

COMPARACIÓ DE DIFERENTS ESTRATÈGIES AVALUATIVES USADES EN LA DOCÈNCIA TEÒRICA DE L'ASSIGNATURA ANATOMIA I FISIOLOGIA HUMANES II DE LA LLICENCIATURA DE FARMÀCIA

Josep Queralt, Francisco J. Pérez-Cano, Carme Pelegrí, Antoni Díez-Noguera, Rut Ferrer, Carles Garriga, Emília Juan, Jordi Vilaplana.

Departament de Fisiologia. Facultat de Farmàcia. Universitat de Barcelona.

Av. Joan XXIII s/n. 08028 Barcelona.

e-mail: josepqueralt@ub.edu

Rebut: maig de 2006. Acceptat: desembre de 2006

ABSTRACT

Human Anatomy and Physiology II is a core subject from Pharmacy syllabus of the University of Barcelona, which is coursed in the third semester of the first cycle. The theoretic exam from this subject includes short-answer questions, with a pre-given space to develop the answer, and multiple choice questions. In the understanding that the evaluation with short-answer questions better reflects the integrative component of learning than multiple choice questions, which mainly evaluate memorized data, the aim of the present study consists in obtaining information about the degree of memoristical and integrative learning reached by the students of this subject. Considering the marks obtained by students in both evaluative strategies, a positive and significant correlation is observed between short answers and multiple choice questions. It is also appreciated that the punctuation obtained in multiple choice questions is generally higher than that obtained in short-answer questions. Thus, students show more difficulty in developing a short answer than in solving multiple choice questions. In the basis of the results of the present study, we find very positive the effort of the teachers of Human Anatomy and Physiology of the Faculty of Pharmacy in promoting the integrative learning of their students.

KEY WORDS: Evaluation, multiple choice, short-answer questions.

RESUM

L'Anatomia i Fisiologia Humanes II és una assignatura troncal de la llicenciatura de Farmàcia de la Universitat de Barcelona que s'imparteix durant el tercer semestre del primer cicle. L'examen teòric d'aquesta assignatura inclou preguntes de desenvolupament amb un espai limitat per a la resposta i preguntes d'elecció múltiple. Entenent que l'avaluació amb

preguntes de desenvolupament reflexa millor el component d'aprenentatge integrador que el basat en preguntes d'elecció múltiple, on majoritàriament s'avaluen les dades que els alumnes han memoritzat, l'objectiu de l'estudi és obtenir informació sobre el grau d'aprenentatge memorístic i integrador assolit pels alumnes de l'assignatura. Analitzant les puntuacions que treuen els alumnes en una i altra estratègia avaluativa, s'observa una correlació positiva i significativa entre elles. S'observa també que les puntuacions obtingudes pels alumnes en les preguntes d'elecció múltiple son generalment superiors a les que treuen mitjançant les preguntes curtes de desenvolupament. En general, doncs, els alumnes mostren més dificultat en desenvolupar una resposta que en solucionar les preguntes d'elecció múltiple. En base als resultats obtinguts, cal valorar positivament l'esforç que estan fent actualment els professors de les assignatures d'Anatomia i Fisiologia Humanes de l'Ensenyament de Farmàcia per tal de potenciar l'aprenentatge integrador d'aquests alumnes.

PARAULES CLAU: Avaluació, elecció múltiple, preguntes curtes de desenvolupament

INTRODUCCIÓ

L'avaluació permet gestionar de forma dinàmica la qualitat dels aprenentatges dels estudiants, i tot i que s'observa l'absència d'un model global d'avaluació a nivell universitari, tots ells han de complir determinades funcions i objectius en el binomi ensenyament-aprenentatge. En aquest sentit, les funcions bàsiques de l'avaluació fan referència a una vessant diagnòstica (detecció de la presència de possibles problemes d'aprenentatge), formativa (amb retroalimentació ininterrompuda per tal de millorar de manera continuada tots els processos implicats) i sumativa (visió global de l'assoliment dels aprenentatges expressats en forma de qualificacions). Les tècniques d'avaluació inclouen entre els seus objectius: facilitar i millorar l'aprenentatge dels estudiants, comprovar l'assoliment dels objectius docents i de les competències bàsiques, així com valorar personalment els estudiants per tal d'optimitzar la docència (Arboix i col., 2003).

La Anatomia i Fisiologia Humanes (AFH), és una assignatura troncal de la llicenciatura de Farmàcia que pretén donar una formació sòlida sobre l'estructura i funció del cos humà que faciliti posteriorment la comprensió de les diferents patologies i la seva terapèutica farmacològica. En la llicenciatura de Farmàcia de la Universitat de Barcelona (UB), aquesta assignatura s'imparteix en dos cursos consecutius: Anatomia i Fisiologia Humanes I i II. Concretament, AFH-I s'imparteix en el segon semestre del primer cicle i inclou continguts generals d'estructura i regulació, sistema nerviós, sang i sistema immunitari. L'AFH-II s'imparteix durant el tercer semestre del primer cicle i inclou com a blocs temàtics els sistemes cardíoc-vascular, respiratori, renal, digestiu, endocrí i reproductor.

L'avaluació dels aprenentatges de l'assignatura troncal AFH-II es realitza a partir d'un examen teòric i un de pràctic. Ambdós exàmens es basen en models avaluatius amb un enfoc metodològic de tipus tradicional, és a dir, basats en la recollida d'informació avaluativa a través de proves objectives, plantejades pel professor, per tal de comprovar l'assoliment dels objectius inicials. En el procés de

construcció d'ambdues avaluacions objectives s'han tingut en compte aspectes fonamentals com selecció de continguts, redacció de les preguntes, correcció, puntuació i presentació de la prova (Arboix i col., 2003). En aquest sentit, les preguntes són representatives dels continguts i mesuren objectius proposats, les preparen els diferents professors que han impartit cada part de l'assignatura, i són sotmeses a la crítica de tots els professors integrants del cos docent de l'assignatura.

Mentre que l'examen pràctic està dirigit a l'avaluació de coneixements pràctics i d'habilitats, l'examen teòric es basa en un procediment mixte d'avaluació que inclou 6 preguntes de desenvolupament amb un espai limitat per a la resposta i 60 preguntes d'elecció múltiple amb resposta única (4 respostes possibles i una única resposta correcta). Així, l'estructura del procediment avaluatiu de continguts teòrics de l'assignatura combina un nombre important de qüestions de tipus test amb qüestions de resposta oberta d'extensió limitada o breu, donat que de forma general es considera que l'èxit de l'aprenentatge pot apreciar-se d'una manera més acurada si el nombre de preguntes o tasques proposades en l'examen és elevat (Mateo, 2000). D'altra banda, tot i que cada tipus d'avaluació captura un nivell d'informació avaluativa diferent i complementària, s'entén que l'avaluació amb preguntes de desenvolupament reflexa millor el component d'aprenentatge integrador que l'avaluació basada en preguntes d'elecció múltiple, on majoritàriament s'avaluen les dades que els alumnes han memoritzat.

En aquest sentit, a partir de les qualificacions obtingudes mitjançant les diferents estratègies avaluatives utilitzades, el present estudi pretén obtenir informació sobre el grau d'aprenentatge memorístic i integrador assolit pels alumnes de l'assignatura. S'analitzarà també l'efecte de les variables gènere, horari i convocatòria sobre la nota final obtinguda per cada alumne.

METODOLOGIA

Per tal d'abordar els objectius plantejats, s'han analitzat un total de 530 exàmens realitzats per alumnes de l'esmentada assignatura durant el curs 2004-2005. D'aquests 530 exàmens, 380 exàmens han estat realitzats per alumnes que han cursat la part teòrica de l'assignatura en horari de matí i 185 que l'han cursada en horari de tarda. 424 exàmens han estat realitzats per persones del gènere femení, predominant en la llicenciatura de farmàcia, i 106 per persones del gènere masculí. 389 exàmens corresponen a la primera convocatòria i 141 a la segona. Cal fer constar que en els grups de matí hi havia un percentatge de 29,5% d'alumnes repetidors i que per tant havien cursat prèviament l'assignatura, mentre que en els grups de tarda aquest percentatge era inferior (25%).

Per a cada alumne examinat s'ha obtingut el valor de tres variables. La variable SPD (sumatori

preguntes desenvolupament) és el resultat de sumar les puntuacions obtingudes amb les preguntes obertes de desenvolupament i la variable SEM (sumatori elecció múltiple) s'obté sumant les puntuacions de les preguntes d'elecció múltiple de resposta única. Com que en l'examen el pes de les preguntes d'elecció múltiple era diferent al de les preguntes de desenvolupament, en SPD i SEM s'ha corregit el desequilibri existent, quedant tant SEM com SPD amb un rang de puntuació de 0 a 100. La variable NF (nota final de l'examen) s'ha redefinit com la mitjana entre SPD i SEM, de manera que les dues estratègies avaluatives tindran igual pes en NF. Per a l'anàlisi estadística de les dades s'ha utilitzat l'ANOVA, s'han fet estudis de correlació entre variables i s'ha usat el test de la t de Student per a dades aparellades mitjançant el programa STATISTICA (Stat Soft®).

RESULTATS

A la Figura 1 es poden observar les gràfiques amb les distribucions percentuals de les notes que han obtingut els alumnes en base a les diferents estratègies avaluatives i en base a la NF. En base a SEM es pot observar una distribució de puntuacions àmplia, que inclou des del rang de puntuacions que va de 0 a 10 fins al que va de 90 a 100, essent el rang predominant (25 % d'alumnes) el dels alumnes que obtenen una nota compresa entre 40 i 50 (Figura 1A). En el cas de les puntuacions obtingudes amb SPD (Figura 1B), el rang amb més percentatge d'observacions (26 %) és el dels alumnes que obtenen una puntuació de 0 a 10. En els següents rangs s'observa gradualment una disminució de percentatges. Les puntuacions més altes, en base a SPD, es troben situades en el rang de puntuacions que va de 60 a 70, i el presenten només un 4% dels exàmens analitzats. Cap alumne ha obtingut més de 70 punts en base a SPD. Analitzant les puntuacions en base a la NF, s'observa una distribució compresa entre els rangs que van de 0 a 10 (2,6%) i el que va de 80 a 90 (0,6%). El pic màxim es troba en el rang de 30 a 40, amb un 24,7 % dels exàmens analitzats.

Analitzant la influència dels factors convocatòria, gènere i horari sobre la variable NF, l'ANOVA indica que el factor gènere no té una influència significativa sobre NF, mentre que els factors convocatòria i horari sí que la tenen ($F_{1,522}=13,96$, $p<0,01$ i $F_{1,522}=6,19$, $p<0,05$ respectivament). S'observa, a més a més, que la interacció entre aquests dos factors és significativa ($F_{1,522}=5,61$, $p<0,05$). Aquest anàlisi, juntament amb les comparacions posteriors amb el test de Scheffé, indiquen que els alumnes dels grups de matí obtenen notes finals superiors que els de la tarda, que en la primera convocatòria els resultats són millors que en la segona i que la nota obtinguda pels grups de matí a la primera convocatòria és superior a l'esperada si els efectes de l'horari i la convocatòria fossin directament additius. En la Figura 3 es mostra la taula dels resultats de l'ANOVA i es poden apreciar les mitjanes (\pm error estàndard i $\pm 1'96$ *error estàndard) de la NF dels subgrups que s'obtenen amb els factors convocatòria i horari. S'observa especialment que són els alumnes de primera convocatòria i

amb horari de matí els que obtenen millor nota.

Les tendències observades analitzant NF són molt similars a les que s'obtenen repetint l'anàlisi amb la variable dependent SEM (Figura 3). L'ANOVA indica també que el factor gènere no té una influència significativa sobre NF, mentre que els factors convocatòria i horari sí que la tenen ($F_{1,522}=14,23$, $p<0,01$ i $F_{1,522}=8,6$, $p<0,01$ respectivament). S'observa, també, que la interacció entre aquests dos factors és significativa ($F_{1,522}=6,24$, $p<0,05$). A l'igual com passava per la variable NF, de forma general, les notes obtingudes en primera convocatòria són superiors a les de la segona i en horari de matí són millors que en horari de tarda, i especialment són els alumnes dels grups de matí i en primera convocatòria els que treuen notes molt superiors a les que s'obtenen en els altres grups.

Analitzant la variable SPD (Figura 4), els resultats difereixen dels obtinguts en els anàlisis anteriors. No s'obtenen diferències significatives pel factor gènere i s'obtenen diferències significatives degudes al factor convocatòria, però no se n'obtenen pel factor horari ni per la interacció d'aquests dos. Es mostra també la taula de resultats de l'ANOVA i la gràfica de les mitjanes (\pm error estàndard i $\pm 1'96$ *error estàndard) de la SPD obtinguda pels diferents subgrups en base als factors convocatòria i horari. En aquest cas, en primera convocatòria les notes són superiors a la segona, i no hi ha diferències significatives pel factor horari. Cal tenir present, però, que en la distribució de notes la variable SPD presentava un biaix cap a l'esquerra respecte les altres variables (SEM i NF), acumulant-se les puntuacions baixes i quedant disminuïda la resolució de l'anàlisi.

Pel que fa a la possible relació existent entre les variables SEM i SPD, aquesta ha estat analitzada mitjançant estudis de correlació i mitjançant la t de Student per a dades aparellades.

L'estudi de la correlació entre les variables SEM i SPD ens indica que entre les dues variables estudiades el coeficient de correlació $r(X,Y)$ és de 0,62, i estadísticament significatiu ($p<0,01$). De forma general podem afirmar, per tant, que quan un alumne té un valor de SEM més elevat que un altre, tendirà a tenir el valor de SPD també més elevat que l'altre. En la figura 5 es poden observar els resultats de l'anàlisi i la gràfica de SEM vs SPD.

Mitjançant la t de Student per a dades aparellades observem que, significativament, la variable SEM té valors superiors als de SPD ($n=530$, $t=36,22$, $p<0,01$). La diferència mitjana entre SEM i SPD és de 25'09 punts (Figura 6). En general, doncs, els alumnes mostren més dificultat en contestar les preguntes de desenvolupament que les preguntes d'elecció múltiple. En la mateixa figura 6 es mostra la distribució de les diferències obtingudes fent per cada alumne la resta SEM – SPD. S'observa que en pràcticament tots els casos (504 de 530 alumnes) els resultats de la nota SEM és superior a la de

SPD. Analitzant mitjançant el test de l'ANOVA els efectes dels factors convocatòria, horari i gènere en la nova variable SEM – SPD, s'observa que no hi ha cap factor que tingui un efecte significatiu ni tampoc hi ha cap interacció positiva (Figura 7).

En l'assignatura Anatomia i Fisiologia Humanes II de la Llicenciatura de Farmàcia de la Universitat de Barcelona es pot concloure, doncs, que els alumnes obtenen millor puntuació mitjançant l'avaluació amb preguntes d'elecció múltiple de resposta única que amb preguntes obertes de desenvolupament, independentment de quina sigui la convocatòria, horari i gènere.

CONCLUSIONS

- Els factors convocatòria i horari tenen influència en la puntuació final i en les puntuacions de les preguntes d'elecció múltiple de resposta única. Per a les dues variables, en la primera convocatòria les puntuacions són superiors a la segona i els alumnes d'horari de matí treuen millors puntuacions que els de tarda.
- El factor gènere no influeix en les puntuacions obtingudes, ja sigui en la puntuació final, en la suma de puntuacions de les preguntes d'elecció múltiple de resposta simple o en la suma de puntuacions de preguntes obertes de desenvolupament.
- Els alumnes obtenen millors resultats quan són avaluats en base a preguntes d'elecció múltiple que quan ho són en base a preguntes curtes de desenvolupament, la qual cosa posa de manifest que cal potenciar en els alumnes la visió integradora dels diferents coneixements que assoleixen.
- Malgrat treure millors resultats en base a les preguntes d'elecció múltiple de resposta única que en base a les preguntes obertes de desenvolupament, la correlació existent entre les dues variables indica que els alumnes que treuen millor puntuació amb un tipus d'estratègia avaluativa també tendeixen a obtenir-la en l'altre.
- En el procés d'avaluació cal tenir present que el pes relatiu de les preguntes de desenvolupament i d'elecció múltiple influirà significativament en les puntuacions finals que obtindran els alumnes, essent aquestes puntuacions inferiors quan més elevat sigui el pes de les preguntes de desenvolupament.
- Cal valorar positivament l'esforç que estan fent actualment els professors de l'assignatura per tal de potenciar l'aprenentatge integrador d'aquests alumnes mitjançant l'ús de noves estratègies docents i aprofitant les noves tecnologies de la informació i la comunicació.
- La utilització en els propers cursos d'exàmens de tipus mixt, amb preguntes d'elecció múltiple de resposta única i preguntes obertes de desenvolupament, permetran observar si l'esforç que realitzen els professors es tradueix en una millora de la formació integradora per part de l'alumne.
- La presència en l'examen d'avaluació de preguntes obertes de desenvolupament pot constituir pels alumnes un estímul positiu des del punt de vista de la seva formació integradora.

FIGURES

Figura 1.- Gràfiques amb les distribucions percentuals de les notes que han obtingut els alumnes en base a SEM (A), en base a SPD (B) i en base a la NF (C).

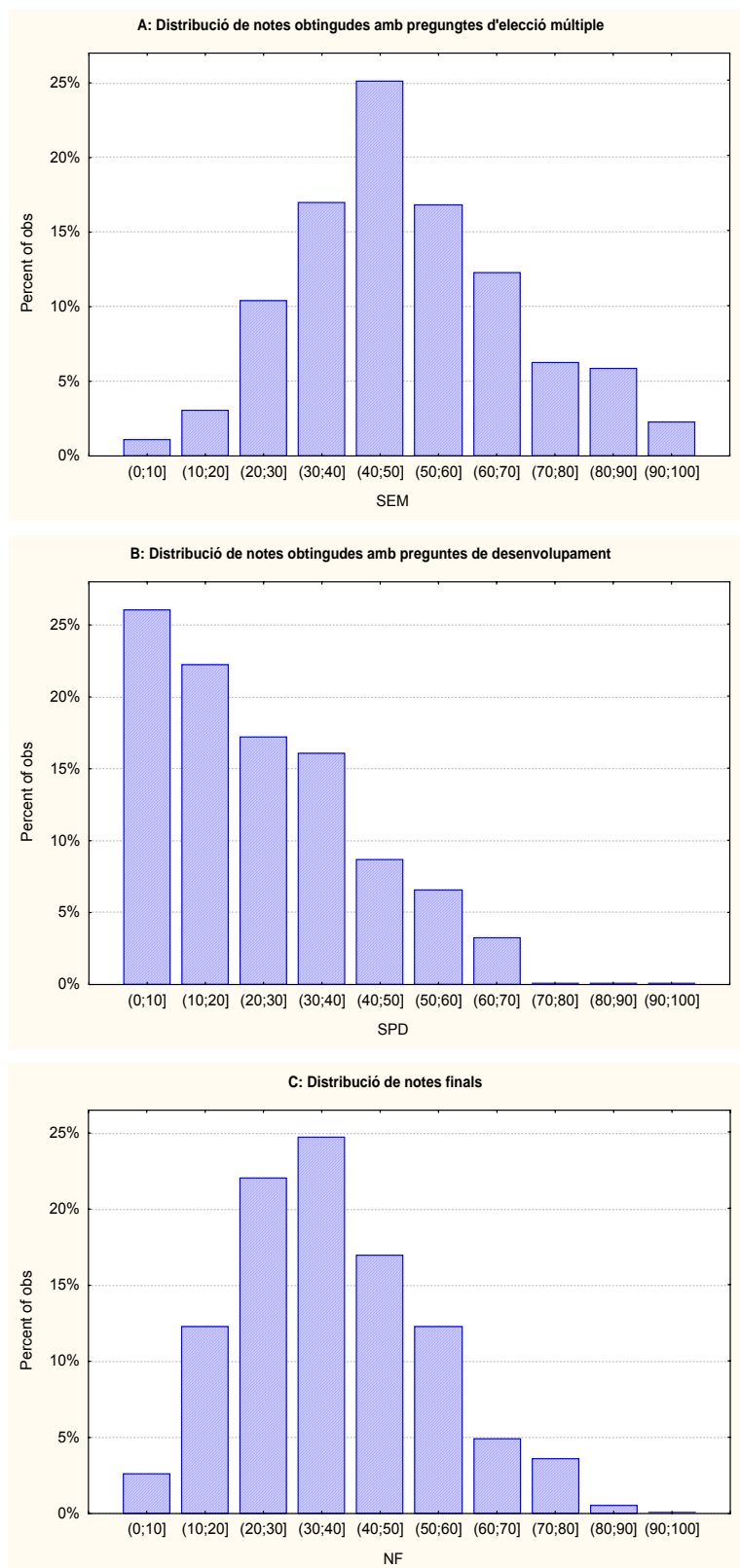


Figura 2.- Taula resum de l'ANOVA realitzat amb la variable independent NF i els factors convocatòria, gènere i horari. Gràfica descriptiva de les mitjanes de NF segons els subgrups generats per les variables convocatòria i horari.

Summary of all Effects; design: (edusfarm.sta)						
GENERAL MANOVA						
1-CONVOCAT, 2-GENERE, 3-HORARI						
Effect	df Effect	MS Effect	df Error	MS Error	F	p-level
1	1	3270,423	522	234,2319	13,96233	,000207
2	1	779,100	522	234,2319	3,32619	,068756
3	1	1450,915	522	234,2319	6,19435	,013126
12	1	3,546	522	234,2319	,01514	,902118
13	1	1315,593	522	234,2319	5,61663	,018153
23	1	227,155	522	234,2319	,96979	,325189
123	1	28,240	522	234,2319	,12057	,728561

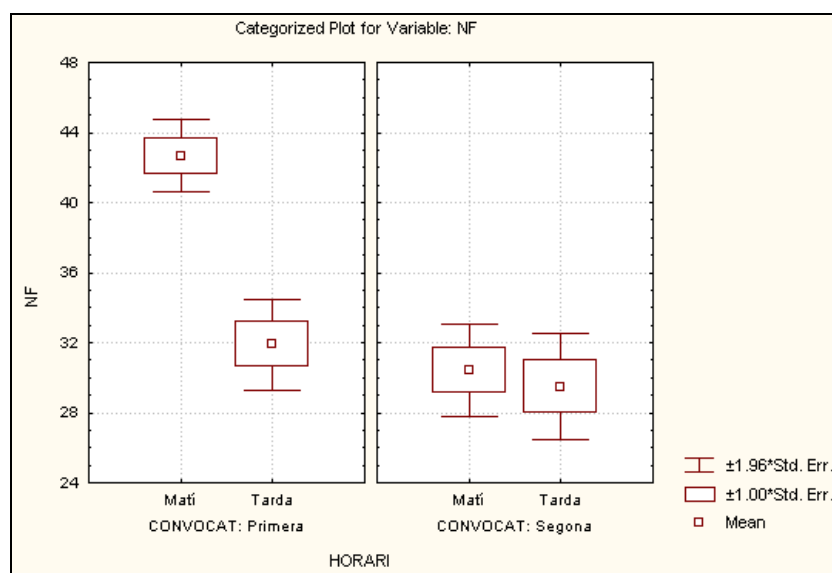


Figura 3.- Taula resum de l'ANOVA realitzat amb la variable independent SEM i els factors convocatòria, gènere i horari. Gràfica descriptiva de les mitjanes de SEM segons els subgrups generats per les variables convocatòria i horari.

Summary of all Effects; design: (edusfarm.sta)						
GENERAL MANOVA						
1-CONVOCAT, 2-GENERE, 3-HORARI						
Effect	df Effect	MS Effect	df Error	MS Error	F	p-level
1	1	4324,902	522	303,8543	14,23347	,000180
2	1	941,052	522	303,8543	3,09705	,079020
3	1	2616,841	522	303,8543	8,61216	,003486
12	1	86,099	522	303,8543	,28336	,594737
13	1	1896,537	522	303,8543	6,24160	,012785
23	1	173,621	522	303,8543	,57139	,450046
123	1	34,646	522	303,8543	,11402	,735748

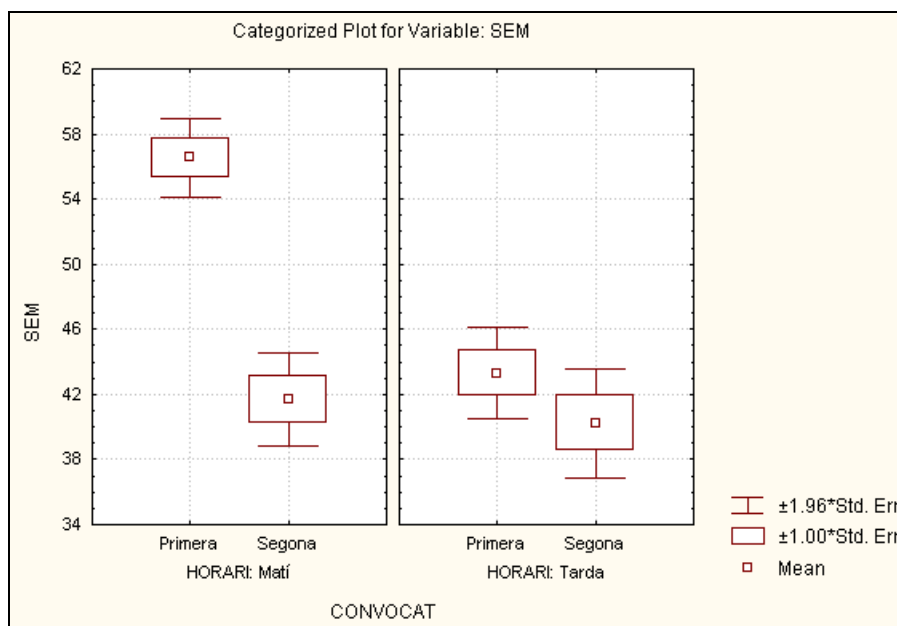


Figura 4.- Taula resum de l'ANOVA realitzat amb la variable independent SPD i els factors convocatòria, gènere i horari. Gràfica descriptiva de les mitjanes de SPD segons els subgrups generats per les variables convocatòria i horari.

Summary of all Effects; design: (edusfarm.sta)						
GENERAL MANOVA	1-CONVOCAT, 2-GENERE, 3-HORARI					
Effect	df Effect	MS Effect	df Error	MS Error	F	p-level
1	1	2363,052	522	289,5312	8,161650	,004449
2	1	632,430	522	289,5312	2,184324	,140025
3	1	626,335	522	289,5312	2,163273	,141946
12	1	30,389	522	289,5312	,104958	,746089
13	1	840,592	522	289,5312	2,903286	,088994
23	1	287,873	522	289,5312	,994272	,319162
123	1	22,489	522	289,5312	,077675	,780584

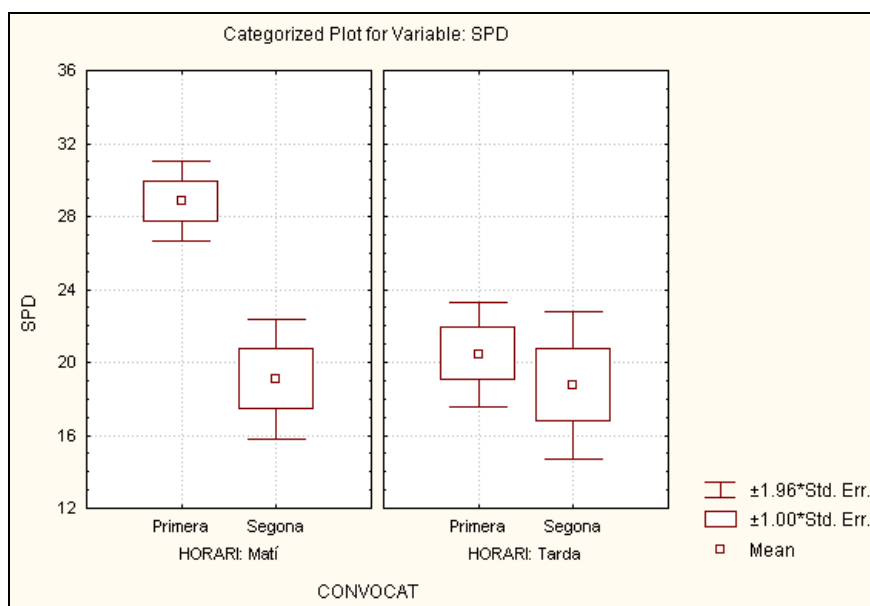


Figura 5.- Taula resum de l'estudi de la correlació existent entre SEM i SPD. Gràfica de SEM enfront SPD.

Correlations (edusfarm.sta)									
BASIC	Marked correlations are significant at $p < .05000$								
STATS	(Casewise deletion of missing data)								
Var. X & Var. Y	Mean	Std.Dv.	r(X,Y)	r ²	t	p	N	Constant dep: Y	
SEM	49,34506	18,91386							
SPD	24,24748	17,60802	,620670	,385232	18,18957	0,00	530	-4,26500	

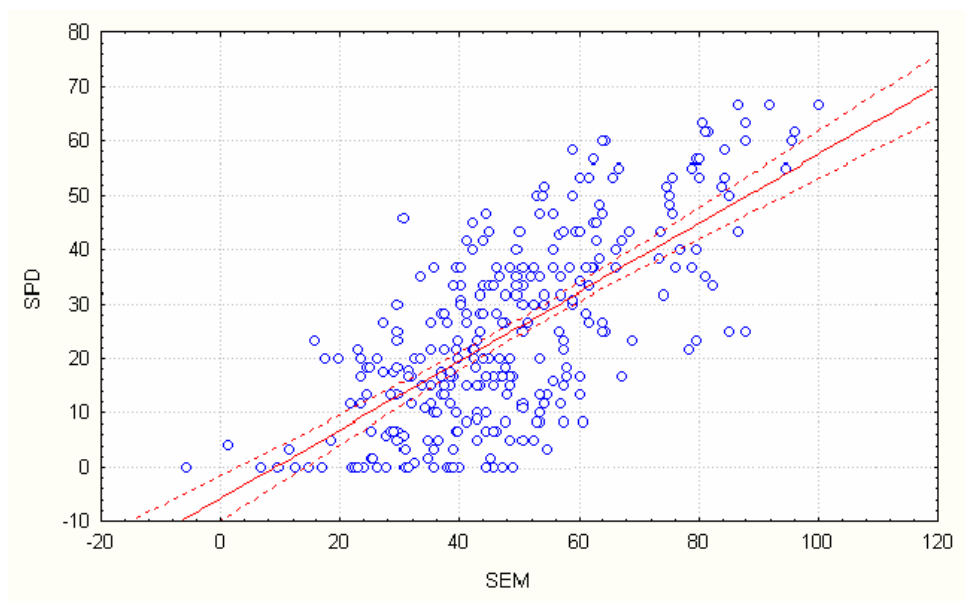


Figura 6.- Taula resum de la t de Student per dades aparellades entre la variable SPD i SEM. Histograma de distribucions de SEM-SPD.

T-test for Dependent Samples (edusfarm.sta)								
BASIC STATS								
Marked differences are significant at p < .05000								
Variable	Mean	Std. Dv.	N	Diff.	Std. Dv. Diff.	t	df	p
SPD	24,24748	17,60802						
SEM	49,34506	18,91386	530	-25,0976	15,94885	-36,2277	529	0,00

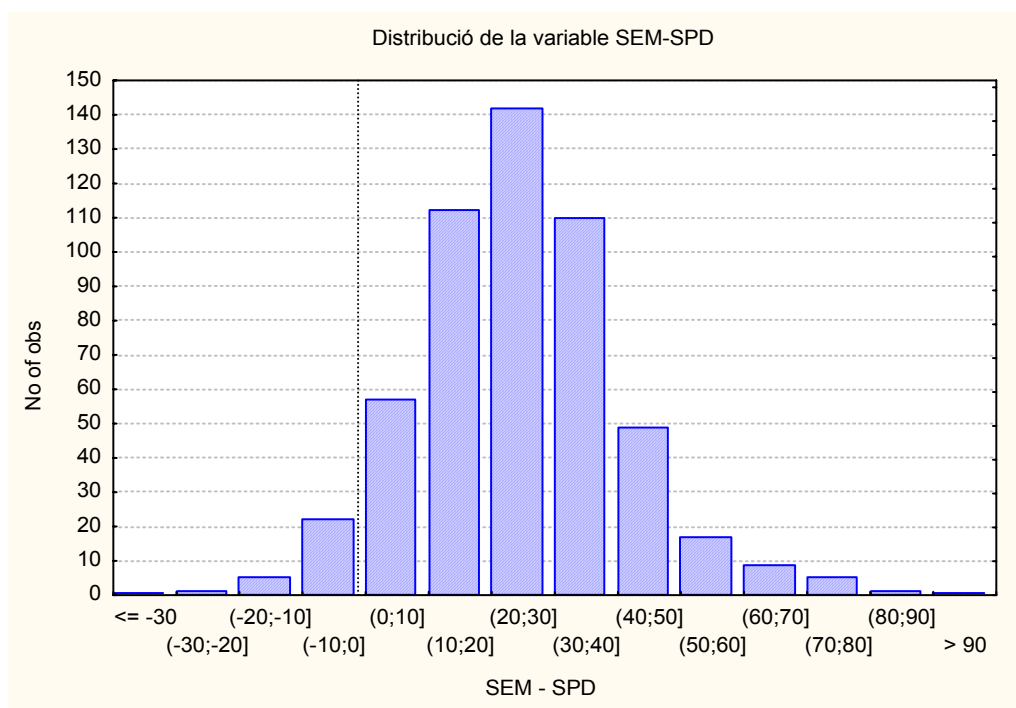


Figura 7.- Taula resum de l'ANOVA realitzat amb la variable independent 'SEM-SPD' i els factors convocatòria, gènere i horari.

Summary of all Effects; design: (edusfarm.sta)						
GENERAL MANOVA						
1-CONVOCAT, 2-GENERE, 3-HORARI						
Effect	df Effect	MS Effect	df Error	MS Error	F	p-level
1	1	294,2169	522	249,8434	1,177606	,278345
2	1	30,5634	522	249,8434	,122330	,726662
3	1	682,6923	522	249,8434	2,732481	,098928
12	1	218,7891	522	249,8434	,875705	,349813
13	1	211,8855	522	249,8434	,848073	,357523
23	1	14,3663	522	249,8434	,057501	,810584
123	1	1,3082	522	249,8434	,005236	,942342

REFERÈNCIES

- Arboix, E., Barà, J., Ferrer, F., Font, J., Forns, M., Mateo, J., Monreal, P., Pérez, J. & Sangrà A. 2003. "Marc general per a l'avaluació dels aprenentatges dels estudiants". Agència per a la Qualitat del Sistema Universitari a Catalunya. Ed. ExLibris, SCCL.
- Mateo J. 2000. "La evaluación del aprendizaje de los alumnos". En: "La evaluación educativa, su práctica y otras metáforas". ICE. Universidad de Barcelona.