

**A VIZUÁLIS KATEGORIZÁCIÓS
FOLYAMATOK IGEDRENDSZERI
HÁTTERÉNEK VIZSGÁLATA
ALACSONY ÉS MAGAS TÉRBELI
FREKVENCIÁJÚ KÉPEK
SEGÍTSÉGÉVEL**

Rokszin Adrienn Aranka

SZTE BTK Neveléstudományi Doktori Iskola, Szeged,
Magyarország

Csifcsák Gábor

SZTE BTK Pszichológiai Intézet, Szeged,
Magyarország

A VIZUÁLIS KATEGORIZÁCIÓS FOLYAMATOK IGEDRENDSZERI HÁTTERÉNEK VIZSGÁLATA ALACSONY ÉS MAGAS TÉRBELI FREKVENCIÁJÚ KÉPEK SEGÍTSÉGÉVEL¹

Absztrakt

Jelen kutatás célja, annak vizsgálata, hogy egy fölérendelt szintű vizuális kategorizációs feladatban milyen hatással van az alacsony, illetve magas térbeli frekvenciák kivonása a viselkedéses mutatókra és az azokkal összefüggő, 0-300 milliszekundum között jelentkező eseményfüggő potenciálokra. A vizsgálat során 12 egészséges személyt kértünk fel állat-jármű kategorizációs feladat végzésére. Az állatokat, illetve járműveket ábrázoló szürkeárnyaltos eredeti képek mellett, azok alacsony és magas térbeli frekvenciájú változatai is bemutatásra kerültek. A feladat végzése alatt 32 csatornás EEG mérést végeztünk. Az elemzés során a három képtípus esetén jelentkező skalpeloszlásokat vetettük össze az inger bemutatását követő 0-300 ms közötti időintervallumban. Az EEG adatok arra utalnak, hogy az alacsony, illetve a magas térbeli frekvenciák eltérően befolyásolják az agyi aktivitást a feldolgozás korai szakaszában, és ez nem feltétlenül tükrözi a képeken szereplő tárgyra adott kategorizációs válaszok eltéréseit.

Kulcsszavak: vizuális kategorizáció, EEG, magnocelluláris pálya, parvocelluláris pálya, térbeli frekvencia

¹ „A kutatás a TÁMOP 4.2.4.A/2-11-1-2012-0001 azonosító számú *Nemzeti Kiválóság Program – Hazai hallgatói, illetve kutatói személyi támogatást biztosító rendszer kidolgozása és működtetése konvergencia program* című kiemelt projekt keretében zajlott. A projekt az Európai Unió támogatásával, az Európai Szociális Alap társfinanszírozásával valósul meg.”

Bevezetés

Manapság széles körben elfogadottak azok az elméletek, melyek a vizuális információ parallel feldolgozását hangsúlyozzák. Az egyik ilyen elmélet a retina különböző méretű ganglionsejt-osztályaiból indul ki, és ez alapján megkülönböztet egymástól magnocelluláris és parvocelluláris sejteket és a hozzájuk tartozó látópályákat. Számos különbség figyelhető meg a két pálya között. Míg a magnocelluláris pálya nagyobb érzékenységet mutat az alacsony térbeli frekvenciájú (ATF), dinamikus és akromatikus ingerekre, addig a parvocelluláris pályát a magas térbeli frekvenciákra (MTF) való érzékenység, a rossz idői felbontás jellemzi, továbbá fontos szerepek játszik a színészlelésben is (Livingstone and Hubel 1988)

Milyen szerepet tölt be a két pálya a vizuális kategorizációban? Bar (2003) modellje szerint komplex ingerek esetén a magnocelluláris pálya által gyorsabban szállított, rosszabb térbeli felbontású információ segítheti az inger ún. „top-down” jellegű gyors és hatékony azonosítását/kategorizációját, ezzel ellentétben a lassabb parvocelluláris pálya a részletazonosításban játszik fontos szerepet, és valójában a magnocelluláris információ által felállított predikciókat („A képen valószínűleg egy macska van”) erősíti („Ez tényleg egy macska”) vagy cáfolja („Ez nem egy macska, hanem egy kutya”).

Napjainkban a human kategorizációs folyamatok idegrendszeri hátterének vizsgálatára az egyik legszélesebb körben alkalmazott mérőeljárás az elektroencefalográfia (EEG), amely az agykérgi neuronok összesített elektromos aktivitását tükrözi. Az EEG mérés egy nem invazív eljárás, mely során az elektromos jelek elvezetése a skalpra helyezett elektródák segítségével történik (Kéri and Gulyás 2003). Vizsgálatunk során mi is ezt az eszközt alkalmaztuk az inger bemutatását követő 0-300 ms között jelentkező agyi aktivitás vizsgálatára.

Érdekes módon számos olyan eredményt találunk, amely szoros kapcsolatot feltételez a vizuális eseményfüggő potenciál két korai komponense, a P1 (az inger bemutatását követően 100 ms-mal, okcipitálisan jelentkező pozitív hullám) és N1 - vagy N70 hullám - (50-100 ms-os latenciával, a parieto-okcipitális területek felett jelentkező negatív hullám) (Kubová et al. 1995; Luck 2005) illetve a parvocelluláris és magnocelluláris pályák között. Úgy tűnik, hogy míg a P1 komponens a magnocelluláris folyamatokkal, addig az N1-es hullám a parvocelluláris folyamatokkal hozható összefüggésbe (Tobimatsu, Tomoda and Kato 1995; Rudvin, Valberg and Kilavik, 2000; Hansen et al. 2011). ElleMBERG és munkatársai (2001) vizsgálatukban ún. szinuszos rácsmintázatok térbeli frekvenciáját és kontrasztját manipulálták. Eredményeik szerint az általuk alkalmazott legalacsonyabb térbeli frekvenciák mellett csak a P1-es volt azonosítható, a térbeli frekvencia fokozatos növelésével azonban egyre dominánsabbá vált az N1-es hullám.

Vizsgálatunkban arra voltunk kíváncsiak, egy fölrendelt szintű vizuális kategorizációs feladatban, milyen hatással vannak az alacsony, illetve magas térbeli frekvenciájú módosított képtípusok a korai, 0-300 milliszekundum között jelentkező eseményfüggő potenciálokra. A korábbi vizsgálatokkal ellentétben, állatokat és járműveket ábrázoló komplex képek alacsony, illetve magas térbeli frekvenciájú változatait alkalmaztuk. A szakirodalom alapján azt feltételezzük, hogy míg a P1 komponens időintervallumában (120-140 ms) jelentkező agyi aktivitást az ATF képek erősebben modulálják, addig az N1 hullám esetében a MTF képekre figyelhető meg eltérő agyi aktivitásmintázat.

Módszerek

Résztevők

A kutatásban 12 egészséges személy vett részt (4 férfi, 8 nő (átlagéletkor= 25,5, SD= 2,236). A résztvevők egyike sem szenvedett fejlődési, pszichiátriai vagy neurológiai rendellenességben. Rákérdeztünk a személyek látásélességére, ami minden esetben szükség szerint korrigálva volt. Minden résztvevő aláírta az etikai bizottság által jóváhagyott tájékoztatás utáni beleegyező nyilatkozatot.

Ingerek

Az ingerek bemutatására, a feladatban nyújtott teljesítmény és reakcióidő rögzítésére az E-Prime számítógépes programot (Psychology Software Tools, Inc., USA) alkalmaztuk.

A vizsgálati személyek állatokat és járműveket ábrázoló szürkeárnyaltos eredeti képek mellett, azok alacsony és magas térbeli frekvenciájú változatairól hoztak állat-jármű kategorizációs döntést (1. ábra). Az eredeti képeket a Li, VanRullen, Koch és Perona (2002) által alkalmazott, az interneten is elérhető adatbázisból (<http://vision.stanford.edu/Datasets/AnimTransDistr.rar>) válogattuk össze. A képek felbontása, mérete és luminanciája megegyezett.

Az eredetileg színes képek szürkeárnyaltossá alakítása, továbbá az alacsony és magas térbeli frekvenciájú módosított képtípusok előállítását az Adobe Photoshop CS5 szoftver (Adobe Systems Inc., San Jose, USA) segítségével történt. Az ATF képek esetében a Gaussian blur szűrőt (sugár: 6,1 pixel), a MTF ingereknél pedig a felüláteresztő szűrőt (sugár: 0,5 pixel) használtuk.

A 3 féle kondícióban (eredeti, ATF és MTF ingerek) állat-jármű kategóriánként 100-100, azaz összesen 600 képet mutattunk be, random sorrendben.



1. ábra: A kutatásban használt eredeti (bal oszlop), ATF (középső oszlop), illetve MTF (jobb oszlop) képekre látható példa, állat (felső sor) és jármű (alsó sor) kategóriában.

Kísérleti elrendezés

A személyek egy elsötétített, hangszigetelt szobában, a monitortól 130 cm-re ülve végezték a feladatot. A válaszadás egy speciálisan erre a célra preparált billentyűzettel, két válaszgomb valamelyikének lenyomásával történt.

A kísérlet során a képernyőn elsőként 1000 ms-ig egy fixációs pont, majd ezt követően a három kondíció valamelyikéből egy kép jelent meg, amelyet a személy 500 ms-ig látott. Ezt követően 1500 ms idő állt rendelkezésre a válaszadáshoz, majd újra egy fixációs képernyő jelent meg. Az adatfelvétel három blokkban történt és megközelítőleg 20 percig tartott.

EEG regisztráció

Az EEG adatok felvételéhez a BioSemi cég (BioSemi B.V., Amsterdam, Hollandia) 32 csatornás EEG készülékét és a rendszerhez tartozó szoftvert használtuk.

Az elektródákat a nemzetközi 10-20-as rendszernek megfelelően helyeztük el a skalpon egy előre gyártott BioSemi EEG sapka segítségével. A regisztráció során a referencia és a föld elektródák a fejtetőn (CMS és DRL elektródák) voltak. Az adatrögzítés 0,01 Hz-es alsó és 100 Hz-es felső frekvenciaszűrők segítségével történt.

EEG elemzés

Az EEG adatok elemzése a Matlab szoftver EEGLAB programcsomagjával (Delorme and Makeig 2004) történt.

Első lépésben az adatokat 1600 ms (az inger bemutatását megelőző 100 ms-os, illetve a bemutatást követő 1500 ms-os szakasz) hosszú szakaszokra szegmentáltuk, majd kondícióknak megfelelően szétválogattuk. Pislogás és egyéb szemmozgással kapcsolatos műtermékeket az általunk használt program segítségével szűrtük ki.

Az EEG elemzéshez három időintervallumot választottunk, az első a 120-140 ms közötti tartomány, mely a P1 komponensnek felel meg, a második időszak 170-190 ms között van, mely az N1 komponensnek feleltethető meg, illetve kiválasztottunk egy kései időintervallumot is, 230 és 290 ms között, melyben a P2 komponenst azonosíthatjuk.

Statisztika

A viselkedéses adatok esetében ismételt méréses varianciaanalízist alkalmaztunk annak vizsgálatára, hogy a KÉPTÍPUS (eredeti, ATF, MTF képek) milyen befolyással van a pontosságra, illetve a reakcióidőre? A reakcióidőt személyenként és kondícióként mediánokban számoltuk ki. A pontosság meghatározására a szignáldetekciós elméletnek megfelelően, a d' értékeket számoltunk, melyet a helyes válaszok és a téves riasztások arányából határozhatunk meg. Az egyes képtípusokra jelentkező pontosság és reakcióidő adatok páronkénti összehasonlítására egyszerű kontrasztanalízist (simple contrast analysis) alkalmaztunk, mely során az egyes módosított képtípusokat az eredetivel hasonlítottuk össze. A jelentkező hatások nagyságának demonstrálására a parciális eta négyzet (η^2) értékeket is feltüntettük.

Az EEG adatok elemzése az EEGLAB-be programozott statisztikai elemző segítségével történt, mely során permutációs elemzéssel vetettük össze az eredeti, illetve módosított ingerekre jelentkező skalpeloszlásokat a három, általunk vizsgált időintervallumban, $p < 0,05$ -os szignifikancia szintet beállítva, és a nagyobb megbízhatóság érdekében false-discovery rate korrekciót használva.

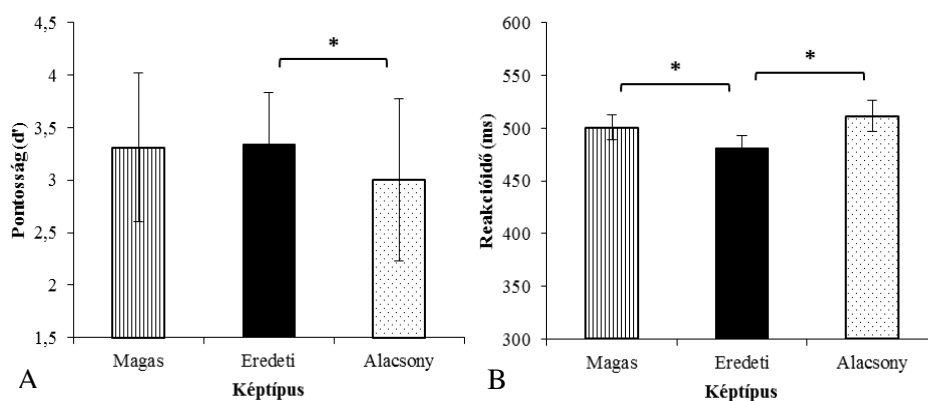
Eredmények

Viselkedéses adatok

A pontosság elemzése során a KÉPTÍPUS főhatás nem volt szignifikáns [$F(2,20)=3,217$, $p=0,061$; $\eta^2=0,243$]. A kontrasztanalízis eredménye szerint

csak eredeti és ATF képek összehasonlításánál figyelhető meg különbség [F(1,10)=5,497, $p=0,041$; $\eta^2=0,355$), az ATF képekre a személyek pontatlanabban teljesítenek (2/A ábra).

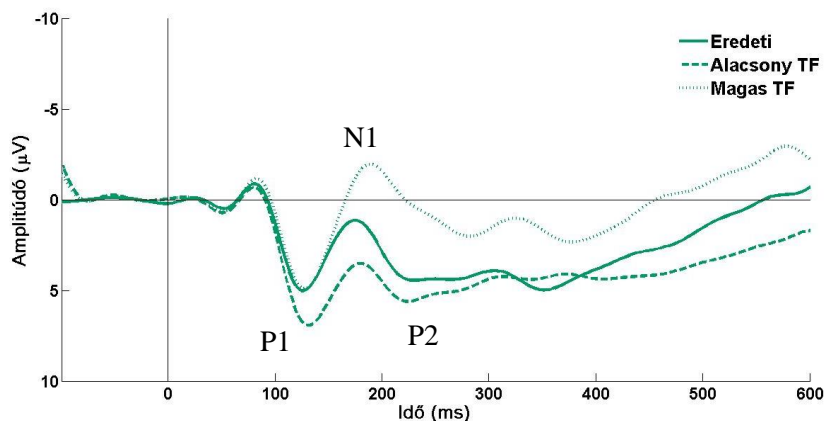
A reakcióidő vizsgálatok a KÉPTÍPUS főhatás szignifikánsnak bizonyult [F(2,20)=22,187; $p<0,00$; $\eta^2=0,689$). Az eredeti és az egyes módosított képtípusok összehasonlításakor mind a két esetben szignifikáns különbséget kaptunk, az eredeti képekre a személyek szignifikánsan gyorsabban válaszoltak (eredeti – MTF: F(1,10)=21,225; $p=0,001$; $\eta^2=0,68$; eredeti – ATF: F(1,10)=34,502; $p<0,001$; $\eta^2=0,075$) (2/B ábra).



2. ábra: Az ábra a három képtípus esetén mért pontosság (A) és reakcióidő (B) adatok átlagait és szórásait mutatja. A csillag a szignifikáns különbséget jelöli 0,05-os szignifikancia szint mellett.

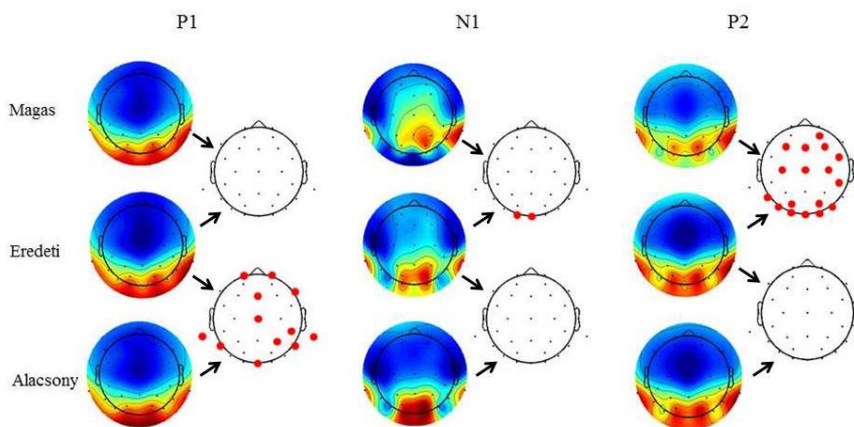
EEG adatok

Az eseményfüggő potenciálokat a három képtípus esetén a 3. ábra szemlélteti.



3. ábra: Az ábra az Oz elektródán regisztrált eseményfüggő potenciálokat mutatja eredeti, ATF, illetve MTF képek esetén.

Az EEGLAB-ben végzett statisztika szerint, a P1 komponens időintervallumában csak az eredeti és az ATF képtípusokra jelentkező agyi aktivitás összehasonlításakor figyelhető meg szignifikáns különbség, nemcsak hátsó területek - a látókéreg - felett, hanem frontálisan is. A P1 komponenssel ellentétben, az N1 esetén csak az eredeti és MTF képek összehasonlítása során tapasztaltunk szignifikáns különbséget, mely csak az okcipitális területeken észlelhető. A P2 komponensnél az N1-hez hasonlóan szintén csak az eredeti és MTF ingerek által kiváltott aktivitás különbözik szignifikánsan, ebben az időintervallumban azonban már az egész agyra kiterjedő szignifikáns képtípus-hatást találtunk. Az eredményeket a 4. ábra mutatja.



4. ábra: Az ábra a vizsgált időintervallumokban a három képtípus esetén jelentkező skalposzlásokat, illetve az EEGLAB-ben végzett statisztikát

mutatja. A piros pöttyök azokat az elektródákat jelölik a skalpon, ahol szignifikáns különbséget kaptunk két képtípus között.

Megvitatás

Jelen kutatás célja annak vizsgálata, hogy egy fölérendelt szintű állatjármű kategorizációs feladatban, az alacsony vagy a magas térbeli frekvenciák kivonása hogyan befolyásolja a korai 0-300 ms között jelentkező eseményfüggő potenciálokat.

Bár a reakcióidő vizsgálatakor mind a két módosított képtípus esetében a személyek lassabban teljesítettek az eredeti ingerekhez képest, a pontosság (d') adatok alapján úgy tűnik, hogy a magas térbeli frekvenciák hiánya jobban befolyásolja a vizuális kategorizációs feladatban nyújtott teljesítményt, mint az alacsony térbeli frekvenciák kivonása.

Az EEG adatok elemzésekor a P1 komponens időintervallumában, 120-140 ms között csak az ATF képek skalpeloszlása mutatott eltérést az eredeti képek által kiváltott agy aktivitással összevetve. Úgy tűnik tehát, hogy a magas térbeli frekvenciák kivonása már nagyon korán, főként a hátulsó területek felett megnyilvánulhat, és ez a pontosság adatokkal is összefügg, hiszen e képtípust kategorizálták a legnehezebben résztvevőink. Ezen eredmények összhangban vannak a szakirodalomban olvasható eredményekkel is, hiszen számos vizsgálat rámutatott arra, hogy míg a P1 komponens a magnocelluláris, folyamatokkal hozható összefüggésbe (pl. Elleberg et al. 2001; Hansen et al. 2011). Fontos azonban hangsúlyozni, hogy a témában olvasható szakirodalmakban az N1 hullámot időben másképp definiálták, pontosabban az N1-et az inger követő 50 és 100 ms közötti legnegatívabb pontként határozták meg. A mi esetünkben, ebben az időintervallumban csak igen kis amplitúdójú N1 volt észlelhető, amely amplitúdója nem változott az egyes kondíciók között (3. ábra). Így kutatásunkban N1 komponensként az inger bemutatását követően megjelenő első negatív hullámot definiáltuk, mely megközelítőleg 170 és 190 ms között volt regisztrálható. Ez nagy valószínűséggel abból adódhat, hogy míg az eddigi vizsgálatok többsége más ingertípusokat, például ún. szinuszos rácsmintázatokat (pl. Elleberg et al. 2001) vagy izolumináns színkontrasztokat alkalmaztak (Foxe et al. 2008), addig a jelen vizsgálatban résztvevő személyek állatokat, illetve járműveket ábrázoló komplex képeken végeztek kategorizációs feladatot. Az általunk észlelt, későbbi N1 összhangban van a komplex ingerek kategorizációjával kapcsolatos irodalommal (pl. Antal et al. 2000). Fontos különbség figyelhető meg továbbá az elvégzendő feladat típusában is. A szakirodalomban olvasható kutatásokkal ellentétben az általunk felkért személyek komplex képeken végeztek kategorizációs feladatot. Vizsgálatunkban a 170-190 ms-ig terjedő időintervallumban, vagyis az N1 komponens időintervallumában csak az eredeti képtípussal és a MTF képekre

jelentkező agyi aktivitásmintázat tért el szignifikánsan, és csak a látókéreg felett elhelyezett elektródák esetében. A P2 komponensnél az egész skalpra kiterjedő képtípus-hatást figyelhetünk meg szintén az eredeti és MTF párok összehasonlításakor. Fontos kiemelni, hogy a szakirodalomban a P1 hullámot követő komponensekre vonatkozóan nem találunk semmilyen adatot, így kutatásunk mindenféleképpen új eredményekkel szolgál ezen témakörben.

Összegezve, eredményeink arra utalnak, hogy az alacsony, illetve a magas térbeli frekvenciák eltérően befolyásolják az agyi aktivitást a feldolgozás korai szakaszában. Míg az alacsony térbeli frekvenciák a P1 hullámot, addig a magas térbeli frekvenciák az N1, és főként a P2 komponenst befolyásolják. Az EEG és viselkedési adatok összevetése arra utal, hogy minél korábban jelentkezik az idegi aktivitás eltérése az eredeti képekhez viszonyítva, annál inkább tapasztalható teljesítményromlás komplex képek kategorizációja esetén.

Referenciák

- Antal, A., Kéri, Sz., Kovács, Gy., Janka, Z. and Benedek, Gy. 2000. Early and late components of visual categorization: an event-related potential study. *Cognitive Brain Research* 9(1):117-119.
- Bar, M. 2003. A Cortical Mechanism for Triggering Top-Down Facilitation in Visual Object Recognition. *Journal of Cognitive Neurosciences* 15(4):600-609.
- Delorme, A. and Makeig, S. 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods* 134(1):9-21.
- Elleberg, D., Hammarrenger, B., Lepore, F., Roy, M.S. and Guillemot, J.P. 2001. Contrast dependency of VEPs as a function of spatial frequency: The parvocellular and magnocellular contributions to human VEPs. *Spatial Vision* 15(1):99-111.
- Foxe, J.J., Strugstad, E.D., Sehatpour, P., Molholm, S., Pasiaka, W., Schroeder, C.E. and McCourt, M.E. 2008. Parvocellular and Magnocellular Contributions to the Initial Generators of the Visual Evoked Potential: High-Density Electrical Mapping of the "C1" Component. *Brain Topography* 21(1):11-21.
- Hansen, B.C., Jacques, T., Johnson, A.P. and Elleberg, D. 2011. From spatial frequency contrast to edge preponderance: the differential modulation of early visual evoked potentials by natural scene stimuli. *Visual Neuroscience* 28(3):221-237.
- Kéri, Sz. and Gulyás, B. 2003. Elektrofiziológiai módszerek a kognitív idegtudományban. in: Pléh, Cs., Kovács, Gy. and Gulyás, B. eds. *Kognitív idegtudomány*. Budapest: Osiris Kiadó. 202-216.
- Kubová, Z., Kuba, M., Spekreijse, H. and Blakemore, C. 1995. Contrast dependence of motion-onset and pattern-reversal evoked-potentials. *Vision Research* 35(2):197-205.

- Li, F.F., VanRullen, R., Koch, C. and Perona, P. 2002. Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences of the United States of America* 99(14):9596–9601.
- Livingstone, M. and Hubel, D. 1988. Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science* 240(4853):740-749.
- Luck, S.J. 2005. *An Introduction to the Event-Related Potential Technique*. Cambridge: The MIT Press.
- Rudvin, I., Valberg, A. and Kilavik, B.E. 2000. Visual evoked potentials and magnocellular and parvocellular segregation. *Visual Neuroscience* 17(4):579–590.
- Tobimatsu, S., Tomoda, H. and Kato, M. 1995. Parvocellular and magnocellular contributions to visual-evoked potentials in humans - stimulation with chromatic and achromatic gratings and apparent motion. *Journal of the Neurological Sciences* 134(1-2):73–82.