UNIVERSITAT JAUME·I

Programa de Doctorat en Ciències

Escola de Doctorat de la Universitat Jaume I

Neuroscience in algebraic problem solving: Studying the reversal error

---

Memòria presentada per Lara Ferrando Esteve per a optar al grau de doctora per la Universitat Jaume I

Doctoranda:
Lara Ferrando Esteve

Directores:
Irene Epifanio López
Noelia Ventura Campos

Castelló de la Plana, octubre 2020

LARA| FERRANDO| ESTEVE
Firmado digitalmente por LARA|FERRANDO| ESTEVE
Fecha: 2020.12.07 13:15:40 +01'00'

IRENE| EPIFANIO|LOPEZ
Firmado digitalmente por IRENE| EPIFANIO|LOPEZ
Nombre de reconocimiento (DN): cn=IRENE|EPIFANIO|LOPEZ, serialNumber=22568201A, givenName=IRENE, sn=EPIFANIO LOPEZ, ou=CIUDADANOS, o=ACCV, c=ES
Fecha: 2020.12.07 13:02:38 +01'00'

Noelia Ventura Campos
Firmado digitalmente por Noelia Ventura Campos
Fecha: 2020.12.07 13:12:32 +01'00'

Finançament rebut

Aquesta tesi està recolzada per les subvencions següents: UJI-A2017-8 i UJI-B2017-13 de la Universitat Jaume I.

A Lia,
pels seus consells diaris.

A David,
per ser el meu suport.

A les tres estreles més brillants del cel i la que brilla des de la terra,
per haver-me donat als meus pares.

Als meus pares,
per tot el que fan per mi.

"No hi ha branca de la matemàtica, per abstracta que siga, que no puga aplicar-se algun dia als fenòmens del món real." *Nikolai Lobachevski.*

"Els problemes són situacions sense una solució òbvia. Si no cal pensar, no hi ha problema." *(OECD, 2014, vol. V, p.1).*

Agraïments

En escriure aquestes línies, sóc conscient de que està finalitzant i a la vegada començant una nova etapa.

M'agradaria agrair en primer lloc a les meues directores de tesi. A tu Noelia, per deixar endinsar-me en un món tan interessant com és el de la neuroeducació, per tots els coneixements que m'has transmés i per haver-me acollit en els braços oberts des d'un primer moment. Gràcies Noelia. A tu Irene, per la grandíssima paciència i els teus consells que m'han ajudat a poder aplegar on estic, i com no, per la teua implicació diària. Eres d'admirar. Gràcies Irene.

No puc oblidar-me de vosaltres, els meus companys d'universitat, a Paco i a Ximo, per totes les hores que hem passat junts i per tot el vostre suport en els moments més difícils. I com no, a tu Aleix, gràcies per aquestos últims anys d'universitat, on amb tu tot és un aprenentatge continu, gràcies per tot amic. Tampoc puc oblidar-me dels meus companys i companyes del grup: Anna, Jesús, Eli, Lidón, Esteban, Naiara i Helena. Sou un exemple a seguir, gràcies per la vostra amistat i per fer que amb vosaltres tot siga més senzill.

Per últim no puc oblidar-me de la meua família. Gràcies mare i pare per confiar amb mi. Gràcies Lia pels teus consells diaris. Gràcies *uelita*, per no soltar-me mai. Gràcies Javier i Óscar per haver-me animat tant. I a tu David, gràcies per ser el culpable d'estar escrivint estes línies ara mateix. De tu no m'oblide Jordi, el dia que lliges aquestes línies, la tia t'explicarà com tan sols amb un somriure em feies que m'inspirara.

Sigles i acrònims

AC Arbres de Classificació
ADF Anàlisi Discriminant Flexible
ADL Anàlisi Discriminant Lineal
ADP Anàlisi Discriminant Penalitzada
ADQ Anàlisi Discriminant Quadràtica
ADR Anàlisi Discriminant Regularitzada
BA Boscos Aleatoris
CART Arbres de classificació i regressió
EI Error Inversió
FBR Funció Base Radial
FDA Anàlisi de Dades Funcionals
FICA Anàlisi de Components Independents Funcional
FPCA Anàlisi de Components Principals Funcional
FWHM Full Width Half Maxium
IPS Sulcus intraparietal
IRM Imatges de Ressonància Magnètica
ISI Interval interestimular
KV K veïns més propers
LCR Líquid Cefaloraquidi
MNI Montreal Neurological Institut
MLG Model Lineal General
MLP Multilayer Perceptron
MVS Màquines de Vectors Suport
no-EI No Error Inversió
PADL Anàlisi discriminant lineal de Fisher penalitzada
PISA Programme for International Student Assessment
RL Regressió Logística
RM Ressonància Magnètica
RMe Ressonància Magnètica Estructural
RMf Ressonància Magnètica Funcional
SB Substància Blanca
SG Substància Grisa
SPHARM Spherical harmonic representation

SPM Statistical Parametric Mapping
VBM Morfometria Basada en el Vòxel
XN Xarxes Neuronals

Índex

Capítol 1

Introducció

1.1 Plantejament

Temps enrere, en l'època dels egipcis, podem imaginar a un dels escribes del faraó plantejant els càlculs matemàtics per construir la coneguda piràmide de Keops, també coneguda com la piràmide de Giza. No obstant, la primera mostra escrita de problemes matemàtics va ser trobada cap a mitat del segle XIX en Thebes. Aquesta mostra fou el Papir de Rhind, també conegut com el Papir d'Ahmes, el qual es va convertir en el document amb continguts matemàtics més antic del món (Robins i Shute, 1987).

Tal i com destaquen Ames et al. (2018) en la seua obra, al principi del papir es data el document i es presenta el nom de l'escriba que el va confeccionar, que va ser Amosis, nom reconegut com Ahmes. Encara que el papir va ser redactat durant el regnat d'Apofis I, fet que ens fa datar-lo durant el 1650-1530 a.C., d'acord amb el que va escriure Ahmes, la còpia original d'aquest tractat matemàtic estaria confeccionada durant el regnat del faraó Amenemes III (1818-1773 a.C.).

En el papir de Rhind es troben diferents problemes tant d'aritmètica com de geometria i medició. Per exemple, un dels problemes que consten en aquest papir era calcular en quin angle havien de ser tallades les pedres per tal de poder recobrir una piràmide amb pedra (Farrington, 1984). Per tant, encara que el papir de Rhind es data cap al 1600 a.C, des que l'ésser humà existeix, en el moment que nasquem, ens apareixen problemes que hem de resoldre, algunes vegades amb més aparença matemàtica que d'altres, però no deixen de ser problemes, els quals necessiten d'un raonament per a la seua resolució.

Cal destacar que la resolució de problemes és una tasca habitual en el nostre dia a dia i un

dels elements centrals en l'ensenyament de les matemàtiques. És per això que, durant l'educació primària i secundària, com a part dels continguts d'aritmètica i àlgebra, solen presentar-se problemes verbals en situacions reals i quotidianes. Aquest fet va produir que la presència d'aquest tipus de problemes fora un dels tres criteris que s'utilitzaren en l'avaluació PISA (Programme for International Student Assessment) 2012 (PISA, 2014).

Des que en l'informe PISA 2003 (OECD, 2005) foren sotmeses les competències que s'inclo-ïen en l'àrea de resolució de problemes, s'ha produït una gran quantitat d'avanços, permetent una major exactitud en l'avaluació. Va ser en l'informe PISA 2012 quan es dissenyaren noves metodologies per a l'avaluació, on destacava conèixer la interacció de l'alumnat amb el proble-ma. En aquest any es va començar a avaluar la competència de l'alumnat per a la resolució de problemes amb la característica de que aquestos problemes es podien presentar en la vida quo-tidiana. Els resultats que es van obtindre mostraren que l'alumnat espanyol mostra diferències significatives respecte a la mitjana de la OCDE, on Espanya, amb 477 punts, es troba per baix de la mitjana de la OCDE la qual està en 500 punts.

Per tant, allò realment interessant seria intentar proporcionar solucions en l'àmbit educatiu, intentant produir així una millora en els resultats de l'alumnat espanyol. Però, per poder proporcionar solucions, s'hauria de localitzar el problema i el primer pas és investigar sobre la resolució de problemes.

Tal i com afirmen Piñeiro et al. (2015), s'ha observat que una de les línies d'investigació en les que s'ha centrat el personal investigador, els últims anys, és la resolució de problemes. Són molts els estudis que es poden trobar sobre la resolució de problemes, com per exemple el de J. Booth et al. (2014), on s'estudien els conceptes erronis al passar d'aritmètica al pensament algebraic; Speedie et al. (1976) on es mostren intervencions per a que es produïsca una millora durant la resolució de problemes; o el de Xin et al. (2008), on es discuteixen les implicacions d'integrar el raonament algebraic en les matemàtiques elementals.

No obstant això, es troben diverses investigacions (Clement (1982), Clement et al. (1981) i Clement et al. (1980), Rosnick i Clement (1980)), les quals destaquen un tipus d'error molt comú a l'hora de resoldre els problemes, que no solament comet l'estudiantat de secundària, sinó que era, i és, un error molt comú en estudiantat universitari i fins i tot en el professorat. Aquest error es caracteritza per la traducció incorrecta de diferents oracions del llenguatge natural a l'algebraic. Aquest error, el qual serà la base de la nostra investigació, va ser anomenat l'Error d'Inversió (EI), ja que al plantejar una sèrie d'oracions com per exemple "Escriu una equació fent ús de les variables E i P per a representar l'enunciat següent: 'Hi ha sis vegades més estudiants que professors en aquesta universitat', fent ús d'E per al nombre d'estudiants i P per al nombre de professors" (Clement, 1982), la majoria de les respostes incorrectes eren $P = 6 \cdot E$, la qual cosa implicava invertir l'ordre de les lletres enfront de la resposta correcta $E = 6 \cdot P$, on el nombre d'estudiants havia de ser 6 vegades el de professors. Per tant, el que feien era multiplicar per 6 el nombre de professors i així obtindre el mateix nombre d'estudiants (veure

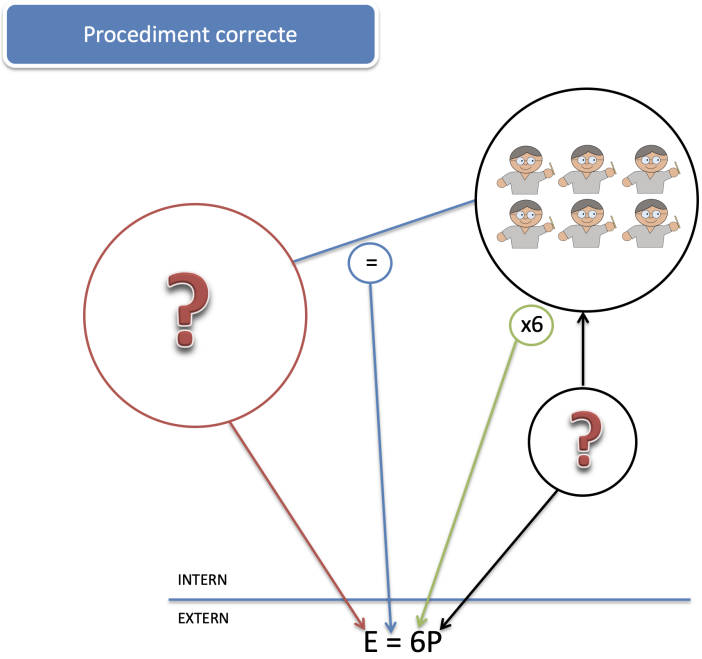


Figura 1.1: Procediment correcte (Clement, 1982).

la figura 1.1).

Diverses investigacions (L. Booth, 1984; Clement, 1982; Kuchemann, 1981; Mestre, 1988; Rosnick i Clement, 1980) han estudiat els errors en la resolució de problemes, alguns d'ells s'han centrat en intentar donar resposta a l'EI. És el cas de Clement (1982), qui va proposar dos models que podien donar una explicació a l'EI. El primer d'ells és el model explicatiu, al qual varen donar el nom de *coincidència en l'ordre de les paraules*, on es planteja com a causa d'aquest error una conversió literal de les paraules de l'enunciat a símbols matemàtics. Suposa una traducció lineal de l'enunciat d'esquerra a dreta, terme a terme, des del llenguatge natural a l'algebraic i proporciona una explicació plausible al fet que a l'enunciat "Hi ha sis vegades tants estudiants com a professors" es done la resposta incorrecta $6 \cdot \mathrm{E} = \mathrm{P}$. El segon model és una interpretació anomenada *comparació estàtica*, on s'afirmava que l'estudiant comprèn que el nombre d'estudiants és major que el de professors, i considera que l'expressió $6 \cdot \mathrm{E}$ vindria donada per l'agrupació major i P de la menor. En aquest cas, la lletra E no estaria sent considerada com una variable que representa el nombre d'estudiants, sinó com una etiqueta o unitat lligada al número 6, mentre que el signe igual representaria una comparació o una associació i no una equivalència exacta. Els estudiants mentalment intenten fer una associació de: *a un professor li corresponen 6 estudiants* i aquesta informació la tradueixen com $1 \cdot P = 6 \cdot E$. Encara que existeixen diverses investigacions per a veure quina d'aquestes interpretacions seria la causa de l'EI, no existeix actualment una evidència clara que ho explique. Aquest error es

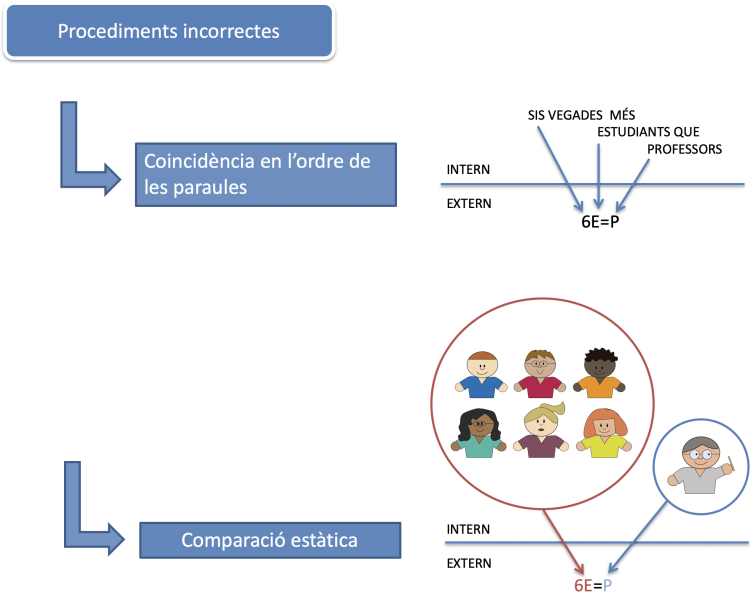


Figura 1.2: Models explicatius de l'error d'inversió (Clement, 1982).

produeix en representar equacions que plasmen tant comparacions multiplicatives com additives. Una explicació gràfica es pot veure a la figura 1.2.

Des de la publicació dels estudis de Clement (1982) i col·laboradors, diversos estudis han intentat determinar el poder explicatiu de la coincidència de l'ordre de les paraules i de la comparació estàtica (Cohen i Kanim, 2005; Cooper, 1986; Davis, 1984; González-Calero et al., 2015; Kirshner et al., 1991; Landy i Goldstone, 2007; López-Real, 1995; Rosnick, 1981). Amb aquest objectiu s'han dissenyat experiments en els quals es feien variar les característiques de les tasques. Entre les variables considerades es pot destacar: la presència de pistes contextuals; l'ús de diferents formats de la comparació; el caràcter additiu o multiplicatiu de la comparació; el format de representació de les quantitats en l'equació; o les operacions permeses en la construcció de l'equació.

Per una banda trobem estudis on s'apunta com a principal causa la comparació estàtica. Rosnick (1981) va plantejar la hipòtesi que la comparació estàtica podria ser conseqüència d'una interpretació incorrecta de les lletres en les quals s'utilitzarien com a unitats de mesura (E seria estudiants en lloc de nombre d'estudiants). Partint d'aquesta idea es van plantejar situacions experimentals en les quals es van emprar lletres que no conduïren a aquesta possible confusió on va concloure que tan sols el 60% dels estudiants identificaven P amb nombre de professors, mentre que la resta no ho feia. A més, Cooper (1986) va trobar que quan apareixia el signe multiplicatiu en l'equació, es produïa una disminució de l'EI. No obstant això, en aquest mateix estudi es va concloure que, l'ús de variables diferents a la lletra inicial de la

quantitat representada no va produir variació en la incidència de l'error (per exemple, canviar E (d'estudiants) i P (de professors) per x i y). A més, per tal d'intentar posar de manifest la idea que la variable representava al número d'estudi, Fisher (1988) en el seu estudi va plantejar que s'utilitzaren les lletres del tipus Ne (fent referència a nombre estudiants) en lloc de E (estudiants). En aquest cas tampoc es va trobar cap relació entre l'ús d'una notació i la major o menor incidència de l'error d'inversió. Seguint la mateixa línia, en Soneira et al. (2018), tampoc es van obtenir diferències en l'aparició de l'EI, on mitjançant l'ús d'una aplicació informàtica, les equacions es construïen utilitzant en lloc de lletres les proposicions escrites *"estudiants"* i *"nombre d'estudiants"*.

D'altra banda trobem els estudis on assenyalen que la coincidència en l'ordre de les paraules és la causant de l'EI. En l'estudi de Fisher et al. (2011), es va veure que quan es demanava als estudiants que escrigueren l'equació utilitzant l'operació inversa (en el cas d'escriure-ho com una multiplicació en l'exemple d'estudiants-professors, fer-ho com una divisió), l'EI es veia reduït. L'explicació que li donen els autors és que quan els indiquem que han de fer ús de la operació inversa, els estudiants ja no fan una lectura lineal de l'enunciat. Aquestos resultats segueixen la línia del que varen observar González-Calero et al. (2015). La investigació més recent sobre l'atribució de l'EI a la coincidència en l'ordre de les paraules és la de González-Calero et al. (2020). En aquest estudi es fa una comparació del fet de fer l'EI en estudiants bilingües bascs-espanyol aprofitant que, en el basc, l'ordre en el que es presenten les quantitats, en una comparació multiplicativa, no produiria l'EI en el cas d'una traducció literal d'esquerra a dreta. Els resultats mostraren una disminució significativa de la tasca d'EI quan els enunciats es proporcionaven en basc, el que suggeriria que la diferència en la incidència de l'EI seria l'efecte d'aplicar la traducció lineal, és a dir, l'ordre en que es presenten les quantitats.

A més, també s'ha vist que l'error es produeix tot i que l'estudiantat reconeix la informació de l'enunciat, és a dir, inclús tenint pistes contextuals. Wollman (1983) va proposar enunciats estructuralment anàlegs al d'*Estudiants i Professors* a l'alumnat universitari, en els quals no existien aspectes contextuals que pogueren influir en la construcció de l'equació, ja que normalment el nombre d'alumnes és major que el de professors. Així, va recórrer a enunciats del tipus: "En una classe hi ha sis vegades més xiquets que xiquetes", on de bestreta sembla impossible preveure si el nombre de xics serà superior al de xiques. L'autor va trobar que l'èxit en la tasca o l'aparició de l'EI no es veia influït pel coneixement contextual durant la traducció.

Encara que al llarg d'aquesta tesi no s'arribarà a determinar quin dels dos models és el que va lligat amb l'EI, sí que estudiarem la perspectiva cognitiva per tal de poder conèixer les diferències cerebrals entre subjectes que cometen o no l'EI, així com les àrees cerebrals associades a aquest error.

Com ja s'ha vist, a la investigació inicial de Clement (1982) li seguiren nombrosos estudis conductuals sobre l'EI (Cohen i Kanim, 2005; Cooper, 1986; Christianson et al., 2012; Don, 2011; Fisher, 1988; González-Calero et al., 2020; González-Calero et al., 2015; Kim et al., 2014; López-Real, 1995; Mangulabnan, 2013; Soneira et al., 2018; Wollman, 1983). En aquestos estudis no es va tindre en compte la importància que té el desenvolupament cerebral de l'alumnat en l'aprenentatge, fet que seria interessant per poder estudiar els processos cerebrals que es produeixen durant la realització de problemes verbals que indueixen a l'error i, així, intentar trobar la relació que hi ha entre cervell i aprenentatge. La ciència que té com a objectiu conèixer el funcionament del cervell i poder aplegar a comprendre com es comporta l'ésser humà davant diferents situacions és la neurociència, la qual unida als principis de l'educació forma el terme de neuroeducació (Portellano, 2018). Tal i com afirma Goswami (2006), la neuroeducació, mitjançant el funcionament del cervell, tracta de millorar l'aprenentatge de l'alumnat. És per això que tota la comunitat educativa hauria de saber des de com aprèn el cervell fins a la fragilitat que té davant determinats estímuls, sent un requisit indispensable per a la innovació pedagògica en l'àmbit educatiu (Campos, 2010).

Per tant, per poder tindre una demostració basada en els processos cerebrals que es produeixen durant la realització de problemes verbals, i poder veure la relació que hi ha entre cervell i aprenentatge i poder trobar una solució, en la nostra investigació es farà ús de la ressonància magnètica. Gràcies a les imatges de ressonància magnètica (IRM), es poden estudiar els processos mentals més complexos com ara és la resolució de problemes (Hanakawa et al., 2003).

Els estudis sobre el desenvolupament han incrementat el nostre coneixement sobre la maduració del cervell humà (Blakemore, 2012; Giedd et al., 1999; Sowell i Jernigan, 1998; Sowell et al., 2003; Sowell et al., 1999; Wierenga et al., 2014). En particular, l'estudi de Giedd et al. (1999) amb IRM revela que el volum de substància grisa (SG) es caracteritza, durant el desenvolupament, per la forma d'una U invertida que aconsegueix el seu punt màxim a diferents edats en diferents regions cerebrals (veure figura 1 de Giedd et al. (1999)). Açò suggereix una trajectòria heterogènia i no lineal, on les competències maduren a diferents velocitats i temps depenent de les regions del cervell que són més importants per a una habilitat donada. Giedd et al. (1999) han mostrat que durant el període de l'adolescència (12 anys aproximadament), el volum de substància blanca (SB) continua augmentant (fins i tot hi ha algunes àrees locals que canvien amb rapidesa) i el volum de SG en les àrees parietals i frontals comença a disminuir a causa de la maduració cerebral. En la mateixa línia l'estudi de Sowell et al. (1999) mostra que hi ha reduccions del volum de SG cerebral entre la infància i l'adolescència, amb els canvis més significants al còrtex dorsal del lòbuls frontal i del parietal.

A més, gràcies a estudis com el de Blakemore (2012), on es fa una revisió de diversos estudis de IRM durant el desenvolupament, s'ha obtingut una major informació sobre la maduració del cervell. Blakemore (2012), després de fer una extensa revisió, afirma que el desenvolupament del cervell no acaba a una edat primerenca, sinó que aquesta s'allarga a l'adolescència aconseguint el seu volum màxim al voltant dels vint-i-cinc anys (Caviness et al., 1996; Sowell i

Jernigan, 1998), produint-se canvis tant en la SG com en la SB, sent la SB la que presenta un major desenvolupament durant aquesta etapa (Ortiz, 2009). Així mateix, l'estudi de Wierenga et al. (2014) revela que el volum de SG de l'estriat (dorsal: el putamen, caudat; i ventral: nucli accumbens) disminueix linealment amb l'edat en el rang de 7 a 23 anys, mentre que el volum d'hipocamp, amígdala, *pallidum* i cerebel mostra una trajectòria de desenvolupament invertida en forma d'U. En concret, el volum del putamen disminueix a partir dels 7 anys, mostrant-se en procés de maduració durant l'adolescència (veure figura 2 de Wierenga et al. (2014)). El desenvolupament de les estructures de l'estriat indica que hi ha canvis dinàmics durant el desenvolupament, estant relacionat amb els canvis en el desenvolupament cognitiu i del comportament. A més, Gogtay et al. (2004) reporta una seqüència de desenvolupament de la SG cerebral, és a dir, de maduració cerebral. Aquest estudi es va realitzar per a 13 subjectes amb edats entre 4 i 21 anys. A més les IRM es van obtindre cada 2 anys durant un període de 8 i 10 anys. Els resultats mostraren que en termes de densitat de SG, les regions sensori-motores són les que maduren més prompte, seguides per la resta del còrtex, que madura (en termes de pèrdua de matèria grisa) en la direcció de regions posteriors a anteriors, és a dir, comença en àrees parietals cap a les àrees frontals, finalitzant en el còrtex temporal. Cal destacar que el còrtex prefrontal-dorsolateral finalitza la seua maduració a la fi de l'adolescència.

Els estudis sobre el desenvolupament del cervell humà han augmentat els nostres coneixements sobre la seua maduració i sobre com aquesta maduració pot dependre o ser important en el procés d'aprenentatge. En referència a l'ensenyança-aprenentatge de les matemàtiques, Dehaene (1997) suggereix que la comprensió intuïtiva de les quantitats s'associa amb l'activitat del Sulcus Intraparietal (IPS). A més, el còrtex parietal participa en diverses tasques matemàtiques des de la comparació numèrica fins a processos més complexos com ara proporcions o raonaments deductius (Kroger et al., 2008; Vecchiato et al., 2013). De Smedt et al. (2011) també mostren que diferents parts del còrtex parietal, com ara l'IPS bilateral i el gir angular esquerre, tenen un paper crucial en el càlcul mental. A més, per a molts investigadors, l'aprenentatge matemàtic implica en gran mesura la memòria de treball (Baddeley, 1997), que s'associa amb zones frontals. La memòria de treball té la capacitat d'emmagatzemar, manipular la informació, permetent així l'execució de tasques cognitives com ara el raonament, la comprensió i la resolució de problemes, recolzades pel manteniment i la disponibilitat d'aquesta informació (Baddeley, 1997).

En referència a les regions cerebrals, com per exemple les àrees prefrontals, semblen madurar relativament més tard i es creu que estan involucrades en la cognició matemàtica i altres processos d'ordre superior que es desenvolupen al llarg de la infància i l'adolescència (Blakemore, 2012). Açò podria ser un punt fort en la transició des de l'aritmètica concreta fins al llenguatge simbòlic de l'àlgebra, on els estudiants han de desenvolupar habilitats de raonament abstracte que els permeten generalitzar, modelitzar i expressar informació en el llenguatge de l'àlgebra (Anderson et al., 2008; Lee et al., 2007; Qin et al., 2004).

Després de realitzar una cerca sobre resolució de problemes, coincidim amb Radford i André

(2009) en què són poques les investigacions que se centren en el cervell i el pensament matemàtic avançat, en concret amb la relació cervell-àlgebra durant el desenvolupament. Entre aquestes caldria destacar l'estudi longitudinal de Qin et al. (2004), qui va observar que, després de la pràctica en resolució de problemes verbals, els patrons que seguien, tant els adolescents com els adults, eren similars, produint-se una reducció del còrtex prefrontal. En canvi, es va observar, només en adolescents, una reducció en l'activació de l'àrea parietal i un increment en el putamen esquerre. Aquests resultats suggereixen que la resposta cerebral en adolescents és més plàstica en aquesta etapa neural del desenvolupament i, per tant, sofreix més canvis deguts a la pràctica i els efectes d'aprenentatge. En conclusió, aquesta major resposta cerebral en les i els adolescents, a causa de la pràctica, sembla indicar que aquest període seria el més apropiat per a l'aprenentatge de l'àlgebra (Qin et al., 2004).

Un altre estudi sobre la resolució de problemes va ser el de Lee et al. (2007), on va trobar que durant la transformació de l'enunciat del problema verbal a l'equació es mostrava activació en àrees cerebrals del còrtex prefrontal i parietal associades amb la memòria de treball i processos atencionals.

A més, investigadors com Anderson et al. (2012) en un dels seus estudis sobre resolució de problemes algebraics, van observar activacions en diferents àrees cerebrals. En primer lloc destacava el còrtex prefrontal, la qual està associada a la recuperació dels procediments algebraics que ja s'havien après anteriorment, a més, el parietal posterior, el qual està relacionat amb la transformació del llenguatge de les equacions algebraiques, i finalment el gir fusiforme el qual està associat amb les regions visuals, aquestes són de gran ajuda per a poder analitzar les equacions i decodificar la informació d'aquestes.

Destaquem la importància d'introduir l'àlgebra en l'edat adolescent, ja que després d'aquesta revisió i basant-nos en els estudis sobre resolució de problemes on es veu que les àrees prefrontals i parietals són les involucrades en aquestos processos cognitius, i veient que segons Giedd et al. (1999) aquestes àrees comencen a madurar als 12 anys. Per tant, és en aquest període quan pareix que les estructures relacionades amb l'aprenentatge estiguen en procés de maduració, existint una major plasticitat cerebral que ajude a l'aprenentatge, tal i com indicava (Qin et al., 2004).

Arribats a aquest punt, queda constància del poc contingut bibliogràfic al voltant dels processos cerebrals lligats al fenomen de l'EI. No es sap si es tracta d'un error persistent o si és una conseqüència de l'efecte dels processos d'ensenyament-aprenentatge durant la maduració en l'adolescència, per tant, el nostre objectiu és intentar proporcionar solucions en els àmbits tant educatius com socials fent una primera aproximació del coneixement dels processos cerebrals que duen a cometre l'EI.

1.2 Organització de la tesi

La distribució d'aquesta tesi és la següent. En aquest primer capítol es presenta la introducció del tema principal, seguit per un segon capítol on es mostren els objectius generals. A continuació, es troba el capítol tercer on es mostra mostra la part teòrica així com les bases de la metodologia utilitzada per als diversos estudis. Al quart capítol es mostren els objectius específics de cadascuna de les aportacions, a més es mostren els quatre estudis que componen la tesi. Al cinquè capítol, es veuran les conclusions i per últim, al capítol sisè el treball futur.

Capítol 2

Objectius de la investigació

2.1 Objectius generals

Els objectius generals d'aquesta tesi són:

- Trobar quines estructures cerebrals estan associades a resolutors competents quan fan una tasca de resolució de problemes amb EI. És a dir, observar les diferències en volum de SG entre el grup de participants que realitzen EI i els que no (no-EI).
- Discriminar entre els subjectes amb EI i no-EI, emprant l'anatomia de la pròpia estructura o estructures com a variables mitjançant l'estudi de forma.
- Proposar nous mètodes estadístics de classificació per dades funcionals multivariants i multi arguments, i comparar-los amb altres alternatives.
- Proposar formes de visualitzar diferències entre grups en problemes de classificació per dades funcionals multivariants i multi arguments.
- Discriminar entre tres grups, els subjectes que no cometen cap error, els que fallen més del 40% de les respostes i els que es troben en mig d'aquestos dos grups. En aquest cas també emprarem les formes de les estructures cerebrals com a variables.
- Proposar nous mètodes estadístics de classificació ordinal per dades funcionals multivariants i multiarguments, i comparar-los amb altres alternatives.
- Determinar les bases neuronals subjacents lligades a l'EI mitjançant la realització d'una tasca dins de la RM.

- Estudiar quines àrees cerebrals determinen una millor competència matemàtica en la resolució de problemes en el grup no-EI, i per tant, un millor aprenentatge algebraic.
- Trobar el mètode òptim per a fer una classificació a partir de les bases neuronals obtingudes a través de la tasca.

Capítol 3

Base teòrica i metodologia utilitzada

3.1 Ressonància Magnètica

Mitjançant la RM, es poden realitzar tant estudis de l'estructura cerebral (RM estructural, RMe), com observar els canvis fisiològics a curt termini, els quals estan relacionats amb el funcionament cerebral (RM funcional, RMf). Aquestos canvis fisiològics tenen relació amb els canvis funcionals d'una regió cerebral. Els canvis fisiològics són la variació de quantitat de des-oxihemoglobina en la sang, mentre que els canvis funcionals fan referència a la intensitat i variació dels canvis d'oxigenació de la sang en una determinada regió cerebral mentre es realitza una tasca.

Per poder comprendre de forma visual les diferències que hi ha entre la RMe i la RMf, es pot visualitzar la figura 3.1.

3.1.1 Ressonància Magnètica Estructural

La RMe tracta de realitzar imatges de teixit biològic mitjançant la utilització de camps magnètics. Aquesta part de l'estudi es va realitzar mitjançant mètodes morfomètrics. Un dels fonamentals és el conegut com a enfocament local, el qual inclou la morfometria basada en el vòxel (VBM) (Ashburner i Friston, 2000). El que es tracta de fer és determinar en cada vòxel [1] la concentració de volum de SG i/o SB mitjançant el processament de segmentació i

[1]Els vòxels, són les unitats en les que s'estableix la localització espacial en una imatge de RMe dels teixits cerebrals. Un vòxel té les tres dimensions (x,y,z) ja que es tracta d'una unitat de volum. Cal destacar que quan més menuts siguen, major serà la possibilitat de delimitar estructures més xicotetes en el cervell. Però té un

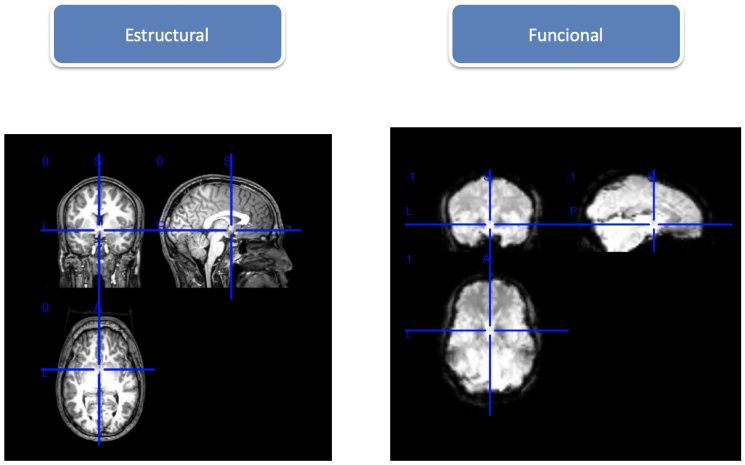


Figura 3.1: Estructural vs Funcional.

normalització de les imatges.

Els passos utilitzats per VBM amb imatges de RMe foren (figura 3.2):

- **Normalització espacial:**

  La normalització es realitza per tal de transformar les imatges de RMe d'un subjecte a un espai estereoestàtic, el qual està definit per un *template*. Amb açò el que volem aconseguir és poder comparar directament els subjectes. El que es fa és establir una relació vòxel a vòxel mitjançant l'aproximació amb *SPM*[2], la qual es realitza mitjançant un registre de mínims quadrats.

  Per a realitzar la normalització, el primer que es farà serà determinar els 12 paràmetres per a una transformació afí òptima. Amb açò aconseguirem que la imatge del subjecte, al qual li aplicarem la normalització, quede ajustada al *template*, Montreal Neurological Institut (MNI)[3] tant en grandària, com en la forma, com en la posició.

  A continuació, *SPM* utilitza la transformada discreta del cosinus per tal de construir les imatges normalitzades, i per últim la utilització de l'estadística bayesiana, per tal d'obtindre un ajust millor. A la figura 3.3, es pot veure una explicació gràfica.

inconvenient, ja que el soroll (*noise*) apareixerà amb la utilització d'imatges amb una alta resolució.

[2]Statistical Parametric Mapping és el paquet de programari que s'ha utilitzat per a l'anàlisi d'imatges de RM, amb ell s'ha dut a terme tant el preprocesament d'imatges cerebrals com la posterior anàlisi estadística de les mateixes.

[3]és un atles creat a partir de 152 subjectes sans. Per obtindre'l, se li va aplicar un corregistre lineal a l'atles anterior (MNI305). L'objectiu era adaptar-lo per a definir les coordenades d'anatomia estàndar per el *Consortium of Brain Mapping (ICBM)*. Aquest atles és utilitzat per SPM des de la versió 99 fins l'actualitat.

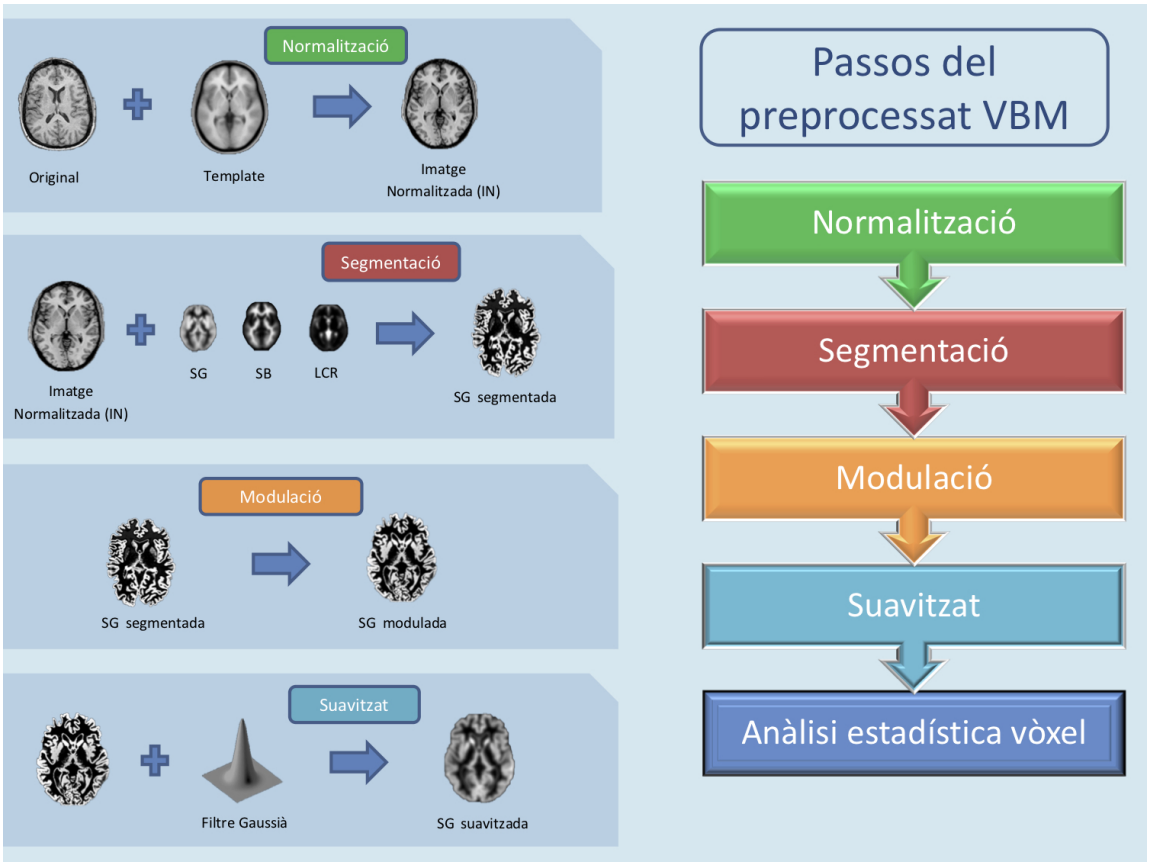


Figura 3.2: Passos del preprocessat amb VBM.

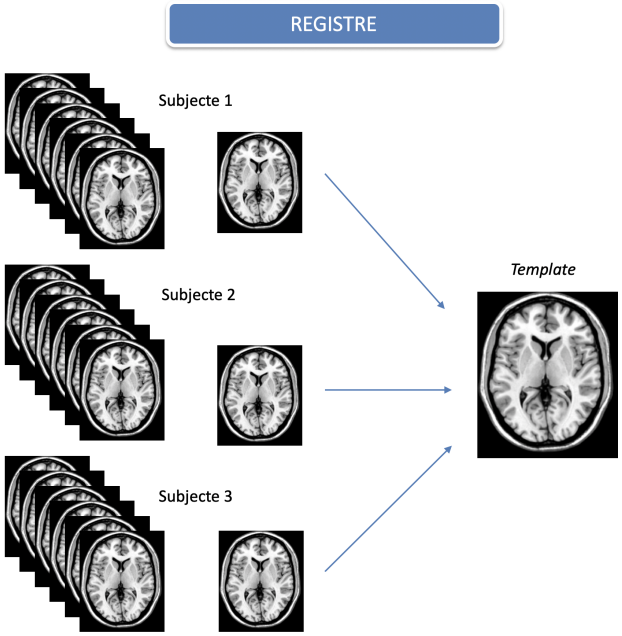


Figura 3.3: Registre de les imatges de tres participants al mateix sistema de coordenades *(template)*.

- **Segmentació:**

  En el pas de la segmentació, aconseguim separar les tres substàncies: la SG, la SB i el líquid cefaloraquidi (LCR) (veure la figura 3.4).

- **Modulació:** Les intensitats de vòxel es modulen mitjançant el determinant jacobià del camp de deformació (warps). Este pas serveix per a corregir els canvis en el volum del cervell, els quals es deuen a la normalització espacial no lineal.

  Es fa una multiplicació de la SG normalitzada pel seu volum relatiu abans i després d'aplicar el warping. On quedaria: $i_B = i_A x[V_A/V_B]$ on

  - V_A és el volum abans de la normalització.
  - V_B és el volum de la plantilla.
  - i_A és la intensitat de la senyal abans de la normalització.
  - i_B és la intensitat de la senyal després de la normalització.

- **Suavitzat espacial:**

  El que volem aconseguir amb el suavitzat és millorar la relació senyal/soroll, fent un increment de la sensibilitat. Els efectes del suavitzat el que fan es difuminar el senyal entre els vòxels veïns, permetent una millor superposició espacial i una millora de la qualitat de les dades per a la posterior anàlisi estadística.

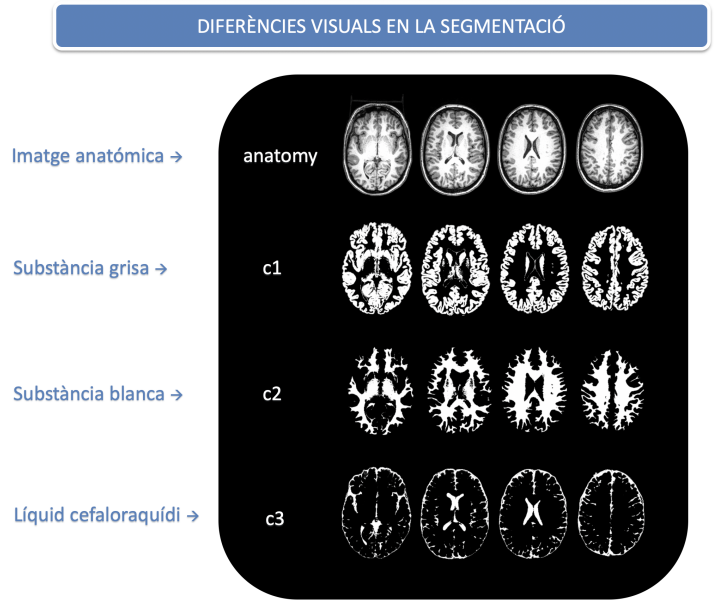


Figura 3.4: Resultats de l'aplicació de la segmentació a les dades de neuroimatge estructural: separació en SG, SB i LCR.

SPM fa ús del filtre Gaussià, mitjançant la mida 'full width half maxium' (FWHM), la qual és l'amplària del nucli (es pot veure a la figura 3.5). Cal destacar que es fan ús de diferents filtres depenent de les anàlisis que es vagen a realitzar, ja que el suavitzat pot causar que algunes regions que funcionalment són diferents es fusionen, modificant les correlacions espacials entre els vòxels.

El VBM compon la normalització espacial de totes les imatges a un mateix espai, d'on s'extrau la SG de les imatges normalitzades i finalment suavitzades. El següent pas és fer l'anàlisi estadística amb les imatges prepocessades, fent ús del Model Lineal General (MLG).

3.1.2 Ressonància Magnètica Funcional

Una mesura indirecta de la resposta neuronal, la qual està presa mitjançant la RMf, seria el senyal que prové del canvi cognitiu relacionat amb l'oxigenació de la sang, on l'activació neuronal produeix un increment del consum d'oxigen en aquelles neurones que estan actives durant la realització d'una tasca cognitiva.

El principi bàsic del mètode subtractiu de Donders (1969) aplicat a RM ens porta a la necessitat d'establir una línia base, la qual coneixerem com a condició de control dins de la

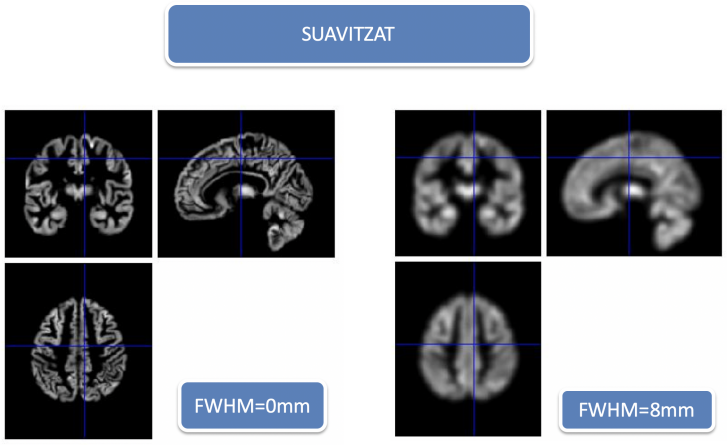


Figura 3.5: Resultats de l'aplicació del suavitzat a les dades de neuroimatge estructural.

RMf. A més de la condició de control, es trobarà l'experimental on es permet estudiar el procés cognitiu durant la realització d'una tasca. Per a que la tasca de control realitze la seua funció de forma correcta, es deuen incloure tots els processos cognitius menys els que a nosaltres ens interessen ja que com el seu nom indica els sustrau l'activació cerebral associada a la tasca de control de l'activació de la tasca experimental.

D'altra banda per als dissenys experimentals de RMf, existeixen principalment dues aproximacions les quals depenen de la forma en la que es presenten els estímuls.

El primer és el **disseny per blocs**. Han sigut els més utilitzats tradicionalment en els experiments, a causa de les limitacions en la resolució temporal de la RMf. Es basen en blocs de control alternats en blocs de tasca, obtenint la suficient quantitat de dades per a realitzar una anàlisi estadística factible.

El segon és el **disseny d'esdeveniments**. Els estímuls es presenten de manera independent, separats per un interval interestimular (ISI), el qual pot ser aleatori o pseudoaleatori dins de l'experiment. Amb la qual cosa, l'estímul es presentarà solament una vegada en un període curt de temps, seguit pel mateix o un estímul diferent.

Basant-se amb els avantatges d'aquestos dos, a dia de hui s'han començat a utilitzar el **dissenys mixtes**. Aquestos dissenys combinen els dos tipus anomenats anteriorment, els blocs de la tasca es van alternant amb els blocs de control. A més, dins de cada bloc de tasca, els estímuls són assignats de forma aleatòria.

A la figura 3.6 es pot veure la presentació dels estímuls.

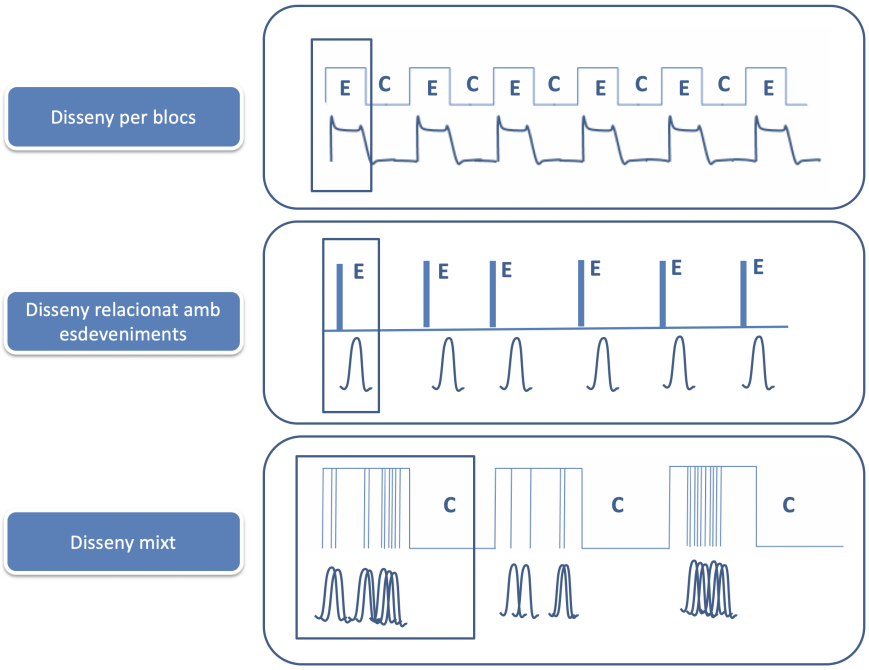


Figura 3.6: Presentació d'estímuls en paradigmes de RMf.

Preprocessat

Caldria destacar que una de les característiques de les imatges de RMf és el soroll. Aquest pot estar determinat per diferents causes.

a) Soroll tèrmic: el qual es manifesta quan es produeix un increment en la temperatura del sistema. L'avantatge és que no resulta important per a l'anàlisi de dades ja que es pot reduir mitjançant la mitjana de les dades.

b) Soroll del sistema: el podem trobar a causa de les fluctuacions en el funcionament del maquinari de RM.

c) Soroll relacionat amb el subjecte, la tasca o tots dos: ocorre quan el subjecte es mou dins de l'escàner, produint que l'activació del vòxel siga contaminada per l'activació dels vòxels veïns (Hajnal et al., 1994).

El preprocessat d'imatges consisteix en la neteja del soroll (encara que no tots els tipus de soroll poden ser eliminats) de les dades anteriors a l'anàlisi estadística. Són diversos els programes que es poden trobar per a realitzar-ho, però en el que ens centrarem per a este

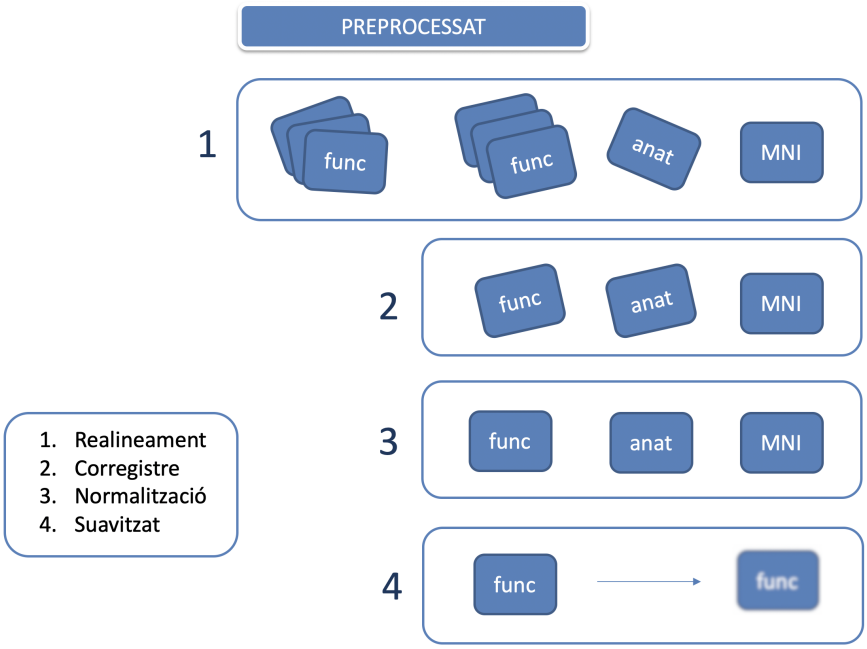


Figura 3.7: Explicació gràfica del preprocessat.

projecte serà el SPM12 (Wellcome Trust Center for Neuroimaging, Londres, UK). A continuació, es podran veure les fases que el componen. A la figura 3.7, trobarem una imatge descriptiva del mateix.

- **Reorientació:** Este procediment consisteix en reorientar totes les imatges al pla de la Comissura Anterior-Comissura Posterior, aconseguint ubicar l'origen en el tall central de la comissura anterior.
- **Realineament. Correcció del moviment:** El realineament és un pas del processat previ, el qual s'aplica perquè es disposa de diverses imatges per a un mateix subjecte (Gispert et al., 2003). Mitjançant l'aproximació de mínims quadrats i una transformació espacial de 3 paràmetres de translació i 3 de rotació, obtindrem la realiniació de les imatges d'un mateix subjecte. Cal destacar que la primera imatge és la que s'utilitza com a *mean image* per a realinear les altres.
- **Corregistre:** Amb el corregistre el que volem aconseguir és re-alinear la imatge anatòmica a les funcionals.
- **Altres:** Tant la **segmentació**, la **normalització espacial** com el **suavitzat**, són els mateixos passos amb les mateixes característiques definides per a les imatges de RMe explicades en el punt 3.1.1.

3.1.3 Anàlisi estadística de IRM

En SPM, l'anàlisi de dades de diversos subjectes sol dividir-se en dues etapes utilitzant models a dos "nivells", ambdós estan basats en el Model Lineal General (MLG) (Winer et al., 1991; Yandell, 2017). Els models de *primer nivell* (també conegut com anàlisi d'efectes fixes) s'utilitzen per implementar una anàlisi intra-subjecte. Normalment hi haurà tants models de primer nivell com subjectes hi hagen. A més, també està el model de *segon nivell*, per tal de fer inferències sobre el conjunt dels subjectes.

En el cas de les imatges de RMf l'anàlisi de dades dels subjectes es divideix en dos passos. Primerament, es fa una l'anàlisi intra-subjecte o efectes fixos, on s'obté la imatge de contrast corresponent a la condició experimental a estudiar. A continuació, aquesta imatge es introduïda en un model d'anàlisis d'efectes aleatoris aconseguint els resultats experimentals entre-subjectes.

Per altra banda, l'anàlisi de dades amb les imatges de RMe es realitza mitjançant l'anàlisi d'efectes aleatoris utilitzant les imatges obtingudes del prepocessat.

Model Lineal General

- **Especificació del model:**

  És el pas fonamental per a l'anàlisi de les imatges, el qual en SPM es realitza vòxel a vòxel. El MLG, explica la variable Y_j en termes d'una combinació lineal de les variables independents més un terme d'error ϵ_j, on $\epsilon_j \sim \mathrm{N}(0,\sigma^2)$ iid.

  $$Y_j = x_{j1}\beta_1 + ... + x_{jl}\beta_l + ... + x_{jL}\beta_L + \epsilon_j. \tag{3.1}$$

  Els β_l són paràmetres desconeguts els quals corresponen a cadascuna de les L variables independents x_{jl}. Per tant, podem dir que és equivalent a la següent forma matricial:

  $$Y = X\beta + \epsilon, \tag{3.2}$$

  on Y és el vector columna d'observacions, X la matriu de disseny de dimensions $J \times L$, on cal destacar que cada fila de la matriu correspon a una observació i cada columna a un paràmetre del model. A més, β és el vector columna de paràmetres, i per últim ϵ és el vector columna del terme d'error.

- **Estimació dels paràmetres:**

  El paràmetres del model seran estimats mitjançant el mètode dels mínims quadrats ordinaris. El conjunt de paràmetres a estimar serà $\tilde{\beta} = [\tilde{\beta}_1, ..., \tilde{\beta}_L]^T$, els quals produiran els valors d'ajust $\tilde{Y} = [\tilde{Y}_1, ..., \tilde{Y}_J]^T = X\tilde{\beta}$, on els errors residuals seran $e = [e_1, ..., e_J]^T =$

$Y - \tilde{Y} = Y - X\tilde{\beta}$. La suma dels quadrats dels residus, és la suma dels quadrats de la diferència entre els valors reals i els ajustats. Este està definit per:

$$S = \sum_{j=1}^{J}(Y_j - x_{j1}\tilde{\beta}_1 - ... - x_{jL}\tilde{\beta}_L)^2. \tag{3.3}$$

Si volem minimitzar els residus, aleshores consideraríem:

$$\frac{\partial S}{\partial \tilde{\beta}_l} = 2\sum_{j=1}^{J}(-x_{jl})(Y_j - x_{j1}\tilde{\beta}_1 - ... - x_{jL}\tilde{\beta}_L) = 0. \tag{3.4}$$

Cal destacar que (2.4), és la l-ésima fila de $X^TY = (X^TX)\tilde{\beta}$, i per tant, $X^TY = (X^TX)\hat{\beta}$. Al MLG, l'estimador de mínims quadrats és el mateix que el de màxima versemblança, a més, és el millor dels no esbiaixats tal i com diu el Teorema Gauss-Markov. Si (X^TX) és invertible, aleshores l'estimació per mínims quadrats serà: $\hat{\beta} = (X^TX)^{-1}X^TY$.

Al *SPM*, s'adopta l'aproximació mitjançant el mètode de la pseudoinversa per als estimadors dels paràmetres. Siga $(X^TX)^-$ la matriu inversa de Moore-Penrose de (X^TX) o matriu pseudoinversa, podem utilitzar $(X^TX)^-$ en lloc de $(X^TX)^{-1}$ en l'equació descrita anteriorment, quedant de la següent forma:

$$\hat{\beta} = (X^TX)^-X^TY = X^-Y. \tag{3.5}$$

- **Inferència:**

  Els estimadors dels paràmetres són normalment distribuits: si X es de rang complet, aleshores $\hat{\beta} \sim N(\beta, \sigma^2(X^TX)^{-1})$. Si λ és un vector columna de L pesos tindrem:

  $\lambda^T\hat{\beta} \sim N(\lambda^T\beta, \sigma^2\lambda^T(X^TX)^{-1}\lambda)$.

  Com que $\hat{\beta}$ y $\hat{\sigma}^2$ són independents, podem avaluar els paràmetres del model $\lambda^T\beta$ de la següent forma:

$$\frac{\lambda^T\hat{\beta} - \lambda^T\beta}{\sqrt{\sigma^2\lambda^T(X^TX)^{-1}\lambda}} \sim t_{J-p}, \tag{3.6}$$

  on cal destacar que t_{J-p} és la distribució t de Student amb J - p graus de llibertat, on $p = rang(X)$.

  És molt important la inferència dels contrastos. Direm que λ és un contrast estimable si i solament si el vector contrast es pot escriure com combinació lineal de les files de X. Amb els contrastos, avaluarem els efectes d'interès i realitzarem la avaluació estadística de la hipòtesi que nosaltres formulem.

Contrast T

El contrast estadístic t, és una mesura senyal-soroll i es calcula mitjançant (Friston et al., 1994):

$$t_{df} = \frac{\lambda\hat{\beta}}{SD(\lambda\hat{\beta})}, \tag{3.7}$$

on *df* representa els graus de llibertat i $SD(\lambda\hat{\beta})$ ens diu la desviació estàndard de $\lambda\hat{\beta}$.

El contrast d'hipòtesi consisteix en $H_0 : \lambda^T\beta = 0; H_1 : \lambda^T\beta > 0$ o $\lambda^T\beta < 0$.

Un dels contrastos més utilitzats en *SPM* és el conegut com lògica *sustractiva*, el qual és una combinació lineal univariada dels estimadors dels paràmetres i és un contrast direccional. Per exemple, en una tasca de RMf on tenim dos condicions (una de control i altra experimental), la nostra matriu de disseny seria $X = [x_1, x_2]$, sent x_1 la condició d'activació i x_2 la de control, i el que volem és veure en quina regió cerebral es produeix un increment lineal en la part de l'activació, utilitzarem el vector de contrast $\lambda^T = [1\ 0]$, metre que si volguérem veure el decreixement fariem ús de $\lambda^T = [\ \text{-}1\ 0]$, i si el que volem és veure quina regió cerebral s'activa més en la condició d'activació que en la de control, utilitzarem el vector de contrast $\lambda^T = [1\ \text{-}1]$.

3.2 Anatomia i segmentació del putamen

Per obtenir les ROIs del putamen esquerre i dret per a cada participant, es va segmentar cada putamen utilitzant la *toolbox imcalc* d'*SPM12*, on es va realitzar una intersecció entre la imatge de SG de cada participant i el ROI dels putamen de l'atles AAL (Tzourio-Mazoyer et al., 2002). Finalment, els talls de cada putamen es van ajuntar utilitzant la funció *isosurface* en MATLAB, la qual retorna les cares i vèrtexs del *triangle mesh*.

3.3 Classificadors

Al llarg d'esta secció, es farà un repàs dels classificadors (Brownlee, 2016; Maroco et al., 2011) utilitzats per a la part de RMf de la tesi. Al llarg del treball, considerarem que les dades es troben a una matriu $\mathbf{X}$ de dimensió $n \times p$, on n és el nombre de individus, la grandària mostral i p el nombre de variables. A més, les dades estan en k possibles grups.

3.3.1 Anàlisi discriminant

L'**anàlisi discriminant lineal** de Fisher (ADL) construeix funcions discriminants $j = min(k-1, p)$ que estimen puntuacions discriminants (D_{ji}) per a cada un dels $i = 1, ..., n$ subjectes classificats en k grups, com $D_{j_i} = w_{i1}X_{1i} + w_{i2}X_{2i} + ... + w_{ip}X_{pi}$ on $i = 1, ..., n$ i $j = 1, ..., min(k-1, p)$.

Els pesos discriminants w_{ij} es calculen per mínims quadrats ordinaris, de manera que la proporció de la variància dins dels grups entre la variància entre els grups siga mínima.

Es poden construir funcions de classificació del tipus $C_{ji} = c_{j0} + c_{j1}X_{1i} + c_{j2}X_{2i} + ... + c_{jp}X_{pi}$ per a cadascun dels grups $j = 1, ..., k$ a partir de les puntuacions discriminants. Els coeficients de la funció de classificació per al grup *j-èsim* són estimats a partir de les matrius de suma de quadrats *"within"* W de les puntuacions discriminants per a cada grup i el vector de mitjanes dels predictors en cada grup M, com $C_j = W^{-1}M$ amb $c_{j0} = logp - \frac{1}{2}C_jM_j$.

L'**anàlisi discriminant quadràtica** (ADQ) utilitza una funció discriminatòria quadràtica de la forma:

$$D_i = \sum_{h=1}^{p} w_{ih}X_h + \sum_{h=1}^{p} q_{ih}X_h^2 + \sum_{h=1}^{p-1} r_{ih}X_hX_{h+1},$$

on $i = 1, ..., min(k-1, p)$.

Amb les funcions classificadores:

$$c_j = c_{0j} \sum_{h=1}^{p} c_{ih}X_h + \sum_{h=1}^{p} o_{ih}X_h^2 + \sum_{h=1}^{p-1} m_{ih}X_hX_{h+1},$$

on $j = 1, ..., k$.

Tant a ADL com a ADQ, el subjecte es classifica al grup pel qual la puntuació de la funció de classificació és més alta.

En la implementació d'ADL i d'ADQ en la part de resultats s'ha utilitzat la funció ADL i ADQ, respectivament, del paquet de R **MASS** (Venables i Ripley, 2002).

3.3.2 Generalització d'anàlisi discriminant

Algunes generalitzacions de l'anàlisi discriminant són les següents. Sols descriurem breument la idea de cadascun d'ells i el paquet de R (R Development Core Team, 2020) on està la seua implementació, per no allargar excessivament el treball. Els detalls teòrics es poden trobar a cadascun dels articles on es van proposar.

Abans de començar a veure-les, una primera modificació d'ADL, seria seleccionar les variables. Especialment pot ser convenient si hi ha moltes variables i pocs individus, com és el cas del nostre problema. Hi ha diferents mètodes per seleccionar variables. Considerarem les funcions del paquet **klaR** (Weihs et al., 2005).

ADF L'Anàlisi Discriminant Flexible proposat per Hastie et al. (1994) és un model de classificació basat en una mixtura de models de regressió lineal, que utilitzen una puntuació òptima per transformar la variable resposta, de manera que les dades estiguin en una millor disposició per a la separació lineal, i MARS (Multiple Adaptive Regression Splines) per generar la superfície discriminant.

En la implementació s'ha usat la funció fda amb MARS del paquet **mda** (Hastie i Tibshirani, 2017).

ADR L'Anàlisi Discriminant Regularitzada va ser proposada per Friedman (1989). L'anàlisi discriminatòria regularitzada és una mena de terme mig entre ADL i ADQ. Recordeu, a ADL, suposem que hi ha una matriu de covariància comú per a totes les classes. ADQ assumeix diferents matrius de covariància per a totes les classes. L'anàlisi discriminatòria regularitzada és una solució intermitja entre ADL i ADQ.

La implementació s'ha fet amb la funció rda del paquet **klaR** (Weihs et al., 2005).

PADL La classificació amb l'anàlisi discriminant lineal de Fisher penalitzada va ser desenvolupada en Witten i Tibshirani (2011) i busca penalitzar els vectors discriminants en el problema discriminant de Fisher d'una manera que condueix a una major interpretabilitat.

La funció PenalizedADL del paquet de R **PenalizedADL** s'ha usat en la implementació (Witten, 2015).

ADP L'Anàlisi Discriminant Penalitzada va ser proposta en Hastie et al. (1995). De forma simple, es pot dir que la matriu de variància dins de les classes (within), Σ_W, és substituïda per una versió regularitzada, és a dir, per $\Sigma_W + \lambda\Omega$, on Ω és una matriu de penalització.

Este procediment es troba implementat en la funció fda amb gen.ridge del paquet **mda** (Hastie i Tibshirani, 2017).

3.3.3 Regressió logística

La **Regressió Logística binomial** (RL) modela la probabilitat d'ocurrència d'un èxit. Una combinació lineal de predictors s'utilitzen per ajustar una transformació logit de la probabilitat d'èxit per a cada subjecte (π_i) com:

$$\hat{\pi}_i = \frac{e^{\beta_0+\beta_1 X_{1i}+...+\beta_p X_{pi}}}{1+e^{\beta_0+\beta_1 X_{1i}+...+\beta_p X_{pi}}}.$$

Si la probabilitat estimada es major que 0.5, el subjecte és classificat dins del grup d'èxits.

Per a la implementació, s'ha fet ús de la funció glm i stepAIC (per fer selecció de variables), del paquet de R **MASS** (Venables i Ripley, 2002).

3.3.4 Xarxes neuronals

Els mètodes de les **Xarxes Neuronals** (XN) s'han utilitzat de forma exhaustiva en els problemes de classificació. Habitualment, trobem una sèrie de capes: la primera és la d'entrada (on es fiquen les variables predictores), la última és la d'eixida, on tenim la sortida, la predicció, en este cas el grup (y_k), i en mig es troben les capes ocultes, on es processen les entrades amb diferents funcions i pesos. De forma col · loquial, les XN funcionen com una caixa negra.

El model de XN és descrit per:

$$\hat{y}_k = f_k(x, w, o, x_0, o_{0_k}, \theta) = f(\sum_{j=1}^{h} o_{kj} \cdot g(\sum_{i=1}^{p} w_{ji} x_i + x_{0j}) + o_{0k}).$$

On:

- x és el vector de p-predictors.
- w és el vector de pesos d'entrada.
- o és el vector de pesos ocults per a la capa oculta .
- x_0 i o_{0k} són constants.

Les funcions g i f són funcions per a la capa oculta i d'eixida, respectivament. Les funcions d'activació pertanyen a una de les famílies generals de funcions lineal, logística, exponencial o gaussiana. Es poden utilitzar diverses topologies de XN en problemes de classificació binària. Dues de les XN més utilitzades són el *Multilayer Perceptron* (MLP) i la *Funció de Base Radial* (FBR). Les principals diferències entre estes dues XN resideixen en les funcions d'activació de la capa oculta.

- Per al MLP, la funció d'activació pertany, generalment, a una funció lineal de la forma

$$f_j(x) = \sum_{i=1}^{p} w_{ij} x_i,$$

  o a la família de funció d'activació logística on:

$$f(x) = \frac{1}{1 + exp(-x)}.$$

- Per a FBR, la funció d'activació pertany a la família Gaussiana:

$$f_j(x) = exp[-\frac{1}{2}(x - \mu_j)' \Sigma_j^{-1}(x - \mu_j)].$$

Normalment, el mètode de XN va ajustant els pesos durant una sèrie d'iteracions, utilitzant un subconjunt de les dades, el conjunt d'entrenament. El vector de pesos sinàptics (w) del mètode, s'actualitza en cada iteració per maximitzar la taxa de classificació correcta o minimitzar una funció dels errors de classificació.

Al final, una volta entrenat el mètode, este s'aplica amb els pesos obtinguts al conjunt de test, que mai ha de ser gastat en el període d'entrenament. El llibre de Hastie et al. (2009) conté una excel · lent explicació sobre com dividir el conjunt de dades en conjunts d'entrenament, validació i test.

Com que hi ha moltes possibles configuracions de xarxes, s'ha optat per gastar la funció ppr (Projection Pursuit), que consisteix en la suma de models lineals transformats de forma no lineal, de forma similar als models de XNs.

3.3.5 Màquines de Vectors Suport

Les **Màquines de Vectors Suport** (MVS) són mètodes de classificació que es basen en projectar els predictors a un altre espai de dimensió superior mitjançant funcions nucli (kernel) lineals o be no lineals. En un problema de classificació binària, és a dir, amb dos grups que

podem anomenar -1 i +1, l'objectiu és trobar un hiperplà lineal de separació $w'\phi(x) + b = 0$ construït a partir del vector x de predicció projectat a un espai de dimensió superior mitjançant una funció no lineal ϕ, un vector w de pesos i una constant b, que classifique qualsevol observació en un dels dos grups {-1, +1}. La funció de classificació és aleshores $f(x) = Sign(w'\phi(x) + b)$.

Atès que, en un problema de classificació binària, hi ha hiperplans de separació infinits, l'objectiu és trobar el pla lineal òptim que separe millor els dos grups. Per trobar-lo, una estratègia és maximitzar la distància o el marge de separació dels plans de suport, respectivament $w'\phi(x) + b \geq +1$ per al grup {+1} i $w'\phi(x) + b \leq -1$ per al grup {-1}. Estos plans de suport s'empenyen fins que arriben a una petita quantitat d'observacions o patrons d'entrenament que respecten les restriccions anteriors i, per tant, es diuen vectors de suport.

L'objectiu de classificació es pot aconseguir maximitzant la distància o marge de separació r entre els dos plans $w'\phi(x) + b = +1$ i $w'\phi(x) + b = -1$ donats per $r = \frac{2}{||w||}$. Això equival a minimitzar la funció de costos

$$C(w) = \frac{||w||^2}{2} + c\sum_{i=1}^{n} \xi_i = \frac{1}{2} w^{'} w + c\sum_{i=1}^{n} \xi_i$$

subjecta a les restriccions $y_i(w^{'}\phi(x_i) + b) \geq 1 - \xi_i$ i $\xi_i \geq 0$, on $c > 0$ és un paràmetre de penalització que equilibra els errors de classificació enfront de la complexitat del model, que està controlat pel marge de separació, i ξ_i, que és l'anomenada variable de folgança. Esta variable és la penalització d'una observació mal classificada i que controla la distància en el costat equivocat de l'hiperplà que un punt pot quedar quan les dades d'entrenament no es poden classificar sense error, és a dir, quan els objectes no es poden separar linealment. La funció ϕ es computa mitjançant funcions no lineals semi-definides positives, K, anomenades nuclis (*kernels*).

La minimització esmentada, generalment es resol a través de:

$$min \frac{1}{2} \sum_{i,j=1}^{n} y_i y_j \alpha_i \alpha_j K(x_i, x_j) - \sum_{i=1}^{n} \alpha_i$$

sotmès a les restriccions lineals:

$$\sum_{i=1}^{n} y_i \alpha_i = 0$$

i $0 \leq \alpha_i \leq C$

On els α_i són multiplicadors de *Lagrange* no negatius i K és la funció nucli. En els problemes de classificació, les funcions habituals del kernel són el nucli lineal $K(x_i, x_j) = x_i' x_j$ o el Gaussià $K(x_i, x_j) = exp(-\gamma||x_i - x_j||^2)$ on γ és el paràmetre del kernel.

Per a la implementació de SVM s'ha gastat la funció del paquet e1071 (Dimitriadou et al., 2017).

3.3.6 Arbres de classificació

Els **Arbres de Classificació** (AC) són classificadors no paramètrics que construeixen arbres de decisió jeràrquics dividint les dades en un determinant node en dos nodes 'fills', d'acord amb una regla *si-aleshores* aplicada a un conjunt de predictors, i partint d'un node arrel que conté tota la mostra. Per tant, els AC poden seleccionar els predictors més importants per a una variable eixida.

Els arbres més habituals, són els anomenats CART, arbres de classificació i regressió (Breiman et al., 1984), on els predictors es dividixen de manera que es minimitze la impuresa del node produït a cada branca t de l'arbre fins que totes les dades es classifiquen en classes k excloents. La mesura d'impureses és l'índex d'impureses de *Gini* definit com:

$$I_G(t) = 1 - \sum_{c=1}^{k} P(c|t)^2 = \sum_{c=1}^{k} \sum_{c \neq j=1}^{k} P(c|t)P(j|t)$$

on $P(c|t)$ és la probabilitat condicional d'una classe c donat el node t. Esta probabilitat es calcula com

$$P(c|t) = \frac{P(c,t)}{P(t)}$$

amb

$$P(c|t) = \frac{\pi(c)n_c(t)}{n_c}$$

i

$$P(t) = \sum_{c=1}^{k} P(c,t)$$

on $\pi(c)$ és la probabilitat d'observar el grup c i $n_c(t)$ és el nombre d'elements del grup c en un node donat t. L'arbre creix fins que no es poden utilitzar més predictors o la impuresa de cada grup en una branca final de l'arbre no es pot reduir encara més. Els predictors no significatius (branques) poden ser "*podats*" des de l'arbre final i eliminats de l'anàlisi.

3.3.7 Boscos aleatoris

Els **Boscos Aleatoris** (BA) van ser proposats per Breiman (2001). Este mètode de classificació construeix una sèrie d'arbres de decisió utilitzant mostres aleatòries *bootstrap* de la mostra

de dades originals. Cadascun d'estos arbres està construït a partir d'un subconjunt aleatori dels predictors que maximitzen els criteris de classificació en cada node.

Es pot obtenir una estimació de la taxa d'error de classificació utilitzant cadascun dels arbres de classificació per predir les dades que no figuren en la mostra *bootstrap (out of bag)* que s'utilitzà per a créixer l'arbre i, a continuació, promediar eixos valors per al conjunt d'arbres, és a dir, per al bosc.

Estes estimacions *out of bag* de la taxa d'error poden ser molt precises si s'han fet crèixer prou arbres. La classificació d'objectes es realitza a partir de la majoria de prediccions donades pels arbres en el bosc aleatori. Tot i que esta estratègia de classificació pot no tenir un avantatge perceptible respecte d'un únic arbre, segons el seu creador Breiman (2001), té una precisió molt destacable entre els algorismes actuals, en concret es comporta molt bé en comparació amb molts classificadors, incloent ADL, XN i MSV.

A més, este mètode és prou fàcil d'utilitzar ja que només té dos paràmetres que la persona usuària ha de definir:

- La quantitat d'arbres aleatoris del bosc.
- El nombre de variables predictores que s'escullen aleatoriàment en cada node de cada arbre.

Estos paràmetres es poden optimitzar fàcilment, encara que els boscos aleatoris no són molt sensibles als seus valors.

Per a boscos aleatoris s'ha gastat la funció randomForest del paquet randomForest (Liaw i Wiener, 2002).

3.3.8 K-Veïns

L'algorisme del k **veïns més propers** (KV) és molt senzill i molt eficaç. Es fan prediccions per a un nou punt mitjançant la cerca en tot el conjunt d'entrenament de les k instàncies més semblants (els veïns) i resumint la variable d'eixida per a estes k instàncies. Per a la regressió, esta podria ser la variable d'eixida mitjana; en la classificació este podria ser el valor de la classe de moda (o més comú).

La estratègia consisteix a determinar la similitud entre les dades. La tècnica més senzilla si els atributs són tots de la mateixa escala (per exemple, tot en polzades) és utilitzar la distància

euclidiana, un nombre que es pot calcular directament en funció de les diferències entre cada variable d'entrada.

L'algorisme pot requerir molta memòria o espai per emmagatzemar totes les dades, però només realitza un càlcul quan es necessita una predicció. També es poden actualitzar les mostres d'entrenament al llarg del temps per tal que les prediccions siguen precises.

Per al mètode de k-veïns s'ha gastat la funció knn del paquet **class** (Venables i Ripley, 2002).

3.4 Dades funcionals

Hi ha ocasions on les dades tenen forma de funció. Encara que les dades es poden recollir de forma discreta, una funció subjau al darrere. Aquest camp es diu en estadística Anàlisi de dades funcionals (FDA). Ramsay i Silverman (2005) és una excel · lent referència en aquest camp.

El cas més habitual i simple és quan les funcions són univariants amb un únic argument, que sol ser el temps. Però, les funcions poden ser multivariants i amb més d'un argument, com és el nostre cas. Al nostre cas, els arguments, són espacials.

En el nostre cas, per mig de la RM s'han anat obtenint talls del cervell. Òbviament, el cervell es continu en l'espai, encara que l'observem en punts discrets de l'espai.

En l'anàlisi de dades funcionals es poden desenvolupar els mateixos anàlisi que en altres branques de l'estadística amb altre tipus de dades.

A l'hora de treballar en funcions es pot fer de manera discreta, gastant els valors on s'han observat les funcions, però no sol ser molt habitual, ja que cada funció pot haver sigut observada en moments/llocs diferents, i a més no estaríem tenint en compte la naturalesa funcional de les dades.

El primer pas habitualment és suavitzar (i registrar si fora necessari) les funcions i obtindre la seua representació en una base apropiada. D'aquesta manera podrem aproximar el valor de la funció en qualsevol punt.

Amb els coeficients de la base podrem realitzar els càlculs que necessitem, sempre tenint en consideració també les propietats de la base de funcions considerada, per exemple, si és o no ortogonal.

Revisarem a continuació alguns conceptes bàsics per al cas de funcions univariants amb un

únic argument, per a després gastar-los i generalitzar-los per al cas de funcions multivariants amb més d'un argument.

Les estadístiques descriptives clàssiques es poden aplicar igualment quan tenim dades funcionals. En el cas de tindre un conjunt de dades funcionals $x_i(t)$ on $i = 1, ..., n$, les corresponents funcions descriptives serien:

- La mitjana funcional:
$$\bar{x}(t) = \frac{\sum_{i=1}^{n} x_i(t)}{n}$$
- La variància funcional:
$$\mathrm{Var}_X(t) = \frac{\sum_{i=1}^{n} [x_i(t) - \bar{x}(t)]^2}{n-1}$$
- La desviació típica funcional:
$$\sigma(t) = \sqrt{\frac{\sum_{i=1}^{n} [x_i(t) - \bar{x}(t)]^2}{n-1}}$$
- La covariància funcional:
$$\mathrm{cov}_X(t_1, t_2) = \frac{\sum_{i=1}^{n} [x_i(t_1) - \bar{x}(t_1)] [x_i(t_2) - \bar{x}(t_2)]}{n-1}$$
- La correlació funcional:
$$\mathrm{corr}_X(t_1, t_2) = \frac{\mathrm{cov}_X(t_1, t_2)}{\sqrt{\mathrm{var}_X(t_1)\,\mathrm{var}_X(t_2)}}$$

Destacar que la mitjana pot no semblar-se massa a les funcions originals si les corbes no estan registrades. A més, les variàncies seran molt altes on s'inclou la variabilitat per amplitud i per fase.

3.4.1 Suavitzat

Les dades funcionals apareixen com un conjunt de valors discrets $y_{i1}, ..., y_{in}$. Per tant, el primer que s'ha de fer és transformar aquestos valors a una funció x_i amb valors $x_i(t)$. En el cas de que els valors discrets es prenguen amb errors observacionals, els quals són necessari eliminar, la conversió de les dades discretes a funció es realitzarà mitjançant el suavitzat. En canvi, si les dades es prenen sense errors, el procediment que es seguirà serà el de la interpolació.

Les combinacions lineals de bases de funcions és el que es sol utilitzar com el mètode principal per a la representació de funcions. Per al nostre cas revisarem la **base de Fourier**.

Base de Fourier:

Podrem aproximar la funció x mitjançant la sèrie de Fourier com:

$$\hat{x}(t) = c_0 + c_1 \sin(\omega t) + c_2 \cos(\omega t) + c_3 \sin(2\omega t) + c_4 \cos(2\omega t) + \dots$$

definida per les bases $\phi_0(t) = 1, \phi_{2r-1}(t) = \sin r\omega t, \phi_{2r}(t) = \cos r\omega t$. Així, la base de Fourier és una base periòdica i el paràmetre ω determina el període $2\pi/\omega$. En el cas de que els valors t_j siguen equiespaciats en un interval acotat on s'estudia x (l'anomenarem T), i el període és igual a la longitud d'aquest interval T, aleshores en dividir la base amb constants apropiades és ortonormal ($\Phi'\Phi = I$).

Destacar que els coeficients es poden estimar per mínims quadrats o per altres mètodes més sofisticats. En la figura 3.8 es mostra la base de Fourier per a 3, 5, 7 i 9 funcions base.

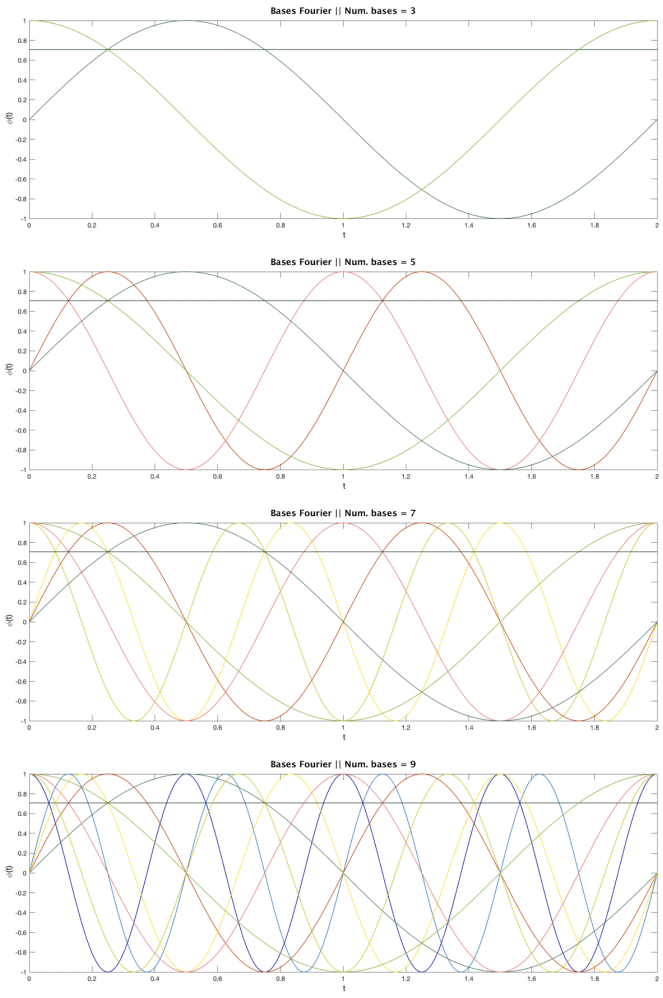


Figura 3.8: Bases de Fourier amb funcions base $\phi_3, \phi_5, \phi_7, \phi_9$.

Capítol 4

Aportacions

A continuació es mostraran els objectius específics per a cadascun dels articles que ha donat lloc aquesta tesi.

- **Aportació 1:** *The brain-anatomy changes in learning algebra problem solving. Studying the reversal error.* Ventura-Campos, Ferrando, Miró-Padilla et al. (2020).

  Aquest és el primer article que forma part de la tesi, el qual està sotmès.

  Una línia important de recerca relacionada amb la resolució de problemes verbals és l'estudi dels processos cognitius implicats quan els subjectes tradueixen problemes del llenguatge a l'àlgebra. Per tant, és essencial saber quines àrees cerebrals estan associades a aquest procés. L'objectiu d'aquest estudi va ser investigar les diferències anatòmiques cerebrals.

  En primer lloc es van obtindre les imatges de RMe per a cadascun dels 33 subjectes, on, per tal d'estudiar les diferències volumètriques cerebrals de SG entre subjectes amb EI i no-EI, es va fer ús del mètode VBM.

  En segon lloc, i per tal de fer la classificació en grups (EI, no-EI), es va fer ús d'una aplicació informàtica similar a la utilitzada en González-Calero et al. (2015). La tasca estava formada per una col · lecció de 16 ítems on se'ls presentaven enunciats de problemes els quals conduïen a fer l'EI. Els enunciats eren diferents: podien aparèixer pistes contextuals o no, quantitats discretes o continues. Així doncs, els subjectes només podien utilitzar multiplicacions/divisions, o bé, adició/subtracció per a expressar l'equació. Per tal de fer la construcció de l'equació se'ls va facilitar les variables i les quantitats que s'havien d'utilitzar on fent 'click' sobre cadascuna de les variables i dels signes construïen l'equació la qual validaven posteriorment.

  A partir de la quantitat de respostes correctes obtingudes amb l'aplicació es van separar els subjectes en dos grups. D'una banda aquells que no cometien cap error i d'altra els

que realitzaven més d'un 40% d'errades.

Els nostres resultats mostraren que el grup amb EI té un volum més gran en el putamen, el que suggereix que aquests subjectes han de fer un major esforç per resoldre problemes, el que suscita la possibilitat de que el putamen no haja madurat de forma correcta. Per al nostre coneixement, aquest és el primer estudi d'anatomia cerebral que mostra l'estructura que dóna suport a l'aprenentatge de la resolució de problemes i la competència algebraica.

- **Aportació 2:** *Detecting and visualizing differences in brain structures with SPHARM and functional data analysis.* Ferrando, Ventura-Campos et al. (2020a).

  Aquest és el segon article que forma part de la tesi, el qual ha sigut publicat a la revista *NeuroImage*.

  En aquest article es presenta un nou mètode per classificar les estructures cerebrals descrites per SPHARM (spherical harmonic representation). El que es va fer fou combinar una tècnica de reducció de dimensions (anàlisi funcional de components principals o anàlisi funcional de components independents) amb selecció de variable pas a pas per a la classificació discriminant lineal.

  Es va comparar el nou mètode que proposàvem amb molts mètodes coneguts en el problema dels subjectes amb EI i no-EI mitjançant l'ús dels putamens esquerres i drets de 33 participants.

  El primer pas fou passar les imatges preprocessades dels putamens (les obtingudes en l'article anterior) a un espai tridimensional (parametrització de les superfícies i representar els putamen amb SPHARM). El segon fou fer la classificació utilitzant l'anàlisi de forma dels putamens on es varen seguir els següents passos:

  En primer lloc, calculàrem la taxa de classificació utilitzant el mètode del *leave-one-out*, és a dir, deixàrem fora un individu cada volta, i a continuació aplicàrem anàlisi de components principals funcional (FPCA) per a la resta. Després obtinguérem les puntuacions del subjecte que havíem deixat fora, i s'utilitzà en el classificador obtingut (discriminant lineal amb selecció de variables) amb la resta d'individus. El procés es va repetir 33 vegades, un per a cada subjecte. En segon lloc es va utilitzar el mètode del *leave-one-out* com abans però aplicant anàlisi de components independents funcional (FICA). El procés es va repetir 33 vegades, un per a cada subjecte.

  La comparació va mostrar que la nostra proposta no només proporcionava un rendiment excepcional en termes de potència predictiva, sinó que també és valuosa en termes d'interpretació, ja que produeix una funció discriminatòria lineal per a les estructures 3D, la qual cosa és un avantatge davant mètodes globals que es centren en els resultats numèrics, però no donen interpretació d'on es produeixen les diferències.

  Destacar que aquest article ha conduit a la publicació d'un altre article *(A neuroimaging data set on problem solving in the case of the reversal error: Putamen data.* Ferrando, Ventura-Campos et al. (2020b)), el qual ha sigut publicat en la revista *Data in Brief*, on s'han publicat les dades associades a aquest estudi.

- **Aportació 3:** *Ordinal classification of 3D brain structures by functional data analysis.* Ferrando, Epifanio et al. (2020).

  Aquest és el tercer article que forma part de la tesi, el qual està en fase de revisió.

  L'objectiu de la classificació ordinal és identificar a quin conjunt de classes ordenades pertany un nou element, sobre la base d'un conjunt de dades d'entrenament. Tot i que la classificació nominal, on les classes estan desordenades, ha rebut molta atenció i hi ha moltes metodologies tant per a dades multivariants com funcionals, aquest no és el cas de la classificació ordinal.

  En aquest estudi s'han introduït quatre mètodes de classificació ordinal per a dades funcionals, concretament dades funcionals multiargument i multivariants. S'ha analitzat el rendiment dels mètodes proposats en diferents conjunts de dades reals i s'han comparat amb altres alternatives. Aquests conjunts de dades pertanyen a un problema neuroeducatiu i un problema neuropatològic, on les estructures cerebrals 3D estan representades per dades funcionals.

  En particular, el conjunt de dades referent al problema neuroeducatiu ha sigut l'estudi del problema de l'EI on les dades que s'han utilitzat en aquest cas han sigut un total de 67 subjectes dividits en tres grups: els subjectes que no cometen cap error, els que fallen més del 40% de les respostes i els que es troben en mig d'aquestos dos grups. D'altra banda, el conjunt de dades referents al problema neuropatològic venen donades per un total de 28 subjectes els quals s'han dividit també en tres grups ordenats depenent de les característiques patològiques de cadascun d'ells.

  Els resultats confirmen que les metodologies propostes, on es té el compte l'ordre, milloren els resultats. S'ha vist que no hi ha cap mètode que siga el millor en tots els casos. No obstant això, s'ha vist que un mètode pot funcionar millor en alguns conjunts de dades i pitjor en altres conjunts de dades. Aquesta és la raó per la qual tenir diferents metodologies alternatives per abordar una classificació en un problema ordinal és una bona opció.

- **Aportació 4:** *The neural basis of the reversal error. How a competent solver solves this algebra problem.* Ventura-Campos, Ferrando, Epifanio et al. (2020).

  Aquest és el quart article que forma part de la tesi, el qual està sotmès. Una versió preliminar d'aquest treball va constituir el meu Treball Final de Grau (Ferrando, 2018) i va ser premiat amb el premi SEA al millor treball final de grau del XVI Concurs Student d'Estadística Aplicada organitzat pel Servei d'Estadística de la Universitat Autònoma de Barcelona i l'Institut d'Estadística de Catalunya. És un premi nacional per a treballs final de grau en Estadística.

  L'objectiu que es va perseguir amb esta investigació va ser determinar les bases neuronals lligades a l'EI, i poder fer una classificació en grups, aquells que fan error *vs* els que no.

  Un total de 20 subjectes participaren en aquesta investigació. En primer lloc, es va passar a l'adquisició de les imatges de RMf mentre els subjectes realitzaven la tasca de resolució de problemes associats a l'EI.

En segon lloc, després de l'adquisició de les imatges de RMf i amb la finalitat de recollir les dades de l'EI es va utilitzar l'aplicació informàtica similar a la de González-Calero et al. (2015).

Amb totes les dades recollides, es van separar els subjectes en dos grups basant-nos en les respostes recollides dins de RM i amb l'aplicació informàtica.

Seguidament, es va dur a terme l'anàlisi de les imatges de RMf mitjançant el programa d'anàlisi estadística *SPM*.

Per tal de complir el segon objectiu d'aquest estudi, que era fer una classificació en grups a través de les àrees cerebrals que diferenciaven als subjectes, es van utilitzar les àrees d'activació com a variables i amb elles es van provar un total de 13 mètodes de classificació.

Els principals resultats d'imatge mostraven que els subjectes que cometen EI requereixen d'una major demanda de recursos cognitius, com són els processos atencionals i memòria de treball, associats a les àrees frontal i parietal bilateral. En canvi, els subjectes que no cometen l'EI mostren una activació en les àrees parietal i frontal lateralitzades a l'esquerra. Açò podria apuntar cap al fet que els subjectes competents tinguen desenvolupats uns coneixements algebraics sòlids quant al significat de variable i d'equacions algebraiques, per la qual cosa no tindrien necessitat d'una gran demanda cognitiva en la resolució de la tasca. L'anàlisi de comparació entre EI *vs* no-EI, activen l'àrea temporal medial dreta, associada amb el processament de nova informació semàntica, és a dir, situacions no familiars o dominades per l'individu que requereixen una decodificació de la situació amb la finalitat d'atorgar-li significat a la translació des del llenguatge natural al llenguatge algebraic.

En referència als resultats dels classificadors, els mètodes que millor classifiquen en el nostre cas, i tenint en compte que és una mostra menuda, eren l'ADF, seguit per l'ADL i la RL els quals tenien el mateix percentatge d'èxits, justament mètodes senzills, com Hand (2006) suggeria que podia ocórrer en problemes reals.

Aportació 1:

The brain-anatomy changes in learning algebra problem solving. Studying the reversal error

Ventura-Campos, Ferrando, Miró-Padilla et al. (2020)

The brain-anatomy changes in learning algebra problem solving. Studying the reversal error

Ventura-Campos, N. [1,2] , Ferrando L. [2], Miró-Padilla, A. [2] and Ávila, C. [2].

[1] Department of Education and Specific Didactics. Universitat Jaume I, Castellón, Spain

[2] Neuropsychology and Functional Neuroimaging Group, Universitat Jaume I, Castellón, Spain.

Abstract

An important line of research related to the resolution of verbal problems is the study of the cognitive processes involved when subjects translate problems into the language of algebra. Therefore, it is essential to know which brain areas are associated with this process. One of the most common errors in problem solving is the reversal error (RE). A case where students recognize the information in the statement, but are not able to build a correct equation. The aim of this neuroeducational study is to investigate the brain anatomy differences between two groups, one group that commits RE and a second group that does not. Magnetic resonance images of 37 normal and healthy participants between the ages of 18–25 years were acquired. Changes in gray matter were assessed using voxel-based morphometry analysis. Our results show that the RE group has larger volume in the putamen, suggesting that these subjects have to make a greater effort to solve problems, which results in the putamen not maturing properly. To our knowledge, this is the first brain anatomy report showing the structure that supports the learning of problem solving and algebraic competence.

Keywords: Neuroeducation, algebra problem solving, reversal error, brain anatomy, mathematical learning.

Introduction

Better understanding of the way the brain learns mathematics has led to a significant number of articles. Most of these studies focus on number sense, arithmetic learning, and difficulties in mathematical learning (Feigenson et al., 2004; Cantlon et al., 2006; Ansari, 2008; Zamarian et al., 2009; Butterworth, 2010 and Butterworth et al., 2011). In the specific case of problem solving, although several behavioral studies have been carried out (Clement, 1982; Clement et al., 1981 and Clement et al., 1980), as far as we know, few studies have used neuroimaging techniques to find out more about the underlying neural processes or structural changes involved in mathematical problem solving. These studies (Qin et al., 2004; Lee et al., 2007; Anderson et al., 2008 and Anderson et al., 2012) could shed light on the transition from arithmetic to the symbolic language of algebra, where students have to develop abstract reasoning skills that allow them to generalize, model, and analyze mathematical equations and theorems.

Problem solving is considered one of the most important components of the study of mathematical knowledge (Castro & Ruiz, 2015). According to Halmos (1980), there are several essential elements of mathematics, such as theorems, demonstrations, formulas, theory, etc., but he suggests that the most important things in mathematics are problems and their solutions. In addition, Kleiner (1986) emphasizes that effort in solving certain problems leads to good development of concepts and mathematical theories. The importance of problem solving is widely accepted by researchers, as Castro (2008) states*:* "Problem solving is not just a scientific activity; it is also a type of educational task that must be given a prominent position in the teaching and learning processes of children, adolescents, and students in general" (p. 2).

A mathematical problem can be considered a combination of subproblems (Nathan et al., 1992). For each subproblem, the student must carry out an analysis-synthesis process. During the analysis,

the activation of a conceptual scheme at the cognitive level will be attempted in order to establish a relationship between quantities that, during synthesis, will be placed in the mathematical operation. Obviously, the difficulty of identifying some of the conceptual schemes can be an obstacle to defining the relationship between the quantities and, therefore, completing the resolution process. However, it is possible that, after the correct identification of the conceptual scheme, an incorrect formulation of the relationship may occur. An example of this situation would be the reversal error (RE).

Some of the investigations on problem solving, such as those by Clement (1982), Clement et al. (1981), and Clement et al. (1980), found that on some comparison problems, both additive and multiplicative, most of the students made mistakes when translating some sentences from natural language to algebraic language. The structure of the sentences these researchers used to find the error was similar to the following example: "Write an equation using the variables S and P to represent the following statement: *There are six times as many students as professors at this university*. Use S for the number of students and P for the number of professors" (Clement, 1982, p.17). In these statements, it was observed that most of the wrong answers were in the form of $P = 6 \cdot S$, which they called RE because students reversed the order of the letters, compared to the correct answer $S = 6 \cdot P$. The initial research by Clement (1982) was preceded by numerous behavioral studies on RE (Wollman, 1983; Cooper, 1986; López-Real, 1995; Fisher, 1988; Cohen & Kanim, 2005; Kim et al., 2014; González-Calero et al., 2015; González-Calero et al., 2019). However, these studies did not consider the importance of human brain development in learning, which is the aim of this study.

Knowledge about the way the brain learns could have a major impact on education. Understanding the brain mechanisms underlying learning and memory, as well as the effects of genetics,

environment, emotion, and age on learning, could transform educational strategies and make it possible to devise programs that optimize learning for people of all ages and with diverse needs (Blakemore & Frith, 2005). Therefore, educational neuroscience could help to better understand the relationship between biological brain development and the development of the human capacity for mathematical cognition, mediated by educational experience (Royer, 2003).

It would be interesting to see the relationship between the brain and learning in the academic field. This relationship can be seen in studies that have used Magnetic Resonance Imaging (MRI), which can study even the most complex mental processes, such as problem solving (Hanakawa et al., 2003). Radford and André (2009) reported that little research has focused on the brain and advanced mathematical thinking, in particular the brain-algebra relationship during development. Among these studies, the longitudinal study by Qin et al. (2004) used functional MRI to see the differences in algebraic learning between adults and teenagers. The results showed that, after practice on verbal problem solving, both adolescents and adults presented reduced activation in the prefrontal cortex, which is involved in mathematical cognition and other higher-order processes that develop throughout childhood and adolescence. Moreover, a reduction in the activation of the left parietal cortex, which holds an image of the equation, and an increase in the left putamen were observed only in adolescents. These results suggest that the brain response in adolescents is more plastic in this neural stage of development and, therefore, undergoes more changes due to practice and learning effects. In conclusion, this increased brain response in adolescents due to practice seems to indicate that this period would be the most appropriate one for learning algebra (Qin et al., 2004). These observations would be consistent with neural changes due to development explored in several brain imaging studies (Blakemore, 2012; Sowell & Jernigan, 1998; Sowell et al., 2003 & 1999a; Giedd et al., 1999; Wierenga et al., 2014).

These MRI studies of children during their development have shown that, during the period of adolescence, the white matter (WM) volume continues to increase (there are even some local areas that change rapidly), and so the grey matter (GM) in some areas begins to lose volume (Sowell et al., 2003 & 1999a). In addition, Blakemore's (2012) extensive review of MRI studies performed during development provides more information about brain maturation. Blakemore (2012) states that brain development does not end at an early age, but rather it extends to adolescence, reaching its maximum volume around the age of 25 (Caviness et al., 1996; Sowell & Jernigan,1998). Changes take place in both the GM and the WM, and the WM has the greatest development during this stage (Ortiz, 2009). In addition, the study by Wierenga et al. (2014) reveals that the GM volume in the striatum (dorsal: the putamen, caudate; and ventral: nucleus accumbens) decreases linearly with age from seven to 23 years old (see Figure 2 from Wierenga et al., 2014). The development of the striatum structures indicates that there are dynamic changes during development that are related to changes in cognitive development, experience, and behavior.

This literature on the areas of algebra learning and brain maturation highlights the importance of introducing algebra at the age when the brain is ready. The brain is prepared when the structures related to learning are in the process of maturation and there is greater brain plasticity, which supports learning, as Qin et al. (2004) indicated. This period is when the students are prepared to absorb all the knowledge and assimilate the new mathematical knowledge. Therefore, the goal of this study is to investigate the anatomical brain differences among subjects while performing a problem-solving task with associated RE, thus exploring the effects of learning problem solving skills and brain maturation. To achieve this objective, we will observe the differences in GM volume between the group of participants who perform RE and those who do not (non-RE), using the Voxel Based Morphometry (VBM) method.

Materials and methods

Participants

In this study, participants were 37 students at the *University Jaume I* with ages ranging from 18-26 years. Informed consent was obtained from each participant before the study. For structural analysis, we used 33 participants divided into two groups: the non-RE group (18 subjects, 9 female; mean age: 22.5, SD: 2.53), who responded correctly 100%, and the RE group (15 subjects, 11 female, mean age: 21.466, SD: 2.1), who failed on more than 50% of the answers on the task. The other four participants were excluded because they did not perform the task correctly or their failure rate was in the 50-100 range (50,100).

The study exclusion criteria were the presence of neurological and medical illness, trauma with loss of consciousness lasting more than one hour, and the typical resonance exclusion criteria such as iron prosthesis and dental implants.

Experimental paradigm

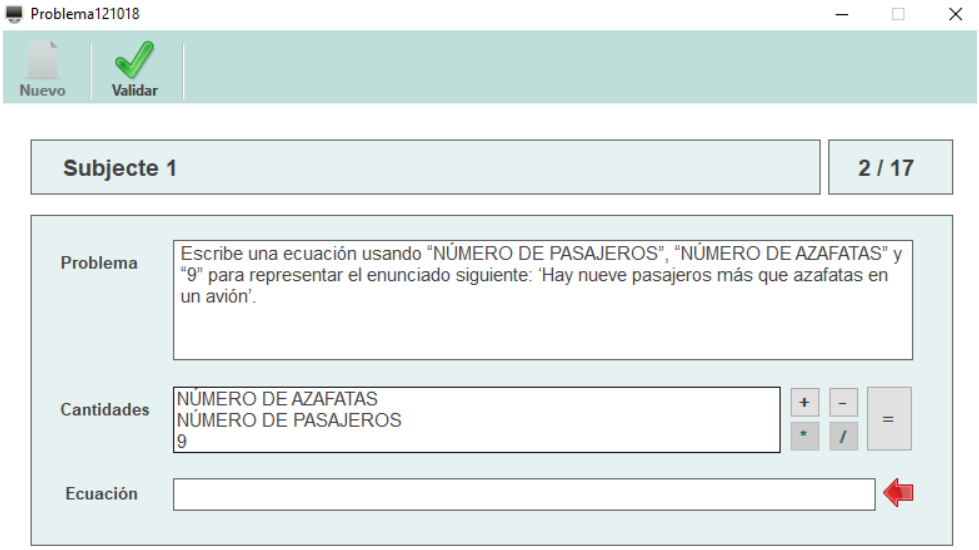
First, the structural image of each subject was obtained with the MRI. Second, outside of the MRI, to obtain the classification of subjects who committed RE vs. non-RE, we used an application similar to González-Calero et al. (2015), where the participants had to build a mathematical equation for each of the statements presented to them.

The task contained a total of 16 statements (2x2x2x2) in Spanish, focusing on: 1) whether the comparisons are multiplicative or additive; 2) whether the comparisons are increasing (times more than, more than) or decreasing (times less than, less than); 3) Contextual or non-contextual clues; and 4) Discrete or continuous amounts. Thus, subjects could use only multiplication/division or

addition/subtraction to express the equation. To construct the equation, we made it easier by giving them the variables and the amounts they had to use (see Figure 1). By clicking on each of the variables and operation signs, they built the equation and later validated it. It should be noted that the response time was unlimited.

Figure 1

Application implemented for data collection in studies on the RE (González-Calero et al, 2015).



Neuroimaging data acquisition

Structural MRI data were acquired using a 1.5 Teslas Siemens Symphony scanner and a 3 Teslas Philips scanner. The parameters of the MPRAGE sequence were as follows: **1)** The Siemens Symphony scanner used a high-resolution T1-weighted, repetition time = 2200 ms, echo time = 3 ms, flip angle = 90º, matrix size =256 x 256 x 160, voxel size = 1 x 1 x 1 mm; and **2)** The Philips scanner used a high-resolution T1-weighted, repetition time = 8.4 ms, echo time = 3.8 ms, matrix size = 320 x 320 x 250, voxel size = 0,75 x 0,5 x 0,8 mm. All the scanner acquisitions were performed in parallel to the anterior commissure-posterior commissure plane (AC-PC), and they

covered the entire brain. Participants were placed in a supine position in the MRI scanner. Their heads were immobilized with cushions to reduce motion degradation, and they were asked to minimize their head movement.

Neuroimaging analysis: Voxel-Based Morphometry (VBM) Analysis

The VBM analysis was performed using the Computational Anatomy Toolbox 12 (CAT12.5, http:// dbm.neuro.uni-jena.de/cat/), a toolbox from the software: Statistical Parametric Mapping (SPM12 (v7219), Wellcome Trust Centre for Neuroimaging, Londres, UK, http:// www.fil.ion.ucl.ac.uk/spm/software/spm12). We used the standard procedure suggested by CAT12, which included: **1)** Normalization and segmentation of the images into GM, WM, and cerebrospinal fluid; **2)** Alignment of the GM and WM between the images; **3)** Using the DARTEL-template in the VBM analysis to obtain the GM and WM tissues, which were normalized to MNI standard space; **4)** Modulation by the "affine + nonlinear" components derived from spatial normalization; **5)** Estimation of the total intracranial volume (TIV) for each subject; **6)** Quality check of the images; no outliers were identified; and **7)** Finally, the images were spatially smoothed using a Gaussian filter (8 mm Full-Width Half-Maximum, FWHM).

Statistical analysis

The smoothed GM images were entered into a statistical analysis using the General Linear Model in SPM12. A two-sample t-test was performed to obtain the differences in GM volume between groups using two covariates of non-interest: i) the TIV to correct different brain sizes; and ii) because the data were obtained from two machines, we created a covariate to remove the effect of

obtaining the MRI from different machines. Finally, we defined the contrast to obtain the results of the analysis. The statistical criterion was set at $p < .05$, using Family-wise error (FWE) cluster-corrected for multiple comparisons (voxel-level uncorrected threshold of $p < .005$ with a critical cluster size).

Results

Behavioral Results

The response time mean was 29.80s (RE group = 30.39s; non-RE group = 29.32s). The results of the t-test showed no significant differences in time between groups. Tables 1 and 2 provide the performance on the task by the RE group and the non-RE group, respectively.

Table 1

Task results for the RE group.

Participant	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
RE	12	11	11	15	14	6	14	12	13	16	7	10	14	8	16
Correct	4	2	2	1	2	9	2	2	3	0	9	6	2	8	0
Machine	1	1	1	1	1	1	1	1	-1	-1	-1	-1	-1	-1	-1

Note. The 'Participant' row contains the ID of each subject. The 'RE' row contains the number of incorrect answers on the task. The 'Correct' row is the opposite of the 'RE' row. The 'Machine' row is a covariate to remove the effect of getting the images from different machines, where 1 is the Philips machine and -1 is the Siemens machine.

Table 2

Task results for the non-RE group.

Participant	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
RE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Correct	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16
Machine	1	1	1	1	1	1	1	1	1	-1	-1	-1	-1	-1	-1	-1	-1	-1

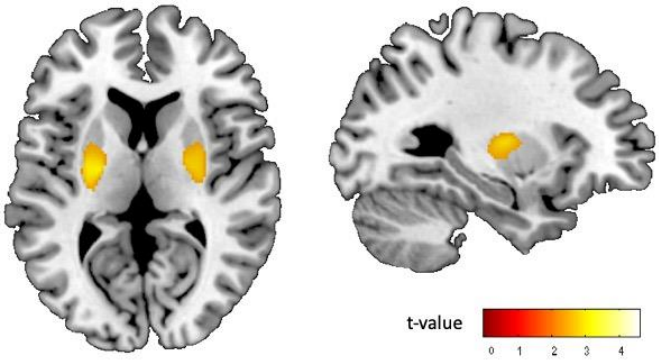
Note. The 'Participant' row contains the ID of each subject. The 'RE' row contains the number of incorrect answers on the task. The 'Correct' row is the opposite of the 'RE' row. The 'Machine' row is a covariate to remove the effect of getting the images from different machines, where 1 is the Philips machine and -1 is the Siemens machine.

MRI Results

To study the difference in GM volume between groups, a two-sample t-test was performed. When studying the RE vs. non-RE contrast, we observed an increase in GM volume in the bilateral putamen in the participants with RE. The MNI coordinates for the left putamen were x=-23, y=-12, z=0, with a Z-value=3.96 (k=938), and the MNI coordinates for the right putamen were x=26, y= -8, z=3, with a Z-value= 3.41 (k=813) (see Figure 2). The opposite contrast did not yield any significant differences.

Figure 2

Neuroimaging results of the two-sample t-test performed between groups.



Note. The figure represents the contrast: RE group vs. non-RE group in our study (FWE cluster-corrected $p < .05$ for multiple comparisons at the whole-brain level, voxel $p < .005$, and a cluster size of $k = 813$ voxels). Results show a higher volume in the bilateral putamen in subjects who commit RE.

Discussion

The present educational neuroscience study focuses on the brain anatomy differences in young people who have performed an algebra problem-solving task that can produce the so-called RE (Clement, 1982).

The algebra problem-solving process often involves multiple cognitive components to reach a solution. Much of the relevant mathematical and pedagogical literature on problem solving was influenced by Polya's (1945) four phases. The first phase in solving a problem is to read and understand the problem and identify the data involved. The second phase is to devise a plan, that is, find the connection between the data and the unknown. This phase requires the skill of choosing an appropriate problem-solving heuristic. The third phase is to carry out the plan, which involves

solving the problem by executing the heuristic. And the final phase, looking back, basically involves checking to whether all the information was used and whether the answer makes sense.

If you look at Polya's first phase, reading comprehension is important in order to know what the problem is asking. Therefore, it is to be expected that the better the reading comprehension, the greater the problem-solving accuracy. The working memory has been shown to predict individual differences in reading comprehension (Daneman & Merikle, 1996). Furthermore, several studies have shown correlations between working memory and problem solving, depending on reading comprehension (Fuchs et al., 2006; Swanson et al., 1993). However, in the study by Lee et al. (2004), the working memory contributed independently to problem solving when reading comprehension was controlled. Therefore, this study suggests that the working memory cannot be attributed solely to its relationship with reading comprehension. This finding agrees with previous studies that relate working memory to mathematical problem solving (Lee et al., 2004; Passolunghi & Pazzaglia, 2004; Swanson, 2006) and predict the performance on a problem-solving task in fifth grade (Lee et al., 2009). The findings by Lee et al. (2009) showed that the working memory plays an important role in text decoding and in constructing a schematic representation of the problem. Moreover, this study suggests that working memory was strongly associated with the selection of the mathematical operations the participants needed to compute the solution. Furthermore, some studies have claimed that a low working memory capacity would lead to a low ability to solve mathematical problems (Díaz, 2010; Ruiz et al., 2012). Thus, the evidence seems to suggest that the working memory is involved in the problem-solving process and, therefore, in all of Polya's phases.

In addition to working memory, other executive functioning domains may also contribute to performance on algebraic problems. Switching is an executive functioning domain that is often

examined in the mathematics literature (Clements et al., 2016). To complete Polya's second phase, the information extracted from a sentence in the algebraic problem has to be integrated with pre-existing knowledge to form a conceptual problem model (Kintsch, 1998). Hence, switching may be important in the representation of the problem and, therefore, in the ability to move between alternative sets of mental operations to choose an appropriate heuristic to solve the problem. This executive function is also a feature that will play an important role in Polya's third phase, where the subjects switch from a problem-solving heuristic to letter-symbolic algebra. Finally, to solve the algebraic problem, it is necessary to have good mathematical skills and knowledge of the mathematical concepts of the problem.

In order to study the brain anatomy of the subjects who performed the algebra problem-solving task with RE, we compared the GM brain volume of the RE group and the volume of the non-RE group using the VBM method. An increase in the GM volume in the bilateral putamen was found in the RE group, compared to the non-RE group. This follows along the lines of the study by Qin et al. (2004), where both adolescents and adults performed a verbal problem-solving task; the results showed that, after performing a practice session, only the adolescents showed increased activation in the left putamen ($x = -26$, $y = -9$, $z = 5$). The authors suggested that the adolescents may have required more effort in performing complex calculations, thus making use of brain regions that are not necessary in adult activity. Furthermore, they also suggest that this might be related to the striatal region changes observed in adolescents during maturation (Sowell et al., 1999b).

Moreover, in different fMRI studies (Hsu & Goh, 2016 and Amalric & Dehaene, 2016), a bilateral activation of the putamen was found after performing mathematical tasks, which leads us to think that the putamen is an important structure in performing mathematical tasks. The study by

Arsalidou and Taylor (2011) showed the significant contributions of the left putamen on number tasks. They suggested that the putamen's involvement in number tasks was related to integrating information by pacing the coordination of top-down and bottom-up items. They also proposed that the putamen may play a role in assigning priority values or sequencing information that needs to be processed during number tasks.

On the other hand, regarding the putamen's association with executive functions that can help to perform algebraic problems, Constantinidis and Klingberg (2016) carried out a review investigating activation during working memory tasks. The authors indicated that, although the putamen (included in the dorsal striatum) has no main function in working memory, activity has been found in this structure, indicating its participation during this cognitive process, with the striatum being one of the brain structures related to working memory. Moreover, the putamen is associated with anticipations of high versus low working memory updating (Yu, FitzGerald & Friston, 2013). Therefore, the evidence seems to indicate that the putamen also plays a relevant role in working memory.

In addition, investigations on the role of the putamen in executive functions suggest that the striatum also appears to be able to perform switching. Sowell et al. (1999b) indicated that the putamen has been involved in the cognitive function of task switching, as in learning of stimulus–response associations (Packard and Knowlton, 2002), which is linked to frontal system function (Rolls, 1994) and improves throughout adolescence (Levin et al., 1991). This suggests that normal brain development has temporal and functional relationships between simultaneous post-adolescent reductions in GM density in frontal and striatal regions (Sowell et al., 1999b). Thus, Sowell et al. (1999b) highlighted the potential importance of frontal/striatal maturation in adult cognition in achieving a temporal and spatial progression of post-adolescent maturation into the

frontal lobes. In addition, neuroimaging evidence suggests that the fronto-striatal network is activated when subjects have to choose from many possible responses (Desmond et al., 1998), and so the striatum may play a role in selecting between alternative responses. Thus, this ability to readily switch between stimulus-response would seem to be a necessary component of many complex executive tasks, such as planning, problem solving, and strategizing. Although these executive functions rely on the prefrontal cortex, the task-switching functions of the striatum may make a critical contribution to executive abilities via the fronto-striatal network (Packard & Knowlton, 2002).

Finally, we understand that the effects of low math skills and problem-solving difficulties in post-adolescent subjects may be due to the type of learning obtained at school during brain development. Therefore, to find out more about the brain development of the putamen during learning, this structure must be examined in detail in preadolescent children. The structural study by Sandman et al. (2014) with 50 preadolescent children using three different methods of analysis, including the VBM method, was conducted to determine and locate areas where the GM volume was associated with poor cognitive performance. The participants were assessed with the WISC-IV test and declarative memory, motor, and executive functioning tasks. Their results showed that the larger the GM volume in the bilateral putamen, the worse the cognitive performance, especially on working memory tasks. Sandman et al. (2014) concluded that larger GM volume in the putamen was associated with impaired cognitive function in typically developing young children.

This inverse relationship between GM volume and learning might be explained from a developmental perspective. According to Kanai and Rees (2011), the reduction in GM volume is thought to reflect pruning, as a process of removing inefficient synapses and neurons underlying brain maturation. Thus, a smaller GM volume may be a consequence of synaptic pruning, which

leads to more efficient processing. This idea is supported by Hartzell et al. (2016). They studied a group of professional Vedic Sanskrit Pandits in India who trained from childhood for about 10 years in an ancient, formalized tradition of oral Sanskrit text memorization and recitation, mastering the exact pronunciation and invariant content of multiple 40.000–100.000 word oral texts. The mean age of the participants was 22 years, and the control group consisted of members of India's National Brain Research Centre community or students from a nearby technical college. The authors found that Pandits showed smaller GM volume than controls in subcortical regions, including the putamen. Although these results were unexpected, they proposed the explanation that they indicate faster maturation of these subcortical regions in Pandits, based on the study by Wierenga et al. (2014). However, we can now add the assumption that greater learning leads to a smaller GM volume in the putamen.

Conclusions

Based on our results showing a higher GM volume in the bilateral putamen in the RE group, we conclude that the RE group requires a large executive function capacity to solve an algebra problem, with the putamen playing a role in this process during adolescence. We tentatively propose that subjects who commit RE have a greater demand for the putamen in algebra problem solving during their learning and development. They need to make a greater effort on problem solving, resulting in the putamen not maturing properly and, thus, not decreasing its GM volume as it should. Overall, our results demonstrate a strong effect of brain maturation on algebra learning and problem solving during development.

To our knowledge, this is the first brain anatomy report showing the structure that supports the learning of problem solving and algebraic competence. In future work, we will continue with this

educational neuroscience study by collecting data from participants through functional MRI while they perform a verbal problem-solving task with RE inside the machine. Our aim is to investigate the underlying neural basis of RE. Moreover, we will study classification methods to identify competent resolvers, using the results of functional neuroimaging as biomarkers.

References

Amalric, M., & Dehaene, S. (2016) Origins of the brain networks for advanced mathematics in expert mathematicians. *PNAS*, 113(18), 4909-4917. DOI:10.1073/pnas.1603205113.

Anderson, J. R., Betts, S., Ferris, J. L., & Fincham, J. M. (2012). Tracking children's mental states while solving algebra equations. *Human Brain Mapping*, *33*(11), 2650–2665. DOI:10.1002/hbm.21391.

Anderson, J. R., Fincham, J. M., Qin, Y., & Stocco, A. (2008). A central circuit of the mind. *Trends in Cognitive Sciences.* 12, 136–143.

Ansari, D. (2008). Effects of development and enculturation on number representation in the brain. *Nature Reviews Neuroscience*, 9, 278-291.

Arsalidou. M. & Taylor M.J. (2011). Is 2+2=4? Meta-analyses of brain areas needed for numbers and calculations. *Neuroimage*, 54, 2382-2393.

Cantlon, J. F., Brannon, E. M., Carter, E. J., Pelphrey, K. A. (2006). Functional imaging of numerical processing in adults and 4-y- old children. *PLoS Biology*, 4:e125.

Blakemore, S. J. (2012). Imaging brain development: The adolescent brain. *NeuroImage*, *61*(2), 397–406. DOI:10.1016/j.neuroimage.2011.11.080.

Blakemore, S. J. & Frith, U. (2005). *The learning brain: lessons for education*. Oxford: Wiley-Blackwell.

Butterworth, B. (2010). Foundational numerical capacities and the origins of dyscalculia. *Trends in Cognitive Science*, 14, 534-541.

Butterworth, B., Varma, S. & Laurillard, D. (2011). Dyscalculia: from brain to education. *Science*, 332, 1049-1053.

Castro, E. (2008). Investigación en Educación Matemática XII. In R. Luengo, B. Gómez, M. Camacho & L.J. Blanco (Eds.), *Resolución de problemas. Ideas, tendencias e influencias en España* (pp. 113- 140). Badajoz: SEIEM.

Castro, E. & Ruiz, J. F. (2015). Matemáticas y resolución de problemas. En P. Flores y L. Rico (Eds.), *Enseñanza y Aprendizaje de las Matemáticas en Educación Primaria* (pp. 89- 107). Madrid: Pirámide.

Caviness, V. S., Jr., Kennedy, D. N., Bates, J. F., & Makris, N. (1996). The developing human brain: A morphometric profile. In R. W. Thatcher, G. R. Lyon, J. Rumsey, & N. Krasnegor (Eds.), *Developmental neuroimaging: Mapping the development of brain and behavior* (pp. 3–14). Academic Press.

Clement, J. (1982). Algebra Word Problem Solutions: Thought Processes Underlying a Common Misconception. *Journal for Research in Mathematics Education*, 13(1), 16–30.

Clement, J., Lochhead, J., & Monk, G. S. (1981). Translation Difficulties in Learning Mathematics. *The American Mathematical Monthly* , 88, 286-290.

Clement, J., Lochhead, J., & Soloway, E. (1980). Positive effects of computer programming on students understanding of variables and equations. *Proceedings of Annual Conference of the*

American Society for Computing Machinery, 467-474.

Clements, D. H., Sarama, J., & Germeroth, C. (2016). Learning executive function and early mathematics: Directions of causal relations. *Early Childhood Research Quarterly, 36*, 79-90. DOI:10.1016/j.ecresq.2015.12.009

Cohen, E., & Kanim, S. E. (2005). Factors influencing the algebra "reversal error". *American Journal of Physics,* 73(11), 1072–1078.

Constantinidis, C., & Klingberg, T. (2016). The neuroscience of working memory capacity and training. *Nature Reviews Neuroscience*, *17*(7), 438–449. DOI: 10.1038/nrn.2016.43.

Cooper, M. (1986). The dependence of multiplicative reversal on equation format. *Journal of Mathematical Behaviour*, *5*(2), 115–120.

Daneman, M. & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin and Review*, 3, 422-433.

Desmond, J.E., Gabrieli, J.D.E. & Glover, G.H. (1998). Dissociation of frontal and cerebellar activity in a cognitive task: evidence for a distinction between selection and search. *Neuroimage*, 7, 368–78

Díaz, R. (2010). *La memoria de trabajo y su relación con habilidad numérica y el rendimiento en el cálculo aritmético elemental* (Thesis). Universidad Pedagógica Nacional Francisco Morazán, Honduras.

Feigenson, L., Dehaene, S., & Spelke, E. (2004) Core systems of number. *Trends in Cognitive Science*, 8, 307-314.

Fisher, K. M. (1988). The students-and-professors problem revisited. *Journal for Research in Mathematics Education*, 19(3), 260–262.

Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., Schatschneider, C. & Fletcher, J. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology*, 98, 29-43.

Giedd, J., Blumenthal, J., Jeffries, N., Castellanos, F., Liu, H., Zijdenbos, A., Paus, T., Evans, A. & Rapoport, J. (1999). Brain development during childhood and adolescence: a longitudinal MRI study, *Nature neuroscience,* 2, 861-863.

González-Calero, J. A., Arnau, D., & Laserna-Belenguer, B. (2015). Influence of additive and multiplicative structure and direction of comparison on the reversal error. *Educational Studies in Mathematics*, 89(1), 133–147.

González-Calero, J. A., Berciano, A. & Arnau, D. (2019). The role of language on the reversal error. A study with bilingual Basque-Spanish students. *Mathematical Thinking and Learning. Advance online publication.* DOI: 10.1080/10986065.2020.1681100.

Halmos, P. R. (1980). The Heart of Mathematics. *The American Mathematical Monthly,* 87(7), 519. DOI:10.2307/2321415.

Hanakawa, T., Honda, M., Okada, T., Fukuyama, H., & Shibasaki, H. (2003). Neural correlates underlying mental calculation in abacus experts: a functional magnetic resonance imaging study. *NeuroImag*e, 19(2), 296–307. DOI: 10.1016/S1053-8119(03)00050-8.

Kanai, R., & Rees, G. (2011) The structural basis of inter-individual differences in human behaviour and cognition. *Nature Reviews Neuroscience,* 12, 231-242.

Hartzell, J. F., Davis B., Melcher, D., Miceli, G., Jovicich, J., Nath, T., Singh, N.C., & Hasson, U. (2016) Brains of verbal memory specialists show anatomical differences in language,

memory and visual systems. *Neuroimage* 131, 181–192 .

Hsu, C. W., & Goh, J. O. (2016). Distinct and Overlapping Brain Areas Engaged during Value-Based, Mathematical, and Emotional Decision Processing. *Frontiers in human neuroscience,* 10, 275. DOI:10.3389/fnhum.2016.00275

Kim, S. H., Phang, D., An, T., Yi, J. S., Kenny, R., & Uhan, N. A. (2014). POETIC: Interactive solutions to alleviate the reversal error in Student-professor type problems. *International Journal of Human Computer Studies,* 72 (1), 12-22.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. NY: Cambridge University Press.

Kleiner, I. (1986). Famous Problems in Mathematics: an Outline of a Course. *For the Learning of Mathematics*, 6(1), 31-38.

López-Real, F. (1995). How important is the reversal error in algebra? In S. Flavel et al. (Eds.), GALTHA (Proceedings of the 18th Annual Conference of the Mathematics Education Research Group of Australasia) (pp. 390–396). Darwin, Australia: MERGA.

Lee, K., Ng, E. L., & Ng, S. F. (2009). The contributions of working memory and executive functioning to problem representation and solution generation in algebraic word problems. *Journal of Educational Psychology,* 101(2), 373-387. DOI:10.1037/a0013843.

Lee, K., Ng, S. F., Ng, E. L., & Lim, Z. Y. (2004). Working Memory and Literacy as Predictors of Performance on Algebraic Word Problems. *Journal of Experimental Child Psychology*, 89, 140-158.

Levin, H. S., Culhane, K. A., Hartmann, J., Evankovich, K., Mattson, A. J., Harward, H., Ringholz, G., Ewing-Cobbs, L. & Fletcher, J.M. (1991). Developmental changes in performance on tests of purported frontal lobe functioning, *Developmental*

Neuropsychology, 7(3), 377-395. DOI: 10.1080/87565649109540499..

Ortiz, T. (2009). *Neurociencia y educación*. Madrid, España: Alianza Editorial.

Packard M. G. & Knowlton, P. J. (2002). Learning and memory functions of the basal ganglia. *Annual Review of Neuroscience*, 25, 563-593.

Passolunghi, M. C. & Pazzaglia, F. (2004). Individual differences in memory updating in relation to arithmetic problem solving. *Learning and Individual Differences*, 14, 219-230.

Polya, G. (1945). *How to Solve It.* Princeton, NJ: Princeton University Press.

Qin, Y., S Carter, C., Silk, E., Andrew Stenger, V., Fissell, K., Goode, A., & R Anderson, J. (2004). The change of the brain activation patterns as children learn algebra equation solving. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 5686–5691.

Radford, L., & Andre, É. (2009). Matemáticas, cognición y cerebro. *Revista Latinoamericana de Investigación En Matemática Educativa*, 12 (2), 215–250.

Rolls, E. T. (1994). Neurophysiology and cognitive functions of the striatum. *Revue Neurologique, 150*(8-9), 648–660

Royer J. M. (ed.). (2003). *Mathematical Cognition*. Portland, OR: Book News, Inc.

Ruiz, M. G., Escotto, E. A., & Sánchez, J. G. (2012). Memoria de Trabajo y Resolución de Problemas matemáticos. *Revista electrónica de psicología* de la *FES Zaragoza-UAM*, 2 (2),43-51.

Sandman C.A., Head, K., Muftuler, L.T., Su, L.,, Buss, C., & Davis, E.P. (2014). Shape of the basal ganglia in preadolescent children is associated with cognitive performance.

Neuroimage, 99, 93-102.

Sowell, E. R., & Jernigan, T. L. (1998). Further MRI evidence of late brain maturation: Limbic volume increases and changing asymmetries during childhood and adolescence. *Developmental Neuropsychology*, *14*(4), 599–617.

Sowell, E. R., Peterson, B. S., Thompson, P. M., Welcome, S. E., Henkenius, A. L., & Toga, A. W. (2003). Mapping cortical change across the human life span. *Nature Neuroscience*, *6*(3), 309–315. DOI:10.1038/nn1008.

Sowell, E. R., Thompson, P. M., Holmes, C. J., Batth, R., Jernigan, T. L., & Toga, A. W. (1999a). Localizing Age-Related Changes in Brain Structure between Childhood and Adolescence Using Statistical Parametric Mapping. *NeuroImage*, *9*(6), 587–597. https://doi.org/10.1006/nimg.1999.0436.

Sowell, E.R., Thompson, P.M., Holmes, C.J., Jernigan, T.L. & Toga, A.W.(1999b). In vivo evidence for post-adolescent brain maturation in frontal and striatal regions, *Nature Neuroscience*, 2, 859-861.

Swanson, H. L. (2006). Cross-Sectional and Incremental Changes in Working Memory and Mathematical Problem Solving. *Journal of Educational Psychology*, 98, 265-281.

Swanson, H. L., Cooney, J. B., & Brock, S. (1993). The influence of working memory and classification ability on children's word problem solution. *Journal of Experimental Child Psychology*, 55, 374-395.

Wierenga, L., Langen, M., Ambrosino, S., van Dijk, S., Oranje, B., & Durston, S. (2014). Typical development of basal ganglia, hippocampus, amygdala and cerebellum from age 7 to 24. *NeuroImage*, *96*, 67–72. DOI: 10.1016/j.neuroimage.2014.03.072.

Wollman, W. (1983). Determining the Sources of Error in a Translation from Sentence to Equation. *Journal for Research in Mathematics Education*, *14*(3), 169–181.

Yu, Y., FitzGerald, T. H., & Friston, K. J. (2013). Working memory and anticipatory set modulate midbrain and putamen activity. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, *33*(35), 14040–14047. DOI:10.1523/JNEUROSCI.1176-13.2013.

Zamarian, L., Ischebeck, A., & Delazer, M. (2009) Neuroscience of learning arithmetic-evidence from brain imaging studies. *Neuroscience & Biobehavioral Reviews*, 33(6), 909-925.

Acknowledgement

This work was supported in part by Universitat Jaume I Grant UJI-A2017-8.

The authors declare that they have no conflict of interest.

Aportació 2.1:

Detecting and visualizing differences in brain structures with SPHARM and functional data analysis

Ferrando, Ventura-Campos et al. (2020a)

NeuroImage 222 (2020) 117209

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Detecting and visualizing differences in brain structures with SPHARM and functional data analysis☆

L. Ferrando[a], N. Ventura-Campos[a,b,*], I. Epifanio[c]

[a] Grup Neuropsicologia i Neuroimatge Funcional, Universitat Jaume I, Spain
[b] Dept. Educació i Didàctiques Específiques. Universitat Jaume I, Spain
[c] Dept. Matemàtiques-IF. Universitat Jaume I, Spain

ARTICLE INFO

Keywords:
Reversal error
Magnetic resonance imaging
Neuroeducation
Functional discriminant analysis
Functional data analysis
Shape analysis

ABSTRACT

A new procedure for classifying brain structures described by SPHARM is presented. We combine a dimension reduction technique (functional principal component analysis or functional independent component analysis) with stepwise variable selection for linear discriminant classification. This procedure is compared with many well-known methods in a novel classification problem in neuroeducation, where the reversal error (a common error in mathematical problem solving) is analyzed by using the left and right putamens of 33 participants. The comparison shows that our proposal not only provides outstanding performance in terms of predictive power, but it is also valuable in terms of interpretation, since it yields a linear discriminant function for 3D structures.

1. Introduction

Nowadays three-dimensional (3D) magnetic resonance imaging (MRI) with high spatial resolution enables the visualization of different brain structures. After their segmentation, anatomical structures of interest, or a region of interest (ROI), are analyzed, since structural abnormality might explain and help to detect certain conditions (Chung et al., 2010). Volumetry is a common marker in many studies, such as those involved in the diagnosis of Alzheimer's disease (Gerardin et al., 2009). However, analysis of a structure's shape can report richer information than volumetry, because ROI-based volumetric measurements do not make explicit if the volume difference occurs over the whole ROI or it is localized within specific zones of the ROI. Gaining insight into morphological changes can provide researchers with a better understanding of the condition (Epifanio and Ventura-Campos, 2014). This is the reason why shape analysis plays an important role in neuroimaging nowadays (Styner et al., 2003). This is also the case in neuroeducation. For example, Sandman et al. (2014) showed that shape analysis may be more sensitive than volumetric analysis when we want to associate brain differences with performance, and they found that deformity of the basal ganglia may be a neurophenotype associated with risk of developmental impairment.

1.1. Motivation

Neuroeducation is another active field of research. There has been growing interest in the support that neuroscience can provide to education. In the specific case of problem solving, some examples are Hanakawa et al. (2003) and Anderson et al. (2012), who observed activation of different areas of the brain while the participants were doing a problem-solving task.

Let us focus on mathematical modeling and the difficulties of translating a practical situation into mathematical notation. Behavioral studies such as Clement (1982), Clement et al. (1981) and Clement et al. (1980) found that most students made mistakes when translating sentences from natural language into algebraic language. When students know the information from the statement, but they are not able to build a correct equation, this is known as reversal error (RE). Clement (1982) showed that the structure of sentences where RE was present was as follows: "Write an equation using the variables S and P to represent the following statement: There are six times as many students as professors at this university. Use S for the number of students and P for the number of professors" (Clement, 1982, p. 17). Most of the wrong answers were $P = 6 \cdot S$, while the correct answer is $S = 6 \cdot P$. Numerous behavioral studies have analyzed this error (Cooper, 1986; González-Calero et al., 2015; Wollman, 1983), but they did not take into account the importance of people's brain development by using MRI.

Ferrando (2019) studied the differences in gray matter (GM) volume that may exist between subjects that make REs versus those who do not. An increase in the volume of the bilateral putamen was found in the group with RE. This follows along similar lines as the study by Qin et al. (2004). Therefore, in this work we will analyze the shape of the left and right putamen in a classification problem. In other words,

☆ The data (putamen surfaces) and code in MATLAB and free software R are available at http://www3.uji.es/~epifanio/RESEARCH/pufda.zip.

* Corresponding author.
E-mail address: venturan@uji.es (N. Ventura-Campos).

https://doi.org/10.1016/j.neuroimage.2020.117209
Received 15 May 2020; Accepted 24 July 2020
Available online 7 August 2020

we will analyze the morphological changes in the left and right putamen between RE and non-RE groups.

1.2. Shape modeling

In neuroimaging studies, shapes have been modeled using different approaches. Some of them are non-parametric, such as medial representation, where the structure is represented by a skeleton (Styner et al., 2003); the distance map approach (Golland et al., 2001); deformation fields (Joshi et al., 1997) and the landmark approach (Park et al., 2008; Shen et al., 2012). However, we opt for a parametric approach: the use of spherical harmonic representation (SPHARM), which has been successfully applied to model several subcortical structures (Chung et al., 2007, 2010; Gerardin et al., 2009; Gerig et al., 2001; Gu et al., 2004; Shen et al., 2004, 2009). Furthermore, previous studies (Styner et al., 2004) have found good concordance of results based on SPHARM and M-rep shape analysis. In SPHARM, we consider the basis functions of spherical harmonics or its weighted version (the weighted spherical harmonic representation), then a set of coefficients weighting the basis functions parametrizes each surface. As a consequence, SPHARM is a way of smoothing functional data.

Functional data analysis (FDA) is the statistical branch that studies functional observations, i.e. when a whole function is a datum. Although typical functional data comprises uni-dimensional functions with only one argument, usually time, in our case we work with trivariate functions with two arguments (angles), which represent spatial locations. Functional data are recorded discretely, but a continuous function lies behind these data. The discrete observations are converted into a true functional form by approximating (smoothing) each function by a weighted sum (a linear combination) of known basis functions. In our case, each surface is initially described by a set of points belonging to the surface, and then it is converted into a functional datum by smoothing it using spherical harmonics.

FDA shares the same objectives as any other branch of statistics. An excellent overview can be found in Ramsay and Silverman (2005), while a non-parametric point of view of FDA is given in Ferraty and Vieu (2006). A recent review of FDA methods can be found in Wang et al. (2016), although it is centered on univariate functional data. As regards applications, Ullah and Finch (2013) review different applications in different fields, and Sørensen et al. (2013) review FDA with medical applications. To the best of our knowledge, in brain imaging studies, FDA has been applied to the analysis of neuroimaging signal, time courses (Lazar, 2008; Tian, 2010; Viviani et al., 2005), but not to the analysis of brain structures, with some exceptions (Epifanio and Ventura-Campos, 2014; Lila et al., 2016). Note that although SPHARM has been used extensively in neuroscience literature, the majority of these works have not been used in the context of FDA and, therefore, functional data techniques have not been exploited in this field.

1.3. Our contributions

We proposed to use FDA for analyzing the shape of 3D brain structures for the first time in Epifanio and Ventura-Campos (2014). In Epifanio and Ventura-Campos (2014) the hippocampus surfaces, for the study of Alzheimer's disease, were described by multivariate (three) functions with two arguments. We extended principal component analysis (PCA) to deal with trivariate functional data with two arguments. Functional independent component analysis (FICA) was also discussed in Epifanio and Ventura-Campos (2014).

Here, as in Epifanio and Ventura-Campos (2014), we deal with a classification problem. The novelty of our contribution is twofold. On the one hand, in this work our proposal will improve on the method presented in Epifanio and Ventura-Campos (2014). In fact, it will improve on the results of other well-known methodologies, both from the numerical performance point of view and from the interpretative and visual point of view. On the other hand, the new methodology will be applied to the study of RE using the putamen, which is a novel classification problem in neuroeducation.

High dimensionality is one of the greatest difficulties in this kind of classification problems. Unless we consider simple features such as volumetry, the number of variables used to describe the anatomical shape of brain structures is always much larger than the number of subjects (observations). This is the case when using SPHARM coefficients as features in the classification problem. To overcome this problem, we can consider two approaches. In the first approach, one can either select or extract a small subset of relevant features, as in Gerardin et al. (2009), where the most discriminative features are selected using a bagging strategy for subsequent classification with a support vector machine (SVM) classifier, or as in Clemmensen et al. (2011), where sparse discriminant analysis is proposed by extending linear discriminant analysis (LDA) to the high-dimensional setting. Instead of selecting features, in the second approach, one can reduce the dimension, i.e. one can use all the features to construct new components which summarize the original variables. This is the case of Boulesteix (2004) or Epifanio and Ventura-Campos (2014). The method in Boulesteix (2004) consists of Partial Least Squares (PLS) dimension reduction and linear discriminant analysis applied to the PLS components. Although it was originally developed in the context of classification with high-dimensional microarray data, it can be used for classification of any high-dimensional data. In Epifanio and Ventura-Campos (2014), linear discriminant analysis is applied to the new components obtained by dimension reduction techniques for functional data.

Our proposal consists of combining both approaches. The idea is simple, but effective. First we reduce the dimension by using trivariate functional principal component analysis (FPCA) with two arguments or other techniques, such as FICA, as described in Epifanio and Ventura-Campos (2014), but we improve on this and go one step further. We then select the most discriminative components (principal components, PCs or independent components depending on whether we use FPCA or FICA) by stepwise variable selection (SVS) (Weihs et al., 2005) for LDA classification. Note that although the dimension has been reduced in the first step, the number of predictor variables may continue to be high, because the number of subjects is usually small in this kind the classification problems. Note that in the classical multivariate case we can use LDA after PCA, assuming the covariance matrix is the same for all groups (Jolliffe, 2002). Nevertheless, the separation between groups does not necessarily have to occur in the PCs with highest variance, but can occur in the last PCs, those with low variance (Epifanio and Ventura-Campos, 2011).

On the other hand, besides the method's predictive accuracy, human interpretability is one of the most valuable characteristics in a classifier. So the point is to propose methods that are interpretable, rather than trying to explain black box machine learning models (Rudin, 2019). Our proposal is based on the application of LDA in the final step. Despite the simplicity of LDA, its performance is typically almost as good as that of more complicated methods (Hand, 2006; Pierola et al., 2016). Furthermore, LDA provides low-dimensional projections of the data onto the most discriminative directions. We have taken advantage of this and we have defined a discriminative function for 3D shapes. From the interpretation point of view, this is very useful because it reveals the exact locations and directions of the main differences between the groups. In other words, we can visualize how the brain structure changes between the groups. Note that our proposed visual representations are not significance maps, which are common visualization tools in neuroscience. In summary, we propose a local shape methodology that can spatially localize shape changes.

The outline of the paper is as follows. Section 2 describes our data and the proposed methodology. Section 3 includes the numeric results obtained by applying this methodology to our database, together with a comparison with other well-known methods. This section also includes the visualization and graphical interpretation of the results using our procedure. In Section 4 conclusions are given.

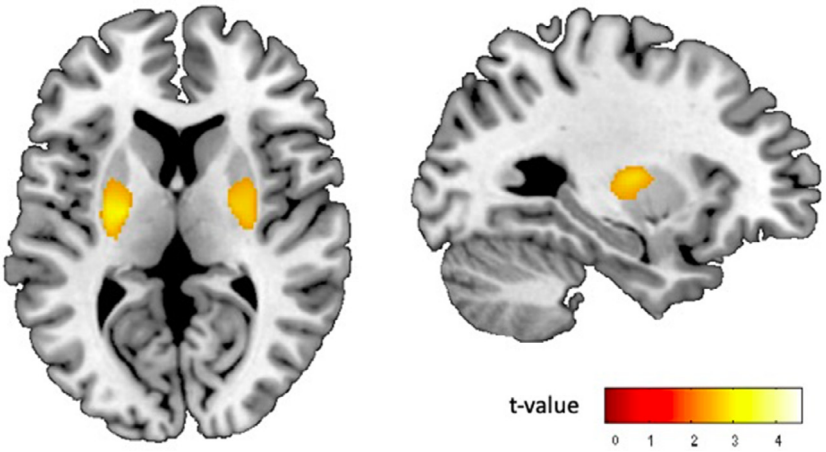


Fig. 1. Neuroimaging results of the two-sample *t*-test performed between groups. It represents the contrast: RE group vs. non-RE group of our study ($p < 0.05$ FWE cluster-corrected using a threshold of $p < 0.005$ at the uncorrected voxel level and a cluster size higher or equal to 813 voxels). The MNI coordinates of the left putamen were x=-23, y=-12, z=0 with Z-value=3.96 (k=938), and the MNI coordinates of the right putamen were x=26, y= -8, z=3 with Z-value= 3.41 (k=813).

2. Materials and methods

The data (putamen surfaces) (Ferrando et al., 2020) and code in MATLAB and free software R (R Core Team, 2020) are available at http://www3.uji.es/~epifanio/RESEARCH/pufda.zip for reproducibility purposes.

2.1. Processing of structural magnetic resonance imaging (smri) brain scans

Thirty-three participants (20 females) with ages ranging between 18 and 26 years (mean age: 22.03, SD: 2.36) were analyzed. All participants were students of Universitat Jaume I. Before participating, they signed a written consent form. All experimental procedures followed the guidelines of the research ethics committee at Universitat Jaume I. The exclusion criteria were trauma with loss of consciousness for more than one hour, typical contraindications to MRI, such as metal implants, and the presence of medical or neurological illness.

Apart from having an MRI scan, the participants carried out a behavioral task. The subjects answered 16 questions, where a mathematical equation had to be built for each of the statements presented. We used an application similar to González-Calero et al. (2015). See details in Ferrando (2019). We established two groups, those who committed REs vs. non-RE, according to the number of errors they made. One group comprised those who failed more than 40% of the equations (RE group: 15 subjects, 4 males, mean age: 21.466, SD: 2.1), and the second group was formed by those who answered 100% of the questions correctly (non-RE group: 18 subjects, 9 male, mean age: 22.5; SD: 2.53).

The sMRI scans were acquired using two scanners. A 3 Tesla Philips scanner and 1.5 Tesla Siemens Symphony scanner (Erlangen, Germany). High-resolution T1-weighted, TR = 8.4 ms, TE = 3.8 ms, matrix size = $320 \times 320 \times 250$ and voxel size = $0.75 \times 0.5 \times 0.8$ mm was used with the Philips scanner. However, high-resolution T1-weighted, TR = 2200 ms, TE = 3 ms, flip angle = 90°, matrix size = $256 \times 256 \times 160$ and voxel size = $1 \times 1 \times 1$mm was used with the Siemens Symphony scanner. The scanner acquisitions covered the entire brain and were performed in parallel to the anterior commissure-posterior commissure plane (AC-PC).

The Voxel Based Morphometry (VBM) analysis was conducted with SPM12 (SPM12 (v7219), Wellcome Trust Centre for Neuroimaging, London, UK, http://www.fil.ion.ucl.ac.uk/spm/software/spm12). We performed the preprocessing steps using the CAT12 toolbox with the default setting (CAT12.5, http://dbm.neuro.uni-jena.de/cat/). We used the GM images to identify the differences in volume between the groups. Having segmented, modulated, and normalized these images, they were smoothed using 8-mm full-width-half-maximum Gaussian smoothing and then fed into a two-sample *t*-test analysis in SPM12. The VBM results showed a greater bilateral posterior putamen volume in the group with RE in comparison to the non-RE group (see Fig. 1). Following these results, we consider the left and right putamen as the Region of Interests (ROIs) to our classification study, since they were the most significant structures. So, the next step was to extract the left and right putamen for each participant. To obtain these ROIs, we segmented each putamen by using the *imcalc* toolbox of SPM12 and performing an intersection between the GM image of each participant and the ROI of the putamen of the AAL atlas. Finally, the slices of each putamen were put together using the *isosurface* function in MATLAB, which returns the faces and vertices of the triangle mesh.

2.2. Surface parametrization

We need three functions with two angular parameters to represent the putamen surfaces: $x(\theta, \phi)$, $y(\theta, \phi)$, $z(\theta, \phi)$ (more details in Shen et al. (2009)). In fact, the surface of these closed 3D objects is mapped onto a unit sphere under a one-to-one. There are several well-known surface flattening techniques that provide this bijective mapping: area preserving mapping (Brechbühler et al., 1995; Shen et al., 2004), conformal mapping (Gu et al., 2004), the deformable surface algorithm (Macdonald et al., 2000) or semi-isometric mapping (Timsari and Leahy, 2000), for instance. Nevertheless, the implementation of these flattening methods is not trivial and, especially, their computational intensity leads us to use the method proposed by Chung et al. (2010) as a better alternative for objects that are almost convex or star-shaped. This method considers the equilibrium state of heat diffusion by tracing the geodesic path of heat equilibrium state from a heat source (the putamen) to a heat sink (the unit sphere). This flattening technique is numerically simpler than the previous methods, since the solution of an isotropic heat equation in a 3D image is computationally trivial and it does not require either to optimize a target function (for more details see Chung et al. (2010)). In any case, any flattening method could be used without altering the subsequent analysis.

Once the surface is projected into the sphere, the angles will act as coordinates for the surfaces of the putamen. Fig. 2 shows an illustration of the surface flattening process for a left putamen with this procedure and the surface parametrization using the angles (θ, φ). The north pole of a unitary sphere coincides with the point θ =0.

2.3. Representing putamens using SPHARM

Although each putamen could be described by a set of points or discrete observations of its surface, each putamen is truly a smooth surface, a function whose two arguments are angles. Smoothing allows us to convert the discrete data to functions and to perform FDA. This smoothing is carried out using a basis system, SPHARM, and considering the coefficients of each putamen in this basis function expansion. As said before, we have chosen SPHARM because it has previously been applied

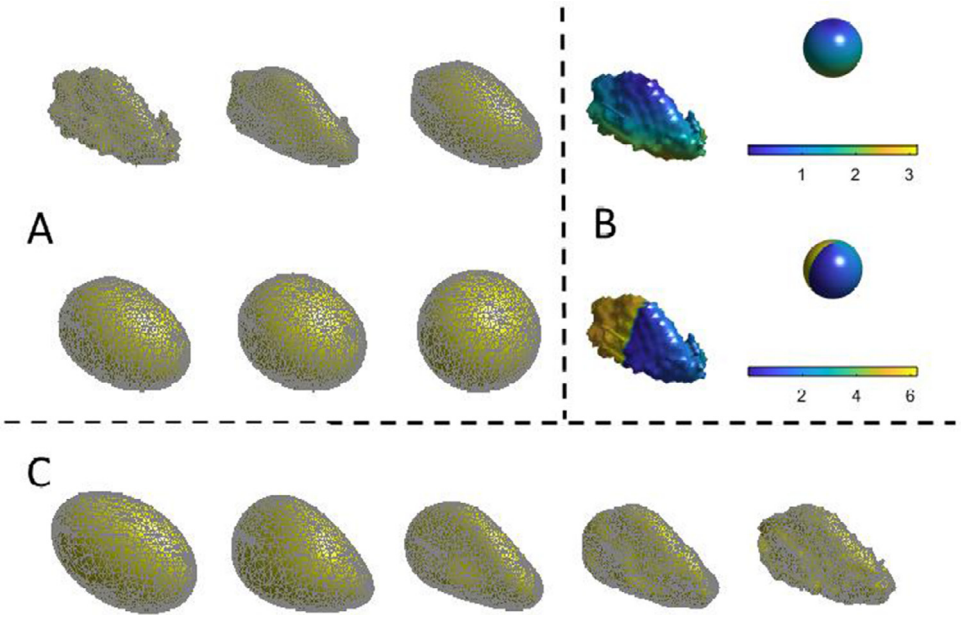


Fig. 2. A left putamen example: (A): The surface flattening process from the original putamen (top left) to the sphere (bottom right). The same level sets as in (Chung et al., 2010) are used (1.0, 0.6, 0.2, -0.2, -0.6, -1.0). (B): The spherical angles are projected on the putamen surface and the unitary sphere, for θ (first row) and φ (second row). (C): The putamen representation using a different number of spherical harmonics: $L = 1, 2, 5, 11, 30$ (from left to right). This will be explained below. The head of the putamen is on the left, while the tail is on the right. This orientation is considered the reference throughout the paper.

successfully to model several subcortical structures. Furthermore, its orthogonality will simplify calculations. Nevertheless, other less common bases, such as spherical splines (Alfeld et al., 1996; He et al., 2005), the weighted Fourier series (Chung et al., 2007) or spherical wavelets (Nain et al., 2007; Yu et al., 2007), could also be used.

We have used the real spherical harmonics as in Chung et al. (2007, 2010), although we could have used spherical harmonics of complex value as in Gerig et al. (2001) or Shen et al. (2004). We prefer to set up a real-valued model because we only need real-valued spherical functions in our application, as in most applications.

A real basis of spherical harmonics, where l is the degree and m is the order, is given by:

$$Y_{lm}(\theta,\varphi) = \begin{cases} \sqrt{2}N_{(l,m)}\cos(m\varphi)P_l^m(cos\theta) & \text{if } m > 0 \\ N_{(l,0)}P_l^0(cos\theta) & \text{if } m = 0 \\ \sqrt{2}N_{(l,|m|)}\sin(|m|\varphi)P_l^{|m|}(cos\theta) & \text{if } m < 0 \end{cases} \quad (1)$$

where P_l^m is the associated Legendre polynomial of order m defined over the range $[-1, 1]$:

$$P_l^m(x) = \frac{(-1)^m}{2^l l!}(1-x^2)^{m/2}\frac{d^{l+m}}{dx^{l+m}}(x^2-1)^l$$

and $N_{(l,m)} = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}$.

Let S^2 be the unit sphere in $\mathbb{R}^3$, and f and $g \in L^2(S^2)$. The inner product is given by:

$$< f, g >= \int_{\theta=0}^{\pi}\int_{\varphi=0}^{2\pi} f(\theta,\varphi)g(\theta,\varphi)d\Omega = \int_{S^2} f(\theta,\varphi)g(\theta,\varphi)d\Omega = \int_{S^2} f g d\Omega \quad (2)$$

where $d\Omega = \sin(\theta)d\varphi d\theta$. The spherical harmonics satisfy the orthonormal condition with respect to the inner product:

$$\int_{S^2} Y_{lm}Y_{l'm'}d\Omega = \delta_{ll'}\delta_{mm'}$$

where δ_{ij} is the Kronecker delta.

Three functions represent each putamen in terms of the spherical harmonics, where L determines the smoothing degree:

- $x(\theta,\varphi) = \sum_{l=0}^{L}\sum_{m=-l}^{l} c_{lm}^x Y_{lm}(\theta,\varphi)$
- $y(\theta,\varphi) = \sum_{l=0}^{L}\sum_{m=-l}^{l} c_{lm}^y Y_{lm}(\theta,\varphi)$
- $z(\theta,\varphi) = \sum_{l=0}^{L}\sum_{m=-l}^{l} c_{lm}^z Y_{lm}(\theta,\varphi)$

This can be expressed as a vector-valued function:

$$F(\theta,\varphi) = (x(\theta,\varphi), y(\theta,\varphi), z(\theta,\varphi))' = \sum_{l=0}^{L}\sum_{m=-l}^{l} \mathbf{c}_{lm}Y_{lm}(\theta,\varphi) \quad (3)$$

where $\mathbf{c}_{lm} = (c_{lm}^x, c_{lm}^y, c_{lm}^z)'$. The coefficients can be estimated by least squares, since we know the values of each function in a sample of points, $\{(\theta_i,\varphi_i)\}_{i=1}^n$. For $x(\theta,\varphi)$ (and similarly for $y(\theta,\varphi)$ and $z(\theta,\varphi)$), $\mathbf{x} = \{x(\theta_i,\varphi_i)\}_{i=1}^n$ is the vector of observations, $\mathbf{Y} = \{Y_{lm}(\theta_i,\varphi_i)\}_{i=1}^n$ is the matrix of basis function values at the observation points and $\mathbf{c}^x$ is the vector with the coefficients c_{lm}^x, which can be least square estimated by $\mathbf{c}^x = (\mathbf{Y}'\mathbf{Y})^{-1}\mathbf{Y}'\mathbf{x}$ or by the iterative residual fitting algorithm (Chung et al., 2007).

Fig. 2 (C) shows a putamen example represented by SPHARM using different L values. For small L values, the surface is highly smoothed and many details are missing, but for high L values, such as $L = 30$, the surface is quite noisy, since noise is also fitted. The value $L = 11$ represents a trade-off between both positions, and it has been chosen for left putamens following the strategy suggested by Ramsay and Silverman (2005, Section 4.6.2) to determine the number of basis. This strategy consists of computing the unbiased estimate of the residual variance and selecting the number of basis that makes this variance decrease substantially (Millán-Roures et al., 2018). Specifically, we compute the residuals as the square Euclidean norm of the difference between the points in smoothed surfaces and the original surface for every spherical mesh vertex. These values are added for all the vertices and divided by the number of vertices minus the number of basis, which is $(L+1)^2$ for the degree value L). Then, we compute the pooled variance for all the individuals for L values from 2 to 23 and the minimum variance estimate was attained at $L = 11$ for left putamens and $L = 12$ for right putamens.

No alignment is necessary, since each putamen was translated to the same point in such a way that its centroid coincided with that point, i.e. location was removed previously. No rotation is needed, since all the putamens had the same orientation. Scaling is not needed either, i.e. we want to keep size since the volume is used as discriminant feature in many classification problems. If size had to be removed, then we could divide each putamen by the size of the centroid at the beginning, as in Epifanio and Ventura-Campos (2011). Note that no registration is necessary, as happened in Chung et al. (2010), because the coordinates $((\theta, \varphi))$ on the surfaces are corresponding pairs, therefore the coefficients match each other.

For this and the previous Section 2.2, the following packages are very helpful: the SurfStat package (http://www.math.mcgill.ca/keith/surfstat) and its extension (http://www.stat.wisc.edu/~mchung/research/amygdala/) (Chung et al., 2010).

2.4. Functional principal component analysis

Before introducing FPCA, let us remember how PCA works in the standard multivariate case. Let $\mathbf{X}$ be the centered data (the mean has been subtracted) matrix with N rows. N indicates the number of individuals. Let $\mathbf{V}$ be the sample variance-covariance matrix, $\mathbf{V} = (N-1)^{-1}\mathbf{X}'\mathbf{X}$, where $\mathbf{X}'$ indicates the transposition of $\mathbf{X}$. The solution of the following eigenequation, where ρ is an eigenvalue and ξ is an eigenvector of $\mathbf{V}$, provides the PCA solution:

$$\mathbf{V}\boldsymbol{\xi} = \rho\boldsymbol{\xi}, \tag{4}$$

The PC scores for the k-th PC are computed by $s_i^k = \sum_j x_{ij}\xi_j^k$.

In FPCA, PCs are not vectors, but functions, and summations change into integrations. Let us begin by recalling FPCA for the functional univariate case with one scalar argument t. Let $\{x_1(t), \ldots, x_N(t)\}$ be the set of observed functions. The mean function is defined by $\bar{x}(t) = N^{-1}\sum_{i=1}^N x_i(t)$, while the variance-covariance function $v(s,t)$ is defined by $v(s,t) = (N-1)^{-1}\sum_{i=1}^N x_i(s)x_i(t)$, once the data have been centered. The functional counterpart of Eq. 4 (see details in Ramsay and Silverman (2005, Chapter 8)) is:

$$\int v(s,t)\xi(t)dt = \rho\xi(s), \tag{5}$$

where ρ is still an eigenvalue and $\xi(s)$ is not an eigenvector, but an eigenfunction. The score for the kth PC for the ith subject is now calculated by using the inner product for functions: $s_{ik} = \int x_i(s)\xi_k(s)ds$.

To solve 5 there are different alternatives (see Ramsay and Silverman (2005, Sec. 8.4.2) for a review). One of them consists of considering the coefficients in a basis functions. In fact, if the basis is orthonormal, FPCA reduces to the classical multivariate PCA of the coefficient matrix, which reduces the computational cost. The functions $\xi_k(t)$ satisfy the orthonormality constraint, as in the multivariate case.

The maximum number of possible functional PCs is limited by $N-1$, although if the number of basis functions M is less than N, then the maximum would be M. Let K be the number of functional PCs considered, then $x_i(t)$ is described by $\sum_{k=1}^K s_{ik}\xi_k(t)$.

2.4.1. FPCA With multiple functions and multiple arguments

Let $\{F_i(\theta,\varphi)\}_{i=1}^N = ((x_i(\theta,\varphi), y_i(\theta,\varphi), z_i(\theta,\varphi))$ be the set of vector-valued functions with two arguments. Each of them represents the putamen of a subject. As previously, we can calculate pointwisely three mean functions, $\bar{x}(\theta,\varphi)$, $\bar{y}(\theta,\varphi)$ and $\bar{z}(\theta,\varphi)$, three covariance functions $v_{XX}((\theta,\varphi),(\vartheta,\phi))$, $v_{YY}((\theta,\varphi),(\vartheta,\phi))$, $v_{ZZ}((\theta,\varphi),(\vartheta,\phi))$, and cross-covariance functions. For example, the cross-covariance function of the centered data for the combination XY is $v_{XY}((\vartheta,\phi),(\theta,\varphi)) = (N-1)^{-1}\sum_{i=1}^N x_i(\vartheta,\phi)y_i(\theta,\varphi)$, it can be computed analogously for the combination XZ and YZ.

The addition of the inner products of the three components (as defined in 2) yields an inner product on the space of vector-valued functions:

$$< F_1, F_2 >=< x_1, x_2 > + < y_1, y_2 > + < z_1, z_2 > . \tag{6}$$

The PC score for the k-th PC is calculated by $s_i^k =< F_i, \xi^k >= \int_{S^2} x_i\xi_X^k d\Omega + \int_{S^2} y_i\xi_Y^k d\Omega + \int_{S^2} z_i\xi_Z^k d\Omega$, where PCs are now a three-vector $\xi=(\xi_X, \xi_Y, \xi_Z)$ of functions, which are solutions of the following eigenequation system $V\xi = \rho\xi$, which is expressed as

$$\begin{aligned}
&\int_{S^2} v_{XX}((\vartheta,\phi),(\theta,\varphi))\xi_X(\theta,\varphi)d\Omega + \int_{S^2} v_{XY}((\vartheta,\phi),(\theta,\varphi))\xi_Y(\theta,\varphi)d\Omega \\
&\quad + \int_{S^2} v_{XZ}((\vartheta,\phi),(\theta,\varphi))\xi_Z(\theta,\varphi)d\Omega = \rho\xi_X(\vartheta,\phi) \\
&\int_{S^2} v_{YX}((\vartheta,\phi),(\theta,\varphi))\xi_X(\theta,\varphi)d\Omega + \int_{S^2} v_{YY}((\vartheta,\phi),(\theta,\varphi))\xi_Y(\theta,\varphi)d\Omega \\
&\quad + \int_{S^2} v_{YZ}((\vartheta,\phi),(\theta,\varphi))\xi_Z(\theta,\varphi)d\Omega = \rho\xi_Y(\vartheta,\phi) \\
&\int_{S^2} v_{ZX}((\vartheta,\phi),(\theta,\varphi))\xi_X(\theta,\varphi)d\Omega + \int_{S^2} v_{ZY}((\vartheta,\phi),(\theta,\varphi))\xi_Y(\theta,\varphi)d\Omega \\
&\quad + \int_{S^2} v_{ZZ}((\vartheta,\phi),(\theta,\varphi))\xi_Z(\theta,\varphi)d\Omega = \rho\xi_Z(\vartheta,\phi).
\end{aligned} \tag{7}$$

As mentioned before, we consider the basis function expansion of the vector-valued functions to solve the eigenequation system. Each F_i is described by the following vector of basis coefficients $\mathbf{c}^i = (\{c_{ilm}^x\}, \{c_{ilm}^y\}, \{c_{ilm}^z\})$, with $l = 0, \ldots, L$ and $m = -l$ to l, and a matrix $\mathbf{C}$ with N rows (one per subject) is built by stacking those vectors. We only need to compute the PCA of the $N \times 3M$ matrix $\mathbf{C}$ since spherical harmonics are orthonormal (M is 144 for $L = 11$, whereas M is 169 for $L = 12$). Once PCs are calculated, the parts corresponding to each coordinate are separated (see Ramsay and Silverman (2005, Sec. 8.5.1) for details in the case of bivariate FPCA with one argument).

As in the multivariate case, the proportion of variance explained by each eigenfunction was given by each eigenvalue ρ divided by the sum of all eigenvalues. In addition, for the j-th principal component $\xi^j = (\xi_X^j, \xi_Y^j, \xi_Z^j)$, we can compute the variation accounted for each coordinate by $< \xi_X^j, \xi_X^j >$, $< \xi_Y^j, \xi_Y^j >$ and $< \xi_Z^j, \xi_Z^j >$ respectively, since their sum is one by definition.

2.5. Functional linear discriminant analysis (FLDA)

In order to obtain meaningful results of LDA with functions, some kind of regularization or filtering is necessary (naively, we could apply the linear discriminant method to the high-dimensional vectors, but this approach does not give meaningful results, see (Ramsay and Silverman, 2002, ch.8) for details about the explanation for this). A common regularization approach consists of carrying out LDA on the first PCs (or other types of dimension reduction techniques), i.e. all the PCs up to a certain number are considered in LDA. This idea has been used in the functional univariate case with one argument (Epifanio, 2008; Hall et al., 2001), with multivariate functions with one argument (Ramsay and Silverman, 2002, ch.8), and with two arguments (Epifanio and Ventura-Campos, 2014).

However, as explained in Section 1.3, this dimension reduction may not be enough, since we are dealing with small sizes. Furthermore, PCA is an unsupervised statistical learning technique, and its application does not ensure that the separation between classes occurs in the first few components, but it can occur in the last PCs (Jolliffe, 2002). On the one hand, if these last PCs were not considered, the accuracy of classification would be affected. On the other hand, if we consider a very high number of components, i.e. a very high number of predictors with few observations, we return to a high-dimensionality problem again. This is the reason why we propose to consider only the most discriminative PCs by SVS for LDA classification.

2.6. Stepwise variable selection for classification

There are many methods for variable section for classification. We consider a stepwise forward variable selection based on Wilks' lambda criterion. The method is implemented in the *greedy.wilks* function of the R package klaR (Weihs et al., 2005). The variable which separates the groups most constitutes the initial model. Then, more variables are added to the model depending on Wilks' lambda criterion: we add the variable that minimizes Wilks' lambda of the model, including the variable if its p-value still shows statistical significance.

Another variable selection method could also be considered, such as the method implemented in the *stepclass* function of the R package klaR (Weihs et al., 2005). However, the selection of variables in this kind of method is based on optimizing a performance measure, such as accuracy, which is estimated by cross-validation (CV). As the sample size is very small in our application, the results of this kind of method are unstable due to the cross-validation step. In other words, the variables selected can change a lot depending on how the data are split. This is why we opt for a deterministic method like *greedy.wilks*.

In summary, our proposal, which is referred to as FPCA-SVS-LDA, consists of SPHARM representing putamen, applying FPCA, selecting the scores of the PCs by SVS using the *greedy.wilks* method and carrying out LDA on this selection.

2.6.1. Linear discriminant function

As in the multivariate case (Mardia et al., 1979, Sect. 11.5), we can define the linear discriminant function in the functional case. The linear discriminant vector function $\lambda^j(\theta, \varphi) = (\lambda_X^j(\theta,\varphi), \lambda_Y^j(\theta,\varphi), \lambda_Z^j(\theta,\varphi))$ is the functional counterpart of the linear discriminant vector or canonical variate in the multivariate case. Therefore, the score or discriminant value of F_i can be obtained by $d_i^j = < F_i, \lambda^j >= \int_{S^2} x_i \lambda_X^j d\Omega + \int_{S^2} y_i \lambda_Y^j d\Omega + \int_{S^2} z_i \lambda_Z^j d\Omega$.

Let us express both functions in the orthonormal base defined by the PCs, then $d_i^j = < F_i, \lambda^j > = < \sum_{k=1}^K s_i^k \xi^k, \sum_{k=1}^K l_k^j \xi^k >$. Due to the orthonormality, $d_i^j = < \mathbf{s}_i, \mathbf{l}^j >$, i.e. the vector $\mathbf{l}^j$ is the j-th canonical variate for the $N \times K$ matrix $\mathbf{S}$ with the scores for the N individuals, where each row in $\mathbf{S}$ is formed by $\mathbf{s}_i$, which has the K scores for the individual i.

In summary, if there are Q groups, each of them with size N_i ($\sum_{i=1}^Q N_i = N$), LDA is applied to $N \times K$ matrix $\mathbf{S}$ of PC scores. The $K \times r$ matrix $\mathbf{L}$, where $r = \min\{K, Q-1\}$ is the number of discriminant functions, contains the linear discriminant vectors $\mathbf{l}^j$, while the $N \times r$ matrix $\mathbf{D} = \mathbf{SL}$ contains the discriminant values, and the linear discriminant function $\lambda^j(\theta, \varphi)$ ($j = 1, \ldots, r$) is $\sum_{k=1}^K l_k^j \xi^k$, where l_k^j is the element (k, j) of the matrix $\mathbf{L}$.

2.7. Visualization of the results

The effect of each functional PC (FPC) or linear discriminant function can be displayed by adding a suitable multiple, which can be positive or negative, of that function to the mean function (mean putamen). This approach is common in the literature on shape analysis (Dryden and Mardia, 1998) and FDA (Ramsay and Silverman, 2002). We can plot a vector map, where vectors are drawn from the mean putamen to the surface formed by the mean plus the multiple of the function in question. Or we can also color the mean putamen using the magnitude (norm) of those vectors. Furthermore, the PCA scores and the discriminant values can be also displayed.

2.8. Functional independent component analysis

Let us introduce the methodology when the FPCA step is exchanged for another dimension reduction technique: FICA.

Let us remember ICA for the multivariate case. The data matrix $\mathbf{X}$ is expressed as a linear combination of non-Gaussian (independent) components: $\mathbf{X} = \mathbf{SA}$, where columns of $\mathbf{S}$ contain the independent components and $\mathbf{A}$ is a linear mixing matrix. ICA seeks to "un-mix" the data by estimating an un-mixing matrix $\mathbf{W}$ such that $\mathbf{XW} = \mathbf{S}$. Under this assumption, the "signals" in $\mathbf{X}$ will be "more Gaussian" than the source components in $\mathbf{S}$ due to the Central Limit Theorem. Therefore, the objective is to find an un-mixing matrix $\mathbf{W}$ that maximizes the non-gaussianity of the sources.

For the functional univariate case, let $x_1(t), \ldots, x_N(t)$ be N linear mixtures of K independent components $s_j(t)$: $x_i(t) = \sum_{j=1}^K a_{ij} s_j(t)$, for all i. Assume the following basis expression for each function: $x_i(t) = \sum_{m=1}^M b_{im} G_m(t)$.

Let x be a vector-valued function with components $x_1, \ldots, x_N$ and G the vector-valued function with components $G_1, \ldots, G_M$, then the simultaneous expansion of all N functions can be expressed by $x = \mathbf{B}G$, where $\mathbf{B}$ is the coefficient matrix, with size $N \times M$. ICA can be performed on $\mathbf{B}'$, thus $\mathbf{B}' = \mathbf{S_b A_b}$ and $x = \mathbf{B}G = \mathbf{A_b' S_b'} G$. In other words, the data x are generated by a process of mixing the K components $I = \mathbf{S_b'} G$ (the independent components are the rows of $\mathbf{S_b'}$).

For any function $\tilde{x}(t)$ that is not contained in the initial data set, its expansion in terms of those ICA components is $\tilde{x}(t) = \sum_{j=1}^K \tilde{a}_j I_j(t)$, where $I_j(t)$ is the j-th component of I. If I and G are estimated in p points ($\{t_k; k = 1, \ldots, p\}$), the $p \times K$ matrix $\mathbf{I}$ and the $p \times M$ matrix $\mathbf{G}$ can be defined, as well as $\mathbf{I} = \mathbf{GS_b}$. Then, we can compute the K-vector $\tilde{\mathbf{a}}$ with the coefficients $\tilde{a}_j$ by least squares fitting (Ramsay and Silverman, 2005): $\tilde{\mathbf{a}} = (\mathbf{I'I})^{-1}\mathbf{I'}\tilde{\mathbf{x}}$, with $\tilde{\mathbf{x}} = \{\tilde{x}(t_k)\}_{k=1}^p$, i. e. $\tilde{\mathbf{a}} = (\mathbf{S_b' G' G S_b})^{-1}\mathbf{S_b' G'}\tilde{\mathbf{x}}$. Similarly, we can estimate the M-vector with the coefficients $\tilde{b}_m$ by $\tilde{\mathbf{b}} = (\mathbf{G'G})^{-1}\mathbf{G'}\tilde{\mathbf{x}}$, where G is the basis and $\tilde{x}(t) = \sum_{m=1}^M \tilde{b}_m G_m(t)$. In the case where the basis G is orthonormal, i.e. $\mathbf{G'G}$ is the identity matrix, then

$$\tilde{\mathbf{a}} = (\mathbf{S_b' G' G S_b})^{-1}\mathbf{S_b' G'}\tilde{\mathbf{x}} = (\mathbf{S_b' S_b})^{-1}\mathbf{S_b'}\tilde{\mathbf{b}}. \tag{8}$$

The same discussion can be adopted when the functions have more than one argument. In the case of multivariate functional data, the coefficients for each function can be concatenated into a single long vector, as done in Section 2.4.1 with multivariate FPCA. In this case, $\tilde{\mathbf{b}}$ would be $(\mathbf{c}^i)'$.

In order to reduce noise and prevent overlearning (Hyvärinen et al., 2001, Section 13.2), data dimension reduction by PCA should be carried out prior to the application of the ICA algorithm (see Hyvärinen et al. (2000, Section 5) for details). Therefore, PCA is computed first with a concrete number of components, then the same number of independent components as the PCA reduced dimension are estimated.

As previously, let Q be the number of groups, with size N_i ($\sum_{i=1}^Q N_i = N$). For computing the linear discriminant vector function $\lambda^j(\theta, \varphi)$, LDA is applied to the $K \times N$ matrix $\mathbf{A}$ with the coefficients of the K ICA components. Then, we obtain a $K \times r$ matrix $\mathbf{L}$ ($r = \min\{K, Q-1\}$ is the number of discriminant functions), which yields the $r \times N$ matrix $\mathbf{D}$ of discriminant values ($\mathbf{D} = \mathbf{L'A}$). According to Eq. 8, $\mathbf{A} = (\mathbf{S_b' S_b})^{-1}\mathbf{S_b' C'}$, where $\mathbf{S_b}$ is the $3M \times K$ matrix that contains the independent components of $\mathbf{C'}$, the $N \times 3M$ matrix with the SPHARM coefficients. As $\mathbf{D} = \mathbf{\Lambda C'}$, where $\mathbf{\Lambda}$ is the $r \times 3M$ matrix with the SPHARM coefficients of the r functions $\lambda^j(\theta, \varphi)$ ($j = 1, \ldots, r$) (λ^j is the j-th row), then $\mathbf{\Lambda} = \mathbf{L'}(\mathbf{S_b' S_b})^{-1}\mathbf{S_b'}$.

In summary, when FPCA is exchanged for FICA in our methodology, while keeping the remaining steps the same, we call this procedure FICA-SVS-LDA.

3. Results and discussion

FPCA-SVS-LDA and FICA-SVS-LDA are applied to our data set. Due to the small sample size of our data set, the findings cannot be conclusive from the neuroeducational point of view. However, the proposed methodology can be applied without modification to a larger data set. The first two FPCs account for 46% of the variance for the left putamens (62% for the right putamens, respectively), while the first 19 FPCs explain 95% of the variation (the first 18 FPCs for the right putamens, respectively). Note that the 1st FPC is not selected by our procedure until step 17 for the left putamens, which demonstrates again that the first PCs are not necessarily the most discriminative ones, as explained in Sect. 1.3. The most discriminating FPCs for the left putamens are the 20th and 11th. Fig. 3 displays the scores for these two components. Blue stars and red circles represent the RE and non-RE subjects, respectively. The separation between both groups can be perceived visually. Fig. 3 also displays the discriminant values for the left putamens. As we only have two classes, only one discriminant function can be defined. In this case, the whole training set is classified correctly.

Fig. 4 shows the effect of the first three FPCs on the mean left putamen by representing their magnitudes with two standard deviations.

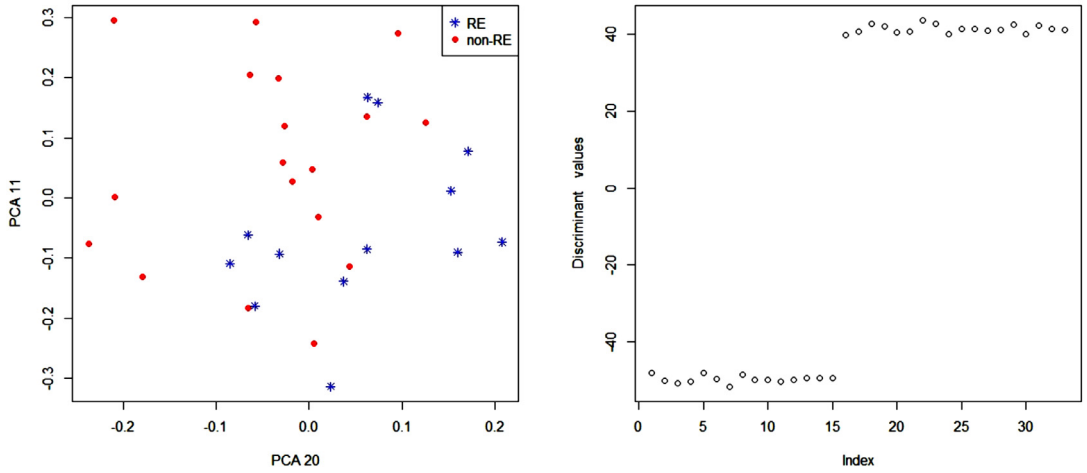


Fig. 3. Left side: Score plot of 20th and 11th PCs for the left putamens. The legend indicates the groups. Right side: Index versus Discriminant values. Note that the first 15 subjects correspond to the RE group, while the remaining individuals belong to the non-RE group.

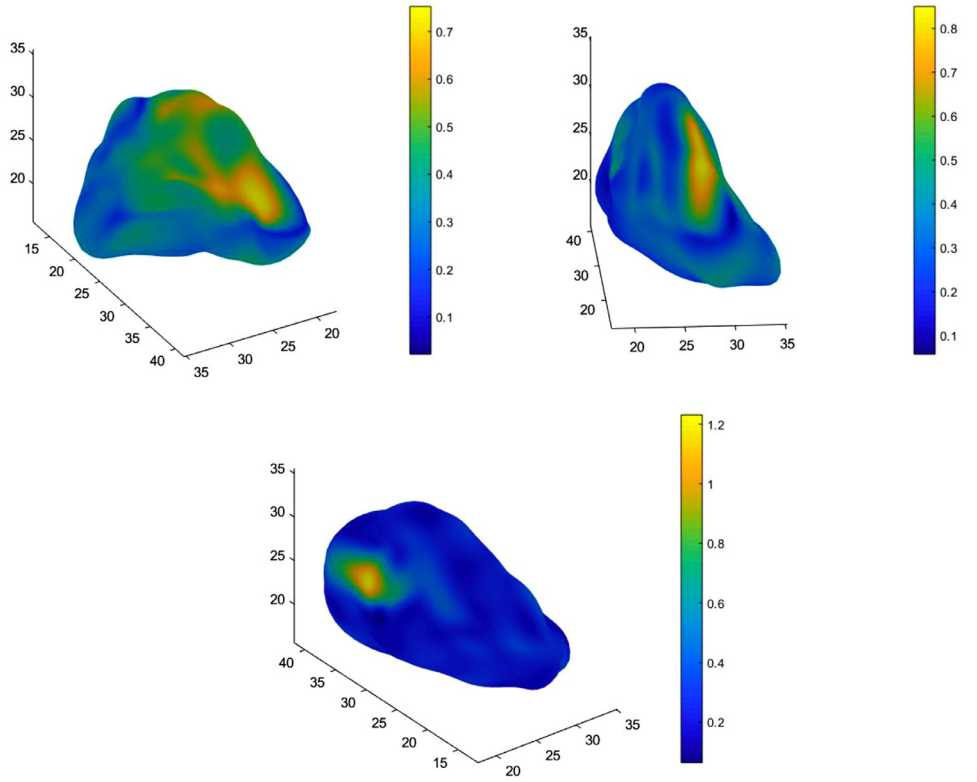


Fig. 4. The effect of first FPCs on the mean shape (from left to right and from top to bottom) for the left putamens. The viewpoint of the first image is defined by azimuth of 150° and elevation of 31°. The viewpoint of the second image is defined by azimuth of -4° and elevation of 20°, while the third one has the same orientation as the reference.

The viewpoints have been selected in order to show the effect of each component better. As the code is available, figures can be reproduced and rotated. The first FPC shows a more global effect than the second and third FPCs, which have a more localized effect. In other words, for the first FPC the effect is distributed throughout a large zone of the putamen; while the effect of the other two components is concentrated on a specific part of the putamen. Fig. 5 displays the linear discriminant function on the mean left putamen by FPCA-SVS-LDA. The directions in which the discriminant score increases fastest are shown by the arrows. The norm of these arrows is displayed by color. The differences between both groups are located in the yellow/orange zones of the putamen. Analogously, Fig. 6 displays the linear discriminant function on the mean right putamen by FICA-SVS-LDA. It seems that the differences are more localized in small zones in the right putamen than in the left putamen, where the differences are more spread out. This also occurred in Sandman et al. (2014) in another study on neuroeducation.

In order to assess the performance of FPCA-SVS-LDA, we estimate it by leave-one-out (LOO) cross-validation. In each trial, one individual is left out, while FPCA-SVS-LDA is applied to the remaining individuals, which constitute the training set of that trial. Then, the FPCA scores for the individual that was left out, which is the test set, are computed and used to predict its class. This procedure is repeated for each individual of the data set. So, finally the performance estimates by LOO are obtained and shown in Table 1 for the left putamen and in Table 2 for the right putamen, together with the LOO performance of FICA-SVS-LDA and other methods explained in Section 3.1. Note that in each trial the FPCs are different, since the training sets are different, and the number of selected FPCs for classification varies for each trial. This is why the

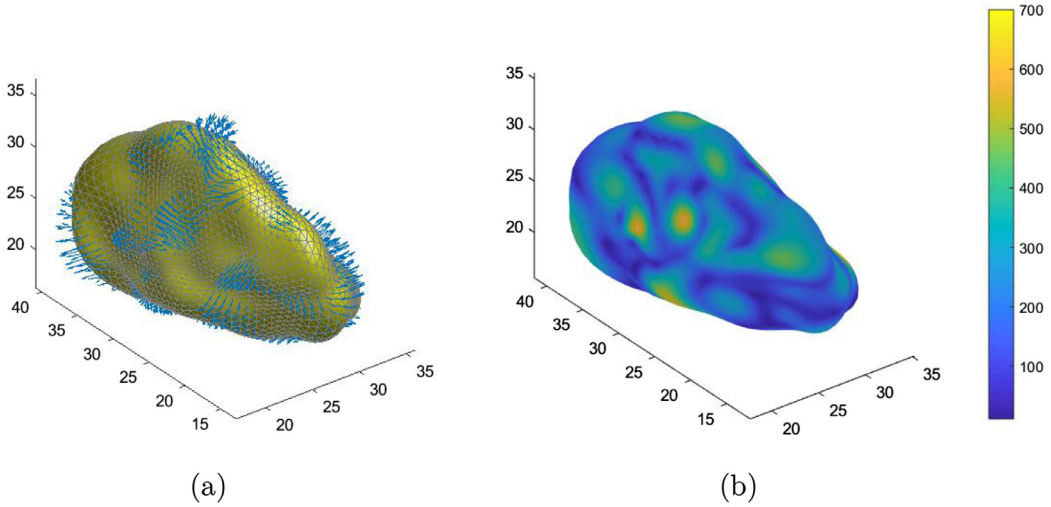


Fig. 5. Representation of the linear discriminating function by FPCA-SVS-LDA, with a vector map (left side) and magnitude map (right side) for the left putamens. The images have the same orientation as the reference (see Fig. 2).

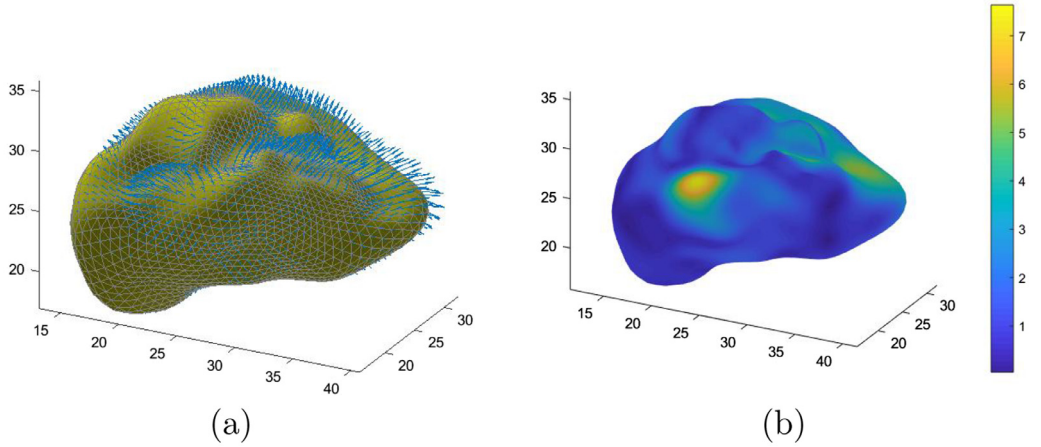


Fig. 6. Representation of the linear discriminating function by FICA-SVS-LDA, with a vector map (left side) and magnitude map (right side) for the right putamens. The viewpoint of the images are defined by azimuth of 63° and elevation of -23°.

Table 1
Left putamen. LOO performance for different methods: accuracy, recall or sensitivity, specificity, precision or positive predictive value, and negative predictive value (NPV), assuming the RE class as the positive class. The maximum value in each column appears in bold.

Method	No. features	Accuracy	Recall	Specificity	Precision	NPV
FPCA-SVS-LDA	14.12	**0.6364**	**0.5333**	**0.7222**	**0.6154**	**0.6500**
FICA-SVS-LDA	1.94	0.5758	0.4667	0.6667	0.5385	0.6
Volume	1	0.4545	0.2	0.6667	0.3333	0.5
SVM	33	0.5152	0.4667	0.5556	0.4667	0.5556
SDA	20	0.4848	0.4	0.5556	0.4286	0.5263
PLS-LDA	10.91	0.4242	0.4	0.4444	0.3750	0.4706
FPCA-LDA	23	0.4848	0.4667	0.5	0.4375	0.5294
FICA-LDA	12	0.4848	0.4667	0.5	0.4375	0.5294

mean of the number of selected FPCs in each trial is shown in the 'No. features' column of those tables.

3.1. Comparison with other methods

We apply different classification methodologies in order to compare the results. The first and simplest one is based on putamen volumetry. The putamen volume is estimated by the sum of the slice areas, i.e. the number of pixels that belong to each slice of the putamen. We perform LDA with this data using LOO cross-validation, and the results are shown in the 'Volume' row of Table 1 and Table 2.

The second methodology is that used in Gerardin et al. (2009). As mentioned previously, an SVM is used to classify the SPHARM coefficients. These are selected with a bagging strategy, where *t*-tests are used for finding the coefficients that best separate the classes. The number of coefficients used is selected by double or nested leave-one-out cross-validation. The results are shown in the 'SVM' row of Table 1 and Table 2.

Table 2
Right putamen. LOO performance for different methods: accuracy, recall or sensitivity, specificity, precision or positive predictive value, and negative predictive value (NPV), assuming the RE class as the positive class. The maximum value in each column appears in bold.

Method	No. features	Accuracy	Recall	Specificity	Precision	NPV
FPCA-SVS-LDA	13.73	0.6364	0.6000	0.6667	0.6000	0.6667
FICA-SVS-LDA	5.42	**0.7273**	**0.667**	**0.7778**	**0.7143**	**0.7368**
Volume	1	0.4242	0	**0.7778**	0	0.4828
SVM	27	0.6364	0.6	0.6667	0.6	0.6667
SDA	20	0.4848	0.4	0.5556	0.4286	0.5263
PLS-LDA	10	0.4848	0.4	0.5556	0.4286	0.5263
FPCA-LDA	7	0.6061	0.4667	0.7222	0.5833	0.6190
FICA-LDA	8	0.6970	0.6	**0.7778**	0.6923	0.7

The third methodology is sparse discriminant analysis (SDA), proposed by Clemmensen et al. (2011). We apply SDA (Sjöstrand et al., 2018) to the SPHARM coefficients. The number of variables is selected by nested leave-one-out cross-validation. The results are shown in the 'SDA' row of Table 1 and Table 2.

The fourth methodology is the method proposed by Boulesteix (2004) and it is implemented in the R package *plsgenomics* (Boulesteix et al., 2018). The choice of the number of latent components is performed by the cross-validation method proposed by Boulesteix (2004). The results are shown in the 'PLS-LDA' row of Table 1 and Table 2.

The fifth and sixth methods are the procedures proposed by Epifanio and Ventura-Campos (2014), where LDA is applied to the coefficients of FPCA and FICA, but without selection of components. The number of components is selected by nested leave-one-out cross-validation. The results are shown in the 'FPCA-LDA' and 'FICA-LDA' rows of Table 1 and Table 2.

For the left putamen, the method that obtains the best results for all the performance measures is our proposal FPCA-SVS-LDA, while the second one is our other proposal, FICA-SVS-LDA. The third best method yields worse results, particularly in terms of accuracy, despite using a high number of features. The importance of considering selection of variables after the dimension reduction step is revealed. Note the great improvement in the measures when SVS is performed. For example, the accuracy goes from 48.48% for FPCA-LDA or FICA-LDA to 63.64% for FPCA-SVS-LDA and 57.58% for FICA-SVS-LDA. It is clear that using a variable selection step after the dimension reduction step has been a success. Note that in this comparison, all the methods except that based on volume, are local shape methods based on SPHARM and the same pre-processing steps have been carried out for all of them. In this way, we have compared the different methods once the SPHARM representation is available.

For the right putamen, the method that obtains the best results for all the performance measures is our proposal FICA-SVS-LDA, which returns better results than those for the left putamen. The second best method in terms of accuracy is FICA-LDA, while the third best are FPCA-SVS-LDA and SVM. As happened with the left putamen, using a variable selection step after the dimension reduction step has improved the results. The accuracy obtained with FICA-SVS-LDA and the right putamen is higher (0.7273) than that obtained with the left putamen.

3.1.1. Multivariate linear model

Although discriminant analysis and testing of mean group difference are different problems, we apply the methodology in Chung et al. (2010) for emphasizing the usefulness of the linear discriminant function applied to brain structures as in Figs. 5 and 6, and for differentiating it from the significance maps of group differences that are commonly used in the neuroimaging literature together with classification results (Gerardin et al., 2009). Multivariate linear modeling (Taylor and Worsley, 2008) is carried out on SPHARM, and the effect of the group variable on the model is tested. Fig. 7 shows the *F*-statistic

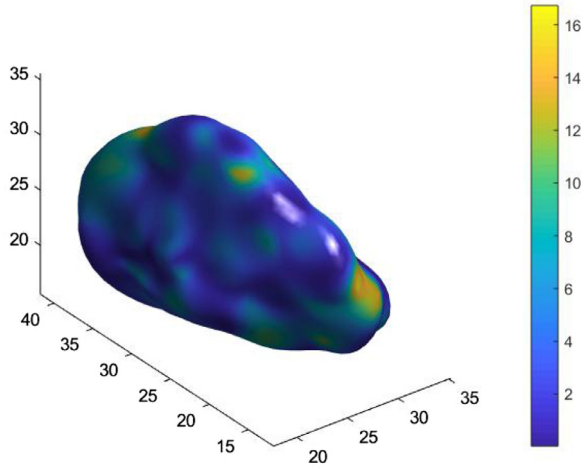


Fig. 7. Left putamen. A map view of the F-statistic of the difference in shape between the subjects in the RE group and the subjects in the non-RE group (the random field-based threshold corresponding to level α=0.05 is 33.93 and for level α=0.1 is 30.26, while the maximum *F*-statistic value is 16.73). The image has the same orientation as the reference (see Fig. 2).

value on the mean left putamen. No statistically significant differences in shape with α=0.05 are found when we test for group differences at each vertex of the putamen surface. However, we can check that discriminant analysis is worthwhile in this problem. In order to check that discrimination is worthwhile in this problem, we apply multivariate analysis of variance (MANOVA) to the PC scores selected to test the hypothesis of equality of group means, as explained in Mardia et al. (1979, Sect. 11.4). In this problem, the $p - value$ is 4-e8, which is small enough to reject the null hypothesis of equality of group means, and discriminant analysis is therefore worthwhile. We have therefore shown the superiority of our approach versus the use of the methodology in Chung et al. (2010) in this problem.

3.1.2. Laplace-Beltrami spectra

We compare our proposal with a methodology for global shape comparison based on the Laplace-Beltrami eigenvalues, as described by Reuter et al. (2006, 2009), Wachinger et al. (2015). The spectra for all the putamens are computed. To compute the eigenfunctions of the Laplace-Beltrami operator, we use the finite element method (FEM) described by Chung and Taylor (2004). Then, we consider a nearest neighbor algorithm (1-NN), with two distances for the eigenvalues that describe each shape, as explained by Wachinger et al. (2015). On the one hand, the Euclidean distance with a linear re-weighting of eigenvalues is considered. We call this procedure LB-WE. On the other hand, the Mahalanobis distance is considered. We call this procedure LB-M. We assess the performance by LOO cross-validation. The number of eigenvalues

used is selected by double or nested leave-one-out cross-validation. For the left putamen, the accuracy is 0.5758 (with 5 eigenvalues) for LB-WE and 0.7272 (with 27 eigenvalues) for LB-M, while for the right putamen, the accuracy is 0.4242 (for any number of eigenvalues considered) for LB-WE and 0.7272 (with 12 eigenvalues) for LB-M. Therefore, the performance of our proposal is better than that of LB-WE. For LB-M, the performance for the right putamen is equal to that of our proposal, but better than our proposal for the left putamen. In any case, the best classification rate is attained for the right putamen with both LB-M and FICA-SVS-LDA. However, LB-WE or LB-M are black box methodologies, since they do not explain why or how these differences occur; they are lacking in human interpretability, unlike our proposal. As discussed previously, it is desirable to have information that provides qualitative understanding (Hastie et al., 2009).

3.2. Limitations

We have seen the advantages of FPCA-SVS-LDA, in terms of both performance and interpretability, especially in comparison with other methodologies. As regards the limitations of our methodology, they may result from the limitations of LDA (Clemmensen et al., 2011) and the variable selection step. On the one hand, a situation where LDA can fail is when the groups cannot be separated by linear boundaries. Then SVS could still be used, but instead of LDA, quadratic discriminant analysis could perhaps be used. A difficult situation occurs when we have unbalanced groups, and the sample size of one of them is very small. In the most extreme case, a single observation per group is insufficient to compute LDA. Furthermore, the confidence in SVS decisions based on very few samples could decrease. On the other hand, the SVS used is based on Wilks' lambda criterion, which may not be the best option for non-Gaussian distributions.

Our methodology is based on local shape analysis rather than global shape analysis. In local shape analysis approaches, one-to-one correspondences between surfaces need to be established. This could be seen as a limitation, since global shape analysis approaches may need fewer pre-processing steps. However, our methodology yields spatially localized results that are straightforward to interpret, unlike global shape analysis approaches. Global shape analysis approaches can be seen as black boxes that do not explain their predictions in a way that humans can understand them easily. Nowadays, it is therefore preferable to use models that are inherently interpretable (Rudin, 2019).

4. Conclusions

We have proposed a methodology based on the use of SPHARM representation of brain structures for classification. The procedure is an improvement on that proposed by Epifanio and Ventura-Campos (2014). We have shown that our proposal not only performs well in terms of predictive power, but also yields interpretable classification in the high-dimensional setting. Furthermore, it has been applied to a novel classification problem in neuroeducation.

Although the procedure has been applied to a binary classification problem, it can be used in multiclass classification problems. In that case, more than one discriminant function can be obtained, as in classical LDA.

As future work, from the practical point of view, the proposed methodology can be applied to any classification problem in neuroscience where the anatomical structures can be expressed with SPHARM coefficients. From the theoretical point of view, other variable selection methods could be studied. Furthermore, we could extend the methodology to the problem of ordinal classification, i.e. when groups (categories or classes) are ordered. In addition, we could extend the methodology to combine not only functional data with SPHARM coefficients, but also multivariate features, such as variables related to education. In other words, we have to define FPCA for hybrid data with vector and multivariate functions, similarly to what Ramsay and Silverman (2005, Chapter 10) did for univariate functions. Finally, the FDA approach could be used not only for classification but also in other problems where the data are 3D brain structures.

Funding

This work has been partially supported by the following grants: DPI2017-87333-R from the Spanish Ministry of Science, Innovation and Universities (AEI/FEDER, EU) and UJI-A2017-8 and UJI-B2017-13 from Universitat Jaume I.

Declaration of Competing Interest

The authors have declared that no competing interests exist.

References

Alfeld, P., Neamtu, M., Schumaker, L.L., 1996. Fitting scattered data on sphere-like surfaces using spherical splines. J. Comput. Appl. Math. 73 (1–2), 5–43.
Anderson, J.R., Betts, S., Ferris, J., Fincham, J., 2012. Tracking children's mental states while solving algebra equations. Hum. Brain Mapp. 33 (11), 2650–2665.
Boulesteix, A.L., 2004. PLS Dimension reduction for classification with microarray data. Stat. Appl. Genet. Mol. Biol. 3 (1), 1–30.
Boulesteix, A.-L., Durif, G., Lambert-Lacroix, S., Peyre, J., Strimmer, K., 2018. plsgenomics: PLS analyses for genomics. R package version 1.5-2. https://CRAN.R-project.org/package=plsgenomics.
Brechbühler, C., Gerig, G., Kübler, O., 1995. Parametrization of closed surfaces for 3-d shape description. Comput. Vision Image Understanding 61 (2), 154–170.
Chung, M.K., Dalton, K.M., Shen, L., Evans, A.C., Davidson, R.J., 2007. Weighted Fourier series representation and its application to quantifying the amount of gray matter. IEEE Trans. Med. Imaging 26 (4), 566–581.
Chung, M.K., Taylor, J., 2004. Diffusion smoothing on brain surface via finite element method. In: 2004 2nd IEEE International Symposium on Biomedical Imaging: Nano to Macro, pp. 432–435. Vol. 1
Chung, M.K., Worsley, K.J., Nacewicz, B.M., Dalton, K.M., Davidson, R.J., 2010. General multivariate linear modeling of surface shapes using surfstat. Neuroimage 53 (2), 491–505.
Clement, J., 1982. Algebra word problem solutions: thought processes underlying a common misconception. J. Res. Math. Educ. 13 (1), 16–30.
Clement, J., Lochhead, J., Monk, G.S., 1981. Translation difficulties in learning mathematics. The American Mathematical Monthly 88, 286. 04
Clement, J., Lochhead, J., Soloway, E., 1980. Positive Effects of Computer Programming on Students Understanding of Variables and Equations. In: Proceedings of the ACM, pp. 467–474.
Clemmensen, L., Hastie, T., Witten, D.B.E., 2011. Sparse discriminant analysis. Technometrics 53 (4), 406–413.
Cooper, M., 1986. The dependence of multiplicative reversal on equation format. Journal of Mathematical Behaviour 5 (2), 115–120.
Dryden, I.L., Mardia, K.V., 1998. Statistical shape analysis. Wiley, Chichester.
Epifanio, I., 2008. Shape descriptors for classification of functional data. Technometrics 50 (3), 284–294.
Epifanio, I., Ventura-Campos, N., 2011. Functional data analysis in shape analysis. Computational Statistics & Data Analysis 55 (9), 2758–2773.
Epifanio, I., Ventura-Campos, N., 2014. Hippocampal shape analysis in alzheimer's disease using functional data analysis. Stat. Med. 33 (5), 867–880.
Ferrando, L., 2019. Classificació d'estructures cerebrals en 3D amb anàlisi de dades funcionals. Universitat Jaume I. Aplicació a problemes verbals amb errors d'inversió. master's thesis.
Ferrando, L., Ventura-Campos, N., Epifanio, I., 2020. A neuroimaging data set on problem solving in the case of the reversal error: Putamen data. Data in Brief (submitted).
Ferraty, F., Vieu, P., 2006. Nonparametric functional data analysis: Theory and practice. Springer.
Gerardin, E., Chetelat, G., Chupin, M., Cuingnet, R., Desgranges, B., Kim, H., Niethammer, M., Dubois, B., Lehericy, S., Garnero, L., Eustache, F., Colliot, O., 2009. Multidimensional classification of hippocampal shape features discriminates alzheimer's disease and mild cognitive impairment from normal aging. Neuroimage 47 (4), 1476–1486.
Gerig, G., Styner, M., Jones, D., Weinberger, D., Lieberman, J., 2001. Shape Analysis of Brain Ventricles Using SPHARM. In: Proceedings of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA'01), pp. 171–178.
Golland, P., Grimson, W.E.L., Shenton, M.E., Kikinis, R., 2001. Deformation Analysis for Shape Based Classification. In: Proceedings of the 17th International Conference on Information Processing in Medical Imaging. IPMI '01, pp. 517–530.
González-Calero, J.A., Arnau, D., Laserna-Belenguer, B., 2015. Influence of additive and multiplicative structure and direction of comparison on the reversal error. Educational Studies in Mathematics 89 (1), 133–147.
Gu, X., Wang, Y., Chan, T.F., Thompson, P.M., Yau, S., 2004. Genus zero surface conformal mapping and its application to brain surface mapping. IEEE Trans. Med. Imaging 23 (8), 949–958.
Hall, P., Poskitt, D., Presnell, B., 2001. A functional data-analytic approach to signal discrimination. Technometrics 43, 1–9.

Hanakawa, T., Honda, M., Okada, T., Fukuyama, H., Shibasaki, H., 2003. Neural correlates underlying mental calculation in abacus experts: a functional magnetic resonance imaging study. Neuroimage 19 (2), 296–307.
Hand, D.J., 2006. Classifier technology and the illusion of progress. Statistical Science 21 (1), 1–14.
Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer Series in Statistics. Springer, New York.
He, Y., Li, X., Gu, X., Qin, H., 2005. Brain Image Analysis Using Spherical Splines. In: Proc. of Energy Minimization Methods in Computer Vision and Pattern Recognition, pp. 633–644.
Hyvärinen, A., Karhunen, J., Oja, E., 2000. Independent component analysis: algorithms and applications. Neural Networks 13, 411–430.
Hyvärinen, A., Karhunen, J., Oja, E., 2001. Independent component analysis. Wiley, New York.
Jolliffe, I.T., 2002. Principal component analysis, 2nd edition Springer.
Joshi, S.C., Miller, M.I., Grenander, U., 1997. On the geometry and shape of brain sub-manifolds. Internation Journal of Pattern Recognition and Artificial Intelligence 11 (8), 1317–1343.
Lazar, N., 2008. The statistical analysis of functional MRI data. statistics for biology and health. Springer, New York.
Lila, E., Aston, J.A.D., Sangalli, L.M., 2016. Smooth principal component analysis over two-dimensional manifolds with an application to neuroimaging. Ann. Appl. Stat. 10 (4), 1854–1879.
Macdonald, D., Kabani, N., Avis, D., Evans, A.C., 2000. Automated 3-d extraction of inner and outer surfaces of cerebral cortex from MRI. Neuroimage 12 (3), 340–356.
Mardia, K., Kent, J., Bibby, J., 1979. Multivariate analysis. probability and mathematical statistics. Academic Press.
Millán-Roures, L., Epifanio, I., Martínez, V., 2018. Detection of anomalies in water networks by functional data analysis. Mathematical Problems in Engineering 2018. (Article ID 5129735), Article ID 5129735
Nain, D., Styner, M., Niethammer, M., Levitt, J.J., Shenton, M., Gerig, G., Tannenbaum, A., 2007. Statistical Shape Analysis of Brain Structures Using Spherical Wavelets. In: Proceedings of the Fourth IEEE International Symposium on Biomedical Imaging, pp. 209–212.
Park, Y., Priebe, C.E., Miller, M.I., Mohan, N.R., Botteron, K.N., 2008. Statistical analysis of twin populations using dissimilarity measurements in hippocampus shape space. Journal of Biomedicine and Biotechnology 2008. (Article ID 694297, doi:10.1155/2008/694297), Article ID 694297
Pierola, A., Epifanio, I., Alemany, S., 2016. An ensemble of ordered logistic regression and random forest for child garment size matching. Computers & Industrial Engineering 101, 455–465.
Qin, Y., Carter, C.S., Silk, E.M., Stenger, V.A., Fissell, K., Goode, A., Anderson, J.R., 2004. The change of the brain activation patterns as children learn algebra equation solving. Proceedings of the National Academy of Sciences of the United States of America 101, 5686–5691. 05
Ramsay, J.O., Silverman, B.W., 2002. Applied functional data analysis. Springer.
Ramsay, J.O., Silverman, B.W., 2005. Functional data analysis, 2nd edition Springer.
Reuter, M., Wolter, F.-E., Peinecke, N., 2006. Laplace-beltrami spectra as “shape-DNA” of surfaces and solids. Comput. Aided Des. 38 (4), 342–366.
Reuter, M., Wolter, F.-E., Shenton, M., Niethammer, M., 2009. Laplace Beltrami eigenvalues and topological features of eigenfunctions for statistical shape analysis. Comput.-Aided Des. 41 (10), 739–755.
Rudin, C., 2019. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence 1, 206–215.
R Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. http://www.R-project.org/
Sandman, C.A., Head, K., Muftuler, L.T., Su, L., Buss, C., Davis, E.P., 2014. Shape of the basal ganglia in preadolescent children is associated with cognitive performance. Neuroimage 99, 93–102.
Shen, L., Ford, J., Makedon, F., Saykin, A., 2004. A surface-based approach for classification of 3D neuroanatomic structures. Intell. Data Anal. 8 (6), 519–542.
Shen, K., Fripp, J., Meriaudeau, F., Chetelat, G., Salvado, O., Bourgeat, P., 2012. Detecting global and local hippocampal shape changes in alzheimer’s disease using statistical shape models. Neuroimage 59 (3), 2155–2166.
Shen, L., Farid, H., McPeek, M., 2009. Modeling 3-dimensional morphological structures using spherical harmonics. Evolution 4 (63), 1003–1016.
Sjöstrand, K., Clemmensen, L., Larsen, R., Einarsson, G., Ersbøll, B., 2018. SpaSM: a MATLAB toolbox for sparse statistical modeling. J. Stat. Softw. 84 (10), 1–37.
Sørensen, H., Goldsmith, J., Sangalli, L.M., 2013. An introduction with medical applications to functional data analysis. Stat. Med. 32 (30), 5222–5240.
Styner, M., Gerig, G., Joshi, S.C., Pizer, S.M., 2003. Automatic and robust computation of 3D medial models incorporating object variability. Int. J. Comput. Vis. 55 (2–3), 107–122.
Styner, M., Lieberman, J.A., Pantazis, D., Gerig, G., 2004. Boundary and medial shape analysis of the hippocampus in schizophrenia. Medical Image Analysis Journal 8 (3), 197–203.
Taylor, J., Worsley, K., 2008. Random fields of multivariate test statistics, with applications to shape analysis. Ann. Stat. 36 (1), 1–27.
Tian, T.S., 2010. Functional data analysis in brain imaging studies. Front. Psychol. 1, 35.
Timsari, B., Leahy, R.M., 2000. Optimization method for creating semi-isometric flat maps of the cerebral cortex. In: Proc. SPIE, Medical Imaging 3979, 698–708.
Ullah, S., Finch, C.F., 2013. Applications of functional data analysis: asystematic review. BMC Med. Res. Methodol. 13 (1), 43.
Viviani, R., Grön, G., Spitzer, M., 2005. Functional principal component analysis of fMRI data. Hum. Brain Mapp. 24 (2), 109–129.
Wachinger, C., Golland, P., Kremen, W., Fischl, B., Reuter, M., 2015. Brainprint: a discriminative characterization of brain morphology. Neuroimage 109, 232–248.
Wang, J.-L., Chiou, J.-M., Müller, H.G., 2016. Functional data analysis. Annu. Rev. Stat. Appl. 3 (1), 257–295.
Weihs, C., Ligges, U., Luebke, K., Raabe, N., 2005. klaR Analyzing German Business Cycles. In: Data Analysis and Decision Support. Springer Berlin, Heidelberg, Berlin, Heidelberg, pp. 335–343.
Wollman, W., 1983. Determining the sources of error in a translation from sentence to equation. J. Res. Math. Educ. 14 (3), 169–181.
Yu, P., Grant, P.E., Qi, Y., Han, X., Ségonne, F., Pienaar, R., Busa, E., Pacheco, J., Makris, N., Buckner, R.L., Golland, P., Fischl, B., 2007. Cortical surface shape analysis based on spherical wavelets. IEEE Trans. Med. Imaging 26 (4), 582–597.

Aportació 2.2:

A neuroimaging data set on problem solving in the case of the reversal error: Putamen data

Ferrando, Ventura-Campos et al. (2020b)

Data in Brief 33 (2020) 106322

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

A neuroimaging data set on problem solving in the case of the reversal error: Putamen data

Lara Ferrando [a], Noelia Ventura-Campos [a,b,*], Irene Epifanio [c]

[a] Grup Neuropsicologia i Neuroimatge Funcional, Universitat Jaume I, Spain
[b] Dept. Educació i Didàctiques Específiques, Universitat Jaume I, Spain
[c] Dept. Matemàtiques-IF, Universitat Jaume I, Spain

ARTICLE INFO

Article history:
Received 6 August 2020
Revised 8 September 2020
Accepted 16 September 2020
Available online 19 September 2020

Keywords:
MRI
Neuroeducation
3D shape
Algebra problem solving
Reversal error

ABSTRACT

Structural Magnetic Resonance Images (sMRI) for a sample of university students were recorded. Out of magnetic resonance, students performed a test of algebra problem solving. As we are interested in reversal errors, the test was prepared to detect this kind of error.

Depending on the number of mistakes made, students were divided into two groups: one group contains 15 students that responded erroneously to more than 60% of the 16 questions, and the other group contains 18 students that did not make any mistake.

We are interested in the more relevant brain structures for this neuroeducation problem. The analysis of these data can be found in Ferrando et al. (2020) [1]. The results of the volumetric analysis showed differences between groups in the right and left putamen. Therefore, both putamens were preprocessed and segmented to use them in the shape analysis. The dataset contains the slices of the left and right putamen and the left putamen of each of 33 subjects, 20 females. It also contains a vector that indicates the group to each subject belongs to.

DOI of original article: 10.1016/j.neuroimage.2020.117209

* Corresponding author at: Grup Neuropsicologia i Neuroimatge Funcional, Universitat Jaume I, Spain.
E-mail address: venturan@uji.es (N. Ventura-Campos).
Social media: (L. Ferrando)

https://doi.org/10.1016/j.dib.2020.106322

Specifications Table

Subject	Neuroscience
Specific subject area	Neuroimaging, education and 3D shape analysis
Type of data	Octave or MatLab file; Rdata file
How data were acquired	A 3 Tesla Philips scanner and 1.5 Tesla Siemens Symphony scanner were used to obtain the images. After, SPM12 (toolbox of MatLab) was used to analyze the images. The images were preprocessing and segmented using the method Voxel Based Morphometry (CAT12).
Data format	Raw
Parameters for data collection	Philips scanner: High-resolution T1-weighted, TR = 8.4 ms, TE = 3.8 ms, matrix size = 320 × 320 × 250 and voxel size = 0.75 × 0.5 × 0.8 mm Siemens Symphony scanner: High-resolution T1-weighted, TR = 2200 ms, TE = 3 ms, flip angle = 90°, matrix size = 256 × 256 × 160 and voxel size = 1 × 1 × 1 mm. Acquisitions covered the entire brain and were performed in parallel to the anterior commissure-posterior commissure plane (AC-PC).
Description of data collection	The slices of the left and right putamens of thirty-three participants (20 females) were collected. Besides a vector indicating the label for each subject (1 for RE-makers and 2 for non-RE makers). The ages of the participants ranges from 18 to 26 years.
Data source location	Institution: Universitat Jaume I City/Town/Region: Castellón Country: Spain
Data accessibility	With the article
Related research article	L. Ferrando, N. Ventura-Campos, I. Epifanio. Detecting and visualizing differences in brain structures with SPHARM and functional data analysis, Neuroimage [1].

Value of the Data

- This is data set about putamen surfaces to the phenomenon of reversal error in the algebra problem solving.
- Data come from a real and important neuroeducational problem like algebra problem solving.
- Data set is useful for reproducibility and to further studies about reversal error problem.
- The data set can be used to benchmark and to compare methods of classification of 3D shapes in general.
- The data set can be beneficial to obtain a big data about the MRI segmentation of putamen.
- The data set can be used to perform a meta-analysis on the association of putamen and mathematical learning.

1. Data Description

The free and open Octave or MatLab file contains three variables: Hleft is a struct MatLab object with the slices of the left putamen for each participant. Hrigh is a struct MatLab object with the slices of the right putamen for each participant (see Fig. 1). The vector g contains the labels that indicates to which group each participant belongs to (1 for RE-makers and 2 for non-RE makers). Data is also provided in Rdata format for the free and open R software, with the same structure.

2. Experimental Design, Materials and Methods

2.1. Participants

In this study were collected data from thirty-three participants (20 females) with 18–26 years (mean age: 22.03, SD: 2.36). The subjects did not have any severe neurological and medical disease, traumatism, loss of consciousness, and the typical exclusion criteria when the magnetic resonance is performing.

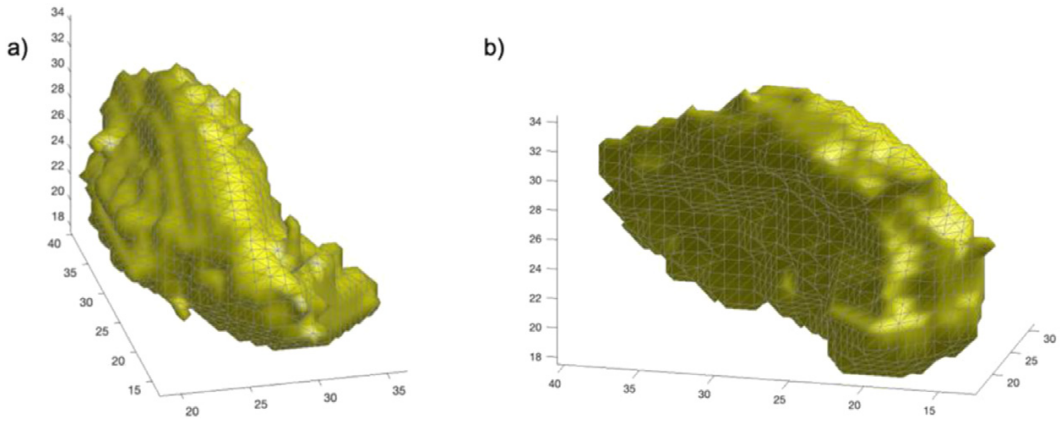


Fig. 1. Representation of the segmentation of Putamen. Example of reconstruction 3D of one participant of slices of a) left putamen and b) right putamen.

2.2. Acquisition images

For the acquisition images were used two scanners. The first Philips scanner (High-resolution T1-weighted, TR = 8.4 ms, TE = 3.8 ms, matrix size = $320 \times 320 \times 250$ and voxel size = $0.75 \times 0.5 \times 0.8$ mm.) and the second Siemens Symphony scanner (High-resolution T1-weighted, TR = 2200 ms, TE = 3 ms, flip angle = 90°, matrix size = $256 \times 256 \times 160$ and voxel size = $1 \times 1 \times 1$ mm). The scanner acquisitions were performed in parallel to the anterior commissure-posterior commissure plane (AC-PC). The MRI scans were acquired while subjects were in rest.

2.3. Data preprocessing

The pre-processing of the images was carried out with SPM (SPM12 (v7219), Wellcome Trust Centre for Neuroimaging, London, UK, http://www.fil.ion.ucl.ac.uk/spm/software/spm12), using the methodology VBM with CAT12 toolbox to perform the pre-processing steps (CAT12.5, http://dbm.neuro.uni-jena.de/cat/). The standard pre-processing suggested in CAT12 manual was performed, where we used an 8-mm full- width-half-maximum Gaussian smoothing.

To segment the putamens, the *imcalc* toolbox of SPM12 was used together with an intersection between the image of each of the subjects with the putamen of the AAL atlas. To obtain bit map formats, MRIcro was used, so that we can get the putamen axial- slices in 2D.

Ethics Statement

All participants were students of Universitat Jaume I. Before participating, they signed a written consent form. All experimental procedures followed the guidelines of the research ethics committee at Universitat Jaume I.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have, or could be perceived to have, influenced the work reported in this article.

Acknowledgments

This work has been partially supported by the following grants: DPI2017- 87333-R from the Spanish Ministry of Science, Innovation and Universities (AEI/FEDER, EU) and UJI-A2017-8 and UJI-B2017-13 from Universitat Jaume I.

Supplementary Materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.dib.2020.106322.

Reference

[1] L. Ferrando, N. Ventura-Campos, I. Epifanio, Detecting and visualizing differences in brain structures with SPHARM and functional data analysis, Neuroimage 222 (2020) 117209, doi:10.1016/j.neuroimage.2020.117209.

Aportació 3:

Ordinal classification of 3D brain structures by functional data analysis

Ferrando, Epifanio et al. (2020)

Received: Added at production | Revised: Added at production | Accepted: Added at production

DOI: xxx/xxxx

ARTICLE TYPE

Ordinal classification of 3D brain structures by functional data analysis

Lara Ferrando[1] | **Irene Epifanio***[2] | **Noelia Ventura-Campos**[1,3]

[1]Grup Neuropsicologia i Neuroimatge Funcional, Universitat Jaume I, Castelló, Spain
[2]Dept. Matemàtiques-IF, Universitat Jaume I, Castelló, Spain
[3]Dept. Educació i Didàctiques Específiques, Universitat Jaume I, Castelló, Spain

Correspondence
*Irene Epifanio, Dept. Matemàtiques, Universitat Jaume I, Castelló 12071, Spain
Email: epifanio@uji.es

Abstract

The aim of ordinal classification is to identify to which of a set of ordered classes a new item belongs, on the basis of a training set of data. Although nominal classification, where classes are unordered, has received a lot of attention and many methodologies exist for both multivariate and functional data, this is not the case for ordinal classification, despite many real-word applications requiring the classification of observations into naturally ordered classes. Here we introduce several ordinal classification methods for functional data, specifically multiargument and multivariate functional data. We analyze the performance of the proposed methods in four real data sets and compare them with other alternatives. These data sets belong to a neuroeducational problem and a neuropathologic problem, where 3D brain structures are represented by functional data. The results confirm that taking into account ordering information improves the performance.

KEYWORDS:
Ordinal data; Supervised statistical learning; Functional data; Neuroeducation; Alzheimer's disease

1 | INTRODUCTION

Supervised learning is one of the most common problem in statistics. Classification is a supervised problem where the objective is to predict a class or label in a set $C_1, \ldots, C_Q$, which is the so-called output, response or dependent variable, from a set of inputs, features or independent variables. Usually classes are considered as unordered, i.e. as levels of a nominal variable and the majority of classification algorithms are designed for this kind of problems. However, classes are ordered, i.e. labels are levels of an ordinal variable, in many real life problems, such as collaborative filtering, credit rating, econometric modeling, information retrieval, medicine, psychology, social sciences, text classification, wind speed prediction, and more (Gutiérrez, Pérez-Ortiz, Sánchez-Monedero, Fernández-Navarro, & Hervás-Martínez 2016). An example of ordered variable would be patient condition (good, fair, serious, critical) or the rating of satisfaction (very low, low, indifferent, high, very high). Nevertheless, the literature about ordered classification methods is not very extensive for multivariate data (Pierola, Epifanio, & Alemany 2016) or high-dimensional data (Hornung 2020; Simó, Ibáñez, Epifanio, & Gimeno 2020), and even less so for functional data (Wang & Shi 2014).

In the multivariate case, Gutiérrez et al. (2016) established a taxonomy according to how the order is taken into account in the classification procedure. They proposed three main approaches. The first one is the naïve approach, which is the simplest and is very common, not only in the multivariate context but also in the functional context. It consists of using standard classification algorithms as if classes were unordered, i.e. nominal classification. In this case, the ordering information is not taken into account and that information is lost. The second approach consists of decomposing the ordinal problem into several binary ones. Then, each of them can be solved by standard classification problem and the results are combined to return a label, as described by Frank and Hall (2001). The third approach includes other methods, which assume that an unobserved continuous

variable underlies the ordinal response. Some of the methods that belong to this approach are: ordinal logistic regression models, such as the cumulative link models used by Pierola et al. (2016), augmented binary classification problems, such as the data replication method by Cardoso and Costa (2007), or ensemble methods, such as ordinal random forests by Hornung (2020).

According to Gutiérrez et al. (2016), although results from the naïve approach can be very competitive in the multivariate case, taking into account the order improves the performance. In the functional case, the few papers that deals with ordinal classification when inputs are functional data are based mainly on the use of the functional generalized linear model (Aguilera & Escabias 2008; Barahona, Centella, Gual-Arnau, Ibáñez, & Simó 2020), and in many cases the order is ignored (Epifanio & Ventura-Campos 2014). To the best of our knowledge, ordinal methods for functional data with the ordinal binary decomposition approach, the second approach, have not been considered until now, and the third approach has not been fully exploited either.

On the other hand, classification problems in neuroscience do not usually exploit the ordering information of the different classes in the multivariate case. For example, few papers use class order in Alzheimer's Disease (AD) diagnosis. Some exceptions are the papers by Fan (2011) and Cruz, Silveira, and Cardoso (2018), which use biomarkers (feature vectors) from brain imaging as inputs.

The objective of this work is to introduce more ordinal methods for functional data, to use them in two neuroscience problems for the first time and to compare these methods. As a novelty, we will present ordinal methods for the second and third approach in the functional context. Note also that the functional data in both neuroscience problems are multivariate functional data with multiple arguments, which are not the classical univariate functional data. Furthermore, functional data analysis (FDA) is not usually applied to the analysis of brain structures, with some exceptions (Epifanio & Ventura-Campos 2014; Lila, Aston, & Sangalli 2016), although it is a natural way to do it. In fact, in the neuroscience literature, many brain structures are represented by a functional basis, such as spherical harmonic (SPHARM) representation, but they do not go any further and functional data procedures are not exploited in this field. To the best of our knowledge, no previous work has considered ordinal classification in neuroscience with functional data as inputs.

The outline of the paper is as follows: Section 2 reviews ordinal classification procedures for the multivariate case and some definitions. Section 3 introduces the motivating problems and describes the data. Section 4 contains the ordinal classification procedures for the functional case, which includes our new proposals. Results are discussed in Section 5. Finally, Section 6 contains conclusions and some ideas for future work.

2 | PRELIMINARY

2.1 | Ranked probability score (RPS)

The RPS for probabilistic forecasts of ordered events (NCAR - Research Applications Laboratory 2015; Wilks 2006) is a squared measure that compares the cumulative density function (CDF) of a probabilistic forecast with the CDF of the corresponding observation over a given number of discrete probability classes. In this way, the ordering information of the levels is considered. It measures how well the probability forecast predicts the class that the instance falls into. RPS ranges from 0 (perfect prediction) to 1. The Brier score is a special case of an RPS with two classes. Unlike accuracy, which simply takes into account whether or not the prediction is correct but not the order information, RPS is more informative.

2.2 | The method of Frank and Hall (FH)

Frank and Hall (2001) proposed to decompose the ordinal classification problem with ordered classes $C_1, ..., C_Q$, into the following binary ones: they discriminated $C_1, ..., C_i$ against $C_{i+1}, ..., C_Q$. Then, for a new instance with input $\mathbf{x}$, the predicted probability values of each of the Q classes for the corresponding output y are estimated by: $P(y = C_1|\mathbf{x}) = 1 - p_1$; $P(y = C_q|\mathbf{x}) = p_{q-1} - p_q$, q = 2, ..., Q − 1; $P(y = C_Q|\mathbf{x}) = 1 - p_{Q-1}$, where $p_q = P(y > C_q \mid \mathbf{x})$ for q = 1, ..., Q − 1. This method was proposed for the multivariate case. Note that this method is applicable as long as the binary classifier produces class probability estimates.

2.3 | Ordered logistic regression

Let $\mathbf{X}$ be an N × K matrix with K inputs in N instances and $\mathbf{y}$ a vector, an ordered factor with Q levels, which contains the outputs. The cumulative link model is explained in detail by Agresti (2002, Ch. 7). The model is $\text{logit}\, P(y <= q|\mathbf{x}) = \zeta_q - \eta$, where the logit link function is the inverse of the standard logistic cumulative distribution function, i.e. $\text{logit}(p) = \log(p/(1-p))$, ζ_q parameters provide each cumulative logit, and η is the linear predictor $\beta_1 x_1 + ... + \beta_K x_K$. In our implementation, we chose the model by a forward stepwise model selection using Akaike's information criterion (AIC). This model predicts the class probabilities for a new instance, once the parameters have been estimated. The polr and extractAIC functions from the R package **MASS** (Venables & Ripley 2002) have been used in the implementation. We refer to this method as POLR.

2.4 | The data replication method

Cardoso and Costa (2007) proposed the data replication method for the multivariate case, where the ordinal classification problem is reduced to binary classification problems by augmenting the features, through their replication. A final classification rule is built based on the results of the binary problems to obtain the prediction of a new instance. We refer to this method as oSVM (ordinal Support Vector Machine), since SVMs are used in the implementation, which is available at `http://www.inescporto.pt/~jsc/ReproducibleResearch/20101030.oSVM.zip`.

2.5 | Ordinal forest (OF)

Hornung (2020) proposed a random forest-based prediction method for ordinal outputs in the multivariate case. The idea of OF is to use optimized score values in place of the category values of the ordinal output. The method is implemented in the R package **ordinalForest** (Hornung 2019).

2.6 | Kernel-Induced Random Forests (KIRF)

Fan, Cao, and Wang (2010) proposed KIRF for functional data. Kernel-induced classification trees are built using kernels of each two different training instances as candidate splitting rules. These trees are used in KIRF. Fan et al. (2010) proposed KIRF for functional data classification by defining some kernels for functional data. Note that only nominal functional classification was considered with univariate and unidimensional functional data. They considered the kernel function of two curves as the squared Euclidean norm of the principal component (PC) scores of the PC expansion with K components.

3 | DATA

3.1 | Neuroeducational problem

Ferrando, Ventura-Campos, and Epifanio (2020a) analyzed the 3D shape of the left and right putamens in a functional binary classification problem. Structural Magnetic Resonance Imaging (sMRI) brain scans for a sample of university students were recorded and are available in Ferrando, Ventura-Campos, and Epifanio (2020b). Furthermore, students performed an algebra problem solving test for detecting reversal errors. Students were divided into two groups according to the number of errors. Class E contains 15 students (11 females, mean age: 21.5, SD: 2.1) who answered more than 40% of the questions incorrectly, while class NE contains 18 students (9 females, mean age: 22.5, SD: 2.5) who did the test perfectly. In this work, we consider another class, ME with mild errors. This is an intermediate class between E and NE and it contains 34 students (18 females, mean age: 22.3; SD: 1.8) who answered more than a half of the 16 questions correctly but with one or more errors. Therefore, the output is an ordinal factor.

Ferrando et al. (2020a) included details about the processing of sMRI brain scans, the surface parametrization, which follows the method by Chung, Worsley, Nacewicz, Dalton, and Davidson (2010) and spherical harmonic (SPHARM) representation of the putamens. The number of basis functions is determined as detailed by Ferrando et al. (2020a), following the method by Millán-Roures, Epifanio, and Martínez (2018). In particular, it is L = 11 for left putamens and L = 12 for right putamens. Then, functional principal component analysis (FPCA) for multivariate functions with multivariate arguments is carried out as explained by Ferrando et al. (2020a). As spherical harmonics are orthonormal, we only need to compute the multivariate PCA of the N $\times$ 3M matrix $\mathbf{C}$ (M = 144 for L = 11 and M = 169 for L = 12), where the matrix $\mathbf{C}$ with N rows (one per student) is defined by stacking the vectors of SPHARM basis coefficients $\mathbf{c}^i = (\{c^x_{ilm}\}, \{c^y_{ilm}\}, \{c^z_{ilm}\})$, with l = 0, ..., L and m = $-$l to l.

An example of one right putamen for each of the three classes appears in Figure 1.

3.2 | Neuropathologic problem

Epifanio and Ventura-Campos (2014) analyzed the 3D shape of the left and right hippocampi in a functional classification problem, with ordinal classes, but treated as nominal classes. A total of 28 subjects were considered: 12 cognitively normal (CN) subjects, 6 individuals with mild cognitive impairment (MCI) and 10 subjects with early AD. Therefore, the output is an ordinal factor. Processing of the brain sMRI scans, surface parametrization and SPHARM representation is performed as in Section 3.1 and is also explained by Epifanio and Ventura-Campos (2014). In this case, L = 15 (M = 256) for left and right hippocampi. FPCA is also computed as in Section 3.1.

An example of one left hippocampus for each of the three classes appears in Figure 2.

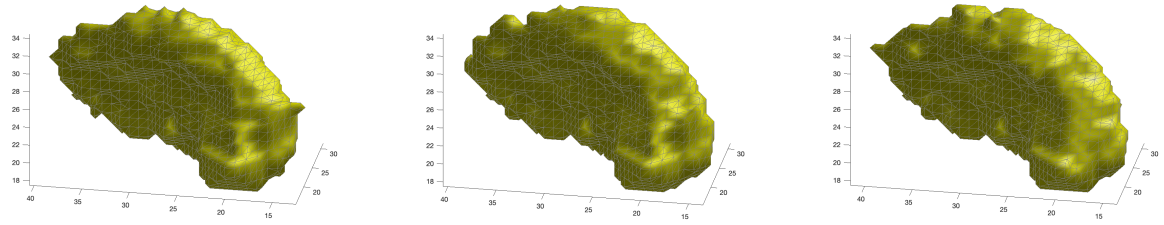


FIGURE 1 A right putamen for each class (from left to right): E, ME and NE.

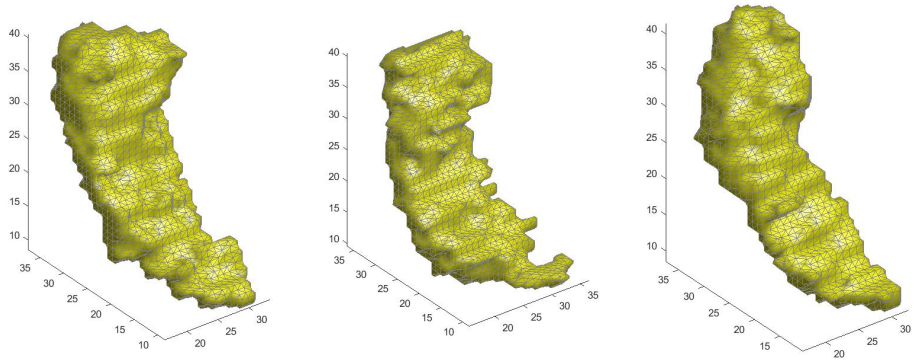


FIGURE 2 A left hippocampus for each class (from left to right): CN, MCI and AD.

4 | METHODOLOGY

4.1 | Previous methodologies

FPCA-SVS-LDA This method is a nominal classification procedure. It was proposed by Ferrando et al. (2020a) and it consists of using FPCA followed by stepwise variable selection for linear discriminant classification. This method belongs to the first approach, where the order is not taken into account.

FPCA-POLR Aguilera and Escabias (2008) proposed the use of FPCA to solve multicollinearity in functional multinomial logit models for ordinal responses. We use FPCA followed by POLR. This method belongs to the third approach.

4.2 | Proposed methodologies

FH-FPCA-SVS-LDA We propose to use the FH method considering FPCA-SVS-LDA as the binary classifier, since this method returns class probability estimates. To the best of our knowledge, this is the first time FH has been used with functional data. This method belongs to the second approach.

FPCA-oSVM We consider the same idea used by Epifanio (2008) and Epifanio and Ventura-Campos (2011): a feature extraction stage, whose resulting features are put forward to a (nominal) classification stage. Feature extraction is a powerful preprocessing method for improving the performance of a learning algorithm (see section on Feature Extraction in Hastie, Tibshirani, and Friedman (2009, pp. 126-127)). In this

case, FPCA is used for the feature extraction stage. As the FPC scores are multivariate, oSVM is used for the ordinal classification stage. This method belongs to the third approach.

FPCA-OF We consider again the previous idea. FPCA is also used for the feature extraction stage, but OF is used for the ordinal classification stage. This method belongs to the third approach.

FPCA-KIOF We consider KIRF for functional data, but in this case OF is used for ordinal classification with multivariate and multiargument functional data. This method belongs to the third approach.

5 | RESULTS AND DISCUSSION

We apply the methods in Section 4 to the data in Section 3. The final class assignation is implemented by choosing the class with the highest probability. The performance of the methods is assessed by leave-one-out (LOO) cross-validation. In each trial, we leave one subject out and FPCA is applied to the remaining subjects, which are the training set of that trial. We compute the FPCA scores of the left-out subject, which is the test set of that trial, and we predict its class and/or save the class probability estimates. We repeat this process for each subject of the data set. In this way, we obtain LOO performance estimates. The performance is evaluated with different measures: the common used measures for nominal classification assessment, such as accuracy, but we also use RPS, which is specific for ordered classification. Tables 1 and 2 show the results for the left and right putamen, respectively, while Tables 3 and 4 show the results for the left and right hippocampi, respectively. RPS can be computed for all the methods, except FPCA-oSVM, since oSVM returns only the predicted class. On the one hand, the neuroeducation problem is a very difficult classification problem, as it is not easy to distinguish between classes with the putamen shape, both with left and right putamens. On the other hand, the neuropathological problem is an easier classification problem. The classes can be distinguished better, especially for the left hippocampus. Therefore, we can see the performance in different kinds of real problems.

TABLE 1 Performance measures for several classifiers, for the left putamens data set: accuracy, recall or sensitivity, specificity, precision or positive predictive value, and negative predictive value (NPV) for each class, and RPS. The best value in each column appears in bold.

Method	Accuracy	Sensitivity			Specificity			Precision			NPV			RPS
		E	ME	NE	E	ME	NE	E	ME	NE	E	ME	NE	
FPCA-SVS-LDA	0.3134	**0.2667**	0.3529	0.2778	0.7115	0.5758	0.6531	0.2105	0.4615	0.22727	0.7708	0.4634	0.7111	0.3922
FPCA-POLR	0.4179	**0.2667**	0.5294	0.3333	0.7308	**0.6667**	0.7143	0.2222	**0.6207**	0.3	0.7755	**0.5789**	0.7447	0.3699
FH-FPCA-SVS-LDA	0.4179	**0.2667**	0.5	**0.3889**	0.7308	0.6364	0.7347	0.2222	0.5862	0.3500	0.7755	0.5526	0.766	0.3562
FPCA-oSVM	0.4478	0.2	0.7059	0.1667	0.9231	0.2727	**0.8163**	0.4286	0.5	0.25	0.8	0.4737	0.7273	-
FPCA-OF	0.4627	0.2	**0.7647**	0.1111	**1**	0.2121	0.7959	**1**	0.5	0.1667	**0.8125**	0.4667	0.7091	**0.1911**
FPCA-KIOF	**0.4776**	0.2	0.6471	**0.3889**	0.8846	0.4242	0.7959	0.3333	0.5366	**0.4118**	0.7931	0.5385	**0.78**	0.1931

According to the results in Table 1, it is clear that taking into account the order improves the performance, since FPCA-SVS-LDA is the method with the lowest accuracy and the highest RPS. The methods with the highest accuracy are FPCA-KIOF, FPCA-OF and FPCA-oSVM, which are also the methods with the lowest RPS: FPCA-OF and FPCA-KIOF. However, many of the subjects are classified in the ME class for FPCA-KIOF, FPCA-OF and FPCA-oSVM. For example, 52 subjects are assigned to the ME group by FPCA-OF. ME is the most numerous group, but only 34 individuals belong to that intermediate class. There is a high percentage in Sensitivity, but a low percentage in Specificity for FPCA-OF in the ME class. Note that 'equal' is the performance function used in *ordfor* of the package **ordinalForest** that computes OF, which is the appropriate function when we are interested in classifying instances from each class with the same accuracy, regardless of the class sizes. In other words, we have not prioritized the global accuracy, and we have not given preference to the larger class at the expense of a lower classification accuracy with respect to smaller classes. We have taken into account the imbalance of the sample sizes of each class.

The same performance pattern is observed in the results in Table 2. Again, it is beneficial to consider the order because the method of the first (naïve) approach returns the worst performance in terms of accuracy and RPS. In this case, the best accuracy is clearly attained by FPCA-oSVM, i.e. the difference in accuracy with the other methods is wider.

As regards the results in Table 3, FH-FPCA-SVS-LDA is the best method is terms of accuracy and RPS. The second best in terms of accuracy is FPCA-POLR. In this case, FPCA-SVS-LDA, which does not take order into account, provides results that are similar to or better than other methods that consider the ordering information. This shows that naïve methods can also be very competitive. The accuracy of FPCA-SVS-LDA is equal to that attained by FPCA-OF and FPCA-KIOF and better than that of FPCA-oSVM. The RPS value for FPCA-SVS-LDA is the second best. Note the difference between the results for putamens and hippocampi. For right putamens the best method was FPCA-oSVM, which is the worst for left hippocampi in terms of accuracy. Therefore, there is no single method which performs best in all possible datasets, as is the case in nominal multivariate classification.

TABLE 2 Performance measures for several classifiers, for the right putamens data set: accuracy, recall or sensitivity, specificity, precision or positive predictive value, and negative predictive value (NPV) for each class, and RPS. The best value in each column appears in bold.

Method	Accuracy	Sensitivity			Specificity			Precision			NPV			RPS
		E	ME	NE	E	ME	NE	E	ME	NE	E	ME	NE	
FPCA-SVS-LDA	0.3582	0.4	0.4118	0.2222	0.8654	**0.4848**	0.6122	0.4615	0.4516	0.1739	0.8333	0.4444	0.6818	0.3282
FPCA-POLR	0.4328	0.3333	0.5588	0.2778	0.8269	**0.4848**	0.7551	0.3571	0.5278	0.2941	0.8113	0.5161	0.74	0.3151
FH-FPCA-SVS-LDA	0.4179	**0.533**	0.5294	0.1111	0.8462	**0.4848**	0.7143	0.5	0.5143	0.125	**0.8627**	0.5	0.6863	0.3232
FPCA-oSVM	**0.5373**	0.2	**0.8529**	0.2222	**0.9808**	0.2424	0.898	**0.75**	0.537	**0.4444**	0.8095	**0.6154**	**0.7586**	-
FPCA-OF	0.4627	0.0667	**0.8529**	0.0556	0.9423	0.1212	**0.9184**	0.25	0.5	0.2	0.7778	0.4444	0.7258	**0.1935**
FPCA-KIOF	0.4328	0	0.6765	**0.3333**	0.8462	0.4242	0.7755	0	**0.5476**	0.3529	0.7458	0.56	0.76	0.2139

TABLE 3 Performance measures for several classifiers, for the left hippocampi data set: accuracy, recall or sensitivity, specificity, precision or positive predictive value, and negative predictive value (NPV) for each class, and RPS. The best value in each column appears in bold.

Method	Accuracy	Sensitivity			Specificity			Precision			NPV			RPS
		CN	MCI	AD	CN	MCI	AD	CN	MCI	AD	CN	MCI	AD	
FPCA-SVS-LDA	0.75	0.8333	0.3333	0.9	0.8125	0.8636	0.9444	0.7692	0.4	**0.9**	0.8667	0.8261	0.9444	0.0884
FPCA-POLR	0.8214	**1**	0.3333	0.9	0.875	0.9546	**0.8889**	0.8571	0.6667	0.8182	**1**	0.84	0.9412	0.0893
FH-FPCA-SVS-LDA	**0.8929**	**1**	**0.5**	**1**	**0.9375**	**1**	**0.8889**	**0.9231**	**1**	0.8333	**1**	**0.88**	**1**	**0.0593**
FPCA-oSVM	0.6071	0.8333	0.1667	0.6	0.75	0.7727	**0.8889**	0.7143	0.1667	0.75	0.8571	0.7727	0.8	-
FPCA-OF	0.75	**1**	0	0.9	0.6875	**1**	**0.8889**	0.7059	NaN	0.8182	**1**	0.7857	0.9412	0.1467
FPCA-KIOF	0.75	**1**	0	0.9	0.8125	0.9546	0.8333	0.8	0	0.75	**1**	0.7778	0.9375	0.0946

For right hippocampi, the results in Table 4 show that FPCA-OF is the best in terms of accuracy, but FH-FPCA-SVS-LDA is the best in terms of RPS. FH-FPCA-SVS-LDA returns the second best in accuracy, as FPCA-POLR. However, the RPS value of FPCA-POLR is the second worst. In this problem, the naïve method, FPCA-SVS-LDA, is again the worst, in terms of both accuracy and RPS.

TABLE 4 Performance measures for several classifiers, for the right hippocampi data set: accuracy, recall or sensitivity, specificity, precision or positive predictive value, and negative predictive value (NPV) for each class, and RPS. The best value in each column appears in bold.

Method	Accuracy	Sensitivity			Specificity			Precision			NPV			RPS
		CN	MCI	AD	CN	MCI	AD	CN	MCI	AD	CN	MCI	AD	
FPCA-SVS-LDA	0.5714	0.9167	0	0.5	**0.875**	0.8182	0.6667	0.8462	0	0.4545	0.9333	0.75	0.7059	0.2043
FPCA-POLR	0.6786	0.9167	0.1667	0.7	**0.875**	0.8636	0.7778	0.8462	0.25	0.6364	0.9333	0.7917	0.8235	0.1789
FH-FPCA-SVS-LDA	0.6786	**1**	**0.3333**	0.5	**0.875**	0.8182	0.8333	**0.8571**	**0.3333**	0.625	**1**	**0.8182**	0.75	**0.1371**
FPCA-oSVM	0.6071	**1**	**0.3333**	0.3	0.625	0.8182	**0.9444**	0.6667	**0.3333**	**0.75**	**1**	**0.8182**	0.7083	-
FPCA-OF	**0.7143**	0.9167	0	**0.9**	**0.875**	**1**	0.6667	0.8462	NaN	0.6	0.9333	0.7857	**0.9231**	0.1529
FPCA-KIOF	0.6071	0.9167	0	0.6	0.6875	0.9546	0.7222	0.6875	0	0.5455	0.9167	0.7778	0.7647	0.162

6 | CONCLUSION

We have introduced several methodologies for ordinal classification of functional data. This problem has hardly been studied. We have considered FH-FPCA-SVS-LDA, a method with the ordinal binary decomposition approach, which has not been considered until now. We have also considered other methods, namely feature extraction plus augmented binary classification (FPCA-oSVM), feature extraction plus ensembles (FPCA-OF) and kernel-induced ordinal random forests (FPCA-KIOF). They have been analyzed in four real neuroscience data sets. The results confirm that taking into account ordering information improves the performance. We have seen that there is no 'number one' method, but a method can perform better in some data sets and worse in other data sets. This is the reason why having different alternative methodologies to address an ordinal classification problem is a good option.

Although the proposed methodologies have been presented for multiargument and multivariate functional data, they can be used for classical univariate functional data. In fact, as future work, ordinal classification for univariate functional data can be studied. We can extend more ordinal classification methodologies from the multivariate case (see (Gutiérrez et al. 2016) for a survey) to the functional case. Another direction of future work would be to consider more applications, not only in the neuroscience field. Many real-world applications include ordinal classification and ordinal information should not be ignored.

ACKNOWLEDGEMENTS

This work has been partially supported by the following grants: DPI2017-87333-R from the Spanish Ministry of Science, Innovation and Universities (AEI/FEDER, EU) and UJI-A2017-8 and UJI-B2017-13 from Universitat Jaume I.

DATA AVAILABILITY STATEMENT

The data sets and code for reproducing the results are available at `http://www3.uji.es/~epifanio/RESEARCH/orderfda.zip`.

References

Agresti, A. (2002). *Categorical data analysis*. Wiley.

Aguilera, A., & Escabias, M. (2008). Solving multicollinearity in functional multinomial logit models for nominal and ordinal responses. In *Functional and Operatorial Statistics* (pp. 7–13). Springer.

Barahona, S., Centella, P., Gual-Arnau, X., Ibáñez, M., & Simó, A. (2020). Generalized linear models for geometrical current predictors. An application to predict garment fit. *Statistical Modelling*, *0*(0). doi: 10.1177/1471082X19885465

Cardoso, J. S., & Costa, J. F. (2007). Learning to classify ordinal data: The data replication method. *Journal of Machine Learning Research*, *8*(Jul), 1393–1429.

Chung, M. K., Worsley, K. J., Nacewicz, B. M., Dalton, K. M., & Davidson, R. J. (2010). General multivariate linear modeling of surface shapes using surfstat. *NeuroImage*, *53*(2), 491-505.

Cruz, R., Silveira, M., & Cardoso, J. S. (2018). A class imbalance ordinal method for Alzheimer's disease classification. In *2018 International Workshop on Pattern Recognition in Neuroimaging (PRNI)* (pp. 1–4).

Epifanio, I. (2008). Shape descriptors for classification of functional data. *Technometrics*, *50*(3), 284–294.

Epifanio, I., & Ventura-Campos, N. (2011). Functional data analysis in shape analysis. *Computational Statistics & Data Analysis*, *55*(9), 2758-2773.

Epifanio, I., & Ventura-Campos, N. (2014). Hippocampal shape analysis in Alzheimer's disease using functional data analysis. *Statistics in Medicine*, *33*(5), 867-880.

Fan, G., Cao, J., & Wang, J. (2010). Functional data classification for temporal gene expression data with kernel-induced random forests. In *2010 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology* (p. 1-5).

Fan, Y. (2011). Ordinal ranking for detecting mild cognitive impairment and Alzheimer's disease based on multimodal neuroimages and CSF biomarkers. In T. Liu, D. Shen, L. Ibanez, & X. Tao (Eds.), *Multimodal Brain Image Analysis* (pp. 44–51). Berlin, Heidelberg: Springer.

Ferrando, L., Ventura-Campos, N., & Epifanio, I. (2020a). Detecting and visualizing differences in brain structures with SPHARM and functional data analysis. *NeuroImage*, *222*, 117209. doi: https://doi.org/10.1016/j.neuroimage.2020.117209

Ferrando, L., Ventura-Campos, N., & Epifanio, I. (2020b). A neuroimaging data set on problem solving in the case of the reversal error: Putamen data. *Data in Brief*, 106322. doi: https://doi.org/10.1016/j.dib.2020.106322

Frank, E., & Hall, M. (2001). A simple approach to ordinal classification. In *European Conference on Machine Learning* (pp. 145–156).

Gutiérrez, P., Pérez-Ortiz, M., Sánchez-Monedero, J., Fernández-Navarro, F., & Hervás-Martínez, C. (2016). Ordinal regression methods: Survey and experimental study. *IEEE Transactions on Knowledge and Data Engineering*, *28*(1), 127-146.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. Springer, New York.

Hornung, R. (2019). ordinalforest: Ordinal forests: Prediction and variable ranking with ordinal target variables [Computer software manual]. Retrieved from `https://CRAN.R-project.org/package=ordinalForest` R package version 2.4.

Hornung, R. (2020). Ordinal forests. *Journal of Classification*, *37*, 4–17.

Lila, E., Aston, J. A. D., & Sangalli, L. M. (2016). Smooth principal component analysis over two-dimensional manifolds with an application to neuroimaging. *The Annals of Applied Statistics*, *10*(4), 1854–1879.

Millán-Roures, L., Epifanio, I., & Martínez, V. (2018). Detection of anomalies in water networks by functional data analysis. *Mathematical Problems in Engineering*, *2018*(Article ID 5129735), Article ID 5129735.

NCAR - Research Applications Laboratory. (2015). verification: Weather forecast verification utilities [Computer software manual]. Retrieved from `http://CRAN.R-project.org/package=verification` R package version 1.42.

Pierola, A., Epifanio, I., & Alemany, S. (2016). An ensemble of ordered logistic regression and random forest for child garment size matching. *Computers & Industrial Engineering*, *101*, 455 - 465.

Simó, A., Ibáñez, M. V., Epifanio, I., & Gimeno, V. (2020). Generalized partially linear models on Riemannian manifolds. *Journal of the Royal Statistical Society Series C*, *69*(3), 641-661.

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (Fourth ed.). New York: Springer.

Wang, B., & Shi, J. Q. (2014). Generalized Gaussian process regression model for non-Gaussian functional data. *Journal of the American Statistical Association*, *109*(507), 1123-1133.

Wilks, D. (2006). *Statistical methods in the atmospheric sciences*. Academic Press.

How to cite this article: Ferrando L., I. Epifanio, and N. Ventura-Campos (2020), Ordinal classification of 3D brain structures by functional data analysis, *Stat*, *.

Aportació 4:

The neural basis of the reversal error. How a competent solver solves this algebra problem

Ventura-Campos, Ferrando, Epifanio et al. (2020)

The neural basis of the reversal error. How a competent solver solves this algebra problem.

Ventura-Campos, N.[1,2], Ferrando, L.[2], Epifanio, I.[3], Arnau D.[4], González-Calero J.A.[5] and Ávila, C.[2]

[1] *Department of Education and Specific Didactics. Universitat Jaume I, Spain*

[2] *Neuropsychology and Functional Neuroimaging Group, Universitat Jaume I, Spain.*

[3] *Department of Mathematics. Universitat Jaume I, Spain.*

[4] *Department of Didactics of Mathematics, University of Valencia, Spain.*

[5] *Department of Mathematics, University of Castilla, Spain.*

Abstract

Problem solving is a core element in learning mathematics. The reversal error in problem solving occurs when students recognize the information in the statement but are unable to build a correct equation. Functional magnetic resonance images were acquired to identify the neural bases associated with the reversal error. We found brain activation in bilateral fronto-parietal areas in the participants who committed reversal errors (RE-group), and only left fronto-parietal activation in those who did not, suggesting that the RE group needed a greater cognitive demand. Additionally, the results showed brain activation in the right middle temporal gyrus when comparing the RE vs non-RE groups. This activation would be associated with the semantic processing and inhibition processes needed to form non-trivial associations.

Keywords: Neuroeducation, algebra problem solving, reversal error, fMRI, classification.

INTRODUCTION AND AIMS

The reversal error was widely described in Clement (1982), where the subjects had to solve tasks such as: “Write an equation using the variables S and P to represent the following statement:" There are six times as many students as professors at this university. "Use S for the

number of students and P for the number of professors." (Clement, 1982, p. 17). The most common incorrect answer among first-year engineering college students was $P = 6 \cdot S$. The fact that the variables P and S were presented in opposite positions to those they should occupy led to the name of the error.

The persistence of the error led to the scientific community's interest in identifying its cause. In these investigations, both quantitative and qualitative experimental designs were used, but the results did not make it possible to conclusively determine the origin of the error. We present the results of an investigation that uses methodologies from the field of neuroimaging to: (1) describe the brain areas activated when constructing an equation, both correct and incorrect, from a word problem that includes an additive or multiplicative comparison; and (2) determine the neural basis of the reversal error.

Explanatory models for the reversal error

In the study presented in Clement (1982), the difficulty of converting the inequality stated in the statement into an equality is identified as a possible source of the error. In the particular case of the Students-Professor problem, the construction of an equation involves converting a first mental image of the situation, in which there is a ratio of one teacher to six students, into a new image in which the number of teachers is multiplied so that there is a teacher for each student (Clement, 1982). This set of actions that end with writing the equation is called the hypothetical active operation. The transformation of the initial image into the final one would involve an important effort by the working memory (Sweller, 1988).

In fact, Clement (1982) identifies two possible paths that would lead to reversal error: static comparison (with a semantic origin) and word-order matching (with a syntactic origin). In both cases, the error would be the consequence of a simplification of the hypothetical active operation. According to the static comparison, the explanation assumes that the students reach

the mental image in which a teacher and six students are represented, but from this articulation they structure the equation as if it represented an abbreviated explanation of the mental image. Thus 1 professor for every 6 students becomes $1P = 6S$ (where P and S act as abbreviations or labels), and from there to $P = 6S$ by cancelling the 1. The explanation based on word-order matching would imply that the student does not build a mental image of the described situation, and he/she is limited to performing a translation from left to right (six (6) times as many students (S) as (=) professors (P)).

In order to determine which of the two models is the usual source of the error, numerous studies have addressed the problem. Thus, for example, Rosnick (1981) hypothesized that the static comparison could be the consequence of an incorrect interpretation of the letters, where they would be used as labels that represent measurement units (S would be students instead of number of students). Based on this idea, experimental set-ups were employed in which different letters were used to avoid this possible confusion. Thus, Cooper (1986) found that the insertion of a multiplication sign in the equation (e.g., $6 \cdot S$ instead of 6S) leads to a decrease in the incidence of the reversal error. However, this study concluded that the use of variables other than the initial letter corresponding to the quantity's name (for example, using x to represent the number of students instead of S) did not produce variations in the incidence of the error. Fisher (1988) proposed a study in which letters such as Ns were used instead of S. The aim of this experiment was to emphasize that the variable represented the number of students and was not just a label to replace 'students'. However, no relationship was found between the use of a type of notation and the greater or lesser incidence of the reversal error. Similarly, in Soneira, et al. (2018), no differences were found in the appearance of the reversal error when the students were constrained to using written propositions such as "students" and "number of students" instead of letters.

To assess the explanatory potential of the static comparison, several studies compared the

greater or lesser presence of reversal errors depending on whether the contextual clues facilitated the mental construction of a typical situation (for example, in any class it is common for the number of students to be greater than the number of professors). In studies such as those by González-Calero, et al (2015) or Wollman (1983), the results showed that the reversal error rate was not affected by the presence or lack of contextual clues. However, the experimental techniques used did not make it possible to conclude whether this information was really taken into account by the solver.

The experimental set-ups used to analyze the importance of word-order matching in the incidence of the reversal error have been conditioned by the restrictions imposed by natural languages when expressing multiplicative comparisons, at least in English. González-Calero, et al. (2020) carried out a comparison of the reversal error commission in Basque-Spanish bilingual students. In this case, the authors tried to take advantage of the fact that, in the Basque language, the order in which the quantities are presented in a multiplicative comparison cannot produce reversal error due to a literal translation from left to right. The results showed a significant decrease in the reversal error rate when the statements were provided in Basque, which would suggest that the difference in the incidence of the error would be the effect of applying word-order matching.

Indirectly, Fisher et al. (2011) attribute the lower incidence of the reversal error to word-order matching when students are asked to construct the equation using the inverse operation (division in the case of the Students-Professors problem). According to these authors, resorting to linear reading is prevented by forcing the use of the inverse operation. These results were also observed in the investigation by González-Calero et al. (2015).

In short, previous studies seem to show that the origin of the reversal error in different individuals is not necessarily the result of an explanatory model. It is possible that the reversal error is a manifestation of a brain function that is a consequence of inadequate learning.

Neuroeducation and algebraic reasoning

Functional Magnetic Resonance Imaging (fMRI) studies can help to elucidate the role that specific brain regions play during mathematical processing and development. Dehaene (1997) suggests that intuitive understanding of quantities is associated with activity in the intraparietal sulcus (IPS), included in the parietal cortex. Additionally, the parietal cortex participates in various mathematical tasks, from numerical comparison to more complex processing such as proportions or deductive reasoning (Kroger et al., 2008; Vecchiato et al., 2013). De Smedt et al. (2011) also show that different parts of the parietal cortex, such as the bilateral IPS and left angular gyrus, have a crucial role in mental calculation. Moreover, for many researchers, mathematical learning largely involves working memory (Baddeley, 1997), which is associated with frontal areas. The working memory has the capacity to store and manipulate the information, thus allowing the execution of cognitive tasks, such as reasoning, understanding, and problem solving, which are supported by the maintenance and availability of this information (Baddeley, 1997). Studies on human brain development have increased our knowledge about its maturation and how this maturation could depend on or be important in the learning process. In their study, Gogtay et al. (2004) report a dynamic sequence of grey matter brain maturation that goes from posterior to anterior areas, beginning in parietal areas towards the frontal areas and ending in the temporal cortex. Some cerebral regions, such as the prefrontal areas, seem to mature later (Gogtay et al. 2004), and some authors believe these areas are involved in mathematical cognition and other higher-order processes that develop throughout childhood and adolescence (Blakemore, 2012). This could be important in the transition from concrete arithmetic, mostly associated with parietal areas, to the symbolic language of algebra, which involves the fronto-parietal areas. At the age of learning algebra (during adolescence), students need to have developed abstract reasoning skills that allow them to generalize, model, and analyze mathematical equations and theorems (Qin et al., 2004; Lee

et al., 2007; Anderson et al., 2008 and Anderson et al., 2012). Studies such as Qin (2004) show that the brain's greatest receptivity for learning takes place during adolescence, thus being conducive to the teaching of algebra. Luna (2004) coincides with this author, stating that, as adults, we would be limited in our ability to "learn", but not in adolescence, even though a brain can experience structural and functional changes due to a learning process at any age (Bransford, 2003).

According to Kieran (2004), the central processes in algebraic problem solving include analyzing the quantitative relationships between the quantities and modelling the structure of these relationships. Several neuroscience studies with adults have examined the processes used for solving word problems, and how these processes are acquired. The Lee et al. (2007) study presented a problem-solving task and showed activation in brain areas of the prefrontal and parietal cortex when the participant transformed the statement into an equation. This agrees with the modules proposed by Anderson et al. (2008) during the process of solving algebraic problems: 1) a retrieval module responsible for the recovery of algebraic rules and procedures that had already been learned, associated with the prefrontal cortex; 2) an imaginal module associated with the posterior parietal cortex, which is related to the transformation of the language into algebraic equations; 3) finally, a visual module associated with the fusiform gyrus that extracts information about the equation. All these cerebral areas are of great help in analyzing the equations and decoding their information. Therefore, this model is important because it can help us to devise methods to track mental states while solving arithmetic-algebraic problems (Anderson et al., 2012).

MATERIALS AND METHODS

Participants

In this study, participants were 37 students at the *University Jaume I* with ages ranging from

18-26 years. Informed consent was obtained from each participant before the study. Finally, only 20 healthy right-handed (10 women) adults with ages ranging between 18 and 25 years participated in this study. These participants were divided into two groups: the non-RE group (10 subjects, 5 female; mean age: 21.7, SD: 2.95), who responded correctly to more than 50% of the answers on the task; and the RE group (10 subjects, 5 female, mean age: 21.3, SD: 2.36), who failed more than 50% of the answers on the task. The other 17 participants were excluded because: 11 had excessive head movement in the MRI, and six seemed to respond randomly inside the MR.

The study exclusion criteria were the presence of neurological and medical illness, trauma with loss of consciousness lasting more than one hour, and the typical resonance exclusion criteria such as iron prostheses and dental implants.

Experimental paradigm

Participants completed the problem-solving task with the associated reversal error (RE-task), based on the experimental procedure by Lee et al. (2007). This task was adapted to an fMRI block design. Visual stimuli were presented electronically using E-Prime software (Psychology Software Tools, Pittsburgh, PA), professional version 2.0, installed in a Hewlett-Packard portable workstation (screen-resolution 800 × 600, refresh rate of 60 Hz). Participants watched the laptop screen through MRI-compatible goggles (VisuaStim, Resonance Technology, Inc., Northridge, CA, USA), and their responses were collected via MRI-compatible response-grips (NordicNeuroLab, Bergen, Norway). The E-Prime's logfile saved each participant's accuracy and reaction time (RTs) to each stimulus.

RE-task

The task consisted of two conditions (control and experimental). In all, participants completed 24 trials, divided into 8 blocks, 4 for each condition. The entire task lasted 9 min 36 s. Within

each block, participants were presented with the experimental and control conditions (three different statements and responses in each block). On each trial, there was a 1s fixation point, 8s when the statement of the word problem in Spanish was presented, and 3s for the equation response, which was correct 50% of the times. Participants had to respond by indicating whether the equation was correct or incorrect depending on the statement previously presented. The incorrect equation was the reversal error. Participants had to give manual responses with their thumb using a response-grip button. To remove brain activation due to the motor response effect from always pressing the button with the same hand for correct and incorrect answers, 50% of subjects had to give their responses with the right/left hand if the answer was correct/incorrect, and the other 50% had to respond with the right/left hand if the answer was incorrect/correct. The block organization and timing details are presented in Figure 1.

The experimental statements used consisted of 8 items (2 x 2 x 2) with discrete quantities: 1) multiplicative or additive; 2) increasing (times more than, more than) or decreasing (times less than, less than) comparisons[1]; and 3) presence or not of contextual clues that make it possible to determine what quantity is typically greater. The control sentences used the same structure as the experimental condition: 8 items with discrete quantities, where the statement does not involve reversal error because the translation of the statement to the equation is direct (for example, in an academy, the number of tutors (T) plus 9 equals the number of students (A) solution: T + 9 = A).

Participants received oral instructions about how to do the task, and they performed a 5 min practice task before performing the RE-task. In the practice task, participants performed four blocks only with the control condition, in order to become familiar with the stimuli presentation

[1] In Spanish the multiplicative comparison “X times as many as” is typically constructed by using “X veces más que” (with a literal translation “X times more than”) and “X veces menos que” (with a literal translation “X times less than”).

and the response buttons. A similar laptop with the same display features and the same hardware for manual responses was used outside of the scanner. Participants were asked to answer as accurately as possible.

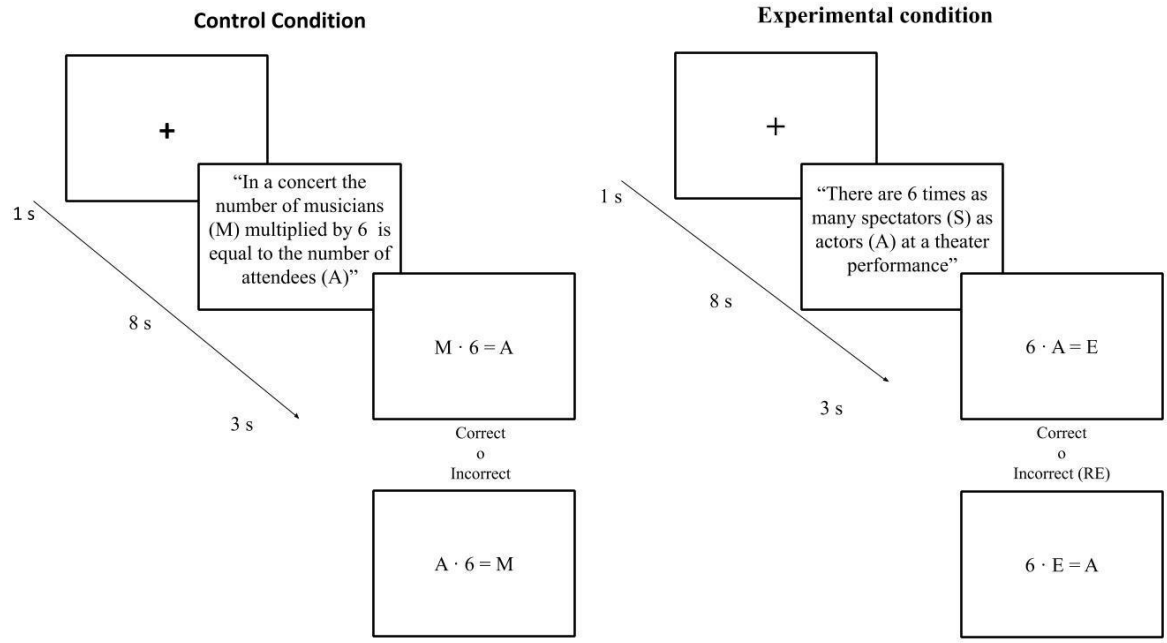


Figure 1. *Schematic representation of the experimental design of the RE-task in fMRI.*

Behavioral task

To collect the behavioral data, we used the computer application similar to the one used in González-Calero et al. (2015), where the participants had to build a mathematical equation for each of the statements presented to them. This process was carried out after the fMRI session.

The task contained a total of 16 statements (2x2x2x2) in Spanish, focusing on: 1) whether the comparisons are multiplicative or additive; 2) whether the comparisons are increasing (times more than, more than) or decreasing (times less than, less than); 3) contextual or non-contextual clues; and 4) discrete or continuous amounts. Thus, subjects could use only multiplication/division or addition/subtraction to express the equation. To construct the

equation, we made it easier by giving them the variables and quantities they had to use (see Figure 2. It should be noted that the response time was unlimited. The App's logfile saved the equations made by each participant and their RTs.

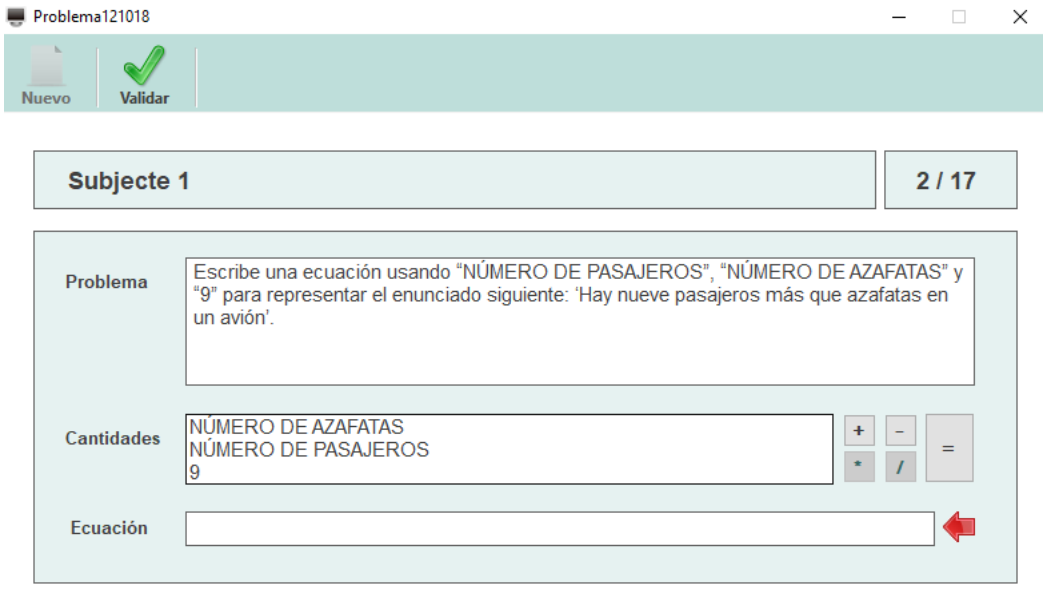


Figure 2. *Application implemented for data collection in studies on the reversal error (González-Calero et al, 2015).* By clicking on each of the variables ("Cantidades") and operation signs (+, - , * , / and =; right to "Cantidades"), participants had to write the equation that corresponded to the statement ("Problema": *Statement:* Write an equation using the "number of passengers", "number of stewardesses" to represent the following statement: "There are nine more passengers than stewardesses on a plane*"*). The rectangle called 'Ecuación' contains the variables they clicked. After that, they validated their equation.

Neuroimaging data acquisition

Functional MRI data were acquired using a 3 Teslas Philips Achiva scanner. For task-fMRI, a gradient-echo T2*-weighted echo-planar image (EPI) MR sequence covering the entire brain was used (TR/TE = 3000/30 ms, matrix = 80 x 80 x 40, flip angle = 90º, voxel size = 3 x 3 x 3.5). A total of 200 volumes were recorded. Before the fMRI sequences, a high-resolution structural T1-weighted MPRAGE sequence was acquired (TR = 8.4 ms, TE = 3.8 ms, matrix size = 320 x 320 x 250, voxel size = .75 x .5 x .8 mm. All the scanner acquisitions were

performed in parallel to the anterior commissure-posterior commissure plane (AC-PC), and they covered the entire brain. Participants were placed in a supine position in the MRI scanner. Their heads were immobilized with cushions to reduce motion degradation, and they were asked to minimize their head movement.

Behavioral analysis

The responses collected from the App were classified as *correct, reversal error,* and *other errors.* In the case of multiplicative comparison (using the example in Figure 2), if we called P (number of passengers) and S (number of stewardesses) the compared quantities, an equation without reversal error would be P = 9 * S, and so any answer that could be reduced to this equation using correct algebraic transformations would be considered correct (e.g., S = P / 9). In a similar way, an equation that could be reduced to S = P * 9 would be classified as a reversal error. Any other equation would be considered another type of error, and the participant who made more than 25% of other errors would be excluded from the sample. In the case of additive comparisons, an analogous criterion was used.

In terms of RTs, participants' performance was processed with the IBM SPSS Statistics software (Version 22 Armonk, New York, USA). A two-sample t-test was conducted to show the differences in RTs between groups in building the equation using the App.

fMRI analysis

Preprocessing

Preprocessing and statistical analyses of fMRI data were conducted with SPM12 (Wellcome Trust Centre for Neuroimaging, London, UK), supported by the MatLab software language. Prior to preprocessing, each subject's fMRI data were aligned to the AC-PC plane by using his/her anatomical image. Then, standard preprocessing was conducted, which included head motion correction, where the functional images were realigned and resliced to fit the mean

functional image. No participant had a head motion of more than 2 mm maximum displacement in any direction or 2° of any angular motion throughout the scan. Afterwards, the anatomical image (T1-weighted) was co-registered to the mean functional image, and the transformed anatomical image was then re-segmented. The functional images were spatially normalized to the MNI (Montreal Neurological Institute, Montreal, Canada) space with 3 mm^3 resolution, and spatially smoothed with an isotropic Gaussian kernel of 8 mm Full-Width at Half-Maximum (FWHM).

Statistical analysis

The experimental effects in each voxel were estimated in the context of the General Linear Model (Friston et al. 1995). In the first-level analysis, we modelled the conditions of interest corresponding to the RE-task (contrast image: experimental > control). The blood- oxygenation level-dependent (BOLD) signal was estimated through the convolution of the stimuli with the canonical hemodynamic response function (HRF). Six motion realignment parameters were included to explain signal variations due to head motion, that is, as covariates of no interest.

The contrast images resulting from the first-level analyses were used for statistical inference in the second-level analyses. In the second-level analysis, a whole-brain one-sample t-test was conducted in order to study the brain regions involved in the main RE-task effects for each group. Next, a two-sample t-test was performed to compare groups at the task level. The statistical criterion was set at $p < .05$, using Family-wise error (FWE) cluster-corrected for multiple comparisons (voxel-level uncorrected threshold of $p < .005$ with a critical cluster size).

Classification analysis

The classification into the RE and no-RE groups or classes was carried out with the R program, where the eight cluster activated brain areas (see Table 3) were used as input variables. Because the database was small, it would be better to use simple models to avoid overfitting. Therefore,

we used simple classifiers such as linear discriminant analysis (LDA) or logistic regression (LG) because they will surely yield better results than more sophisticated classifiers. Nevertheless, we also applied a large number of classifiers and evaluated their predictive capacity, particularly, accuracy.

As Hand (2006) explains, the simplest classical methods can often work better than more recent and sophisticated methods due to uncertainties and arbitrariness. Note that in our problem we considered a threshold for defining the classes. Moreover, different types of people (making more or less reversal errors) can be found within the same class. However, the output variable was finally reduced to two groups.

To estimate the performance, we tested 13 classification methods, using leave-one-out cross-validation (one element was left out each time, the model was adjusted without this data, and finally the prediction was made for that element). Therefore, 20 (the sample size) different models were estimated for each of these methods: LDA (with and without selection of variables), quadratic discriminant analysis (QDA), flexible discriminant analysis (FDA), regularized discriminant analysis (RDA), penalized classification using Fisher's linear discriminant (PLDA), penalized discriminant analysis (PDA), LR, neural networks (NN), support vector machine (with default parameters and cross-validation estimated parameters) (SVM), random forests (RF), and K-neighbors (KN) (Hastie et al., 2009) .

RESULTS

Behavioral results

The data collected with the App and fMRI responses are shown in Tables 1 and 2. These tables show the classification of participants into two groups according to their responses. In both tables, there are two columns, one for the App responses (outside-MR) and the other for those obtained in the fMRI (inside-MR). Table 1 shows the results for the group that commits

reversal error (RE-group). As the table reveals, there are five participants who commit reversal error inside-MR but not outside-MR. The use of fMRI images could impose certain restrictions on task resolution. One of them would be that the participants had a limited time to answer. This could translate into a higher number of participants who commit reversal error than in a pencil and paper environment. As explicitly pointed out in studies such as Wollman (1983) or Pawley et al. (2005), the validation of the proposed equation is a step that competent solvers usually undertake and that can cause certain participants to correct the reversal error a posteriori. However, in our research, the restriction of the task resolution time is aligned with the purpose of determining which neural processes are involved in the equation construction process when the reversal error is committed, and not in post-hoc monitoring processes.

Therefore, to analyze the fMRI, we based the classification of the groups on the data obtained inside-MR because the brains of these five participants work like solvers who commit reversal error. The App responses helped us to select those participants who committed reversal error when building the equation and remove those who made other types of errors that were not the objective of this research. Table 2 shows the results for the group that did not commit reversal error (non RE-group).

Two-sample t-tests conducted on the App's data revealed no significant differences in RTs between groups (t_{18} = .282; p-value = .781).

Table 1

RE-group's data obtained from App and fMRI responses

Subject	**Outside-MR**		**Inside-MR**	
	RE	**Correct**	**RE**	**Correct**
1	*2*	*14*	*19*	*5*
2	*12*	*4*	*20*	*4*

3	*1*	*15*	*18*	*6*
4	*11*	*2*	*17*	*7*
5	*5*	*10*	*20*	*4*
6	*15*	*1*	*22*	*2*
7	*5*	*11*	*22*	*2*
8	*3*	*13*	*15*	*9*
9	*14*	*2*	*18*	*6*
10	*12*	*2*	*19*	*5*

Note. Shaded participants are those who committed reversal error inside-MR but not outside-MR. There were 16 items to respond to outside-MR and 24 items inside-MR.

Table 2

Non RE-group's data obtained from App and fMRI responses

Subject	**Outside-MR**		**Inside-MR**	
	RE	**Correct**	**RE**	**Correct**
1	*0*	*16*	*8*	*16*
2	*3*	*13*	*6*	*18*
3	*0*	*16*	*7*	*17*
4	*0*	*16*	*7*	*17*
5	*1*	*15*	*6*	*18*
6	*0*	*16*	*2*	*22*
7	*0*	*16*	*3*	*21*
8	*2*	*14*	*7*	*17*
9	*0*	*16*	*2*	*22*
10	*0*	*16*	*5*	*19*

Note. There were 16 items to respond to outside-MR and 24 items inside-MR.

fMRI results

A whole brain one-sample t-test analysis showed significant bilateral activations in frontal and

parietal lobes, as well as the Supplementary Motor Area (SMA)/ cingulate gyrus, in the RE group. For the non-RE group, significant activations were obtained only in the left hemisphere in the frontal lobe extending to the insula and parietal lobe (see Table 3 and Figure 3 for details).

The two-sample t-test analysis of the brain differences between groups showed significant brain activation in a single cluster that encompasses the right posterior Middle Temporal Gyrus (pMTG) (x = 51 , y = -43 , z = -4 , with a Z-value = 3.59) in the RE-group when compared to the non-RE group (see Figure 4). No significant differences were found between the non RE-group and the RE group.

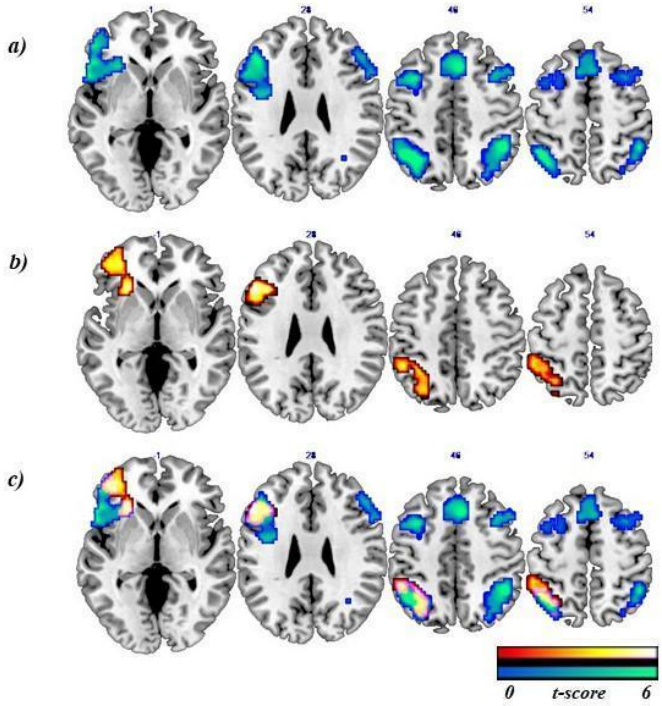


Figure 3. *Main effects of the fMRI RE-task***.** Results of: a) Whole brain one-sample t-test in RE-group (blue-green bar); b) Whole brain one-sample t-test in non RE-group (red-orange bar); c) Common regions between RE and non RE-groups (violet-pink bar). Results were $p < .05$ FWE cluster-corrected using a threshold of $p < .005$ at the uncorrected voxel level, and a cluster size of k = 199 voxels and k = 163 voxels, respectively. Coordinates are in the MNI space. Color bars express t-scores. The left/right of the image corresponds to the left/right brain hemisphere.

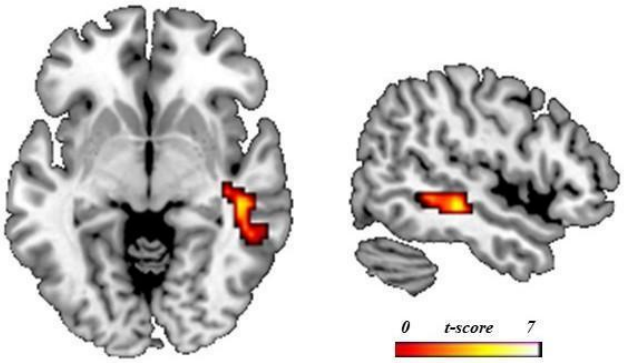


Figure 4. *Differences in brain activation in the fMRI RE-task group analysis***.** Results of two-sample t-test between the RE and non-RE groups. Figure represents the RE-group > non RE-group contrast. Results were $p < .05$ FWE cluster-corrected using a threshold of $p < .005$ at the uncorrected voxel level, and a cluster size of $k = 258$ voxels. Coordinates are in the MNI space. Color bars express t-scores. The left/right of the image corresponds to the left/right brain hemisphere.

Table 3

List of brain activations as a result of RE-task

Regions / AAL	**BA**	**cluster**	**MNI coordinates**	**z-value**
RE-group				
L Inferior parietal lobule	40	392	-36 -55 41	4.46
L Inferior parietal lobule/ Angular			-39 -56 40	4.15
L Superior parietal lobule/Intraparietal Sulcus			-30 -61 47	3.91
L Superior parietal lobule	7		-33 - 70 50	3.81
L Inferior frontal gyrus/frontal inferior triangularis	45	1188	-54 23 5	4.45
L Middle frontal gyrus/frontal inferior triangularis	46		-45 32 23	4.09
L Inferior frontal gyrus/frontal inferior orbital			-39 23 -7	3.84

L Middle frontal gyrus/Frontal middle	10		-39 47 14	3.79
L Anterior insula			-27 20 8	3.78
L Precentral gyrus/frontal inferior operculum	44		-60 17 14	3.64
L Middle frontal gyrus/frontal inferior triangularis	9		-48 14 29	3.16
R Inferior parietal lobule	40	304	45 -49 47	4.38
R Angular gyrus			30 -58 41	3.28
R Superior parietal lobule/ Intraparietal Sulcus			30 -70 50	3.71
R Medial frontal gyrus/Supplementary motor area		425	6 23 47	3.99
L Superior frontal gyrus	6		-15 17 68	3.54
R Cingulate gyrus	32		9 23 35	2.73
R Middle Frontal gyrus/frontal inferior operculum	9	294	57 20 35	3.31
R Middle frontal gyrus	8		51 20 44	3.28
R inferior frontal gyrus/ frontal inferior operculum	45		60 20 20	2.83
Non RE-group				
L Middle frontal gyrus / frontal inferior triangularis	9	163	-51 26 26	4.11
L Inferior frontal gyrus/ Anterior insula		184	-27 26 -4	3.93
L Middle frontal gyrus/ frontal middle orbital			-33 47 -1	3.37
L Inferior frontal gyrus/ frontal middle	10		-39 47 2	3.15
L Inferior parietal lobule	40	237	-54 - 43 53	3.57
L Superior parietal lobule/ Intraparietal sulcus			-30 -61 47	3.05

Note. L Left, R right, BA Brodmann Area. All the analyses were corrected for multiple comparisons FWE cluster-corrected at $p < .05$.

Classifications results

Table 4 shows the LOU results for the 13 classifiers. The best classifier is the FDA, with a correct classification of 16 of the 20 subjects, corresponding to an 80% success rate. The second and third best classifiers are the LDA and LR, both with a 70% success rate, classifying 14 of the 20 subjects correctly. Note that the simplest methods yielded better results.

Table 4

LOU results for the 13 classifiers. The Accuracy column contains the number of correct predictions, whereas the Proportion column contains the success rate.

Method	Accuracy	Proportion
FDA	16	0.8
LDA	14	0.7
LR (variable selection)	14	0.7
NN (Projection Pursuit)	12	0.6
LDA with stepclass	12	0.6
RDA	11	0.55
PLDA	10	0.5
RF	9	0.45
1-KN	9	0.45
QDA	8	0.4
SVM (default parameters)	8	0.4
PDA (Hastie et al. 1995)	8	0.4
SVM (adjusted parameters)	7	0.35

DISCUSSION

The main goals of this study were, first, to investigate the underlying neuronal bases associated with the reversal error phenomenon and, second, find the optimal classification method for these neuronal data to determine whether the activity of the brain areas associated with the two types of solvers could predict the reversal error. Therefore, we aimed to find out whether these areas were able to identify competent solvers. To accomplish these objectives, we separated participants who performed an algebraic problem-solving task with an associated reversal error into two groups (RE-group and non RE-group), depending on their responses given inside-MR. To achieve the first objective, we analyzed the fMRI of participants who performed the RE-task, in order to obtain the main brain effects derived from committing the reversal error. Two main results were obtained. First, the neural activation elicited by the RE-task differed in the two groups when they processed the task. The RE-group showed brain activation in the bilateral inferior and middle frontal gyrus, superior and inferior parietal gyrus including the IPS and angular gyrus, left insula, and SMA/cingulate gyrus. The non RE-group showed activation only in the left hemisphere, in the inferior and middle frontal gyrus, inferior and superior parietal gyrus including the IPS, and the insula. Second, to identify differences between groups in specific problem-solving processes when a reversal error is committed, we compared the RE-group and the non RE-group. We found activation in the right pMTG.

As expected, both groups obtained greater activation in frontal and parietal areas. These regions have previously been associated with processes that have higher working memory or attentional demands (Wager and Smith, 2003, Owen et al., 2005) and with quantitative processing (Dehaene et al., 2003), which occurs in processes involving symbolic algebra. With regard to the role of parietal and frontal regions in algebra problem solving, the results of Danker and Anderson (2007) confirm their involvement, but they also highlight how tightly they are intertwined. Although the parietal areas have greater activation in imaginal operations,

and frontal areas have greater activation in retrieval operations, both regions were involved in both the transformation and retrieval stages of the problem-solving task (Danker & Anderson, 2007).

Regarding the frontal activations, as Owen et al. (2005) also reported, on the one hand, in our study, the middle frontal gyrus (BA 9/46) was found to be involved in monitoring and manipulation within working memory, response selection, and implementation of strategies to facilitate memory. Therefore, it played an essential role in increasing task performance, organization of material before encoding, and verification and evaluation of representations retrieved from long-term memory (Owen, 1997; Petrides,1994; Rowe et al.,2000; Bor et al.,2003,2004; Fletcher et al.,1998; Dobbins et al.,2002 and Rugg et al.,1998). On the other hand, the inferior frontal gyrus (BA 45) has been specifically implicated in a similarly diverse but distinct set of cognitive processes, including the selection, comparison, and judgment of stimuli held in short-term and long-term memory, holding non-spatial information on-line, stimulus selection, and the elaborative encoding of information into episodic memory (Petrides, 1994; Courtney et al., 1997; Goldman-Rakic,1994; Rushworth et al.,1997; Henson et al.,1999; Wagner et al.,1998). Thus, evidence seems to indicate that the bilateral frontal areas are highly involved in processes implicated in the transformation of information embedded in word problems into equations. Furthermore, left frontal activation is frequently associated with verbal working memory tasks (Wager & Smith, 2003). Therefore, this association and the behavioral findings of Lee et al. (2004), who show that verbal working memory tasks predict better algebraic word problem performance, would explain the left lateralized activation of the non-RE group during the performance of the RE-task.

Activation in the parietal cortex has been shown in various mathematical tasks ranging from numerical comparison to more complex processing, such as proportions or deductive reasoning (Kroger et al., 2008; Vecchiato et al., 2013). It is also involved in the implementation of

stimulus response mapping (Andersen & Buneo, 2003; Corbetta & Shulman, 2002; Dreher & Grafman, 2003; Kimberg et al., 2000; Miller & Cohen, 2001; Rushworth et al., 2001) and the storage of working memory contents (Jonides et al.,1997), as well as executive manipulation of acquired facts (Koenigs et al. 2009). Furthermore, the IPS results were expected because this brain area is associated with quantitative processing (Dehaene et al., 2003) on numerosity habituation tasks (Piazza et al., 2004), mental arithmetic tasks involving symbolic versus non-symbolic conditions (Venkatraman et al., 2005), and magnitude comparison processing (Fias et al., 2003), the process used by participants to help solve the RE-task. Activation of the IPS has also been found on algebraic problem-solving tasks that show the brain differences between algebraic equations and mental number line conditions (Terao et al., 2004). Terao et al. (2004) found that the bilateral IPS was activated in the mental number line condition, whereas the algebra equation condition activated the IPS, but largely left lateralized. Regarding our results, we found that the finding by Terao et al. (2004) has great relevance in our study because only the competent solvers, the non RE-group, who selected the correct equation (which coincided with the equation they had previously created in their head), activated the left IPS. On the other hand, the Terao et al. (2004) findings also reveal that the IPS is sensitive to condition on mathematical tasks. Thus, in our results, we can rule out, as in Lee et al. (2007), that the IPS activation merely reflects exposure to numbers, given that similar numeric stimuli were presented in the experimental and control conditions. We found IPS activation when we subtracted the control condition from the experimental condition (experimental > control contrast).

On the one hand, the cingulate gyrus (BA6/32) has been related to error detection (Bush et al. 2000) and integration of information (Devue et al., 2007), and it is often involved in increased effort, complexity, or attention (Callicott et al., 1999; Duncan & Owen, 2000). Arsalidou and Taylor (2011) suggested that the cingulate gyri implement cognitive goals by integrating

available information. On the other hand, the insula are associated with error processing (Hester et al., 2004) and execution of responses (Huettel et al., 2001). Thus, the cingulate gyrus and the insula can work together in initiating attentional control signals and detecting important stimuli (Menon and Uddin, 2010).

The results for the RE-group during the RE-task overlap with the bilateral fronto-parietal network, with the cingulate gyrus and anterior insula, found in the study by Lee et al., (2007; see Table 2) using the two methods studied. Their research studied the similarities and differences in the cognitive processes for representing algebraic word problems with two methods: the model method (Singapore method) and the symbolic method. On the task, the experimental condition required more extensive magnitude comparison (only additive) and working memory engagement than the control condition, as in our RE-task. It should be noted that the Lee et al. (2007) task has the reversal error associated in the statement, but the solutions they propose in choosing the answers facilitate decision-making, thus avoiding making this type of error with the symbolic method. For example, when they say "fewer than", they are expected to choose the subtraction option, which is the correct one, instead of the addition option (see Figure 2 in Lee et al. 2007). They are not given the option that addition (using algebraic transformation) could be a correct answer, or the subtraction option with the variables presented in opposite positions, and, therefore, commit a reversal error. Therefore, the final results include both types of participants, those who make reversal errors and those who do not. For this reason, the RE-group results are very similar to those obtained by Lee et al. (2007) because in the fMRI, the main effects are the union of all the significant group activations. This finding gives greater reliability to our results and reveals the network associated with solving complex algebra word problems.

Several neuroimaging studies have examined algebraic word problems (e.g. Qin, et al. 2004; Lee et al. 2007; Anderson, 2008, 2012), and our findings are consistent with the areas involved

in these problem-solving tasks. However, no fMRI studies have examined the phenomenon of the reversal error. In relation to differences between groups found in the main effects of the RE-task, our results revealed that the non-RE group activated the fronto-parietal network only in the left hemisphere, showing fewer areas needed to perform the RE-task. In previous studies on learning/training, lower activation in areas associated with the task after learning has often been interpreted as an indication of better neural efficiency, which may allow participants to respond by making fewer mistakes (Buschkuehl et al., 2014). Moreover, this effect of decreased activation is typically observed after training on higher cognitive tasks, which could be the case of our RE-task, and lower activation is associated with increased neural efficiency, which means that fewer neurons are needed to give a fast and accurate answer to the task (Kelly et al., 2006). Although our research is not a longitudinal study where learning or training has been carried out, we understand that a formal learning process of algebra has taken place during elementary school, high school, and university, with subjects who perform problem-solving without committing reversal errors and, therefore, are competent solvers, and others who, although having received similar teaching-learning, commit reversal errors when performing the task.

The comparison of the RE and non-RE groups showed activation in the right pMTG in the RE group, compared to the non-RE group, associated with semantic processing. Although the left hemisphere is dominant in language processing, the right temporal cortex processes the novel semantic information and novel meaning of idioms (Faust & Mashal, 2007, Mashal et al., 2008, Pobric et al., 2008), and the right pMTG plays a crucial role in verbal insight problem solving (Bowden & Jung-Beeman, 2003, Jung-Beeman et al., 2004, Zhang et al., 2011, Zhou et al., 2011, Zhao et al. 2014). More specifically, Shen, et al. (2017) point out that "the right pMTG may undergo sustained activation for weak meanings of knowledge nodes or distant conceptual associations that are essential for insight" (p. 363). Parsons and Osherson (2001) conclude that

the right pMTG is used when subjects solve problems by using deductive reasoning and in the review conducted. Therefore, this result seems to show that the right pMTG is involved in processing new semantic information, possibly because the situations presented are not familiar, and the subject has to decode them in order to give them meaning in the translation into algebraic language.

Regarding the goal of classification, after testing with 13 classifiers, the methods that best classify in our case, taking into account that the sample is small, would be the FDA, followed by the LDA and LR, which have the same percentage of success. These are simple methods, and our results confirm Hand's (2006) suggestion that the simplest classical methods often work better than more recent and sophisticated methods, due to uncertainties and arbitrariness, and that this could especially occur in real problems. Moreover, the results show that the brain areas activated and introduced as classification variables could be considered good biomarkers to help to identify competent solvers. Moreover, these results provide valuable information to continue with this educational neuroscience research by using these areas in a functional connectivity study and finding the brain network of competent solvers.

CONCLUSION

Our findings show that, on the one hand, the RE group needed more resources and a greater cognitive demand, and so there was greater activation of the bilateral network associated with working memory, attention processes, and executive function. Furthermore, in this group, the right pMTG was required to understand the statement and translate its meaning into the construction of the equation. On the other hand, the non-RE group showed activation in the left lateralized areas, associated with greater activation during a verbal working memory task involved in better algebraic word problem performance and symbolic processing in the algebra problem solving condition. This result seems to show that these participants are competent solvers and have developed solid algebraic knowledge, in terms of the meaning of the variable

and the logic of the process of constructing an equation.

BIBLIOGRAPHY

Anderson, J. R., Fincham, J. M., Qin,Y., and Stocco, A. (2008) A central circuit of the mind. *Trends in Cognitive Sciences*, 12(4),136 - 143.

Anderson, J. R., Betts, S., Ferris, J. L., and Fincham, J. M. (2012). Tracking children's mental states while solving algebra equations. *Human Brain Mapping 33*, 2650–2665. doi: 10.1002/hbm.21391

Baddeley, A. (1997). Human Memory: Theory and Practice (Revised Edition). Psychology Press, East Sussex.

Blakemore, S. J. (2012). Imaging brain development: The adolescent brain. *NeuroImage*, 61(2), 397- 406.

Bowden, E.M. and Jung-Beeman, M. (2003). Aha! Insight experience correlates with solution activation in the right hemisphere. *Psychonomic Bulletin & Review*, 10, 730-737

Clement, J. (1982) Algebra word problem solutions: Thought processes underlying a common misconception. *Journal for Research in Mathematics Education*, 13(1), 16-30.

Cooper, M. (1986). The dependence of multiplicative reversal on equation format. *Journal of Mathematical Behavior, 5*, 115–120.

Danker, J. F., and Anderson, J R. (2007) The roles of prefrontal and posterior parietal cortex in algebra problem solving: A case of using cognitive modeling to inform neuroimaging data. *Neuroimage*, 35, 1365-1377. http: //doi.org/10.1016/j.neuroimage.2007.01.032

De Smedt, B., Holloway, I. D., y Ansari, D. (2011). Effects of problem size and arithmetic operation on brain activation during calculation in children with varying levels of arithmetical fluency. *NeuroImage*, *57*(3), 771–781

Dehaene, S. (1997). *The Number Sense*. New York, NY: Oxford University Press.

Devue, C., Collette, F., Balteau, E., Degueldre, C., Luxen, A., Maquet, P., and Bredart, S. (2007). Here I am: the cortical correlates of visual self-recognition. *Brain Research*. 1143, 169–

182.

Fisher, K. M. (1988). The students-and-professors problem revisited. *Journal for Research in Mathematics Education, 19*(3), 260-262.

Fisher, K., Borchert, K., & Bassok, M. (2011). Following the standard form: Effects of equation format on algebraic modeling. *Memory & Cognition, 39*(3), 502-515.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.P., Frith, C. D., & Frackowiak, R. S. J. (1994) Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, 2(4), 189-210.

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, 20, 487-506.

Faust, M., and Mashal, N. (2007) The role of the right cerebral hemisphere in processing novel metaphoric expressions taken from poetry: a divided visual field study. *Neuropsychologia*, 45, 860-870.

Gogtay, N., Giedd, J.N., Lusk, L., Hayashi, K.M., Greenstein, D., Vaituzis, A.C.,... Thompson P.M. (2004) Dynamic mapping of human cortical development during childhood through early adulthood. Proc. Natl. Acad. Sci. U. S. A. *101*, 8174–8179

González-Calero, J. A., Arnau, D. & Laserna-Belenguer, B. (2015) Influence of additive and multiplicative structure and direction of comparison on the reversal error. *Educational Studies in Mathematics*, 89, 133-147.

González-Calero, J. A., Berciano, A., & Arnau, D. (2020). The role of language on the reversal error: A study with bilingual Basque-Spanish students. *Mathematical Thinking and Learning, 22*(3), 214–232.

Hand, D. J. (2006). Classifier technology and the illusion of progress, *Statistical Science*, 21, 1–14.

Hastie, T., Buja, A., and Tibshirani, R. (1995). Penalized discriminant analysis. *Annals of Statistics*, 23, 73–102.
Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Series in Statistics, Springer New York.

Hester, R., Fassbender, C., and Garavan, H. (2004). Individual differences in error processing:

a review and reanalysis of three event-related fMRI studies using the GO/NOGO task. *Cerebral Cortex,* 14, 986–994.

Huettel, A.S., Guzeldere, G., and McCarthy, G. (2001). Dissociating the neural mechanisms of visual attention in charge of detection using functional MRI. *Journal of Cognitive Neuroscience* 13, 1006–1018.

Jung-Beeman, M. J, Bowden, E.M., Haberman, J., Frymiare, J.L., Arambel-Liu, S., Greenblatt, R., Reber, P.J., and Kounios J. (2014). Neural activity when people solve verbal problems with insight. *PLoS Biology,* 2, E97

Kieran, C. (2004). Algebraic thinking in the early grades: what is it? *The Mathematics Educator*, 8 ,139-151

Kroger, J. K., Nystrom, L. E., Cohen, J. D., y Johnson-Laird, P. N. (2008). Distinct neural substrates for deductive and mathematical processing. *Brain Res., 1243*, 86– 103. doi:10.1016/j.brainres.2008.07.128

Lee, K., Lim, Z. Y., Yeong, S. H., Ng, S. F., Venkatraman, V., y Chee, M. W. (2007). Strategic differences in algebraic problem solving: neuroanatomical correlates. *Brain Res., 1155*, 163–171.

Mashal, N., Faust, M., Hendler, T. and Jung-Beeman, M. (2008) Hemispheric differences in processing the literal interpretation of idioms: converging evidence from behavioral and fMRI studies. *Cortex,* 44, 848-860

Owen, A.M., McMillan, K.M., Laird, A.R., and Bullmore, E. (2005). N-back working memory paradigm: a meta-analysis of normative functional neuroimaging. *Human Brain Mapping*, 25, 46-59

Parsons, L. M. & Osherson, D. (2001). New evidence for distinct right and left brain systems for deductive versus probabilistic reasoning. *Cerebral Cortex, 11*(10), 954-965. https://doi.org/10.1093/cercor/11.10.954.

Pobric, G., Mashal, N., Faust, M. and Lavidor, M. (2008). The role of the right cerebral hemisphere in processing novel metaphoric expressions: a transcranial magnetic stimulation study. *Journal of Cognitive Neuroscience*, 20, 170-181

Qin, Y., Carter, C. S., Silk, E. M., Stenger, V. A., Fissell, K., Goode, A., Anderson, J.R (2004). The change of the brain activation patterns as children learn algebra equation solving. Proc. Natl. Acad. Sci. U. S. A. 101, 5686–5691

Shen, W., Yuan, Y., Liu, C., & Luo, J. (2017). The roles of the temporal lobe in creative insight: an integrated review. *Thinking & Reasoning, 23*(4), 321-375. https://doi.org/10.1080/13546783.2017.1308885.

Soneira, C., González-Calero, J. A., & Arnau, D. (2018). An assessment of the sources of the reversal error through classic and new variables. *Educational Studies in Mathematics, 99*(1), 43–56.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257-285.

Susac, A., and Braeutigam, S. (2014). A case for neuroscience in mathematics education. *Frontiers in Human Neuroscience*, 8, 314. DOI=10.3389/fnhum.2014.00314

Terao, A., Koedinger, K.R., Sohn, M.H., Qin, Y., Anderson, J.R., Carter, C.S., (2004). An fMRI study of the interplay of symbolic and visuo-spatial systems in mathematical reasoning Proceedings of the Twenty-sixth Annual Conference of the Cognitive Science Society. Lawrence Erlbaum Associates, Mahwah, NJ.

Vecchiato, G., Susac, A., Margeti, S., De Vico Fallani, F., Maglione, A. G., Supek, S., Planinic, M., and Babiloni, F. (2013). High- resolution EEG analysis of power spectral density maps and coherence networks in a proportional reasoning task. *Brain Topografy*. 26, 303–314. doi: 10.1007/s10548-012-0259-5

Wollman, W. (1983). Determining the sources of error in a translation from sentence to equation. *Journal for Research in Mathematics Education, 14*(3), 169-181.

Zhang, M., Tian, F., Wu, X., Liao, S., and Qiu, J. (2011). The neural correlates of insight in Chinese verbal problems: an event related-potential study. Brain Research Bulletin, 84, 210-214

Zhao, Q., Zhou, Z., Xu, H., Fan, W., and Han.,L. (2014) Neural pathway in the right hemisphere underlies verbal insight problem solving. *Neuroscience*, 256, 334-341

Zhou ZJ, Xu HB, Zhao QB, Zhao LL, Liao MJ (2011). The processing of novel semantic association in Chinese: Converging evidence from behavior and fMRI studies. In: The 4th international conference on image and signal processing (CISP 2011), Shanghai. Vol. 3, pp 1588–1592.

Capítol 5

Conclusions

L'objectiu principal de la present tesi era aprofundir en els coneixements sobre l'EI, els quals són prou escassos respecte a com el cervell processa la informació. Al llarg d'ella hem aconseguit conèixer més al voltant de l'EI.

En la primera aportació es van estudiar les diferències volumètriques cerebrals de SG entre subjectes amb EI i no-EI, mitjançant l'ús del mètode VBM. Es va mostrar un major volum de SG en el putamen bilateral en el grup amb EI, associat a la planificació de la resposta. Basant-nos en els nostres resultats on es mostra un major volum de SG en els putamen bilaterals en el grup EI, arribem a la conclusió que el grup EI requereix una gran capacitat de funció executiva per resoldre un problema d'àlgebra, amb el putamen dret i el putamen esquerre jugant un paper molt important en aquest procés. Aquest resultat segueix la línia amb l'estudi de Qin et al. (2004), on es plantejava tant a adolescents com a adults una pràctica sobre resolució de problemes verbals, qui va mostrar que tan sols els adolescents després d'un entrenament en resolució de problemes mostraven un increment d'activació en el putamen esquerre. Este fet suggeria que els adolescents podrien necessitar un major esforç per a realitzar càlculs complexos, recolzant-se en les regions del cervell que no són necessàries en l'exercici dels adults.

Tenint en compte l'estudi del desenvolupament cerebral de Wierenga et al. (2014), que indicava que el volum del putamen disminueix amb l'edat (rang 7-23 anys) estant en procés de maduració durant l'adolescència. Suposem que els subjectes que realitzen l'EI necessiten fer un major esforç en la resolució de problemes, on el putamen pareix no seguir els passos d'una maduració normal, és a dir, no disminueix el seu volum de SG durant el seu desenvolupament des de l'adolescència a l'edat adulta. És per açò que pensem que aquestos subjectes necessiten una major demanda del putamen en la resolució de problemes d'àlgebra durant el seu aprenentatge i desenvolupament. En general, els nostres resultats pareixen demostrar un fort efecte de la maduració cerebral en l'aprenentatge d'àlgebra i la resolució de problemes durant

el desenvolupament.

Amb els resultats de la primera aportació, on mostraven que el putamen era l'estructura cerebral que diferia entre els subjectes EI i els no-EI, ens vàrem centrar en discriminar entre grups mitjançant la pròpia estructura, és a dir, mitjançant la forma del putamen (tant dret com esquerre). Per això vam proposar la metodologia basada en l'ús de la representació per SHARPM d'estructures cerebrals per tal de fer la classificació. Aquesta suposava una millora del que es va presentar en Epifanio i Ventura-Campos (2014).

S'ha demostrat que la nostra metodologia aconsegueix uns bons resultats de classificació en un entorn d'alta dimensió que millorava altres alternatives que hem considerat en una comparativa. Però, a més també hem proposat una metodologia per visualitzar on es troben les diferències entre els grups. És a dir, la metodologia proposada en la tesi no sols obté bons resultats predictius, sinó que també torna resultats interpretables.

Després d'aconseguir aquestos resultats i veure que amb la classificació entre els grups s'obtenia un gran poder predictiu, ens varem preguntar si continuaria havent eixe poder si es tenien en compte tots aquells subjectes que es quedaven en un interval mitjà, és a dir aquells subjectes que encertaven més del 60% de les respostes sense aplegar al 100%. El fet que ens va conduir a planificar esta part de la investigació era que vam veure una gran pèrdua de mostra.

Per aquest motiu es va plantejar fer una discriminació en este cas en tres grups ordenats, encara que serviria per més de tres grups ordenats. Proposàrem quatre nous mètodes estadístics de classificació ordinal per a dades funcionals multivariants i multiarguments. Així, el que es va trobar era que les nostres metodologies tenint en compte l'ordre, milloraven el rendiment a l'hora de fer la classificació, respecte que quan no es considerava l'ordre de la variable eixida.

No obstant això, no es va trobar cap mètode que fora el millor per a diferents problemes, és a dir, es va observar que alguns mètodes podien funcionar millor en alguns conjunts de dades i pitjor en altres. És per això, que és molt important tindre diferents metodologies alternatives per tal d'abordar problemes de classificació ordinal funcional.

Per altra banda, per a complir un dels objectius principals d'aquesta tesi, el qual era conèixer les bases neuronals subjacents associades a l'EI, era necessari veure les àrees que s'activaven durant la resolució d'una tasca que induïra a l'EI. És per això que es va utilitzar la tècnica de RMf, la qual ens va permetre conèixer les activacions funcionals durant la realització de la tasca referida.

Després de la realització d'una tasca amb diferents enunciats de problemes que induïen a l'EI durant l'adquisició de les imatges de RMf, arribem a la conclusió que pareix ser que els nostres resultats mostren que el grup amb EI necessitava més recursos i una major demanda cognitiva, de manera que hi havia una major activació de la xarxa bilateral associada a la memòria

de treball, els processos d'atenció i la funció executiva. A més, requereixen del gir temporal mitjà posterior dret, suggerint una major activació en aquesta àrea amb la finalitat d'entendre l'oració i ser capaços de traduir el seu significat en la construcció de l'equació. D'altra banda, el grup no-EI va mostrar una activació a les àrees lateralitzades a l'esquerra, associades amb una major activació en tasques de memòria de treball verbal, implicada en un millor rendiment de problemes verbals algebraics i processament simbòlic en la condició de resolució de problemes d'àlgebra. Aquest resultat sembla demostrar que aquests participants són resolutors competents i han desenvolupat un sòlid coneixement algebraic, pel que fa al significat de la variable i les equacions.

Ja per últim, i en referència a l'objectiu de trobar el mètode òptim per a fer una classificació a partir de les bases neuronals, s'ha vist que després de provar amb 13 mètodes, es pot concloure que els mètodes que millor classifiquen en el nostre cas, i tenint en compte que és una mostra menuda, serien l'ADF, seguit per l'ADL i la RL, els quals tenen el mateix percentatge d'èxits, justament mètodes senzills, com Hand (2006) suggeria que podia ocórrer en problemes reals.

Amb aquesta tesi no solament aportem nou coneixement a la societat sinó que a més és un primer pas per a començar a implementar noves metodologies educatives que puguen corregir l'EI, ja que aquest error no és sols important dins de l'àmbit acadèmic, sinó que com bé sabem les matemàtiques ens envolten en el nostre dia a dia. Dins de l'àmbit acadèmic, en concret en graus de ciències l'EI pot donar lloc a dificultats en la realització de diferents problemes d'enginyeria, química, física... que són aplicables a problemes reals. Per tant, esta problemàtica es pot trobar posteriorment en l'àmbit laboral.

Per tant, amb aquesta tesi multidisciplinar hem aconseguit aportar nous coneixements a tres grans camps com son la neurociència, la didàctica de la matemàtica, unint-les en un sol com és la *Neuroeducació matemàtica*, i l'estadística on s'han aportat noves metodologies les quals no es poden utilitzar únicament en problemes neuroeducatius, sinó en molts més. La neuroimatge ens ha donat les ferramentes necessàries per a estudiar un problema existent en el camp de l'educació matemàtica, en concret, en la resolució de problemes verbals que indueixen a l'EI, i a la vegada hem pogut utilitzar les dades obtingudes per a provar les noves metodologies estadístiques propostes. Amb açò hem aportat al camp de l'educació més coneixement sobre l'EI, és a dir, aquesta tesi ens dona a conèixer quines àrees son les implicades en el procés de resolució de problemes tant quan es comet l'EI com quan no. Sent el primer estudi que reporta aquesta informació tant important, ja que per a poder ensenyar, primer s'ha de conèixer el cervell, el qual és el factor principal de l'aprenentatge. A més, al camp de les matemàtiques hem aportat noves metodologies de classificació que es poden aplicar tant a problemes neuroeducatius com d'altres tipus, i on s'ha vist que les àrees cerebrals són bons biomarcadors per a fer la classificació.

Capítol 6

Treball futur

Com a treball futur tenim previst realitzar les següents tasques, on algunes d'elles ja s'estan produint.

Una vegada realitzades les investigacions i en vista de que alguns subjectes no feien EI amb l'aplicació similar a la de González-Calero et al. (2015) però sí que en feien durant la RM, pensàrem que tal vegada alguns subjectes comprovaven la equació que escrivien en l'aplicació informàtica abans d'enviar-la, modificant la seua resposta. Aquest fet ens va dur a començar a realitzar millores en l'aplicació per tal de veure si els subjectes modifiquen l'equació que escriuen abans de validar-la. Amb aquestes millores estaríem simulant una tasca del tipus llapis i paper, la qual es recomanada per Wollman (1983) ja que d'aquesta manera podríem saber si realment són resolutors competents des d'un primer moment o no. Agraïm al Dr. David Arnau i al Dr. José Antonio González-Calero, la seua col·laboració.

A més, es vol realitzar una tasca de llenguatge on l'objectiu principal serà veure el procés neuronal de l'EI en la comparació estàtica i la de la coincidència en l'ordre de les paraules. La idea és plantejar dues condicions experimentals amb enunciats comparatius, en el que un es pot resoldre amb traducció literal i lineal i en l'altre caldria llegir tot l'enunciat per poder crear l'equació amb la impossibilitat de traduir-lo linealment. Els enunciats plantejats serien en primer lloc, enunciats de forma lineal ("En un despatx hi han tants PC nous, com bolígrafs vells", on la solució seria: PC nous = bolígrafs vells) i en segon lloc, enunciats no lineals ("En una casa hi ha tantes cadires com taules, sent respectivament blaves i negres", on la solució seria: cadires blaves = taules negres).

La tasca es presentarà seguint els mateixos temps i criteris que la tasca d'EI, és a dir, l'enunciat del problema tindrà una presència en pantalla de 8 s. A continuació, es presentarà la solució expressada com a equació algebraica en la qual els participants hauran de respondre

si la solució presentada és correcta o incorrecta i tindran 3 s per a respondre.

Amb l'objectiu de poder implementar una proposta educativa que ajude a superar l'EI, es realitzarà un estudi on els participants amb aquest error es sotmetran a un procés d'aprenentatge. Aquest estudi estarà dividit en tres parts. Una primera part de pre-entrenament, on tots els participants seran sotmesos a una avaluació conductual i de RMe i RMf (tasca d'error d'inversió, tasca de llenguatge i estat de repòs). Una segona part d'aprenentatge, la qual està prevista que tinga una duració de 3 setmanes. El programa d'aprenentatge serà una adaptació del programa *POETIC* de Kim et al. (2014), el qual està basat en un entorn web interactiu que es va utilitzar per a millorar l'EI en estudiants, mitjançant representacions visuals dels problemes i la notació matemàtica. Basant-nos en aquest estudi crearem la nostra pròpia ferramenta interactiva que ajude a reduir aquest error i ens permeta clarificar quines àrees cerebrals canvien o es modifiquen a partir d'aquest aprenentatge. La tercera i última part serà una post-evaluació on els subjectes seran sotmesos de nou a RMe i la RMf.

A més, s'està realitzant actualment una col·laboració amb el grup d'investigació del Dr. Jorge Sepulcre de la Universitat de Harvard, on estem realitzant un meta-anàlisi per vore les àrees i estructures implicades en el raonament matemàtic. Aquest meta-anàlisi ens servirà per a poder correlacionar-lo amb les àrees que hem trobat.

Amb el mateix grup d'investigació es preveu un altra col·laboració, on, es realitzarà un estudi de les bases neuronals en estat de repòs per a les diferències individuals en estructura i en tasca, l'estudi de les xarxes intrínseques del cervell, així com la capacitat predictiva de les àrees cerebrals associades a l'aprenentatge.

Pel que fa a la part de metodologies estadístiques, com treball futur hi ha diferents vies obertes.

Des del punt de vista pràctic, la metodologia proposada a Ferrando, Ventura-Campos et al. (2020a) es pot aplicar a qualsevol problema de classificació en neurociències on les estructures anatòmiques es puguen expressar amb coeficients SPHARM. Des del punt de vista teòric, es podrien estudiar altres mètodes de selecció de variables. A més, podríem ampliar la metodologia per combinar, no només dades funcionals amb coeficients SPHARM, sinó també característiques multivariants, com ara variables relacionades amb l'educació. En altres paraules, hauríem de definir FPCA per a dades híbrides amb funcions vectorials i multivariants, de manera similar al que van fer Ramsay i Silverman (2005, Cap. 10) per a funcions univariants. Finalment, l'enfocament de la FDA es podria utilitzar no només per a la classificació, sinó també en altres problemes on les dades són estructures cerebrals en 3D.

Pel que fa a les metodologies proposades en Ferrando, Epifanio et al. (2020), podem estendre al cas funcional més metodologies de classificació ordinal des del cas multivariant (vegeu Gutiérrez et al. (2016) per una revisió). Una altra direcció de futur per eixe treball consistiria

a considerar més aplicacions, no només en el camp de les neurociències, i amb funcions univariants. Moltes aplicacions del món real inclouen classificació ordinal i la informació ordinal no s'ha d'ignorar.

Bibliografia

Ames, C., Sagristani, M., Moreno, A. i Moreno-Leoni, Á. (2018). *Estudios interdisciplinarios de historia antigua V*. Córdoba: Universidad Nacional de Córdoba.

Anderson, J. R., Betts, S., Ferris, J. i Fincham, J. (2012). Tracking Children's Mental States While Solving Algebra Equations. *Human Brain Mapping*, *33*(11), 2650 - 2665.

Anderson, J. R., Fincham, J. M., Qin, Y. i Stocco, A. (2008). A central circuit of the mind. *Trends in Cognitive Sciences*, *12*(4), 136 - 143.

Ashburner, J. i Friston, K. (2000). Voxel-Based Morphometry—The Methods. *NeuroImage*, *11*(6), 805 - 821.

Baddeley, A. (1997). *Human Memory: Theory and Practice (Revised Edition)*. Psychology Press, East Sussex.

Blakemore, S. J. (2012). Imaging brain development: The adolescent brain. *NeuroImage*, *61*(2), 397 - 406.

Booth, J., Barbieri, C., Eyer, F. i Paré-Blagoev, E. (2014). Persistent and Pernicious Errors in Algebraic Problem Solving. *Journal of Problem Solving*, *7*, 10 - 23.

Booth, L. (1984). Algebra: Children's Strategies and Errors. *Secondary Mathematics Project*.

Breiman, L., Friedman, J., Olshen, R. i Stone, C. (1984). *Classification and Regression Trees*. Wadsworth; Brooks.

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5 - 32.

Brownlee, J. (2016). Machine Learning Algorithms Mini-Course.

Campos, A. (2010). Neuroeducación: uniendo las neurociencias y la educación en la búsqueda del desarrollo humano. *La educ@ción. Digital Magazine*.

Caviness, V., Kennedy, D., Bates, J. i Makris, N. (1996). The developing human brain: A morphometric profile. A R. W. Thatcher, G. R. Lyon, J. Rumsey i N. Krasnegor (Ed.),

Developmental neuroimaging: Mapping the development of brain and behavior (pp. 3 - 14). Academic Press.

Clement, J. (1982). Algebra Word Problem Solutions: Thought Processes Underlying a Common Misconception. *Journal for Research in Mathematics Education*, *13*(1), 16 - 30.

Clement, J., Lochhead, J. i Monk, G. S. (1981). Translation Difficulties in Learning Mathematics. *The American Mathematical Monthly*, *88*, 286.

Clement, J., Lochhead, J. i Soloway, E. (1980). Positive effects of computer programming on students understanding of variables and equations. *ACM*, 467 - 474.

Cohen, E. i Kanim, S. (2005). Factors influencing the algebra "reversal error". *American Journal of Physics*, *73*(11), 1072 - 1078.

Cooper, M. (1986). The dependence of multiplicative reversal on equation format. *Journal of Mathematical Behaviour*, *5*(2), 115 - 120.

Christianson, K., Mestre, J. i Luke, S. (2012). Practice Makes (Nearly) Perfect: Solving 'Students-and-Professors'-Type Algebra Word Problems. *Applied Cognitive Psychology*, *26*.

Davis, R. B. (1984). Learning mathematics. The cognitive science approach to mathematics education. *London: Croom Helm.*

De Smedt, B., Holloway, I. D. i Ansari, D. (2011). Effects of problem size and arithmetic operation on brain activation during calculation in children with varying levels of arithmetical fluency [Special Issue: Educational Neuroscience]. *NeuroImage*, *57*(3), 771 - 781.

Dehaene, S. (1997). *The Number Sense*. NewYork, NY: Oxford University Press.

Dimitriadou, E., Hornik, K., Leisch, F., Meyer, D. i Weingessel, A. (2017). e1071: Misc Functions of the Department of Statistics (e1071), TU Wien. R package version 1.6-8.

Don, G. E. A. (2011). Secondary School Students' Misconceptions in Algebra.

Donders, F. (1969). On the speed of mental processes. *Acta Psychologica*, *30*, 412 - 431.

Epifanio, I. i Ventura-Campos, N. (2014). Hippocampal shape analysis in Alzheimer's disease using functional data analysis. *Statistics in medicine*, *33 (5)*, 867 - 880.

Farrington, B. (1984). *Ciencia y filosofía en la antigüedad*. Editorial Planeta, S. A.

Ferrando, L. (2018). Mètodes de classificació aplicats a problemes verbals amb errors d'inversió en dades de neuroimatge. *Treball fi de grau. Universitat Jaume I.*

Ferrando, L., Epifanio, I. i Ventura-Campos, N. (2020). Ordinal classification of 3D brain structures by functional data analysis, sotmès.

Ferrando, L., Ventura-Campos, N. i Epifanio, I. (2020a). Detecting and visualizing differences in brain structures with SPHARM and functional data analysis. *NeuroImage*, *222*, 117209.

Ferrando, L., Ventura-Campos, N. i Epifanio, I. (2020b). A neuroimaging data set on problem solving in the case of the reversal error: Putamen data. *Data in Brief*, *33*, 106322.

Fisher, K., Borchert, K. i Bassok, M. (2011). Following the standard form: Effects of equation format on algebraic modeling. *Memory and Cognition*, *39*, 502 - 15.

Fisher, K. (1988). The Students-and-Professors Problem Revisited. *Journal for Research in Mathematics Education*, *19*(3), 260 - 262.

Friedman, J. H. (1989). Regularized Discriminant Analysis. *Journal of the American Statistical Association*, *84*(405), 165 - 175.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D. i Frackowiak, R. S. J. (1994). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, *2*(4), 189 - 210.

Giedd, J., Blumenthal, J., Jeffries, N., Castellanos, F., Liu, H., Zijdenbos, A., Paus, T., Evans, A. i Rapoport, J. (1999). Brain development during childhood and adolescence: a longitudinal MRI study. *Nature neuroscience*, *2*(10), 861 - 863.

Gispert, J., Pascau, J., Reig, S., García-Barreno, P. i Desco, M. (2003). Mapas de estadísticos paramétricos (SPM) en medicina nuclear. *Revista Española de Medicina Nuclear*, *22*, 43 - 53.

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., Nugent, T. F., Herman, D. H., Clasen, L. S., Toga, A. W., Rapoport, J. L. i Thompson, P. M. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences*, *101*(21), 8174 - 8179.

González-Calero, J., Berciano, A. i Arnau, D. (2020). The role of language on the reversal error. A study with bilingual Basque-Spanish students. *Mathematical Thinking and Learning*, *22*(3), 214 - 232.

González-Calero, J. A., Arnau, D. i Laserna-Belenguer, B. (2015). Influence of additive and multiplicative structure and direction of comparison on the reversal error. *Educational Studies in Mathematics*, *89*(1), 133 - 147.

Goswami, U. (2006). Neuroscience and Education: From Research to Practice? *Nature reviews. Neuroscience*, *7*, 406 - 11.

Gutiérrez, P. A., Pérez-Ortiz, M., Sánchez-Monedero, J., Fernández-Navarro, F. i Hervás-Martínez, C. (2016). Ordinal Regression Methods: Survey and Experimental Study. *IEEE Transactions on Knowledge and Data Engineering*, *28*(1), 127 - 146.

Hajnal, J., Myers, R., Oatridge, A., Schwieso, J., Young, I. i Bydder, G. (1994). Artifacts due to stimulus correlated motion in fuctional imaging of the brain. *Magnetic Resonance in Medicine*, (31), 283 - 291.

Hanakawa, T., Honda, M., Okada, T., Fukuyama, H. i Shibasaki, H. (2003). Neural correlates underlying mental calculation in abacus experts: a functional magnetic resonance imaging study. *NeuroImage*, *19*, 296 - 307.

Hand, D. (2006). Classifier Technology and the Illusion of Progress. *Statistical Science*, *21*(1), 1 - 14.

Hastie, T. i Tibshirani, R. (2017). *mda: Mixture and flexible discriminant analysis. R port by Leisch, F., Hornik, K. and Ripley, B. D.* [R package version 0.4-10].

Hastie, T., Tibshirani, R. i Buja, A. (1994). Flexible Discriminant Analysis by Optimal Scoring. *JASA*, 1255 - 1270.

Hastie, T., Tibshirani, R. i Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer New York.

Hastie, T., Buja, A. i Tibshirani, R. (1995). Penalized Discriminant Analysis. *Ann. Statist.*, *23*(1), 73 - 102.

Kim, S.-H., Phang, D., An, T., Yi, J. S., Kenney, R. i Uhan, N. A. (2014). POETIC: Interactive solutions to alleviate the reversal error in student–professor type problems. *International Journal of Human-Computer Studies*, *72*(1), 12 - 22.

Kirshner, D., Awtry, Y., McDonald, J. i Gray, E. (1991). The cognitivist caricature of mathematical thinking: The case of the students and professors problem. *Thirteenth Annual Conference of the North America Chapter of the International Group for the Psychology of Mathematics Education, Blacksburg, Va.*

Kroger, J., Nystrom, L., Cohen, J. i Johnson-Laird, P. (2008). Distinct neural substrates for deductive and mathematical processing. *Brain research*, *1243*, 86 - 103.

Kuchemann, D. (1981). Children's understanding of mathematics. *Journal of Mathematical Behaviour*, *3 (1)*, 102 - 119.

Landy, D. i Goldstone, R. (2007). How abstract is symbolic thought? *Journal of experimental psychology. learning, memory and cognition*, *33(4)*, 720 - 33.

Lee, K., Lim, Z. Y., Yeong, S. H., Ng, S. F., Venkatraman, V. i Chee, M. W. (2007). Strategic differences in algebraic problem solving: neuroanatomical correlates. *Brain Res.*, (1155), 163 - 171.

Liaw, A. i Wiener, M. (2002). Classification and Regression by random forest. *R News*, *2*(3), 18 - 22.

López-Real, F. (1995). How Important is the Reversal Error in Algebra? *Darwin, Australia: MERGA*, 390 - 396.

Mangulabnan, P. A. T. M. (2013). Assessing Translation Misconceptions inside the Classroom: A Presentation of an Instrument and Its Results. *US-China Education Review*, *3*(6), 365 - 373.

Maroco, J., Silva, D., Rodrigues, A., Guerreiro, M., Santana, I. i de Mendonça, A. (2011). Data mining methods in the prediction of Dementia: A real-data comparison of the accuracy, sensitivity and specificity of linear discriminant analysis, logistic regression, neural networks, support vector machines, classification trees and random forests. *BMC Research Notes*, *4*(1), 299.

Mestre, J. (1988). The Role of Language Comprehension in Mathematics and Problem Solving. A R. Cocking i J. Mestre (Ed.), *Linguistic and Cultural Influences on Learning Mathematics* (pp. 201 - 220). Taylor; Francis.

OECD. (2005). *Problem Solving for Tomorrow's World: First Measures of Cross-Curricular Competencies from PISA 2003*. OECD Publishing, Paris.

Ortiz, T. (2009). *Neurociencia y educación*. Alianza Editorial.

Piñeiro, J., Pinto, E. i Díaz-Levicoy, D. (2015). ¿Qué es la resolución de problemas? *Boletín REDIPE*, *4*, 6 - 14.

PISA. (2014). *PISA 2012. Resolución de problemas de la vida real. Resultados de matemáticas y lectura por ordenador. Informe español*. Ministerio de Educación.

Portellano, J. A. (2018). *Neuroeducación y funciones ejecutivas*. Ciencias de la educación preescolar y especial, Madrid.

Qin, Y., S Carter, C., Silk, E., Andrew Stenger, V., Fissell, K., Goode, A. i R Anderson, J. (2004). The change of the brain activation patterns as children learn algebra equation solving. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 5686 - 91.

R Development Core Team. (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.

Radford, L. i André, M. (2009). Cerebro, Cognición y Matemáticas. *RELIME. Revista latinoamericana de investigación en matemática educativa*, *12*(2), 215 - 250.

Ramsay, J. O. i Silverman, B. W. (2005). *Functional Data Analysis* (2nd). Springer.

Robins, G. i Shute, C. (1987). *The Rhind mathematical papyrus: an ancient Egyptian text*. Trustees of the British Museum.

Rosnick, P. i Clement, J. (1980). Clearning without understanding: The effect of tutoring strategies on algebra misconceptions. *Journal of Mathematical Behaviour*, *3 (1)*.

Rosnick, P. (1981). Some misconceptions concerning the concept of variable. *The Mathematics Teacher*, *74*(6), 418 - 450.

Soneira, C., González-Calero, J. A. i Arnau, D. (2018). An assessment of the sources of the reversal error through classic and new variables. *Educational Studies in Mathematics*, *99*(1), 43 - 56.

Sowell, E. R. i Jernigan, T. L. (1998). Further MRI evidence of late brain maturation: Limbic volume increases and changing asymmetries during childhood and adolescence. *Developmental Neuropsychology*, *14*(4), 599 - 617.

Sowell, E. R., Peterson, B., Thompson, P., Welcome, S., L Henkenius, A. i W Toga, A. (2003). Mapping cortical change across the lifespan. *Nature neuroscience*, *6*, 309 - 15.

Sowell, E. R., Thompson, P. M., Holmes, C. J., Batth, R., Jernigan, T. L. i Toga, A. W. (1999). Localizing Age-Related Changes in Brain Structure between Childhood and Adolescence Using Statistical Parametric Mapping. *NeuroImage*, *9*(6), 587 - 597.

Speedie, S., Treffinger, D. i Houtz, J. (1976). Classification and evaluation of problem-solving tasks. *Contemporary Educational Psychology*, *1*(1), 52 - 75.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B. i Joliot, M. (2002). Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain. *NeuroImage*, *15*(1), 273 - 289.

Vecchiato, G., Susac, A., Margeti, S., De Vico Fallani, F., Maglione, A., Supek, S., Planinic, M. i Babiloni, F. (2013). High-Resolution EEG Analysis of Power Spectral Density Maps and Coherence Networks in a Proportional Reasoning Task. *Brain topography*, *26*.

Venables, W. N. i Ripley, B. D. (2002). *Modern Applied Statistics with S* (Quarta ed.). Springer.

Ventura-Campos, N., Ferrando, L., Epifanio, I., Arnau, D., González-Calero, J. i Ávila, C. (2020). The neural basis of reversal error. How a competent solver solves this algebra problem, sotmès.

Ventura-Campos, N., Ferrando, L., Miró-Padilla, A. i Ávila, C. (2020). The brain-anatomy changes in learning algebra problem solving. Studying the reversal error, sotmès.

Weihs, C., Ligges, U., Luebke, K. i Raabe, N. (2005). klaR Analyzing German Business Cycles. A D. Baier, R. Decker i L. Schmidt-Thieme (Ed.), *Data Analysis and Decision Support* (pp. 335 - 343). Springer-Verlag.

Wierenga, L., Langen, M., Ambrosino, S., van Dijk, S., Oranje, B. i Durston, S. (2014). Typical development of basal ganglia, hippocampus, amygdala and cerebellum from age 7 to 24. *NeuroImage*, *96*, 67 - 72.

Winer, B., Brown, D. i Michels, K. (1991). *Statistical Principles in Experimental Design*. McGraw-Hill.

Witten, D. (2015). *penalizedLDA: Penalized Classification using Fisher's Linear Discriminant* [R package version 1.1].

Witten, D. i Tibshirani, R. (2011). Penalized classification using Fisher's linear discriminant. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *73*(5), 753 - 772.

Wollman, W. (1983). Determining the Sources of Error in a Translation from Sentence to Equation. *Journal for Research in Mathematics Education*, *14*(3), 169 - 181.

Xin, Y. P., Wiles, B. i Lin, Y.-Y. (2008). Teaching Conceptual Model-Based Word Problem Story Grammar to Enhance Mathematics Problem Solving. *The Journal of Special Education*, *42*(3), 163 - 178.

Yandell, B. (2017). *Practical Data Analysis for Designed Experiments*. CRC Press.