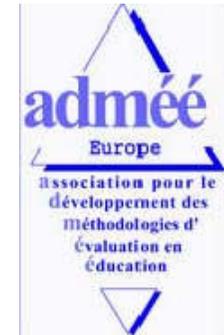




*Colloque ADMEE 2005*



# Le concept de cohérence spectrale appliqué à l'analyse d'items de tests standardisés

*Jean-Luc Gilles*

Symposium « Recherches récentes sur les degrés de certitude »  
Colloque ADMEE Europe – IUFM-Ardenne – 24-25 octobre 2005



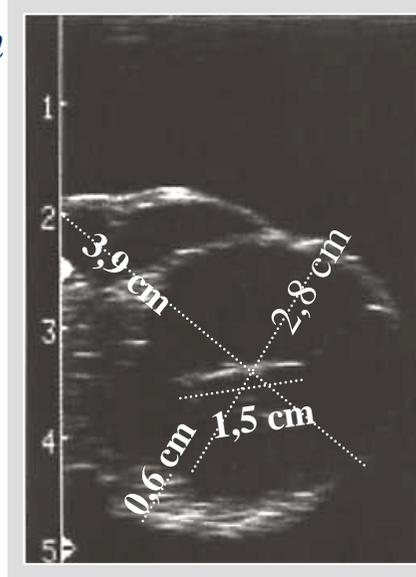
*Système Méthodologique d'Aide à la Réalisation de Tests*  
*Service de didactique des disciplines et intervention éducative*  
*Université de Liège - Belgique*



## Introduction:

*Exemple de question à choix multiple :*  
(Hansen & al, 2000)

*Exemple d'échelle de degrés de certitude (Leclercq & Gilles, 2000):*  
L'étudiant est invité à accompagner son choix par le pourcentage de chance qu'il attribue à sa réponse d'être correcte.



Cette échographie de l'ovaire gauche d'une vache pie noire de 6 ans, a été réalisée une semaine après l'oestrus. Votre diagnostic :

1. Un follicule de de Graaf
2. Un kyste folliculaire
3. Un Corps Jaune
4. Un Corps Jaune Cavitaire
5. Un artefact
6. Aucune
7. Toutes

Si vous considérez que votre réponse a une probabilité d'être correcte de

	cochez
0 %	0
20 %	1
40 %	2
60 %	3
80 %	4
100 %	5

*Avec un dispositif de lecture optique de marques il est possible de lire automatiquement les réponses et les certitudes de certaines d'étudiants :*

11	1	2	3	4	5	6	7	8	9
R	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C	0	1	2	3	4	5			
	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>				

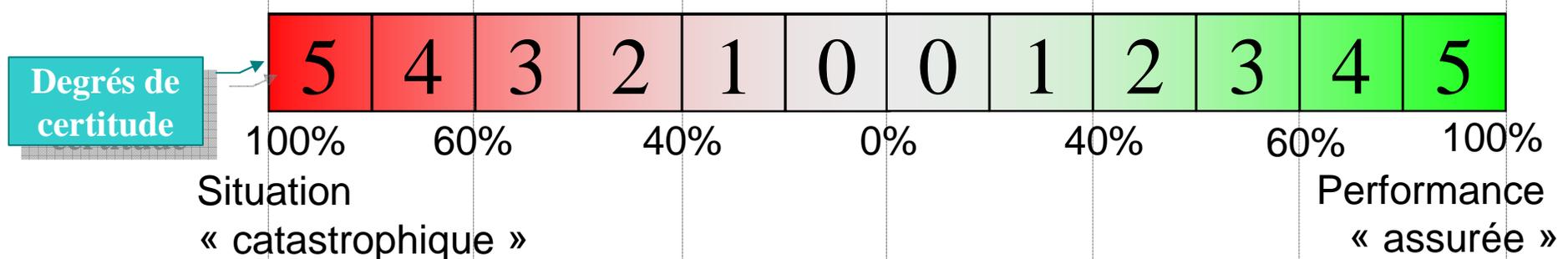
## Introduction:

- Sans l'utilisation des degrés de certitude, la solution choisie par l'étudiant est soit correcte, soit incorrecte.

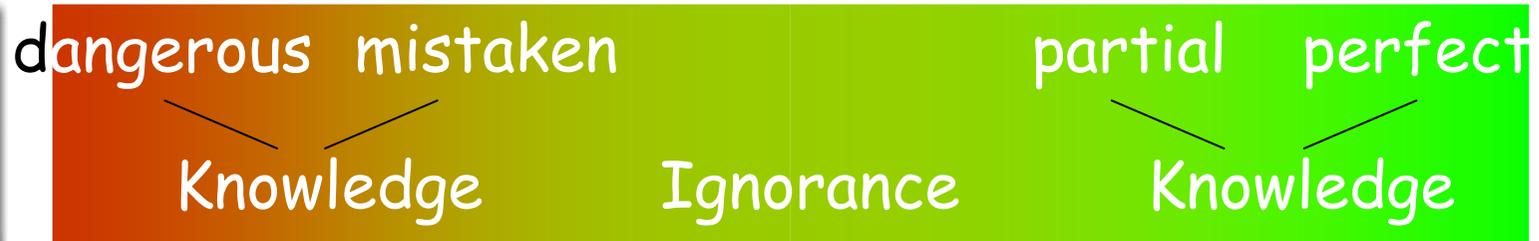
La solution est :



- Avec la technique des degrés de certitude, il est possible de dépasser le caractère binaire de l'évaluation des acquis et de proposer une conception plus spectrale des performances d'une personne



- Avec les pourcentages de certitude il est possible de fournir des feedbacks plus nuancés à propos des performances de chaque personne !



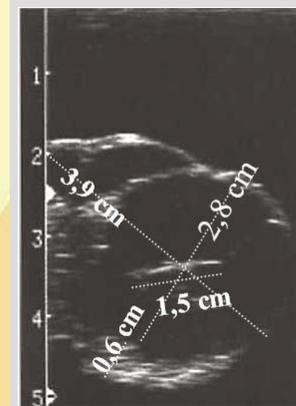
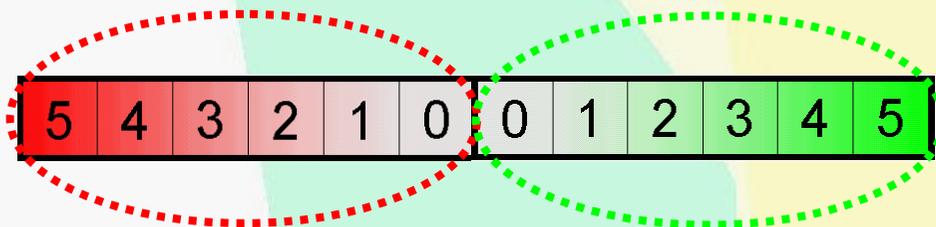
(Jans & Leclercq, 1999)

## Intuition de départ : utiliser les informations spectrales pour mettre en évidence les propositions des QCM qui contiennent des anomalies

Degrés de certitude des étudiants qui ont fourni une réponse incorrecte

Degrés de certitude des étudiants qui ont fourni une réponse correcte

Nous avons exploité les degrés de certitude des étudiants pour fournir de l'information "spectrale" à propos de la qualité des questions (et non à propos de la qualité des performances des étudiants).



Cette échographie de l'ovaire gauche d'une vache pie noire de 6 ans, a été réalisée une semaine après l'oestrus. Votre diagnostic :

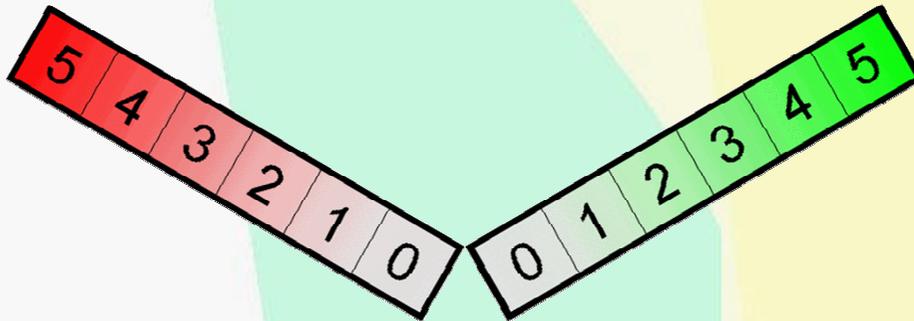
1. Un follicule de de Graaf
2. Un kyste folliculaire
3. Un Corps Jaune
4. Un Corps Jaune Cavitaire
5. Un artefact
6. Aucune
7. Toutes

# Qualité spectrale d'une question

**Intuition de départ : utiliser les informations spectrales pour mettre en évidence les propositions des QCM qui contiennent des anomalies**

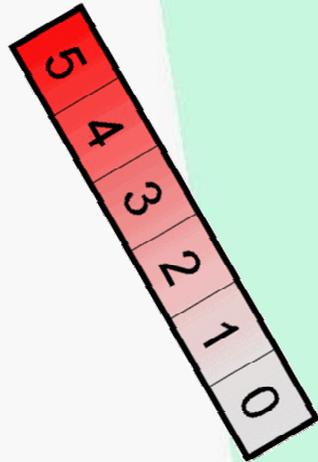
Degrés de certitude  
des étudiants qui ont  
fourni une réponse  
incorrecte

Degrés de certitude  
des étudiants qui ont  
fourni une réponse  
correcte

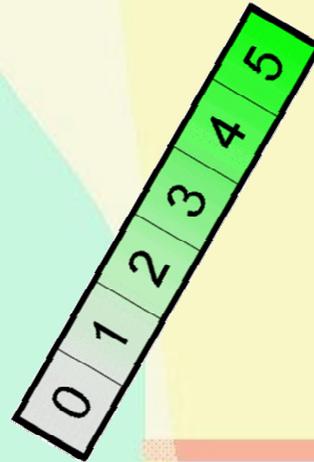


**Intuition de départ : utiliser les informations spectrales pour mettre en évidence les propositions des QCM qui contiennent des anomalies**

Degrés de certitude  
des étudiants qui ont  
fourni une réponse  
incorrecte

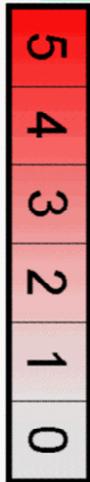


Degrés de certitude  
des étudiants qui ont  
fourni une réponse  
correcte



**Intuition de départ : utiliser les informations spectrales pour mettre en évidence les propositions des QCM qui contiennent des anomalies**

Degrés de certitude  
des étudiants qui ont  
fourni une réponse  
incorrecte



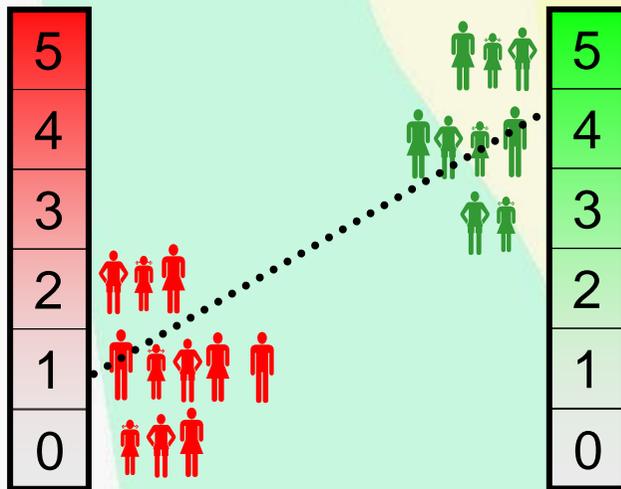
Degrés de certitude  
des étudiants qui ont  
fourni une réponse  
correcte



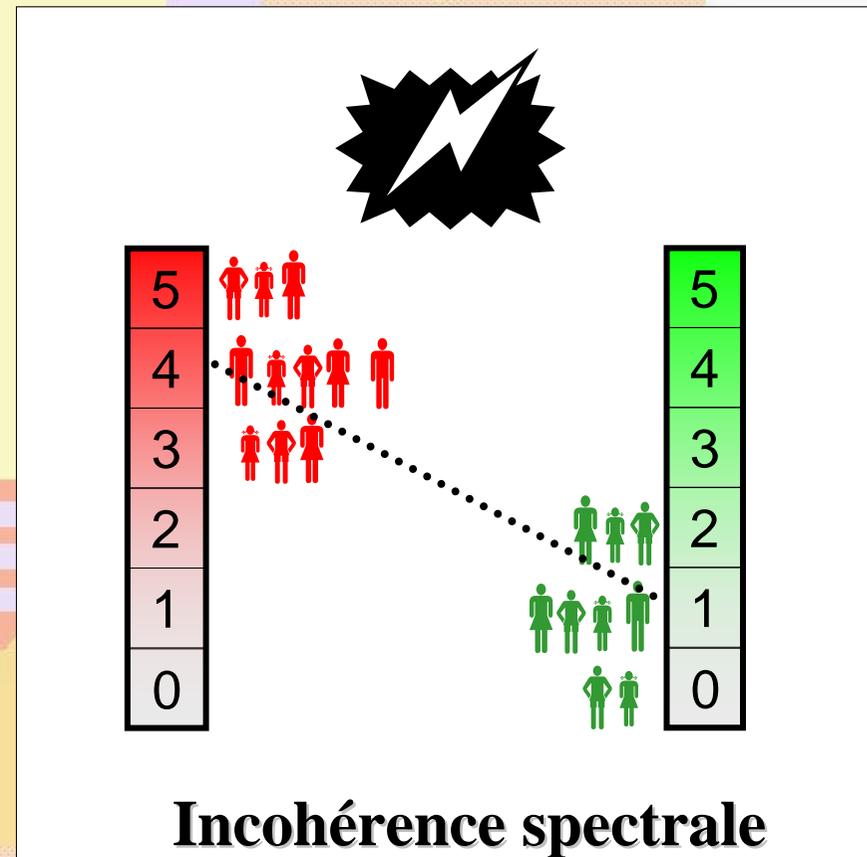
**Intuition de départ : utiliser les informations spectrales pour mettre en évidence les propositions des QCM qui contiennent des anomalies**

Degrés de certitude des étudiants qui ont fourni une réponse incorrecte

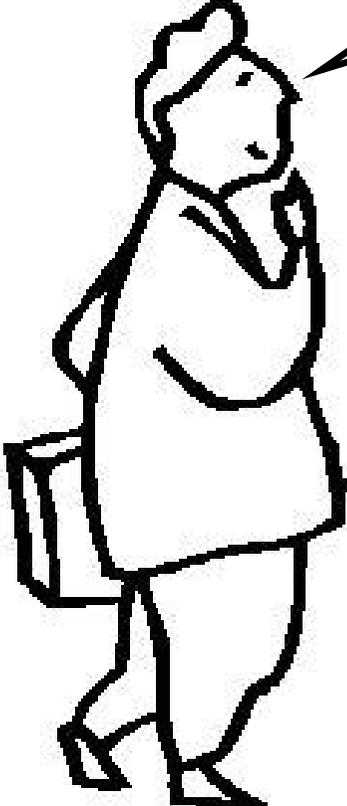
Degrés de certitude des étudiants qui ont fourni une réponse correcte



**Cohérence spectrale**



**Incohérence spectrale**



Comment calculer le  
niveau de cohérence spectrale  
des items d'une QCM?

Nous proposons de distinguer  
le cas de la cohérence spectrale  
d'une **réponse correcte** de celui  
de la cohérence spectrale  
d'une **réponse incorrecte**...

# Corrélation point bisériale Spectrale Contrastée d'une solution correcte :

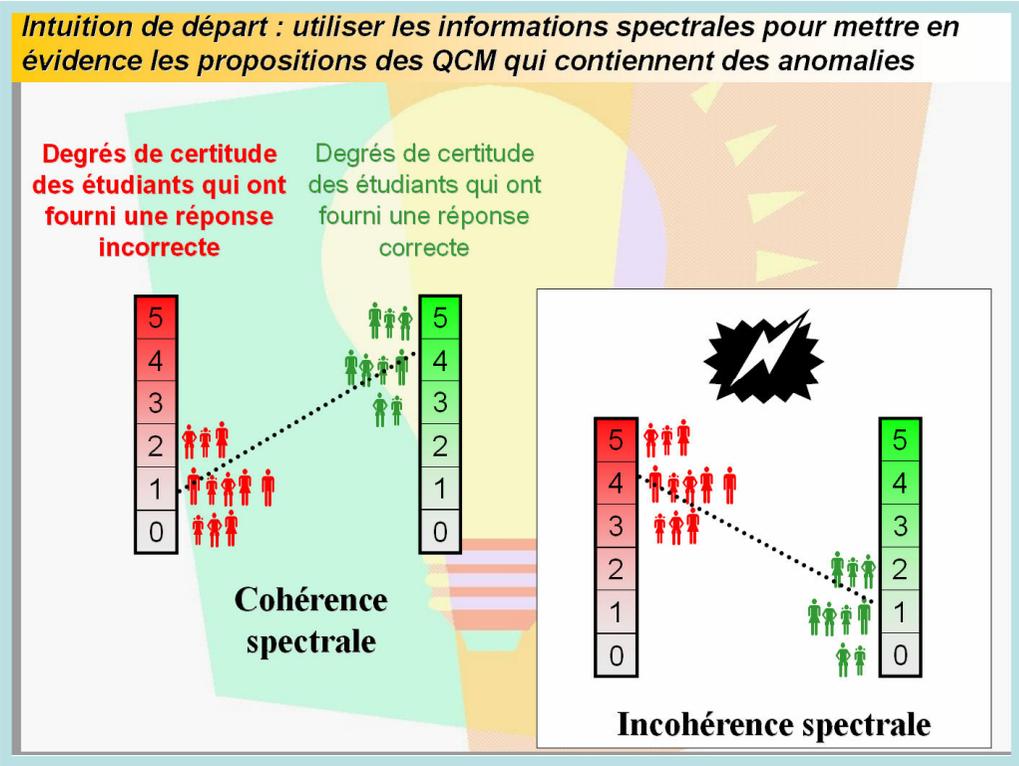
$$r_{pb} SC^c$$

Complete:  
 He ..... a heart attack while he ..... for the bus.

1. was having / was waiting
2. had / waited
3. had / was waiting ← ..... Réponse correcte
4. was having / waited

	OM	1. was having / was waiting S1	2. had / waited S2	3. had / was waiting S3	4. was having / waited S4	DC
Student 1	0	0	0	0	1	0
Student 2	0	1	0	0	0	1
Student 3	0	0	0	1	0	4
Student 4	0	0	0	1	0	5
Student 5	1	0	0	0	0	0
Student 6	0	0	0	1	0	3
Student 7	1	0	0	0	0	0
Student 8	0	0	0	0	1	3
Student 9	0	0	0	1	0	2
Student 10	0	0	1	0	0	0
Student 11	0	0	0	0	1	1
Student 12	0	1	0	0	0	2
Student 13	0	0	0	1	0	5
Student 14	0	1	0	0	0	1
Student 15	0	0	1	0	0	0
Student 16	0	0	1	0	0	3
Student 17	0	0	1	0	0	2
Student 18	0	0	0	0	1	2
Student 19	0	1	0	0	0	0
Student 20	0	0	0	1	0	4

Degrés de certitude (DC) des étudiants.



# Corrélation point bisériale Spectrale Contrastée d'une solution correcte :

$$r_{pb} SC^c$$

Quels sont les degrés de certitude des étudiants qui ont choisi la solution correcte « S3 » ?

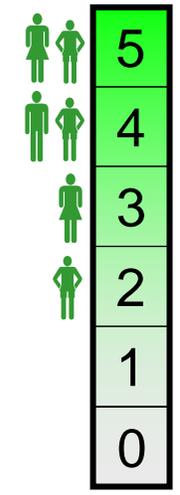
Complete:  
 He ..... a heart attack while he ..... for the bus.

1. was having / was waiting
2. had / waited
3. had / was waiting ← Réponse correcte
4. was having / waited

	S3	DC
Student 1	0	0
Student 2	0	1
Student 3	1	4
Student 4	1	5
Student 5	0	0
Student 6	1	3
Student 7	0	0
Student 8	0	3
Student 9	1	2
Student 10	0	0
Student 11	0	1
Student 12	0	2
Student 13	1	5
Student 14	0	1
Student 15	0	0
Student 16	0	3
Student 17	0	2
Student 18	0	2
Student 19	0	0
Student 20	1	4

Degrés de certitude (DC) des étudiants.

Degrés de certitude liés à la solution correcte



# Corrélation point bisériale Spectrale Contrastée d'une solution correcte :

$$r_{pb} SC^c$$

Degrés de certitude des étudiants qui ont répondu incorrectement ?

Complete:  
 He ..... a heart attack while he ..... for the bus.

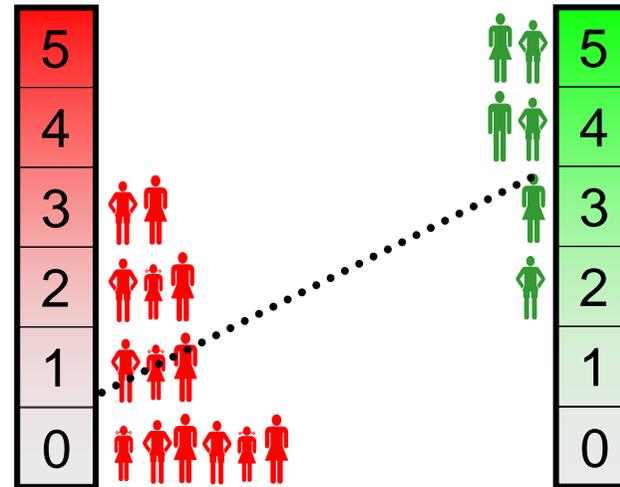
1. was having / was waiting
2. had / waited
3. had / was waiting ← Réponse correcte
4. was having / waited

Degrés de certitude (DC) des étudiants.

Student	DC	CD
Student 1	0	0
Student 2	0	1
Student 3	1	4
Student 4	1	5
Student 5	0	0
Student 6	1	3
Student 7	0	0
Student 8	0	3
Student 9	1	2
Student 10	0	0
Student 11	0	1
Student 12	0	2
Student 13	1	5
Student 14	0	1
Student 15	0	0
Student 16	0	3
Student 17	0	2
Student 18	0	2
Student 19	0	0
Student 20	1	4

Degrés de certitude liés aux solutions incorrectes

Degrés de certitude liés à la solution correcte



## Corrélation point bisériale Spectrale Contrastée d'une solution correcte :

$r_{pb} SC^c$

$$r_{pb} SC^c = \frac{Dx^c - Da^c}{\sigma} \sqrt{pq}$$

avec

$Dx^c$  = la moyenne des Degrés de certitude des étudiants qui ont choisi la solution correcte

$Da^c$  = la moyenne des Degrés de certitude des autres étudiants qui n'ont pas choisi la solution correcte

$\sigma$  = l'écart type des degrés de certitude de tous les étudiants

$p$  = la proportion d'étudiants qui ont choisi la solution correcte de la question

$q$  = la proportion d'étudiants qui n'ont pas choisi la solution correcte de la question ( $q = 1 - p$ )

# Corrélation point bisériale Spectrale Contrastée d'une solution correcte :

$$r_{pb} SC^c$$

Complete:  
 He ..... a heart attack while he ..... for the bus.

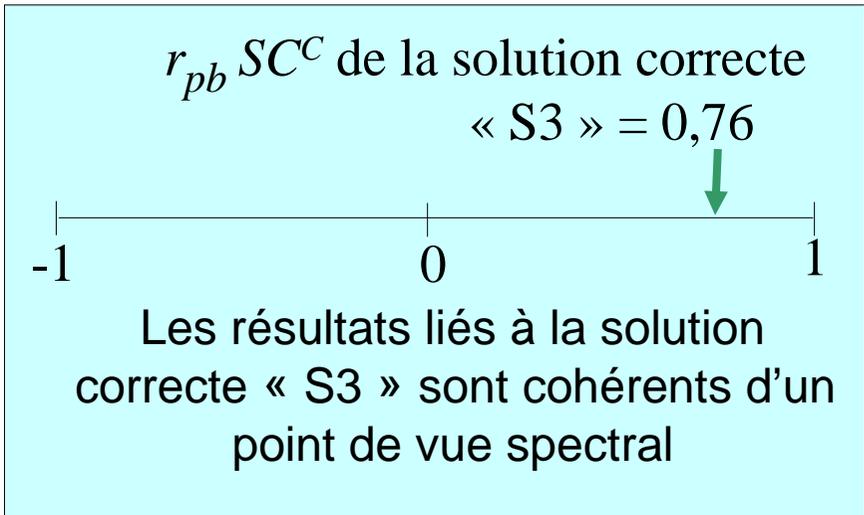
1. was having / was waiting
2. had / waited
3. had / was waiting ← Réponse correcte
4. was having / waited

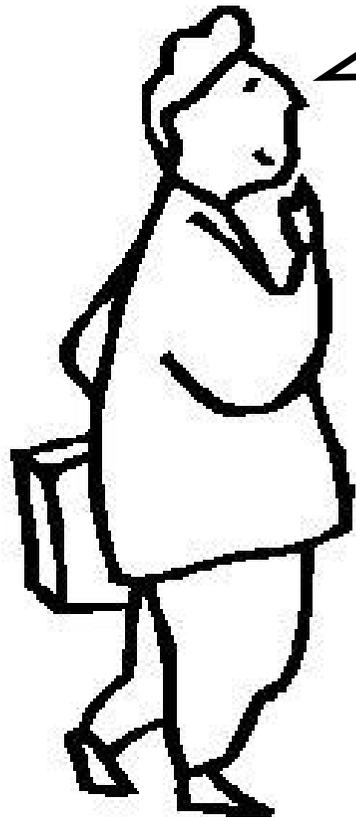
Degrés de certitude (DC) des étudiants.

En appliquant la formule :

$$r_{pb} SC^c = \frac{Dx^c - Da^c}{\sqrt{pq}}$$

Student	Choice	CD
Student 1	0	0
Student 2	0	1
Student 3	1	4
Student 4	1	5
Student 5	0	0
Student 6	1	3
Student 7	0	0
Student 8	0	3
Student 9	1	2
Student 10	0	0
Student 11	0	1
Student 12	0	2
Student 13	1	5
Student 14	0	1
Student 15	0	0
Student 16	0	3
Student 17	0	2
Student 18	0	2
Student 19	0	0
Student 20	1	4





...et en ce qui concerne  
la mesure de la cohérence spectrale  
d'une solution incorrecte ?

C'est ici que nous proposons  
un traitement **contrasté** !

# Corrélation point bisériale Spectrale Contrastée d'une solution incorrecte:

$$r_{pb} SC^I$$

Solution incorrecte  
S1

Solution correcte  
S3

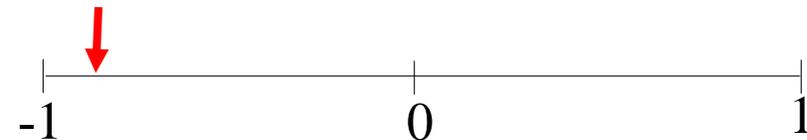
DC

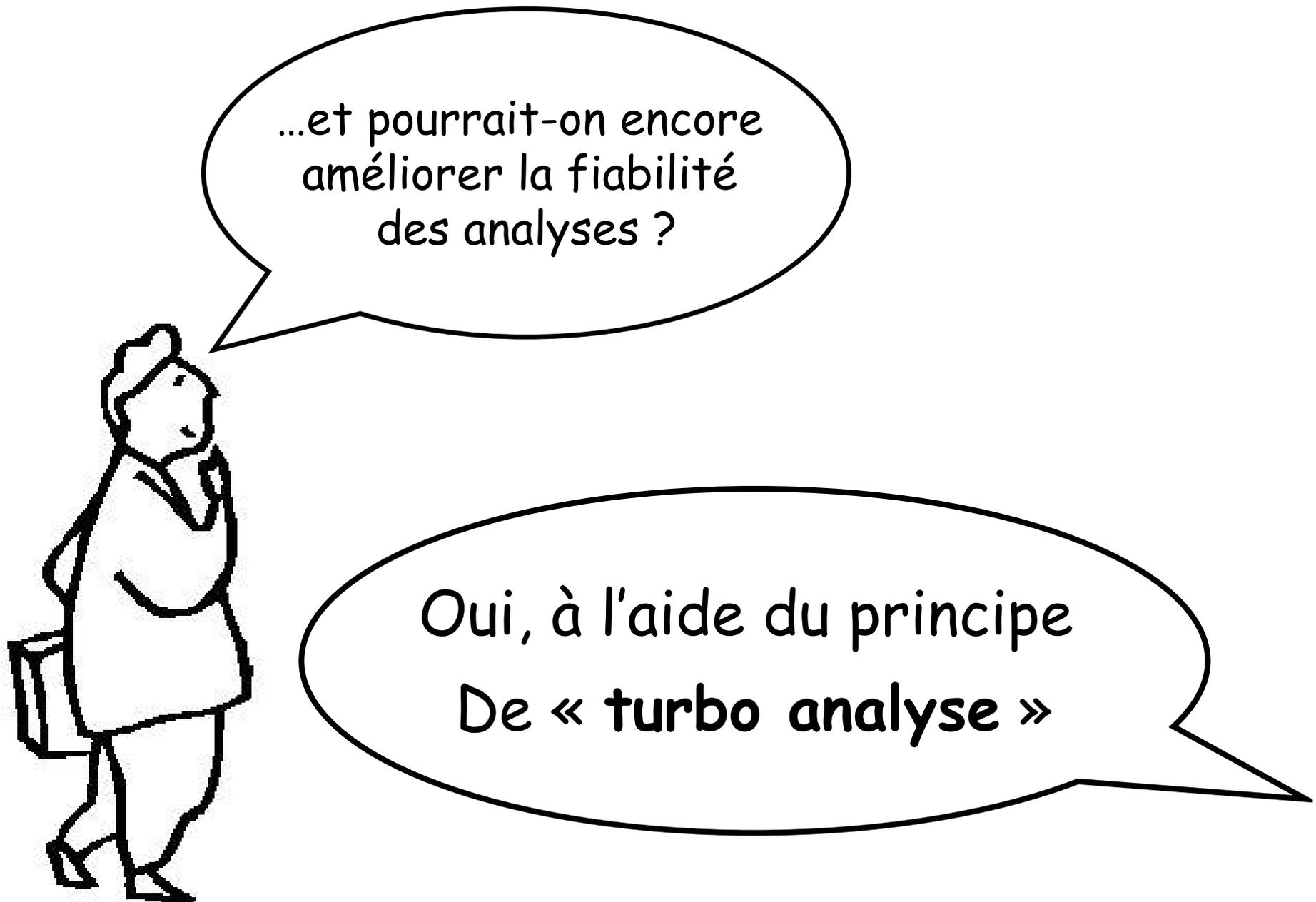
	S1	S3	DC
Student 2	1	0	1
Student 3	0	1	4
Student 4	0	1	5
Student 6	0	1	3
Student 9	0	1	2
Student 12	1	0	2
Student 13	0	1	5
Student 14	1	0	1
Student 19	1	0	0
Student 20	0	1	4

Traitement contrasté  
=  
Elimination du  
« bruit de fond »  
généralisé par les données des  
autres solutions incorrectes

Après traitement contrasté,  
nous calculons la corrélation  
point bisériale en utilisant les  
colonnes « S1 » et « DC »:

$$r_{pb} SC^I \text{ de « S1 »} \\ = -0.83$$





...et pourrait-on encore  
améliorer la fiabilité  
des analyses ?

Oui, à l'aide du principe  
De « turbo analyse »

# Principe de turbo analyse et intégration du calcul du réalisme dans le cadre des corrélations point bisérialles Spectrales Contrastées ( $r_{pb}SCT$ ) :

Complete:

He ..... a heart attack while he ..... for the bus.

1. was having / was waiting
2. had / waited
3. had / was waiting
4. was having / waited

← Réponse correcte

	S3	DC	Réalisme (Rs)
Student 1	0	0	55
Student 2	0	1	86
Student 3	1	4	81
Student 4	1	5	92
Student 5	0	0	83
Student 6	1	3	74
Student 7	0	0	43
Student 8	0	3	76
Student 9	1	2	43
Student 10	0	0	93
Student 11	0	1	82
Student 12	0	2	66
Student 13	1	5	89
Student 14	0	1	82
Student 15	0	0	94
Student 16	0	3	34
Student 17	0	2	72
Student 18	0	2	87
Student 19	0	0	86
Student 20	1	4	84

Réalisme ( $R_s$ ) de chaque étudiant

← Très bon réalisme !

← Mauvais réalisme !

# Principe de turbo analyse et intégration du calcul du réalisme dans le cadre des corrélations point bisérialles Spectrales Contrastées ( $r_{pb}SCT$ ) :

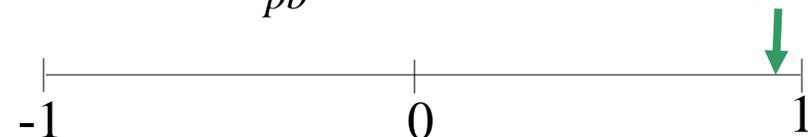
**Turbo analyse = Augmentation de la fiabilité des indices spectraux par la sélection des données des sujets qui commettent peu d'erreurs dans leurs auto-estimations**

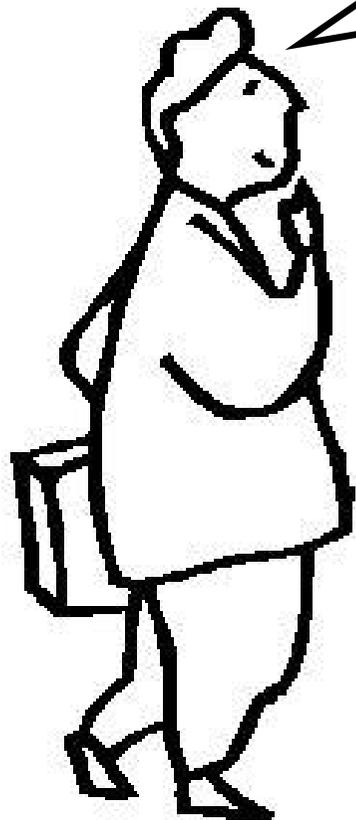
Elimination des données des étudiants qui obtiennent un Réalisme ( $R_s$ ) inférieur à 80...

	S3	CD	REALISM ( $R_s$ )
Student 1	0	0	55
Student 2	0	1	86
Student 3	1	4	81
Student 4	1	5	92
Student 5	0	0	83
Student 6	1	3	74
Student 7	0	0	43
Student 8	0	3	76
Student 9	1	2	43
Student 10	0	0	93
Student 11	0	1	82
Student 12	0	2	66
Student 13	1	5	89
Student 14	0	1	82
Student 15	0	0	94
Student 16	0	3	34
Student 17	0	2	72
Student 18	0	2	87
Student 19	0	0	86
Student 20	1	4	84

...Calcul de la corrélation point bisériale Spectrale Contrastée

$$r_{pb}SCT80 \text{ of } \ll S3 \gg = 0,94$$





...et ces indices spectraux  
ont-ils été testés sur  
de vraies données ?

Oui, à l'aide des  
milliers de réponses et certitudes  
des épreuves **MOHICAN**

## **“Full-size” test**

Nous avons testé ces nouveaux indices de corrélation point bisériale Spectrale Contrastée à l'aide des données liées aux milliers de réponses et certitudes collectées dans le cadre des dix tests standardisés du projet MOHICAN soutenu par la Communauté française de Belgique et coordonné par le Service de Technologie de l'Education de l'Université de Liège.

Le but du projet MOHICAN (Leclercq & al, 2003) était de faire le point sur les connaissances des étudiants entrant à l'Université.

## **Banque de questions du projet “MOHICAN” :**

- Vocabulaire (VOC) – 45 QCM
- Syntaxe et articulation logique (SYN) – 12 QCM
- Compréhension de texte (COM) – 6 QCM
- Lecture de documents en géographie (GEO) – 10 QCM
- Mathématiques (MAT) – 22 QCM
- Physique (PHY) – 10 QCM
- Chimie (CHE) – 8 QCM
- Biologie (BIO) – 10 QCM
- Connaissances artistiques (ART) – 25 QCM
- Connaissances en histoire et Socio-économie (HIS) – 25 QCM

## **Population:**

Les institutions qui ont participé au projet :

- l'Université Catholique de Louvain (UCL);
- l'Université Libre de Bruxelles (ULB),
- l'Université de Liège (ULg);
- les Facultés Universitaires Notre-Dame de la Paix de Namur;
- l'Université de Mons-Hainaut (UMH);
- les Facultés Universitaires Catholiques de Mons (FUCAM);
- les Facultés Universitaires Saint-Louis (FUSL);
- les Facultés Polytechniques de Mons (FPM);
- les Facultés Universitaires Agronomiques de Gembloux (FUSAG).

## Population:

	<i>n</i>
Vocabulary test (VOC)	3,846
Syntax and logical structuring of ideas test (SYN)	3,739
Text comprehension test (COM)	3,420
Reading of maps, graphics and tables in geography test (GEO)	3,688
	<i>n</i>
Mathematics test (MAT)	2,516
Physics test (PHY)	2,497
Chemistry test (CHE)	2,501
Biology test (BIO)	2,507
	<i>n</i>
General knowledge in arts and culture test (ART)	1,392
General knowledge in History and Socio-economy test (HIS)	1,410



## Résultats des analyses spectrales de la qualité des QCM MOHICAN : Taille des populations des tests aux 10 niveaux de turbo analyse

*Size of the 10 tests' population (VOC, SYN, GEO, HIS, ART, MAT, BIO, CHE and PHY) at 10 levels of turbo analysis (from T0 to T90 with a step of 10)*

10 niveaux de turbo analyse ↓

	VOC	SYN	COM	GEO	HIS	ART	MAT	BIO	CHE	PHY
	<i>n</i>									
T0	3.846	3.739	3.420	3.685	1.410	1.392	2.516	2.507	2.501	2.497
T10	3.846	3.734	3.413	3.683	1.409	1.392	2.514	2.507	2.500	2.494
T20	3.843	3.720	3.394	3.678	1.407	1.391	2.499	2.497	2.482	2.472
T30	3.838	3.699	3.348	3.649	1.405	1.387	2.465	2.485	2.461	2.458
T40	3.815	3.658	3.208	3.600	1.399	1.380	2.434	2.442	2.419	2.394
T50	3.773	3.553	2.860	3.474	1.391	1.363	2.398	2.340	2.341	2.342
T60	3.602	3.318	2.279	3.116	1.340	1.323	2.278	2.036	2.096	2.129
T70	3.089	2.691	1.481	2.303	1.101	1.137	1.956	1.511	1.661	1.724
T80	1.806	1.611	756	1.162	584	681	1.276	757	909	1.007
T90	350	439	251	253	73	124	339	151	232	275

**Résultats des analyses spectrales de la qualité des QCM MOHICAN :  
 Comparaison des valeurs des  $r_{pb}$ ,  $r_{pb}^{SC}$  et  $r_{pb}^{SCT80}$  à une question  
 particulièrement discriminante**

*Values obtained by the 36th question of the vocabulary test*

		OM	P1	P2	P3
VOC Q36	$n$	157	* 2.080	323	1.252
	%	4	* 54	8	33
	$r_{pb}$ →	-0.27	* 0.51	-0.16	-0.30
	$r_{pb}^{SC}$ →	xxx	* 0.41	-0.22	-0.35
	$n^{T80}$	103	* 1.199	135	612
	% <sup>T80</sup>	5	* 58	7	30
	$r_{pb}^{SCT80}$ →	xxx	* 0.58	-0.28	-0.47

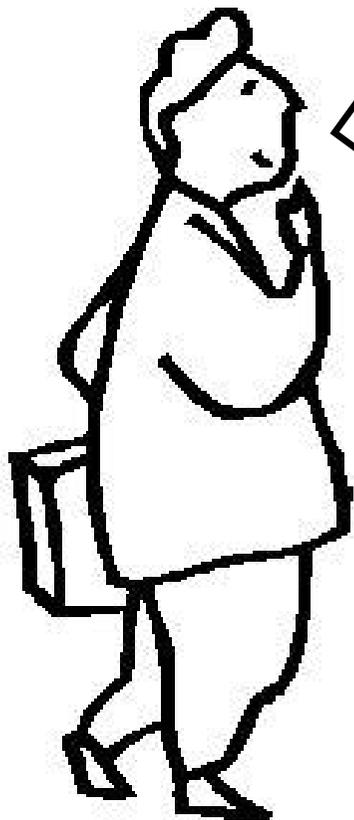
*Note.* OM = omission; P1, P2 and P3 = MCQ's solutions. Values obtained by the correct solution are preceded by a star (\*).

**Résultats des analyses spectrales de la qualité des QCM MOHICAN :  
 Comparaison des valeurs des  $r_{pb}$ ,  $r_{pb}^{SC}$  et  $r_{pb}^{SCT80}$  à une question  
 particulièrement discriminante**

*Values obtained by the 3<sup>rd</sup> question of the General knowledge in History and Socio-economy test*

		OM	P1	P2	P3	P4	P5	P6
HIS Q3	$n$	90	501	132	197	104	258	* 128
	%	6	36	9	14	7	18	* 9
	$r_{pb}$ →	-0.21	0.36	-0.02	-0.10	-0.06	-0.17	* 0.01
	$r_{pb}^{SC}$ →	xxx	0.49	0.05	0.07	0.04	0.19	* -0.15
	$n^{T80}$	55	218	67	97	52	111	* 54
	% <sup>T80</sup>	8	33	10	15	8	17	* 8
	$r_{pb}^{SCT80}$ →	xxx	0.38	0.00	0.02	-0.03	0.06	* -0.06

*Note.* OM = omission; P1 to P6 = MCQ's solutions. Values obtained by the correct solution are preceded by a star (\*).



L'analyse spectrale  
permet-elle de  
mieux mettre en évidence  
les questions problématiques  
que les analyses classiques ?

## CONCLUSIONS:

### Résumé des analyses spectrales et classiques :

- Les 10 tests MOHICAN = 173 QCM
- 16 QCM présentent des valeurs anormales aux corrélations classiques et/ou spectrales
- Les experts pointent 5 QCM qui présentent de réels problèmes

#### Concernant la DETECTION des 5 QCM pointées par les experts :

Corrélation classique ( $r_{pb}$ )	Détection de 2 QCM problématiques parmi les 5.
Corrélation spectrale ( $r_{pb SC}$ )	Détection des 5 QCM pointées par les experts.
Corrélation spectrale avec turbo analyse ( $r_{pb SCT80}$ )	Détection de 2 QCM problématiques parmi les 5.

#### Concernant les FAUSSES ALARMES (détection de problèmes là où les experts ne détectent pas d'anomalies particulières) :

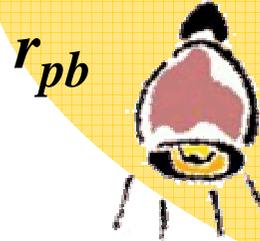
Corrélation classique ( $r_{pb}$ )	Déclenche 10 fausses alarmes.
Corrélation spectrale ( $r_{pb SC}$ )	Déclenche 7 fausses alarmes.
Corrélation spectrale avec turbo analyse ( $r_{pb SCT80}$ )	Déclenche 1 fausse alarme.

## CONCLUSIONS:

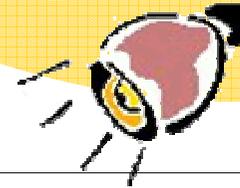
*Analyse qualitative :  
opinions des experts a posteriori*



*Analyse classique de discrimination :*



*Analyse spectrale :*  
 $r_{pb} SC$  &  $r_{pb} SCT$



① Du point de vue de la « **détection** »:

▶  $r_{pb} SC$  est plus efficace

② Du point de vue de la  
« **diminution de fausses alarmes** »:

▶  $r_{pb} SCT80$  est le plus efficace

▶  $r_{pb} SC$  est plus efficace que les  $r_{pb}$

③ L'analyse spectrale est « **test free** »

Merci pour votre attention !

Pour plus d'informations concernant  
l'analyse de la qualité spectrale  
des tests standardisés,

n'hésitez pas à me contacter :

[jl.gilles@ulg.ac.be](mailto:jl.gilles@ulg.ac.be)