

Plans $S < A > *B$

Enoncé 33 *Données Conrad*

Dans une reprise d'une expérience de Conrad (1971), on veut mettre en évidence l'hypothèse de recherche suivante: "les enfants jeunes n'utilisent pas un codage phonologique en mémoire à court terme". Pour ce faire, on sélectionne cinq enfants de 5 ans et 5 enfants de 12 ans (Variable A , avec deux modalités). On montre à chaque enfant un certain nombre de paires d'images représentant des objets dont on s'est assuré auparavant qu'ils sont nommés d'une seule manière par les enfants. On montre les images aux enfants. Puis on retourne les images (les enfants ne voient plus que le dos des images). Ensuite, on donne aux enfants une paire d'images identiques à celles retournées. Enfin, on leur demande de placer ces nouvelles images comme les images retournées sur la table. Pour la moitié des paires d'images les noms des objets se ressemblent (e.g., noix et doigt). Pour l'autre moitié, les noms des objets ne se ressemblent pas (e.g., maison et cheval). Conrad prédit que les enfants les plus vieux réussiront dans l'ensemble mieux que les enfants les plus jeunes, mais également que les enfants les plus vieux utiliseront un codage phonologique comme mnémotechnique (i.e., "la parole intérieure"). De ce fait, les enfants les plus vieux devront commettre plus d'erreurs lorsque les noms se ressemblent acoustiquement que lorsque les noms diffèrent. On présente à chaque enfant cinquante paires d'images correspondant à la modalité b_1 (disstmcorreh acoustique), et cinq paires d'images

correspondant à la modalité b_2 (resstmcorreh acoustique) de la Vcorrevariable D

ires d'images correspondantes. L'ordre de présentation est (Pourquoi cette précaution?).

fiction sur les sources de vcorreariation de

e (qui est
ou si vous
interaction

Résultats de l'expérience		
a_1b_1	a_1b_2	Somme
s_1 15	14	29
s_2 23	20	43
s_3 12	11	23
s_4 16	17	33
s_5 14	13	27
80	75	155
a_2b_1	a_2b_2	Somme
s_6 40	33	73
s_7 38	23	61
s_8 31	21	52
s_9 36	26	62
s_{10} 30	22	52
175	125	300
255	200	455

Calcul en 13 points

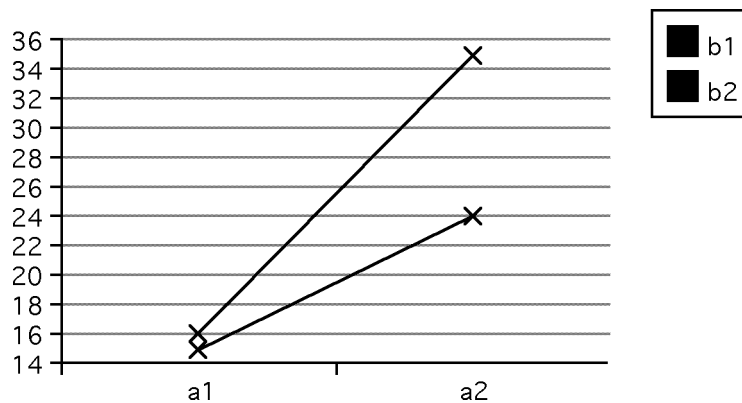
$$\begin{aligned}
 Q1 &= 15 + \dots + 22 = 455 \\
 Q2 &= 15^2 + \dots + 22^2 = 11945 \\
 Q3 &= (15^2 + 300^2)/10 = 11402.5 \\
 Q4 &= (255^2 + 200^2)/10 = 10502.5 \\
 Q5 &= (80^2 + \dots + 125^2)/5 = 11655 \\
 Q6 &= (29^2 + \dots + 52^2)/2 = 11669.5 \\
 Q7 &= 455^2/20 = 10351.25 \\
 Q8 &= SC_A = Q3 - Q7 = 1051.25 \\
 Q9 &= SC_B = Q4 - Q7 = 151.25 \\
 Q10 &= SC_{S(A)} = Q6 - Q3 = 267 \\
 Q11 &= SC_{AB} = Q5 - Q3 - Q4 + Q7 = 101.15 \\
 Q12 &= SC_{BS(A)} = Q2 - Q5 - Q6 + Q3 = 23 \\
 Q13 &= SC_{total} = Q2 - Q7 = 1593.75
 \end{aligned}$$

Source	ddl	SC	CM	F_{cal}	$Pr(F_{cal})$
<i>Entre les sujets</i>					
A	1	1051.25	1051.25	31.5 **	.00050
S(A)	8	267.00	33.38	—	
<i>Dans les sujets</i>					
B	1	151.25	151.25	52.6 **	.00009
AB	1	101.25	101.25	35.2 **	.00035
BS(A)	8	23.00	2.86	—	
Total	19	1593.75			
ns : suspension du jugement au seuil .05					
* : p inférieur à .05; ** : p inférieur à .01.					

Ainsi, les prédictions de Conrad se réalisent. On note un effet principal attribuable à l'âge ($F_{cal}(1,8) = 31.5$; p inférieur à .01), une interaction entre l'âge et la ressemblance acoustique ($F_{cal}(1,8) = 35.2$; p inférieur à .01). On obtient également un effet principal de la ressemblance ($F_{cal}(1,8) = 52.6$, p inférieur à .01), mais cet effet est d'interprétation délicate du fait de l'interaction.

Retrouver les résultats précédents à l'aide de méthodes de comparaison de moyennes sur des protocoles dérivés convenablement choisis.

Représenter graphiquement l'interaction entre les deux facteurs.



Le résultat concernant l'âge peut être retrouvé à l'aide d'un test de comparaison de moyennes sur deux groupes indépendants, en utilisant le protocole dérivé des sommes ou des moyennes par sujet.

Le résultat relatif à l'interaction peut être retrouvé en construisant le protocole des différences individuelles et en faisant une comparaison des moyennes de ce protocole sur les deux groupes.

Enoncé 34 Données Bahrick

Dans une reprise partielle d'une expérimentation de Bahrick (1984), on demande à dix sujets (cinq étudiants et cinq étudiantes) de participer à l'expérience suivante :

On montre aux sujets vingt portraits en noir et blanc (composés de dix portraits d'hommes et dix portraits de femmes). On demande aux sujets d'essayer de "mémoriser" ces vingt portraits afin de pouvoir les reconnaître lors d'un test ultérieur. Les sujets accomplissent ensuite pendant environ une demi-heure diverses tâches. Puis on leur présente vingt paires de photographies composées d'un portrait vu pendant la phase d'apprentissage et d'un portrait inconnu des sujets; et on leur demande d'identifier dans chaque paire de photographies le portrait connu.

On donne ci-dessous le nombre de portraits correctement identifiés en fonction des sujets et du "sexe" des portraits. (Les sujets sont identifiés par un prénom).

Nom du sujet	Portrait masculin	Portrait féminin
Albert	6	6
Henri	6	6
Jules	5	5
Paul	5	5
Octave	5	6
Albertine	6	8
Henriette	7	8
Julie	6	6
Paule	7	7
Octavie	6	6

Indications de solution. L'expérience est menée selon le plan $\mathcal{S}_5 < \mathcal{X}_2 > * \mathcal{P}_2$. Le tableau d'analyse de variance est donné par :

Source	ddl	SC	CM	F_{cal}	$Pr(F_{cal})$
<i>Entre les sujets</i>					
\mathcal{X}	1	7.2	7.2	10.28 *	.0124
$\mathcal{S}(\mathcal{X})$	8	5.6	0.7	—	
<i>Dans les sujets</i>					
\mathcal{P}	1	0.8	0.8	3.2 NS	.11
$\mathcal{X}\mathcal{P}$	1	0.2	0.2	0.8 NS	.40
$\mathcal{P}\mathcal{S}(\mathcal{X})$	8	2	0.25	—	
Total	19	15.8			
ns : suspension du jugement au seuil .05					
* : p inférieur à .05; ** : p inférieur à .01.					

Enoncé 35 Dossier “King”

En 1986, King a étudié l’activité motrice chez le rat après injection d’un médicament appelé midazolam. La première injection du médicament entraîne généralement une diminution nette de l’activité motrice. Mais une certaine tolérance se développe rapidement. King souhaitait savoir si cette tolérance acquise pouvait s’expliquer sur la base d’une tolérance conditionnée.

Il a utilisé trois groupes et n’a recueilli les données (présentées dans le tableau ci-dessous) que le dernier jour, jour du test. Durant le pré-test, deux groupes d’animaux ont reçu à plusieurs reprises des injections de midazolam réparties sur plusieurs jours, tandis que le groupe témoin recevait des injections d’une solution saline physiologique.

Le jour du test, un groupe (le groupe “même”) a reçu une injection de midazolam dans le même environnement qu’auparavant. Le groupe “différent” a également reçu une injection de midazolam, mais dans un environnement différent. Enfin, le groupe témoin a reçu, pour la première fois, une injection de midazolam. Ce groupe témoin devrait donc manifester la réaction initiale classique au médicament (comportement ambulatoire réduit), tandis que le groupe “même” devrait présenter l’effet normal de tolérance. Par contre, si King a raison, le groupe “différent” devrait réagir de la même façon que le groupe témoin; en effet, ces animaux allaient cette fois recevoir l’injection dans un environnement différent, et les éléments nécessaires pour susciter une tolérance conditionnée ne seraient pas présents. La variable dépendante du tableau ci-dessous est une mesure du comportement ambulatoire, en unités arbitraires.

Comme le médicament se métabolise sur une période d’environ 1 heure, King a enregistré ses données par blocs (ou intervalles) de 5 minutes. Le tableau 8 donne les valeurs observées pour les 6 premiers blocs de données.

1) a) Quelles sont les variables indépendantes (ou facteurs de variation) prises en compte? Quel est le nombre de niveaux de chacun des facteurs?

b) Quelle est la variable dépendante?

c) Ecrire le plan d’expérience correspondant.

2) a) Calculer les moyennes correspondant aux 18 conditions expérimentales définies par les combinaisons des variables “groupe” et “intervalle”.

b) Réaliser un graphe illustrant une éventuelle interaction entre ces variables. Commenter le graphe obtenu.

3) *Analyse de variance.*

Le tableau d’analyse de variance relatif aux données observées se présente ainsi :

Sources de var.	ddl	SC	CM	F
<i>Entre les sujets</i>				
Groupes	2	285815	142907	...
$\mathcal{S}(\mathcal{G})$	21	384722	18320	...
<i>Dans les sujets</i>				
Intervalles	5	399736
$\mathcal{I} \times \mathcal{G}$	10	80820
Résidu	105	281199	2678	...
Total	143	1432293		

a) Compléter ce tableau en calculant les carrés moyens et les statistiques F de Fisher qui sont remplacés par “...” dans le tableau ci-dessus.

	Intervalles					
	1	2	3	4	5	6
Témoïn	150	44	71	59	132	74
	335	270	156	160	118	230
	149	52	91	115	43	154
	159	31	127	212	71	224
	159	0	35	75	71	34
	292	125	184	246	225	170
	297	187	66	96	209	74
	170	37	42	66	114	81
Même	346	175	177	192	239	140
	426	329	236	76	102	232
	359	238	183	123	183	30
	272	60	82	85	101	98
	200	271	263	216	241	227
	366	291	263	144	220	180
	371	364	270	308	219	267
	497	402	294	216	284	255
Différent	282	186	225	134	189	169
	317	31	85	120	131	205
	362	104	144	114	115	127
	338	132	91	77	108	169
	263	94	141	142	120	195
	138	38	16	95	39	55
	329	62	62	6	93	67
	292	139	104	184	193	122

Tableau 8: Données King

b) En utilisant un seuil de 5%, étudier quelles sont les sources de variation dont l'effet est significatif.

4) *Comparaisons de moyennes*

a) Déterminer le protocole dérivé obtenu en calculant le score moyen observé sur les intervalles 2 à 6 pour chacun des sujets des groupes "témoïn" et "différent".

b) Comparer le comportement des deux groupes à l'aide d'un test sur les moyennes de ce protocole dérivé.

Enoncé 36 Dossier "Termites"

Dans une étude expérimentale (d'après Catherine Venturelli: la dynamique du creusement chez *Reticulitermes santonensis*, 1990), on étudie le comportement de creusement de 72 groupes de termites (d'où le facteur groupe G à 72 modalités). Chaque groupe comprend 50 termites. Ces 72 groupes sont répartis dans 12 conditions expérimentales différentes; d'où le facteur Condition C à 12 modalités. A chacune des conditions expérimentales on affecte 6 groupes, chacun des 72 groupes étant affecté à une seule condition expérimentale. Pour chaque groupe on observe le nombre de centimètres de galeries creusés par le groupe en 12 heures. L'expérience se déroule sur 15 jours, avec 2 relevés par jour correspondant

à deux périodes d'activité: un relevé le soir, où l'on observe le nombre de centimètres creusés pendant la journée (période p1), un relevé le matin, où l'on observe le nombre de centimètres creusés pendant la nuit (période p2). D'où le facteur Jour J à 15 modalités et le facteur Période d'activité P à 2 modalités (p1 et p2).

Les données suivantes concernent une partie des observations, les résultats de 12 groupes de termites. Six groupes ont été placés dans du sable humide (h1), 6 groupes dans du sable peu humide (h2). D'où le facteur Humidité à 2 modalités. Il s'agit des deux relevés (p1 et p2) du premier jour.

Données "Termites" et protocoles dérivés

		p1	p2	p2-p1	(p1+p2)/2
G(h1) humide	g1	0.0	48.80	48.80	24.40
	g2	36.5	79.3	42.80	57.90
	g3	16.2	46.3	30.10	31.25
	g4	17.8	69.3	51.50	43.55
	g5	34.4	96.8	62.40	65.60
	g6	29.4	81.2	51.80	55.30
G(h2) peu humide	g7	0.0	0.0	0.00	0.00
	g8	9.8	51.0	41.20	30.40
	g9	9.4	64.0	54.60	36.70
	g10	13.1	76.9	63.80	45.00
	g11	2.2	30.7	28.50	16.45
	g12	4.4	7.3	2.90	5.85

Moyennes

	p1	p2
h1	22,38	70,28
h2	6,48	38,32

Effets intra

H*P	p1 (jour)	p2 (nuit)	diff p1,p2
h1 (humide)	22,38	70,28	-47,90
h2 (peu humide)	6,48	38,32	-31,84
diff h1,h2	15,90	31,96	-16,06

- 1) En moyenne, jour et nuit confondus, les termites creusent-ils plus en sable humide ou en sable peu humide?
- 2) Le jour, les termites creusent-ils plus en sable humide ou en sable peu humide?
- 3) Les termites creusent-ils plus la nuit que le jour?
- 4) Lorsqu'ils sont placés en sable peu humide, les termites sont-ils influencés par le jour et la nuit?
- 5) La différence entre le jour et la nuit est-elle la même, quelle que soit l'humidité du sable?

Indications de solutions.

Les données fournies correspondent au plan d'expérience $S_6 < H_2 > *P_2$.

1) On raisonne ici sur le protocole dérivé des moyennes par individu. Au niveau descriptif: $\bar{x}_1 = 46.33$, $\bar{x}_2 = 22.4$, $s_{c,1} = 16.1352$, $s_{c,2} = 17.8343$. Les termites semblent creuser plus en condition H_1 qu'en condition H_2 .

Un test de comparaison de deux moyennes sur des groupes indépendants donne: $t_{obs} = 2.43$, et $ddl = 10$, valeur significative au seuil de 5% unilatéral.

2) On procède ici comme dans la question précédente, mais en utilisant la première colonne de données. On obtient: $\bar{x}_1 = 22.38$, $\bar{x}_2 = 6.48$, $s_{c,1} = 13.8070$, $s_{c,2} = 5.0598$. $t_{obs} = 2.65$, et $ddl = 10$, valeur significative au seuil de 5% unilatéral.

Un autre paramètre descriptif intéressant est l'effet calibré du facteur H défini par:

$EC = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_{1,c}^2 + s_{2,c}^2}} = 1.08$. Un écart calibré supérieur à 1 indique un effet important.

3) On utilise ici le protocole dérivé des effets individuels (colonne des différences). On a: $\bar{d} = 39.87$ et $s_c = 20.99$. Une comparaison de moyennes sur deux groupes appariés donne: $t_{obs} = 6.57$ et $ddl = 11$, valeur significative au seuil de 5% unilatéral.

4) On procède comme dans la question précédente, en se limitant aux individus statistiques du deuxième groupe. On a: $\bar{d} = 31.83$ et $s_c = 26.42$. Une comparaison de moyennes sur deux groupes appariés donne: $t_{obs} = 2.95$ et $ddl = 5$, valeur significative au seuil de 5% unilatéral.

5) Il s'agit ici d'étudier l'interaction entre le facteur P et le facteur H . On pourra tracer un diagramme d'interaction à partir du tableau des moyennes donné dans l'énoncé.

On obtient: $Moy((P_2 - P_1)/H_1) = 47.90$, $s_{c,1} = 10.79$, $Moy((P_2 - P_1)/H_2) = 31.83$, $s_{c,2} = 26.42$. On peut aussi calculer l'effet calibré (cf. question 2). Ici, $EC = 0.563$, ce qui est assez faible.

Une comparaison de moyennes sur les deux groupes correspondant aux modalités h_1 et h_2 du facteur H donne: $t_{obs} = 1.38$, valeur non significative aux seuils traditionnels.

Remarque. Le tableau d'analyse de variance est ici le suivant:

Source	ddl	SC	CM	F_{cal}	Pr
H	1	3437	3437	5.94	.035
$S < H >$	10	5784	578.4		
P	1	9536	9536	46.84	.00005
Interaction	1	387.2	387.2	1.90	.20
Résidu	10	2036	203.6		
Total	23	21180			

On retrouve ainsi les résultats des questions 1, 3 et 5, en remarquant que $F_{cal} = t_{obs}^2$.

En particulier, $t_{obs} = 2.43$ correspond à $F_{cal} = 5.94$ (question 1), tandis que $t_{obs} = 1.38$ correspond à $F_{cal} = 1.90$ (question 5).

Autres plans

Enoncé 37 Dossier "Neglige"

Une recherche a porté sur la "pseudo-négligence" qu'on observe chez des sujets normaux. Ce nom provient des similarités qu'elle présente avec l'héminégligence (atteinte de la moitié du champ visuel) de sujets atteints d'une lésion cérébrale. La tâche des sujets consiste à déterminer le milieu subjectif d'une baguette de 24cm avec la seule aide d'informations kinesthésiques. La pseudo-négligence se traduit par une déviation systématique

vers la droite (pour les droitiers) de ce milieu subjectif par rapport au milieu objectif de la baguette.

Les données portent sur 24 femmes droitrières (facteur S) réparties selon 2 conditions (12 sujets pour chacune): active (c1) où le sujet peut librement déplacer son doigt posé sur un curseur mobile le long de la baguette; ou passive (c2) où le sujet commande un moteur déclenchant le mouvement de la baguette dans un sens ou dans l'autre, alors que son doigt ne bouge pas (facteur C). Chaque sujet exécute cette tâche dans 6 situations expérimentales obtenues par le croisement de la main utilisée, gauche (m1) ou droite (m2) et l'orientation du regard, 30° à gauche (o1), 0° (o2) ou 30° à droite (o3) (facteurs M et O). Pour chaque sujet et chaque situation on mesure la déviation en cm entre le milieu subjectif et le milieu objectif de la baguette. Une déviation à droite est notée par une valeur positive, à gauche par une valeur négative.

L'objectif principal de l'expérience est l'étude de l'effet de la condition sur la déviation, et des possibles variations de cet effet selon l'orientation.

Les données sont présentées dans le tableau ci-dessous.

Données "Négligence" et protocoles dérivés (PD1 à PD3)

	m1o1	m1o2	m1o3	m2o1	m2o2	m2o3	PD1	PD2	PD3
s1c1	1.95	0.95	0.55	0.15	-0.80	-0.65	0.36	1.80	1.58
s2c1	3.00	3.10	1.55	-0.10	-0.30	-0.60	1.11	3.10	2.88
s3c1	1.00	1.20	-0.65	-1.25	-2.20	-2.15	-0.68	2.25	2.38
s4c1	0.25	0.20	0.85	-0.85	0.05	-0.05	0.08	1.10	0.72
s5c1	1.15	0.55	0.95	-0.10	-0.40	-1.95	0.03	1.25	1.70
s6c1	1.85	0.75	-1.65	-0.35	0.15	-0.35	0.07	2.20	0.50
s7c1	2.05	1.75	-1.50	1.05	0.05	-0.90	0.42	1.00	0.70
s8c1	1.75	-0.50	0.25	-0.25	-0.05	0.90	0.35	2.00	0.30
s9c1	0.40	1.85	-0.10	-0.45	-0.40	-1.05	0.04	0.85	1.35
s10c1	-0.80	4.10	3.00	-1.15	-0.20	-2.25	0.45	0.35	3.30
s11c1	2.50	-0.75	0.30	0.20	-1.25	-1.20	-0.03	2.30	1.43
s12c1	1.80	1.65	0.55	1.00	-1.00	1.30	0.88	0.80	0.90
s13c2	-0.30	-0.10	-0.55	1.30	-1.90	0.75	-0.13	-1.60	-0.37
s14c2	1.40	-1.00	0.95	-0.20	0.00	-0.50	0.11	1.60	0.68
s15c2	0.25	0.75	-0.70	0.75	0.55	0.60	0.37	-0.50	-0.53
s16c2	0.75	1.10	1.40	0.25	-1.55	0.05	0.33	0.50	1.50
s17c2	-0.30	-0.70	-0.80	0.85	-0.10	-1.15	-0.37	-1.15	-0.47
s18c2	-2.10	3.45	-1.85	0.95	1.50	2.80	0.79	-3.05	-1.92
s19c2	1.85	-0.55	2.25	-0.05	1.55	-0.30	0.79	1.90	0.78
s20c2	1.65	-0.75	0.05	-0.60	-1.55	-1.75	-0.49	2.25	1.62
s21c2	-0.75	1.25	-0.25	-0.20	-0.05	2.35	0.39	-0.55	-0.62
s22c2	1.80	1.15	1.95	0.80	0.85	0.90	1.24	1.00	0.78
s23c2	-0.95	0.15	0.80	-0.15	-0.30	-1.05	-0.25	-0.80	0.50
s24c2	0.20	-0.45	-0.80	0.65	2.20	-0.10	0.28	-0.45	-1.27

Réf: Chokron, Imbert (1993) - *Egocentric reference and asymmetric perception of space. Neuropsychologia* 31, 3, 267-275. D'après J.M. Bernard (1994) - *Structure des données, données planifiées Mathématiques, Informatique et Sciences humaines n°126* 7-18.

1) a) Pour chacun des 4 cas ci-dessous, écrire au moyen des symboles <> et * la relation entre les deux facteurs S et C ; M et O ; C et M ; C et O.

b) Ecrire au moyen des mêmes symboles la relation entre les trois facteurs S, M et O.

2) a) On calcule la moyenne de toutes les valeurs observées. On trouve 0.256. Que signifie cette valeur. Que peut-on en conclure?

b) On s'intéresse aux effets moyens et intra du facteur C.

	m1	m2	Moyennes
c1	+0.996	-0.483	0.256
c2	+0.285	+0.226	0.256
Moyennes	0.640	-0.128	0.256

Commenter l'effet moyen du facteur C.

Commenter l'effet observé du facteur C lorsque c'est la main gauche (m1) qui est utilisée.

Calculer la valeur de cet effet

Construire un graphe représentant l'interaction entre les facteurs C et M.

3) Les trois dernières colonnes du tableau général (PD1 à PD3) indiquent trois protocoles pouvant être dérivés du protocole de base. Indiquer, pour chacun des effets ci-dessous, quel est le protocole dérivé pertinent et indiquer quel calcul a permis d'obtenir la première valeur de ce protocole dérivé:

Effet de C ; effet de M ; effet de M.C ; effet de M/o1 .

4) On s'intéresse à l'effet de la condition (C) avec la main gauche (m1) et avec l'orientation o2. Le protocole dérivé pertinent est la deuxième colonne du tableau principal (m1o2). A l'aide d'un test de comparaison de moyennes, déterminer si cet effet est significatif.

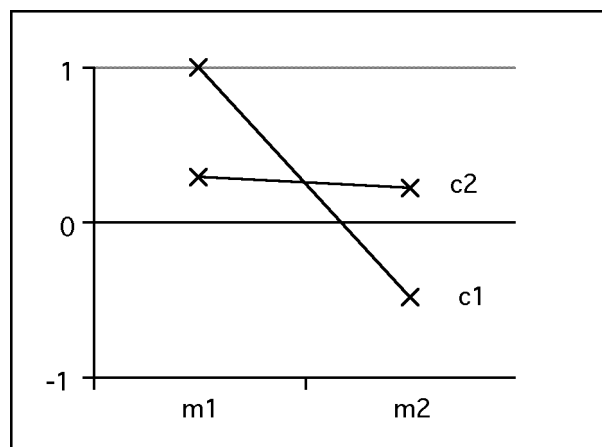
Réponses : 1) a)

<i>S et C</i>	<i>S<C></i>
<i>M et O</i>	<i>M*O</i>
<i>C et M</i>	<i>C*M</i>
<i>C et O</i>	<i>C*O</i>

1) b) Entre S, M et O : $S*M*O$.

2) a) La moyenne générale représente l'effet de la pseudo-négligence, indépendamment des effets des autres facteurs.

b) L'effet moyen du facteur C est nul. En revanche, l'effet observé du facteur C dans la modalité m1 est $0.996 - 0.285 = 0.711$. Le graphe d'interaction pourra être représenté par :



c)

	PD	Calcul
Effet de C	PD1	$(1.95 + 0.95 + 0.55 + 0.15 - 0.80 - 0.65)/6$
Effet de M	PD3	$(1.95 + 0.95 + 0.55)/3 - (0.15 - 0.8 - 0.65)/3$
Effet de M.C	PD3	id.
Effet de M/o1	PD2	$1.95 - 0.15$

3) Il s'agit d'une comparaison de moyennes sur groupes indépendants. On obtient les résultats suivants :

	Moyenne	écart type	écart type cor.
c1	1.238	1.333	1.392
c2	0.358	1.211	1.265

$t_{obs} = 1.6191$ et $ddl = 22$. Au seuil de 5%, la différence n'est pas significative.

Il s'agit ici d'un plan $S < C_2 > * M_2 * O_3$. Les éléments vus en cours ne permettent pas de prévoir la structure du tableau d'analyse de variance, qui est ici assez complexe :

Source	ddl	SC	CM	F_{cal}	Pr
C2	1	0.000	0.000	0.00	0.997
S(C2)	22	31.658	1.439		
M2	1	21.275	21.275	13.50	0.001
C2 × M2	1	18.169	18.169	11.53	0.003
M2 × S(C2)	22	34.677	1.576		
O3	2	5.342	2.671	2.25	0.118
C2 × O3	2	3.236	1.618	1.36	0.267
O3 × S(C2)	44	52.348	1.190		
M2 × O3	2	1.357	0.678	0.57	0.571
C2 × M2 × O3	2	0.495	0.248	0.21	0.814
Résidu	44	52.586	1.195		
Total	143	221.142			

Ce tableau peut être obtenu à l'aide de Minitab. On saisit les données, à raison d'une colonne pour chaque facteur et d'une colonne pour la variable dépendante, et on utilise le menu Stats - Anova - Modèle linéaire généralisé. On complète ensuite le dialogue en indiquant dans la zone d'édition "Modèle" :

Condition Sujet(Condition) Main Main*Condition Main*Sujet(Condition)

Orientation Orientation*Condition Orientation*Sujet(Condition)

Main*Orientation Main*Orientation*Condition

et en indiquant Sujet comme facteur aléatoire.

Ce tableau ne fait d'ailleurs que confirmer ce que nous avons établi par ailleurs : seuls le facteur "Main" et l'interaction "Main × Condition" ont des effets significatifs.

Enoncé 38 Dossier "Tapping"

Une tâche de "Tapping" consiste à demander à des sujets droitiers d'appuyer avec l'index sur un bouton le plus rapidement possible durant une période limitée (20s). On note le nombre d'appuis effectués durant cette période. Huit garçons (sexe 1) et huit filles (sexe 2) ont passé une expérience dans laquelle on comparait les performances selon que la tâche était effectuée avec la main gauche (m1) ou avec la main droite (m2), et selon que les sujets avaient à résoudre une tâche concurrente (c1) ou non (c2). Chaque sujet

est confronté à chacune des 4 conditions définies par le croisement des facteurs M et C. Les données présentées dans le tableau 9 correspondent à la moyenne des performances obtenues sur 5 essais.

Critères d'importance de l'effet:

Critère sémantique: On considérera qu'un effet est faible s'il est inférieur à 1, important s'il est supérieur à 2.

Critère psychométrique: On prendra les critères habituels ($d/s < 1/3$ et $d/s > 2/3$) comme limites d'un effet faible et d'un effet important.

	m1c1	m2c1	m1c2	m2c2
s1x1	70.25	79.50	61.75	68.00
s2x1	57.50	65.00	62.25	71.00
s3x1	71.00	91.00	60.00	68.00
s4x1	63.75	76.25	58.25	70.50
s5x1	72.50	60.25	61.25	60.50
s6x1	61.25	80.00	64.25	87.00
s7x1	58.00	59.75	62.75	69.00
s8x1	57.50	64.25	74.25	85.00
s9x2	62.00	66.50	66.25	72.50
s10x2	74.00	78.25	74.25	87.25
s11x2	63.25	77.50	61.50	86.50
s12x2	88.25	91.50	93.50	83.25
s13x2	60.75	67.25	60.75	69.25
s14x2	70.75	84.00	73.25	79.50
s15x2	54.50	59.00	64.50	68.25
s16x2	48.50	51.00	62.50	70.25

Tableau 9: Données "Tapping"

1) Pour chacun des 5 effets X , M , C , $X.M$, $M.C$ indiquer :

– quel est, parmi les 4 protocoles dérivés du tableau 10, le protocole dérivé pertinent pour l'étude de cet effet

– comment a été calculée la première valeur (sujet s1) de ce protocole dérivé pertinent.

2) On s'intéresse à l'effet du facteur Sexe (X), c'est-à-dire aux différences de performances obtenues par les garçons (x1) et les filles (x2). On trouve une différence de performance $d_{obs} = 3.086$ en faveur des filles. On veut tester l'hypothèse H_0 d'une absence d'effet parent pour cette comparaison.

– Indiquer la formule du test T de Student à utiliser dans ce cas.

– On trouve $t_{obs} = 0.788$. Indiquer le nombre de degrés de liberté et le résultat du test. Donner une conclusion inférentielle sur l'effet de X .

3) On s'intéresse à l'effet du facteur C, c'est-à-dire à l'effet de la tâche concurrente sur les performances au tapping. En moyenne, on trouve $m_{c_1} = 68.27$ et $m_{c_2} = 70.53$ d'où $d_{obs} = 2.26$. On souhaite savoir si l'effet de la tâche concurrente est différent pour les garçons (C/x1) et les filles (C/x2). Pour cela, on dérive le protocole des moyennes de support $X*C$ représenté ci-dessous.

	PD1	PD2	PD3	PD4
s1x1	-3.00	7.750	-10.000	69.8750
s2x1	1.25	8.125	5.375	63.9375
s3x1	-12.00	14.000	-17.000	72.5000
s4x1	-0.25	12.375	-5.625	67.1875
s5x1	11.50	-6.500	-5.500	63.6250
s6x1	4.00	20.750	5.000	73.1250
s7x1	4.50	4.000	7.000	62.3750
s8x1	4.00	8.750	18.750	70.2500
s9x2	1.75	5.375	5.125	66.8125
s10x2	8.75	8.625	4.625	78.4375
s11x2	10.75	19.625	3.625	72.1875
s12x2	-13.50	-3.500	-1.500	89.1250
s13x2	2.00	7.500	1.000	64.5000
s14x2	-7.00	9.750	-1.000	76.8750
s15x2	-0.75	4.125	9.625	61.5625
s16x2	5.25	5.125	16.625	58.0625

Tableau 10: Protocoles dérivés des données "Tapping"

	c1	c2
x1	67.984	67.734
x2	68.563	73.328

- Pour chacun des deux effets C/x1 et C/x2, calculer les valeurs des effets, puis conclure sur l'importance des effets en utilisant le critère sémantique.

- Construire un graphique représentant l'interaction entre C et X.

4) La première colonne du tableau 10 (protocole PD1) est un sous-protocole de structure S8<X2>.

- Calculer sur ce protocole les moyennes m_{x_1} et m_{x_2} , les variances corrigées ($s_{1,c}^2$ et $s_{2,c}^2$) et les écarts-types corrigés de chacun des deux groupes x1 et x2.

- Calculer la variance corrigée intra et l'écart-type corrigé intra.

5) Pour les mêmes données que la question précédente, de structure S8<X2>, on peut calculer la variance inter (variance des moyennes des deux groupes x1 et x2) et la variance intra (moyenne des variances des deux groupes).

- Indiquer comment augmenter la variance intra sans modifier la variance inter.

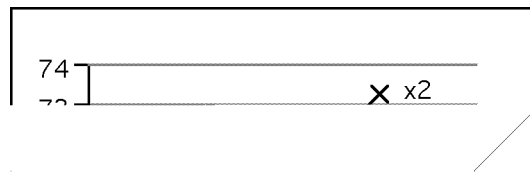
- Indiquer comment augmenter la variance inter sans modifier la variance intra.

Indications de réponses : 1)

Effet	Protocole	Calcul
X	PD4	$(70.25 + 79.50 + 61.75 + 68)/4$
M	PD2	$[(79.50 - 70.25) + (68.00 - 61.75)] / 2$
C	PD3	$[(61.75 + 68.00) - (70.25 + 79.50)] / 2$
X.M	PD2	
M.C	PD1	$(68.00 - 61.75) - (79.50 - 70.25)$

2) Comparaison de moyennes sur groupes indépendants. $ddl = 14$. Pas d'effet significatif au seuil de 5%.

3) Le graphe d'interaction est donné par :



4) On obtient :

$m_{x_1} = 1.25$, $m_{x_2} = 0.9025$, $s_{1,c}^2 = 46.66$, $s_{2,c}^2 = 64.570$, $s_{1,c} = 6.83$, $s_{2,c} = 8.04$.

La variance corrigée intra est la demi-somme des variances précédentes (car il s'agit de groupes équilibrés). D'où $s_{intra}^2 = 55.615$ et $s_{intra} = 7.45$.

5) Pour augmenter la variance intra sans modifier la variance inter, on modifie les scores de façon à augmenter la dispersion dans l'un des groupes sans modifier les moyennes des groupes. Par exemple, on diminue de 10 points le score du sujet 3 et on augmente de 10 points celui du sujet 5.

Pour augmenter la variance inter sans modifier la variance intra, on peut, par exemple, augmenter d'une même quantité tous les scores du groupe présentant la moyenne la plus élevée.

Enoncé 39 Dossier "TR"

TR= Temps de réaction

Les données présentées ci-dessous sont extraites d'une expérience de temps de réaction (Holender et Bertelson, 1975). La tâche du sujet (adulte) est de réagir le plus rapidement possible à la présentation d'un stimulus. Deux facteurs expérimentaux sont en jeu: le facteur F, fréquence du stimulus, à 2 modalités: f1: stimulus fréquent (de fréquence 0.75) f2: stimulus rare (de fréquence 0.25) et le facteur D, durée de la période préparatoire (délai entre la présentation d'un signal avertisseur et celle du stimulus), à 2 modalités également: d1: période courte (0.5 secondes) d2: période longue (5 secondes). On introduit ici un facteur supplémentaire (non présent dans l'expérience originale) en supposant que les sujets sont repartis en 2 groupes de 4 sujets chacun, d'où le facteur Groupe, $G = \{g1, g2\}$. Le facteur sujet, S, a 8 modalités.

Chaque sujet effectue plusieurs essais dans chacune des 4 conditions expérimentales correspondant aux modalités du croisement $F2 * D2$. Le protocole de base présenté ci-dessous donne pour chaque sujet, et chacune des conditions expérimentales, la moyenne des temps de réaction pour les différents essais (en millisecondes). D'où un protocole numérique de 32 observations.

		Protocole de base				Protocoles dérivés		
		f1d1	f2d1	f1d2	f2d2	F/d1	F/d2	Moy.
S(g1)	s1	387	435	416	473	48	57	427.75
	s2	321	336	343	368	15	25	342
	s3	333	362	358	390	29	32	360.75
	s4	344	430	352	393	86	41	379.75
S(g2)	s5	368	432	432	504	64	72	434
	s6	357	367	394	411	10	17	382.25
	s7	336	346	340	421	10	81	360.75
	s8	387	454	438	496	67	58	443.75

Tableau des moyennes (m) et des écarts types corrigés (s):

		Protocole de base				Protocoles dérivés		
		f1d1	f2d1	f1d2	f2d2	F/d1	F/d2	Moy.
g1	m	346.250	390.750	367.250	406.000	44 500	38.750	377.56
	s	28.745	49.406	33.079	46.036	30.795	13.817	36.84
g2	m	362.000	399.750	401.000	458.000	37 750	57.000	405.19
	s	21.307	51.461	45.092	48.778	32.066	28.296	40.08
Ens.	m	354.125	395.25	384.125	432			

Facteur F - Facteur D

	f2	f1	f2 - f1
s1	454	401.5	52.5
s2	353	332	20
s3	376	345.5	30.5
s4	411.5	348	63.5
Moy.			41.625
$s_{1,c}$			19.90
s5	468	400	68
s6	389	375.5	13.5
s7	383.5	338	45.5
s8	475	412.5	62.5
Moy.			47.375
$s_{2,c}$			24.53
Moy.			44.5
s_c			20.91

	d2	d1	d2 - d1
s1	444.5	411	33.5
s2	355.5	328.5	27
s3	374	347.5	26.5
s4	372.5	387	-14.5
Moy.			18.125
$s_{1,c}$			21.98
s5	468	400	68
s6	402.5	362	40.5
s7	380.5	341	39.5
s8	467	420.5	46.5
Moy.			48.625
$s_{2,c}$			13.28
Moy.			33.75
s_c			23.42

- 1) Quel est le plan d'expérience utilisé?
- 2) Etudier l'effet du facteur G (groupe).
- 3) Etudier l'effet principal du facteur F.
- 4) Etudier l'effet principal du facteur D.
- 5) Construire un graphe illustrant l'interaction entre les les facteurs D et F. Le commenter.
- 6) Construire un graphe illustrant l'interaction entre les facteurs D et G. Quel est le protocole dérivé pertinent pour étudier cette interaction? Cette interaction est-elle significative? Conclure à l'aide d'un test de comparaison de moyennes.

Indications de correction.

1) Le plan d'expérience est $S_4 < G_2 > *D_2 * F_2$.

2) Le protocole dérivé pertinent est ici celui des moyennes par sujet (premier tableau). En utilisant les données calculées dans le second tableau, on obtient $t_{obs} = 1.015$, résultat non significatif aux seuils traditionnels.

3) L'effet principal du facteur F peut être étudié à l'aide d'une comparaison de moyennes sur deux groupes appariés, en utilisant les données du tableau 3. On obtient : $t_{obs} = 4.03$ et $ddl = 7$, résultat significatif au seuil de 5% unilatéral.

4) L'effet principal du facteur D peut être étudié à l'aide d'une comparaison de moyennes sur deux groupes appariés, en utilisant les données du tableau 4. On obtient : $t_{obs} = 6.01$ et $ddl = 7$, résultat significatif au seuil de 5% unilatéral.

5) On utilise les moyennes figurant dans la dernière ligne du tableau 2. Le graphe ne montre aucune interaction entre D et G .

6) On utilise ici le tableau 4. L'interaction peut être étudiée à l'aide d'une comparaison de deux moyennes sur des groupes indépendants. On obtient : $t_{obs} = 1.67$ et $ddl = 6$, résultat non significatif au seuil de 5% unilatéral.

Enoncé 40 Données "Figures"

Les données qui suivent sont empruntées à une expérience de ségrégation perceptive sur des figures de GOTTSCHALDT. La tâche du sujet consiste à retrouver et à retracer les contours d'une figure simple dans des figures complexes. Le matériel comportait plusieurs planches de difficulté croissante. Pour l'analyse qui suit, on retiendra seulement les résultats relatifs aux planches p1 et p2.

Les sujets de l'expérience ont été affectés au hasard à l'une des deux conditions d'apprentissage a1 et a2 décrites plus loin. A l'intérieur de chaque condition, les sujets passent toutes les planches, dans tous les ordres possibles, donc s'agissant des planches p1 et p2: soit dans l'ordre 1-2, soit dans l'ordre 2-1.

Les conditions a1 et a2 sont définies de la manière suivante:

- condition a1: les sujets effectuent la tâche après avoir procédé à un apprentissage prolongé de figures simples;
- condition a2: les sujets effectuent la tâche sans apprentissage préalable.

La variable dépendante de base est le temps (en secondes) mis pour repasser une figure simple (temps moyen par sujet et par planche).

Pour l'analyse qui suit, on a retenu le sous-protocole suivant relatif à 24 sujets (6 sujets pour chacun des 4 groupes) et aux deux planches p1 et p2.

condition a1			
		planche p1	planche p2
g11 ordre 1-2	1	15	38
	2	9	40
	3	17	20
	4	8	13
	5	16	41
	6	9	35
g21 ordre 2-1	7	7	27
	8	7	36
	9	12	35
	10	4	24
	11	7	36
	12	7	30

condition a2			
		planche p1	planche p2
g12 ordre 1-2	13	6	27
	14	7	15
	15	25	24
	16	26	30
	17	10	31
	18	6	10
g22 ordre 2-1	19	6	22
	20	10	27
	21	15	18
	22	10	22
	23	9	25
	24	9	5

1) Chacun des sujets de l'expérience a été affecté au hasard à l'un des quatre groupes g11, g21, g12, g22. Par ailleurs chaque sujet passe les deux planches p1 et p2. On appellera:

P2 = {p1, p2} le facteur "Planches";

G4 = {g11, g21, g12, g22} le facteur "Groupe";

S le facteur "sujets".

Donner une formule verbale et une justification de chacune des écritures suivantes:

S<G4> ; S*P2 ; S<G4>*P2

2) On considère les facteurs:

Apprentissage A2 = {a1, a2} avec a1: apprentissage préalable, a2: pas d'apprentissage préalable.

Ordre O2 = {o1, o2} avec o1: planche p1 puis planche p2, o2: planche p2 puis planche p1

Donner une formulation verbale et une justification de l'écriture suivante: O2*A2.

Donner enfin une formulation verbale de l'écriture suivante: S<O2*A2>*P2

3) A l'intérieur de chacun des 4 groupes, on a calculé la moyenne et la variance corrigée relatives à chaque planche. Dans le tableau suivant, on trouvera les résultats relatifs à la planche 1 (les résultats relatifs à la planche 2 n'étant pas nécessaires pour les calculs demandés par la suite); dans chaque case, la première valeur est la moyenne, la deuxième la variance corrigée.

condition a1		
	planche p1	planche p2
g11	12,3 16,67	
g21	7,3 6,67	

condition a2		
	planche p1	planche p2
g12	13,3 91,07	
g21	9,8 8,57	

L'objectif principal de l'expérience est d'examiner s'il y a une différence entre les conditions a1 et a2. On se bornera ici à cet examen pour la planche p1 seulement, à l'intérieur de chacun des ordres de passation. Donc, on examinera les 2 comparaisons g11, g12 et g21, g22.

a) Que suggère l'examen à vue des données, indépendamment de toute procédure d'inférence statistique?

b) Examiner d'un point de vue inférentiel chacune des 2 comparaisons indiquées; on pourra notamment procéder à un test de comparaison de 2 moyennes. Commenter brièvement les résultats obtenus.

Réponses : 1) $S < G_4 >$: chaque sujet est affecté à un seul groupe; le facteur "sujet" est emboîté dans le facteur "groupe".

$S * P_2$: chaque sujet passe par les deux niveaux du facteur "planche". Le facteur "sujet" est croisé avec le facteur "planche".

2) $O_2 * A_2$: chaque niveau du facteur "apprentissage" est combiné avec chaque niveau du facteur "ordre". Les facteurs O_2 et A_2 sont croisés.

$S < O_2 * A_2 > * P_2$: Chaque sujet est affecté à un ordre donné et un type d'apprentissage donné. Chaque sujet passe par les deux modalités du facteur "planche".

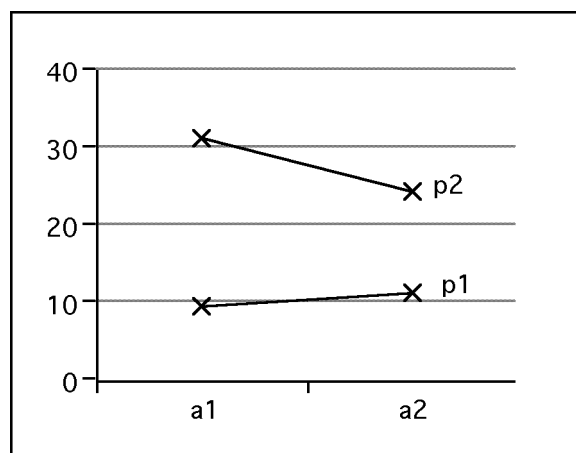
3) b) On compare g_{11} et g_{12} à l'aide d'une comparaison de moyennes sur des groupes indépendants. $t_{obs} = -0.24$, qui n'est pas significatif d'une différence entre les deux groupes. De même, la comparaison des groupes g_{21} et g_{22} aboutit à $t_{obs} = -1.56$, dont le niveau de significativité n'est que de 7.4%.

Remarques. On obtiendrait des résultats plus significatifs en étudiant les scores obtenus sur la planche p2.

Le tableau complet d'analyse de variance donne ici :

Source	ddl	SC	CM	F_{cal}	Pr
O_2	1	96.33	96.33		
A_2	1	200.08	200.08	2.75	11%
$O_2 A_2$	1	2.08	2.08		
$S(O_2 A_2)$	20	1457.51	72.88		
P_2	1	2914.08	2914.08	89.22	**
$O_2 P_2$	1	24.08	24.08		
$A_2 P_2$	1	408.33	408.33	12.5	0.2%
$O_2 A_2 P_2$	1	16.33	16.33		
Résidu	20	653.17	32.66		
Total	47	5772			

L'effet du facteur "planche" est très significatif. L'interaction entre les facteurs "apprentissage" et "planche" est également très significatif et peut être illustré par un graphe d'interaction.



Énoncé 41 Dossier “Syssau”

Dans un article publié en 1996, A. Syssau et D. Brouillet étudient le rôle de la nature et de la valeur affective d’un texte dans la récupération du souvenir chez les personnes âgées. L’expérience a concerné 80 sujets répartis en 2 groupes : 40 sujets présentant un déficit mnésique et 40 sujets non déficitaires. Quatre textes ont été utilisés, correspondant au croisement des deux modalités du facteur “type de texte” (narratif v/s descriptif) et des deux modalités du facteur “connotation” (affective v/s neutre). On a constitué 8 groupes de 10 sujets : chaque texte est proposé à un groupe de sujets déficitaires et un groupe de sujets non-déficitaires.

La performance des sujets est mesurée par le nombre de propositions correctement rappelées (score de 0 à 20).

Les auteurs font les hypothèses suivantes :

Quel que soit le groupe, les textes narratifs seront mieux restitués que les textes descriptifs, avec des performances, dans le groupe non déficitaire, supérieures à celles du groupe déficitaire. De plus, la charge affective facilitera la restitution des textes, dans le groupe déficitaire. Ainsi, l’association “charge affective et structure narrative” aurait pour conséquence une atténuation des différences de performances entre le groupe déficitaire et non déficitaires.

Les résultats observés dans une reprise de cette expérience sont rassemblés dans le tableau 11.

- 1) a) Quelles sont les variables indépendantes (ou facteurs de variation) prises en compte ? Quel est le nombre de niveaux de chacun des facteurs ?
- b) Quelle est la variable dépendante ?
- c) Écrire le plan d’expérience correspondant.
- 2) a) Compléter le tableau 13 en calculant les carrés moyens et les statistiques F de Fisher qui sont remplacés par “...” dans l’énoncé.
- b) Utiliser ce tableau pour déterminer si les effets principaux des facteurs déficit mnésique, type de texte et connotation affective sont significatifs.

3) *Étude des interactions*

- a) Calculer les scores moyens des 4 groupes obtenus en croisant les deux facteurs déficit et type de texte, indépendamment de la connotation affective du texte.
- b) Réaliser un graphe montrant l’absence d’interaction entre ces deux facteurs.
- c) Réaliser une étude analogue montrant l’interaction entre le déficit mnésique et la connotation affective du texte.
- d) Utiliser le tableau d’analyse de variance (tableau 13) pour confirmer au niveau inférentiel les résultats précédents.

4) Les auteurs affirment :

Quel que soit le texte, lorsque la connotation affective est présente, les sujets déficitaires ont des performances améliorées.

- a) Justifier cette affirmation, pour les textes descriptifs, à l’aide d’un test unilatéral de comparaison de moyennes.

Les auteurs poursuivent :

Ils ont alors des performances comparables à celles des sujets non déficitaires.

b) Justifier cette affirmation en s'appuyant sur l'un des tableaux d'analyse de variance fournis.

Les auteurs affirment également :

Quel que soit le texte, si la connotation est neutre, les sujets non déficitaires obtiennent des résultats supérieurs aux sujets déficitaires.

c) L'un des tableaux d'analyse de variance permet de justifier en partie cette conclusion. Lequel? Pourquoi la justification n'est-elle que partielle?

5) Au vu des résultats obtenus, les hypothèses des auteurs sont-elles vérifiées?

	texte narratif		texte descriptif	
	connotation affective	connotation neutre	connotation affective	connotation neutre
déficitaires	0	2	1	1
	2	0	3	0
	1	5	0	0
	8	9	4	1
	2	1	1	0
	4	0	2	1
	1	0	2	1
	0	7	1	1
	6	3	0	0
	9	0	0	0
non déficitaires	0	0	0	6
	4	2	0	4
	2	3	3	2
	5	6	4	0
	1	3	2	0
	2	8	0	1
	6	6	0	1
	2	4	2	3
	0	3	1	0
	2	5	0	5

Tableau 11: Données observées

		texte narratif		texte descriptif	
		connotation affective	connotation neutre	connotation affective	connotation neutre
déficitaires	\bar{x}	3.3	2.7	1.4	0.5
	s	3.13	3.10	1.28	0.50
	s_c	3.30	3.27	1.35	0.53
non déficitaires	\bar{x}	2.4	4.0	1.2	2.2
	s	1.91	2.19	1.40	2.09
	s_c	2.01	2.31	1.48	2.20

Tableau 12: moyennes et écarts types par groupe

Sources de var.	ddl	SC	CM	F
Déficit \mathcal{D}	1	4.513	4.513	...
Type texte \mathcal{T}	1	63.012
Connotation \mathcal{C}	1	1.513
Interaction $\mathcal{D} \times \mathcal{T}$	1	1.513
Interaction $\mathcal{D} \times \mathcal{C}$	1	21.013
Interaction $\mathcal{T} \times \mathcal{C}$	1	1.013
Inter. $\mathcal{D} \times \mathcal{T} \times \mathcal{C}$	1	0.112	0.112	0.02
Résidu	72	360.700	5.010	
Total	79	453.388		

Tableau 13: Analyse de variance (ensemble des facteurs)

Sources de var.	ddl	SC	CM	F	Prob.
Déficit \mathcal{D}	1	3.02	3.02	0.64	0.43
Type texte \mathcal{T}	1	24.02	24.02	5.07	0.03
Interaction $\mathcal{D} \times \mathcal{T}$	1	1.22	1.22	0.26	0.61
Résidu	36	170.50	4.74		
Total	39	198.77			

Tableau 14: Analyse de variance (connotation affective)

Sources de var.	ddl	SC	CM	F	Prob.
Déficit \mathcal{D}	1	22.50	22.50	4.26	0.046
Type texte \mathcal{T}	1	40.00	40.00	7.57	0.009
Interaction $\mathcal{D} \times \mathcal{T}$	1	0.40	0.40	0.08	0.785
Résidu	36	190.20	5.28		
Total	39	253.10			

Tableau 15: Analyse de variance (connotation neutre)