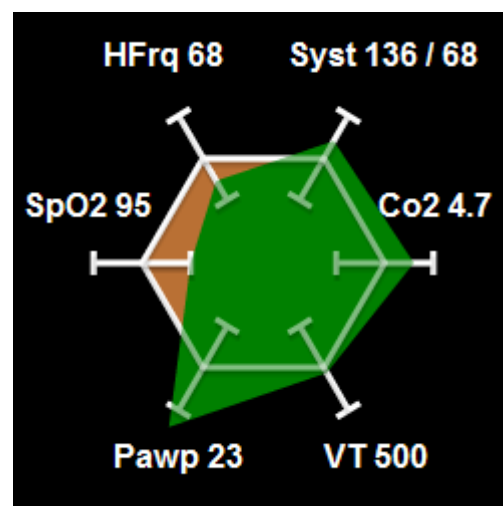
2.2. Stimuli

Gewenningsfase en testfase. De virtuele bewakingsmonitors bestonden uit een groene zeshoek, die op een oranje regelmatige zeshoek werd geprojecteerd. Elke hoek representeerde een vitale functie. De zes hoeken stonden respectievelijk voor bloeddruk, CO₂-gehalte van het bloed, ademvolume, luchtdruk in de longen, O₂-verzadiging van het bloed en hartfrequentie. Als de hoek naar buiten bewoog, betekende dat een stijging in de corresponderende waarde. Bewoog de hoek naar binnen, dan betekende dat een daling. Er waren voor elke vitale functie buiten- en binnengrenzen getekend. Bij elke hoek stond ook de getalswaarde van de vitale functie aangegeven. Zie figuur 2.1 en 2.2.

Negen van deze bewakingsmonitors werden gelijktijdig afgebeeld in een 3-bij-3-opstelling. Op elke monitor werd een fictieve patiënt gemonitord, elke vijf seconden werd het beeld vernieuwd. Voor de testfase werden dezelfde negen scenario's gebruikt als voor de gewenningsfase, maar deze hadden aselect een



Figuur 2.1: De zeshoekmonitor. Dit is een voorbeeld van de gebruikte zeshoekmonitor van Doesburg (2011). De groene zeshoek geeft de waarde van zes vitale functies weer. De oranje regelmatige zeshoek fungeert als referentie, evenals de witte binnen- en buitengrenzen. De patiënt verkeert hier in normale toestand.



Figuur 2.2: De zeshoekmonitor terwijl een longprobleem optreedt. De O₂-saturatie daalt, het CO₂-gehalte en de luchtdruk in de longen stijgen.

Syst (bloeddruk)	CO2 (CO2-gehalte)	VT (adem-volume)	PAWP (luchtdruk longen)	SpO2 (O2-gehalte)	HFrq (Hartslag)	Complicatie
	↑		↑	↓		Longprobleem
↑					↑	Pijn
↓	↑		↑	↓		Spanningspneumotorax
	↓			↓		Circulatiefalen
↓						Hypovolemie

Tabel 2.1: De tabel die de ene groep proefpersonen in de leerfase aangeboden kreeg. De vijf rijen staan voor vijf complicaties, de kolommen voor de zes vitale functies. Pijltjes omhoog of omlaag betekenen een stijging resp. daling van een vitale functie, en geven zo aan wat de symptomen van elke complicatie zijn.

andere positie in de 3-bij-3-opstelling gekregen zodat de proefpersonen dit niet zouden opmerken.

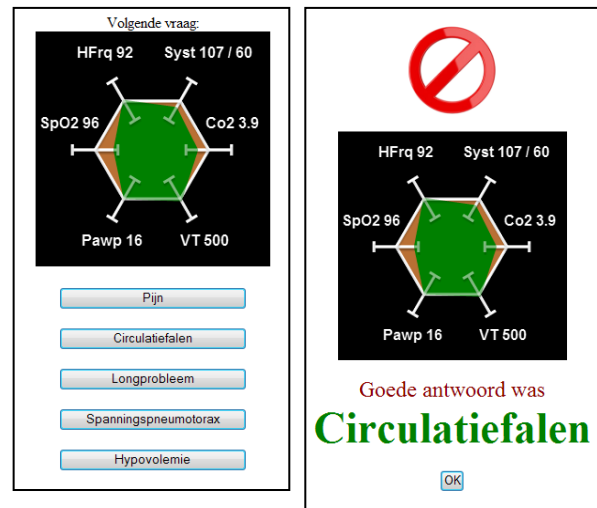
Leerfase. De proefpersonen uit groep A kregen tabel 2.1 te zien. De tabel bestond uit vijf rijen (voor elke complicatie één) en zeven kolommen (zes voor de vitale functies en één voor de complicatiennaam). Met omhoog of omlaag gerichte pijltjes werd aangegeven hoe de vitale functies zich bij elke complicatie gedragen. De proefpersonen uit groep B kregen steeds een momentopname van een optredende complicatie te zien en moesten raden welke complicatie het betrof (figuur 2.3).

2.3. Opzet & Procedure

Het eerste deel van het experiment – de **gewinningsfase** – duurde 15 minuten, en verspreid over deze tijd (en over de negen monitors) traden in totaal twaalf complicaties op. Als de proefpersonen de indruk hadden dat het mis ging met een patiënt, moesten ze op de corresponderende knop drukken. Ter bevestiging kleurde de betreffende monitor rood, wat betekende dat een fictieve dokter onderweg was.

Vervolgens moesten de proefpersonen in de **leerfase** vijf complicaties leren, aan de hand van de bijbehorende symptomen. We pasten een tussen-proefpersonenopzet toe. De proefpersonen waren in twee groepen verdeeld, deze worden hieronder gespecificeerd.

Groep A moest vanaf de tabel leren. De proefpersonen werden van tevoren geïnstrueerd om net zolang door te gaan tot ze de inhoud kenden en de complicaties dachten te kunnen



Figuur 2.3. Het interactieve leerprogramma. Links een screenshot met daaronder de vijf opties, in willekeurige volgorde. Rechts een voorbeeld van feedback.

benoemen. Direct na het leren werd kort getest of dit inderdaad het geval was: elke complicatie werd eenmaal in tabelvorm weergegeven, en de proefpersoon moest aanklikken welke complicatie het was. Maakte hij een fout, dan moest hij verder leren van de tabel.

De proefpersonen van groep B kregen een screenshot van een complicatie te zien. Ze moesten aanklikken welke complicatie ze dachten dat het was. Na elke screenshot werd aan de proefpersoon teruggekoppeld of zijn antwoord juist of onjuist was (zie figuur 2.3). Er waren drie unieke screenshots per complicatie (in totaal vijftien afbeeldingen), die in willekeurige volgorde werden aangeboden. Tussen twee screenshots van dezelfde complicatie werden – indien mogelijk – altijd twee andere complicaties aangeboden. Als een complicatie driemaal op rij correct benoemd was, werd deze niet langer gevraagd. Dit herhaalde zich totdat de complicaties elk driemaal achtereenvolgens correct benoemd waren.

De **testfase** was een replicatie van de gewinningsfase, met de volgende aanpassing: Zodra de proefpersoon het gevoel kreeg dat er iets niet pluis was bij een proefpersoon moest hij op de corresponderende knop drukken, de monitor kleurde dan geel. Als het verder misging, en hij vond dat er een dokter moest komen, moest hij nogmaals op dezelfde knop drukken. De monitor kleurde dan rood. Bovendien moest hij hardop zeggen welke complicatie hij dacht dat er optrad, dit werd

door de experimentator genoteerd. De testfase duurde eveneens 15 minuten. Hierna was het experiment ten einde en vroegen we de proefpersoon kort naar zijn ervaring van het experiment.

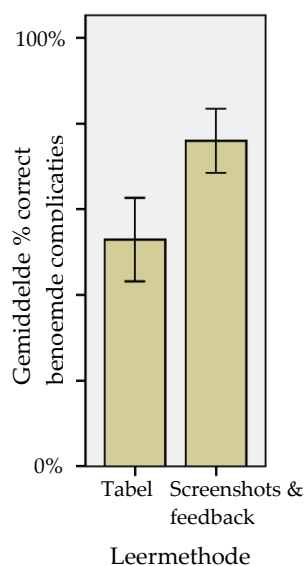
2.4. Apparatuur

Ons experiment was geprogrammeerd in PHP en Javascript. De dataopslag (MySQL) en aansturing van het programma gebeurde door een server elders in het gebouw, de precisie-tijdsmeting en grafische rendering van het experiment gebeurde op het lokale werkstation. Dat was een Samsung N150, Intel Atom N450 1660Mhz, 2GB DDR2. Het experiment draaide in Internet Explorer 9 op Windows 7 (beide van Microsoft). Het werd afgebeeld een 19" iiyama CRT-monitor (1024x786, 100Hz), invoer geschiedde d.m.v. een standaard optische muis en het numerieke blok van een standaard toetsenbord.

2.5. Dataverzameling

Gedurende het experiment werd van elke toetsdruk opgeslagen wanneer (in ms) en op welke toets gedrukt was. De genoteerde complicatienamen zijn achteraf toegevoegd aan de opgeslagen data. We hebben voor alle twaalf complicaties een starttijdstip vastgesteld, ongeveer 60 seconden voor het optreden van de complicatie. Alle detectie- en herkenningstijden zijn gerekend vanaf deze tijdstippen.

De verstreken tijd tussen dit starttijdstip en



Figuur 3.1: Algehele accuratesse van beide leergroepen.

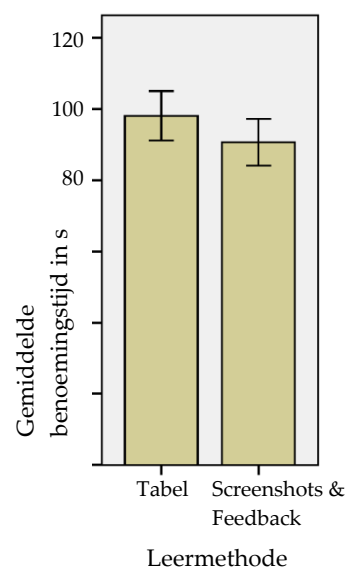
de (eerste) toetsdruk door de proefpersoon noemen we de *detectietijd*. Verder hebben we in de testfase ook de *herkenningstijd* berekend, dat is de verstreken tijd tussen het bovengenoemde starttijdstip en de tweede toetsdruk. Deze databewerking is gedaan d.m.v. een script geschreven in Borland Turbo Basic. De tijd die elke proefpersoon voor de leerfase nodig had werd ook berekend en opgeslagen. We berekenden ook het percentage "sterfgevallen" oftewel niet-gedetecteerde complicaties.

De statistische analyse is gedaan met PASW Statistics 18 (SPSS Inc.). We gebruikten de t-toets voor twee onafhankelijke steekproeven en de meerwegs variantieanalyse om verschillen tussen beide groepen te testen.

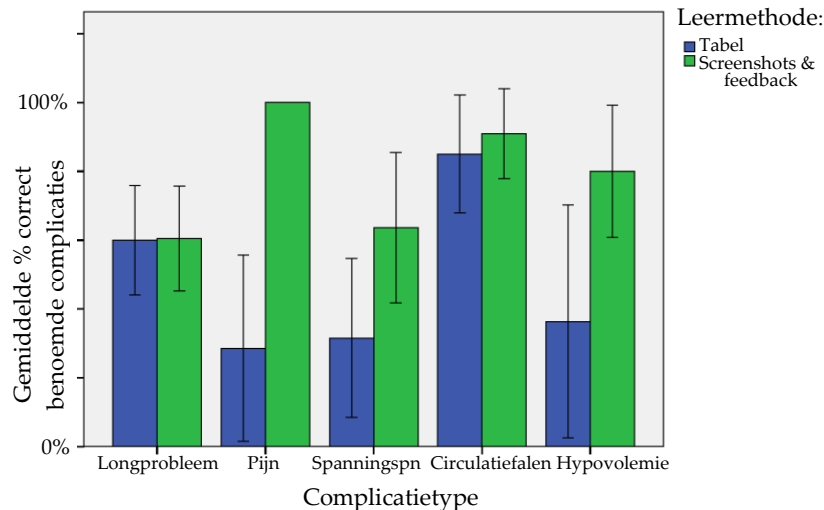
3. Resultaten

3.1. Accuratesse

We hebben een meerwegs variantieanalyse uitgevoerd om het effect van de variabelen groep en complicatietype te testen. We vonden een significant effect van groep op het percentage correct benoemde complicaties in de testfase ($F(1,232) = 4.92, p = 0.001$). Groep A (die van de tabel had geleerd) herkende gemiddeld 53% van de complicaties correct, groep B (die het impliciete leertraject doorlopen had) herkende gemiddeld 76% correct (zie figuur 3.1). Verder vonden we een significant interactie-effect tussen groep en complicatietype ($F(4,232) = 4.97, p = 0.001$). Dit effect is duidelijk waarneembaar in figuur 3.2. Bij de ene complicatie is het verschil



Figuur 3.3: Herkenningstijd van beide leergroepen.



Figuur 3.2: Accuratesse van beide leergroepen, opgesplitst op complicatietype. Het verschil tussen de beide groepen is erg groot bij de complicatie ‘pijn’, maar nihil bij de complicatie ‘longprobleem’.

tussen beide groepen nihil, terwijl het verschil bij een andere complicatie aanzienlijk is.

3.2. Herkenningstijd

Een meerwegs variantieanalyse liet een significant effect zien van groep op herkenningstijd ($F(1,232) = 4.42, p = 0.037$). Groep B herkende een complicatie gemiddeld iets eerder dan groep A (zie figuur 3.3). We vonden geen significant interactie-effect tussen groep en herkenningstijd.

3.3. Sterfgevallen

Er is geen significant effect gevonden van groep op het percentage sterfgevallen (oftewel niet-gedetecteerde complicaties). Een t-toets voor twee onafhankelijke steekproeven liet wel zien dat het aantal sterfgevallen in de testfase significant lager was dan in de gewinningsfase ($t(250) = -0.92, p = 0.361$). In de gewinningsfase werd 14% van de complicaties over het hoofd gezien, in de testfase bedroeg dit 6%.

4. Discussie

Het aantal sterfgevallen is na het leren significant lager dan voor het leren. Hieruit mogen we concluderen dat het leren in ieder geval niet voor niets is geweest, er werden na het leren meer complicaties gedetecteerd dan vóór het leren. Verder bevestigen de resultaten onze hypothese: De groep die het interactieve leerprogramma had doorlopen was sneller en beter in het herkennen van complicaties. We vroegen ons af of er ook een verband bestond

tussen de tijdsduur van het leren en de prestatie tijdens de testfase, maar we hebben geen significante correlatie gevonden.

De belangrijkste bevinding van ons onderzoek is dat de accuratesse en de snelheid van complicatieherkenning significant verschilt tussen de beide groepen. De proefpersonen aan wie 15 screenshots in een trial-and-error concept werden aangeboden wisten een complicatie vaker correct te benoemen, en deden dat ook nog eens sneller. Een opvallend resultaat, want juist de groep die de expliciete instructie kreeg – “leer dit!” – heeft minder goed gepresteerd.

Blijkbaar levert het leren van een tabel voor deze taak niet het optimale resultaat op, ondanks dat de proefpersonen ten tijde van het leren al wisten wat de taak inhield. Het is de vraag wat precies de oorzaak is van het prestatieverschil tussen beide groepen.

Mogelijk wordt dit primair veroorzaakt is door het gebruik van screenshots in het leertraject. Per complicatie waren er drie screenshots. Doordat deze onderling verschilden, was het niet afdoende om slechts één exemplaar van elke complicatie te onthouden. We gaan er daarom vanuit dat de proefpersonen een enigszins abstracte representatie van een complicatie hebben gevormd. Wellicht stelt dit mechanisme ook in staat om een complicatie in een vroeger stadium te detecteren dan het stadium waarin de screenshots genomen waren.

Het is – anderzijds – ook mogelijk dat het gevonden verschil primair aan het testing effect

is toe te schrijven. De proefpersonen die vanaf de tabel leerden moesten zich actief tot leren zetten, terwijl het leren van de andere groep onbewust vanzelf gebeurde. Een replicatie van ons experiment zou uit kunnen wijzen in hoeverre al deze factoren een rol spelen, door ze dan uit te splitsen over verschillende groepen.

Bij de proefpersonen die het interactieve leerprogramma doorliepen kregen we de indruk dat ze in de testfase het best presteerden op de complicaties waar ze in de leerfase de meeste moeite mee hadden. Zou dit kunnen komen doordat deze complicaties vaker zijn herhaald in het leerproces? Of doordat ze langer in het leerprogramma opgenomen waren, en dus verser in het geheugen lagen? Verder onderzoek is nodig om dit te bepalen, en uit te wijzen hoe de leerprestaties op dit punt gemaximaliseerd kunnen worden.

Bij dezelfde groep deden we ook nog de volgende observatie. De eerste keer dat proefpersonen in het leertraject een complicatie te zien kregen, moesten ze simpelweg raden welke het was. Dit ging meestal niet in één keer goed. Maar eigenaardig genoeg lieten veel proefpersonen gedurende de rest van het leertraject de muiscursor altijd eerst nog even op deze initiële (foute) keuze hangen, voordat ze uiteindelijk de goede (geleerde) keuze aanklikten. Hun eerste gedachte bij een bepaalde complicatie werd blijkbaar niet vergeten, ondanks het verdere leertraject. Als we hierop inspelen zijn de leerprestaties wellicht verder te verbeteren, bijvoorbeeld door, bij aanvang van het leerprogramma, eerst voor elke complicatie eenmaal het juiste antwoord aan te bieden.

Het belang van snel en accuraat herkennen van complicaties is groot, en kan mensenlevens redden. Ons experiment was gebaseerd op de zeshoekmonitor van Doesburg (2011) en Moes (2012), maar misschien treedt het prestatieverschil ook op bij gebruik van de conventionele anesthesische bewakingsmonitor. Als vervolgonderzoek dit aantoont, dan heeft dat volgens ons concrete consequenties voor medische opleidingen. Medici leren over het algemeen door hun leerboeken te bestuderen en door het stampen van lijstjes, wat overeenkomsten vertoont met het leren van de complicaties vanaf een tabel, zoals dat in ons experiment gebeurde. Een onderzoek naar de

trainingsmethoden voor anesthesisten en IC-verpleegkundigen is daarom wenselijk. Als we hierdoor letsel kunnen voorkomen of wellicht mensenlevens kunnen redden, is het de moeite zeker waard.

Referenties

- Cook, R.I., Potter, S.S., Woods, D.D., & McDonald, J.S. (1991). Evaluating the human engineering of microprocessor-controlled operating room devices. *Journal of Clinical Monitoring*, 7(3), 217-226.
- Cooper, J.B., Newbower, R.S., & Kitz, R.J. (1984). An analysis of major errors and equipment failures in anesthesia management: Considerations for prevention and detection. *Anesthesiology*, 60(1), 34-42.
- Doesburg, T. (2011). Mobile monitoring in anaesthesia using smartphones. *Unpublished master thesis*, Rijksuniversiteit Groningen, Nederland.
- Drews, F.A., & Westenskow, D.R. (2006). The right picture is worth a thousand numbers: Data displays in anesthesia. *Human Factors*, 40(1), 59-71.
- Effken, J.A., Kim, N., & Shaw, R.E. (1994). Making the relationships visible: Testing alternative display design strategies for teaching principles of hemodynamic monitoring and treatment. *Journal of the American Medical Informatic Association*, 949-953.
- Eichhorn, J.H. (1989). Prevention of intraoperative anesthesia accidents and related severe injury through safety monitoring. *Anesthesiology*, 70(4), 572-577.
- Gurushanthaiah, K., Weinger, M.B., & Englund, C.E. (1995). Visual-display format affects the ability of anesthesiologists to detect acute physiologic changes - A laboratory study employing a clinical display simulator. *Anesthesiology*, 83(6), 1184-1193.
- Kang, S.H. (2010). Enhancing visuospatial learning: The benefit of retrieval practice. *Memory & Cognition*, 38(8), 1009-1017.
- Kang, S.H., McDaniel, M.A., & Pashler, H. (2011). Effects of testing on learning of functions. *Psychonomic Bulletin & Review*, 18(5), 998-1005.
- Michels, P., Gravenstein, D., & Westenskow, D. R. (1997). An integrated graphic data display improves detection and identification of

- critical events during anesthesia. *Journal of Clinical Monitoring*, 13(4), 249-259.
- Moes, P. (2012). Using a polygonal display in a mobile anesthesia monitor. *Unpublished master thesis*, Rijksuniversiteit Groningen, Nederland.
- Van Amsterdam, K., Cnossen, F., Ballast, A., & Struys, M.M. (submitted). Integration of visual metaphors in an anesthesia monitor. Rijksuniversiteit Groningen, Nederland.
- Wickens, C.D., & Hollands, J.G. (1999). *Engineering Psychology and Human Performance (Third Edition)*. New Jersey: Prentice-Hall Inc.