UNIVERSITE DE LIEGE

Institut de Psychologie et des Sciences de l'Education

Année académique 1975-1976

L'évaluation subjective de la probabilité d'exactitude des réponses en situation pédagogique

Thèse de doctorat
en Sciences de l'Education
présentée par

Dieudonné LECLERCQ
Licencié en Sciences de l'Education

Exemplaire avec annotations de L. D'HAINAUT

UNIVERSITE DE LIEGE

Institut de Psychologie et des Sciences de l'Education

Année académique 1975-1976

L'évaluation subjective de la probabilité d'exactitude des réponses en situation pédagogique

Thèse de doctorat
en Sciences de l'Education
présentée par

Dieudonné LECLERCQ
Licencié en Sciences de l'Education

Tout au long de ce travail, nous avons été guidé et soutenu par la confiance et l'intérêt que le professeur G. DE LANDSHEERE nous a témoignés. Suivant son exemple, nous avons tenté de donner à notre démarche un maximum de rigueur. C'est à lui que vont nos remerciements les plus vifs.

Nous tenons à exprimer notre reconnaissance au professeur L. D'HAINAUT qui, lors de contacts chaleureux, a marqué notre formation de chercheur par des apports toujours efficaces.

Notre gratitude s'adresse aussi

au professeur H. BRENY qui nous a révélé l'importance des probabilités conditionnelles dans les sciences du comportement et a accepté de nous éclairer dans ce travail;

au professeur Ch. HEUCHENNE qui nous a ouvert sa bibliothèque, - si précieuse pour notre propos -, qui nous a permis de le consulter à maintes reprises et nous y a aidé;

au professeur R. PIRET qui a aimablement accepté de présider le jury du doctorat;

au professeur M. RICHELLE qui nous a initié, dans son Laboratoire, à l'analyse expérimentale du comportement et a déterminé nos options scientifiques fondamentales;

à D. DEFAYS, critique très attentif et constructif, qui nous a aidé tant sur le fond que sur la forme du travail;

à J. DONNAY, pour sa solidarité quotidienne;

à G. HENRY qui nous a fait profiter de sa lucidité dans la conduite de recherches expérimentales;

à R. DE BAL, B. GIOT et M.-T. WANNYN qui, durant une année, ont rendu moins lourdes nos charges d'enseignement programmé;

aux nombreux professeurs, étudiants et collègues qui se sont prêtés aux diverses expériences, souvent peu engageantes;

au colonel P. MOONEN et au commandant P. VAN EYLEN qui nous ont permis de réaliser la banque de questions gérée par ordinateur, instrument de base d'une partie importante de notre étude;

au commandant P. VAN ROY, pour le dynamisme avec lequel il accroît sans cesse l'efficacité du système BANKETFA et pour les données dont il nous a permis de disposer;

à Mlle D. COMANE qui a perforé les milliers de cartes-données relatives au chapitre 6;

à Mme C. ENGLEBERT qui a aimablement réalisé les nombreux tableaux de chiffres du présent volume et nous a aidé à élaborer maints documents;

à Mme N. SAENEN qui, avec beaucoup de patience, a dactylographié nos protocoles expérimentaux et des versions provisoires du texte actuel;

à Mme G. NIBUS qui a assuré avec talent la réalisation du texte définitif;

aux membres de notre famille, F. et J. BOXUS, D. et L. DUCHATEAU, soutiens permanents de notre vie quotidienne;

enfin à notre femme, E. BOXUS qui, en plus de son travail de recherche, a pris plus que sa part dans nos charges familiales.

SOMMAIRE

Approche individuelle : style cognitif

Analyse clinique

Vous n'utilisez pas les variable de personnalité

Il n'est pas dit que d'un individu à l'autre, à des degré de certitude égaux correspondent des degré de connaissance égaux.

"L'évaluation (...) a toujours, directement ou indirectement, rapport avec le progrès, en extension ou en qualité, de l'apprentissage."

G. DE LANDSHEERE, 1971, p. 14.

INTRODUCTION

- Objectifs et méthodes de l'étude.
- Mise au point terminologique.
- L'intérêt du problème pour l'enseignement.
- L'intérêt du problème pour la recherche.
- Les fondements épistémologiques de la méthode proposée.
- La validité, la fidélité et la sensibilité des mesures obtenues à partir de l'expression de la certitude.
- Les grands types de notation de la certitude.
- Les matrices de conséquences.
- Les conditions d'une étude valide des certitudes.
- Thèses et hypothèses générales.

Remarques préliminaires

Plusieurs expressions fréquemment employées sont fort longues. Afin d'alléger la présentation du texte, nous utilisons des sigles dont la liste se trouve sur une page volante.

Cette notation est particulièrement utile lorsqu'il s'agit de variables indicées. En outre, les sigles (noms de variables et indices) sont directement transposables en FORTRAN.

La plupart des valeurs numériques présentées sont comprises entre 0 et 1, puisqu'il s'agit de probabilités. Nous nous permettrons d'utiliser souvent la notation américaine (.75 pour 0,75), plus compacte.

N'y a-t-il pas liaison entre 1 degré de certitude et le degré de connaissance ?

OBJECTIFS ET METHODES DE L'ETUDE

> *"La probabilité qu'une unité de comportement donnée survienne à un moment donné constitue, dans une science du comportement, un fait de base."*
>
> *(B. SKINNER, 1971, p. 109)*

Le présent travail s'inscrit dans le mouvement de recherches axées sur la probabilité d'apparition d'un comportement, qui sera appelée PUR : probabilité d'utilisation d'une réponse par un sujet donné, à un moment donné.

L'étude porte essentiellement sur les relations entre cette probabilité et la probabilité d'exactitude d'une réponse ou probabilité de succès d'un comportement (PER). Nous nous sommes limité au seul cas où la réponse est correcte (1) ou incorrecte (0), sans nuance possible.

Les deux probabilités considérées (PUR et PER) sont étroitement liées : plus le succès d'une réponse (PER) est probable, plus l'utilisation (PUR) de cette réponse devrait l'être. La liaison entre la PER et la PUR est, cependant, plus complexe et a fait l'objet de plusieurs études, spécialement en sciences économiques et, plus généralement, dans de nombreux cas où il importe d'optimiser les décisions. Dans ces situations, le décideur connaît souvent les taux de succès observés dans le passé. Il arrive aussi qu'un modèle théorique préside à l'attribution de valeurs numériques à la probabilité d'exactitude de la réponse (PER). Par exemple, la connaissance des propriétés géométriques d'un dé à jouer (six surfaces égales délimitées par douze arêtes arrondies de la même façon) permet d'attribuer une probabilité de 1/6 à l'occurrence du caractère favorable (face 5 par exemple) lors d'un jet donné. La probabilité d'exactitude (PER) de ce jet est donc .166.

Je préférerais "estimation" à évaluation

Cette évaluation de la PER, que l'on parte de données expérimentales ou d'un modèle (le plus souvent mathématique) sera appelée EXPER : évaluation EXpérimentale de la PER. Remarquons que les données expérimentales et le modèle peuvent être connus de plusieurs décideurs et revêtent donc un caractère public. Dans le cas d'un dé à jouer (non pipé), tous les spécialistes s'accordent sur le modèle théorique et sur les résultats statistiques. Dans un tel cas (idéal, il est vrai), les probabilités (PER) sont dites objectives.

L'élève répondant à une question d'un test peut être comparé à un décideur. La décision, qu'elle soit consciente ou inconsciente, porte sur la réponse à fournir. Dans une telle perspective, la probabilité d'exactitude (PER) de chacune des réponses possibles importe grandement, car il s'agit de fournir la meilleure d'entre elles.

Or, dans une situation de testing pédagogique, l'élève ne connaît pas les taux de réussite des diverses réponses, calculés à partir d'expériences préalables. Donner à l'élève une telle information reviendrait à lui fournir la réponse correcte à la question. Chaque sujet possède cependant une connaissance personnelle (qui correspond au "modèle théorique" évoqué ci-dessus) des probabilités de succès des diverses solutions possibles.

Les dispositifs habituels d'évaluation pédagogique (procédures anticopiage) tentent d'empêcher que cette connaissance individuelle soit influencée par celle des autres individus. Le contrôle de l'exactitude des réponses fournies par l'individu, à l'aide d'ouvrages de référence ou par consultation de personnes compétentes, n'aura lieu qu'après la production de la réponse. Avant, l'élève ne dispose que d'éléments privés et subjectifs pour évaluer les PER. Selon une suggestion du professeur G. De Landsheere, une telle évaluation sera appelée ESPER : Evaluation Subjective de la PER.

Il paraît simple de demander à des étudiants - d'un certain âge (qui ont atteint le stade des opérations formelles) et possédant un quotient intellectuel suffisant - d'évaluer subjectivement la probabilité d'exactitude (ESPER) de diverses réponses à une question d'un test pédagogique (relevant du domaine cognitif). Pour faciliter la tâche, on permet aux élèves d'indiquer des degrés de certitude représentant des zones de probabilités. Par exemple, la certitude 1 pourrait signifier une ESPER comprise entre 0 et 20 %, une certitude 2 une ESPER comprise entre 20 et 40 %, etc.

Un certain nombre de recherches (en Italie, aux Etats-Unis et aux Pays-Bas principalement) se sont développées sur le sujet entre 1960 et 1972, mais ne semblent ni avoir abouti (1) ni avoir connu des prolongements auxquels on aurait pu s'attendre.

On se trouve dans une impasse car l'obstacle rencontré est d'importance : on ne dispose, jusqu'à présent, ni d'un cadre théorique, ni de données expérimentales permettant d'estimer la validité, la fidélité et la sensibilité de l'ESPER.

La validité de l'ESPER est, à raison, mise en doute : d'autres variables (personnalité, attrait ou peur du risque, etc.) que la probabilité subjective d'exactitude (ESPER) expliquent la certitude exprimée par le sujet. Divers auteurs ont abordé ce problème, soit sur le plan théorique (C.H. COOMBS) (2), soit sur le plan expérimental (S.S. JACOBS) (3). Le simple bon sens donne raison aux sceptiques : un indice de certitude ne peut être considéré comme l'expression sans biais de l'ESPER.

---

(1) Les quelques auteurs que nous avons pu consulter personnellement (R.F. VAN NAERSSEN, H. ROUANET, R.K. HAMBELTON, B. CHOPPIN) ont, provisoirement du moins, abandonné ce domaine de recherche.

(2) La théorie du dépliage développée par C.H. COOMBS (1950, 1960, 1967, 1970) s'attaque notamment au problème des préférences individuelles relatives à l'ampleur du risque.

(3) S.S. JACOBS (1968) a calculé la corrélation entre les scores individuels aux quatre échelles du CPI (California Psychological Inventory) et un indice de "confiance douteuse" (unwarranted confidence). La corrélation vaut .39. L'expérience portait sur 130 questions et 72 sujets.

En opposition avec la théorie d'Ausubel
pour qui le savoir c'est aller du vague au précis.
= en opposition avec la gestalt.

"En cas d'erreur, le faire bénéficier de sa franchise"
c'est mêler la reconnaissance d'un mérite à l'évaluation d'une
connaissance, c'est aussi substituer un jugement de valeur
à une tentative de mesure.

Le manque de validité des indices de certitude est sans doute la raison principale du ralentissement des recherches dans ce domaine.

Or, on le verra plus loin, l'ESPER est un outil potentiellement intéressant tant pour la pratique scolaire que pour la recherche en éducation. Malheureusement, mesurer l'ESPER de façon valide constitue un défi méthodologique. L'objet principal de notre travail consiste, précisément, à relever ce défi. En fait, notre thèse repose sur trois convictions.

1. La connaissance partielle (ni nulle, ni parfaite) existe et l'ESPER tient une place importante dans les processus cognitifs.
2. Trop souvent, en pédagogie, des consignes sommaires contraignent les élèves à répondre sans nuance "Je sais" ou "Je ne sais pas". Un tel appauvrissement des comportements liés à l'apprentissage n'est ni inéluctable ni souhaitable. Il est possible d'effectuer des mesures valides, fidèles et sensibles de la connaissance partielle qu'un individu a d'un contenu.
3. L'étudiant peut être conscient de l'imperfection de sa réponse. Il faut lui laisser exprimer ce doute et, en cas d'erreur, le faire bénéficier de sa franchise. Il importe, non seulement de cesser de pénaliser le doute, mais d'entraîner les étudiants à l'évaluer et à l'exprimer.

Avant d'illustrer la proposition ci-dessus et de l'exprimer en termes plus techniques, il importe de lever une ambiguïté relative au mot "doute". Dans le langage courant, ce mot recouvre deux concepts assez différents : l'ignorance ou la méconnaissance, d'une part, la remise en cause, d'autre part.

Le premier cas correspond à la situation de l'individu qui doute de l'exactitude de sa réponse, mais accepte, quand on la lui rappelle, la solution notoirement correcte.

Il s'agit plutôt de la construction d'une attitude de doute.

Le second cas correspond à la situation d'un individu qui, connaissant très bien la solution généralement acceptée comme correcte, en conteste la pertinence :

> "... c'est remettre en cause les choses dont on est le plus sûr qui fait faire à la science ses plus grands progrès." (L. D'HAINAUT, 1974, p. 60)

Une connaissance assurée est non seulement compatible avec leur remise en cause, mais en constitue fréquemment une condition nécessaire. En effet, ce sont, en général, des personnes très informées qui font progresser les sciences. Comment un étudiant ignorant la statistique saurait-il la critiquer valablement ?

Le présent travail ne portera que sur le premier type de doute (méconnaissance ou connaissance imparfaite); dans les cas qui seront envisagés, le sujet se rallie à la "solution communément acceptée". On ne traitera donc que des questions pour lesquelles on peut, sans ambiguïté, déterminer si une réponse est correcte ou non. L'exactitude de la réponse sera établie soit par un groupe d'experts, soit par un seul professeur, soit encore en consultant un ouvrage de référence généralement considéré comme sûr.

MISE AU POINT TERMINOLOGIQUE

Le terme "probabilité", trop souvent employé sans adjectif ou sans complément déterminatif, est polysémique.

Si l'on prenait la peine d'écrire toujours "probabilité de succès" d'une réponse, il n'y aurait pas d'ambiguïté. Mais souvent, on oublie de mentionner "de succès". Dans la littérature anglo-saxonne (1), l'expression "probabilité subjective" se rapporte à la probabilité subjective de succès.

Afin d'alléger la notation et de lui imposer une plus grande rigueur, il serait commode de remplacer l'expression "probabilité de succès" par un terme unique. Divers termes ont été envisagés avec le professeur G. De Landsheere. Deux mots ont retenu notre attention : faillibilité et fiabilité. La traduction anglaise de faillibilité est *fallibility*; celle de faillible est *fallible* (2). Faillibilité désigne, malheureusement, la probabilité d'échec et non de succès. Or, c'est cette dernière que les sujets doivent exprimer. Il faudrait dès lors employer l'antonyme infaillibilité.

Robert définit l'infaillibilité comme suit : "1. Vx. caractère de ce qui ne peut manquer de se produire. V. certitude, ex. : l'infaillibilité d'un succès. 2. Mod. caractère de ce qui ne peut manquer de réussir. Ant. faillibilité, fragilité."

---

(1) A de rares exceptions près, comme ATKINSON (1964).

(2) Ces deux termes anglais viennent de to fall, Harraps indique :
fall away = faire défection
fall down (of plans) = échouer
fall through (of scheme) = ne pas aboutir, échouer, avorter.

Ce terme présente deux inconvénients. Tout d'abord, l'adjectif dont il est issu, infaillible, ne semble pas accepter de nuances. Robert indique : "Qui ne peut se tromper; qui a des conséquences certaines, des résultats assurés. V. parfait, souverain. (Cf. Qui réussit à coup sûr)."

Le second inconvénient, le plus important à nos yeux, réside dans la morphologie même d'infaillibilité. La racine (faillir), concept négatif, est corrigée par un préfixe négatif lui aussi, ce qui, au total, revient à un concept positif. Il y a là une lourdeur regrettable.

Le professeur L. D'Hainaut nous a suggéré "fiabilité". Ce terme plus récent ne figure pas au Robert. Larousse (1966) en donne la définition suivante : "Probabilité pour qu'un appareillage complet et, plus généralement, un équipement, fonctionne de manière adéquate pendant une période de temps déterminée dans les conditions opérationnelles spécifiées".

ASRATIAN et SIMONOV, auteurs de l'ouvrage La fiabilité du cerveau, définissent la fiabilité comme suit : "la probabilité de conserver les indices de qualité au cours d'un temps donné...".

Le verbe "se fier" a pour origine le terme latin populaire *fidare* (confier) et l'adjectif *fidus* (fidèle).

Dans son emploi moderne, fiabilité a pour correspondant anglais *reliability*; hélas, ce dernier terme a une signification très précise en théorie des tests et sa traduction française est fidélité. Les trois mots anglais les plus proches (1), *trustiness* (fidélité, loyauté), *truthfulness* (véracité, véridicité, fidélité) et *trustworthiness* (loyauté, fidélité, crédibilité, exactitude (d'un témoignage) rendent un sens légèrement différent et peuvent donc difficilement être retenus.

(1) MANSION, J.E., Harrap's Shorter: French and English Dictionary, Bordas, Paris, 1967.

Le terme "fiabilité" pourrait désigner la probabilité d'exactitude d'une réponse et peut-être s'imposera-t-il. Malheureusement, nous n'avons pas trouvé de correspondant pour la probabilité d'émission (ou d'utilisation) d'une réponse.

Faute de solution plus élégante, nous préférons employer des expressions longues mais claires et, au besoin, recourir à l'usage des sigles.

* c'est le problème

L'INTERET DU PROBLEME POUR L'ENSEIGNEMENT

L'enseignement fait acquérir un grand nombre de connaissances, parfois éphémères, comme l'orthographe française, les lois physiques, etc. (1). Quand des questions relatives à ces connaissances sont posées en classe, une série de conventions, tacitement passées entre le professeur et les élèves, interviennent. Ainsi, une question de chimie sur le tableau périodique des éléments, porte, sauf avis contraire, sur le tableau généralement utilisé au XXe siècle. En l'absence d'autre précision, le professeur qui pose la question 2 + 2 = considère implicitement que l'on travaille en base 10. On pourrait multiplier de tels exemples où une seule classe de réponses est considérée comme correcte.

Dans la transmission des connaissances, le pédagogue doit viser la certitude, mais respecter le doute. En effet, on peut mettre les connaissances des contemporains à la disposition des élèves sans les leur enseigner directement. Il est plus fécond d'entraîner les étudiants à retrouver par eux-mêmes les sources d'information, et à s'y référer.

Quoi de plus navrant que ces élèves qui, entourés de moyens d'information efficaces, ne savent pas résoudre un problème parce qu'il n'a pas été "vu" en classe ? S'orienter vers des ouvrages de référence ou des personnes susceptibles d'apporter des informations utiles, voilà des comportements à enseigner en classe ! Qu'en l'absence de tout support l'élève doute, quoi de plus normal ? La conscience de ce doute est le point de départ de comportements destinés à diminuer l'incertitude ou, comme disait A. Kaufmann, "faire chuter l'entropie" (2).

(1) L'enseignement ne transmet pas, loin s'en faut, que des connaissances. La présente étude est, par conséquent, limitée à un seul des aspects du processus pédagogique.

(2) L'entropie est, en thermodynamique, une mesure de désordre qui prend sa valeur maximale dans le cas du "désordre" total où toutes les situations sont équiprobables. Toute structuration qui rend plus probables certaines situations que d'autres, fait "chuter l'entropie" du système.

x Il ne s'agit pas de punir mais de mesurer !

Quel est le but de la cotation ?

Que veut-on évaluer ?

Pourquoi mêler - la connaissance

et - la conscience de l'ignorance

Le doute n'est donc pas condamnable en soi, il est même assez souvent le signe de la réflexion :

> "L'ennui, dans ce monde, c'est que les idiots sont sûrs d'eux et les gens sensés pleins de doute." (B. RUSSEL)

Dans une perspective formative, le doute et, à la limite, l'omission, doivent être débarrassés de la valeur négative qui leur est trop souvent attribuée dans l'enseignement. Pourquoi punir l'omission aussi sévèrement qu'une erreur ? De même, ce serait pénaliser la connaissance exacte mais peu assurée que de l'assimiler à une omission.

Or l'incertitude peut - et doit - être le point de départ d'une série de comportements positifs : recherche d'indices, vérification, en général "prise d'informations". Quand on doute de l'orthographe d'un mot, la réaction adaptée consiste à ne pas l'écrire directement, mais à vérifier au dictionnaire. Pourquoi un tel comportement devrait-il être humiliant ? Savoir que l'on ignore est précieux. Un individu qui, sûr de ses réponses, se tromperait une fois sur deux constituerait (aussi bien dans un hôpital que sur les routes, dans l'industrie ou dans l'enseignement) une source constante de désagréments et de dangers pour lui-même et pour les autres. Estimer correctement sa certitude et agir en conséquence sont des comportements qui devraient être appris à l'école. Cette préoccupation revient à dépénaliser l'incertitude et à enseigner les actions adaptées à cette situation.

Le raisonnement ci-dessus amène à proposer une technique de cotation qui accorde à la réponse exacte et estimée certaine par l'élève le maximum des points. Une réponse exacte, mais dont l'élève doute, doit être récompensée, mais plus légèrement. Par contre, une réponse inexacte et cependant considérée par l'élève comme certaine doit être fortement pénalisée pour décourager cette confiance abusive.

Encore faut-il que les élèves soient capables d'évaluer leur degré de certitude. Aucun apprentissage systématique n'est prévu à ce sujet, vraisemblablement à cause du mépris qui, dans les milieux scolaires, s'attache encore aujourd'hui au doute.

?

et le cas d'1 E qui omettrait systématiquement ?

Or les indices de certitude peuvent augmenter la fécondité de l'évaluation. Outre qu'ils contribuent au calcul des scores (voir les *corrections for guessing* au chapitre 3), ils fournissent des renseignements spécifiques sur la confiance qu'un étudiant éprouve vis-à-vis de ses propres connaissances et sur le bien-fondé de cette confiance. Il devient alors possible de repérer les étudiants toujours très sûrs d'eux-mêmes, mais dont les réponses sont souvent erronées et, à l'inverse, les élèves qui réussissent très bien, mais doutent constamment.

Voici, par exemple, les résultats obtenus à un test de trente questions, passé le 21 octobre 1971 par seize élèves, à l'Ecole Technique de la Force Aérienne Belge.

	NOM	MR	OM	BR	Mauvaises réponses —3	—2	—1	OM	Bonnes réponses +1	+2	+3	Cote totale
1	TOM	9	0	21	8	1	0	0	0	2	19	11.63
2	BUL	4	6	20	4	0	0	6	0	0	20	13.47
3	VOS	9	2	19	9	0	0	2	2	1	16	7.96
4	FLO	4	0	26	4	0	0	0	0	2	24	21.43
5	DAN	6	4	20	5	0	1	4	1	2	17	10.61
6	STI	4	0	26	2	1	1	0	0	0	26	20.00
7	WAE	4	2	24	4	0	0	2	0	0	24	16.53
8	KAM	7	1	22	7	0	0	1	0	0	22	12.86
9	BAR	11	0	19	10	1	0	0	0	0	19	9.39
10	LEC	6	0	24	4	2	0	0	0	3	21	17.14
11	PLA	3	0	27	3	0	0	0	0	2	25	23.27
12	VER	6	0	24	2	2	2	0	1	9	14	15.10
13	LAZ	7	1	22	5	1	1	1	3	0	19	10.41
14	FNN	11	0	19	3	5	3	0	1	1	17	10.61
15	KER	4	0	26	3	0	1	0	3	3	20	20.20
16	VER	4	3	23	2	2	0	3	4	8	11	12.24
TOT.	480	99	19	362	75	15	9	19	15	33	314	M = 14.55 0 = 4.7

. Les élèves 3 (VOS) et 9 (BAR) répondent presque toujours avec la certitude 3, ... mais se trompent souvent et ont, par conséquent, les cotes les plus basses de la série. La procédure de cotation défavorise donc ceux qui sont toujours sûrs d'eux-mêmes alors qu'ils se trompent.

Dans un atelier, un hôpital, une administration, les élèves 3 (VOS) et 9 (BAR) seraient de véritables catastrophes. La procédure de cotation les pénalise (voir leur cote).

. Les élèves 2 (BUL) et 16 (VER) ont omis plus de réponses que les élèves 1 (TOM) et 14 (ENN), et ont pourtant des totaux supérieurs. La procédure de cotation ne défavorise pas les élèves qui omettent (voir leur cote).

Les programmes scolaires font de plus en plus souvent figurer l'autoévaluation parmi leurs objectifs. On aurait tort d'en sous-estimer l'importance. Habituer l'élève à porter un jugement rétrospectif sur ses propres performances ne devrait-il pas être un des buts de toute évaluation scolaire ? L'estimation par le sujet lui-même de ses capacités n'est-elle pas de nature à rendre les projets des enfants plus compatibles avec la réalité ? Il ne s'agit pas seulement d'éviter ou d'atténuer des frustrations, mais aussi d'évaluer plus exactement la probabilité de succès d'entreprises (individuelles ou collectives) pour les appuyer efficacement. De plus en plus, l'apprentissage est géré par l'étudiant lui-même (méthodes actives, projets, enseignement programmé, enseignement assisté par ordinateur). Une bonne estimation de ses potentialités sera, pour chaque étudiant, un secours appréciable. Or, à l'école, tout est occasion d'auto-évaluation.

La pratique récente qui consiste à communiquer aux élèves les objectifs précis d'une intervention pédagogique, est particulièrement favorable à l'auto-évaluation. Amené, dès le départ, à prévoir ses performances finales, l'individu confrontera tout naturellement ses estimations et son comportement réel. S'auto-estimer mieux pour mieux s'élever, tel pourrait être le but de l'apprentissage qui consiste à évaluer de plus en plus correctement les probabilités de succès des actions.

Ainsi, un domaine nouveau s'ouvre à la réflexion pédagogique.

En fait, il n'y a pas d'information du tout sur le sujet mais une
information sur d'autres sujets et comme ce que l'on mesure, c'est
l'estimation que fait le sujet de la force de liaison entre chacun
de ces phénomènes communs et le sujet de la question

Si je dis "la capitale du Canada est Quebec",
je ne fais pas une plus petite ou plus grande erreur
que de dire Ouagadougou.

L'INTERET DU PROBLEME POUR LA RECHERCHE

Outre cet aspect formatif, la connaissance, par un observateur, de l'ESPER d'un sujet permet de mieux décrire et expliquer les comportements. L'exemple qui suit illustre cet apport.

Imaginons une situation d'évaluation où quatre expérimentateurs (enseignants ou chercheurs) interrogent un même étudiant sur un même contenu, mais par quatre techniques différentes.

La question posée est la suivante : "<u>Quel était le statut politique de la Malaisie en 1939 ?</u>". L'étudiant envisage plusieurs possibilités de réponse qui lui paraissent dignes d'être retenues :

- <u>colonie française</u> (vu la présence française au Cambodge et au Laos);
- <u>colonie anglaise</u> (vu la présence anglaise en Inde et en Birmanie);
- <u>colonie japonaise</u> (vu les incursions nombreuses des japonais sur le continent asiatique);
- <u>colonie hollandaise</u> (vu la présence hollandaise en Indonésie);
- <u>colonie chinoise</u> (vu la proximité de la Chine);
- <u>état souverain</u> (comme la Thaïlande).

Entre parenthèses sont notées les raisons qui, aux yeux du sujet, justifient chaque réponse envisagée. Sans connaître la réponse exacte, le sujet est donc capable de proposer les solutions les plus probablement correctes. Incontestablement, le sujet "sait" un certain nombre de choses relatives à la question posée, mais il ne sait pas tout. DRESSEL et SCHMID (1953) et COOMBS (1965) parlent, pour décrire cette situation, de <u>connaissance partielle</u> (*partial information*).

Mais le sujet est capable de faire mieux que de dresser la liste des réponses probables : il peut les ordonner selon leur degré de plausibilité, tel qu'il le perçoit. Si on le lui demande, il peut même attribuer à chacune des réponses un "pourcentage de chances d'être correcte".

Admettons que notre sujet estime les probabilités d'exactitude (PER) de la façon suivante :

	ESPER
- Colonie anglaise	25 %
- Colonie hollandaise	20 %
- Colonie japonaise	15 %
- Colonie française	15 %
- Etat souverain	10 %
- Colonie chinoise	5 %
	90 %
- Autre réponse	10 %
	100 %

Selon les exigences des quatre expérimentateurs, le sujet va maintenant fournir des réponses différentes.

. L'expérimentateur A demande une réponse sans aucun commentaire. Le sujet propose "colonie anglaise", c'est-à-dire la solution dont la PER lui paraît la plus élevée. La démarche du sujet est raisonnable : il peut ainsi espérer le meilleur score (1).

. L'expérimentateur B souhaite lui aussi une seule réponse, mais demande que l'élève indique, de plus, dans quelle mesure il la croit correcte. Le sujet répond : "Colonie anglaise (avec 25 % de certitude)". Bref, l'expérimentateur B permet au sujet d'indiquer son ESPER.

(1) On peut facilement vérifier que les sujets agissent bien ainsi. Il suffit de présenter des questions à choix multiple. Dans un premier temps, on demande de dire, pour chaque solution proposée, quelle est la probabilité de succès, le total devant être 100 %. Dans un second temps, on demande au sujet de proposer une seule réponse. Invariablement, le sujet désigne celle dont l'ESPER est la plus élevée. Ce phénomène (voir chapitre 6) est décrit par B. CHOPPIN (1971) et s'explique bien par la théorie des décisions (voir chapitre 1). Il s'agit de l'application d'une stratégie pure qui sera définie ci-après.

Ce qui est très dangereux, dans cette manière de procéder, c'est de faire associer à l'élève une probabilité, un degré de plausibilité à un fait qui est certain.

Si on avait demandé (avant de la connaître) "La face cachée de la Lune est-elle pareille à sa face visible ?" il serait normal de pénaliser la réponse certaine.

On substitue la conviction à la connaissance.

Conviction

Affectif — Cognitif

. L'expérimentateur C demande qu'on lui soumette les différentes réponses envisagées dans l'ordre d'ESPER décroissantes. Par exemple : colonie anglaise, colonie hollandaise, colonie japonaise, colonie française, état souverain, colonie chinoise.

. L'expérimentateur D enregistre les différentes réponses jugées plausibles et les ESPER qui y sont associées. Le sujet fournira donc le tableau complet ci-dessus.

Les valeurs numériques données par le sujet dépendent, entre autres, de la consigne imposée par l'expérimentateur D (1). On voit combien les mesures qu'il effectue sont plus précises que celles de l'expérimentateur C (2).

Nous venons de voir comment les quatre expérimentateurs ont mesuré la connaissance du sujet "avant" information (au prétest). Dans les travaux concernant la révision des probabilités après information, on parle de "connaissance a priori"qui s'oppose évidemment à l'état cognitif du sujet au post-test, appelé "connaissance a posteriori".

Imaginons que le sujet dont il vient d'être question reçoive des informations au moyen d'un film sur la Malaisie. Dans ce dernier, il n'est pas dit ou écrit que la Malaisie était une colonie anglaise (depuis 1824) et qu'elle a obtenu l'indépendance en 1957. Admettons que notre sujet retire du film les deux indices suivants :

a) L'invasion de la Malaisie, en 1941, par les troupes japonaises.
b) L'inscription "HAIR DRESSOR" sur la devanture d'une boutique.

---

(1) L'expérimentateur aurait pu demander au sujet d'exprimer son évaluation sous forme de rapports : une chance sur deux, une chance sur dix, etc., par pourcentages arrondis aux dizaines ou au moyen d'une échelle ordinale (quasi certain, très probable, probable, peu probable, quasi impossible). Dans le littérature, l'expression confidence weighting désigne l'ensemble de ces procédures. On trouvera un inventaire des consignes possibles dans DE FINETTI (1965) et LECLERCQ (1973).

(2) Si l'expérimentateur C attribuait une probabilité moyenne de 1/6 à chaque solution proposée, il commettrait deux erreurs. Tout d'abord, il n'obtiendrait que des estimations trop grossières des probabilités réelles. Ensuite, comme il ignorerait que la somme des probabilités fait .90 et non 1.00, il calculerait 1/6 (soit 16,6 %) au lieu de .9/6 (soit 15 %).

Encore une fois, ce qu'on mesure n'est pas la connaissance,
mais l'aptitude – à déduire une information d'indices suffisants
– à estimer la plausibilité de la déduction.

On peut dire aussi qu'en fait, on mesure l'attirance individuelle des réponses et non la connaissance.
Cette attirance peut être bien superficielle !

Ces renseignements vont amener le sujet à réviser les ESPER d'au moins deux solutions. Il est désormais pratiquement exclu (1) que la Malaisie ait été colonie japonaise en 1939 puisque ce sont les troupes japonaises qui l'envahissent en 1941. La solution colonie anglaise se voit nettement plus accréditée par l'inscription HAIR DRESSOR. Cet indice ne permet cependant pas une certitude totale, mais fait passer l'ESPER de 25 % à 70 %. En conséquence, les ESPER des autres solutions baissent nettement.

Voici l'état *a posteriori* des ESPER du sujet:

- Colonie anglaise	70 %
- Colonie hollandaise	10 %
- Colonie française	7 %
- Etat souverain	5 %
- Colonie chinoise	3 %
- Colonie japonaise	1 %
	96 %
- Autre solution	4 %
	100 %

On remarque que le total des ESPER des diverses solutions est 96 % contre 90 % *a priori*. Après cette prise d'information par le sujet, nos quatre expérimentateurs vont remesurer sa connaissance.

. A l'expérimentateur A, le sujet répond de nouveau "colonie anglaise". L'expérimentateur peut conclure (abusivement) que la valeur informative du film (concernant cette question précise) est nulle, car il n'observe aucune modification du comportement.

. A l'expérimentateur B, le sujet répond "colonie anglaise, à 70 % de chances". L'expérimentateur conclut que le film a provoqué une modification de la certitude de l'individu. On remarque que l'expérimentateur B ignore totalement un des deux effets du film (recul de la solution "colonie japonaise").

(1) Le sujet n'exclut pas totalement la solution "colonie japonaise" en considérant la possibilité que les Japonais aient été chassés de leur colonie entre 1939 et 1941.

. A l'expérimentateur C, le sujet énumérera cinq solutions possibles dans le même ordre qu'auparavant, à l'exception de "colonie japonaise". L'expérimentateur C peut donc observer cet effet du film, contrairement à l'observateur B. L'effet de l'information sur la solution "colonie anglaise" échappe cependant au premier et tous deux ignorent que la somme des ESPER vaut maintenant .96.

. A l'expérimentateur D, le sujet fournit l'ensemble du tableau *a posteriori*. La comparaison de celui-ci avec le tableau *a priori* permet de mesurer bien mieux les effets cognitifs (relatifs à cette question) provoqués par le film.

L'ensemble des données de l'expérience figure dans le tableau ci-dessous.

	AVANT information	Information	APRES information	
A	Colonie anglaise	Vision du film sur la Malaisie Deux indices : - Inscription HAIR DRESSOR. - Invasion de la Malaisie par les Japonais en 1941.	Colonie anglaise	A
B	Colonie anglaise (cert.: 25 %)		Colonie anglaise (cert.: 70 %)	B
C	1. Colonie anglaise 2. Col. hollandaise 3. Col. japonaise 4. Col. française 5. Etat souverain 6. Colonie chinoise		1. Colonie anglaise 2. Col. hollandaise 3. Col. française 4. Etat souverain 5. Col. chinoise 6. Col. japonaise	C
D	Col. anglaise (25 %) Col. holland. (20 %) Col. japon. (15 %) Col. franç. (15 %) Etat souverain(10 %) Col. chinoise (5 %)		Col. anglaise (70 %) Col. holland. (10 %) Col. franç. (7 %) Etat souverain (5 %) Col. chinoise (3 %) Col. japonaise (1 %)	D

Réponses recueillies par quatre observateurs différents (A, B, C et D) qui posent la même question ("Quel était le statut politique de la Malaisie en 1939 ?") à un même sujet, mais en lui permettant quatre modes de réponses différents.

Dans l'exemple ci-dessus, les expérimentateurs A, B et C sous-estiment l'effet d'un medium sur les connaissances du sujet. L'observateur A va même jusqu'à ignorer un phénomène qui existe. Il n'est pas difficile d'imaginer, pour la même question, une surestimation des phénomènes, due simplement aux méthodes des expérimentateurs. Il suffirait, par exemple, qu'un autre sujet ait formulé, avant information, les ESPER suivantes :

Colonie hollandaise ...... 45 %
Colonie anglaise .... 40 %

A l'expérimentateur A, le sujet répondrait "colonie hollandaise". On imagine facilement ses réponses aux autres expérimentateurs. Après vision du film, ses ESPER seraient :

Colonie anglaise .... 50 %
Colonie hollandaise ...... 45 %

Pour l'expérimentateur A, qui conclurait à un changement radical de l'état cognitif du sujet, sa réponse *a posteriori* serait donc "colonie anglaise". Or cette modification est loin d'être aussi tranchée. L'expérimentateur C pourrait, lui aussi, être abusé par ses "observations" car il sera fort tenté de conclure à une diminution importante de la crédibilité de la solution "colonie hollandaise". Or il n'en est rien, et seul l'expérimentateur D pourra observer l'état réel de la connaissance.

On voit combien, selon les méthodes de notation, les conclusions des expérimentateurs diffèrent et même s'opposent. Dans l'exemple ci-dessus, recueillir (comme l'expérimentateur A) la seule réponse, c'est simplifier abusivement la réalité psychologique et aboutir à des conclusions grossières, voire fausses. Or c'est la procédure la plus répandue dans la pratique scolaire, ainsi que dans les laboratoires pédagogiques. Les approches statistiques qui multiplient les sujets ou les questions ne contrebalancent pas valablement le manque de précision du phénomène observé.

Il ne faudrait pas favoriser la connaissance confuse vis-à-vis de la connaissance claire.

Ainsi, pour améliorer les caractéristiques statistiques des mesures, on est classiquement amené à multiplier le nombre de questions, et des formules célèbres expriment les gains de validité et de fidélité (J.P. GUILFORD, 1965, pp. 458, 465, 485) entraînés par un accroissement du nombre d'items (toutes choses restant par ailleurs égales).

Or, dans les situations les plus critiques, testant des hypothèses élaborées, on ne dispose souvent que de très peu de questions (une ou deux) et on ne peut les multiplier à volonté. Ainsi, dans l'expérience de PIAGET sur la conservation de la quantité de matière où un liquide a été transvasé dans un récipient d'une autre forme, on pose <u>la</u> question "Où y a-t-il le plus à boire ?" Dire : "Laquelle préfères-tu boire ?" n'est qu'une autre façon (peut-être plus dangereuse) de poser la <u>même</u> question. Il est difficile de poser une autre question. <u>Par contre, l'expérimentateur creusera la réponse</u>. Il ne se contentera pas du seul "ici". Il testera, par exemple, la résistance à la contradiction : "Pourtant, hier, un autre enfant m'a dit le contraire". Il demandera des justifications. Si l'enfant répond que c'est "parce que l'eau arrive plus haut", il fera remarquer que "le tube est plus étroit", etc. Cette évaluation ne procède pas par une simple multiplication des questions, mais un approfondissement de la réponse initiale : est-elle stable, justifiée pour le sujet, etc. ? La méthode piagétienne exploite ingénieusement les possibilités de l'approche clinique.

A l'opposé, l'évaluation pédagogique courante se prive souvent de possibilités d'approfondissement de la réponse. On met, par exemple, beaucoup de soin à contrôler de plus en plus finement des variables indépendantes telles que la lisibilité des textes, les composantes audiovisuelles des émissions de télévision scolaire, les interactions maître-élève, la progression minutieuse d'un cours programmé, etc., tandis que chaque item de la variable dépendante (le rendement) reste dichotomique : l'élève "sait" ou il "ne sait pas".

Il est regrettable qu'à la minutie déployée dans les plans d'analyse, les formules de lisibilité ou les grilles de codage ne corresponde pas une analyse fine du rendement.

Dans certains cas, la valeur explicative des recherches serait faible, non par manque d'outils explicatifs adéquats, mais par manque d'épreuves de rendement nuancées. A quoi sert-il, en effet, de construire trois séries parallèles d'émissions télévisées, si les différences - pourtant réelles - ne peuvent être mesurées que par des épreuves de rendement trop frustres ?

Quand les réponses brutes sont incapables de révéler les phénomènes sous-jacents, il faut modifier la procédure expérimentale ou mesurer un aspect encore ignoré de la réponse. La recherche de TERRACE (1963) sur l'acquisition de comportements complexes par des pigeons illustre bien cette démarche.

La procédure d'apprentissage sans erreur ne se révélait <u>apparemment</u> pas différente (quant au rendement) des procédures classiques d'apprentissage. Le recours à des épreuves d'extinction a seul permis de démontrer la nette supériorité de la nouvelle procédure.

Pour appréhender des aspects ignorés de la réponse, on peut mesurer la rapidité, la puissance ou les répétitions. Cependant, il faut chaque fois tenir compte de l'interférence des variables individuelles et circonstancielles comme le temps de réaction, la puissance musculaire, le degré de vigilance... Dans certains contextes (en psychophysique par exemple), il est possible de multiplier les mesures, soit parce qu'on se trouve en laboratoire, soit parce que les stimuli se prêtent à la replication. Ces deux conditions sont souvent absentes en pédagogie. Il faut, dès lors, songer à une autre approche pour approfondir et nuancer les réponses elles-mêmes.

L'ESPER paraît être une solution générale s'appliquant aux cas les plus divers. Néanmoins, elle reste un échelon intermédiaire dans l'approfondissement des connaissances. Ainsi, les données obtenues par l'expérimentateur D dans l'exemple de la Malaisie peuvent être à leur tour approfondies. On demandera, par exemple, ce qui justifie une réponse et la certitude qui l'accompagne (le contenu des parenthèses dans l'exemple). Mais la justification à son tour peut être nuancée par une ESPER... et ce, de façon quasi infinie. Ce caractère intermédiaire de l'ESPER est lié à sa généralité.

Il semble possible, néanmoins de recourir à cette mesure pour prédire le comportement d'un sujet donné à un essai donné, quand certaines conditions expérimentales sont remplies. Le chapitre 7 est, en partie, consacré à des essais de prédiction de ce genre.

LES FONDEMENTS EPISTEMOLOGIQUES DE LA METHODE PROPOSEE

Le caractère probabiliste de l'approche n'est pas fortuit, car la tendance actuelle des recherches concernant le comportement est de moins en moins déterministe :

> "Les modèles considérés dans notre ouvrage (SUPPES et ATKINSON, 1960) sont entièrement probabilistes. C'est là une propriété qu'ils partagent avec la plupart des modèles mathématiques actuellement à l'étude dans les sciences du comportement. Dans la mesure où notre compréhension de la neurophysiologie croît, il se peut que l'on construira de plus en plus de modèles déterministes, mais les modèles probabilistes semblent devoir jouer un rôle majeur pendant les années à venir, et pour un certain temps encore." (p. 284)

Néanmoins, l'approche probabiliste entraîne des difficultés épistémologiques, comme le souligne B. DE FINETTI (1970) :

> "Au départ, les interprétations probabilistes de la mécanique classique furent considérées par beaucoup comme de petits subsituts aux "vraies" (mais non connues avec précision) lois déterministes de la physique. L'introduction de la physique quantique a d'abord été perçue commme une menace pour le déterminisme et, en tant que telle, pour tout l'édifice de la science. Aujourd'hui, ces approches sont reconnues non comme menaces pour la science, mais comme nécessités pour un développement significatif des idées scientifiques."

En mécanique quantique, on ne parle plus de la présence d'un objet à un endroit donné à un moment donné, mais de la probabilité de cette présence. Un chercheur peut être frustré par ce modèle "non déterministe" qui, cependant, présente un aspect positif : il permet d'expliquer plus de phénomènes et de guider l'action de manière plus féconde.

Mais l'approche probabiliste comporte des difficultés intrinsèques :

> "La probabilité est un concept difficile. A bien des fins, nous pouvons nous contenter du débit de réponse, mais celui-ci n'a guère de sens lorsque nous attribuons à plus d'une variable un cas isolé de comportements. Des problèmes analogues surgissent avec beaucoup d'autres, quand on infère <u>la probabilité de la présence ou de l'absence d'une réponse dans un "essai donné"</u> (SKINNER, 1971, p. 129).

Ces difficultés sont rencontrées également par les mathématiciens qui élaborent des modèles afin de simuler l'activité intellectuelle :

> "Dans une telle théorie (stochastique), on ne fait pas de prédiction sur une réponse précise d'un sujet précis à un essai précis." (NEWELL et SIMON, 1963, pp. 368-369).

Les problèmes rencontrés sont identiques parce qu'une même option est à la base des deux courants de recherche : la fréquence de réponses est l'observation de base :

> "(Seule est prédite la distribution des réponses) calculée à partir d'un certain nombre d'essais ou un certain nombre de sujets." (NEWELL et SIMON)
>
> "L'analyse expérimentale traite de cette probabilité (qu'un comportement survienne) en termes de fréquence ou de débit de réponses." (SKINNER, 1971, p. 109)

Il nous paraît possible de dépasser les difficultés décrites ci-dessus, en recourant au concept d'ESPER. A partir de ce nouvel outil, il devient possible de prédire la probabilité d'apparition de ce comportement (PUR), pour un sujet donné, à un moment donné.

Quoique cette notion soit implicite dans divers écrits, son importance, cruciale à nos yeux, n'a jamais été soulignée. Il est symptomatique de constater que, jusqu'à présent, les auteurs n'ont pas réssenti le besoin d'utiliser un terme spécifique pour désigner la probabilité de succès. Probabilité est indifféremment employé dans les deux sens : probabilité d'émission et probabilité de succès. Une telle ambiguïté n'est plus admissible parce que, en paraphrasant SKINNER, on peut dire que la probabilité d'exactitude de la réponse d'un individu donné, à un moment donné, constitue, dans une science du comportement, et tout spécialement en éducation, un fait de base.

Or la probabilité d'exactitude peut expliquer, dans une large mesure, la probabilité d'apparition, comme montre l'exemple médical (imaginaire) qui suit.

Un médecin se trouve face à une épidémie qui a touché 1000 personnes. Il dispose de trois médicaments (X, Y et Z).

Statistiquement, X assure la guérison dans 70 % des cas;
Y assure la guérison dans 40 % des cas;
Z assure la guérison dans 20 % des cas.

Le médecin possède X, Y et Z en quantités suffisantes, mais les médicaments s'excluent : un même sujet ne peut en absorber deux différents. Que fera le médecin ? Il donnera le médicament X à tous les malades, car l'espérance mathématique des guérisons est : 0,70 x 1000 = 700 personnes.

Admettons qu'il ait adopté une autre stratégie : donner 50 % des médicaments X et 25 % de chacun des deux autres. L'espérance mathématique des guérisons aurait alors été :

(0,70 x 500) = 350 personnes
(0,40 x 250) = 100 personnes
(0,20 x 250) = 50 personnes

soit, au total : 500 personnes

Ce nombre est largement inférieur aux 700 personnes sauvées par la première stratégie. Mathématiquement, il est facile de démontrer que, s'il a comme objectif de maximiser le nombre de guérisons, le médecin doit toujours répondre de la même façon, c'est-à-dire administrer le remède X à tous les sujets. La probabilité d'utilisation du comportement est maximale (PUR = 1). Néanmoins, la probabilité de succès n'est que de 70 % (PER = .7) et le médecin le sait.

La fréquence de distribution du médicament X est donc une mesure distincte du taux de succès de ce médicament, mais en dépend directement.

On dira du médecin qu'il a adopté une stratégie pure pour maximiser le nombre de guérisons. La stratégie pure (*pure strategy*) doit être opposée à une stratégie d'ajustement (*matching strategy*). Les expériences de S. SIEGEL et al. (1961) à ce propos sont éclairantes.

Dans ces expériences, on présente à des enfants de quatre ans deux bouteilles opaques dont une seule contient un petit objet toujours différent et attractif pour l'enfant (voiture en réduction, bille, etc.). A chaque essai, l'enfant désigne l'une des deux bouteilles placées devant lui et l'expérimentatrice la retourne; si un objet en sort, l'enfant peut le garder. En respectant une séquence au hasard,

l'expérimentateur a décidé de placer pour 75 % des cas l'objet dans la bouteille de gauche. Au début, l'enfant désigne à peu près autant de fois le flacon de droite que celui de gauche. Mais les résultats rendent compte de la différence de répartition. L'enfant pourrait dès lors adopter une stratégie de correspondance (ou d'ajustement) qui consiste à présenter un comportement parallèle aux probabilités observées (*matching strategy*). Dans le cas présent, il s'agirait de désigner la bouteille de gauche dans 75 % des cas. L'espérance mathématique de réussite de cette stratégie est .62 :
(.75 x .75) + (.25 x .25) = .56 + .06 = .62

Or les enfants de quatre ans adoptent spontanément (après plusieurs dizaines d'essais) une stratégie pure (*pure strategy*) qui consiste à désigner invariablement la bouteille de gauche. L'espérance mathématique de cette attitude vaut .75 : (.75 x 1) = .75. La stratégie pure offre donc une espérance mathématique plus élevée que la stratégie d'ajustement.

S. SIEGEL et J. McMICHAEL (1961) ont également soumis d'autres enfants de quatre ans à une situation expérimentale identique, mais la bouteille contenait invariablement le même objet (un petit bouton de gilet). Dans ce cas, les enfants ont adopté la stratégie de correspondance décrite plus haut, car le renforcement n'était plus constitué par la possession de l'objet stéréotypé, mais par le seul aspect ludique de la prédiction.

Les expériences faites par S. SIEGEL et son équipe sur adultes confirment leurs observations : quand il importe de maximiser les gains (les renforcements étant monétaires), les sujets adoptent une stratégie pure, malgré la monotonie motrice (désigner invariablement le même stimulus). Par contre, lorsqu'une récompense tangible n'est plus associée à la réponse, bref, quand les sujets ne sont pas exposés aux conséquences de leurs actes, il tendent à adopter une stratégie de correspondance, plus variée au point de vue moteur.

Quand l'objectif est de maximiser les conséquences, la stratégie pure est mathématiquement supérieure à la stratégie de correspondance. En économie, la théorie des décisions utilise largement ce principe. On constate que, soumis à une procédure operante, les humains adoptent aussi cette stratégie; ils le font graduellement et inconsciemment (on peut en effet raisonnablement penser que les enfants de quatre ans sont incapables de calculer l'espérance mathématique de leur stratégie).

La distinction entre les deux types de probabilité (PER et PUR) peut être féconde dans le domaine de la recherche pédagogique et dans la pratique scolaire. Ainsi, on rencontre fréquemment un emploi abusif de la PUR pour valider les distracteurs d'une question à choix multiple. On sait que, classiquement, on base une telle validation sur les fréquences d'utilisation de chaque solution, ces fréquences constituant des estimations de la PUR.

Admettons, par exemple, que tous les étudiants accordent les probabilités de succès suivantes aux quatre solutions proposées pour une question à choix multiple :

Question	ESPER.
1	10 %
2	40 %
3	20 %
4	30 %

Si chaque étudiant fait la même répartition que ci-dessus, tous les étudiants choisiront la proposition 2, parce que c'est la solution la plus probable. Classiquement (et à tort, selon nous), on considérera que les autres propositions sont de mauvais distracteurs, car elles ne sont pas choisies. C'est faux. Seule la façon de recueillir la réponse peut le laisser croire. Choisie ou non, la solution 3 a une attractivité de 20 %, ce qui entraîne en partie l'affaiblissement de la certitude accordée à la réponse correcte (ici 2, par exemple).

Pq ne pas parler d'attirance
en se plaçant d'un ∆ phénoménologique

Voici, extraits du chapitre 7, deux exemples réels de questions à choix multiple présentant trois solutions (la consigne précisait que l'une d'elle est correcte).

	Sujets :	1	2	3	4	5	6	% de choix	ESPER moyenne
Question 10	Solution 1	.15	.10	.20	.25	.15	.05	0 %	.166
	Solution 2	.70	.80	.75	.50	.50	.90	100 %	.691
	Solution 3	.15	.10	.05	.25	.25	.05	0 %	.142

	Sujets :	1	2	3	4	5	6	% de choix	ESPER moyenne
Question 6	Solution 1	.80	.45	.55	0	.50	.60	75 %	.483
	Solution 2	.10	.10	.15	.30	.10	.20	0 %	.158
	Solution 3	.10	.45	.30	.70	.40	.20	25 %	.358

Dans le cas précis des questions à choix multiple, nous proposons de recourir à la probabilité de succès pour la validation; il suffit que chaque sujet indique son ESPER pour chaque solution proposée (voir annexe 2).

Il peut sembler paradoxal de parler de "probabilité de succès d'une réponse" alors qu'en fait, elle ne peut être que correcte (totalement) ou incorrecte (totalement). Peut-on attribuer des valeurs intermédiaires face à ce phénomène dichotomique ?

En fait, la préoccupation ci-dessus est, au niveau du score à une question, parallèle à la recherche du score vrai (*true score*) au niveau du score total à un test. Ce dernier souci est rencontré par un grand nombre de chercheurs contemporains (1), comme R.D. BOCK ou C.K. FISCHER (approche classique), ou M.R. NOVICK, P.H. JACKSON (approche bayésienne), ou encore L.J. CRONBACH et J. CARDINET (théorie de la généralisabilité).

(1) Voir les actes du Second International Symposium on Educational Testing, Montreux, juillet 1975.

Comment estimer le SCORE VRAI d'un sujet à une question ? (1)

Si on représente par X_{is} le score observé d'un sujet s à un item i, X_{is} ne peut valoir que 0 ou 1, puisque la réponse sera correcte ou incorrecte (l'omission étant assimilée à une réponse incorrecte).

Classiquement, on considère que

$$\boxed{X_{is} = \mathcal{E}X_{is} + E_{is}} \qquad \text{(définition 1)}$$

où $\mathcal{E}X_{is}$ = le score vrai
et E_{is} = une erreur de mesure.

Selon les cas, l'erreur de mesure est favorable ou défavorable au sujet. Dans le premier cas, la réponse est exacte (1) alors que les connaissances ne sont pas parfaites. Dans le second cas, la réponse est inexacte (0) alors que les connaissances ne sont pas nulles.

La répétition de la situation permettraient d'annuler les effets du hasard. Imaginons, par exemple, que le score vrai d'un sujet soit 0,70. Dans 70 % des cas, il fournirait la réponse correcte et l'erreur de mesure vaudrait 0,30. Dans 30 % des cas, la réponse serait incorrecte et l'erreur de mesure serait -0,70.

Les effets du hasard s'annulent.

$$\mathcal{E}E_{is} = (0,30 \times .70) + (-0,70 \times .30) = 0$$

$$\boxed{\mathcal{E}E_{is} = 0} \qquad \text{(définition 2)}$$

(1) D. DEFAYS a apporté, dans l'élucidation et la formulation du problème, une aide déterminante.

un autre facteur intervient : la rigidité de la réponse.

Ce raisonnement est théorique, car on pose rarement (1) plusieurs fois la même question à un même sujet, dans les mêmes conditions. C'est la raison pour laquelle on s'intéressera à des items différents mais qui présentent des caractéristiques communes pour un sujet donné : les items <u>parallèles</u> et les items <u>équivalents</u>.

On pose généralement que deux items (indépendants) sont <u>parallèles pour un sujet donné</u> s'ils ont le même score vrai et le même écart-type de l'erreur de mesure. (définition 3)

> Si $\mathcal{E}\, X_{is} = \mathcal{E}\, X_{js}$
>
> et $\sigma_{(E\ is)} = \sigma_{(E\ js)}$
>
> les items i et j sont <u>parallèles pour s</u>.

?

Si l'on considère la réponse comme fixe et l'ESPER comme une variable aléatoire, on peut poser que deux items (indépendants) sont <u>équivalents pour un sujet donné</u> s'ils ont la même ESPER moyenne et la même variance d'ESPER. (définition 4)

> Si $\mathcal{E}\,(ESPER_{is}) = \mathcal{E}\,(ESPER_{js})$
>
> et $\sigma\,(ESPER_{is}) = \sigma\,(ESPER_{js})$
>
> les items i et j sont équivalents pour s

---

(1) Sauf peut-être dans le cas des <u>universe-referenced tests</u> dont la théorie est en élaboration (voir entre autres SHOEMAKER, 1973).

?

Si des questions équivalentes sont parallèles, alors

Postulat : Le vrai score est égal à l'ESPER moyenne

$$\mathcal{E} X_{is} = \mathcal{E} \text{ (ESPER } _{is})$$

Corollaire : En moyenne, la proportion des réponses correctes à N questions équivalentes d'ESPER π est π .

Ce corollaire peut être démontré :

$$\mathcal{E} \left(\frac{1}{N} \sum_{i=1}^{N} X_{is}\right) = \frac{1}{N} \sum_{i1}^{N} \mathcal{E} X_{is} = \pi$$

Ce corollaire peut être directement soumis à la vérification expérimentale.

LA VALIDITE, LA FIDELITE ET LA SENSIBILITE DES MESURES OBTENUES A PARTIR DE L'EXPRESSION DE LA CERTITUDE

Les mesures "avec certitude" sont-elles plus valides, plus fidèles et plus sensibles que les mesures "sans certitude" ? Le premier problème (celui de la validité) doit être envisagé d'abord. Diverses études expérimentales ont été menées pour y répondre, par exemple : VAN NAERSSEN et VAN BEAUMONT (1965), MELLENBERGH (1967), SANDBERGEN (1968, 1971 et 1972), BAKER (1969), RIPPEY (1970), HAMBLETON, ROBERTS et TRAUB (1970), KOEHLER (1971), JACOBS (1971). Nous pensons cependant que le problème de la validité ne peut être résolu par la seule expérimentation. Il doit faire l'objet de raisonnements psychométriques.

A. La validité concurrente.

La validité concurrente est, classiquement, appréciée par la corrélation entre deux tests. Les scores avec certitude (SC) sont fortement corrélés avec les scores simples (SS) (fonction linéaire du nombre de réponses correctes). La raison en est simple : l'importance de la certitude est conditionnée par l'exactitude de la réponse.

La corrélation entre les deux scores (SS et SC) dépend étroitement du poids relatif donné à la certitude dans la matrice des conséquences. Prenons un exemple. Soit quatre degrés de certitude : d'une part, 0 (l'omission) et, d'autre part, 1, 2 et 3 que l'on peut associer à chaque réponse.

Voici trois barêmes de cotation possibles :

	Barême A			Barême B			Barême C	
	MR	BR		MR	BR		MR	BR
OM	0	0	0	0	0	0	0	0
Cert. 1	-1,9	+1,9	1	- 1	+ 1	1	-0,1	+0,1
Cert. 2	-2	+2	2	- 2	+ 2	2	-2	+2
Cert. 3	-2,1	+2,1	3	- 3	+ 3	3	-3,9	+3,9

On voit que le poids relatif (1) de la certitude croît, du barême A au barême C. Anticipant sur le chapitre 3, où une méthode de calcul est proposée, ces poids relatifs de la certitude sont estimés à .25, .35 et.53 pour les barêmes A, B et C respectivement. Les corrélations (théoriques) entre les scores simples et les scores avec certitude vaudraient .865, .806 et .684.

Or, une corrélation n'est d'aucun secours pour déterminer lequel des deux scores est le plus valide.

Le score avec certitude (SC) ne peut prétendre à une validité intrinsèque puisque, on vient de le voir, il dépend du poids relatif de la certitude. Les adversaires du score avec certitude avanceront qu'une telle mesure est faussée par une variable que l'on maîtrise mal et que, par conséquent, on augmente le "bruit" plutôt que l'information. Les partisans du score avec certitude, dont nous sommes, répliquent par les deux arguments suivants. D'une part, la variable introduite (l'ESPER) peut être contrôlée avec une précision suffisante, au prix, il est vrai, d'un important effort méthodologique (voir chapitres 1 à 4). D'autre part, cette variable dûment maîtrisée constitue pour une large part, la solution au problème de l'erreur de mesure. Les considérations théoriques examinées plus haut (les six définitions et les deux postulats) fondent ce deuxième argument.

et les variables de personnalité

à voir plus loin

(1) Poids de la certitude comparé au poids de l'exactitude de la réponse. Les valeurs fournies sont des "pourcentages du poids total".

Mais tout le problème est que ces deux informations
sont mélangées et que l'on ne fait pas "chuter l'entropie",
qu'on la fait croître au contraire

Nous avons vu plus haut que les valeurs de la matrice de conséquences entraînent des corrélations différentes entre les scores simples et les scores avec certitude. Le choix de la matrice doit être guidé par les objectifs de l'enseignant (et pourquoi pas de l'enseigné ?) et le poids relatif qu'y prend l'auto-évaluation. Dès lors, la valeur numérique précise d'un score avec certitude (SC) ne peut prétendre à une validité intrinsèque. La richesse pédagogique réside dans les deux types d'informations ainsi recueillies : l'exactitude de la réponse et la certitude du sujet.

C'est là un argument très favorable à l'utilisation de l'ESPER, car, comme le dit THORNDIKE (1) : "L'utilité prédictive d'un test dépend non seulement de la mesure dans laquelle il est corrélé avec un critère, mais aussi de la quantité d'informations nouvelles qu'il apporte". C'est l'intérêt principal qu'y voit L.J. CRONBACH (2).

Le score total qui doit être communiqué au sujet ou utilisé dans une perspective de sélection est le score avec certitude (SC). Ce score, influencé par les ESPER, constituera un renforcement (3) à leur utilisation la plus pertinente possible.

B. La validité prédictive.

La validité prédictive est établie, comme la validité concurrente, par une corrélation, mais cette fois entre une mesure 'x' et une mesure 'y' portant sur un autre contenu. En éducation, ces mesures 'y' concernent assez souvent des données socialement utiles (rendement scolaire, compétence professionnelle, performance sportive, qualité du produit final, etc.), et on s'attache à les prédire efficacement.

---

(1) Voir Thorndike (1971) et surtout le chapitre 6 (pp. 162-209) de THORNDIKE et HAGEN (1969) : "Quelles qualités rechercher dans toute mesure ?".

(2) Communication personnelle, Montreux, juillet 1975.

(3) Cet aspect opérant sera étudié par la suite (chapitres 2 et 4).

Tout ceci n'est pas convaincant.
Revue insuffisante de la littérature.
Il faudrait voir si la validité prédictive est plus grande pour
prédire des valeurs – en rapport avec la personnalité
– en rapport avec le savoir

Revenons aux "simples" :
Une mesure est valide si elle mesure bien la grandeur
qu'elle prétend mesurer.
– Que mesure exactement le score avec certitude ?
– le mesure-t-il réellement ?

Si on répond "la connaissance",

Dans une telle perspective, la sensibilité peut être d'une importance cruciale pour les prédictions, comme le montre l'exemple amusant ci-après. Admettons qu'un sujet soit capable d'exprimer (fidèlement) diverses durées en intervalles de cinq minutes. Par exemple, 31' (au chronomètre) est exprimé par '7', car cette durée s'étend sur le septième intervalle de cinq minutes. Une telle "mesure" de la durée n'est pas moins valide que celle d'un chronomètre; elle est simplement moins sensible. Or cette sensibilité a un impact direct, <u>dans certaines situations</u>, sur la validité prédictive.

Ainsi, il est facile de démontrer, expérimentalement, que le chronomètre s'indique quand il s'agit de cuire des oeufs à la coque. L'estimateur est impuissant, alors que le chronomètre est efficace.

Cet argument est celui de la fécondité pratique, de la pertinence des mesures dans l'action. Mais, même si les nouvelles mesures (avec certitude) s'avéraient moins prédictives, il faudrait se demander si, combinées dans la prédiction avec d'autres variables (avec lesquelles elles seraient moins corrélées), elles ne permettraient pas des coefficients de corrélation multiple plus élevés.

C. <u>La validité de contenu</u>.

L'exemple de la Malaisie développé plus haut était essentiellement destiné à montrer la validité de contenu des mesures effectuées à partir d'un indice de certitude. De plus, à propos de la validité concurrente, nous avons avancé que l'ESPER est un outil particulièrement propre à réduire l'erreur de mesure. La raison en est évidente, puisque c'est ce que précise la consigne aux étudiants :

Estimez la probabilité qu'a votre réponse d'être correcte.

Vous diminuez le facteur d'erreur (la part de "deviné" qu'il y a dans la réponse) mais vous en introduisez d'autres (variable de personnalité).

$$X_{is} = K_{is} + f(R_{is})_s + f(S)_i + f(\text{certitude})_{is} + E$$

connaissance

goût du risque

fonction de la stratégie liée au test (ou à l'item si elle n'est pas la même pour chaque item)

certit. du sujet pour la question

La consigne joue un rôle capital dans la validité de contenu puisqu'elle résume la procédure de mesure qui fonde cette validité. A partir d'une échelle ordinale, par exemple, on pourrait moins prétendre réduire l'erreur de mesure :

Dites si vous êtes peu certain (certitude 1),
moyennement certain (certitude 2),
très certain (certitude 3),
de votre réponse.

Cette consigne ne pose pas le problème de la mesure de façon compatible avec les définitions et les postulats décrits ci-dessus.

D. La fidélité.

à voir

La fidélité des mesures avec indice de certitude est un problème qui ne relève que de l'expérimentation. Le nombre de degrés de certitude (la sensibilité de l'échelle), la consigne, la stabilité (test-retest) de la certitude sont vraisemblablement de nature à influencer la fidélité des nouvelles mesures. Le chapitre 6 traitera d'une partie de ces problèmes par des voies expérimentales.

E. La sensibilité.

à voir

La sensibilité des mesures avec certitude est plus grande que celle qu'on obtient par des procédures classiques. Mais la limite de la sensibilité doit, de nouveau, être déterminée par l'observation. Ce problème sera aussi envisagé au chapitre 6.

LES GRANDS TYPES DE NOTATION DE LA CERTITUDE

Il n'est pas étonnant que les recherches sur la connaissance partielle aient été menées à partir de l'utilisation des questions à choix multiple. Cette procédure, en effet, met en cause la fidélité de la réponse et permet de masquer une ignorance par le choix d'une solution au hasard (cf. L. CRONBACH, 1950; HOFFMAN, 1962; BLACK, 1963; GOSLIN, 1966).

Ce choix peut cacher l'ignorance... ou une connaissance peu assurée. Pourquoi, dès lors, ne pas considérer que toute connaissance se situe sur une échelle d'ESPER, quelle que soit la façon dont la question est posée (qu'elle soit ouverte ou à choix multiple) ?

La technique la plus simple qui traite ce problème est l'utilisation d'indices de certitude (en anglais *confidence marking*). EBEL (1965a, p. 49) en donne la définition suivante :

> "C'est un mode de réponse spécial aux questions d'un test objectif et un mode spécial de cotation de ces réponses. En bref, le sujet doit indiquer non seulement ce qu'il croit être la réponse correcte à une question, mais aussi quelle est sa certitude dans l'exactitude de sa réponse. Au moment de la cotation, le sujet reçoit plus de points pour une réponse correcte avec certitude que pour une réponse accompagnée d'un doute. Mais la pénalisation d'une réponse incorrecte avec certitude est suffisamment lourde pour décourager les déclarations de confiance non fondées."

Cette définition appelle plusieurs remarques terminologiques. Le terme test objectif fait référence aux tests à choix multiple. C'est une définition trop restrictive, car les indices de certitude s'appliquent aussi bien aux questions ouvertes.

L'expression degré de certitude est la plus courante, mais on trouve aussi "assurance", "confiance", "degré de conviction" (IRWIN, c'est assez - très différent

1953), car l'utilisation de la certitude a précédé les formulations théoriques. La *confidence weighting* est l'indication de la certitude de la réponse choisie et le *probabilistic testing* est l'indication de l'ESPER de chaque solution proposée d'une question à choix multiple. Ces procédures sont étudiées par STANLEY et WANG (1970) et discutées par JACOBS (1971).

A. Les pionniers.

Les premiers utilisateurs systématiques de tels indices sont HENMON (1911), HOLLINGWORTH (1913), TROW (1923) et HEVNER (1932). JACOBS (1968) trace l'historique du problème et ALLGREN (1967) établit une synthèse de plusieurs recherches qui les emploient. Nous n'en citerons que quelques-unes.

B. La procédure la plus élémentaire.

Elle consiste à mettre l'accent (en soulignant, en entourant, en accompagnant d'un astérisque, etc.) sur une réponse "dont on est certain". Ainsi, aux Pays-Bas, VAN NAERSSEN et VAN BEAUMONT (1965) présentent leurs questions à choix multiple comme suit :

Z1
Z2
Z3
Z4

NB.: Z = la première lettre de *zekerheid* (certitude).

Les auteurs proposent deux façons de choisir une solution :

Z 1 (sans certitude)
Z 1 (avec certitude)

C. Certitudes exprimées ordinalement.

Une autre procédure consiste à demander au sujet d'indiquer sa certitude sur une échelle ordinale.

Exemples :

1	2	3
(peu sûr) ou doute ou non sûr	(sûr)	(très sûr) certitude

Le nombre de degrés de cette échelle ne peut être très élevé en raison même de leur imprécision : les termes qui les décrivent font l'objet d'interprétations très variées.

D. L'ESPER.

Aux consignes ordinales, on préférera les consignes qui se réfèrent à des échelles d'intervalle et tout spécialement à l'échelle (de 0 à 100 %) de l'ESPER. En effet, l'échelle de l'ESPER correspond strictement à l'échelle des proportions de réussites, directement observables par tous.

Dans les consignes de l'ESPER, le sujet doit exprimer le pourcentage de chance d'exactitude qu'il attribue à sa réponse. Cette ESPER peut s'exprimer directement en pourcentages (ex.: 75 %), en rapports (ex.: 1 chance sur 1000), ou en degrés de certitude représentant (de façon codée), soit des zones de pourcentage (exemple 1 ci-dessous), soit des rapports (exemple 2).

Exemple 1.

Certitude	A	B	C	D	E
0	20 %	40 %	60 %	80 %	100 %

Exemple 2.

Certitude	A	B	C	D	E	F	G	H	I
	1/1000	1/100	1/10	1/4	1/2	3/4	9/10	99/100	999/1000

Le lecteur trouvera ailleurs (LECLERCQ, 1973) de nombreuses modalités pratiques du même type. B. DE FINETTI a attaché son nom à certaines d'entre elles et, par exemple, le *continuous confidence marking (C.C.M.)* ou notation de la probabilité avec autant de décimales que le souhaite le sujet et le *five stars system* ou attribution d'étoiles aux différentes solutions envisagées (le sujet n'a que cinq étoiles à attribuer par question !). B. DE FINETTI (1965) proclame :

> "Ce n'est que la probabilité subjective qui peut donner une signification objective à chaque réponse et à chaque méthode de cotation."

Nous partageons cette conviction de B. DE FINETTI qui pousse le raisonnement jusqu'au bout :

> "Nous ne nous intéressons pas à une incertitude possible qui serait inhérente à la question. Même s'il y avait des controverses parmi les personnes informées, la probabilité de n'importe quelle solution correspondrait à une estimation individuelle basée sur une pondération soigneuse de toutes les démonstrations et opinions..." (p. 88)

Dans la majorité des recherches, deux conséquences sont associées à chaque degré de certitude : une en cas de réponse correcte et une autre en cas de réponse incorrecte. L'ensemble des degrés de certitude nécessite donc une matrice de conséquences.

LES MATRICES DE CONSEQUENCES

Les points gagnés et perdus étant fonction des degrés de certitude éprouvée, ces derniers interviennent directement dans la cotation des sujets.

VAN NAERSSEN et VAN BEAUMONT (1965) utilisent la matrice de points suivante :

	Réponse correcte	Réponse incorrecte
Sans certitude	+ 1/2	0
Avec certitude (Z)	+ 1	- 1/2

Cette matrice est appelée *pay-off matrix* en sciences économiques. L'expression "matrice de conséquences", commune à l'économie et à la psychologie, a été retenue ici.

S. SANDBERGEN (1971) utilise la même matrice, à une constante près : tous les coefficients sont doublés.

S.S. JACOBS (1972) utilise une autre matrice :

	Réponse correcte	Réponse incorrecte
Je devine	+ 1	0
J'ai moyennement confiance en ma réponse	+ 2	- 2
J'ai très confiance en ma réponse	+ 3	- 3

Dans la première expérience relatée au chapitre 2, la matrice de points est la suivante :

	Réponse correcte	Réponse incorrecte	
Incertain (peu sûr)	+ 1	- 1	Certitude 1
Certitude normale	+ 2	- 2	Certitude 2
Très certain (très sûr)	+ 3	- 3	Certitude 3

Les points de toutes ces matrices sont attribués selon le double critère de la facilité mnémotechnique et de la facilité de calcul. Notre matrice est optimale à ces deux points de vue.

Malheureusement, ces matrices n'ont pas été construites selon un modèle mathématique directement inspiré de la théorie des décisions (voir chapitres 1 et 3). Or l'indice de certitude se définit (EBEL, 1965) tout autant par sa matrice des gains et des pertes que par les consignes données aux élèves. De même, le souci de décourager les déclarations mensongères est fondamental (SODERQVIST l'éprouvait déjà en 1936).

Tout le problème de la matrice de cotation se trouve dans l'expression d'EBEL : "Une pénalisation assez lourde pour décourager les déclarations de confiance non fondées." Quelle est la valeur optimale de la pénalisation ? Elle ne doit en tous cas pas être trop lourde, pour ne pas décourager l'expression de certitudes fortes. Au chapitre 3, on verra que VAN NAERSSEN (1962) et DE FINETTI (1965), SHUFORD, ALBERT et MASSENGILL (1966) ont développé des "procédures admissibles de mesure des probabilités subjectives". Nous en avons tiré des versions adaptées à la pédagogie.

Malgré toutes les raisons théoriques qui amènent à préférer des systèmes de cotation sophistiqués, il faut aussi être conscient de leurs faiblesses, comme le montre l'expérience de RIPPEY.

R.M. RIPPEY (1970) a présenté un test à un millier d'étudiants de l'enseignement secondaire américain. La consigne était la suivante (p. 167) : "Attribuez un nombre de 0 à 9 à chaque solution proposée, selon que vous pensez que la réponse est plus ou moins correcte. Si vous pensez qu'une seule solution proposée est correcte, indiquez un 9 en face de cette réponse et indiquez un 0 pour les autres."

Après cette consigne (proposant une échelle typiquement ordinale), l'auteur montre au sujet des exemples variés de réponses. La consigne continue alors comme suit : "Votre copie sera cotée de telle façon que vous obteniez le score le plus élevé en estimant votre degré de certitude et en le transcrivant avec précision. Deviner d'une manière ou d'une autre diminuera votre cote. Si vous ne savez rien sur une question et que vous n'avez pas de préférence, vous obtiendrez votre score le plus élevé en distribuant honnêtement votre certitude parmi toutes les options."

Remarquons que la dernière phrase concerne les sujets en situation d'ignorance totale; la consigne indique au sujet comment répondre là où l'omission s'imposerait dans une situation pédagogique.

L'originalité de l'expérience de RIPPEY est de ne pas dévoiler aux élèves la matrice de cotation. En fait, RIPPEY va utiliser cinq fonctions différentes et donc obtenir cinq scores différents pour chaque élève.

La première fonction de calcul est la probabilité attribuée à la réponse correcte (p_c = probabilité réponse correcte).

La deuxième est logarithmique. Si la probabilité de la bonne réponse est inférieure à .01, le score vaut 0. Sinon, il vaut $(2 + \log_{10} p_c) / 2$.

La troisième est sphérique. Le score vaut $pc \Big/ \left(\sum_{i=1}^{n} (pi)^2 \right)^{1/2}$ avec pi = probabilité des diverses solutions.

La quatrième est euclidienne. Le score vaut :

$$1 - \left[\left(\sum_{i=1}^{n} (pi - ki)^2\right)^{1/2}\right] / \sqrt{2}$$

avec ki = probabilité de référence attribuée à la solution i par un ensemble de juges.

La cinquième est le choix inféré : le score vaut 1 si la probabilité attribuée à la réponse correcte est plus élevée que les autres probabilités, sinon le score est 0.

RIPPEY souligne que la première procédure est la plus intuitive à présenter et qu'elle permet les calculs les plus simples. Les fonctions logarithmiques et sphériques partagent l'intéressante propriété de permettre au sujet de maximiser son score si - et seulement si - il ne devine pas mais ajuste ses réponses à ses ESPER. La fonction euclidienne permet de corriger les questions qui admettent plus d'une solution correcte. La procédure du choix inféré est similaire à la cotation la plus simple qui existe (1 point par réponse correcte, 0 par réponse incorrecte ou par omission), car on se fonde - à juste titre - sur le principe de la stratégie pure (voir plus haut).

Des comparaisons de fidélités ont amené RIPPEY à conclure que la première fonction donnait la fidélité la plus élevée et donc qu'en l'absence d'information sur le système de cotation, les sujets attribuent leur certitude sur la base du modèle de gains et de pertes le plus simple. En outre, la fidélité décroît à mesure que les scores calculés s'éloignent du score simple (fonction 1). La deuxième fonction quant à la fidélité, s'avère être la fonction euclidienne. La moins bonne est la fonction 5, ou cotation classique. En fait, les moyennes des indices de fidélité (1) sont les suivants :

- 1. Probabilité de la bonne réponse .69
- 2. Logarithmique .50
- 3. Sphérique .49
- 4. Euclidienne .58
- 5. Choix inféré .47

(1) Estimés par l'analyse de variance de HOYT. La moyenne a été calculée après transformation de r en z.

RIPPEY (p. 169) conclut que "Les scores calculés à partir des fonctions ésotériques présentent une composante d'erreur due au manque de compréhension du système de cotation de la part des élèves."

Il illustre cette composante d'erreur par la surface hachurée dans les graphiques ci-dessous. L'abscisse représente la probabilité attribuée à la bonne réponse. L'ordonnée représente le score calculé.

Si le sujet croit que son score sera égal à la probabilité qu'il attribue à la bonne réponse et que le score est, de fait, calculé ainsi, les observations se mettront sur la diagonale; c'est le cas pour la fonction 1. Si les scores sont calculés autrement, il existe une discordance plus ou moins importante qui ne manquera pas d'avoir un effet sur la composante d'erreur du score.

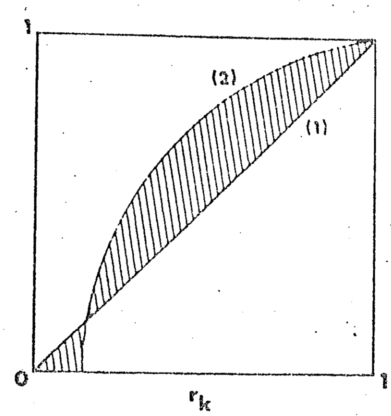


Figure 1. Graph of Scores for Functions 1 and 2.

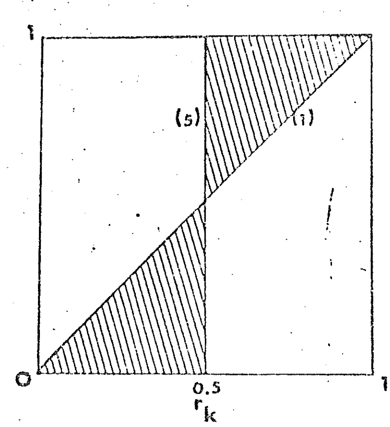


Figure 2. Graph of Scores for Functions 1 and 5.

Nous serons attentif à cet aspect "limpidité"du système.

LES CONDITIONS D'UNE ETUDE VALIDE DES CERTITUDES

L'indication de sa certitude par un sujet est, *a priori*, suspecte d'être, volontairement ou non, biaisée par des attitudes (crainte ou attrait du risque), par des stratégies (audacieuses ou timorées), sans rapport avec la connaissance réelle. Ainsi, le besoin d'accomplissement (McClelland, 1955; ATKINSON, 1964) peut influencer considérablement le comportement. Ces attitudes et ces stratégies du sujet font le jeu des détracteurs des systèmes de notation de l'ESPER, mais ils en tirent des conclusions trop radicales. Une grande partie de nos expériences met en lumière "l'interférence de facteurs de personnalité et de stratégie dans l'évaluation d'une connaissance" (1). Cette démonstration n'est plus à faire depuis les travaux de WILEY et TRIMBLE (1936), HEVNER (1932), SWINEFORD (1941) ou, plus récents, ceux de JACOBS (1971).

Nous sommes donc d'accord avec L. D'HAINAUT (1974, p. 59) lorsqu'il dit :

> "En introduisant des coefficients d'assurance dans les réponses aux questions fermées, on évalue donc autre chose que ce qu'on voulait mesurer et un gain éventuel de fidélité perd sa valeur s'il s'accompagne d'une diminution de la validité : on peut même penser que c'est la composante de personnalité introduite dans l'épreuve qui contribue pour beaucoup à l'accroissement de fidélité (*). On pourrait arguer que cette composante apporte une information supplémentaire et met en évidence une plus grande variété dans l'aptitude comme le rapporte ECHTERNACHT (1965). C'est juste, sans doute, mais l'information est mélangée si bien qu'en termes d'information utile, il y a, au contraire, une diminution : au lieu d'étudier séparément les facteurs, on les mélange sans avoir isolé celui qu'on visait, la connaissance ou plus exactement la compétence."
>
> (*) Pendant que cet article était sous presse, Hopkins K.D., Hakstian R.A. et Hopkins R.B. ont précisément montré que la pondération des réponses par l'assurance entraînait un léger accroissement de fidélité mais une diminution de validité.

(1) D'HAINAUT, 1974, p. 59.

Et sa conclusion s'applique, hélas, à bien des études sur le sujet.

> "En fin de compte, nous ne sommes pas loin de penser comme Lord et Novick (1968) que l'intérêt de l'évaluation pondérée par la confiance réside surtout dans son attrait intellectuel et nous ajouterons même qu'à l'examen, ce beau flacon se révèle une bouteille à encre : on y a tout mêlé pour une illusion." (p. 60)

Notre thèse vise à dépasser l'apparente impossibilité d'utiliser l'ESPER, non en infirmant les constatations qui précèdent, mais en changeant la perspective comme l'ont suggéré (sans le réaliser) nombre d'expérimentateurs. Les constatations faites en une séance de testing unique, sur des sujets non entraînés et qui ne sont pas exposés aux conséquences de leurs actes, ne peuvent être considérées avec le même intérêt que les résultats de séances répétées dans une procédure operante.

Avec l'école de SKINNER, on peut penser que l'exposition à un programme de renforcement apporte une efficacité supplémentaire à l'explication "intellectuelle" de consignes. Autrement dit, le comportement dirigé par une règle et contrôlé par un programme de renforcement est un comportement qui permettra vraisemblablement de recueillir des mesures plus valides, plus fidèles et plus sensibles.

> "Le nombre d'organismes étudiés est généralement beaucoup plus réduit que dans les plans d'expérience prévus pour satisfaire aux règles de la statistique, mais, par contre, chaque organisme est soumis à l'observation pendant beaucoup plus de temps." (SKINNER, 1971, p. 118).

Dans cette perspective, nous avons recueilli des observations continues où les mêmes sujets étaient soumis à la procédure de cotation durant une année entière. Ces sujets ne sont pas nombreux, mais ils ont vécu plusieurs situations expérimentales (réponses suivies de l'exposition aux conséquences réelles). Ces observations se distinguent donc radicalement des autres expériences au cours desquelles des centaines de sujets ont été testés simultanément : 535 par KOEHLER (1971), un millier par RIPPEY (1970), un millier par SANDBERGEN (1972), une centaine par MELLENSBERGH (1967), une centaine par SANDBERGEN (1968).

1. Chez De Finetti, il faut utiliser un ordinateur pour calculer les notes

2. L'intensité de la motivation de maximisation n'est pas la m. chez tous les sujets.

Dans ces recherches, les sujets n'ont été, dans aucun cas, "spécialement entraînés aux intrications des procédures de notation de la certitude" (KOEHLER, 1971, p. 302). Pourtant, tous les auteurs sont d'accord : "L'entraînement devrait diminuer l'influence des facteurs d'attitudes" (VAN NAERSSEN et VAN BEAUMONT, 1965; SANDBERGEN, 1971).

Bruno DE FINETTI (1965) a le premier posé clairement le problème et envisagé les aspects mathématiques, psychologiques et méthodologiques de la probabilité subjective dans son important article : *Methods for Discriminating Levels of Partial Knowledge Concerning a Test Item*. Les trois points importants de son approche sont les suivants :

> "1. Les sujets doivent connaître aussi bien la méthode de cotation que les modalités de réponse permises... Ils doivent en comprendre pleinement les implications, notamment en cas d'incertitude dans la réponse.
>
> 2. Le sujet doit désirer obtenir un score total élevé, et le maximiser.
>
> 3. Les élèves doivent être entraînés à traduire leurs certitudes en probabilités numériques". (p. 89)

Les procédures expérimentales devront, elles aussi, s'imposer des contraintes, comme le montre l'expérience suivante. HAMBLETON, ROBERTS et TRAUB (1970) trouvent que l'utilisation d'un indice de certitude augmente la validité prédictive de .62 à .72 et que le test classique devrait être huit fois plus long pour atteindre la validité et la fidélité de la mesure obtenue avec l'indice de certitude. Cependant, de l'avis même des auteurs, il y a trois raisons d'émettre d'importantes réserves. Tout d'abord, le petit nombre de sujets (environ 70) ne permet pas de généraliser ces conclusions. Ensuite, la procédure "avec certitude" dure 5 à 10 minutes de plus que l'autre. Enfin, la facilité moyenne des questions est de 75 % et, dans 77 % des cas, les sujets ont noté la solution choisie d'une certitude de 100 %. Une telle facilité rend le test peu propice à l'étude envisagée. Nous serons sensible à ce dernier problème dans les expériences qui suivent.

Qu'appelez-vous "Connaissance partielle"
Il faudrait préciser.

Il n'existe pas dans tous les cas et il ne faut pas confondre avec l'impression de connaissance partielle.

Par exemple, on peut avoir l'impression de savoir partiellement ce qu'est un insecte si on n'est pas sûr qu'une araignée n'en est pas un. En réalité, on ne sait pas alors ce qu'est un insecte, on sait ce qu'est une bestiole.

THESES ET HYPOTHESES GENERALES

Notre première thèse générale est la suivante :

> La connaissance partielle existe. En pédagogie, il est utile et possible de l'évaluer, au moyen de l'évaluation subjective de la probabilité d'exactitude d'une réponse (ESPER). Les conditions optimales d'étude de la validité, de la fidélité et de la sensibilité de l'ESPER sont les suivantes :
>
> 1.- La procédure de testing doit être operante : les sujets doivent être exposés aux conséquences de leurs actes (des points ou toute autre valeur jugée utile par les sujets doivent dépendre des réponses fournies).
> 2.- Les sujets doivent avoir été entraînés à l'attribution de leurs certitudes (par exemple en étant exposés de façon répétée au programme de renforcement).
> 3.- Les conséquences doivent être calculées selon la théorie des décisions afin que la stratégie optimale consiste à exprimer le plus exactement possible son ESPER.
> 4.- La consigne doit présenter des certitudes correspondant à une échelle des probabilités d'exactitude (PER).

Les points 3 et 4 de notre thèse ne nous sont apparus qu'en cours de travail, à la lumière des résultats expérimentaux. Cela explique que les expériences décrites aux chapitres 2 et 4 ne présentent pas toutes les conditions de la thèse ci-dessus.

Notre deuxième thèse générale est que

> L'ESPER s'avère féconde dans l'explication de phénomènes tels que la modification des comportements après information et la prise d'information en situation d'incertitude.

L'hypothèse qui découle de cette deuxième thèse est tout naturellement :

Le théorème de BAYES (1) s'applique (éventuellement à une constante près) dans les situations cognitives proches des situations pédagogiques.

Autrement dit, la corrélation entre les valeurs prédites par le théorème et les valeurs observées doit être élevée.

Les notes théoriques et les expériences décrites dans les chapitres 1, 2, 3 et 4 sont consacrées à la validité de l'ESPER. Les chapitres 5 et 6 sont consacrés à l'étude de la fidélité (test-retest) et de la sensibilité de l'ESPER. Le chapitre 7 relate des expériences sur la révision des ESPER après information, notamment par le théorème de BAYES. Ce dernier chapitre ne rend compte que de deux expériences mais elles sont indicatives des voies de recherche nouvelles qui s'ouvriraient si les six premiers chapitres s'avèraient convaincants.

Dans chacun de ces chapitres, les hypothèses seront formulées de façon plus précise.

Dans les chapitres 2 et 4, nous ferons part d'expériences répétées, suivies d'effets réels et immédiats, menées dans le milieu scolaire lui-même. Pour exposer un grand nombre d'élèves aux conséquences de leurs (nombreuses) réponses nuancées, il fallait disposer d'une banque de questions gérée par ordinateur. Bien que l'idée d'une telle banque ne soit pas neuve et qu'elle connaisse actuellement un certain succès, l'important était surtout d'en faire fonctionner une dans la réalité, en résolvant les problèmes qui y sont attachés, depuis la formation des professeurs qui l'utilisent jusqu'à la rédaction du programme d'ordinateur, en passant par des problèmes d'organisation non négligeables.

(1) Le théorème de BAYES (voir chapitre 6) lie dans une équation les probabilités du sujet <u>avant</u> information, ses probabilités <u>après</u> information et la vraisemblance de l'information.

Style cognitif

Nous avons eu le plaisir de créer la première banque de questions en Belgique et d'en lancer deux autres. La description du fonctionnement de ces banques et les options docimologiques et technologiques ont été décrites ailleurs (D. LECLERCQ, 1971, 1973, 1975).

La validité, la fidélité et la sensibilité de l'ESPER varient vraisemblablement avec l'âge, le degré de scolarité et l'entrainement des sujets, avec la matière considérée, avec les documents disponibles, etc. Les expériences qui vont être décrites dans le présent travail ne constituent, on le comprendra aisément, que des réponses très parcellaires à l'ensemble des questions. C'est surtout sur le plan méthodologique que nous souhaitions faire oeuvre utile.

CHAPITRE 1

EXAMEN DE L'ESPER A LA LUMIERE DES THEORIES ECONOMIQUES ET PSYCHOLOGIQUES

PREMIERE PARTIE : *La théorie des décisions*.

A. Actions, états et conséquences
B. Prise de décision et utilité des conséquences
C. L'utilité des scores aux épreuves scolaires est-elle linéaire ?
D. Certitude, incertitude et risque
E. Stratégies possibles en situation d'incertitude
F. La connaissance partielle et la théorie moderne de l'utilité (E.S.U.)

DEUXIEME PARTIE : *L'approche psychologique*.

A. Prise de risque et besoin d'accomplissement
B. La théorie du dépliage de C.H. COOMBS

Les coefficients de certitude fournis par le sujet mesurent-ils la probabilité d'exactitude de la réponse (PER) et rien que cette probabilité ?

On suspecte à raison les indices de certitude d'être biaisés par divers facteurs étrangers à la connaissance. Tout d'abord, la situation se présente comme un jeu où divers paris sont disponibles, comme à la roulette. Le sujet a, en effet, le choix entre plusieurs indices de certitude auxquels sont attachés des gains et des pertes. Le critère de choix d'un indice de certitude devrait être la probabilité d'exactitude de la réponse (PER) et elle seule. Il n'en est pas toujours ainsi et, dans le présent chapitre, les autres critères de choix seront examinés. Il arrive que certaines stratégies soient objectivement plus efficaces que d'autres et, même si ce n'est pas le cas, le sujet, selon sa personnalité, peut être persuadé de la supériorité de l'une d'entre elles et l'adopter. Ainsi, les "risqueurs" utiliseront systématiquement des indices de certitude élevés alors que les "modérés" se cantonneront dans les certitudes faibles.

Dans les expériences des chapitres 2 et 4, ces stratégies et une série d'autres seront systématiquement recherchées dans les *patterns* comportementaux de tous les sujets à tous les tests. Si de telles stratégies sont répandues et perdurent, il faut renoncer à interpréter la certitude en termes d'ESPER.

Les stratégies possibles lors d'un choix en situation d'incertitude ont été décrites par la théorie des décisions en économie et par la théorie du dépliage en psychologie mathématique. Le présent chapitre expose ces théories, en adaptant les exemples et les raisonnements au cas particulier de l'utilisation d'indices de certitude dans l'évaluation de performances cognitives. De la théorie des décisions, seuls les concepts et les outils relatifs aux jeux "contre la nature" seront examinés, contrairement aux jeux "avec partenaires".

Première partie : La théorie des décisions

A. ACTIONS, ETATS ET CONSEQUENCES

Avant d'examiner les raisonnements mathématiques, précisons les termes qui sont utilisés dans la théorie des décisions.

> "Un problème préliminaire qui se pose est d'établir la liste des actions disponibles (...). La conséquence du choix de l'une des actions possibles doit dépendre de l'"état de la nature". Rappelons que cette expression, devenue conventionnelle, désigne les situations que le preneur de décision peut avoir à considérer, et sur lesquelles il n'exerce en principe aucun pouvoir... L'hésitation dans les choix des actions naît notamment de l'incertitude relative aux états de la nature... Le tableau des résultats ou conséquences (pay-off matrix) donne les résultats, gains ou coûts, résultant du fait de choisir les actions A1, A2,... lorsque les états de la nature sont E1, E2 (...) respectivement (...)." (Ch. DE BRUYN, 1973, 7, pp. 3-4)

Appliquons les principes décrits par les économistes à la pédagogie. Partons de l'exemple d'une question à choix multiple qui propose quatre solutions possibles, la consigne précisant aux sujets qu'il n'y a qu'une solution correcte et qu'elle se trouve parmi ces quatre solutions.

Si l'on veut savoir quelle stratégie l'élève utilise pour choisir une solution, on peut représenter le problème comme suit :

- Quatre actions disponibles :

  A1 = Choisir la solution 1
  A2 = Choisir la solution 2
  A3 = Choisir la solution 3
  A4 = Choisir la solution 4

- Quatre états possibles (on dit aussi "états de la nature") :

  E1 = La solution 1 est correcte
  E2 = La solution 2 est correcte
  E3 = La solution 3 est correcte
  E4 = La solution 4 est correcte

On peut alors dessiner une matrice 4 x 4 de conséquences qui représente les seize situations à envisager.

Nous allons concentrer notre attention non pas sur la solution choisie, mais sur l'indice de certitude choisi après que la réponse ait été donnée. Il s'agit donc d'une action conditionnelle. La notation "certitude|R" (certitude étant donné que la réponse est déjà fournie) est lourde et ne sera pas répétée par la suite; le lecteur voudra bien s'en souvenir.

On précise aux étudiants qu'ils sont obligés de répondre à toutes les questions (1) et qu'ils doivent associer un des trois indices de certitude suivants à chacune de leurs réponses :

Certitude 1 = faible
Certitude 2 = moyenne
Certitude 3 = forte.

Une fois la réponse donnée, l'étudiant se trouve donc devant trois actions disponibles :

A1 = choix de la certitude 1
A2 = choix de la certitude 2
A3 = choix de la certitude 3.

Les états possibles sont au nombre de deux (2) :

E1 = la solution choisie est correcte
E2 = la solution choisie est incorrecte.

La matrice de conséquences *(pay-off matrix)* contient les gains et les pertes réalisés dans ces six situations. Nous avons fourni dans l'introduction un certain nombre d'exemples de telles matrices.

---

(1) Il s'agit ici d'un exemple propre à rendre la démonstration mathématique simple. Nous discuterons plus loin de la place à faire à l'omission.

(2) C'est aussi arbitrairement que l'on a choisi de proposer, dans le présent exemple, des solutions entièrement fausses ou entièrement correctes.

Nous en rappelons un seul ici :

		Etats La solution choisie est correcte	incorrecte
Actions	A1 = certitude 1	+ 1	- 1
	A2 = certitude 2	+ 2	- 2
	A3 = certitude 3	+ 3	- 3

Les états possibles de la qualité de la réponse sont fixés par le correcteur; l'élève n'a aucun pouvoir sur eux, mais il doit en connaître le nombre et les modalités. Par contre, l'élève choisit en toute liberté une action parmi celles qui sont disponibles, et il doit aussi connaître celles-ci avant de prendre sa décision.

B. PRISE DE DECISION ET UTILITE DES CONSEQUENCES

Empruntons de nouveau la définition du terme "décision" à un économiste.

> "Une personne est placée devant un problème de décision chaque fois qu'il y a un choix entre au moins deux actions. Une femme qui n'a qu'une robe n'a pas à décider laquelle porter. La première chose à faire dans toute situation de décision est de considérer quelles sont les actions disponibles. Il n'est pas nécessaire de faire une distinction entre la décision et l'action. La décision de porter une certaine robe n'a pas besoin d'être distinguée de l'action de porter la robe : en effet, si la décision n'a pas conduit au port de la robe, c'était probablement parce que quelque chose est intervenu pour l'empêcher, et un nouveau problème de décision se pose..." (LINDLEY, 1971, p. 4).

De même que, en psychologie, la notion de débit de réponses ne prend sa valeur que lorsqu'elle est associée à la notion de programme de renforcement, de même la notion de décision réclame la notion corollaire d'<u>utilité</u>.

> "L'utilité est un nombre qui mesure l'attrait d'une conséquence - plus l'utilité est élevée, plus la conséquence est désirable - la mesure étant faite sur une échelle de probabilité." (LINDLEY, 1971, p. 70)

L'expression "Echelle de probabilité" est généralement comprise dans l'acception qui suit :

> "On attribue généralement l'utilité 1 à la meilleure conséquence (...) et l'utilité 0 pour la plus mauvaise conséquence. Ainsi, les utilités posées de la sorte s'échelonnent nécessairement entre 0 et 1. Mais, si nous multiplions toutes les utilités u par n'importe quelle constante positive, et ajoutons n'importe quelle autre constante, les nombres résultants conduisent aux mêmes décisions que les utilités originales... Ce fait peut être exprimé en disant que les échelles d'utilité sont uniques à une transformation linéaire croissante près." (DE BRUYN 1973, <u>9</u>, p; 12).

Ceci est en fait, le cinquième théorème de VON NEUMANN et MORGENSTERN (COOMBS, DAWES et TVERSKY, 1970, p. 126) (1).

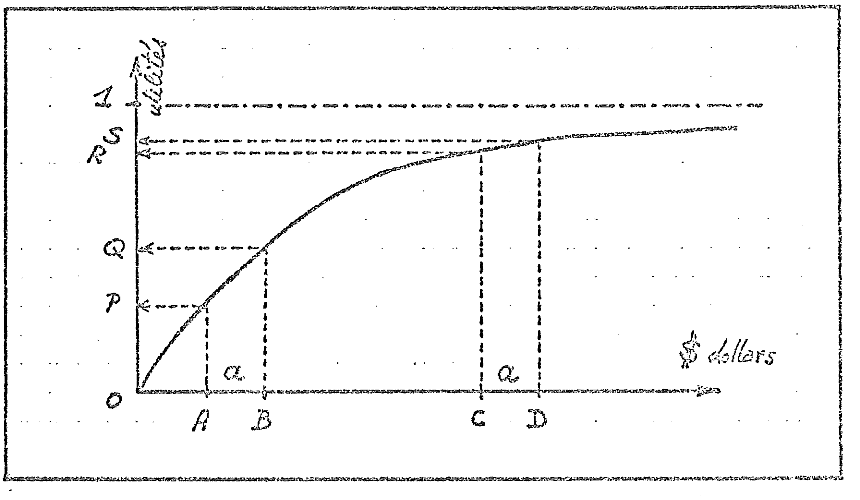
---

(1) Démontré par LUCE et RAIFFA (1966, pp. 23-31).

Les économistes considèrent particulièrement les conséquences financières et expriment les utilités en valeurs monétaires (dollars ou francs), à condition qu'il y ait linéarité entre l'utilité et la valeur en monnaie.

Exprimer les utilités en argent ou en points n'est légitime que si la fonction d'utilité considérée est linéaire. Or, pour l'argent, par exemple, c'est rarement le cas, comme le montre le tableau 1.1 (LINDLEY, 1971, p. 77). Cette fonction (une ogive) est croissante, concave et a une asymptote horizontale.

Tableau 1.1



Relation classique entre les utilités et les valeurs monétaires (d'après D. LINDLEY, 1971).

On constate que (B-A) = (D-C) = a = augmentation marginale de a dollars. Pour une même augmentation <u>marginale</u> (a) de A en B, on obtient un accroissement de l'utilité marginale de Q-P, tandis que cet accroissement n'est que de S-R pour une même augmentation de C à D.

L'explication de cette curvilinéarité est psychologique :

> "Supposez que vous considériez une conséquence à laquelle est assuré le gain, disons de cinquante millions de francs (...), il vous serait difficile de distinguer entre cinquante millions et cent millions. Bien sûr, la dernière somme est préférable, mais la première vous permettrait de faire toutes ces choses merveilleuses dont vous rêviez depuis longtemps et d'acheter tous ces merveilleux gadgets que vous n'espériez pas pouvoir obtenir un jour, si bien que cinquante millions en plus de ceux-là ne feraient que dorer une situation déjà très attrayante. Pour nous exprimer autrement, il arrivera un moment où le capital supplémentaire cessera de vous émouvoir."
> (LINDLEY, 1971, p. 77)

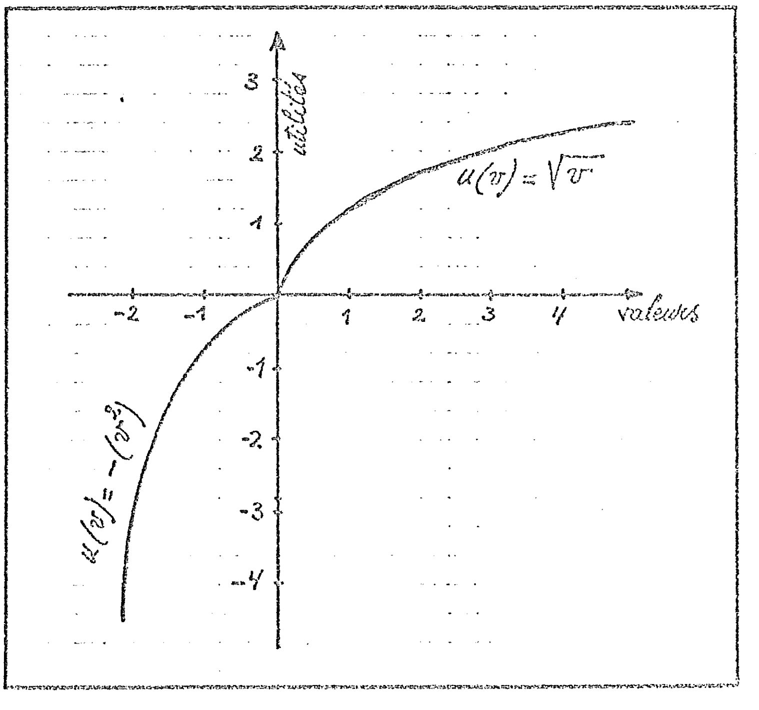
Ce sont les utilités, plus que les valeurs elles-mêmes, qui guident le comportement humain. Considérons un jeu où, sur le lancer d'une seule pièce de monnaie, on peut gagner ou perdre 50000 francs. Les chances de gagner et de perdre sont égales et devraient donc se compenser. Pourtant, très nombreuses sont les personnes qui refuseront de jouer à ce jeu, car, subjectivement pour elles, la perte possible ne compense pas le gain possible.

Ces considérations nous amènent à faire remarquer que les fonctions d'utilité marginale des pertes sont assez souvent différentes de celles des gains. COOMBS, DAWES et TVERSKY (1970, p. 120) en donnent un bel exemple théorique, où l'utilité vaut la racine carrée de la valeur pour les gains et le carré (négatif) de la valeur pour les pertes :

$$\text{Utilité de } v = u(v) = \begin{cases} \sqrt{v} & \text{si } v \geq 0 \\ -(v^2) & \text{si } v \leq 0 \end{cases}$$

Sur le tableau 1.2, ces deux fonctions apparaissent bien différentes.

Tableau 1.2



Fonctions d'utilité des pertes et des gains (Everyman's utility function).

On peut imaginer d'autres fonctions, par exemple logarithmiques, mais toujours concaves; la fonction d'utilité est négativement accélérée, comme les fonctions classiques de perception en psychophysique.

L'ensemble des deux fonctions décrites ci-dessus a été proposé, dès le XVIIIe siècle, comme la fonction d'utilité de tout homme (*Everyman's utility function*), idée qu'a défendue S. STEVENS (1959), un des plus grands psychophysiciens contemporains, sur la base de résultats expérimentaux.

Ces similitudes avec la loi de FECHNER ne sont pas accidentelles.

En effet, dès 1738, Daniel BERNOUILLI, pour résoudre des difficultés apparaissant dans certains problèmes (1), formula le principe de l'utilité attendue qui devait remplacer l'échelle objective des valeurs par l'échelle subjective des utilités. Plus de cent années plus tard, la notion bernoullienne d'échelle subjective devient la pierre angulaire de la psychophysique fondée par G.T. FECHNER.

BERNOUILLI avait introduit la notion de valeur morale (que nous appelons aujourd'hui "utilité"), par opposition à la valeur faciale de la monnaie. Il supposait que la première était égale au logarithme de l'autre. Cette notion donna à Fechner l'idée de relier la valeur morale du stimulus, c'est-à-dire de la sensation, à sa valeur physique, par la même fonction logarithmique.

(1) Par exemple, le paradoxe dit "de Saint-Pétersbourg" (parce qu'il fut publié dans les Annales de l'Académie de Saint-Pétersbourg). On en trouvera la description dans F. BRESSON (1965, p. 238).

C. L'UTILITE DES SCORES AUX EPREUVES SCOLAIRES EST-ELLE LINEAIRE ?

Il nous paraît raisonnable de supposer que la relation entre la valeur des notes scolaires et leur utilité est linéaire, du moins quand la marge de variation possible (différence entre les valeurs extrêmes possibles) est assez restreinte.

Ce problème a été peu abordé expérimentalement. A notre connaissance, l'étude la plus précise menée sur ce sujet a été réalisée par l'équipe hollandaise de VAN NAERSSEN, SANDBERGEN et BRUYNIS (1966). Le titre de leur article (*Is de utiliteitscurve van examenscores een ogief ?*) décrit une hypothèse aussi plausible que celle de la linéarité. Leur procédure expérimentale et leurs résultats méritent d'être décrits en détails.

Ils ont proposé à deux groupes de sujets des épreuves où les points sont gagnés (1) non seulement par le nombre de réponses correctes, mais, en plus, par un indice de certitude noté Z *(Zekerheid)*. Opérationnellement, l'hypothèse de l'ogive est formulée comme suit :

> "Les étudiants dont le nombre d'items corrects est faible, anxieux quant à leurs résultats, devraient indiquer relativement plus souvent Z (certitude) que les candidats qui ont répondu correctement à un grand nombre d'items. Ce dernier groupe ayant un moins grand besoin de points, on pourrait s'attendre à ce qu'il utilise moins l'indice de certitude."

Cette hypothèse a été testée séparément, mais selon une procédure identique, au cours de deux épreuves :

- 1964 : sur 147 sujets qui avaient répondu à 50 items;
- 1965 : sur 117 sujets qui avaient répondu à 80 items.

---

(1) La matrice des conséquences est celle de VAN NAERSSEN et VAN BEAUMONT (1965) qui est décrite dans l'introduction.

Voici les étapes du traitement des résultats :

1) Les sujets ont été répartis en trois groupes, selon leur nombre total de réponses correctes.

   A = forts, B = moyens, C = faibles.

2) Les pourcentages de difficulté de chaque question (p) et le pourcentage de réponses "certaines" (Z) ont été calculés séparément pour chacun des trois groupes A, B et C.

3) On a constitué des paires d'items (souvent différents) de difficulté équivalente dans deux groupes pris deux à deux. Dans la première épreuve, on n'a pu constituer que

   - 39 paires pour les groupes A et B
   - 34 paires pour les groupes B et C
   - 26 paires pour les groupes A et C.

4) On a alors créé, dans chaque série d'items, trois groupes de difficulté, pour présenter graphiquement les résultats: items difficiles (1), moyens (2) et faciles (3) dont on a calculé les moyennes $\bar{p}$ et $\bar{z}$.

5) Les neuf valeurs disponibles (A1, A2, A3, B1, B2, B3, C1, C2, C3) de $\bar{p}$ et $\bar{z}$ ont été disposées dans un graphique (tableau 1.3).

Tableau 1.3

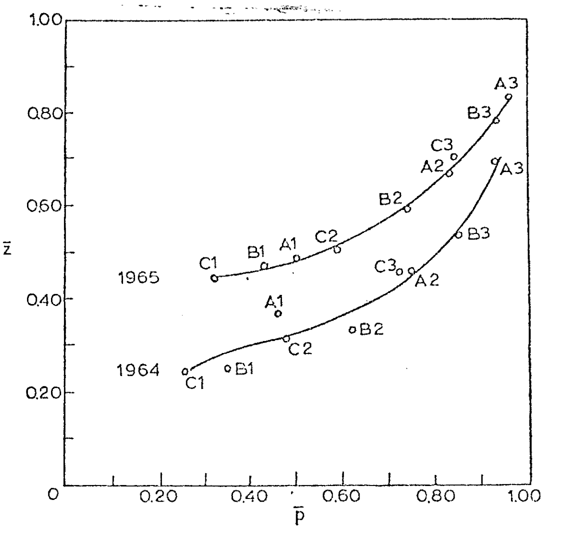


Fig. 1. Gemiddelde p- en z-waarden van een volgens „totaal aantal items goed" hoge personengroep (A), een middengroep (B) en een lage groep (C) berekend bij 3 itemgroepen: moeilijk (1), gemiddeld (2) en gemakkelijk (3). Tentamens 1965 en 1964.

On peut faire les constatations suivantes :

1) Les courbes se présentent sous la forme d'un S en miroir (dont on ne voit pas la portion gauche par manque d'items difficiles).
2) Les items difficiles à p = 50 correspondent à une proportion de certitudes de Z = .50 (pointillés).
3) Les courbes sont plus horizontales vers les items difficiles.

Admettons un instant que l'hypothèse selon laquelle la fonction de l'utilité est une ogive soit correcte. On devrait alors s'attendre à ce que les valeurs B soient approximativement situées sur la courbe, les A en dessous et les C au-dessus. Or ni le graphique, ni les tests de signification ne laissent penser que l'hypothèse se vérifie. Par conséquent, les auteurs ne voient pas d'objection à l'emploi d'un indice de certitude tel qu'il est décrit dans l'expérience (1).

Sans disposer de plus ample démonstration, nous avons admis, jusqu'à preuve du contraire, que la fonction d'utilité des points d'une épreuve scolaire est linéaire, et, dès lors, nous parlerons des valeurs (les points gagnés ou perdus) en lieu et place des utilités. Ce problème mériterait un approfondissement, quoiqu'il ne semble pas avoir joué un grand rôle dans nos expériences; nous nous garderons d'être trop catégorique à ce sujet.

---

(1) D'autres résultats intéressants ont été mis en évidence lors de la même expérience. Ainsi, la corrélation entre p et z était de .85 dans l'expérience de 1964 et .72 dans celle de 1965.

D. CERTITUDE, INCERTITUDE ET RISQUE

> "Le domaine de la prise de décision est généralement partitionné selon que la décision est placée dans une situation de certitude, de risque, ou d'incertitude.
>
> Il y a certitude si l'on sait que chaque action envisagée conduit invariablement à un résultat spécifié.
>
> Il y a risque si chaque action conduit à un élément d'un ensemble de résultats possibles, chaque résultat se présentant avec une probabilité connue du preneur de décision...
>
> Il y a incertitude si les actions ont pour conséquence un ensemble de résultats spécifiques possibles, mais les probabilités attachées à ces résultats sont totalement inconnues ou, éventuellement, n'ont pas de signification."
> (LUCE and RAIFFA, 1957, chap. 2, 1)

Luce et Raiffa introduisent une quatrième catégorie : une combinaison d'incertitude et de risque, considérée à la lumière de résultats expérimentaux. C'est ce type de situation qui correspond le mieux aux problèmes pédagogiques.

> "L'école conduite par SAVAGE (1954) considère que, par le traitement de sa propre information partielle (comme, par exemple, répondre à une série de questions hypothétiques simples du type oui-non), on peut générer une distribution de probabilité *a priori* sur les états de la nature, distribution appropriée à la prise de décision. Ceci réduit le problème de l'incertitude à un problème de risque." (LUCE and RAIFFA, 1957, p. 300)

Soulignant l'intérêt d'une explicitation du risque pour l'administration des affaires, DE BRUYN (1973, 6, p. 4) énumère les raisons qui entourent le rejet d'une telle explication. La plus courante des raisons est "... le défaut de pratique d'un langage adéquat et non ambigu. Les termes de la statistique ont une signification dans les modèles formalisés, mais correspondent-ils à ceux du langage courant ?" A titre d'exemple, DE BRUYN compare, avec J.S. LAMBIN, certaines expressions et les probabilités auxquelles elles correspondent.

C'est inexact, car on ne considère qu'un item.

Expressions	Probabilités
c'est impossible	0
peu probable	0 à 0,05
douteux	0,5 à 0,25
vraisemblable	0,25 à 0,45
possible	0,45 à 0,55
très vraisemblable	0,55 à 0,75
très possible	0,75 à 0,95
sûr	0,95 à 0,99
certain	1

Des stratégies de décision ont été mises au point pour les cas où le sujet doit répondre sous incertitude totale quant aux états possibles. Nous allons décrire ces stratégies à partir d'un exemple pédagogique simple. L'omission, qui pose un problème particulier, sera étudiée plus tard.

La *correction for guessing* classique est une application simple de l'espérance mathématique en situation d'ignorance totale :

Décisions	E1 (réussite) p1 = .2	E2 (échec) p2 = .8
Choisir solution 1	+ 4	- 1
Choisir solution 2	+ 4	- 1
Choisir solution 3	+ 4	- 1
Choisir solution 4	+ 4	- 1
Choisir solution 5	+ 4	- 1

Ceci n'est pas exact !

Les chances d'obtenir la bonne réponse (E1) en répondant totalement au hasard sont de 1/5, soit .2 (p1). Les chances de se tromper (E2) sont évidemment: 1 - .2 = .8 (p2). La valeur attendue ($\bar{v}$) de chacune des i actions (ici 5) se calcule par la formule :

$$\boxed{\bar{v}_i = p_{i1} \cdot v_{i1} + p_{i2} \cdot v_{i2}}$$

Dans ce cas-ci, $\bar{v}_1 = (.2 \times 4) + (.8 \times -1) = .8 - .8 = 0$

Tous les $\bar{v}_i$ sont égaux : $\bar{v}_1 = \bar{v}_2 = \bar{v}_3 = \bar{v}_4 = \bar{v}_5 = 0$
L'élève n'a donc pas plus de raison de choisir une solution plutôt qu'une autre, s'il s'en remet au seul hasard. En outre, puisque la valeur attendue vaut 0, il peut aussi bien omettre de répondre.

La valeur attendue (*expected utility*) est aussi appelée l'espérance mathématique de la valeur (de chaque action).

Quand le nombre d'événements possibles dépassé 2 (j 2), la formule générale de calcul est :

$$\boxed{\bar{v}_i = \sum_{j=1}^{n} v_{ij}\, p_j}$$

où j est l'indice des n événements potentiels (Ej)
et i est l'indice de m actions disponibles (a_i).

Quand il n'y a que deux événements possibles (n = 2), où p1 = prob. de réussite et p2 d'échec, on utilise souvent la notation p pour p1 et q pour p2.

Les formules qui viennent d'être évoquées servent à calculer la valeur attendue (ou espérance mathématique) des conséquences de chaque action et, par conséquent, à choisir l'action optimale, c'est-à-dire celle dont la valeur attendue est la plus élevée. Cette procédure est connue depuis longtemps.

L'apport de BERNOUILLI et des suivants a été de remplacer, dans la formule, les valeurs par les utilités :

$$\boxed{\bar{u}_i = \sum_{j=1}^{n} u_{ij}\, p_j}$$

Depuis VON NEUMANN et MORGENSTERN (1947), le critère de l'utilité attendue désigne le choix de l'action (i) qui fournit le $\overline{u}$ maximal. Une modification d'importance a ensuite été introduite par des chercheurs dits "BAYESIENS" (1), comme SAVAGE, DE FINETTI, RAIFFA, etc., qui ont substitué aux probabilités des probabilités subjectives. Depuis, ce critère est connu sous le nom de *subjective expected utility criterion (S.E.U.)* que nous traduisons par "critère de l'espérance (mathématique) subjective de l'utilité"(E.S.U.) :

$$\boxed{\text{E.S.U. (i)} = \sum_{j=1}^{n} u_{ij}\ ps_j}$$

Nos plans d'expériences ont été conçus pour que les élèves se conforment à ce seul critère. Mais il importe de vérifier si les résultats obtenus ne sont pas compatibles (2) avec d'autres critères. qui expliqueraient le comportement des sujets vis-à-vis des indices de certitude. Avant de répondre à cette question, nous examinerons les différents critères (ou stratégies) bien connus en théorie des décisions.

---

(1) RAIFFA (1970) proclame : "... j'appartiens à un parti minoritaire, les "bayésiens" (...) ou subjectivistes qui veulent introduire les jugements intuitifs et les impressions directement dans l'analyse formelle d'un problème de décision. Les non-bayésiens ou objectivistes pensent qu'il est préférable de tenir ces aspects subjectifs à l'écart de l'analyse formelle et ne doivent être utilisés (...) que pour combler le fossé existant entre le monde réel et les résultats objectifs obtenus en utilisant un modèle formel."

(2) Il est impossible de prouver que le recours à un critère a provoqué le comportement. Tout au plus, pouvons nous dire que cela est possible ou ne l'est pas. Dans ce dernier cas, on peut rejeter le critère comme explication du comportement.

l'intervalle n'est pas constant dans le négatif et l'est dans le positif.

E. STRATEGIES POSSIBLES EN SITUATION D'INCERTITUDE

Revenons à l'exemple où l'on demande au sujet, une fois sa réponse donnée (construite ou à choix multiple, peu importe), d'y adjoindre un des trois indices de certitude suivant :

A1 = certitude 1 = Je ne suis pas sûr de ma réponse.
A2 = certitude 2 = Ma réponse me semble bonne.
A3 = certitude 3 = Je suis tout à fait sûr de ma réponse.

Les conséquences ont été fixées arbitrairement par le professeur :

	Réponse correcte	Réponse incorrecte
Certitude 1	+ 1	0
Certitude 2	+ 2	- 2
Certitude 3	+ 3	- 3

Nous appliquerons les diverses stratégies à cet exemple. Examinons les critères qui donnent lieu à différentes stratégies.

1° Le critère du profit maximum (DE BRUYN, 1973) ou *maximum utility criterion* (COOMBS, DAWES and TVERSKY, 1971, p. 141) ou encore MAXIMAX.

Ce critère consiste à choisir la solution dont la valeur la plus élevée est maximale. Ce principe est optimal si l'événement le plus favorable se produit toujours. C'est donc un critère optimiste.

Dans notre exemple, ce critère conduit à choisir la certitude 3 qui permet d'obtenir un profit de +3 points.

2° Le critère de WALD ou *maximin utility criterion* (LUCE and RAIFFA, 1966) ou encore MAXIMIN.

C'est un critère pessimiste qui conduit à choisir la décision qui procure le profit le plus élevé dans le cas où les événements sont le plus hostiles. On s'efforce d'éliminer le risque de ruine, d'assurer une certaine sécurité en choisissant *the best worst state* ou, si l'on veut, le meilleur des sorts dans la situation la plus adverse. Ce critère porte aussi les noms de perte minimale ou "max des pertes" (mais cette dernière expression est ambiguë).

Dans notre exemple, il reviendrait à choisir la certitude 1 où la perte ne vaut que 0.

°

° °

COOMBS, DAWES et TVERSKY (1971, p. 141) font remarquer que :

> "Parce que les choses ne sont d'habitude ni aussi mauvaises qu'on les craint, ni aussi bonnes qu'on les souhaite, il est conseillable de pondérer le meilleur et le pire."

Cette recommandation est mise en application par les deux critères qui suivent.

3° Le critère de LAPLACE.

Ce critère, attribué à Jacob BERNOUILLI (1654-1705) est fréquemment appelé *Principle of insufficient reason*, principe de complète ignorance ou encore principe d'équiprobabilité.

Dans ce critère, on considère que les divers types d'événements (ou états), ici E1 et E2, ont autant de chances de se présenter. En conséquence, on choisit la décision à laquelle correspond la moyenne ou la somme) la plus élevée :

Certitude 1 : (1 + 0)/2 = 0,5
Certitude 2 : (2 - 2)/2 = 0
Certitude 3 : (3 - 3)/2 = 0

On choisirait donc la certitude 1.

4° Le critère de HURWICZ (1951) ou *pessimism-optimism index criterion.*

HURWICZ introduit une fonction de décision qui est une moyenne pondérée entre les gains minimum (E3) et maximum (E1) pour chaque décision. Ces gains sont pondérés par la probabilité que le sujet attribue à l'occurrence de ces événements en général. Il ne s'agit pas de la probabilité subjective que le sujet attribue à sa réponse, puisque la probabilité dont il s'agit ne change pas d'une question à l'autre.

Si, par exemple, le sujet a un caractère pessimiste, il attribuera une probabilité plus faible aux profits élevés (à la réussite, dans notre exemple).

Si ce poids est, par exemple, de 1/3 pour la réussite et 2/3 pour l'échec, on aura :

Certitude 1 : (0,33 x 1) + (0,66 x 0) = 0,33
Certitude 2 : (0,33 x 2) + (0,66 x -2) = 0,66 - 1,32 = -0,66
Certitude 3 : (0,33 x 3) + (0,66 x -3) = 0,99 - 1,98 = -1

On choisirait donc la solution 1.

La valeur attendue calculée par le critère de HURWICZ est, pour chaque action :

$$H(a_i) = p\ (E1) + (1-p)\ (E2)$$

Remarquons que si p = 1, le critère de HURWICZ se confond avec le critère maximax, et si p = 0, il se confond avec le critère maximin.

5° Le critère du "regret minimax" de SAVAGE (1951) ou *minimax risk criterion*, ou encore *minimax regret criterion*.

Ce critère est une amélioration du critère maximin, donc pessimiste puisque focalisé sur les événements défavorables. SAVAGE construit une nouvelle matrice appelée "matrice des regrets" dont chaque élément est l'écart par rapport au gain correspondant à la décision qui eût été la meilleure, si l'événement avait été connu.

Dans notre exemple, les profits maximum sont respectivement :

En cas de réponse correcte : +3
En cas de réponse incorrecte : 0

La matrice des regrets se présente comme suit :

	Réponse correcte	Réponse incorrecte
Profits maximum	+ 3	0
Regret avec certitude 1	- 2	0
avec certitude 2	- 1	- 2
avec certitude 3	0	- 3

On remarque que les regrets sont toujours exprimés par des nombres négatifs ou par 0 (le cas le plus favorable).

Pour chaque action possible (examen horizontal de la matrice), le regret maximal (le plus grand en valeur absolue) est souligné et la plus petite des trois valeurs soulignées est - 2, les certitudes 1 et 2 seraient donc choisies.

° ° °

REMARQUES.

Des modifications au critère de regret minimax, proposées par LUCE et RAIFFA (1966, p. 282), ne réfutent pas complètement trois objections faites par CHERNOFF (1954).

Les cinq critères qui viennent d'être décrits peuvent aussi être placés dans des systèmes axiomatiques comme ceux de CHERNOFF (1954) et MILNOR (1954). On trouvera une synthèse de ces travaux dans LUCE et RAIFFA (1966, p. 297 sq.).

COOMBS, DAWES et TVERSKY (1971) proposent un exemple numérique théorique (p. 142) où des actions différentes sont choisies en fonction des cinq critères :

		Etats			Critère de décision et l'action choisie :
		E1	E2	E3	
Actions	A1	5	5	5	maximin (WALD)
	A2	10	0	0	maximax
	A3	9	2	2	p = 1/2 (HURWICZ)
	A4	8	0	8	équiprobabilité (LAPLACE)
	A5	6	1	4	regret minimax (SAVAGE)

LA CONNAISSANCE PARTIELLE ET LA THEORIE MODERNE DE L'UTILITE

1) La théorie de l'utilité attendue de SAVAGE.

La théorie moderne de l'utilité a été développée pour la première fois par von Neumann et Morgenstern (1947). Une des conséquences de leurs axiomes nous importe particulièrement :

> "Si les préférences d'un individu satisfont les dits axiomes, alors son comportement peut être décrit, ou rationnalisé, comme la maximisation de l'utilité attendue."

L'espérance mathématique est l'utilité subjective attendue. L'approche (1) consiste à attribuer *a priori* des probabilités subjectives sur les événements possibles. On s'efforce donc de placer le problème dans un univers probabiliste. Le critère est de maximiser la moyenne des profits, pondérés par la probabilité que chacun se réalise.

Dans notre exemple, si le sujet attribue une probabilité de .7 à sa réponse d'être correcte (2), l'utilité attendue sera :

Certitude 1 : $(.7 \times 1) + (.3 \times 0) = .7$
Certitude 2 : $(.7 \times 2) + (.3 \times -2) = 1.4 - .6 = .8$
Certitude 3 : $(.7 \times 3) + (.3 \times -3) = 2.1 - 0{,}9 = 1.2$

Le présent critère, ou E.S.U. (3), se distingue fondamentalement du critère de HURWICZ. Ce dernier, en effet, attribue (par

---

(1) Cette approche est aussi connue sous le nom de "bayésienne", mais nous n'utiliserons pas ce terme, sur la recommandation du professeur Ch. HEUCHENNE, car il évoque trop le fameux théorème du révérend Thomas BAYES (1763).
Le professeur Lee CRONBACH regrettait, à un symposium à Montreux, en 1975, que le terme "bayésien" soit devenu de plus en plus polysémique. Il n'a pas été contredit sur ce point par les tenants de la statistique bayésienne (M. NOVICK, R. JACKSON, H. ROUANET).

(2) Et, par conséquent, .3 d'être incorrecte.

(3) En anglais : Subjective expected utility. Nous avons traduit par l'Espérance (mathématique) Subjective de l'Utilité.

optimisme ou pessimisme) un même coefficient aux états de la nature, quelle que soit la question. Le critère de l'E.S.U., par contre, attribue à chacun des états des probabilités différentes selon les questions. On comprend, dès lors, que l'E.S.U. soit le seul critère compatible avec l'utilisation d'indices de certitude dans l'évaluation pédagogique.

Comme SAVAGE (1964), DE FINETTI (1965), RAIFFA (1970) et bien d'autres, nous pensons qu'il est bien des situations où seul le critère de l'espérance subjective de l'utilité est compatible avec les comportements observés, contrairement aux autres critères. L'expérience que nous rapporterons au chapitre 2 en est un exemple.

Il reste néanmoins un point important à envisager : celui de la cohérence, chez un même sujet, de décisions distinctes.

2) La cohérence des décisions.

Il faut d'abord considérer deux axiomes présidant à l'évaluation subjective des probabilités :

1. Toutes les probabilités doivent être positives (ou nulles).
2. Leur somme doit valoir 1.

Si les probabilités d'un joueur contredisent l'un de ces deux axiomes, il sera amené à reviser ses probabilités, pour rester cohérent en probabilité.

Mais le problème de la cohérence est aussi logique, comme le montre un exemple adapté de L. SAVAGE (1954, p. 103) :

> "Un homme qui achète une voiture pour 113.500 FB est tenté de la commander avec la radio installée, ce qui amène le prix total à 121.500 FB; la différence semble négligeable. Mais, quand l'acheteur réfléchit au fait que, s'il avait déjà la voiture, il ne dépenserait certainement pas 8000 FB pour acquérir la radio, il comprend qu'il a fait une erreur."

> "Tandis que la réaction initiale peut fort bien refléter une loi de base de la psychologie de la perception et peut être une ligne de comportement dans la vie réelle pour beaucoup de gens, c'est la seconde réaction qui est une base saine de la théorie normative." (DE BRUYN, 1973, 9, p. 11)

LINDLEY (1971, p. 21) fait remarquer que, bien entendu, "nous sommes tous incohérents, mais nous corrigeons nos probabilités quand notre incohérence devient flagrante."

DE FINETTI (1965, p. 88), lui aussi, met en avant l'aspect normatif du concept de cohérence :

> "Considérer, comme nous le ferons, que les évaluations de probabilités doivent être cohérentes, (...) ce n'est pas faire une constatation empirique (qui serait fausse : en fait, personne n'est infaillible dans ses raisonnements avec la logique de la certitude ou de l'incertitude). "La cohérence est simplement l'évitement de décisions inadmissibles; c'est-à-dire celles qui mènent (...) à des résultats plus mauvais qu'une autre décision disponible" (WALD, 1955). Par conséquent, dans une étude sur la façon dont les gens <u>doivent</u> se comporter, nous supposerons qu'ils ne violent la cohérence ni en logique, ni en probabilité."

Une autre règle de la cohérence logique est la <u>transitivité</u>.

Si U1 $>$ U2 et que U2 $>$ U3, alors U1 $>$ U3

(où $>$ signifie "est préféré à").

LINDLEY (1971, pp. 20-21) fournit une jolie démonstration par l'absurde de cette propriété : si un sujet préfère l'utilité 2 à l'utilité 1 (U1 $<$ U2), l'utilité 3 à l'utilité 2 (U2 $<$ U3), l'utilité 1 à l'utilité 3 (U3 $<$ U1), c'est-à-dire si, pour lui, U1 $<$ U2 $<$ U3 $<$ U1, ce sujet sera prêt à payer une certaine somme (même minime) pour que l'on remplace U1 par U2, puis U2 par U3, puis U3 par U1. On se retrouve ainsi dans la situation initiale (cercle vicieux). Cette procédure ferait de ce sujet ce que Lindley appelle *"a perpetual money making machine"*. La position incohérente est intenable. S'il en allait autrement, on verrait souvent les gens jet~~t~~er littéralement l'argent par les fenêtres.

F. BRESSON (1965, p. 228) rapporte "un conte populaire montrant que la transitivité des préférences est requise même par une conscience naïve, et qu'il est banal, mais blâmable de s'en écarter".

La meilleure décision est celle qui consiste à choisir l'utilité attendue maximale. Ceci n'est pas évident, car bien d'autres stratégies de choix (ou critères de jugement des conséquences) peuvent être utilisées, comme nous l'avons vu plus haut.

°

° °

Quelle stratégie est utilisée, en fait, dans les situations pédagogiques ? C'est à cette question que nous nous efforcerons de répondre dans les chapitres 2 et 4.

DEUXIÈME PARTIE : L'APPROCHE PSYCHOLOGIQUE

A. PRISE DE RISQUE ET BESOIN D'ACCOMPLISSEMENT

En plus des stratégies inspirées de la théorie de la décision, il importe de considérer les stratégies qui seraient inspirées par une attitude liée à l'ampleur du risque. Il s'agit du même problème, mais il est simplement vu sous un autre angle. Nous venons de le voir avec les yeux des économistes, voyons-le maintenant avec les yeux des psychologues.

Déjà I. FISCHER (1906), puis M. ALLAIS (1953 a,b) et d'autres auteurs ont suggéré que la différence entre les conséquences possibles (1) peut être un critère plus important que l'utilité attendue ($\bar{u}$). Chaque sujet se situerait à un certain endroit de l'échelle des risques, selon des caractéristiques personnelles de motivation. Les recherches entreprises sur ce sujet doivent être examinées à la lumière des théories de McCLELLAND et, notamment, du "besoin d'accomplissement" (*Need of achievement* ou *Need ach*). Ce besoin est défini comme "un affect en relation avec une performance évaluée" (2), jaugée, susceptible de s'inscrire dans une échelle d'excellence (3). McCLELLAND mesure ce besoin au moyen d'un *T.A.T. – Need ach.*, aussi appelé *McCLELLAND Achievement Motivation Test (MAMT)*, où sont observées les fréquences d'apparition de thèmes liés au succès, à l'atteinte d'un but, au progrès, etc. (4).

(1) On appelle aussi "ampleur du risque" la différence entre la conséquence favorable (réussite) et la conséquence défavorable. Dans notre exemple, avec la certitude 1, la différence (entre +1 et 0) vaut 1. Avec la certitude 2, cette différence vaut 4 et, avec la certitude 3, elle vaut 6. Ces nombres représentent l'ampleur du risque lié à chaque action possible.

(2) "Affect in connection with evaluated performance" (McCLELLAND, 1953, p. 79).

(3) Th. PROGNEAUX (1974, p. 32).

(4) On lira les critères de correction dans Th. PROGNEAUX (1974), notamment dans les annexes.

La proportionnalité est très discutable.
La théorie d'ATKINSON est plus complexe et
sa mathématisation douteuse.

Dans le cadre des recherches sur le besoin d'accomplissement, J.W. ATKINSON (1964) a développé un modèle théorique, dans lequel toute situation de risque présenterait deux composantes : l'espoir du succès (*hope of success*) et la crainte de l'échec (*fear of failure*). Son hypothèse de base est que le sujet retire d'autant plus de satisfaction de l'accomplissement d'une tâche que celle-ci est difficile (1). La facilité d'une tâche est représentée par la probabilité de succès (P.S.) qui lui est associée (nous préférerions dire "par l'ESPER").

Ainsi, si P.S. vaut .90, la tâche est facile et si P.S. vaut .10, la tâche est difficile.

La deuxième hypothèse est que la valeur attractive (*incentive value of success, I.S)* est inversément complémentaire à 1 de PS.

$$\boxed{IS = 1 - PS}$$

Pour Atkinson, la tendance (2) à réussir dans une situation donnée (TS) peut s'écrire comme la résultante de variables de l'individu et de la situation :

$$\boxed{TS = Ms \times (PS \times IS)}.$$

où Ms = la force du besoin de réussir du sujet considéré.

On voit bien que (PS x IS) sera maximum pour des tâches de difficulté moyenne (p = q = .50 de sorte que pq = .25).

---

(1) Cette hypothèse est inspirée des travaux de S. ESCALONA (1940) et de L. FESTINGER (1942).

(2) Ou motivation, ou encore "besoin actualisé".
Tendance à réussir = Tendance to approach success.

Les conséquences ne sont pas si simples !

Les individus dont le besoin d'accomplissement est élevé devraient donc entreprendre des tâches de difficulté moyenne. Par contre, ceux qui ont une motivation élevée d'éviter l'échec devraient rechercher les tâches faciles.

Il est frappant de constater le parallélisme entre ce raisonnement psychologique et le raisonnement économique de la maximisation de l'utilité attendue. Il ne faut cependant pas perdre de vue qu'une situation de testing courante ne permet pas "d'entreprendre des tâches d'une difficulté donnée" mais impose ces tâches (les questions). Il est néanmoins des situations de testing (moins courantes, il est vrai) où l'étudiant peut choisir un certain nombre de questions pour arriver à un certain total (1). Dans ces situations particulières, on pourrait envisager de tester le modèle d'Atkinson. Nous n'avons pas entrepris une telle recherche car, on le verra par la suite, nous nous sommes tenus à la procédure de testing courante où chaque sujet doit répondre à chaque question.

cela a déjà été fait.

Les chercheurs ont été amenés à distinguer deux types d'influence systématique de la personnalité dans les situations de choix :

1. Les préférences d'un sujet pour certaines probabilités;
2. Les préférences d'un sujet pour certaines ampleurs de risques.

1. *PREFERENCES DE CERTAINES PROBABILITES* *(probabilities preference)*

ATKINSON (1964) cite (p. 210) à l'appui de son modèle les résultats acquis par W. EDWARDS (1953). Ce dernier propose deux situations ayant une valeur attendue ($\overline{V}$) égale.

- La situation A comporte 80 % de chances de gagner 10 F
  et 20 % de chances de ne rien gagner.
- La situation B comporte 50 % de chances de gagner 16 F
  et 50 % de chances de ne rien gagner.

---

(1) Ce système est appliqué par le service du professeur P. LASZLO, en première candidature à l'Université de Liège, dans des cours de chimie.

Dans les deux situations, la valeur attendue (8 F) est la même. Dans une telle situation, les sujets préfèrent surtout les risques intermédiaires (dans la figure ci-dessous, les pics se situent au rapport 50-50, ici la probabilité 4/8).

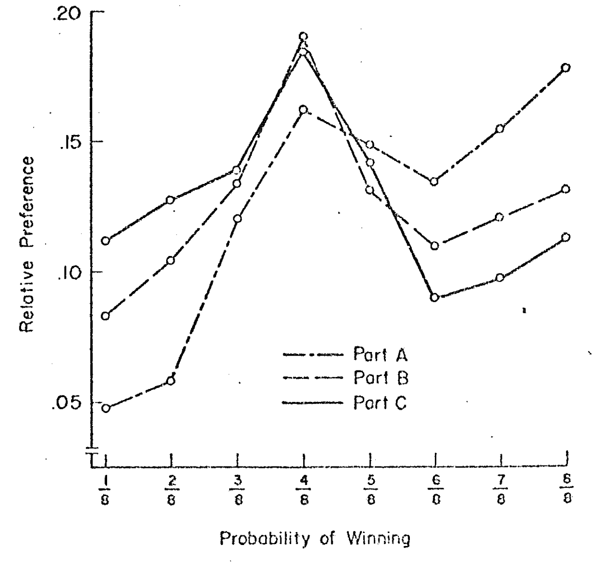


FIGURE 8.3 Preference as a function of probability for positive expected value bets (*From Edwards, "Probability-preferences in gambling,"* The American Journal of Psychology, *Vol. 66, 1953, p. 355*)

Cette constatation n'est valable que lorsque les valeurs attendues sont positives. EDWARDS a aussi fait l'expérience avec des valeurs attendues négatives (pertes). En voici un exemple.

- La solution A comporte 80 % de chances de perdre 10 F et 20 % de chances de gagner 20 F (la valeur attendue est - 3F).
- La solution B comporte 50 % de chances de perdre 20 F et 50 % de chances de gagner 14 F (la valeur attendue est - 3F).

La figure ci-dessous montre les préférences marquées lors de séries d'expériences où les valeurs attendues étaient négatives.

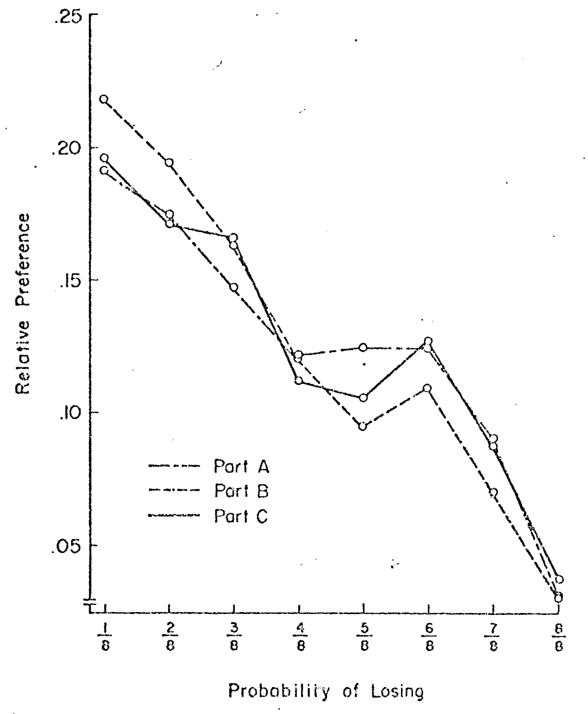


FIGURE 8.4 Preference as a function of probability for negative expected value bets (*From Edwards, "Probability-preferences in gambling,"* The American Journal of Psychology, *Vol. 66, 1953, p. 355)*

On voit que les sujets préfèrent les probabilités faibles de perdre des sommes importantes aux probabilités fortes de perdre des sommes minimes.

Th. PROGNEAUX (1974) cite, p. 55, une série d'autres expériences confirmant les hypothèses ci-dessus (1), mais elle signale que l'association entre un score Need Ach. élevé et une prise de risques calculés semble être spécifique aux situations où le sujet peut influencer ses résultats par des efforts personnels. Ainsi, dans un jeu de hasard (résultat indépendant de la compétence), les sujets dont le besoin d'accomplissement est élevé prennent généralement des risques bas (2).

(1) ATKINSON et LITWIN (1960), CLARK et al. (1956), SMITH (1963), ISAACSON (1964), MYERS (1965).
(2) LITTIG (1954), HANCOCK et TEEVAN (1964), RAYNOR et SMITH (1965), selon Th. PROGNEAUX (1974).

Th. PROGNEAUX commente ces données comme suit :

> "Dans un jeu de hasard, la compétence du sujet n'étant pas sollicitée, la valeur attractive du succès n'est plus liée à la difficulté de la tâche : elle est faible et constante pour toutes les valeurs de p (...) En gros, dans une telle situation, le "high achiever" semble se comporter comme s'il préférait être sûr de gagner, même peu."

Dans les expériences relatives à la certitude, ne peut-on pas s'attendre à ce que certains sujets réagissent comme il vient d'être décrit plus haut ? Bien entendu, préférer une probabilité donnée (par exemple intermédiaire) consiste à ignorer l'ESPER et un tel comportement serait très défavorable à l'utilisation des indices de certitude. Les "préférences de probabilités" aboutissent à des critères (maximax, maximin, et surtout celui de HURWICZ) déjà examinés dans les théories économiques.

Les "préférences de variance" qui vont être décrites maintenant débouchent sur des *patterns* comportementaux entièrement originaux, dont la théorie du dépliage (C.H. COOMBS) doit rendre compte.

2. *PREFERENCES DE CERTAINES AMPLEURS DE RISQUES (risk variance preference)*

COOMBS et PRUITT (1960) ont montré que les sujets ont aussi des préférences de variance, lorsque les probabilités de gain et les valeurs attendues sont égales. En voici un exemple :

Loterie A	Loterie B
1/6 de chances de gagner 180 F	1/6 de chances de gagner 360 F
5/6 de chances de perdre 36 F	5/6 de chances de perdre 72 F
La valeur attendue est 0.	La valeur attendue est 0.

Cependant, la variance diffère :

Loterie A : Variance = $(180-36)^2.(1/6).(5/6) = 16416 \times .138 = 2279$

Loterie B : Variance = $(360-72)^2.(1/6).(5/6) = 82944 \times .138 = 601.043$

Pour C.H. COOMBS, chaque sujet se situerait à un certain endroit d'une "échelle des variances de risques". Si les données justifiaient une telle explication, il faudrait renoncer à utiliser l'indice de certitude pour estimer les ESPER des sujets.

La méthode la plus élaborée permettant de tester l'hypothèse des "préférences d'ampleur de risque" est incontestablement la théorie du dépliage (*unfolding theory*) de C.H. COOMBS (1950).

Exposée par son auteur, puis reprise dans des traités (1) ou dans des discussions (2), elle a été peu décrite en français. Afin d'en faciliter la compréhension, nous l'expliquerons à partir de notre exemple.

Outre les problèmes de décision, la théorie du dépliage intéresse la psychophysique et les mesures des attitudes (cf. les *scaling theories*).

---

(1) COOMBS, DAWES et TVERSKY, 1970, P. 137; COOMBS C.H., 1967.
(2) GREENBERG G., 1963.

B. LA THEORIE DU DEPLIAGE DE C.H. COOMBS

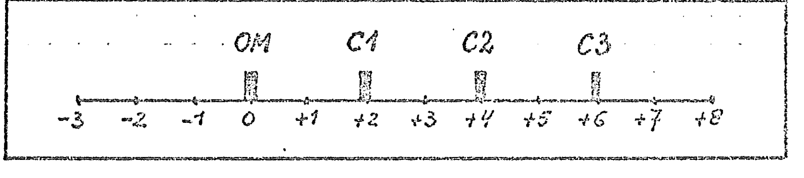
1. *L'ECHELLE DES STIMULI*.

Considérons que les ampleurs des risques inhérents aux quatre actions possibles (omission, certitude 1, certitude 2 et certitude 3) constituent des stimuli pour le sujet. Ces stimuli représentent respectivement 0 point, 2 points, 4 points et 6 points.

	B.R.	M.R.	Ampleur du risque
Omission	0	0	0
Certitude 1	+1	-1	2
Certitude 2	+2	-2	4
Certitude 3	+3	-3	6

Ces stimuli peuvent être portés sur une échelle (tableau 1.6).

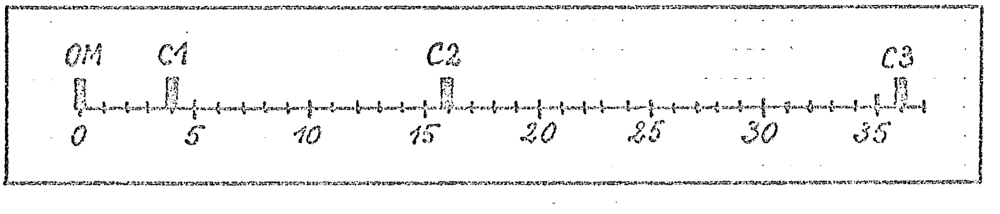
Tableau 1.6



Remarquons qu'on pourrait traduire l'ampleur du risque non par ces nombres, mais par leur logarithme, leur carré, etc., ce qui donnerait, évidemment, une échelle différente.

Ainsi, si l'on considère les carrés des risques courus (variances des risques), on obtient l'échelle présentée dans le tableau 1.7.

Tableau 1.7



C'est à partir de cette dernière échelle (1) que nous expliquerons la théorie du dépliage.

2. *LE POINT IDEAL PERSONNEL.*

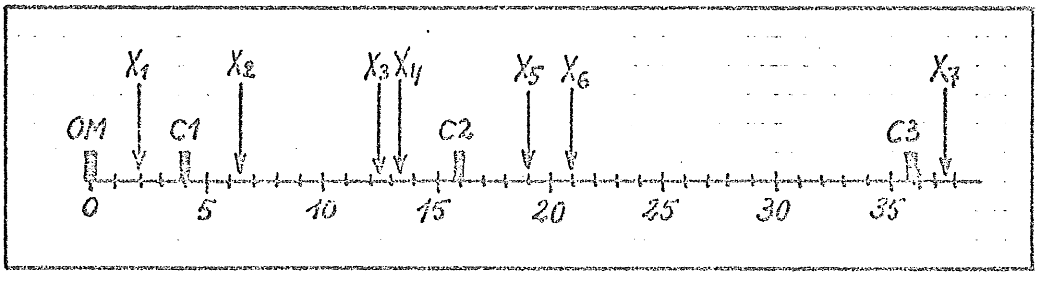
L'échelle des risques est la même pour tous les étudiants; elle est objective. COOMBS part du principe qu'il existe pour chaque étudiant un point idéal sur cette échelle, son "degré idéal du risque".

Les étudiants les plus timorés ne prennent qu'un risque faible et se situent sur la gauche de l'échelle, tandis que les plus audacieux se situent sur la droite. Remarquons qu'un point idéal situé à gauche de 0 n'a pas de sens dans le cas qui nous occupe. Par contre, un point idéal pourrait être situé à droite de 36.

3. *L'ECHELLE JOINTE (J Scale).*

Si l'on connaissait (2) le point idéal de plusieurs étudiants, on pourrait porter ces points idéaux (3) sur une seule échelle. Le tableau 1.8 présente un exemple (imaginaire) pour sept sujets :

Tableau 1.8



Une telle échelle est appelée échelle jointe (*J scale*) parce qu'on y porte à la fois la position des stimuli et la position (idéale) des étudiants.

(1) En raison de son asymétrie, elle permet de mieux expliquer la théorie du dépliage. On raisonne comme si cette échelle était effectivement à la base des choix, ce qui n'est aucunement établi a priori.

(2) Ce qui n'est évidemment pas le cas en pratique.

(3) Avec le signe conventionnel X.

4. *L'ORDRE INDIVIDUEL DES PREFERENCES.*

Considérons l'étudiant 5 dont le point idéal (X5) est 19. S'il doit choisir entre deux stimuli (ici deux des actions possibles), cet étudiant optera pour le stimulus le plus rapproché de son point idéal.

Ainsi, s'il doit choisir entre C1 et C2, il choisira C2;
entre C1 et OM, il choisira C1;
entre C2 et C3, il choisira C2;
entre OM et C3, il choisira C3.

Si l'on fait toutes les comparaisons (deux à deux) possibles, on constate que l'ordre de préférence de l'étudiant est :

C2 C1 C3 OM	(1)

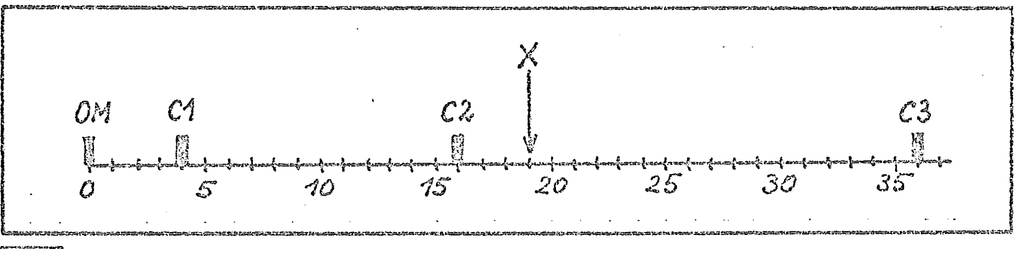
Cet ordre de préférence caractérise un étudiant déterminé. L'étudiant 6 dont le point idéal (X6) est 21 réagirait dans l'ordre suivant : C2 C3 C1 OM.

5. *LE PLIAGE D'UNE ECHELLE JOINTE (J Scale).*

La théorie du dépliage de COOMBS tient son nom d'une analogie mécanique : l'ordre de préférence apparaît lorsqu'on plie l'échelle jointe au point idéal du sujet.

Par exemple, pour le sujet dont le point idéal est 19, on obtient l'échelle jointe du tableau 1.9 (où X représente le point idéal).

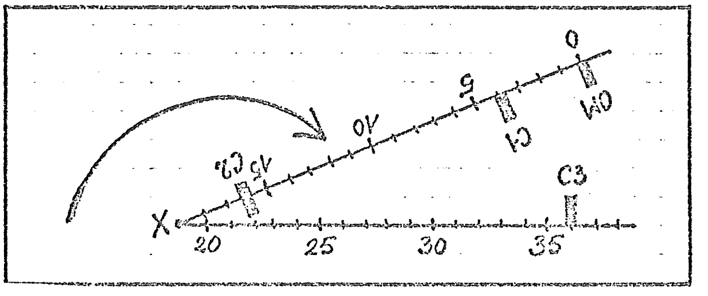
Tableau 1.9



(1) L'ordre se lit de gauche à droite, du stimulus le plus proche du point idéal (ici C2) au stimulus le plus éloigné (ici OM).

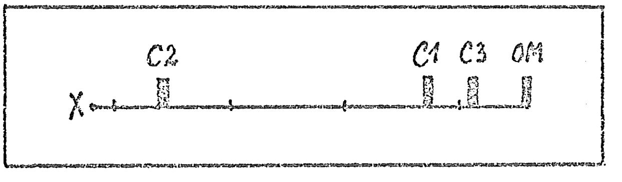
On la plie (1) avec, pour charnière, le point X, comme le montre le tableau 1.10.

Tableau 1.10



Le tableau 1.11 représente l'échelle individuelle (*I Scale*) obtenue.

Tableau 1.11



Cette échelle rend directement compte de l'ordre des préférences : C2 C1 C3 OM, et est strictement individuelle.

6. LE DEPLIAGE D'UNE ECHELLE INDIVIDUELLE (*I Scale*).

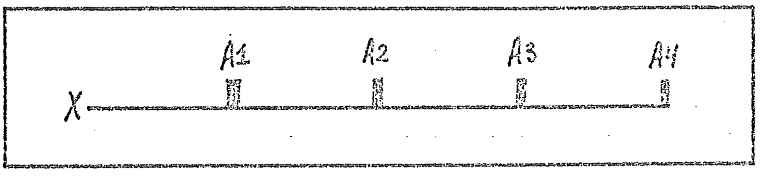
Dans la réalité, on ne connaît pas le "point idéal" d'un sujet (qui serait d'ailleurs bien en peine de le situer si on le lui demandait). On peut cependant obtenir du sujet qu'il réponde à une série de comparaisons deux à deux. Par exemple, dans une expérience où quatre actions sont disponibles avec les ampleurs A1, A2, A3 et A4, le sujet devra préciser s'il préfère A3 à A1, A2 à A1, etc.

(1) COOMBS (1967, p. 80) utilise une représentation légèrement différente.

Par une série de questions semblables, il n'est pas difficile de construire, pour chaque sujet, une échelle individuelle (*I scale*), avec cette importante restriction qu'elle ne sera pas métrique, mais simplement ordinale.

Il sera possible d'établir, pour un sujet donné, son ordre ou échelle des préférences, par exemple celle qui est présentée dans le tableau 1.12.

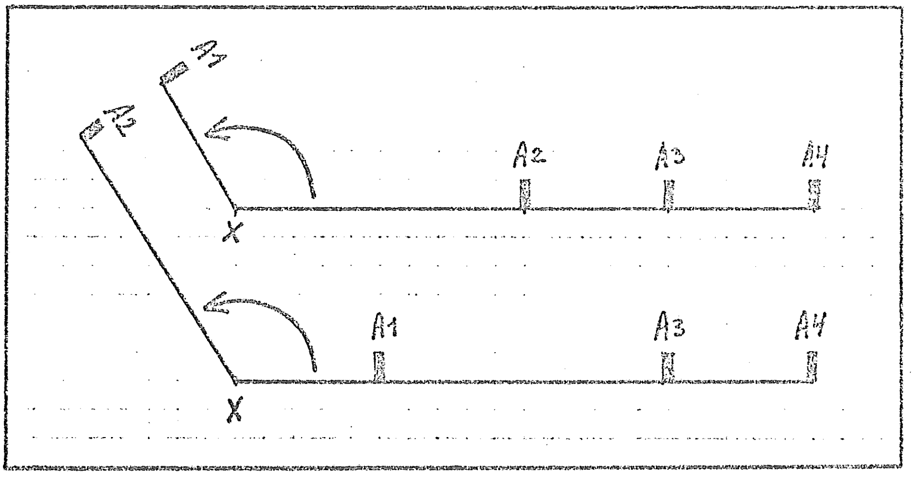
Tableau 1.12



Rappelons que, les distances étant inconnues, nous les avons toutes représentées (arbitrairement) égales. On écrira, conventionnellement A1 A2 A3 A4, en considérant que le stimulus de gauche est le plus proche du point idéal.

De quelle échelle jointe cette échelle individuelle peut-elle découler ? Autrement dit, y a-t-il une façon de déplier cette échelle individuelle pour obtenir une échelle jointe ? En fait, il y a de nombreuses façons de déplier, selon l'endroit où l'on fixe le point idéal. Le tableau 1.13 contient deux façons différentes de déplier l'échelle I du tableau 1.12.

Tableau 1.13

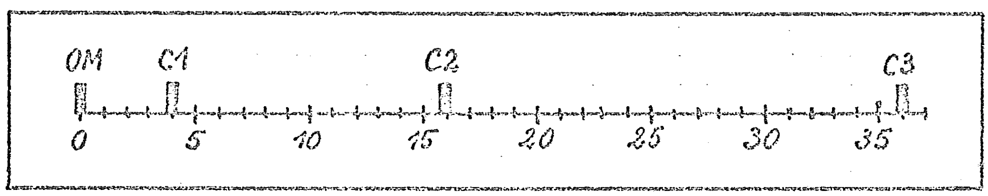


7. *LA LISTE DES ECHELLES INDIVIDUELLES ISSUES D'UNE MEME ECHELLE JOINTE.*

Dans la mesure des attitudes, il n'est pas rare que l'on doive examiner toutes les façons possibles de déplier l'échelle individuelle et il y a d'autant plus de solutions possibles que le nombre de stimuli sur l'échelle est grand (voir COOMBS, 1965, p. 457).

Le cas qui nous occupe est heureusement plus simple, puisque nous connaissons l'échelle jointe de départ (tableau 1.7).

Tableau 1.7



Etant donné le petit nombre de stimuli (4) sur cette échelle jointe, il est assez facile de faire l'inventaire de toutes les échelles individuelles qui pourraient en provenir. Il suffit pour cela d'envisager toutes les positions possibles du point idéal (X) - de gauche à droite - sur l'échelle jointe.

Envisageons deux positions du point idéal représentées dans les tableaux 1.14 et 1.15.

Tableau 1.14

cas_1

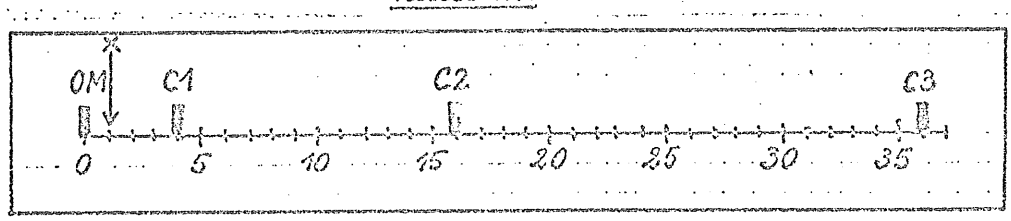
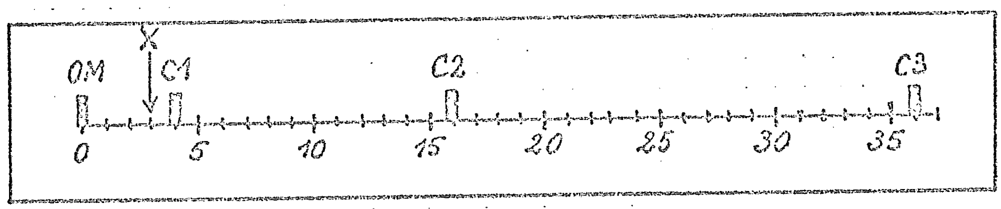


Tableau 1.15

cas_2



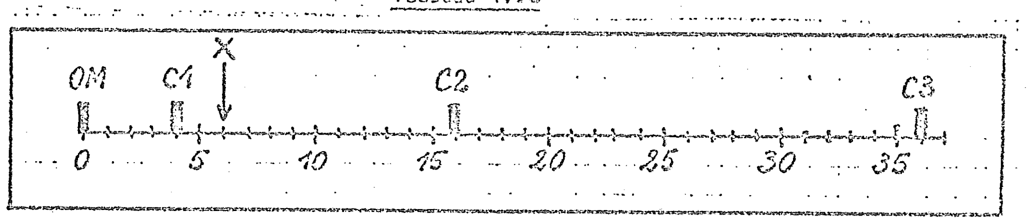
Lorsque le point idéal (X) se trouve exactement entre OM et C1, l'ordre des préférences est : OM=C1 C2 C3 (cas 3).

Nous appellerons MO1 le point central (*midpoint*) entre OM et C1. A gauche de MO1, on retrouve le cas 1 ci-dessus. A droite de MO1, on rencontre le cas 2. MO1 correspond au cas 3.

On voit que c'est le passage du point idéal X d'une zone à une autre qui provoque la modification dans l'ordre des préférences.

Tableau 1.16

cas 3



Ordre des préférences = C1 OM C2 C3.

Ainsi, pour la position de X dans le tableau 1.16, l'ordre des préférences est le même que pour le cas 2, alors que X est situé de l'autre côté de C1. Ce stimulus (C1) n'est donc pas une limite de zone. Par contre, la mi-distance (MO1), elle, constitue une telle limite de zone.

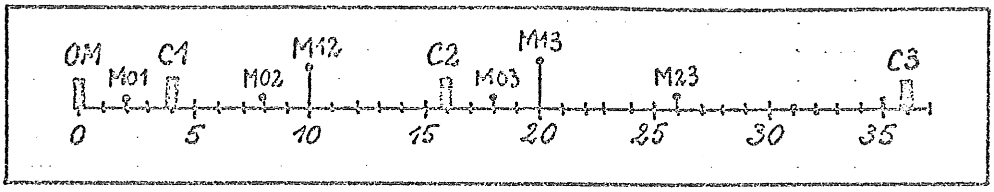
8. LES MI-DISTANCES, LIMITES DE ZONES.

Sur notre échelle jointe, portons :

MO1 = point à mi-distance entre OM et C1
MO2 = point à mi-distance entre OM et C2
MO3 = point à mi-distance entre OM et C3
M12 = point à mi-distance entre C1 et C2
M13 = point à mi-distance entre C1 et C3
M23 = point à mi-distance entre C2 et C3.

On obtient le tableau 1.17.

Tableau 1.17



Ce sont ces points à mi-distance, et non les stimuli eux-mêmes qui constituent les limites de zone. Il est maintenant possible de décrire exhaustivement la liste des différentes échelles individuelles découlant des positions de X dans les différentes zones. Ces correspondances entre zones et ordre de préférence sont synthétisées dans le tableau 1.18.

Tableau 1.18

Positions de X (1) sur l'échelle jointe	Ordre de préférence (échelle individuelle)	N°
Si X < M01	OM C1 C2 C3	1
Si X = M01	OM=C1 C2 C3	2
Si M01 < X < M02	C1 OM=C2 C3	3
Si X = M02	C1 OM C2 C3	4
Si M02 < X < M12	C1 C2 OM C3	5
Si X = M12	C1=C2 OM C3	6
Si M12 < X < M03	C2 C1 OM C3	7
Si X = M03	C2 C1 OM=C3	8
Si M03 < X < M13	C2 C1 C3 OM	9
Si X = M13	C2 C1=C3 OM	10
Si M13 < X < M23	C2 C3 C1 OM	11
Si X = M23	C2=C3 C1 OM	12
Si X > M23	C3 C2 C1 OM	13

(1) Le signe < signifie ici "situé en deçà de" ou "plus petit que";
le signe > signifie ici "situé au-delà de" ou "plus grand que".

Il existe donc 13 (et seulement 13) échelles individuelles possibles à partir de notre échelle jointe. Inversément, n'importe lequel des 13 ordres de préférence indique deux choses :

1° Cet ordre des préférences pourrait être issu de notre échelle jointe;

2° Le point idéal (X) peut être situé dans une zone précise sur l'échelle jointe.

Dans les chapitres 2 et 4, les résultats expérimentaux seront confrontés à la théorie du dépliage.

Il existe au moins un autre modèle relatif au risque : celui de SLOVIC et LICHTENSTEIN (1968). Nous venons d'en apprendre l'existence et nous n'avons pas pu le confronter aux résultats.

CONCLUSIONS

Dans les chapitres 2 et 4, les réponses des élèves seront systématiquement confrontées avec ces modèles théoriques décrivant des stratégies qui, toutes, à l'exception du critère E.S.U., sont contraires à une étude valide de la certitude.

Nous formulons l'espoir que le critère E.S.U. s'avérera plus plausible que les autres, puisqu'il est, selon notre thèse, le responsable du comportement. Malheureusement, nous ne pouvons le démontrer, car, comme le dit H.M. BLALOCK (1971, p. 21) :

> "Il y aura toujours un certain nombre d'autres modèles qui donneront les mêmes prédictions (...) que le modèle étudié. Nous pouvons seulement procéder par <u>élimination</u> des modèles inadéquats qui donnent des prédictions incorrectes, car il est d'habitude impossible d'éliminer toutes les alternatives plausibles sur la base des données disponibles.
>
> Dans ce sens, on ne peut jamais "établir" aucun modèle causal."

CHAPITRE 2

UNE EXPERIENCE CONTINUE, AVEC CERTITUDES ORDINALES ET CONSEQUENCES EMPIRIQUES

LE PROBLEME DE LA VALIDITE DE L'ESPER

1. PRESENTATION DE L'EXPERIENCE

A. *LE PROBLEME*.

Dans l'introduction, plusieurs études utilisant une matrice de conséquences associées à un ou des indices de certitude ont été décrites. L'expérience qui suit (ainsi que celle qui fera l'objet du chapitre 4) se distingue des recherches précédentes par trois aspects importants :

1.- Dans les expériences antérieures, les scores avec certitude ne sont pas utilisés pour gratifier ou pénaliser les sujets (1). Ceux-ci, dès lors, pouvaient réagir au gré de leur fantaisie. Nous ne pensons pas que ce fait invalide profondément les résultats déjà obtenus. Néanmoins, dans les deux expériences qui suivent, les sujets seront mis dans une situation operante (ils seront exposés aux conséquences de leurs actes). Cette procédure est conforme à la recommandation de B. DE FINETTI (1965, p. 83) : "Le sujet doit avoir été intéressé dans l'obtention d'un score total élevé et dans la maximisation de ce score."

2.- Dans les expériences antérieures, la matrice des conséquences n'est jamais systématiquement confrontée à la théorie des décisions, et tout spécialement au critère de l'E.S.U. (2). Nos matrices de conséquences seront analysées afin de décrire les stratégies possibles (voir chapitre 1) auxquelles les réponses de chaque élève à chaque test seront confrontées. Evaluer le plus exactement possible les PER est une stratégie parmi les autres, mais c'est, bien sûr, celle que nous souhaitons voir appliquer par les élèves.

---

(1) Sauf dans l'expérience de VAN NAERSSEN et al. (1966).

(2) Espérance subjective de l'utilité développée par la théorie moderne de l'utilité (VON NEUMANN et MORGENSTERN, 1947).

3.- Les expériences antérieures se présentent sous la forme d'une séance unique de testing. Cette procédure ne permet pas la mise en évidence d'une stabilité ou d'une évolution éventuelles du comportement des sujets. De nombreux chercheurs sont cependant très conscients du problème.

> "... La cotation avec un indice de certitude rend la note plus fidèle. On constate également qu'il existe un lien positif entre l'exactitude des réponses et le sentiment de certitude des sujets. Cette corrélation peut sans doute être augmentée par un entraînement préalable avec feedback avant l'administration du test même... On suppose que l'entraînement diminuera l'influence de facteurs d'attitude." (VAN NAERSSEN et VAN BEAUMONT, 1965, p. 314).

Dans les expériences qui vont être relatées, plusieurs classes d'étudiants ont été suivies pendant toute une année scolaire sur une même matière (la mécanique). La première expérience s'est déroulée pendant l'année scolaire 1971-72 et la seconde pendant l'année scolaire 1972-73, à l'Ecole Technique de la Force Aérienne belge (ETFA) située à Saffraanberg.

Des conséquences réelles (positives et négatives) ont été appliquées (1) immédiatement (délai maximum = 5 jours) à des étudiants selon leur score total avec certitude. Répéter le processus à intervalles réguliers (de 15 en 15 jours) n'a été possible qu'avec des moyens pédagogiques et informatiques importants.

Il a fallu développer un système de banque de questions géré par ordinateur, prendre des décisions docimologiques, concevoir et mettre au point les programmes FORTRAN. Il a, en outre, été nécessaire de former les enseignants (par des petits manuels illustrés) à la création d'objectifs et de questions et à l'exploitation des résultats. Enfin, il a fallu élaborer un système de fonctionnement de l'ensemble.

---

(1) Cette procédure était déjà en vigueur avant notre entreprise. Les scores simples servaient alors de base aux conséquences.

Ces aspects du travail ont fait l'objet de publications (1). Ils n'ont pu être réalisés que grâce à la collaboration de nombreux professeurs de l'ETFA et membres de la section "Contrôle des Etudes" créée par le Col. Moonen. Le Commandant P. Van Roy, qui nous a beaucoup aidé dans la mise sur pied de l'entreprise et qui la dirige depuis lors, nous a permis de récolter les données des chapitres 2 et 4. Son apport a été particulièrement précieux.

Rappelons que, selon notre thèse générale (voir introduction), les conditions optimales d'utilisation et d'étude des ESPER (Evaluation subjective de la probabilité d'exactitude de la réponse) sont au nombre de quatre :

1. Une situation operante.
2. Des sujets entraînés.
3. Une matrice des conséquences calculées selon le critère E.S.U.
4. Des certitudes exprimées, par la consigne, dans une échelle de probabilités.

Seules les deux premières conditions seront réunies dans l'expérience de 1971-1972 car, à l'époque, l'importance des deux dernières ne nous apparaissait pas. Ce sont les résultats expérimentaux de cette première étude longitudinale qui nous ont permis de comprendre l'impact des valeurs numériques des conséquences sur les décisions.

Nous avions, dans le cadre de la théorie des décisions, construit involontairement une contre-expérience dont les résultats, à la lumière des hypothèses actuelles, s'avèrent plus éclairants encore que l'expérience elle-même. En effet, avec une matrice des conséquences non conformes au critère E.S.U., les comportements évoluent, se modifient, ce qui n'est pas le cas avec une matrice orthodoxe.

expliquer ?

(1) LECLERCQ, 1971, 1973, 1975.
Voir aussi l'annexe 2.

B. *PRECISIONS TERMINOLOGIQUES.*

Avant d'entrer dans le détail des hypothèses et des résultats, il importe de préciser et de justifier les notations dont nous ferons usage. Les noms de variables et d'indices ont été choisis pour être aisément utilisés, le cas échéant, dans un programme FORTRAN. Il arrivera cependant, pour abréger les symboles au maximum, que des lettres grecques (deux seulement) soient utilisées en lieu et place d'une variable lourdement indicée. La correspondance, alors, sera chaque fois faite avec précision.

INDICES

Indice i = individu (ni = nombre d'individus)

Indice q = question (nq = nombre de questions)

Indice t = test (nt = nombre de tests)

Indice m = moment (plusieurs applications d'un même test) (nm = nombre de moments)

Indice r = réponse. Dans nos expériences (sauf au chapitre 7), l'individu ne peut donner qu'une seule réponse par question. Pour cette raison, l'indice r est fréquemment omis. (nr = nombre de réponses)

VARIABLES

ER_{riqtm} = exactitude de la réponse r émise par l'individu i à la question q du test t au moment m.

C'est une matrice à cinq dimensions contenant des 1 (exact) et des 0 (inexact).

PER_{riqtm} = probabilité d'exactitude de la réponse... (mêmes indices que ER).

C'est une matrice à cinq dimensions contenant des nombres réels compris entre 0 et 1 (inclus).

$ESPER_{riqtm}$ = évaluation subjective de la PER (mêmes indices et même type de variables).

NB.: La PER et l'ESPER diffèrent peu, car toute probabilité est subjective, ne serait-ce que par l'hypothèse communément acceptée que "les fréquences observées dans le passé se reproduiront dans une certaine tranche du futur" (A. KAUFMANN).

C_{riqtm} = indice de certitude. On préférera cette notation à ESPER lorsque la PER est codée dans une échelle de certitude à nc catégories. Dans ce dernier cas, la matrice C a cinq dimensions et contient des entiers compris entre 1 et nc inclus.

C_{criqtm} = une autre façon de noter la certitude utilisée à la réponse r, par le sujet i, à la question q...

C'est alors une matrice à six dimensions (c va de 1 à nc) contenant des 0 (non utilisé) ou des 1 (utilisé).

SIGNES D'OPERATIONS.

Le signe $\sum_{i=1}^{ni}$ désigne la somme sur l'indice i des valeurs de la variable qu'il précède.

Le signe $\mathcal{M}|_i$ désigne la moyenne observée pour chacun des i individus des valeurs (sommées sur tous les autres indices) de la variable qu'il précède (1). Ainsi, $\mathcal{M}|_{ij}$ implique que (ixj) moyennes seront calculées.

Le signe $\mathcal{E}|_i$ désigne l'espérance mathématique conditionnelle à i des valeurs de la variable qu'il précède.

EXPRESSIONS PARTICULIERES.

L'expression $TUC|_{ct}$ désigne le taux d'utilisation conditionnel à t de chaque degré de certitude c, par rapport à l'ensemble des degrés de certitude. Ici, $TUC|_{ct}$ est une matrice à deux dimensions contenant des réels compris soit entre 0 et 1 (inclus), soit entre 0 et 100 (inclus).

On emploira aussi la lettre grecque τ (tau) pour désigner cette variable.

L'expression $\mathcal{M}|_i(ER|C=c)$ désigne la Moyenne, conditionnelle à i, des exactitudes des réponses lorsque la certitude associée à cette réponse vaut c. On emploira aussi la lettre grecque μ (mu) pour désigner cette variable.

L'expression $\mathcal{M}|_i(C|ER=1)$ désigne la Moyenne, conditionnelle à i, des certitudes utilisées lorsque la réponse est correcte.

Les trois expressions ci-dessus résultent, on le voit, du traitement de sous-ensembles des matrices décrites plus haut.

(1) Cette notation (moyennes conditionnelles) est inspirée des cours du professeur BRENY.

C. *LA PROCEDURE DE RENFORCEMENT DE L'EXPERIENCE 1971-1972.*

Environ deux cents étudiants, répartis en plusieurs classes, ont été soumis, de septembre 1971 à juin 1972, à une trentaine de tests successifs (de 25 questions en moyenne) portant sur des matières scientifiques et techniques.

En plus de leur réponse, les étudiants doivent indiquer l'un des trois indices de certitude suivants :
- la certitude 1 (Je ne suis pas sûr de ma réponse)
- la certitude 2 (Ma réponse me semble bonne)
- la certitude 3 (Je suis très sûr de ma réponse).

L'omission est conseillée en cas d'ignorance totale.

On remarquera que la consigne présente une échelle ordinale. Elle revient à dire aux étudiants : "Sur votre échelle de certitude, considérez que OM < certitude 1 < certitude 2 < certitude 3." On est loin, évidemment, d'une échelle d'intervalles.

La matrice des conséquences (en points) est la suivante :

Tableau 2.1

	E1 (réponse correcte)	E2 (réponse incorrecte)
Omission	0	0
Certitude 1	+ 1	- 1
Certitude 2	+ 2	- 2
Certitude 3	+ 3	- 3

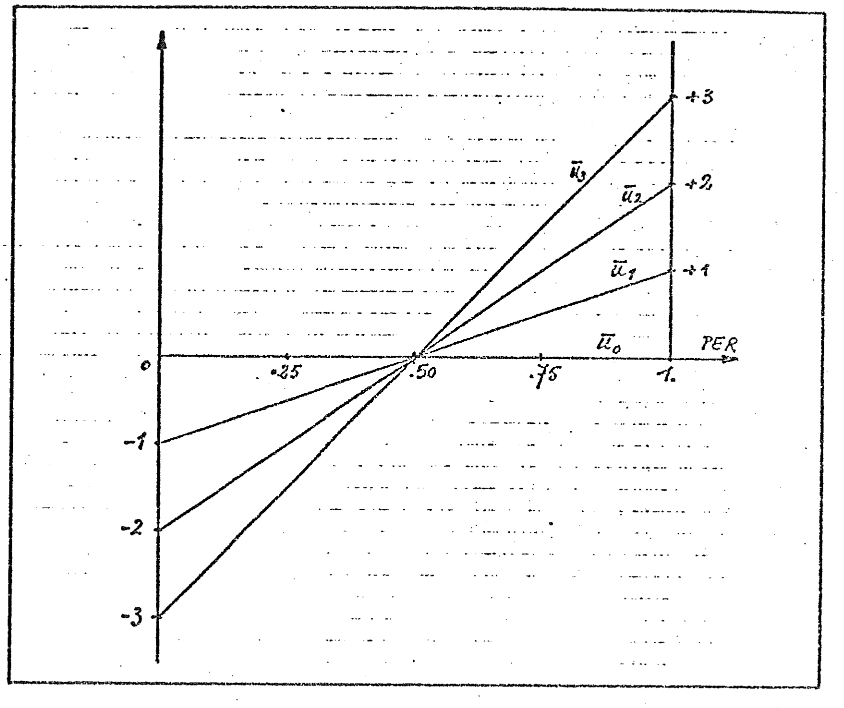
Matrice des conséquences utilisée durant l'année scolaire 1971-1972.

Ces valeurs ont été choisies pour leur simplicité mnémotechnique et pour faciliter les calculs, en conformité avec l'opinion de R.M. RIPPEY (1970) : "La plupart des scores calculés en utilisant des fonctions ésotériques présenteront une composante d'erreur due au manque

de compréhension du système de cotation par le sujet." (p. 169)
La recommandation de RIPPEY d'utiliser des "fonctions de cotation transparentes pour l'intuition des sujets" (p. 166) a été appliquée pour les conséquences elles-mêmes.

Le programme de renforcement mis en oeuvre consiste donc à accorder ou retirer les points selon quatre fonctions d'utilités attendues ($\bar{u}$) représentées dans le graphique qui suit :

Tableau 2.2



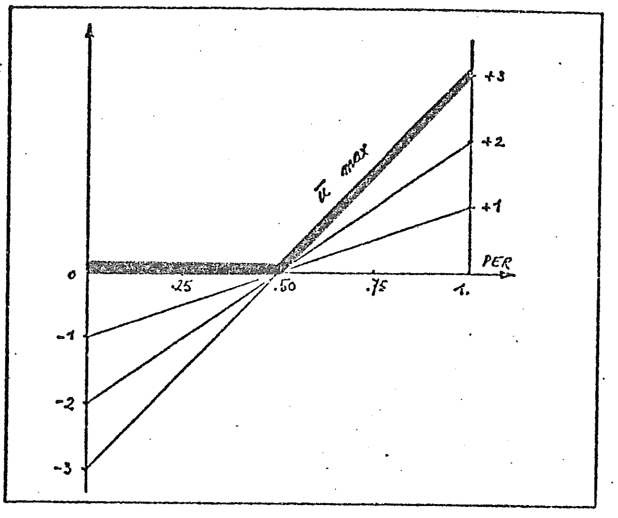
Fonctions d'utilité attendue ($\bar{u}$) pour chacun des quatre indices de certitude

Les conséquences (en points) sont portées en ordonnées et la PER (probabilité d'exactitude de la réponse) en abscisse. Les obliques représentent les $\bar{u}$ (utilités attendues) pour chaque action (degré de certitude choisi).

La <u>fonction de maximisation</u> se trouve représentée en traits gras dans le tableau 2.3.

Ceci revient à dire "réponds juste ou abstiens toi"

Tableau 2.3



La fonction de maximisation des utilités attendues à partir de la matrice de conséquences du tableau 2.1

La technique de représentation ci-dessus, particulièrement parlante, a été conçue par P. VAN ROY et P. DE SCHUTTER, dont l'aide nous a été très efficace dans le cadre de la section "contrôle des études de l'Ecole Technique de la Force Aérienne".

Ces fonctions de maximisation peuvent être décrites schématiquement par la règle suivante : "Si tu es certain à plus de 50 %, réponds avec la certitude 3. Sinon, abstiens-toi." Suivre ces recommandations, c'est respecter le principe de l'E.S.U., c'est-à-dire maximiser l'espérance subjective de l'utilité par l'emploi des certitudes.

Une façon beaucoup plus classique de décrire le même phénomène consiste à calculer la matrice des utilités attendues. Le tableau 2.4 présente, pour quatre valeurs de PER (.125, .375, .625 et .875), les utilités attendues pour chacune des actions possibles.

Tableau 2.4

Réponse avec certitude	Valeur de la PER			
	.125	.375	.625	.875
0	0	0	0	0
1	-.75	-.25	+.25	+.75
2	-1.45	-.50	+.50	+1.45
3	-2.25	-.75	+.75	+2.25
Action optimale	omission	omission	certit. 3	certit. 3

Matrice des utilités attendues pour quatre valeurs de PER, à partir de la matrice des conséquences du tableau 2.1

Le tableau 2.4 révèle la difficulté qu'il y aurait à se faire une représentation d'ensemble des utilités maximales et des actions optimales qui en découlent. Or les quatre valeurs des PER sont très distantes l'une de l'autre. La confusion entre stratégies est d'autant plus grande que la difficulté de l'item se rapproche de .50. On comprendra que, même si les étudiants utilisent le critère de l'E.S.U., les certitudes 1 et 2 ne disparaîtront pas totalement.

Le total des points détermine l'obtention de permissions de sorties anticipées et le passage dans une classe supérieure en fin d'année. L'utilisation de ces renforcements, immédiats et lointains, a évidemment pour but de répondre à l'exigence formulée par DE FINETTI, c'est-à-dire appliquer une procédure operante à l'expérience.

Il ne s'agit pas d'une certitude forte mais moyenne

D. LES HYPOTHESES.

Quand les quatre conditions optimales sont réunies, même d'une façon imparfaite, on peut vérifier si les sujets utilisent leur certitude de façon cohérente. Il s'agit en fait d'étudier la validité de l'ESPER .

On distinguera deux types de cohérence :

- La cohérence dans le taux d'utilisation de chaque degré de certitude (cohérence en τ);
- La cohérence dans l'exactitude des réponses fournies avec les divers degrés de certitude (cohérence en μ).

Ces deux types de cohérence donnent lieu à plusieurs hypothèses précises dans lesquelles on distinguera les certitudes fortes (ici 2 et 3) et les certitudes faibles (ici 0 et 1).

COHERENCES DANS LE TAUX D'UTILISATION DE CHAQUE DEGRE DE CERTITUDE (cohérences en τ)

1. La cohérence globale en taux d'utilisation.

H1. Le taux d'utilisation de certitudes maximales (ici certitude 3) est plus élevé et le TUC minimales (ici certitude 1) est plus faible pour les réponses correctes (ER = 1) que pour les réponses incorrectes (ER = 0).

$(TUC_3 | ER = 1) > (TUC_3 | ER = 0)$

$(TUC_1 | ER = 1) < (TUC_1 | ER = 0)$

2. La cohérence entre individus en fréquence d'utilisation.

Les individus plus compétents que d'autres ($\mathcal{M}|_i$ ER élevée) devraient utiliser plus souvent les certitudes fortes, d'où l'hypothèse H2.

H2. Le taux de bonnes réponses d'un sujet ($\mathcal{M}|_i$ER) est positivement corrélé avec le taux d'utilisation de certitudes fortes.

$(TUC_2|_i + TUC_3|_i)$

3. La cohérence entre épreuves en fréquence d'utilisation.

Les tests les plus faciles devraient entraîner une utilisation plus grande des indices de certitude élevés, d'où l'hypothèse H3.

H3. Il existe une corrélation positive et élevée entre la facilité de chaque test $\mathcal{M}|_t ER$ et le taux d'utilisation des certitudes fortes $(TUC_{2|t} + TUC_{3|t})$.

COHERENCES DANS L'EXACTITUDE DES REPONSES FOURNIES AVEC LES DIVERS DEGRES DE CERTITUDE (cohérences en μ)

1. La cohérence globale en taux d'exactitude.

Plus les réponses sont accompagnées d'un degré de certitude (ESPER ou C) élevé, plus leur succès $(\mathcal{M}|_c ER)$ doit être élevé, d'où l'hypothèse H4.

H4. Il existe, pour l'ensemble des réponses, une corrélation positive et élevée entre le degré de certitude (C) utilisé et le taux d'exactitude des réponses auxquelles il a été associé : $\mathcal{M}(ER \mid C = c)$ ou μ_c.

2. La cohérence individuelle en taux d'exactitude.

Au niveau de chaque individu, les degrés de certitude faibles doivent être attribués à des réponses dont les taux de succès seront faibles, d'où l'hypothèse H5.

H5. Il existe, pour chaque sujet, une relation croissante entre le degré de certitude utilisé (C) et le taux d'exactitude des réponses auxquelles il a été associé : $\mathcal{M}|_i(ER \mid C = c)$ ou μ_{ic}.

3. La stabilité entre individus du taux d'exactitude (d'un degré de certitude donné).

Quelle que soit la compétence des sujets, le taux d'exactitude des réponses auxquelles un degré de certitude a été associé ne devrait pas varier, d'où l'hypothèse H6.

H6. Il existe une corrélation nulle entre le taux d'exactitude des réponses $M|_i\,(ER|C = c)$ auxquelles un degré de certitude a été associé par un sujet et le taux d'exactitude des réponses de ce sujet, $M|_i\,ER$.

On peut déjà raisonnablement prédire que la consigne ordinale va amener une variabilité regrettable entre les individus pour leurs moyennes de réussite pour chaque degré de certitude $M|_i(ER|C = c)$ ou μ_{ic}

4. La cohérence en taux d'exactitude à l'intérieur d'un test.

Quelles que soient les caractéristiques d'un test et notamment sa facilité $M|_t\,ER$, les degrés de certitude faibles doivent correspondre à des réponses dont le taux d'exactitude $M|_t(ER|C = 0 \text{ ou } 1)$ est faible, d'où l'hypothèse H7.

H7. Il existe, pour chaque test, une corrélation positive et élevée entre C et $M|_t\,(ER|C = c)$ ou μ_{tc}.

5. La stabilité en taux d'exactitude entre épreuves.

Quelle que soit la facilité du test ou l'ordre du test dans la séquence, le taux de succès des réponses correspondant à une certitude donnée devrait ne pas varier, d'où les hypothèses H8 et H8'.

H 8. La corrélation entre $M|_t\,ER$ et $M|_t\,(ER|C = c)$ est nulle.

H 8'. La corrélation entre l'ordre des tests et $M|_t\,(ER|C = c)$ est nulle.

Il importe de considérer trois aspects de l'expérience qui contribueront à sous-estimer les hypothèses H1 à H8 décrites ci-avant.

?

Aspect 1. Un point ne peut respecter totalement les principes du conditionnement operant : les résultats ne sont fournis que quelques jours après les tests, et ce, en bloc pour la vingtaine de questions. Cet inconvénient est lié à la technologie scolaire actuelle : chaque étudiant ne dispose pas encore des services d'un terminal d'ordinateur qui lui communiquerait immédiatement les conséquences de chaque réponse isolée dès qu'elle est produite.

Aspect 2. Les sujets n'ont pas été réellement "entraînés". Ils ont simplement été exposés aux conséquences de leurs actes au cours des tests successifs, mais aucun exercice systématique préalable ne leur a été fourni. S'il en avait été autrement, il eût été impossible de mesurer l'impact de la seule situation operante.

Aspect 3. La matrice des conséquences n'a pas été construire selon le critère de l'E.S.U.; elle a été choisie pour sa facilité mnémotechnique et pour la simplicité des calculs.

°
° °

Rappelons un autre volet de notre thèse générale. Quand les sujets sont exposés aux conséquences de leurs actes (par exemple, si les points - ou toute autre valeur jugée utile par les sujets - sont dépendants des réponses fournies), ils présentent un comportement directement contrôlé par le programme de conditionnement operant (après une période d'exposition à ce programme).

L'hypothèse générale qui en résulte est la suivante : "Dans l'expérience présente, les sujets utilisent la stratégie qui leur donne le maximum de points au total (par application du critère de l'espérance subjective de l'utilité) et aucune autre stratégie."

Cette interrogation très générale sera décomposée en une série d'hypothèses particulières relatives à chaque stratégie décrite par la théorie des décisions et par la théorie du dépliage. Ensuite, d'autres hypothèses concerneront la stratégie E.S.U. et un approfondissement des résultats de l'expérience présente; ces hypothèses seront présentées au moment même de leur vérification.

E. LES DONNEES DE BASE.

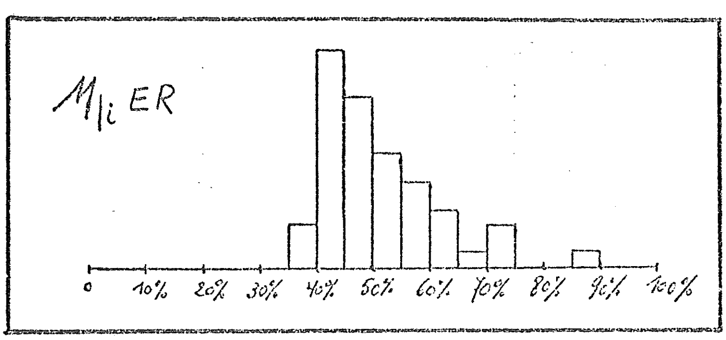
L'expérience de l'année scolaire 1971-1972 a été menée à l'Ecole Technique de la Force Aérienne belge à Saffraanberg (ETFA), par le Commandant P. VAN ROY et nous-mêmes (1). Plus de deux cents sujets (2), répartis en plusieurs classes, ont été soumis, de septembre 1970 à juin 1971, à une trentaine de tests successifs (de 25 questions en général) portant sur la mécanique, la chimie, la physique, l'électricité et l'électronique.

Puisqu'il importe d'étudier l'évolution du comportement, on n'a retenu que la plus longue série de résultats successifs pour les mêmes sujets, sur une même matière.

Il s'agit de 62 étudiants répartis en quatre classes qui ont, en neuf mois (septembre 1970 - juin 1971) reçu 14 tests de 25 questions en moyenne, portant sur la mécanique. Seuls 53 résultats complets sont disponibles car neuf étudiants ont dû modifier le cours de leurs études à Noël (3). Parmi les 53 sujets, 14 ont été absents lors d'un test, 3 absents lors de deux tests, 2 lors de trois tests, 2 lors de quatre tests, un lors de cinq tests.

Les pourcentages de réussite des divers sujets sont présentés dans le tableau 2.5.

Tableau 2.5



Distribution des pourcentages individuels de réussite
année scolaire 1971-1972

(1) Nous tenons à remercier le Col. MOONEN et le Maj. VAN LEUVEN pour leur collaboration dans cette entreprise, les enseignants qui ont mis leur confiance dès le début dans le "Bank System", ainsi que l'équipe technique de la section "contrôle des études" que le Col. MOONEN a créée à notre demande.

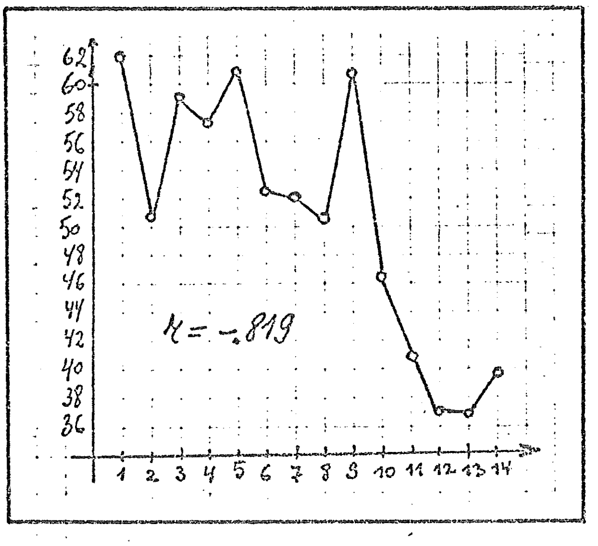
(2) Age modal : 17 ans.

(3) Cette fonction de réorientation est un aspect tangible de la procédure operante.

Le pourcentage moyen de réussite est de 50,16 % et l'écart type vaut 10,25. Calculés sur 62 sujets, ces résultats évitent donc l'inconvénient rencontré par HAMBLETON, ROBERTS et TRAUB (1970) de la trop grande facilité des questions. On trouvera, en annexe 4, les tableaux de résultats par test et par sujet.

La facilité des tests décroît au cours de l'année (tableau 2.6). Ce fait est plutôt regrettable dans l'optique du présent travail. L'impact de cette variable sera contrôlé systématiquement, notamment au niveau des hypothèses H3 et H8.

Tableau 2.6



Facilité des 14 tests successifs
(année scolaire 1971-1972)

F. REMARQUES METHODOLOGIQUES.

REMARQUE 1.

Dans les analyses qui suivent, c'est le coefficient de Bravais-Pearson qui est utilisé pour toutes les corrélations. On pourrait s'étonner du calcul de corrélations et d'équations de régression pour n = 14. Rappelons que les 14 valeurs sont des moyennes calculées sur les mêmes sujets à chaque test. En plus de cette homogénéité dans la population, on est assuré d'une grande homogénéité quant aux circonstances (consignes, temps disponible, matière, locaux...). Les nombreuses précautions prises en amont nous permettent de penser que les analyses statistiques ne dénatureront pas les résultats. Les variables sont entremêlées d'une telle façon que, sans l'usage d'outils statistiques (corrélations partielles entre autres), il serait impossible de rendre aux phénomènes leur importance exacte.

REMARQUE 2.

Il est une variable pour laquelle le ρ de Spearman s'impose: c'est la variable "numéro d'ordre des tests". Dans son chapitre sur les séries temporelles; KENDALL (1961) présente la corrélation (de Spearman) comme un moyen d'évaluer l'évolution temporelle d'un phénomène. Les valeurs des deux types de corrélations sont proches, comme le montre le tableau 2.7.

Tableau 2.7

Corrélation	Spearman	Bravais-Pearson
Ordre-omissions	.964	.939
Ordre - C1	-.518	-.624
Ordre - C2	-.639	-.757
Ordre - C3	-.239	-.212
Ordre-facilité	-.802	-.819

Comparaison des valeurs calculées par deux formules de corrélation.

Afin de ne pas mêler, par exemple dans le calcul des corrélations partielles, des corrélations calculées par les deux méthodes, nous avons systématiquement utilisé une seule méthode pour toutes les variables : celle de Bravais-Pearson.

REMARQUE 3.

Il y a deux façons de contrôler l'influence de la variable perturbante "facilité du test".

1) On calcule les corrélations partielles entre l'ordre des tests et le taux d'utilisation de chaque indice, l'influence de la facilité étant "contrôlée".

   La formule est : $$r_{12.3} = \frac{r_{12} - r_{13}\, r_{23}}{\sqrt{1 - r_{13}^2}\sqrt{1 - r_{23}^2}}$$

   où la variable perturbante (ici la facilité) est la variable 3.

2) On considère non pas le taux absolu (d'omissions par exemple) mais son taux relatif (parmi les certitudes faibles). Cette façon de faire ne neutralise pas complètement l'influence de la variable "facilité". En effet, la corrélation entre les certitudes fortes et la facilité est élevée (.886), mais non parfaite. Pour cette raison, l'autre procédure (les corrélations partielles) fera l'objet de plus d'intérêt.

2. VALIDITE ET COHERENCES DANS LE TAUX D'UTILISATION DE CHAQUE DEGRE DE CERTITUDE
(cohérences en τ)

1. *LA COHERENCE GLOBALE EN TAUX D'UTILISATION.*

H1. Le taux d'utilisation de certitudes maximales (ici cert. 3) est plus élevé et le $TUC_{minimales}$ (ici cert. 1) est plus faible pour les réponses correctes (ER = 1) que pour les réponses incorrectes (ER = 0).

$(TUC_3 | ER = 1) > (TUC_3 | ER = 0)$

$(TUC_1 | ER = 1) < (TUC_1 | ER = 0)$

Sur les 18147 items, on a observé :

9314 réponses correctes (51,32 %)
4968 réponses incorrectes (27,37 %)
3865 omissions (21,29 %)

Parmi les réponses correctes, on a observé :

594 certitudes 1 (soit 6,37 %)
1432 certitudes 2 (soit 15,37 %)
7288 certitudes 3 (soit 78,24 %)

Parmi les réponses incorrectes, on a observé :

943 certitudes 1 (soit 18,98 %)
1287 certitudes 2 (soit 25,90 %)
2738 certitudes 3 (soit 55,11 %)

L'hypothèse H1 est donc pleinement confirmée, puisque 6,37 % < 18,98 % et 78,24 % > 55,11 %.

2. *LA COHERENCE ENTRE INDIVIDUS EN FREQUENCE D'UTILISATION.*

Les individus plus compétents que d'autres ($\mathcal{M}|_i$ ER élevée) devraient utiliser plus souvent les certitudes fortes, d'où l'hypothèse H2).

H2. Le taux de bonnes réponses d'un sujet ($\mathcal{M}|_i$ ER) est positivement corrélé avec le taux d'utilisation de certitudes fortes, τ_F.

Le tableau 2.8 présente les taux d'utilisation et les taux de bonnes réponses, sujet par sujet.

Tableau 2.8

Code	-3	-2	-1	OM	+1	+2	+3	% BR (1) μ	% BR (2)	% BR (3)	NBR	NR	Cert. fortes	% cert. fortes τ_F	% BR
1602	83	24	21	31	8	29	168	27	54	66	205	364	304	83,5	56,31
1603	67	34	4	83	1	19	156	20	35	69	176	364	276	75,8	48,35
1604	99	28	20	6	5	24	182	20	46	64	211	364	333	91,48	57,96
1606	86	8	2	25	1	7	235	33	46	73	243	364	336	92,3	66,75
1607	83	17	9	57	2	13	183	18	43	68	198	364	296	81,31	54,39
1608	36	31	7	81	1	78	130	12	71	78	209	364	275	75,54	57,41
1610	73	10	1	46	2	13	194	66	56	72	209	339	290	85,54	61,65
1611	65	26	15	69	17	23	149	53	46	69	189	364	263	72,25	51,92
1612	94	14	0	27	0	11	193	0	44	67	204	339	312	92,03	60,17
1613	75	8	1	49	1	25	181	50	75	70	206	339	289	85,25	60,76
1701	79	36	23	64	7	22	133	23	37	62	162	364	270	74,17	44,50
1702	59	24	21	83	8	23	145	27	48	71	176	363	251	69,14	48,48
1703	85	47	14	56	6	36	120	30	43	58	162	364	288	79,12	44,50
1704	56	28	42	90	15	23	85	26	45	60	123	339	192	56,63	36,28
1705	37	55	52	74	23	42	81	30	43	68	146	364	215	59,06	40,10
1707	38	47	32	85	16	49	97	33	51	71	162	364	229	62,91	44,50
1708	44	36	67	54	36	44	83	34	55	65	163	364	207	56,86	44,78
1709	39	29	19	41	9	24	104	32	45	72	137	265	196	73,96	51,69
1710	58	32	34	61	22	21	136	39	39	70	179	364	247	67,85	49,17
1711	58	14	32	87	18	15	140	36	51	70	173	364	227	62,36	47,52
1712	40	66	19	33	12	60	134	38	47	77	206	364	300	82,41	56,52
1713	71	24	16	106	6	13	128	27	35	64	146	363	236	65,01	40,22
1714	34	46	19	89	13	100	73	40	68	75	186	364	243	66,75	51,09
1715	64	31	9	112	6	15	104	40	32	61	125	341	214	62,75	36,65
1801	67	6	18	101	13	0	138	41	0	61	151	363	231	63,63	41,59
1802	48	23	23	9	20	27	213	46	54	81	260	354	311	87,85	73,44
1803	62	27	27	48	18	36	120	40	57	65	174	338	245	72,48	51,47
1804	35	46	38	81	24	48	91	38	51	72	163	363	220	60,60	44,90
1805	45	83	4	75	6	57	93	60	40	67	156	363	278	76,58	42,97
1807	4	64	33	91	23	75	48	41	53	92	146	338	191	56,50	43,19
1810	36	8	32	27	31	16	213	49	66	85	260	363	273	75,20	71,62
1812	20	18	31	66	21	37	52	40	67	72	110	245	127	51,83	44,89
1813	47	54	28	66	17	39	112	37	41	70	168	363	252	69,42	46,28
1814	15	19	50	42	31	60	123	38	75	89	214	340	217	63,82	62,94
1815	74	6	0	116	5	9	153	100	60	67	167	363	242	66,66	46,00
1816	33	0	0	10	0	0	295	0	89	0	295	338	328	97,04	87,27
1817	41	41	23	112	17	38	91	42	48	68	146	363	211	58,12	40,22
1818	32	9	14	60	5	14	58	26	60	64	77	192	113	58,85	40,10
1901	40	18	17	95	11	26	131	39	59	76	168	338	215	63,60	49,70
1902	47	5	1	110	4	3	168	80	37	78	175	338	223	65,97	51,77
1903	39	0	0	106	1	2	141	100	100	78	144	289	182	62,97	49,82
1904	28	1	0	77	0	0	195	0	0	87	145	301	224	74,41	64,78
1905	70	5	4	95	1	5	183	20	50	72	189	363	263	72,45	52,06
1906	53	2	5	108	8	7	130	61	77	71	145	313	192	61,34	46,32
1908	43	18	17	117	23	23	97	57	56	69	143	338	181	53,55	42,30
1909	30	27	7	68	1	37	142	12	57	82	180	312	236	75,64	57,69
1910	53	2	10	145	8	2	143	44	50	72	153	363	200	55,09	42,14
1911	60	9	7	106	0	6	176	0	40	74	182	364	251	68,95	50,00
1913	50	6	21	87	19	7	148	47	53	74	174	338	211	62,42	51,47
1914	22	45	9	110	11	83	84	55	64	79	178	364	234	64,28	48,90
1915	42	13	3	26	5	12	214	62	48	83	231	315	281	89,20	73,33
1916	21	11	24	125	31	27	99	56	71	82	157	338	158	46,74	46,44
1918	48	6	18	77	5	7	103	21	53	68	115	264	164	62,12	43,56

VENTILATION DES CERTITUDES INDIVIDUELLES

AU TOTAL DES 14 EPREUVES DE MECANIQUE (ANNEE SCOLAIRE 1971-72)

Il y a 1 contradiction à supposer que la certitude ou la connaissance est continue d'appeler "forte" une certitude 2 (dans une échelle de 1 à 3); elle devrait être moyenne.

La corrélation entre le taux d'utilisation des certitudes fortes (2 et 3) et le taux de bonnes réponses vaut .8159. Avec les certitudes faibles, la corrélation est évidemment -.8159.

3. *LA COHERENCE ENTRE EPREUVES EN FREQUENCE D'UTILISATION.*

Les tests les plus faciles devraient entraîner une utilisation plus grande des indices de certitude élevés, d'où l'hypothèse H3.

H3. Il existe une corrélation positive et élevée entre la facilité de chaque test ($M|_t$ ER) et le taux d'utilisation des certitudes fortes ou τ_F.

Tableau 2.9

Dates		Erreurs -3	-2	-1	Omissions 0	Réussites +1	+2	+3	Total	τ U1	τ U2	τ U3	BR	μ % Br.
27- 9-70	1	84	100	97	157	70	160	482	1150	167	260	566	712	61,91
11-10-70	2	148	167	108	173	57	137	412	1202	165	304	560	606	50,41
25-10-70	3	160	116	79	156	53	146	540	1250	132	262	700	739	59,12
8-11-70	4	172	78	59	219	41	113	568	1250	100	191	740	722	57,76
6-12-70	5	305	128	103	472	93	240	1229	2570	196	368	1534	1562	60,77
16-12-70	6	213	71	73	216	37	67	523	1200	110	138	736	627	52,25
24- 1-71	7	176	94	48	245	34	81	495	1173	82	175	671	610	52,00
7- 2-71	8	212	71	38	247	23	66	487	1144	61	137	699	576	50,34
6- 3-71	9	179	47	33	236	33	75	645	1248	66	122	824	753	60,33
14- 3-71	10	242	86	55	295	39	87	457	1261	94	173	699	583	46,23
17- 4-71	11	264	93	54	312	23	57	421	1224	77	150	685	501	40,93
3- 5-71	12	224	88	68	390	30	75	350	1225	98	163	574	455	37,14
15- 5-71	13	164	65	57	345	32	58	279	1000	89	123	443	369	36,90
5- 6-71	14	195	83	71	402	29	70	400	1250	100	153	595	499	39,92
		2738	1287	943	3865	594	1432	7288	$\boxed{18147}$	1537	2719	10026	9314	50,43 (1) 51,3[illegible] (2)

UTILISATION DE CHAQUE DEGRE DE CERTITUDE LORS DES
14 EPREUVES SUCCESSIVES DE MECANIQUE (ANNEE SCOLAIRE 1971 - 1972)

Les Nombres ont été obtenus à partir de 53 sujets.
(1) Valeur calculée à partir de la somme des 14 % BR.
(2) Valeur calculée par l'opération 9314/18147.

Plus le test est facile, plus les sujets utilisent les certitudes fortes, au dépend des certitudes faibles. La facilité des tests est corrélée à .886 avec les taux d'utilisation des certitudes fortes (et par conséquent à -.886) avec les taux d'utilisation des certitudes faibles (tableau 2.10).

La relation est curvilinéaire →

Tableau 2.10

	Omissions τ_0	Certitude 1 τ_1	Certitude 2 τ_2	Certitude 3 τ_3	τ_f Certitudes faibles (OM+C1)	τ_F Certitudes fortes (C2+C3)	Facilité
1	13,65	14,52	22,60	49,21	28,17	71,81	61,91
2	14,39	13,72	25,29	46,58	28,11	71,87	50,41
3	12,48	10,56	20,96	56,00	23,04	76,96	59,12
4	17,52	8,00	15,28	59,20	25,52	74,48	57,76
5	18,36	7,62	14,31	59,68	25,98	73,99	60,77
6	18,00	9,16	11,50	61,33	27,16	72,83	52,25
7	20,88	6,99	14,91	57,20	27,87	72,11	52,00
8	21,59	5,33	11,97	61,10	26,92	73,07	50,34
9	18,91	5,28	9,77	66,02	24,18	75,79	60,33
10	23,39	7,45	13,71	55,43	30,84	69,14	46,23
11	25,49	6,29	12,25	55,96	31,78	68,21	40,93
12	31,83	8,00	13,30	46,85	39,83	60,15	37,14
13	34,50	8,90	12,30	44,30	43,40	56,60	36,90
14	32,16	8,00	12,24	47,60	40,16	59,84	39,92
M1	21,69	8,56	15,03	54,75	30,21	69,77	50,43
M2	21,29	8,47	14,98	55,25	29,77	70,23	51,32

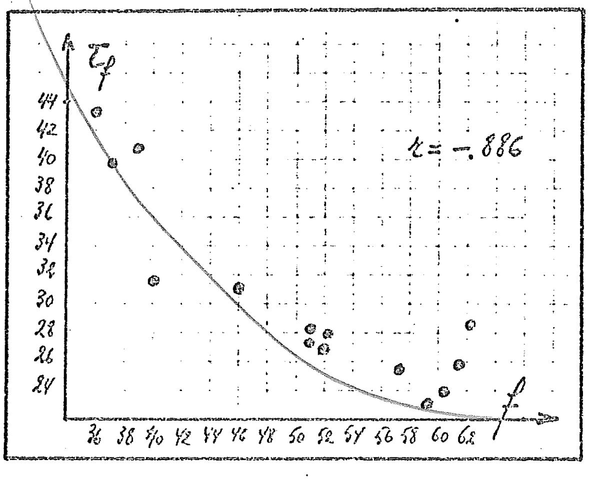
Pourcentage d'utilisation de chaque degré de certitude lors des 14 épreuves successives de mécanique (année scolaire 1971-1972)

M1 = Division de la somme de chaque colonne du présent tableau par 14

M2 = Division des sommes du tableau 2.9 par 18147

On utilisera surtout M1 (régression de chaque variable sur l'ordre des tests, par exemple).

Tableau 2.11



Liaison entre le taux d'utilisation des certitudes faibles (τ_f) et la facilité du test

Année scolaire 1971-1972

L'hypothèse H3 est donc pleinement confirmée.

3. VALIDITE ET COHERENCES DANS L'EXACTITUDE DES REPONSES FOURNIES AVEC LES DIVERS DEGRES DE CERTITUDE
(cohérences en μ)

1. *LA COHERENCE GLOBALE EN TAUX D'EXACTITUDE.*

Plus les réponses sont accompagnées d'un degré de certitude (ESPER ou C) élevé, plus leur succès doit être élevé, d'où l'hypothèse H4.

H4. Il existe, pour l'ensemble des réponses, une corrélation positive et élevée entre le degré de certitude (C) utilisé et le taux d'exactitude des réponses auxquelles il a été associé $M(ER|C=c)$ ou μ_c.

Sur les 18147 items, les résultats sont les suivants :

- 3865 omissions (soit 21,29 %)
- 1537 certitudes 1 (soit 8,46 %)
- 2719 certitudes 2 (soit 14,98 %)
- 10026 certitudes 3 (soit 55,24 %)

Parmi les réponses avec certitude 1, 594 sont correctes (μ_1 = 38,6 %)
certitude 2, 1432 sont correctes (μ_2 = 52,6 %)
certitude 3, 7288 sont correctes (μ_3 = 72,7 %)

La relation $\mu_1 < \mu_2 < \mu_3$ est donc vérifiée, puisque
38,6% < 52,6% < 72,7%.

Le rho de Spearman (1) vaudrait 1. L'hypothèse H4 est donc vérifiée. Le caractère nettement différencié de ces trois pourcentages montre, *a posteriori*, le bien fondé du choix de trois échelons distincts (un grand nombre d'expériences antérieures n'utilisent que deux échelons de certitude).

---

(1) Le caractère ordinal de la consigne rend préférable, en toute rigueur, le rho au r de Bravais-Pearson.

Ces résultats sont confirmés sur des séries moins longues de tests (20 au total) en chimie, physique et électricité, présentés à d'autres étudiants de seize ans. Ces vingt tests et les quatorze précédents totalisent 44684 réponses.

Le tableau 2.12 montre les valeurs de μ_1 , μ_2 et μ_3 pour ces groupes de tests (1).

Tableau 2.12

Echantillon	Nombre de réponses	μ_1	μ_2	μ_3
14 tests	19.396 *	37,8 %	51,4 %	71,4 %
20 tests	25.288	46,5 %	60,0 %	75,3 %
14+20 tests	44.684	42,5 %	56,1 %	73,5 %

Moyennes de réussite de chaque indice de certitude pour 34 tests de matières différentes appliqués à des individus différents (année scolaire 1971-1972).

(*) Ce nombre est calculé sur les 62 sujets.

(1) $\mu_1 = M(ER \mid C = 1)$

2. *LA COHERENCE INDIVIDUELLE EN TAUX D'EXACTITUDE.*

Au niveau de chaque individu, les degrés de certitude faibles doivent être attribués à des réponses dont les taux de succès seront faibles, d'où l'hypothèse H5.

H5. Il existe, pour chaque sujet, une relation croissante entre le degré de certitude utilisé (C) et le taux d'exactitude des réponses auxquelles il a été associé $M|_i(ER \mid C=c)$ ou μ_{ic}.

Appelons (1) <u>cohérence forte</u> la situation $\mu_{i1} \leqslant \mu_{i2} \leq \mu_{i3}$
<u>cohérence simple</u> la situation $\mu_{i1} < \mu_{i3}$
<u>non cohérence</u> les autres situations $\mu_{i1} \geqslant \mu_{i3}$

Le tableau 2.8 permet de vérifier l'hypothèse H5.

Sur les 53 étudiants, on dénombre 38 cohérences fortes (72 %)
12 cohérences simples (23 %)
3 non cohérences (5 %).

Si l'on ne considère que les sujets qui ont utilisé chaque degré de certitude plus de dix fois, on dénombre, sur 34 sujets :
31 cohérences fortes (91 %)
3 cohérences simples (9 %)

Ces résultats sont très favorables à l'hypothèse 5.

---

(1) Cf. les définitions de "transitivités stochastiques fortes" et "transitivités stochastiques faibles", dans F. BRESSON (1965, p. 234).

3. *LA STABILITE ENTRE INDIVIDUS DU TAUX D'EXACTITUDE (d'un degré de certitude donné).*

Quelle que soit la compétence des sujets, le taux d'exactitude des réponses auxquelles un degré de certitude a été associé ne devrait pas varier, d'où l'hypothèse H6.

H6. Il existe une corrélation nulle entre le taux d'exactitude des réponses $M|_i(ER|C=c)$ auxquelles un degré de certitude a été associé ou μ_{ic} et le taux d'exactitude des réponses de ce sujet $M|_i ER$ ou %BR.

On peut déjà raisonnablement prédire que la consigne ordinale va amener une variabilité regrettable entre les individus pour leurs moyennes de réussite pour chaque degré de certitude.

Les données de base sont contenues dans le tableau 2.8.

Les paramètres sont les suivants :

Tableau 2.13

	M	S^2	S
μ_1	37,84	441,9	21,02
μ_2	49,85	293,56	5,41
μ_3	72,21	60,01	7,75
% BR	51,30	450,8	21,23

Les corrélations avec le taux de réponses correctes sont :

(de la certitude 1) : -.0762

(de la certitude 2) : -.0441

(de la certitude 3) : .2925

Ces trois valeurs sont très favorables à l'hypothèse H6. Les deux premières sont nettement voisines de 0; quant à la troisième, elle est faible.

est-elle significative ?

On ne peut cependant pas accorder un même poids à un μ calculé sur n = 5 et à un μ calculé sur n = 50. Dès lors, une représentation d'ensemble pourrait être basée sur le principe suivant : à la moyenne d'exactitude (μ) de chaque degré de chaque sujet est attribuée une répétition égale au nombre de réponses sur lequel il a été calculé. Les répétitions obtenues sont alors portées sur un histogramme (classes regroupées).

L'histogramme de chacun des trois degrés peut être dressé à partir du tableau 2.14.

Tableau 2.14

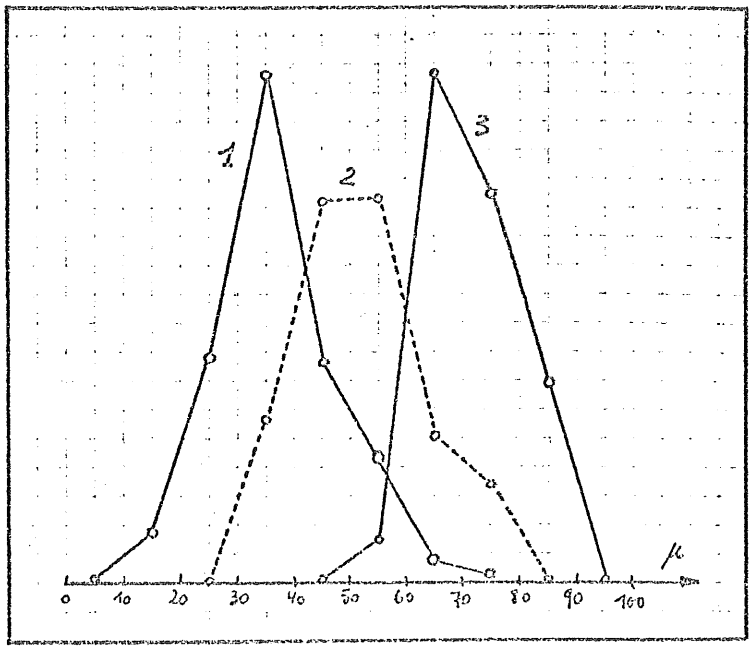
μ	Certitude 1	Certitude 2	Certitude 3
91-100 %	0,36	0,15	0,51
81-90 %	0	0	17,42
71-80 %	0,31	9,85	34,34
61-70 %	15,6	12,98	44,26
51-60 %	10,21	30,89	3,45
41-50 %	19,12	30,85	0
31-40 %	44,17	15,04	0
21-30 %	19,77	0	0
11-20 %	4,03	0	0
0-10 %	0,4	0	0
Moyennes sigmas	38,6 11,17	52,6 10,2	72,7 8,18

Répartition des μ de chaque portion de PER, après multiplication des μ individuels par les répétitions correspondantes (année scolaire 1971-1972)

Les sigmas restent élevés par rapport à l'hypothèse H6. Un entraînement systématique permettrait vraisemblablement de réduire ces variabilités.

Le tableau 2.14 peut être représenté graphiquement comme suit (tableau 2.15).

Tableau 2.15



"Histogrammes" des μ de chaque degré de certitude obtenus par multiplication des μ par les répétitions correspondantes

Année scolaire 1971-1972

4. *LA COHERENCE EN TAUX D'EXACTITUDE A L'INTERIEUR D'UN TEST.*

Quelles que soient les caractéristiques d'un test et notamment sa facilité ($M|_t$ ER), les degrés de certitude faibles doivent correspondre à des réponses dont le taux d'exactitude est faible $M|_t (ER|C=0$ ou 1 d'où l'hypothèse H7.

H7. Il existe, pour chaque test, une corrélation positive et élevée entre C et $M|_t(ER|C=c)$ ou μ_{tc}

?

Les pourcentages au cours des 14 tests successifs sont présentés dans le tableau 2.16.

Tableau 2.16

	μ_1	μ_2	μ_3
Test	Certitude 1	Certitude 2	Certitude 3
1	41	61	85
2	34	45	74
3	40	55	77
4	41	59	76
5	47	65	80
6	33	48	71
7	41	46	74
8	37	48	70
9	50	61	78
10	41	50	65
11	29	38	61
12	30	46	61
13	35	47	63
14	29	45	65
Moyenne	38,6	52,6	72,7

μ de chaque degré de certitude lors des 14 tests de l'année scolaire 1971-1972

On ne constate, à aucun test, une interversion des μ_{tc} de deux degrés de certitude. Le rho de Spearman vaut 1 à chaque test.

5. *LA STABILITE EN TAUX D'EXACTITUDE ENTRE EPREUVES*.

Quelle que soit la facilité du test ou l'ordre du test dans la séquence, le taux de succès des réponses correspondant à une certitude donnée devrait ne pas varier, d'où les hypothèses H8 et H8'.

Les données de base figurent au tableau 2.9.

H 8. La corrélation entre $\mu|_t$ ER et $\mu|_t$ (ER | C = c) est nulle.

Les pourcentages de réussite ($\mu|_t$(ER | C = c) ou μ_{tc} sont corrélés comme suit avec la facilité du test ($\mu|_t$ER) :

pour μ_{t1} (certitude 1) : r = .791
pour μ_{t2} (certitude 2) : r = .818
pour μ_{t3} (certitude 3) : r = .948

Ces résultats, contraires à l'hypothèse H8, s'expliquent par le caractère grossier de l'échelle des certitudes. La certitude 3, par exemple, représente aussi bien les ESPER juste supérieures à 75 % que les ESPER proches de 100 %. Il est raisonnable d'imaginer que, dans un test facile, les certitudes 3 représentent, en moyenne, des ESPER plus élevées que dans un test difficile. Cette hypothèse pourrait être vérifiée par une expérience où des degrés de certitude plus nombreux seraien disponibles.

H 8'. La corrélation entre $\mu|_t$(ER | C = c) et l'ordre du test est nulle.

Les corrélations brutes entre chacune de ces trois μ_t et le numéro d'ordre du test sont :

pour μ_{t1} (certitude 1) : -.430
pour μ_{t2} (certitude 2) : -.524
pour μ_{t3} (certitude 3) : -.808

Ces corrélations, contraires à l'hypothèse H8', résultent en partie de l'effet parasite de la facilité.

elles sont loin d'être négligeables

on peut conclure que H8 est infirmée

Après contrôle de la variable "facilité du test", les corrélations partielles entre les μ_t et le numéro d'ordre du test sont bien plus faibles que les corrélations de départ:

μ_{t_1} (certitude 1) = .614
μ_{t_2} (certitude 2) = .439
μ_{t_3} (certitude 3) =-.179

Il est intéressant, néanmoins, de constater que l'évolution de μ_t en fonction de l'ordre est très peu importante en ampleur (tableau 2.18).

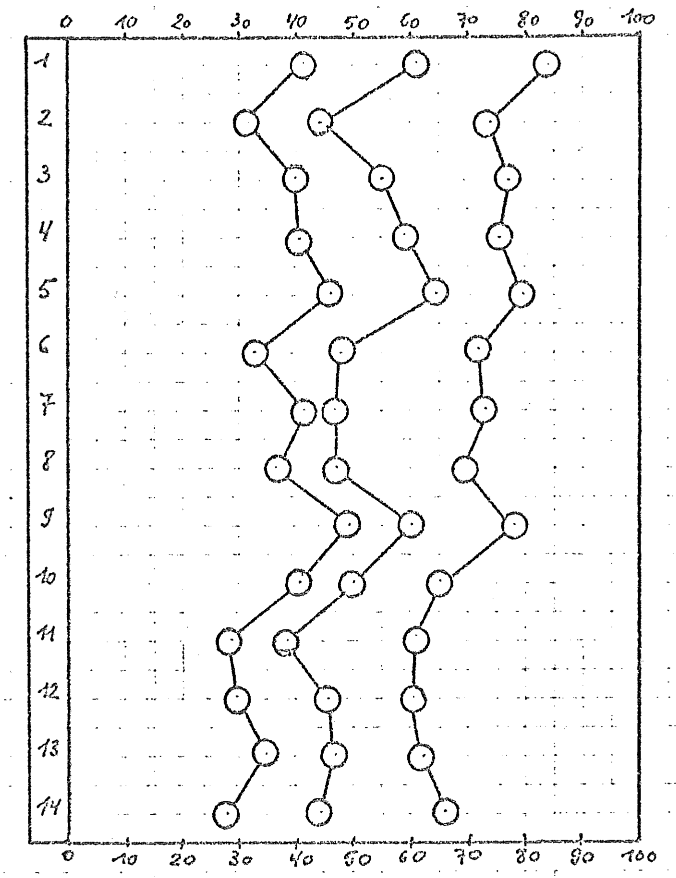
Le taux d'exactitude μ_{t_i} des réponses fournies avec chaque degré de certitude a été prédit à partir de la facilité du test.

Les équations de prédiction sont les suivantes :
pour μ_{t_1} (certitude 1) : 0,568 facil.+ 9.090
pour μ_{t_2} (certitude 2) : 0,715 facil. + 14,956
Pour μ_{t_3} (certitude 3) : 0,780 facil. + 32,039

Est-ce bien le même problème ? La constante d'équilibre est effectivement une constante par rapport aux concentrations.

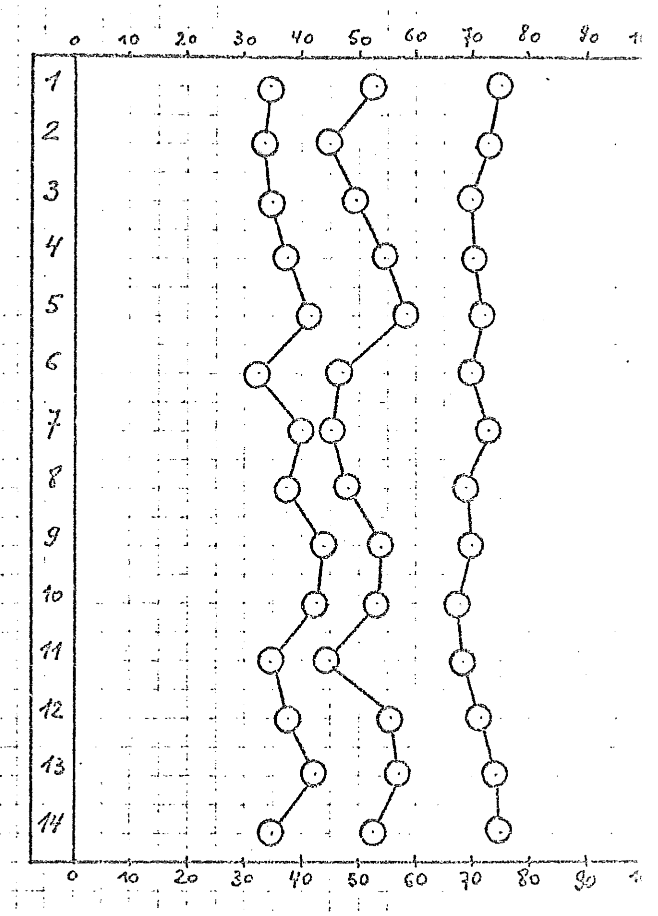
Les μ_{tc} de départ ont été portés sur le tableau 2.17.

Tableau 2.17



Moyennes d'exactitude des réponses fournies avec chaque indice de certitude au cours des 14 tests de l'année scolaire 1971-1972

Tableau 2.18



Moyennes d'exactitude des réponses fournies avec chaque degré de certitude au cours des 14 tests de l'année scolaire 1971-1972 après élimination de la facilité du test

Les résidus de la prédiction des μ_{tc} par la facilité (ajoutés à la moyenne des μ_{tc} de chaque degré) ont été portés sur le tableau 2.18. La différence de variation entre les scores originaux et les scores calculés est très nette, spécialement pour la certitude 3.

Que le taux d'exactitude des réponses fournies avec un certain degré de certitude soit lié à la facilité du test complique le problème, mais ne constitue nullement une contre-indication pour en continuer l'étude (1).

---

(1) Ce n'est pas parce que la constante de ~~concentration~~ varie avec la température que les chimistes ont renoncé au concept, ni même au terme "constante".

d'équilibre

4. STRATEGIES NON COMPATIBLES AVEC LA VALIDITE DES INDICES DE CERTITUDE

Les *patterns* individuels pour chaque test constituent la base des analyses qui suivent. Par *pattern* comportemental, nous entendons la distribution des degrés de certitude. Par exemple, lors d'un test à 20 questions, un sujet peut fournir trois omissions, six réponses avec certitude 1, huit réponses avec certitude 2, trois réponses avec certitude 3. L'ensemble des quatre nombres (3, 6, 8, 3) constitue le *pattern* de l'individu i au test t.

Dans les pages qui suivent, on cherche à savoir si ces *patterns* individuels sont compatibles (1), et ce de façon continue au fil des tests, avec les différentes stratégies décrites par la théorie des décisions. Seule la stratégie qui utilise le critère E.S.U. est compatible avec l'hypothèse H2. Quand des *patterns* comportementaux sont compatibles avec certaines stratégies, cela ne signifie pas qu'ils soient incompatibles avec toutes les autres stratégies (2).

On ne dispose pas de 742 *patterns* (53 sujets à 14 tests), mais bien de 703 *patterns* à cause des 39 absences.

---

(1) L'examen des seuls patterns ne peut apporter la preuve qu'un critère a été utilisé; par contre, il permet de prouver que certains critères n'ont pas été utilisés (cf; la phrase de BLALOCK, 1971, p. 21).

(2) Il existe en effet des patterns compatibles avec plusieurs critères simultanément.

HYPOTHÈSE H9. Les étudiants n'utilisent pas la stratégie 1 (application du critère maximax.

Utiliser le critère maximax consisterait à ne fournir que des réponses avec la certitude 3 (car elle donne le gain maximal en cas d'événement favorable).

Dix étudiants (sur 53) ont présenté ce *pattern* mais, pour sept de ces dix étudiants, lors d'un seul test et, pour les trois autres étudiants, lors de trois tests seulement. Il est donc bien possible que la stratégie 1 ait été utilisée, mais par peu de sujets et dans les tout premiers tests.

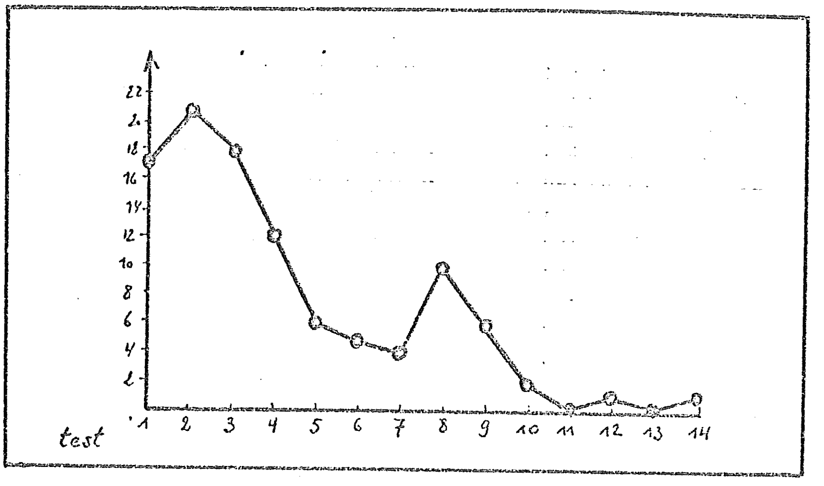
Un *pattern* moins strict, inspiré lui aussi du critère maximax, ne s'observerait-il pas ? Si l'on élargit l'interprétation du critère maximax, l'utilisation des seules certitudes 2 et 3 peut être considérée comme un *pattern* comportemental compatible avec la stratégie 1. Ce *pattern* s'observe chez neuf étudiants, mais pour cinq d'entre eux, une seule fois sur quatorze tests et, pour les autres sujets, jamais plus de quatre fois.

En élargissant encore l'interprétation du critère maximax, le *pattern* "absence d'omissions" est pris en considération. Ce *pattern* s'observe chez 29 sujets. Pour un seul d'entre eux, dix fois sur 14 tests chez aucun des 28 autres sujets, ce *pattern* ne s'observe plus de cinq fois sur 14 tests.

Quand on rassemble ces trois *patterns* (interprétations possibles - quoique très larges - du critère maximax), ils se présentent surtout au début de la série des tests, mais disparaissent quasi complètement par la suite. Un tel résultat semblerait indiquer que la procédure operante a bien eu pour effet de diminuer le recours à cette stratégie non operante.

Ce sont des *patterns* "de début" qui tendent à disparaître au cours des tests successifs. Le processus de disparition ne s'amorce qu'au quatrième test (tableau 2.19).

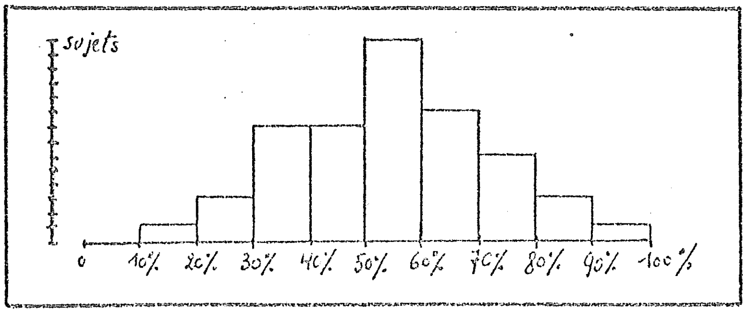
Tableau 2.19



Evolution du nombre de patterns compatibles avec une interprétation très large du critère maximax
Année scolaire 1971-1972

On peut considérer le critère maximax avec moins de rigueur encore et accepter des *patterns* où 95 % de certitude 3 apparaissent. Cette approche, elle non plus, n'est pas favorable au critère maximax. En effet, voici (tableau 2.20) comment se répartissent, pour les 53 sujets, les taux d'utilisation de la certitude 3 sur le total des 14 tests :

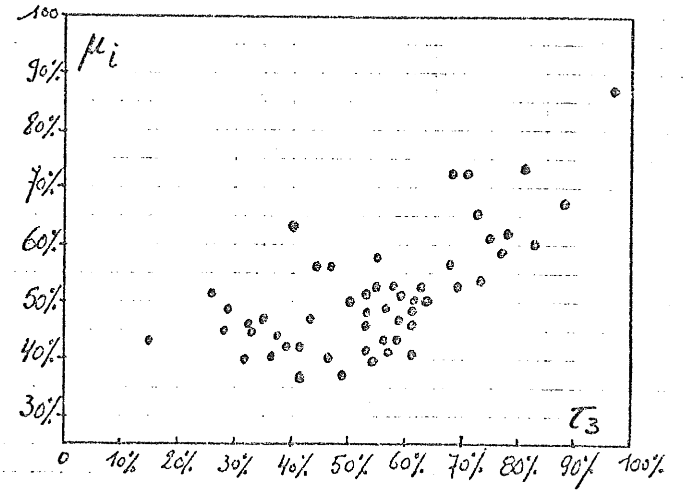
Tableau 2.20



Taux individuel d'utilisation de la certitude 3
Année scolaire 1971-1972

Le tableau 2.20 montre qu'un seul sujet (97 % de certitude 3) relèverait du critère maximax (élargi). Cependant, l'aspect gaussien de la courbe porte à croire qu'il s'agit tout simplement du cas extrême de la distribution normale. Cette dernière explication est confirmée par la forte corrélation (.68) existant entre le taux d'utilisation de la certitude 3 ($\bar{\tau}_{i3}$) et le pourcentage de réponses correctes fournies (μ_i) pour un même sujet, comme le montre le tableau 2.21.

Tableau 2.21



Lien entre la moyenne de réponses correctes (p) de chaque sujet et le taux individuel d'utilisation de certitude 3 (T3)

Année scolaire 1971-1972

Sur le tableau 2.21 il apparaît que le sujet qui présente 97 % de réponses avec la certitude 3 est justement celui qui fournit le plus de réponses correctes (87,2 %).

HYPOTHÈSE H10. Les étudiants n'utilisent pas la stratégie 2 (application du critère maximin), qui consiste à omettre sans cesse puisque c'est le comportement qui donne la pénalisation la moins grave (0) en cas d'échec.

Aucun étudiant n'a présenté ce pattern à aucun test.

En considérant ce critère de la façon la plus complaisante possible, on pourrait accepter le pattern où l'omission est plus utilisée que tout autre type de réponse. Un tel pattern s'observe chez un seul sujet et l'avantage de l'omission est bien faible :

Omission	Certitude 1	Certitude 2	Certitude 3
36 %	16 %	11 %	35 %

Une telle étude test par test est quasi impossible vu les petits nombres sur lesquels elle porterait.

H11. Les étudiants n'utilisent pas la stratégie 3 (application du critère d'équiprobabilité).

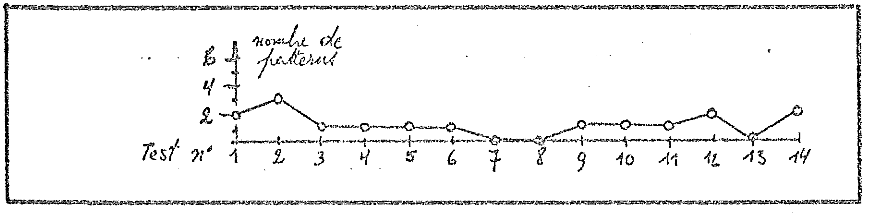
Ce critère consiste à attribuer une probabilité égale à l'exactitude et à l'inexactitude de la réponse fournie.

Les espérances mathématiques (utilités attendues) sont alors :
- Certitude 1 : $(0,5 \times 1) + (0,5 \times -1) = 0,5 - 0,5 = 0$
- Certitude 2 : $(0,5 \times 2) + (0,5 \times -2) = 1 - 1 = 0$
- Certitude 3 : $(0,5 \times 3) + (0,5 \times -3) = 1,5 - 1,5 = 0$

Les utilités attendues sont les mêmes pour les quatre actions disponibles. Dès lors, le critère ne peut guider le comportement : il n'en découle aucune stratégie.

On pourrait cependant, avec beaucoup de complaisance, considérer comme indicatif de la stratégie 3 le *pattern* qui consiste à fournir un nombre égal de réponses de chacun des quatre types (omission, certitude 1, certitude 2 et certitude 3). Aucun sujet n'a présenté, à aucun test, le *pattern* correspondant au critère de Laplace. Si l'on interprète de façon nettement moins stricte, on retiendra les cas où chacune des actions disponibles est utilisée entre 15 et 35 %. Ce *pattern* s'est présenté chez 15 sujets, mais 14 surviennent aussi bien lors des derniers que des premiers tests, ce qui permet de penser que ces 16 *patterns* sur 689, soit 2 %, sont dus au hasard (tableau 2.22):

Tableau 2.22



Evolution du nombre de patterns compatibles avec le critère de LAPLACE (interprétation large) lors des 14 tests de l'année scolaire 1971-1972

HYPOTHÈSE H12. Les étudiants n'utilisent pas la stratégie 4 (application du critère de HURWICZ).

Ce *pattern* de réponse correspondant au critère de HURWICZ consiste à ne fournir que des réponses avec la certitude 3 (si l'on est optimiste) ou alors à omettre sans cesse (si l'on est pessimiste).

Puisque le *pattern* "omettre sans cesse" (examiné à propos de la stratégie 2) n'a pas été observé une seule fois, il faudrait reconnaître qu'il n'existe pas de sujet pessimiste dans l'échantillon considéré. Quant aux sujets optimistes, ils auraient fourni les *patterns* déjà examinés à propos de la stratégie 1.

Il est raisonnable de penser que les *patterns* "répondre sans cesse avec certitude 3" s'expliquent à la fois par le critère du maximax et par le critère de HURWICZ (surestimation de la probabilité de l'événement favorable).

L'évolution au cours des 14 tests de ce *pattern* comportemental a été décrite à propos de la stratégie 1.

Il reste cependant vraisemblable que certain sujets sont "plutôt optimistes" ou "plutôt pessimistes", ce qui les mènerait à surévaluer ou sous-évaluer (mais de façon modérée) leurs PER.

Il est d'autant plus intéressant de constater que le comportement évolue dans un sens favorable sous l'effet de la procédure operante.

HYPOTHÈSE H13. Les étudiants n'utilisent pas la stratégie 5 (application du critère de SAVAGE).

Le *pattern* de réponse correspondant au critère de SAVAGE (regret minimax) consiste à ne fournir que des réponses avec la certitude 1 et 2.

En effet, les profits maximum sont :

- En cas de réponse correcte (E1) : + 3 points
- En cas de réponse incorrecte (E2) : 0 point.

Dans la matrice des regrets (donc valeurs négatives ou nulles) ci-dessous (tableau 2.23), le regret maximum pour chaque action possible est souligné.

Tableau 2.23

	E1 (réponse correcte)	E2 (réponse incorrecte)
Omission	- 3	0
Certitude 1	- 2	- 1
Certitude 2	- 1	- 2
Certitude 3	0	- 3

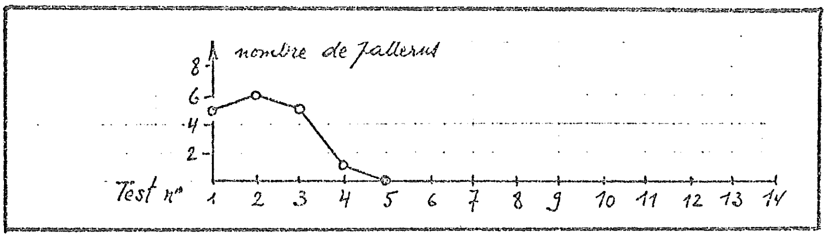
Matrice des regrets en application du critère de SAVAGE à la matrice de conséquences du tableau 2.1

Le plus petit des quatre regrets maxima est - 2 (qui se produit pour les certitudes 1 et 2).

Onze sujets ont présenté ce *pattern*. Sept sujets l'ont présenté lors d'un seul test, deux sujets lors de deux tests et deux sujets lors de trois tests.

Ce *pattern* n'a jamais été observé au-delà du quatrième test (tableau 2.24).

Tableau 2.24



Evolution des patterns compatibles avec le critère du regret minimax de SAVAGE au cours des 14 tests de l'année scolaire 1971-1972

On constate donc à nouveau que les trois premiers tests présentent des *patterns* qui pourraient être expliqués par des stratégies non compatibles, mais que ces *patterns* tendent à disparaître dès le quatrième test.

HYPOTHÈSE H14. Les étudiants n'utilisent pas la stratégie 6 (utilisation du critère de COOMBS : l'ampleur du risque).

Rappelons que cette stratégie peut être observée à la lumière de la théorie du dépliage de C.H. COOMBS, exposée dans le chapitre 1. Dans l'expérience, les ampleurs des risques sont les suivantes (tableau 2.25) :

Tableau 2.25

	Evénements possibles Réussite	Echec	Ampleur du risque
Omission	0	0	0
Rép. cert. 1	+1	-1	2
Rép. cert. 2	+2	-2	4
Rép. cert. 3	+3	-3	6

Ampleur des risques pour chaque degré de certitude de la matrice du tableau 2.1

Pour mettre le plus possible d'éléments favorables du côté de la stratégie 6, diverses fonctions de l'ampleur du risque ont été considérées. D'autres échelles jointes ont donc été testées

- à partir du carré de l'ampleur du risque (r^2);
- à partir d'une fonction "inverse" : $f(R_{i+1}) = f(R_i) + \frac{1}{R_{i+1}}$

  avec $f(R_0) = 0$;
- à partir d'une fonction logarithmique : $\log_{10}(r+1)$ ou $\log_{n}(r+1)$.

Les valeurs sont calculées dans le tableau 2.26 :

Tableau 2.26

	r	r^2	"inverse"	log. 10 (r+1)	log. n (r+1)
Omission	0	0	0	0	0
Rép. cert. 1	2	4	0,5	0,48	1,1
Rép. cert. 2	4	16	0,75	0,70	1,6
Rép. cert. 3	6	36	0,91	0,84	1,95

Echelles jointes obtenues pour certaines ampleurs de risque (r) pour la matrice du tableau 2.1

Quand on dessine les échelles jointes, les milieux se placent dans l'ordre suivant (tableau 2.27).

Tableau 2.27

Echelles jointes	Positions												
	1	2	3	4	5	6	7	8	9	10	11	12	13
r		M01		M02		M12		M03		M13		M23	
r^2		M01		M02		M12		M03		M13		M23	
inverses et log.		M01		M02		M03		M12		M13		M23	
zones	1—2—3—4—5—6—7—8—9—10—11—12—13												

Situation des 13 positions distinctes pour les divers types d'échelle jointe (tableau 2.28)

Les échelles jointes pour log 10 (r+1) et log n (r+1) sont les mêmes que pour 1/r. On constate que seules les positions relatives de M12 et M03 diffèrent d'une échelle à l'autre. C'est donc de ces zones (zones 6, 7 et 8) que proviendront des échelles individuelles différentes pour les trois échelles jointes. Les autres zones fourniront des ordres de préférence communs aux différentes échelles jointes. C'est ce que montre l'extrait du programme FORTRAN (tableau 2.28) qui confronte les *patterns* observé aux échelles individuelles pouvant provenir d'une des échelles jointes décrites ci-dessus.

Tableau 2.28

```
C-----------TEST DES I SCALES
C------MATRICE DES CONSEQUENCES NON BAYESIENNE
C----------PATTERNS COMMUNS A R, R2, 1/R
      K=0
      IF(OM.GT.C1.AND.C1.GT.C2.AND.C2.GT.C3)K=1
      IF(OM.EQ.C1.AND.C1.GT.C2.AND.C2.GT.C3)K=2
      IF(C1.GT.OM.AND.OM.GT.C2.AND.C2.GT.C3)K=3
      IF(C1.GT.OM.AND.OM.EQ.C2.AND.C2.GT.C3)K=4
      IF(C1.GT.C2.AND.C2.GT.OM.AND.OM.GT.C3)K=5
      IF(C2.GT.C1.AND.C1.GT.C3.AND.C3.GT.OM)K=9
      IF(C2.GT.C1.AND.C1.EQ.C3.AND.C3.GT.OM)K=10
      IF(C2.GT.C3.AND.C3.GT.C1.AND.C1.GT.OM)K=11
      IF(C2.EQ.C3.AND.C3.GT.C1.AND.C1.GT.OM)K=12
      IF(C3.GT.C2.AND.C2.GT.C1.AND.C1.GT.OM)K=13
      K2=K
      KN=K
C---------PATTERN ORIGINAL POUR R
      IF(C1.EQ.C2.AND.C2.GT.OM.AND.OM.EQ.C3)K=7
C---------PATTERNS ORIGINAUX POUR R2
      IF(C1.EQ.C2.AND.C2.GT.OM.AND.OM.GT.C3)K2=6
      IF(C2.GT.C1.AND.C1.GT.OM.AND.OM.GT.C3)K2=7
      IF(C2.GT.C1.AND.C1.GT.OM.AND.OM.EQ.C3)K2=8
C----------PATTERNS ORIGINAUX POUR 1/R
      IF(C1.GT.C2.AND.C2.GT.OM.AND.OM.EQ.C3)KN=6
      IF(C1.GT.C2.AND.C2.GT.C3.AND.C3.GT.OM)KN=7
      IF(C1.EQ.C2.AND.C2.GT.C3.AND.C3.GT.OM)KN=8
```

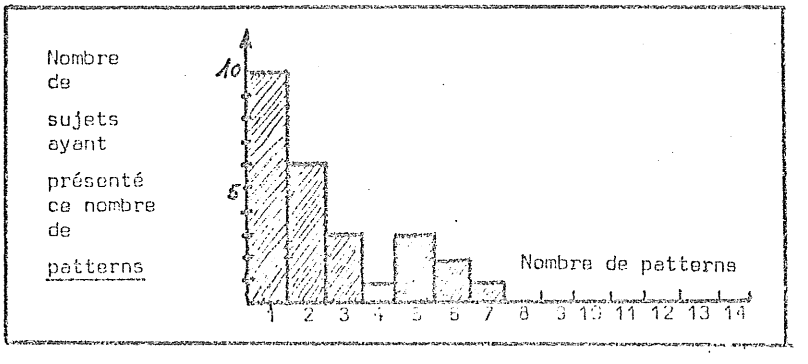
Extrait du programme FORTRAN relatif à la théorie du dépliage

Des *patterns* compatibles avec une des fonctions (r, r^2, inverse et log.) apparaissent-ils chez un grand nombre de sujets d'une façon stable ?

Sur les 703 *patterns* observés, 69 seulement sont "compatibles", c'est-à-dire qu'ils pourraient provenir d'une échelle jointe r, r^2, ou inverse, ou logarithmique. Quatre *patterns* supplémentaires ne sont compatibles qu'avec l'échelle inverse ou logarithmique. Ainsi donc, à peine 10 % des *patterns* observés sont compatibles avec la théorie du dépliage.

La moitié des sujets (N = 27) n'a jamais présenté de *pattern* compatible au cours des 14 tests. Pour l'autre moitié (N = 26), les nombres de *patterns* présentés sont les suivants (tableau 2.29).

Tableau 2.29



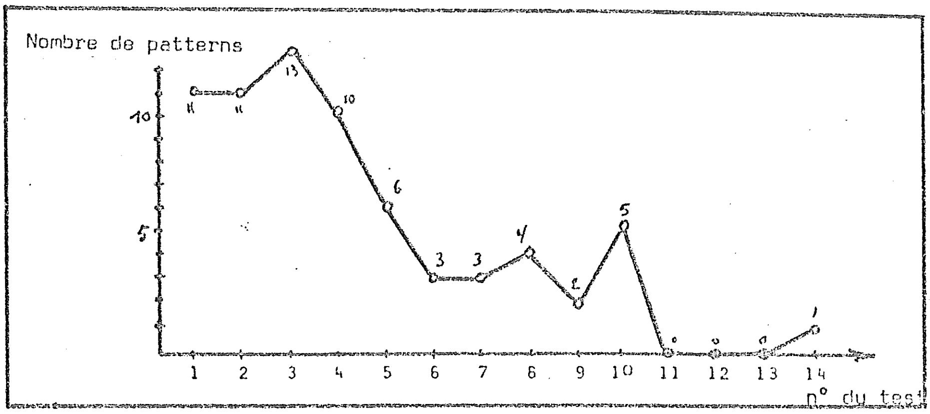
Répartition du nombre de patterns compatibles avec la théorie du dépliage parmi les 26 sujets qui les ont présentés (année scolaire 1971-1972)

Parmi ces 69 *patterns*, on dénombre :

46 *patterns* '13' (chez 21 sujets différents)
14 *patterns* '11' (chez 8 sujets différents)
6 *patterns* '12' (chez 5 sujets différents)
2 *patterns* '10' (chez 2 sujets différents)
1 *pattern* '9'.

Si l'on observe le moment où se sont présentés ces *patterns* on constate (tableau 2.30) qu'il s'agit d'un phénomène "de début" et que, si des *patterns* compatibles se présentent, ils sont progressivement abandonnés (la disparition des *patterns* n'est sensible qu'au cinquième test).

Tableau 2.30



Nombre de patterns compatibles avec la théorie du dépliage
année scolaire 1971-1972

L'hypothèse de stabilité des *patterns* doit être rejetée. En effet, si l'on observe les *patterns* des six étudiants qui en ont présenté au moins cinq, on constate :

Etudiant 1604 : 5 *patterns* 13

Etudiant 1703 : 2 *patterns* 12 et 3 *patterns* 13

Etudiant 1708 : 1 *pattern* 10
2 *patterns* 11
2 *patterns* 13

Etudiant 1712 : 4 *patterns* 11
1 *pattern* 12
1 *pattern* 13

Etudiant 1802 : 6 *patterns* 13

Etudiant 1814 : 2 *patterns* 11
1 *pattern* 12
4 *patterns* 13

Ces résultats, de toute façon très peu nombreux, indiquent que si certains sujets (1604 et 1802) ont, pour certains tests (moins de la moitié), un point idéal, pour les quelques autres, il faudrait plutôt parler d'une zone idéale (11-12-13).

Conclusions relatives à toutes les stratégies incompatibles.

On constate que si certaines d'entre elles ont été utilisées, ce n'est que chez une minorité de sujets et pour les trois premiers tests.

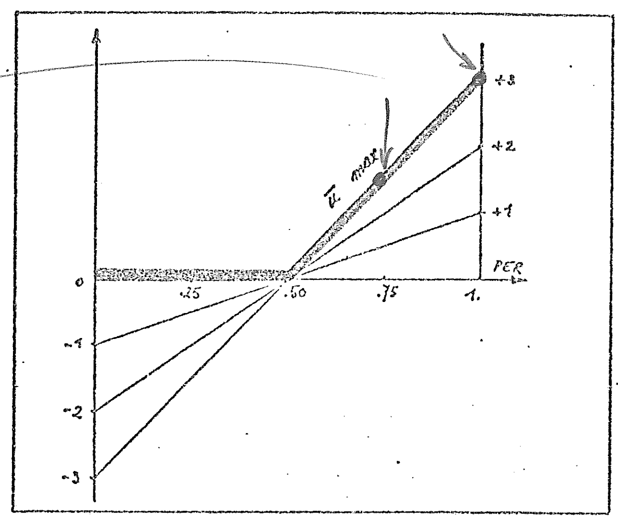
La seule exposition à la procédure operante les fait régresser nettement, surtout à partir du quatrième test. Nous pensons qu'un entraînement à partir d'exercices spécialement conçus dans cette intention réduirait plus nettement encore ces *patterns*, dès le premier test.

Il ne s'agit pas d'une fonction continue !

5. LA STRATEGIE COMPATIBLE AVEC LA VALIDITE DES INDICES DE CERTITUDE (ESPER)

Le *pattern* individuel de réponse correspondant parfaitement au critère de l'E.S.U. (maximisation de l'utilité attendue selon l'ESPER) consisterait <u>à ne présenter que des omissions</u> (quand la PER est inférieure à .50) <u>et des réponses avec certitude 3</u> (quand PER $\geqslant$ .50). La démonstration de l'efficacité de cette stratégie réside dans un petit graphique (tableau 2.3) :

Tableau 2.3



La fonction de maximisation des utilités attendues à partir de la matrice de conséquences du tableau 2.1

Les certitudes 1 et 2 ne sont jamais optimales et devraient donc disparaître. Néanmoins, près de la valeur-charnière de 50 %, leur utilité attendue ($\bar{u}$) est proche de celle des autres. C'est pourquoi leur élimination totale est peu probable, d'autant plus que les étudiants n'ont jamais connu le graphique ci-dessus et ne l'ont jamais soupçonné. Tout au plus, certains étudiants interrogés à la fin de l'année scolaire pouvaient-ils avancer (timidement) que "l'omission et la certitude 3, c'est ce qui rapporte le plus" (sic).

Hypothèses

H15. Les *patterns* compatibles avec le critère E.S.U. sont plus nombreux que les *patterns* compatibles avec d'autres critères.

H16. Les taux moyens d'omission et de certitude 3 augmentent avec le numéro d'ordre du test, tandis que les taux des certitudes 1 et 2 diminuent.

H17. Si le taux d'utilisation des certitudes fortes est très lié (cf. H2) à la facilité, alors les taux relatifs de l'omission parmi les certitudes faibles et de la certitude 3 parmi les certitudes fortes augmentent avec le temps.

H18. La disparition des certitudes 1 et 2 est décroissante et tend vers 0 pour un nombre de tests suffisamment grand.

H19. Chaque sujet présente une courbe de disparition des certitudes 1 et 2 uniformément décroissante.
Si l'on se base sur les courbes d'évitement enregistrées chez les animaux, on doit, au contraire, s'attendre à des types d'évolutions différentes selon les individus. Pour certains, la chute sera brusque (précoce, tardive ou intermédiaire); pour d'autres, on observera une décroissance continue. D'autres enfin présenteront une évolution en dents de scie avec tendance à la diminution.

Les trois hypothèses (H16, H17, H18) s'entendent "toutes choses par ailleurs égales". Or la facilité des tests décroît ($r = -.819$) au cours de l'année et ce de façon nette : elle passe de 60 % environ à 40 % environ. La chute de facilité s'oppose (H2) à l'augmentation des certitudes 3 (H16). Il s'agit là d'une variable parasite, d'où les trois hypothèses suivantes.

H 16 B. L'hypothèse H16 ne se vérifie pleinement qu'après le contrôle de la variable "facilité du test".

H 17B. Idem pour H17.

H 18B. Idem pour H18.

VÉRIFICATION DES HYPOTHÈSES

HYPOTHÈSE H15. Les patterns compatibles avec le critère E.S.U. sont plus nombreux que les patterns compatibles avec d'autres critères.

Le nombre de *patterns* individuels E.S.U. ("ne présenter que des réponses avec certitude 3 et des omissions") est plus important que le nombre de *patterns* incompatibles. En effet, 28 sujets (soit plus de 50 %) ont présenté ce *pattern*. Sept sujets l'ont présenté sept fois ou plus (soit lors de plus de la moitié des tests).

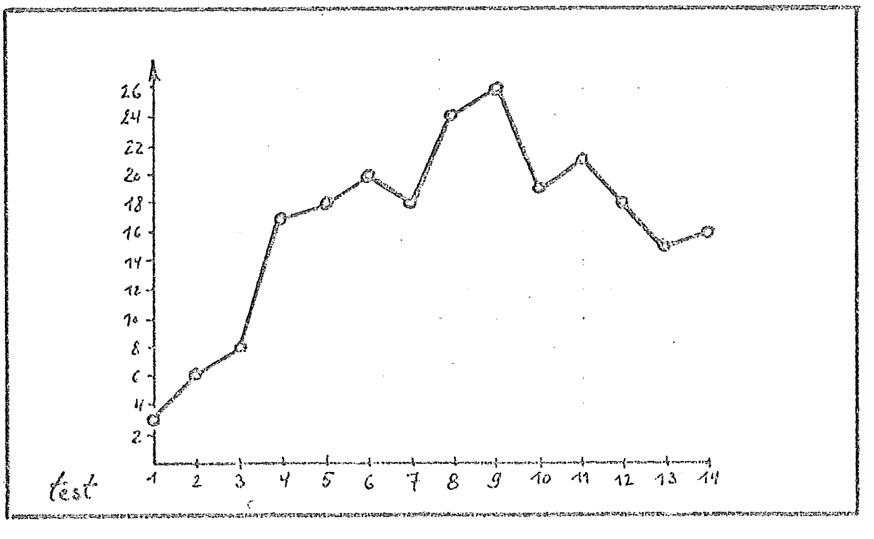
Si l'on considère le critère moins strictement, en admettant les *patterns* où la somme des réponses avec certitude 1 et 2 ne totalise pas 10 %, 37 sujets (soit 76 % des sujets) ont présenté ce *pattern*; 16 d'entre eux l'ont présenté dans la moitié des tests et plus.

Si on admet les *patterns* où la somme des réponses avec certitude 1 et 2 ne totalise pas plus de 20 %, on constate que 50 sujets sur 53 présentent ces *patterns*, dont 25 dans la moitié des tests et plus.

L'hypothèse H15 est donc vérifiée.

Voici l'évolution temporelle de ces *patterns* (tableau 2.31) :

Tableau 2.31



Nombre de patterns E.S.U. (où certitude 1 + certitude 2 $\leq$ 10 %)

Année scolaire 1971-1972

On constate à nouveau une grande différence entre les trois premiers tests et les suivants. Il semble donc que trois tests constituent la "période d'exposition au programme de renforcement" dont il est question dans l'hypothèse H19.

Quand on admet les *patterns* où la somme des certitudes 1 et des certitudes 2 n'excède pas 20 %, on assiste à nouveau à une montée qui culmine au test 9 puis à une descente (sans toutefois retomber au niveau des trois premiers tests).

La chute dans les cinq derniers tests paraît provenir de la chute de facilité des tests. Ce problème va être étudié à propos des hypothèses suivantes.

HYPOTHÈSE H16. <u>Les taux moyens d'omission et de certitude 3 augmentent avec le numéro d'ordre du test, tandis que les certitudes 1 et 2 diminuent.</u>

Le tableau 2.9 contient les données de base.

Tableau 2.9

Dates		Erreurs -3	-2	-1	Omissions 0	Réussites +1	+2	+3	Total	τ U1	τ U2	τ U3	BR	μ % Br.
27- 9-70	1	84	100	97	157	70	160	482	1150	167	260	566	712	61,91
11-10-70	2	148	167	108	173	57	137	412	1202	165	304	560	606	50,41
25-10-70	3	160	116	79	156	53	146	540	1250	132	262	700	739	59,12
8-11-70	4	172	78	59	219	41	113	568	1250	100	191	740	722	57,76
6-12-70	5	305	128	103	472	93	240	1229	2570	196	368	1534	1562	60,77
16-12-70	6	213	71	73	216	37	67	523	1200	110	138	736	627	52,25
24- 1-71	7	176	94	48	245	34	81	495	1173	82	175	671	610	52,00
7- 2-71	8	212	71	38	247	23	66	487	1144	61	137	699	576	50,34
6- 3-71	9	179	47	33	236	33	75	645	1248	66	122	824	753	60,33
14- 3-71	10	242	86	55	295	39	87	457	1261	94	173	699	583	46,23
17- 4-71	11	264	93	54	312	23	57	421	1224	77	150	665	501	40,93
3- 5-71	12	224	88	68	390	30	75	350	1225	98	163	574	455	37,14
15- 5-71	13	164	65	57	345	32	58	279	1000	89	123	443	369	36,90
5- 6-71	14	195	83	71	402	29	70	400	1250	100	153	595	499	39,92
		2738	1287	943	3865	594	1432	7288	$\boxed{18147}$	1537	2719	10026	9314	50,43 (1) 51,3[illegible] (2)

UTILISATION DE CHAQUE DEGRE DE CERTITUDE LORS DES 14 EPREUVES SUCCESSIVES DE MECANIQUE (ANNEE SCOLAIRE 1971 - 1972)

Les Nombres ont été obtenus à partir de 53 sujets.
(1) Valeur calculée à partir de la somme des 14 % BR.
(2) Valeur calculée par l'opération 9314/18147.

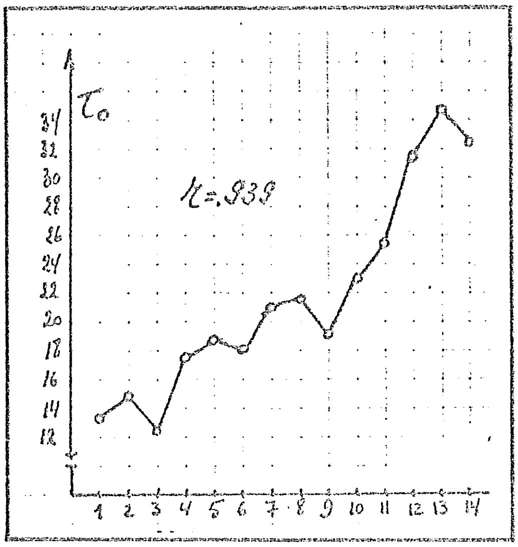
A partir du tableau 2.9, on peut élaborer les tableaux 2.32, 2.33, 2.34 et 2.35.

L'augmentation progressive de l'omission et la décroissance des certitudes 1 et 2 vont dans le sens de l'hypothèse H16a. Néanmoins, le taux d'utilisation de la certitude 3 évolue d'une façon inattendue (on aurait pu s'attendre à un accroissement continu). En fait, ces pourcentages bruts doivent être examinés à la lumière de l'hypothèse H16b.

η n'a aucune signification
dans une courbe présentant
un maximum

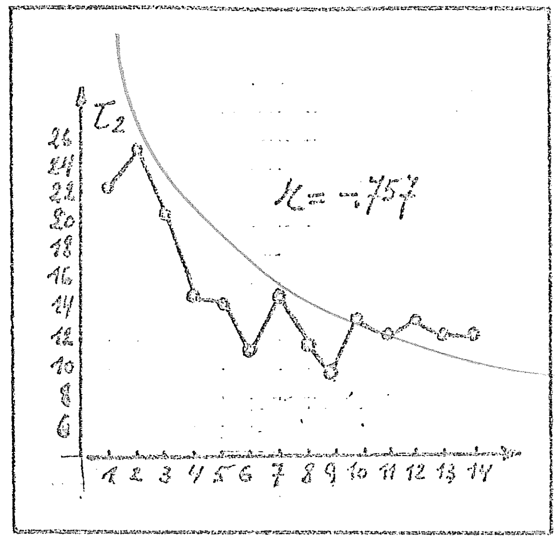
il fallait calculer les rapports de corrélation

Tableau 2.32



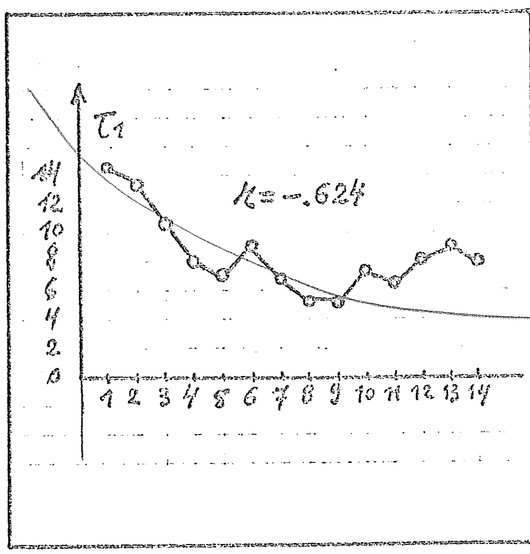
Taux d' utilisation de l'omission au cours des 14 tests de l'année scolaire 1971-1972

Tableau 2.34



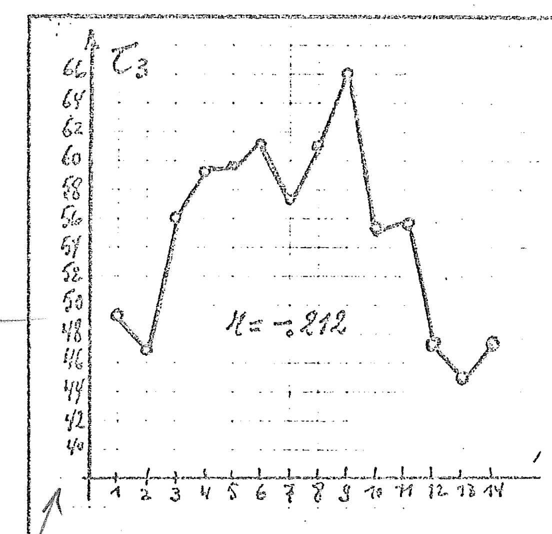
Taux d' utilisation de la certitude 2 au cours des 14 tests de l'année scolaire 1971-1972

Tableau 2.33



Taux d' utilisation de la certitude 1 au cours des 14 tests de l'année scolaire 1971-1972

Tableau 2.35



Taux d' utilisation de la certitude 3 au cours des 14 tests de l'année scolaire 1971-1972

que signifie r ici ?

Vous ne pouvez pas calculer de corrélations partielles sur des courbes qui s'écartent autant de la linéarité.

Que signifie une équation de régression dans ces conditions ?

HYPOTHÈSE H16B. L'hypothèse H16a ne se vérifie qu'après contrôle de la variable "facilité du test".

Voici les corrélations partielles (tableau 2.36) entre le taux d'utilisation de chaque degré de certitude et l'ordre des tests, avec suppression de l'influence de la facilité.

Tableau 2.36

	r de départ	r partielles
omissions	.939	.810
certitude 1	-.624	-.782
certitude 2	-.757	-.861
certitude 3	-.211	.595

rappel
r facilité-ordre = -.891

Corrélations avec l'ordre des tests au cours des 14 épreuves successives de l'année scolaire 1971-1972.

Les corrélations partielles sont bien plus en accord avec l'hypothèse H16a que les corrélations de départ. L'évolution du taux de certitude 3 prend une tout autre allure. Les hypothèses H16a et H16b sont donc confirmées.

Pour visualiser l'impact réel de l'ordre des tests sur l'utilisation de la certitude 3, on peut calculer les taux d'utilisation (TUC | t) par l'équation de régression suivante :

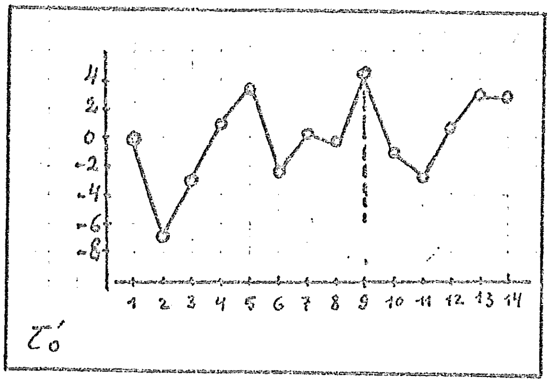
$$\tau_3' = 0,443\ x + 32,408$$

où τ_3 = pourcentage observé d'utilisation de la certitude 3
τ_3' = pourcentage calculé d'utilisation de la certitude 3
x = facilité du test.

Voici les graphiques (tableaux 2.37 à 2.40) de résidus des équations de régression :

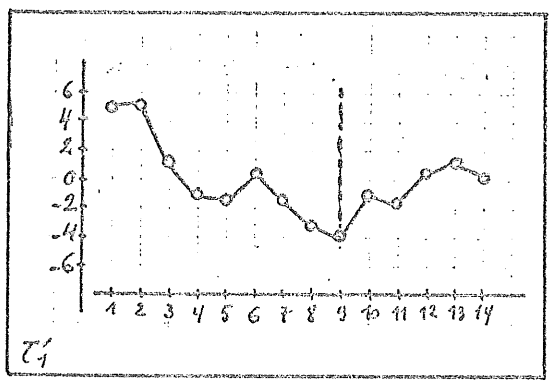
$$\tau_0' = 0,699\ x + 56,913$$
$$\tau_1' = 0,070\ x + 5,040$$
$$\tau_2' = 0,168\ x + 5,642$$
$$\tau_3' = 0,443\ x + 32,408$$

Tableau 2.37



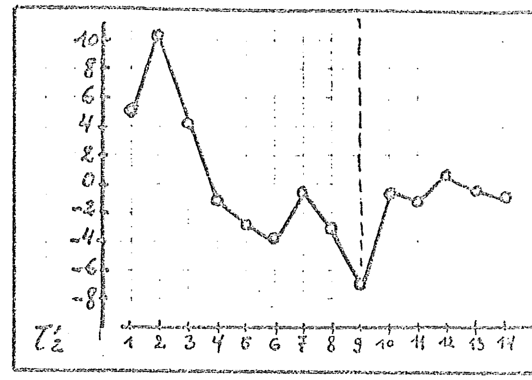
Evolution des résidus de la prédiction de τ_0'' (taux d'utilisation de l'omission) par la facilité du test (année scolaire 1971-1972)

Tableau 2.38



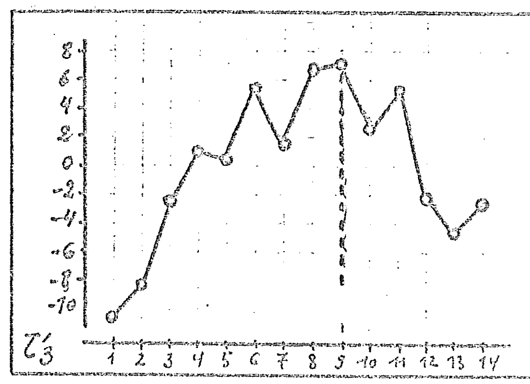
Evolution des résidus de la prédiction de τ_1 (taux d'utilisation de la certitude 1) par la facilité du test (année scolaire 1971-1972)

Tableau 2.39



Evolution des résidus de la prédiction de τ_2 (taux d'utilisation de la certitude 2) par la facilité du test (année scolaire 1971-1972)

Tableau 2.40



Evolution des résidus de la prédiction de τ_3 (taux d'utilisation de la certitude 3) par la facilité du test (année scolaire 1971-1972

L'examen de ces graphiques révèle l'importance-charnière du test 9. Pour les quatre graphiques, il constitue le point d'inversion des pentes. Du point de vue de la difficulté, il constitue aussi un moment charnière : les neuf premiers tests sont supérieurs à la facilité-charnière (50 %), alors que les cinq derniers tests sont inférieurs.

Malgré ces inversions de tendance, on constate à nouveau la différence systématique entre les trois premiers tests et les autres. La signification des coefficients (a) de pente figure au chapitre 4.

HYPOTHÈSE H 17A. Les taux relatifs de l'omission (parmi les certitudes faibles) et de la certitude 3 (parmi les certitudes fortes) augmentent avec le temps.

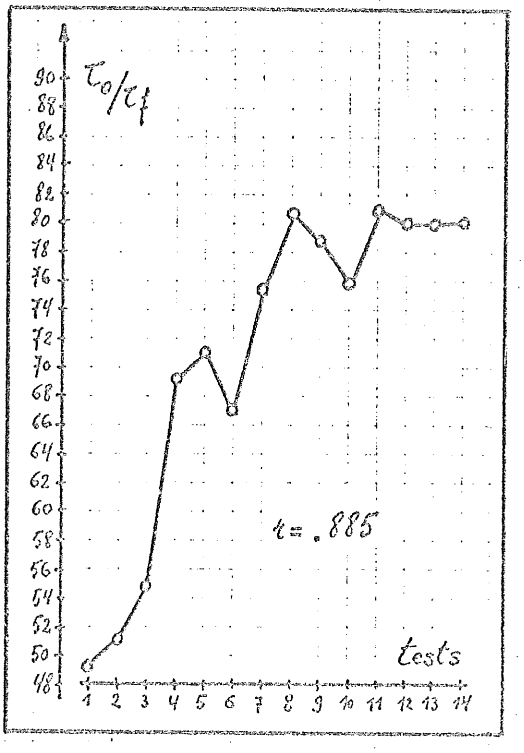
Le tableau 2.41 permet de retracer l'évolution de l'importance relative de l'omission (tableau 2.42) et de la certitude 3 (tableau 2.43).

	OM.	Cert.1	Cert. faibles	% OM	% Cert.1	Cert.2	Cert.3	Cert. fortes	% Cert.3
1	13,65	14,52	28,17	48,45	51,54	22,60	49,21	71,81	68,52
2	14,39	13,72	28,11	51,19	48,81	25,29	46,58	71,87	64,81
3	12,48	10,56	23,04	54,16	45,84	20,96	56,00	76,96	72,76
4	17,52	8,00	25,52	68,65	31,35	15,28	59,20	74,48	79,48
5	18,36	7,62	25,98	70,66	29,34	14,31	59,68	73,99	80,65
6	18,00	9,16	27,16	66,27	33,73	11,50	61,33	72,83	84,20
7	20,88	6,99	27,87	74,91	26,09	14,91	57,20	72,11	79,32
8	21,59	5,33	26,92	80,20	19,80	11,97	61,10	73,07	83,61
9	18,91	5,28	24,18	78,20	22,80	9,77	66,02	75,79	87,10
10	23,39	7,45	30,84	75,84	24,16	13,71	55,43	69,14	80,17
11	25,49	6,29	31,78	80,20	19,80	12,25	55,96	68,21	82,04
12	31,83	8,00	39,83	79,91	20,09	13,30	46,85	60,15	77,88
13	34,50	8,90	43,40	79,49	20,51	12,30	44,30	56,60	78,26
14	32,16	8,00	40,16	80,07	19,93	12,24	47,60	59,84	79,54
Total				71,54	28,46			70,23	

(*) Ce pourcentage est obtenu par la division totale (3665/5402).

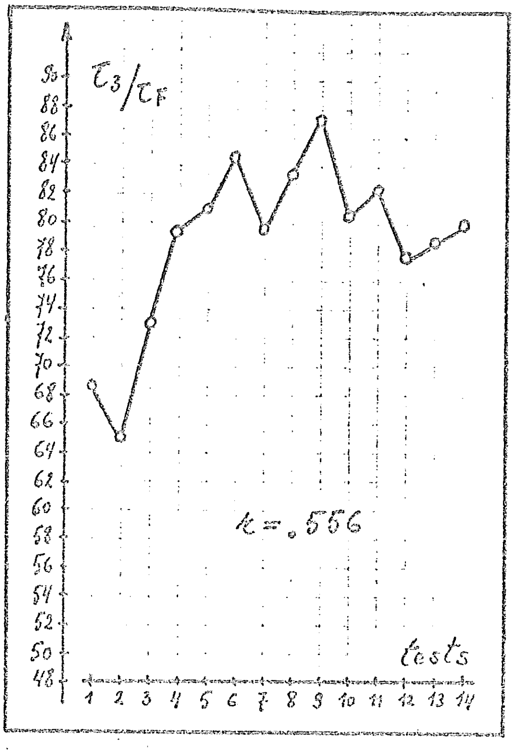
"Les changements dans le débit de réponse ne se laissent jamais facilement décrire. Ils se déroulent nécessairement dans le temps, et même un observateur n'arrive à les voir que s'ils ont été réduits sous une forme graphique."
(SKINNER, 1971, p. 161)

Tableau 2.42



Taux relatifs d'omission (sur les certitudes faibles) au cours des 14 tests de l'année scolaire 1971-1972

Tableau 2.43



Taux relatifs de certitude 3 (sur les certitudes fortes) au cours de l'année scolaire 1971-1972

L'hypothèse H17a est donc confirmée.

HYPOTHÈSE H 17B. L'hypothèse H17a ne se vérifie pleinement qu'après le contrôle de la variable "facilité du test".

Les corrélations brutes sont les suivantes (tableau 2.44) :

Tableau 2.44

	Ordre	% relatif d'OM	% relatif de C3
Ordre	1	.885	.566
Facilité	-.819	-.594	-.111

Corrélations relatives aux 14 tests de l'année scolaire 1971-1972

Les r partielles avec l'ordre sont :

% relatif d'OM : .859 (au lieu de .885)

% relatif de cert. 3 : .803 (au lieu de .556).

L'hypothèse H17b est donc confirmée et montre toute sa pertinence au niveau de la certitude 3.

HYPOTHÈSE H18A. La disparition des certitudes 1 et 2 est décroissante et tend vers 0 pour un nombre de tests suffisamment grand.

TUC_1 vaut 14,52 % lors du premier test et 8 % lors du dernier, avec un minimum (5,28) lors du neuvième test.

TUC_2 vaut 22,6 % lors du premier test et 12,24 % lors du dernier, avec un minimum (9,77) lors du neuvième test.

L'hypothèse H18a ne se confirme pas. Il y a stabilisation après le troisième test et il est vraisemblable qu'une augmentation du nombre de tests ne changerait pas cette constatation.

HYPOTHÈSE H18B. H18a ne se vérifiera qu'après contrôle de la facilité des tests.

Quand on examine l'évolution des taux relatifs d'omissions et de certitude 3, on constate deux choses :

1° L'accroissement le plus important a lieu lors du quatrième test, ce qui confirme à nouveau le nombre de "trois" comme nombre minimal de tests d'exposition à la procédure.

2° La part relative de l'omission et de la certitude 3 se stabilise à 80 % dans les quatre derniers tests.

On peut donc penser que, dans cette expérience, la valeur-plafond (difficile à dépasser) pour ces pourcentages relatifs est située entre 80 et 90 %, contrairement à l'hypothèse H18a selon laquelle les taux relatifs doivent atteindre 100 %. Les hypothèses H18a et H18b doivent donc être infirmées.

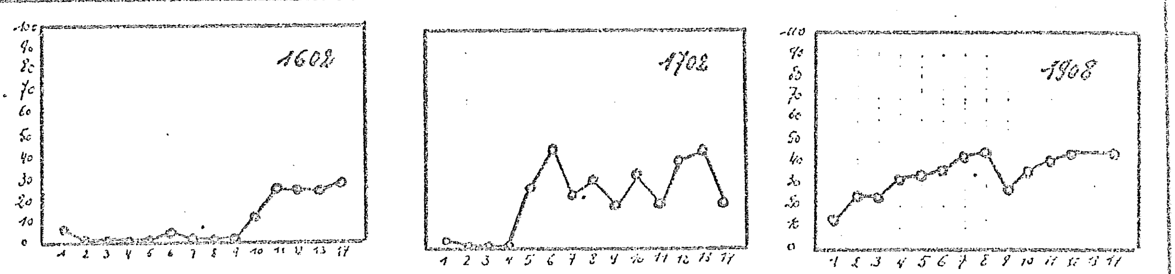
HYPOTHÈSE H 19. Chaque sujet présente une courbe de disparition des certitudes 1 et 2 uniformément décroissante.

Si l'on se base sur les courbes d'évitement enregistrées chez les animaux, on doit, au contraire, s'attendre à des types d'évolution différentes selon les individus. Pour certains, la chute sera brusque (précoce, tardive, ou intermédiaire). Pour d'autres, on observera une décroissance continue. D'autres enfin présenteront une évolution en dents de scie avec tendance à la diminution.

"Une courbe moyenne fournit rarement une image correcte d'aucun des cas sur lesquels elle est fondée." (SIDMAN, 1970)

En fait, on constate des *patterns* d'évolution très différents selon les sujets (tableau 2.45). C'est la raison pour laquelle nous fournissons ci-après l'évolution individuelle des pourcentages d'omission (tableau 2.46).

Tableau 2.45



Modification brusque en fin de la série des tests (sujet 1602)

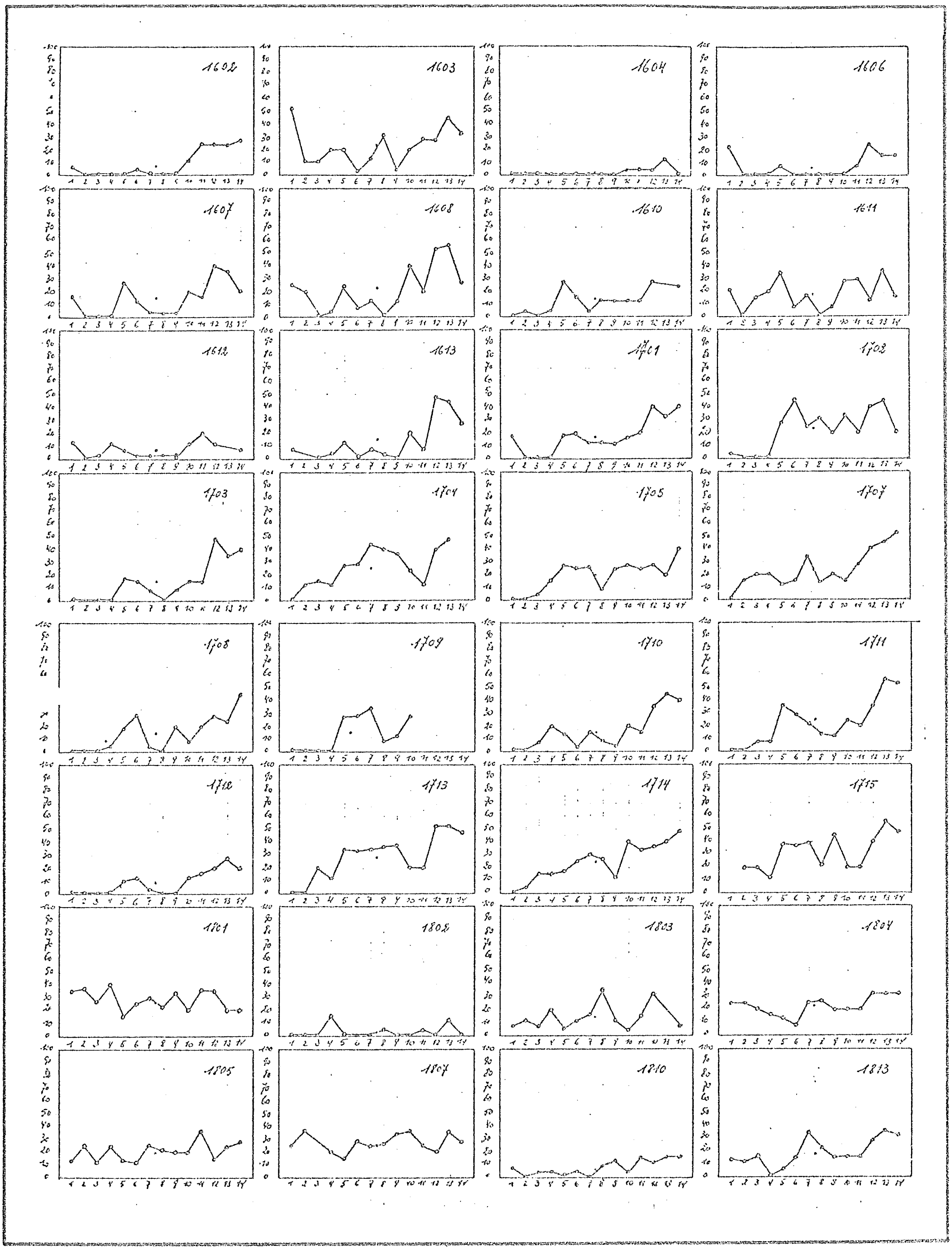
Modification brusque en début de la série des tests (sujet 1702)

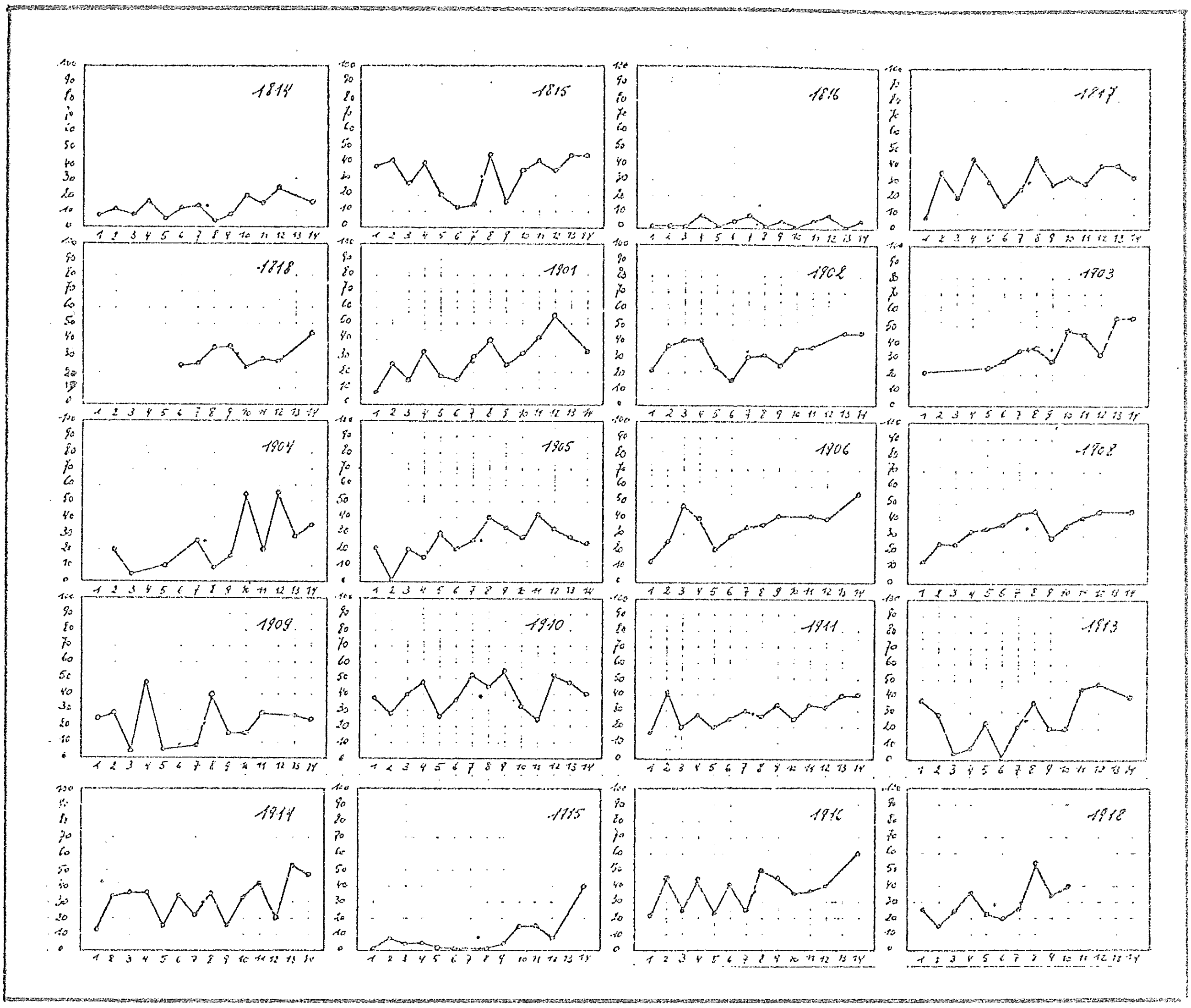
Modification très progressive (sujet 1908)

Trois exemples individuels contrastés d'Evolution des taux d'utilisation de l'omission lors des 14 tests de l'année scolaire 1971-1972

On constatera que trois sujets de la classe 16 (1604, 1606 et 1612) et sept sujets de la classe 18 (1801, 1802, 1803, 1805, 1807, 1815 et 1816), soit 18,8 % des sujets, ne présentent pas d'accroissement des omissions.

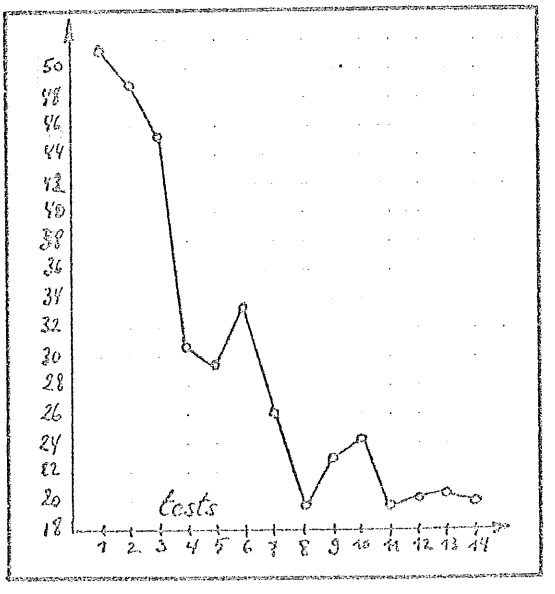
Tableau 2.46


1602
1603
1604
1606
1607
1608
1610
1611
1612
1613
1701
1702
1703
1704
1705
1707
1708
1709
1710
1711
1712
1713
1714
1715
1801
1802
1803
1804
1805
1807
1810
1813


1814
1815
1816
1817
1818
1901
1902
1903
1904
1905
1906
1908
1909
1910
1911
1913
1914
1915
1916
1918

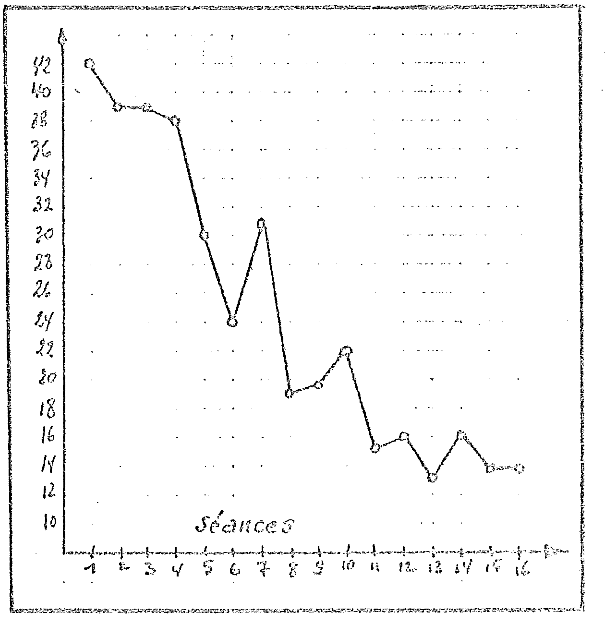
Les courbes relatives aux certitudes 1 et 2 sont très typiques de l'évitement. Ainsi, le graphique du tableau 2.47 représente la part relative de la certitude 1 lors des 14 épreuves successives. Le graphique du tableau 2.48 représente le nombre de chocs reçus (en moyenne par périodes de quinze minutes) par un rat blanc au cours de seize séances successives (1) dans une procédure d'évitement (2).

Tableau 2.47



Evolution du τ_1 (taux d'utilisation de la certitude 1) au cours des 14 tests de l'année scolaire 1971-1972

Tableau 2.48



Evolution du nombre moyen de chocs reçus (par période de 15 minutes) par un rat blanc au cours de 16 séances d'une procédure d'évitement, in A. BEAUJOT, M. DIDELEZ O. FONTAINE et D. LECLERCQ, 1966

Comparer des courbes humaines et animales peut paraître choquant dans le domaine du comportement. Pourtant, personne ne s'étonnerait de voir comparer l'électrocardiogramme d'un homme avec celui d'un animal. Chacun reconnaît que, dans le domaine physiologique, il existe des lois valables aussi bien pour l'homme que pour les animaux.

---

(1) Voir A. BEAUJOT, M. DIDELEZ, O. FONTAINE et D. LECLERCQ, 1966.

(2) Le rat pouvait éviter le choc électrique en se déplaçant dans une cage circulaire.

Le conditionnement operant a mis en évidence des mécanismes communs dans le domaine comportemental. Comme le souligne M. RICHELLE (1970) en réponse au biologiste français CHAUVIN (1969),

> "La confusion capitale (...) consiste à voir dans le conditionnement une catégorie de comportements alors qu'il s'agit d'un mécanisme... Si le mécanisme du conditionnement est, dans son principe, extrêmement simple, il ne s'ensuit pas que les comportements qui en résultent sont simples, eux aussi." (p. 29).

CONCLUSIONS

1. Dans les conditions de la présente expérience, l'ESPER s'est avérée valide sur la base des constatations suivantes :

   a) En ce qui concerne le taux d'utilisation des différents degrés de certitude, on observe :
   - une cohérence globale (H1),
   - une cohérence entre individus (H2),
   - une cohérence entre épreuves (H3).

   b) En ce qui concerne le taux d'exactitude des différents degrés de certitude, on observe :
   - une cohérence globale (H4),
   - une cohérence individuelle (H5),
   - une stabilité entre individus des taux d'exactitude (H6),
   - une cohérence à l'intérieur d'un test (H7),
   - une stabilité entre épreuves (H8).

   c) En ce qui concerne les stratégies présidant aux choix des indices de certitude, on observe que :
   - des stratégies incompatibles avec la validité de l'ESPER ont vraisemblablement été utilisées (surtout le critère minimax et celui du regret minimax), mais
   - en petit nombre,
   - dans les tout premiers tests.
   - Ces stratégies incompatibles ont pratiquement disparu après le troisième test et ont fait place à des *patterns* compatibles avec le seul critère de l'E.S.U. (les critères de la théorie des décisions et ceux de la théorie du dépliage ne pouvant expliquer qu'un nombre très petit de *patterns*).

2. La procédure operante et l'aspect longitudinal de l'expérience ont été déterminants, sur la base des constatations suivantes.

a) Les étudiants ont été soumis à un programme de renforcement sans qu'ils connaissent exactement les contingences en jeu, mais ils ont modifié leur comportement en conséquence. Interrogés, ils pouvaient à peine décrire leur comportement : "J'utilise moins les certitudes 1 et 2", mais ne pouvaient expliquer pourquoi, sauf au travers de déclarations vagues de type : "On y gagne".
Cette constatation est tout à fait normale dans la mesure où "un sujet peut rarement décrire avec précision la manière dont il a été renforcé" (SKINNER, 1971, p. 160). "Les contingences sont néanmoins efficaces même lorsque le sujet n'est pas capable de les décrire." (SKINNER, 1971, p. 173).

b) Les conclusions relatives à la validité de l'ESPER eussent été négatives sur la base des résultats du premier test. Elles eussent été meilleures sur la base des derniers tests. Néanmoins, ce n'est qu'à la lumière de l'évolution comportementale à travers l'ensemble des tests que des conclusions pertinentes peuvent être tirées.

c) Trois tests constituent le nombre minimum de situations auxquelles il faut soumettre les étudiants pour que disparaissent les stratégies non désirables.

3. Les recherches sur l'ESPER peuvent être poursuivies et étendues.

Les contraintes méthodologiques sont lourdes, mais pas insurmontables. Il s'impose en premier lieu de répéter une expérience semblable à celle-ci, mais à partir d'une matrice des conséquences calculées selon le critère de l'E.S.U.

Le chapitre 4 est entièrement consacré à une telle étude.

CHAPITRE 3

COMMENT CALCULER LA MATRICE DE CONSEQUENCES SELON LE PRINCIPE DE L'UTILITE ATTENDUE

A. Le problème

B. Principes et techniques de solution

- *1. La solution graphique*
- *2. L'algorithme de la solution numérique*
- *3. La solution FORTRAN*
- *4. Vérifications par la matrice des utilités attendues*
- *5. Les lieux d'indifférence*

C. Le choix des paramètres générant la matrice

- *1. Principes généraux*
- *2. Faut-il diviser l'axe en parties égales ?*
- *3. Le poids relatif de la certitude*

D. Conclusions

A. LE PROBLEME

Pour une question à choix multiple, on peut décider arbitrairement d'accorder ou de retirer n'importe quel nombre de points pour chacune des solutions proposées. Mais, si l'on veut que l'espérance mathématique soit nulle, la matrice des conséquences doit être calculée selon des règles précises. Contraindre les sujets à estimer leurs PER et à les exprimer sans biais est chose plus complexe que rendre nulle l'espérance mathématique d'une réponse au hasard[1] à une question à choix multiple. Quelles règles précises doit-on suivre pour atteindre ce but ?

Tout d'abord, l'examen des diverses stratégies décrites par la théorie des décisions révèle que si les sujets veulent maximiser leurs scores, ils doivent avoir recours au seul critère de l'espérance subjective de l'utilité (E.S.U.). Pour amener les sujets à vouloir maximiser leurs scores, il faut y associer des conséquences réelles et, si possible, immédiates. Consciemment ou non, les sujets devraient adopter spontanément le critère E.S.U., parce qu'exposés à la réalité de façon répétée. Ces conditions expérimentales de conditionnement operant ont été réalisées dans les expériences des chapitres 2 et 4.

Il reste à amener les sujets à exprimer leur ESPER sans biais. Or il est possible de calculer des matrices de conséquences telles que les sujets ont intérêt à fournir des indices de certitude les plus proches possibles des PER. C'est à la construction de telles matrices qu'est consacré le présent chapitre. Dans une première partie, les écrits les plus décisifs sont explicités car, en version originale, ils ont un caractère mathématique assez ardu, aux dires mêmes de leurs auteurs. Dans une seconde partie, nous proposons des matrices spécialement adaptées à la situation pédagogique et des procédures originales de génération de telles matrices.

---

(1) Plus exactement : "l'espérance mathématique des valeurs des conséquences d'une réponse au hasard."

Le principe mathématique fondamental.

On s'accorde généralement à attribuer à VAN NAERSSEN (1962), DE FINETTI (1965), SHUFFORD, ALBERT et MASSENGILL (1966) les apports décisifs relatifs à la construction des matrices de conséquences (1).

Dans son article "Une échelle pour la mesure de la probabilité subjective" (1965), VAN NAERSSEN présente les raisonnements de base. Nous leur avons donné, dans l'exposé ci-dessous, une signification pédagogique qu'ils n'ont pas dans le texte original.

Admettons qu'un sujet doive choisir, pour nuancer sa réponse, l'une des k catégories de la PER (dans le texte original : p). S'il choisit p_i, il reçoit x_i points en cas de réponse correcte et y_i points en cas de réponse incorrecte. Le problème consiste maintenant à exprimer x et y en fonction de p.

Si p_i est choisi, E.S.U. $= p_i\, x_i + (1-p_i)\, y_i$ (1)

Cette valeur de l'E.S.U. doit être supérieure à l'E.S.U. obtenue au moyen de n'importe quelle autre paire de x et de y, avec le même p_i. Autrement dit (2), $p_i\, x_j + (1-p_i)\, y_j < p_i\, x_i + (1-p_i)\, y_i$ (2)

et $p_j\, x_i + (1-p_j)\, y_i < p_j\, x_j + (1-p_j)\, y_j$ (3)

de (1) et (2), il découle que

$$\frac{p_i}{1-p_i}(x_j-x_i) < y_i-y_j < \frac{p_j}{1-p_j}(x_j-x_i) \qquad (4)$$

En annexe, nous expliquons cette transformation présentée par VAN NAERSSEN. Nous y ajoutons aussi des exemples.

Quand on a décidé des séries $p_1 \ldots p_i \ldots p_k$ et des séries $x_1 \ldots x_i \ldots x_k$, on peut trouver des séries $y_1 \ldots y_j \ldots y_k$ au moyen de cette inégalité (4).

---

(1) On lira aussi avec intérêt les travaux de A.H. MURPHY (1966 à 1969), WINKLER (1967 à 1970) et EPSTEIN (1967 et 1969) centrés sur les prédictions météorologiques.

(2) j = i + 1.

Il peut être utile, ici, de prendre un exemple.

Soit k = 4, $p_1 = .125$, $p_2 = .375$, $p_3 = .625$ et $p_4 = .875$

$x_1 = 1$, $x_2 = 2$, $x_3 = 3$ et $x_4 = 4$.

On aura, pour i = 3 : $\frac{.625}{.375}(4-3) < y_3 - y_4 < \frac{.875}{.125}(4-3)$

dont $1,666 < y_3 - y_4 < 7$

La différence entre yi et yj peut donc prendre n'importe quelle valeur entre 1,66 et 7.

Pour i = 2 : $\frac{.375}{.625}(3-2) < y_2 - y_3 < \frac{.625}{.375}(3-2)$

donc $0,6 < y_2 - y_3 < 1,666$

Pour i = 1 : $\frac{.125}{.875}(2-1) < y_1 - y_2 < \frac{.375}{.625}(2-1)$

donc $0,142 < y_1 - y_2 < 0,6$

On voit qu'il est avantageux de fixer les x avec un incrément de 1, car cela réduit la formule (4) en :

$$\frac{p_i}{1-p_i} < y_i - y_j < \frac{p_j}{1-p_j} \qquad (4')$$

On remarque que $y_i - y_j$ augmente avec i et que

$$y_i - y_j \leqslant y_j - y_l \qquad \text{où } l = j + 1$$

On voit qu'il suffit de fixer $y_1 = 0$, puis de prendre les valeurs inférieures, centrales ou supérieures des intervalles pour obtenir automatiquement des séries y. Dès ce moment, le travail peut être confié à l'ordinateur.

Dans l'exemple ci-dessus, nous aurions obtenu les valeurs que voici :

Valeurs inférieures

i	x	y
1	1	0
2	2	-0,142
3	3	-0,742
4	4	-2,408

Valeurs supérieures

i	x	y
1	1	0
2	2	-0,6
3	3	-2,22
4	4	-9,22

VAN NAERSSEN s'attache alors à rendre les séries de points symétriques. Son raisonnement est le suivant. Si le sujet attribue une probabilité p_i au succès, il recevra x_i points en cas de succès et y_i points en cas d'échec. Mais on peut également dire que le sujet donne une probabilité $1 - p_i$ que la réponse soit fausse (1). Il est dès lors désirable que le sujet, s'il choisit la probabilité $1-p_i$ gagne le même nombre de points x_i s'il a raison.
Autrement dit :

$$x_p = y_{1-p} \quad \text{et} \quad y_p = x_{1-p} \qquad (5)$$

où p = la probabilité subjective (2).

Après intégration (voir VAN NAERSSEN, 1965, p. 162), l'auteur présente la solution la plus simple :

$$x = A - Bq^2 \quad \text{et} \quad y = A - Bp^2 \qquad (6)$$

où A et B sont des constantes arbitraires et $q = 1 - p$.

C'est la solution quadratique.

La construction d'une échelle est facile si p progresse par intervalles égaux. Voici un exemple :

Type 1

p =	05	15	25	35	45	55	65	75	85	95
x =	0	9	17	24	30	35	39	42	44	45
y =	45	44	42	39	35	30	24	17	9	0

où A = 45,125 et B = 50.

Voici deux autres exemples où B = 12,5 :

Type 2

p	0-20	20-40	40-60	60-80	80-100
x	0	4	7	9	10
y	10	9	7	4	0

où A = 10,125.

(1) Nous ne pouvons suivre constamment ce raisonnement de VAN NAERSSEN qui n'est établi que pour les questions à deux solutions strictes (VRAI-FAUX, par exemple).

(2) Probabilité subjective est la terminologie de VAN NAERSSEN. Pour nous, p = ESPER.

Type 3

p	0-20	20-40	40-60	60-80	80-100
x	-7	-3	0	2	3
y	3	3	0	-3	-7

où A = 3,125.

VAN NAERSSEN estime que ce dernier exemple peut être choisi si les ESPER *a priori* du sujet valent 50 %, car, sile sujet donne à sa réponse 50 % d'être correcte et 50 % d'être incorrecte, il n'a rien dit (1) et, par conséquent, il ne gagnera aucun point. EDWARDS (1954 et 1955) a montré que l'utilité du succès n'équivaut pas à l'utilité de l'échec (2), et, dès lors, propose que les x et y soient tous positifs ou tous négatifs. Très précautionneusement, VAN NAERSSEN estime alors qu'il se pourrait que les échelles du type 2 soient parfois préférables aux échelles du type 3, bien que ces dernières soient plus claires et présentent des nombres plus petits.

C'est ici que nous voudrions ajouter aux travaux de VAN NAERSSEN une dimension pédagogique. Les trois types de matrices ci-dessus sont inadmissibles dans le contexte scolaire. Des réponses erronées ne doivent pas pouvoir entraîner des gains de points, or c'est le cas dans les trois matrices (lignes y). De même, on ne peut admettre que des réponses correctes aboutissent à des pertes de points, comme c'est le cas dans la matrice 3 (ligne x).

Il nous paraît qu'en pédagogie :
- L'omission doit être permise et entraîner la note 0.
- Le succès doit entraîner des points positifs.
- L'erreur doit entraîner des points négatifs.

Ces trois exigences ne sont nullement incompatibles avec les principes de VAN NAERSSEN; elles complexifient légèrement le problème.

---

(1) Rappelons que le raisonnement de VAN NAERSSEN porte sur des questions à deux solutions possibles.
(2) Ce qui confirme le principe (théorique) des courbes exposées dans le tableau 1.2.

Dans le fond, ceci est contraire à l'idée d'un continuum

ignorance totale → connaissance parfaite totale

Dans la deuxième partie du présent chapitre, le lecteur trouvera des procédures conformes aux raisonnements de VAN NAERSSEN, mais qui permettent de satisfaire aux trois exigences pédagogiques ci-dessus.

Il existe une solution différente de celle de VAN NAERSSEN, présentée par DE FINETTI dans un autre article important : "Les fondements logiques de la mesure de la probabilité subjective" (1970).

Son raisonnement est le suivant. Imaginons qu'un individu doive estimer la probabilité d'un événement E; p est la probabilité subjective de succès <u>estimée</u> et x la probabilité subjective <u>exprimée</u>. Logiquement, x = p et DE FINETTI va montrer comment on peut donner l'utilité maximale à x = p.

Dans les lignes qui suivent, les raisonnements de DE FINETTI sont développés avec une autre terminologie. Nous écrirons ESPER pour p, la probabilité subjective <u>estimée</u> par le sujet, et C (certitude) pour x, la probabilité subjective <u>exprimée</u> par le sujet (ou telle que l'a recueillie l'observateur). Considérons ici que C (comme ESPER) peut prendre n'importe quelle valeur entre 0 et 1. On pose que E (l'événement) ne peut valoir que 1 (quand il est vrai) ou 0 (quand il est faux).

La technique la plus simple qui oblige à révéler le véritable sentiment à propos d'un événement E consiste, pour DE FINETTI, à donner une punition égale à $(E - C)^2$. Cette punition vaudra $(1 - C)^2$ si E est vrai et $(0 - C)^2$, ou C^2, si E est faux.

Une erreur e entre ESPER et C a une influence sur C^2. En effet, C = ESPER + e, et $C^2 = (ESPER + e)^2 = ESPER^2 + 2e.ESPER + e^2$. Commettre une erreur e revient donc à augmenter C^2 par 2 e ESPER + e^2 (quand E est faux) et à diminuer $(1 - C)^2$ par 2 e (1 - ESPER) − e^2 (1).

Quand C = .50, la punition = .25, que E soit vrai ou faux.

---

(1) Si C = ESPER, $(1 - C)^2 = 1 + ESPER^2 - 2\ ESPER$ (1)

Si C = ESPER + e, $(1 - C)^2 = (1 - (ESPER + e))^2$
$= 1 - 2\ (ESPER + e) + (ESPER + e)^2$
$= 1 - 2\ ESPER - 2\ e + ESPER^2 + 2\ e\ ESPER + e^2$ (2)

Si l'on compare (1) et (2), ils ont en communs $1 + ESPER^2 - 2\ ESPER$

Par conséquent (2) = (1) $- 2\ e + 2\ e\ ESPER + e^2$

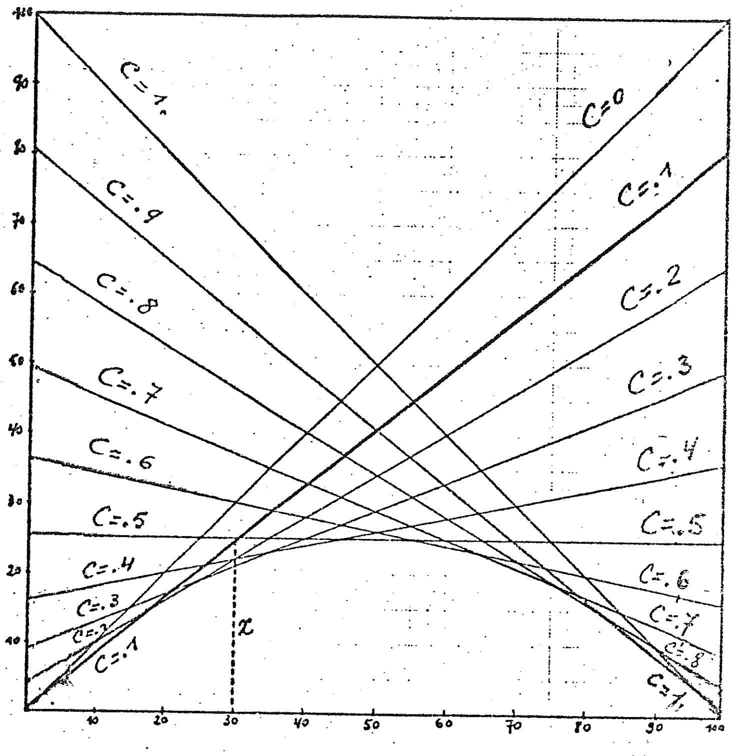
(2) = (1) $- 2\ e\ (1 - ESPER) + e^2$

DE FINETTI présente alors (tableau 3.2) une série de tangentes à une parabole dont l'équation est $Z = p\ (1 - p)$. En abscisse, en effet, on a placé les p ou proportions moyennes de réussite (avec $0 \leqslant p \leqslant 1$). Dans toute la suite des raisonnements, on considérera que ESPER = p. En ordonnée, on place les π ou punitions (avec $0 \leqslant \pi \leqslant 1$). La tangente en ESPER = p à la parabole est la droite $y = p - 2\ C\ p + C^2$.

Tableau 3.1

C	y(p=0)	y(p=1)
.10	.01	.81
.20	.04	.64
.30	.09	.49
.40	.16	.36
.50	.25	.25
.60	.36	.16
.70	.49	.09
.80	.64	.04
.90	.81	.01
1.	100.	0

Tableau 3.2



Graphique dressé à partir des valeurs du tableau 3.1, extrait de DE FINETTI (1970)

Puisque, conventionnellement, ESPER = p, la punition moyenne est

$$\begin{aligned} \pi &= p\ (1 - C)^2 + (1 - p)\ C^2 \\ &= p\ (1 - 2C + C^2) + C^2 - pC^2 \\ &= p - 2\ pC + pC^2 + C^2 - pC^2 \\ &= p - 2pC + C^2 \end{aligned}$$

L'intersection de la verticale en ESPER = p et de la tangente calculée pour C = p est un point dont l'ordonnée est la punition moyenne π.

Le tableau 3.1 fournit les valeurs d'y quand p = 0 et quand p = 1, pour onze valeurs de C : 0, .1, .2, .3, .4, .5, .6, .7, .8, .9 et 1. Ces données permettent de tracer, sur le tableau 3.2, onze tangentes correspondant aux onze valeurs différentes de C.

La pénalisation la plus élevée (π = 1) se produit pour C = 0 et p = 1 et pour C = 1 et p = 0. La pénalisation la plus faible (π = 0) se produit pour C = p = 0 et pour C = p = 1.

Pour C = p = .5, π = .25. Dans un contexte éducatif, où l'utilité attendue de l'omission est nulle, cela serait regrettable, car, si C = p, il ne devrait y avoir aucune pénalisation. On voit que les matrices de DE FINETTI ne peuvent être appliquées telles quelles en pédagogie.

En répondant C ≠ p (ou C ≠ ESPER), on augmente immanquablement la punition. Ainsi, si p = .3 et C = .1, la pénalisation vaut .25 (1) au lieu de .21 (2) quand C = .3 et au lieu de .09 (3) quand C = .1 et p = .1. Cet exemple (C = .1 et p = .3) est représenté (dans le tableau 3.2), par une verticale (π) en traits pointillés qui rencontre la tangente (C=.1)en un point d'abscisse p =.3 . Le point de rencontre est situé au-dessus de la parabole de pénalisation maximale.

Dans les pages qui suivent, une procédure conforme aux principes de VAN NAERSSEN et DE FINETTI est développée. Elle se veut plus directement parlante pour les enseignants.

---

(1) π = .3 - 2 (.3 .1) + .01 = .31 - .06 = .25
(2) π = .3 - 2 (.3 .3) + .09 = .39 - .18 = .21
(3) π = .1 - 2 (.1 .1) + .01 = .11 - .02 = .09

B. PRINCIPES ET TECHNIQUES DE SOLUTION

1. *LA SOLUTION GRAPHIQUE*.

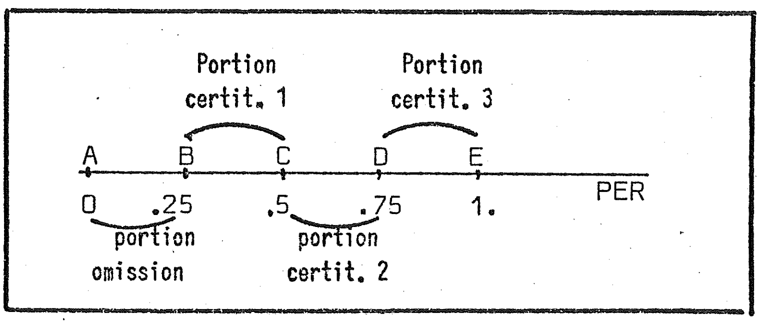
Il importe, par rapport à l'expérience de 1971-1972, de calculer une nouvelle matrice des conséquences telle que les diverses actions disponibles soient utilisées en relation avec l'ampleur de l'ESPER. En d'autres termes, la matrice doit forcer les sujets à répondre la vérité car elle est la plus "payante". Dans le graphique de notre exemple, cela revient à créer quatre droites d'utilité attendue telles que chacune d'entre elles soit maximale pour une portion donnée de l'ESPER. Rappelons que, dans le graphique de l'expérience précédente (1971-1972), deux droites n'étaient jamais maximales : l'utilité attendue pour la certitude 1 et l'utilité attendue pour la certitude 2.

Pour simplifier la procédure, deux conventions de travail sont respectées :

1. On crée sur l'axe des PER autant de portions que d'actions disponibles (omission comprise).
2. Ces portions sont égales entre elles.

Dans l'exemple ci-dessus, puisqu'il y a quatre actions disponibles, l'axe des PER (qui va de 0 à 1) sera divisé en quatre portions (convention 1) égales (convention 2), comme le montre le tableau 3.3.

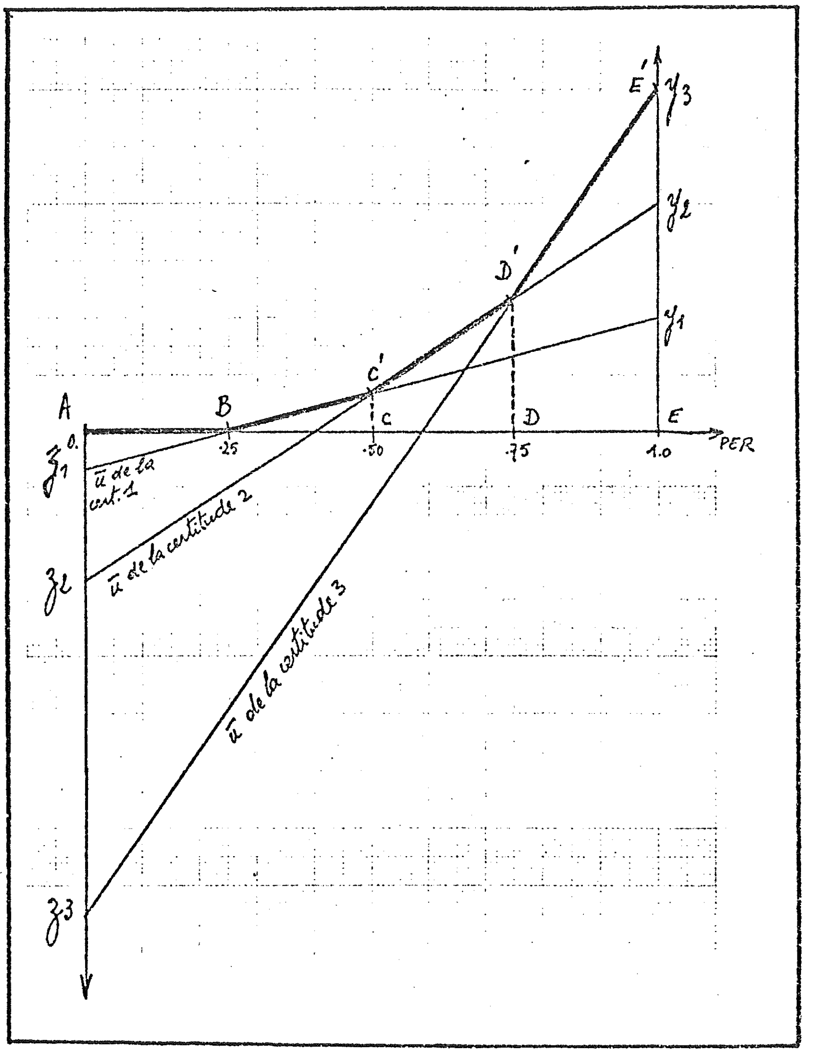
Tableau 3.3



Les conséquences doivent être calculées de manière à obtenir un graphique (1) du type suivant (tableau 3.4).

que signifie z ? (pénalisation)

Tableau 3.4



Les droites de maximisation dans le cas de quatre niveaux de certitude

On remarquera que le segment de droite

$\overline{AB}$ est maximal entre 0 et .25

$\overline{BC'}$ est maximal entre .25 et .5

$\overline{C'D'}$ est maximal entre .5 et .75

$\overline{D'E'}$ est maximal entre .75 et 1.

---

(1) C'est au cours de recherches menées dans le cadre de la banque de questions de l'Ecole Technique de la Force Aérienne à Saffraanberg que cette solution graphique a été mise au point par P. VAN ROY et P. DE SCHUTTER.

pourquoi ? —

Il existe donc des conséquences (1) qui rendent l'utilité attendue d'une réponse supérieure à celle des autres réponses à l'intérieur d'une zone fixée des PER.

Ainsi, c'est l'utilité de l'omission ($\overline{u}$ OM) qui est maximale pour 0 $<$ PER $<$ 25 %. De même :

$\overline{u}$ max = $\overline{u1}$ pour 25 $<$ PER $<$ 50 %

$\overline{u}$ max = $\overline{u2}$ pour 50 $<$ PER $<$ 75 %

$\overline{u}$ max = $\overline{u3}$ pour 75 $<$ PER $<$ 100 %

<u>Comment calculer les conséquences y1, y2, y3, z1, z2 et z3 ?</u>

Avant de présenter une procédure générale, valable pour des nombres variables d'actions disponibles et pour des options variables, examinons comment ces conséquences ont été calculées dans l'exemple qui précède.

<u>Etape 1.</u> On a fixé arbitrairement y2[p = 1.] à + 2 points
y1[p = 1.] à + 1,75 points
y3[p = 1.] à + 2,25 points
y0[p = 1.] à 0 point (omission)

<u>Etape 2.</u> On a dessiné les droites d'utilité attendue :

- la droite 0 se confond avec l'axe des abscisses;
- la droite 1 part du point (1. 1,75) et coupe l'axe des abscisses en B (2);
- la droite 2 part du point (1. 2) et coupe la droite 1 en C' (3);
- la droite 3 part du point (1. 2,25) et coupe la droite 2 en D' (4).

<u>Etape 3.</u> On a ensuite lu les valeurs des ordonnées à l'origine pour chacune de ces droites. Dans l'exemple ci-dessus, ces valeurs sont voisines (5) de -0,5 (droite 1), -0,75 (droite 2) et -1,5 (droite 3).

(1) Ici, ces conséquences sont les points à accorder (y1, y2 et y3) et à retirer (z1, z2 et z3).

(2) Situé à .25 sur l'axe PER.

(3) Situé sur la droite 1 à la verticale de .5.

(4) Situé sur la droite 2 à la verticale de .75.

(5) En fait, elles valent exactement - 0,53, - 0,83 et - 1,57.

La nouvelle matrice des conséquences peut donc s'écrire comme suit (tableau 3.5).

Tableau 3.5

	La réponse est correcte	La réponse est incorrecte
Omission	0	0
Certitude 1	+1,75	-0,5
Certitude 2	+2	-0,75
Certitude 3	+2,25	-1,5
	état favorable	état défavorable

Matrice des conséquences E.S.U. pour $y_0 = 0$, $y_1 = 1,75$, $y_2 = 2$, $y_3 = 2,25$ en $p = 1$.

2. *L'ALGORITHME DE LA SOLUTION NUMERIQUE.*

Nous avons développé une série de procédures expérimentales "jeux de certitudes" permettant d'étudier l'expression de la certitude (Chapitre 5). Dans ces jeux, les actions disponibles sont variables et les matrices de conséquences sont calculées selon certains principes. Ainsi, pour certains jeux, l'omission n'est pas autorisée alors qu'elle l'est toujours dans nos applications en milieu scolaire normal.

Un programme FORTRAN (tableau 3.6) calcule les conséquences et la matrice (1) des utilités attendues ($\bar{u}$), pour des options variables, mais avec des maxima. Rappelons que par "état favorable" il faut entendre : "la réponse est correcte" et par "état défavorable" : "la réponse est incorrecte".

Voici les tâches que doivent effectuer (dans l'ordre) l'expérimentateur et l'ordinateur.

<u>Etape 1</u>. Fixer le nombre (n) d'actions disponibles (max= 10). L'axe des PER sera donc divisé en n parties égales (2).

<u>Etape 2</u>. Décider de permettre ou non l'omission. Dans le premier cas, les conséquences qui sont associées à l'omission valent conventionnellement 0 et 0.

<u>Etape 3</u>. Fixer la différence <u>minimale</u> (D) entre la moins bonne des conséquences (non nulles) dans l'état favorable et la moins mauvaise dans l'état défavorable (éventuellement nulle).

<u>Etape 4</u>. Définir la différence (d) entre les conséquences successives dans l'état favorable (d= le pas ou l'incrément).

<u>Etape 5</u>. Calculer les différentes valeurs des conséquences dans l'état favorable.

<u>Etape 6</u>. Calculer les différentes valeurs des conséquences dans l'état défavorable.

<u>Etape 7</u>. Calculer la matrice des $\bar{u}$.

<u>Etape 8</u>. Vérifier que les $\bar{u}$ maximales correspondent bien aux portions souhaitées sur l'axe des PER.

---

(1) On utilise cette matrice pour vérifier si les $\bar{u}$ max se produisent bien dans la diagonale, comme souhaité.

(2) Il est aussi raisonnable, dans certaines conditions, de ne pas diviser l'axe PER en parties égales. Nous discuterons de ce point en fin du présent chapitre.

Les étapes 1 à 4 sont assurées par le chercheur, puisqu'il s'agit de décisions de principe. Les étapes 5, 6 et 7 sont assurées par un programme FORTRAN, sur base des décisions 1 à 4.

L'étape 8 consiste à interpréter la matrice des $\overline{u}$.

3. *LA SOLUTION FORTRAN.*

L'exemple suivant illustre la procédure. Dans un de nos jeux, les sujets devaient répondre en choisissant parmi dix indices de certitude (allant de 1 à 10). La signification était la suivante :

- Certitude 1 : mon ESPER est comprise entre 0 et .1;
- Certitude 2 : mon ESPER est comprise entre .1 et .2;

⋮

- Certitude 10 : mon ESPER est comprise entre .9 et 1.

Voici un exemple d'application de la procédure :

Etape 1 : nombre d'actions disponibles = 10 Etape 2 : pas d'omission permise Etape 3 : différence minimale = 21 points Etape 4 : incrément = 1 point	Décisions prises par l'expérimentateur

↓

Etapes 5, 6 et 7 prises en charge par le programme FORTRAN (tableau 3.6).

Le tableau 3.7 présente les résultats imprimés (par le programme FORTRAN de l'étape 6 :

Tableau 3.7

VALEURS DES UTILITES PAR REPONSE

	BR	MR	RISQUE	COEFF. A
1	21.00	0.0	21.00	2.10
2	22.00	-0.11	22.11	2.21
3	23.00	-0.36	23.36	2.34
4	24.00	-0.79	24.79	2.48
5	25.00	-1.46	26.46	2.65
6	26.00	-2.46	28.46	2.85
7	27.00	-3.96	30.96	3.1C
8	28.00	-6.29	34.29	3.43
9	29.00	-10.29	39.29	3.93
10	30.00	-19.29	49.29	4.93

Série des résultats fournis par le programme FORTRAN (tableau 3.6)

Les valeurs des conséquences liées à l'état favorable (BR = bonne réponse) sont connues à partir de la différence minimale et de l'incrément.

Les valeurs des conséquences liées à l'état défavorable (MR = mauvaise réponse) résultent, entre autres, du calcul des coefficients de pente (a) des dix différentes droites.

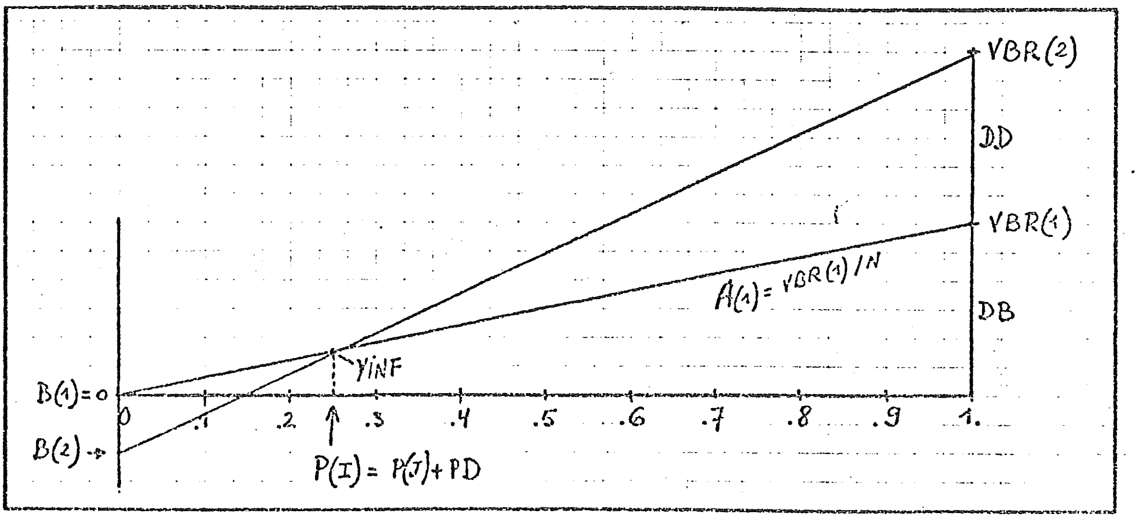
Tableau 3.6

```
C                      ---------------------
C-------------------  JEU DES CERTITUDES  -------------------
C                      ---------------------
C
C       CE PROGRAMME SERT A CALCULER
C                   - LES VALEURS DES UTILITES DES BONNES REPONSES = VBR
C                   - LES VALEURS DES UTILITES DES MAUVAISES REPONSES = VMR
C                   - LE COEFFICIENT DE PENTE DE CHAQUE OBLIQUE = A
C                   - LES AMPLEURS DE RISQUE = R
C        QUAND ON A DETERMINE ARBITRAIREMENT....
C                   ...LE NOMBRE D'INDICES DE CERTITUDE DISPONIBLES = N
C                   ...LA DIFFERENCE DE BASE (GRAND D) = DB
C                   ...L'INCREMENT (PETIT D) = DD
C                   ...LA PROBABILITE DE BASE = PB
C                   ...L'INCREMENT EN PROBABILITE = PD
C
C       LE PRINCIPE DU PROGRAMME CONSISTE A TROUVER LES VALEURS DE B (OU VMR)
C           PAR LA FORMULE  B=Y-AX
C           A PARTIR DE L'EQUATION DE CHAQUE DROITE  Y=AX+B
C                         YINF = POINT D'INTERSECTION DE DEUX OBLIQUES
C
C
      REAL MAT(50,50)
      DIMENSION P(50),Q(50),VBR(50),VMR(50),A(50),B(50),R(50)
    6 READ(5,102,END=5) N,DB,DD,PB,PD
  102 FORMAT(I3,10F5.2)
C
C-----------CALCUL DES DIFFERENTES VALEURS DE VBR ET DE P
      P(1)=PB
      VBR(1)=DB
      DO 7 I=2,N
      J=I-1
      P(I)=P(J)+PD
    7 VBR(I)=VBR(J)+DD
C-------------------------CALCUL DE B(1)
      B(1)=0
      A(1)=VBR(1)/N
C------------------------CALCUL DE B(2) A B(N)
      DO 2 I=2,N
      J=I-1
C----CALCUL DU POINT D'INTERSECTION EN P(I)
      YINF=(A(J)*J)+B(J)
C---- CALCUL DE LA PENTE POUR VBR(I)
      A(I)=(VBR(I)-YINF)/(N-J)
C---- LES VALEURS DE B SE CALCULENT TOUJOURS SUR UN GRAND TRIANGLE(DE BASE N)
    2 B(I)=VBR(I)-(A(I)*N)
C------------------ECRITURE DES UTILITES DES REPONSES ET DES RISQUES
      WRITE(6,600)
  600 FORMAT('1','VALEURS DES UTILITES PAR REPONSE'/'0',9X,'BR',6X,
     1'MR',4X,'RISQUE    COEFF. A')
      DO 1 I=1,N
      VMR(I)=B(I)
      R(I)=VBR(I)-VMR(I)
    1 WRITE(6,601)I,VBR(I),VMR(I),R(I),A(I)
  601 FORMAT('0',I2,4(F9.2))
C-----------------------CALCUL DE LA MATRICE DES UTILITES ATTENDUES
      DO 3 I=1,N
      DO 3 J=1,N
      Q(J)=1-P(J)

    3 MAT(I,J)=(VBR(I)*P(J))+(VMR(I)*Q(J))
C----------------ECRITURE DE LA MATRICE DS UTILITES ATTENDUES
      M=0
   10 M=M+10

      L=M-9
      WRITE(6,602)(P(I),I=L,M)
  602 FORMAT('0','MATRICE DES UTILITES ATTENDUES'/' ',30('-')/,'OPROB.
     1SUBJECTIVE',30X,'PROB. OBJECTIVE'/'0',13X,10F8.2)
      DO 4 I=1,N
    4 WRITE(6,603)I,(MAT(I,J),J=L,M)
  603 FORMAT('0',4X,I3,6X,10F8.2)
      IF(N.GT.M)GOTO 10
      GOTO 6
    5 STOP
      END
```

Programme FORTRAN du calcul des conséquences en cas de réponse incorrecte (z) quand on a décidé des valeurs des conséquences en cas de réponse correcte (y)

Le schéma suivant (tableau 3.6') permet d'interpréter directement les expressions mathématiques du programme FORTRAN (tableau 3.6).

Tableau 3.6'



Variables employées dans le programme FORTRAN du tableau 3.6

4. *VERIFICATIONS PAR LA MATRICE DES UTILITES ATTENDUES ($\overline{u}$).*

A partir de la matrice des conséquences, on peut calculer l'utilité attendue de n'importe quelle réponse quand on connaît la PER.

La formule est simple et s'explique par la théorie des décisions (chapitre 1) :

$$\boxed{\overline{v}_i = \sum_{j=1}^{n} v_{ij}\, p_j}$$

Ainsi, pour une PER de .35, le sujet a intérêt à choisir la certitude 4 (qui signifie : PER comprise entre .3 et .4). En utilisant la certitude 4, l'utilité attendue ($\overline{u}4$) vaut :

$\overline{u}4 = (24 \times .35) + (-.79 \times .65) = 7{,}887$

car 24 = la conséquence en cas de bonne réponse et

-.79= la conséquence en cas de mauvaise réponse.

Le sujet <u>a intérêt à</u> choisir la certitude 4 parce que <u>7,887 est l'utilité attendue maximale</u> ($\overline{u}$ max) <u>quand la probabilité objective vaut .35</u>. En effet, si le sujet utilise la certitude 2, $\overline{u}2 = (22 \times .35) + (-.11 \times .65) = 7{,}628$. Et s'il utilise la certitude 6, $\overline{u}6 = (26 \times .35) + (-2.46 \times .65) = 7{,}503$. On constate que $\overline{u}4 > \overline{u}2$ et que $\overline{u}4 > \overline{u}6$.

Pour une PER de .35, les $\overline{u}$ des dix certitudes disponibles sont les suivantes (tableau 3.8).

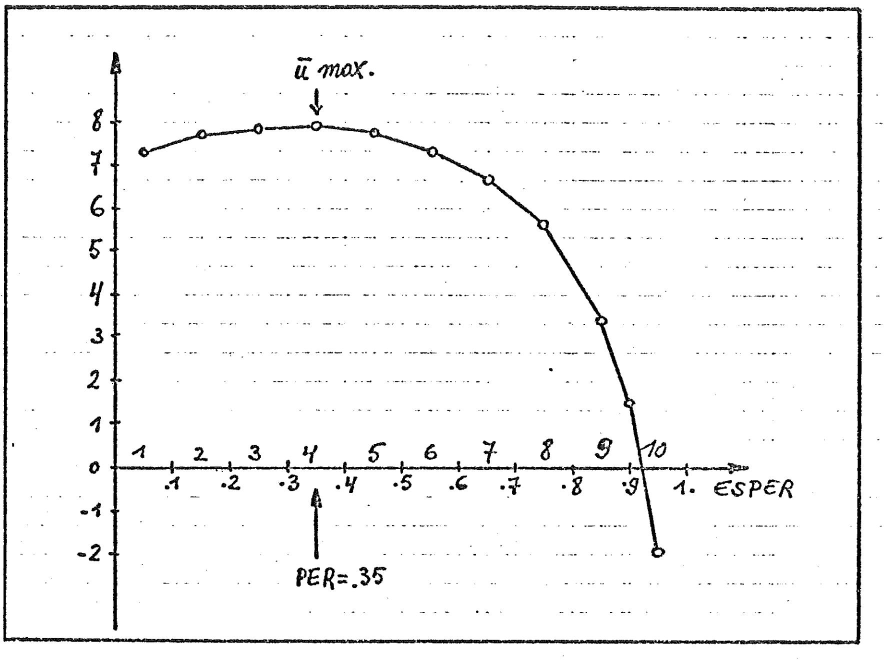
<u>Tableau 3.8</u>

certitude	$\overline{u}$	
1	7.350	
2	7.628	
3	7.815	
4	7.887	$\overline{u}$ max
5	7.803	
6	7.503	
7	6.878	
8	5.712	
9	3.462	
10	-2.038	

Utilités attendues des dix certitudes disponibles de la matrice du tableau 3.7 pour une PER de .35

Cette constatation confirme donc que les valeurs des conséquences ont été bien choisies puisqu'elles maximisent l'utilité attendue pour la certitude qui correspond à la PER, comme le montre le graphique ci-dessous (tableau 3.9).

Tableau 3.9



Représentation graphique des utilités attendues du tableau 3.8

Cette représentation met en évidence que des surestimations importantes (certitudes 8, 9 et 10) entraîneront des conséquences néfastes pour le sujet.

La matrice du tableau 3.10 indique les utilités attendues ($\bar{u}$) (1) des dix actions disponibles (dix ESPER) dans chacune des dix portions de PER (chaque portion est représentée par sa valeur centrale; ainsi, .25 est le centre de la portion de PER qui va de .20 à .29).

---

(1) Dans ces exemples, on fait l'hypothèse de linéarité entre utilité et valeur.

Tableau 3.10

CERTITUDE	PER (μ) .05	.15	.25	.35	.45	.55	.65	.75	.85	.95
1	1.050	3.150	5.250	7.350	9.450	11.550	13.650	15.750	17.850	19.950
2	0.994	3.206	5.417	7.628	9.839	12.050	14.261	16.472	18.683	20.894
3	0.807	3.143	5.479	7.815	10.151	12.488	14.824	17.160	19.496	21.832
4	0.450	2.929	5.408	7.887	10.366	12.845	15.324	17.803	20.282	22.760
5	-0.133	2.512	5.158	7.803	10.449	13.895	15.740	18.386	21.032	23.677
6	-1.033	1.812	4.658	7.503	10.349	13.195	16.040	18.886	21.732	24.577
7	-2.408	0.687	3.783	6.878	9.974	13.070	16.165	19.261	22.357	25.452
8	-4.575	-1.146	2.288	5.712	9.141	12.570	15.999	19.428	22.857	26.286
9	-8.325	-4.396	-0.467	3.462	7.391	11.320	15.249	19.178	23.107	27.036
10	-16.825	-11.896	-6.967	-2.038	2.891	7.820	12.749	17.678	22.607	27.536

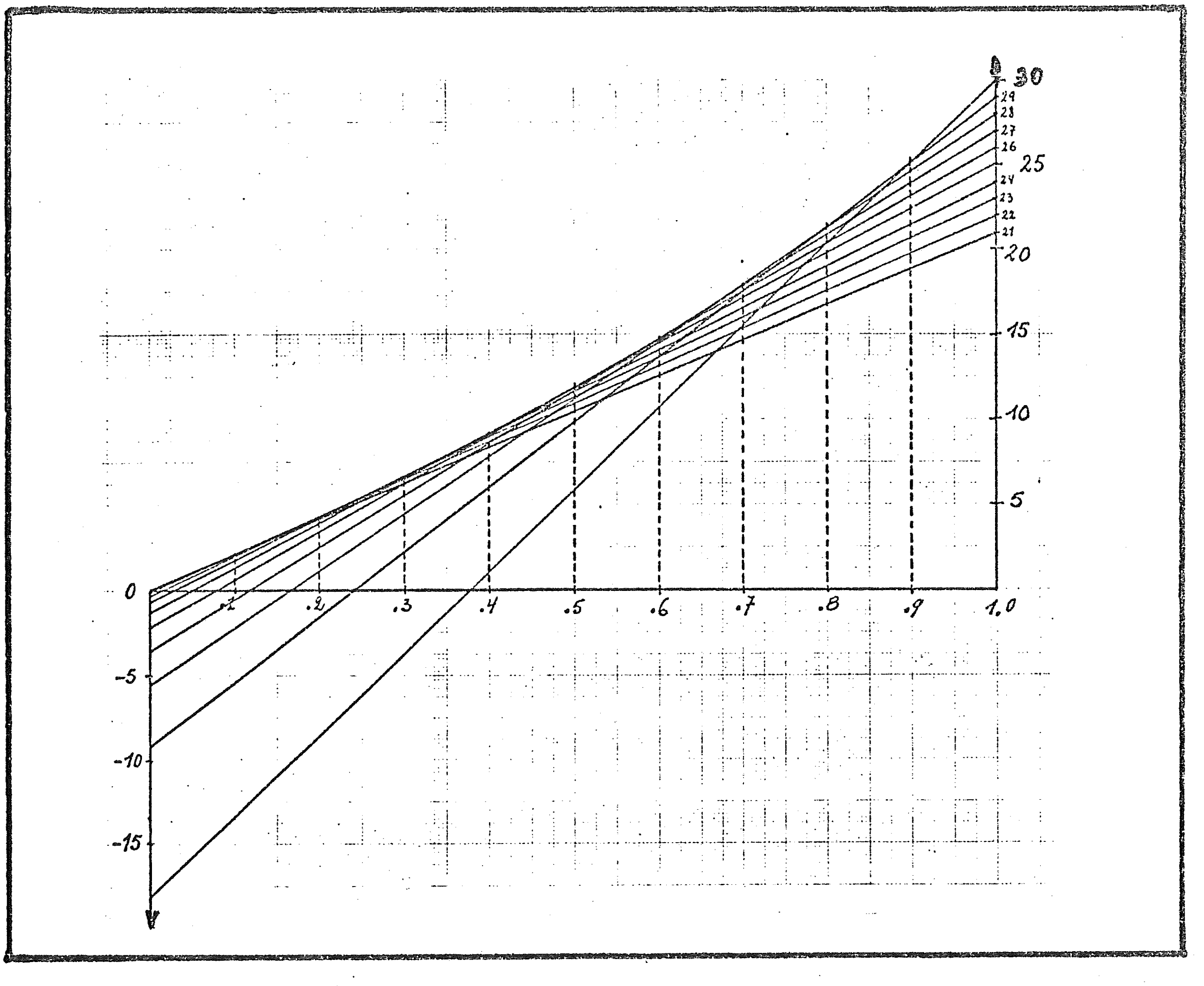
Matrice des valeurs attendues calculée à partir de la matrice des conséquences (tableau 3.7) et des probabilités objectives du centre de chaque zone de certitude

Fournie par le programme FORTRAN à partir de la matrice des conséquences (tableau 3.7), la matrice des utilités (tableau 3.10) permet de vérifier que les $\overline{u}$ max se produisent bien dans la diagonale, comme on le souhaitait. L'opération 8 confirme donc le choix des conséquences.

5. *LES LIEUX D'INDIFFERENCE.*

Dans le tableau 3.11, les dix droites représentent l'utilité attendue de chaque action disponible ($\overline{u}1$, $\overline{u}2$, $\overline{u}3$ ... $\overline{u}10$). Chacun de ces segments est maximal à l'intérieur d'une portion bien délimitée de l'axe PER.

Tableau 3.11



Les dix droites de maximisation calculées à partir de la matrice des conséquences (tabeau 3.7)

Ces segments se coupent en des points dits d'indifférence, puisqu'en ces points, deux actions sont également attractives.

Dans l'exemple décrit ci-dessus, les points d'indifférence se produisent pour les valeurs .1, .2, .3, .4, .5, .6, .7, .8 et .9 de la PER. Pour ces valeurs de PER situées exactement à l'intersection de deux droites, la matrice des $\overline{u}$ des certitudes est la suivante (tableau 3.12).

Tableau 3.12

CERTITUDE	PER (μ) .1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
1	2.10	4.20	6.30	8.40	10.50	12.60	14.70	16.80	18.90	21.00
2	2.10	4.31	6.52	8.73	10.94	13.16	15.37	17.58	19.79	22.00
3	1.98	4.31	6.65	8.98	11.32	13.66	15.99	18.33	20.66	23.00
4	1.69	4.17	6.65	9.13	11.61	14.08	16.56	19.04	21.52	24.00
5	1.19	3.83	6.48	9.13	11.77	14.42	17.06	19.71	22.35	25.00
6	0.39	3.24	6.08	8.93	11.77	14.62	17.46	20.31	23.15	26.00
7	-0.86	2.24	5.33	8.43	11.52	14.62	17.71	20.81	23.90	27.00
8	-2.86	0.57	4.00	7.43	10.86	14.28	17.71	21.14	24.57	28.00
9	-6.36	-2.43	1.50	5.43	9.36	13.28	17.21	21.14	25.07	29.00
10	-14.36	-9.43	-4.50	6.43	5.36	10.28	15.21	20.14	25.07	30.00

Matrice des valeurs attendues calculée à partir de la matrice des conséquences (tableau 3.7) et des probabilités objectives de l'intersection de deux zones de certitude

On remarque deux valeurs maximales identiques par colonne, ce qui illustre bien l'indifférence : pour une probabilité de .4, il est indifférent de choisir la certitude 4 ($\overline{u}$ = 9.13) ou la certitude 5 ($\overline{u}$ = 9.13).

C. LE CHOIX DES PARAMETRES

Dans les exemples qui suivent, seule la "configuration pédagogique" des conséquences est développée. Les conséquences associées à l'état favorable sont toujours positives, les omissions entraînent toujours le score 0 et les points liés à l'état défavorable sont toujours négatifs. Il s'agit, bien entendu, d'une option méthodologique. Le raisonnement reste valable pour des matrices de conséquences telles que l'on gagne toujours, même en omettant, ou telles que l'on perde toujours, même en omettant. De tels exemples n'ont pas été développés car ils ne font que refléter le problème de l'acceptation de tels jeux par le sujet. Les matrices conçues ici sont destinées à des situations scolaires où les sujets seront relativement contraints d'accepter les options. Il importe de rendre ces dernières à la fois acceptables par le sujet et éducatives par leur fonctionnement. C'est cette double préoccupation qui préside à l'établissement de conséquences à "configuration pédagogique".

1. *PRINCIPES GENERAUX*.

Rappelons les options relatives aux paramètres qui sont prises lors des différentes étapes :

Etape 1 : Fixer le nombre (n) d'actions disponibles.

Cette option aura des conséquences sur la finesse des réponses. Dans des situations scolaires, avec des étudiants de 16 à 20 ans, nous avons utilisé trois niveaux de certitude (omission permise). Les résultats montrent que ces étudiants sont capables de discriminer au moins entre ces trois niveaux. Dans des situations à caractère plus expérimental , nous avons proposé à des adultes d'utiliser jusqu'à 40 degrés de certitude (omission non permise). On lira, au chapitre 6, les résultats de ces expériences.

Etape 2 : Permettre (ou non) l'omission.

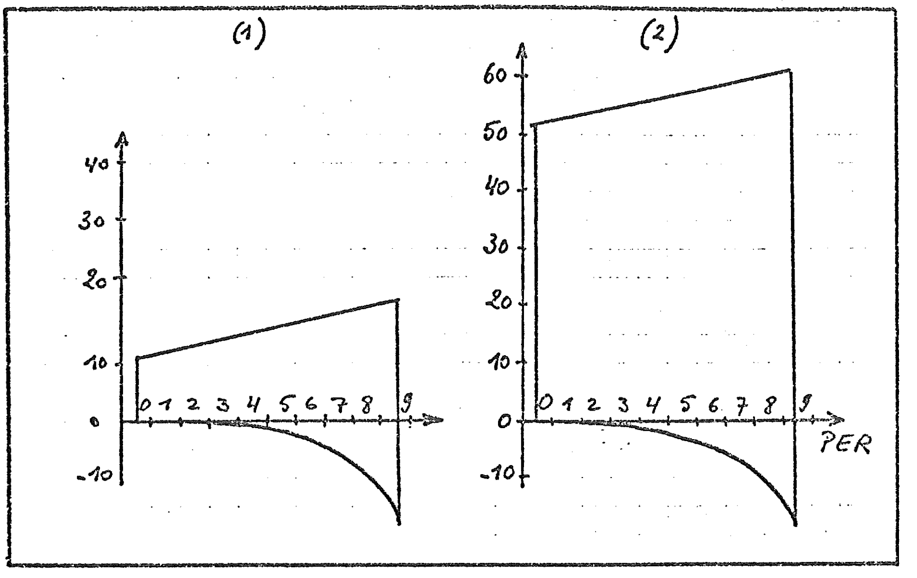
Dans une situation scolaire, l'omission doit être autorisée pour sa valeur formative au niveau des habitudes intellectuelles. Comme elle ne permet ni l'utilisation d'un indice de certitude, ni la vérification de l'exactitude de la réponse, elle constitue une source de données peu utilisables pour l'étude de l'ESPER.

C'est pourquoi, en laboratoire, on préférera ne pas permettre l'omission.

Etape 3 : Fixer la différence (D) minimale entre la moins bonne des conséquences (non nulles) dans l'état favorable et la moins mauvaise des conséquences dans l'état défavorable (éventuellement nulle).

Cette option va fixer l'ampleur minimale du risque qu'aura à considérer le sujet. Si l'omission n'est pas autorisée, le sujet devra courir au moins ce risque pour chaque réponse. Dans le tableau 3.13, on verra mieux la conséquence d'une telle option. Dans le premier cas, on a fixé la différence D à 11 points (1). Dans le second cas, on l'a fixée à 51 points (2).

Tableau 3.13



Ampleurs des risques d'une matrice à 10 degrés de certitude
(n = 10) où d = 1

Dans les deux cas, l'incrément (d) vaut 1, ce qui amène des conséquences défavorables identiques, et des conséquences favorables identiques à une constante près.

Ces deux exemples montrent bien que D est surtout fonction du prix que l'institution de recherche (ou d'enseignement) est prête à engager dans le jeu. Ce prix peut être lui-même fonction de la facilité plus ou moins grande avec laquelle le sujet accepte de jouer au jeu considéré.

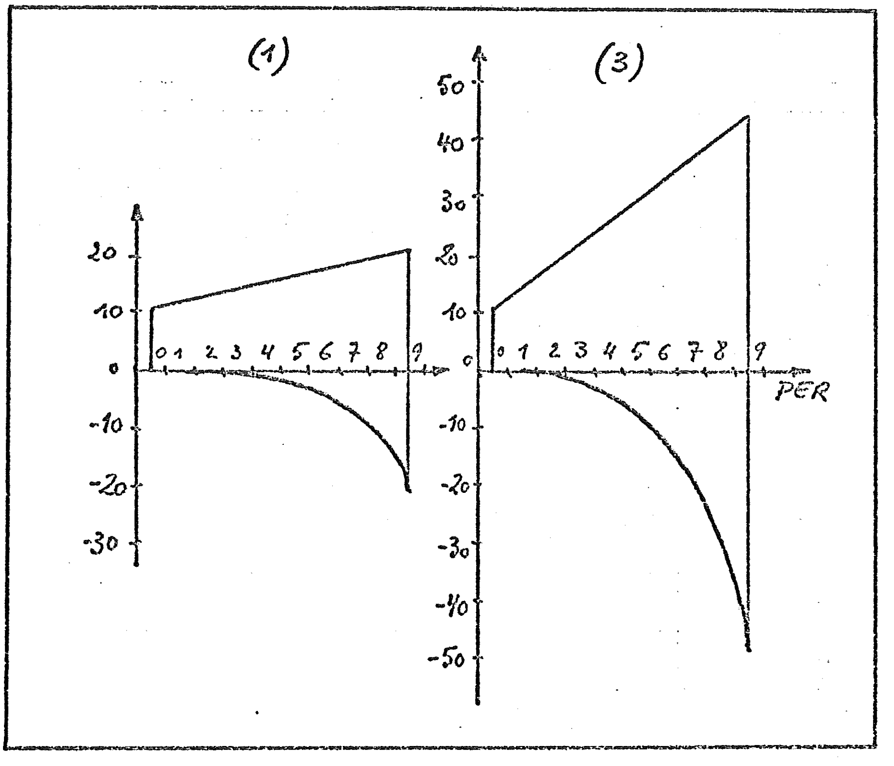
A valeurs égales, la valeur D affecte en outre d'autres paramètres : les rapports de risques. Ainsi, dans le cas 1, l'ampleur du risque couru avec la certitude 1 vaut 11 points et l'ampleur du risque couru avec la certitude 10 vaut 39,29 points. Le rapport entre ces deux risques est : $\frac{r10}{r1} = \frac{39,29}{11} = 3,57$

Puisqu'il s'agit des certitudes extrêmes, cet indice sera appelé "rapport des risques maximal". Par contre, dans le cas 2, $\frac{r10}{r1} = \frac{29,29}{51} = 1,55$. Une telle différence dans les rapports de risque n'est pas négligeable, comme on le verra ci-après.

Etape 5 : Définir la différence (d) entre les conséquences dans l'état favorable. Cette option a une influence prépondérante sur les ampleurs de risques et, par conséquent, sur les rapports de risques.

Dans les deux graphiques du tableau 3.14, D vaut 11 points. Dans le cas 1, d vaut 1 point, tandis que, dans le cas 3, d vaut 5 points.

Tableau 3.14



Ampleurs des risques pour n = 10 et D = 11

Dans le cas 1,

$$\frac{r10}{r1} = \frac{39,29}{11} = 3,57$$

Dans le cas 3,

$$\frac{r10}{r1} = \frac{143}{11} = 13$$

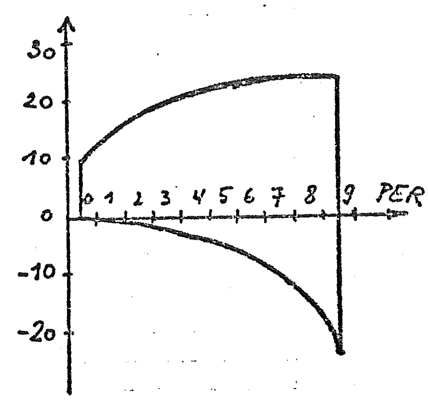
On voit que d affecte également le prix que l'institution de recherche (ou d'enseignement) est prête à engager dans le jeu. Néanmoins, ce sont essentiellement les rapports de risques qui sont affectés par d.

Dans l'exemple ci-dessus, d est constant. Rien n'empêche cependant d'employer pour d une fonction du type "inverse" ou logarithmique, ou exponentielle, etc.

Les tableaux 3.15 et 3.16 donnent de tels exemples.

Tableau 3.15

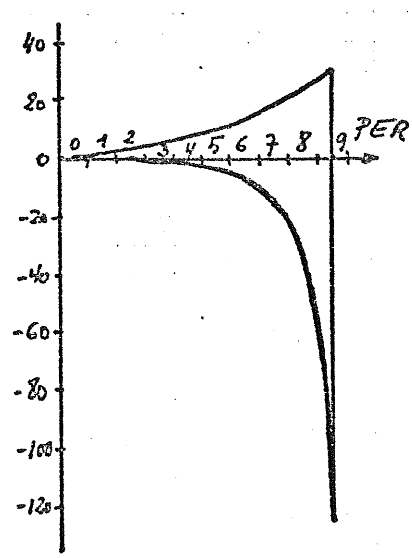
Cert.	BR	MR
1	10.0	0.0
2	15.0	-0.56
3	18.3	-1.38
4	20.5	-2.32
5	22.5	-3.66
6	24.2	-5.36
7	25.6	-7.46
8	26.9	-10.49
9	28.0	-14.89
10	29.0	-23.89



Progression inverse où n = 10, D = 10, d(i) = D/i
i = degré de certitude - Rapport de risques maximal = 5,19

Tableau 3.16

Cert.	BR	MR
1	1.	0.0
2	2.1	-0.12
3	3.3	-0.42
4	4.8	-1.07
5	6.5	-2.2
6	8.5	-4.2
7	10.7	-7.5
8	14.0	-15.2
9	19.0	-35.2
10	29.0	-125.2



Progression où n = 10, D = 1, d(i) = 10/ ((n + 1) - i)
Rapport de risques maximal = 154, 2.

Cette quatrième et dernière option (fixer les valeurs de d) prend toute son importance dans l'ampleur du risque et, notamment, dans les expériences destinées à tester la théorie du dépliage de C.H. COOMBS (voir chapitre 1).

Enfin, on peut se demander comment calculer le rapport des risques maximal quand l'omission est permise.

2. *FAUT-IL DIVISER L'AXE EN PARTIES EGALES ?*

Toutes les considérations qui précèdent concernent les réponses ouvertes : jamais le cas des questions à choix multiple n'a été considéré. Or ce cas pose un problème supplémentaire. En effet, si le sujet a l'assurance (par la consigne) que la réponse correcte est, par exemple, l'une des trois solutions proposées (1), son ESPER est "automatiquement" portée à .33 (en moyenne) avant même d'avoir lu la question. On voit que, dans une telle situation, il n'a aucun intérêt à omettre si l'omission correspond à $0 < \text{ESPER} < .25$. Avec une question à deux solutions (oui/non ou vrai/faux), l'ESPER "automatique" est portée à .50, ce qui supprime presque l'utilisation de la certitude 1 (où $.25 < \text{ESPER} < .50$).

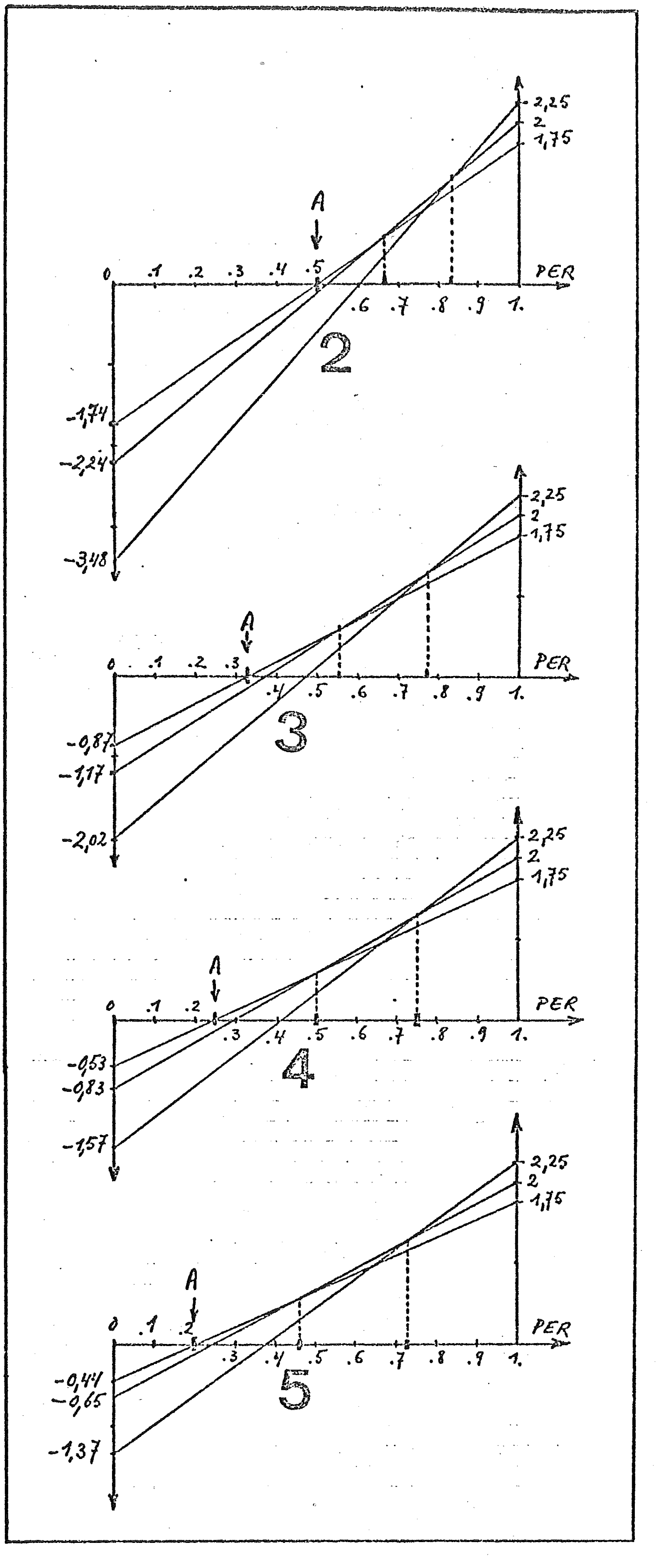
Pour parer à ce problème, on peut porter sur l'axe des PER cette valeur "automatique", puis diviser le reste de l'axe en autant de parties égales que de degrés de certitude (2). Les exemples présentés dans le tableau 3.17 ne considèrent que le cas d'une échelle de certitude à quatre degrés (omission + 3 indices de certitude). Les quatre "segments d'utilité maximale" apparaissent dans le tableau 3.17, pour les questions à choix multiple présentant respectivement 2, 3, 4 et 5 solutions strictes.

---

(1) On dit aussi : Question à choix multiple à trois solutions strictes, précisant ainsi que le rejet des solutions proposées n'est pas une réponse autorisée.

(2) C'est cette option qu'a adoptée depuis 1974, le commandant P. VAN ROY, pour la banque de questions de l'Ecole Technique de la Force Aérienne. Les coefficients (tableau 3.17) ont été calculés par son service.

Tableau 3.17



Droites d'utilité maximale pour n = 4, dans les cas de 2, 3, 4 et 5 solutions proposées, établies selon le principe de la PER automatique

Cette procédure est séduisante, mais comporte des désavantages. En effet, un indice de certitude donné a ainsi des limites différentes (vers le haut et vers le bas) sur l'axe des PER, selon le nombre de solutions proposées. On ne voit pas, dès lors, comment rédiger la consigne qui, en toute rigueur, devrait se présenter comme suit :

"Répondez avec la certitude 1 si votre ESPER est comprise entre
.50 et .66 quand deux solutions sont proposées
.33 et .55 quand trois solutions sont proposées
.25 et .50 quand quatre solutions sont proposées
.20 et .466 quand cinq solutions sont proposées
.166 et .444 quand six solutions sont proposées."

et

"Répondez avec la certitude 2 si votre ESPER est comprise...(etc.)."

Une consigne ordinale évite cette difficulté, mais elle enlève à l'ESPER une part de sa fécondité potentielle, puisqu'elle est contraire à sa validité.

Les raisonnements ci-dessus concernent le cas où le sujet est assuré de trouver la bonne réponse parmi les solutions proposées. On peut à l'opposé avertir les élèves que, dans certaines questions, aucune solution proposée n'est correcte, et construire un nombre raisonnable (1) de ces questions. Dans une telle perspective, les "valeurs automatiques" doivent être calculées pour n+1 et non pour n.

Enfin, la notion de valeurs automatiques de la PER découle d'un modèle fort simple et incorrect de l'activité mentale de l'élève. Ce modèle a été critiqué à plusieurs reprises (B. DE FINETTI, 1965; B. CHOPPIN, 1971; D. LECLERCQ, 1973). Dans ce modèle élémentaire, on confond l'attirance théorique et l'attirance réelle des solutions d'une question à choix multiple. Les exemples concernant la Malaisie fournis dans l'introduction ou les questions 8 et 10 d'une expérience du chapitre 7 montrent quelles différences peuvent exister entre ces deux types d'attirance. L'ESPER combine l'ensemble des informations

---

(1) Si on avertit les élèves de cette possibilité, mais que la banque de questions ne contient, en fait, aucune question de ce type, les élèves auront tôt fait de négliger la consigne pour s'en tenir aux faits.

d'un sujet sur la solution choisie (et l'attirance théorique ne représente qu'une de ces informations). Le recours à des indices de certitude permet, entre autres, de résoudre d'une façon plus générale le problème des choix aléatoires que les *corrections for guessing* classiques.

3. *LE POIDS RELATIF DE LA CERTITUDE.*

Dans l'introduction, trois matrices contrastées ont été présentées (tableau 3.18). Appelons barèmes A', B' et C' les matrices A, B et C mais amputées de la ligne OM (omissions).

Tableau 3.18

	Barème A BR	Barème A MR	Barème B BR	Barème B MR	Barème C BR	Barème C MR
OM	0	0	0	0	0	0
Cert.1	+1,9	-1,9	+ 1	- 1	+0,1	-0,1
Cert.2	+2	-2	+ 2	- 2	+ 2	- 2
Cert.3	+2,1	-2,1	+ 3	- 3	+3,9	-3,9

Trois barèmes de conséquences (A, B et C) constrastés quant à l'importance relative de la certitude. Les barêmes A', B' et C' sont, simplement, les barêmes A, B et C amputés de l'omission

A simple vue, on peut se rendre compte que la certitude a un poids (relatif) croissant, du barème A au barème C. Ce poids relatif peut être exprimé sous la forme du modèle général d'analyse de la variance

Variance totale (SCT) = variance due à la réponse (SCR) correcte ou fausse.
+ variance due à la certitude (SCC).
+ variance d'interaction R $\times$ C (SCI).
+ variance d'erreur due aux répétitions (SCE).

où SC = somme des carrés (des écarts à la moyenne).

La SCE vaudra 0 dans le modèle théorique puisqu'on considère que la matrice est orthogonale et que l'attribution des points est objective et automatique.

On aura, pour chacun des six barèmes, la décomposition suivante de la variance (tableau 3.19).

Tableau 3.19

Barème	SCT	= SCR	+ SCC	+ SCI	+ SCE
A	24,04	18	0	6,04	0
A'	24,04	24	0	0,04	0
B	28	18	0	10	0
B'	28	24	0	4	0
C	38,44	18	0	20,44	0
C'	38,44	24	0	14,44	0

Analyse de la variance des six barêmes du tableau 3.18

SCC vaut 0. Cela n'est pas étonnant, puisque l'impact de la certitude est conditionnel à la qualité de la réponse. La variance introduite par l'indice de certitude se retrouve donc au seul niveau de l'interaction : "qualité de la réponse x certitude".

Les valeurs du tableau 3.19 permettent de calculer les pourcentages de variance due à la qualité de la réponse (SCR) et à la certitude (SCC + SCI) (tableau 3.20).

Tableau 3.20

Barème	% de variance due à		Corrélation théorique
	la réponse	la certitude	
A (om.)	.7487	.2513	.865
A'	.9983	.0017	.999
B (om.)	.6428	.3571	.802
B'	.8570	.1430	.927
C (om.)	.4682	.5317	.684
C'	.6243	.3757	.790

Expression du tableau 3.19 en termes de pourcentages de variance épuisée

Dans le tableau 3.20, la colonne de droite (corrélation théorique) représente la racine carrée du pourcentage de variance due à la réponse. On peut considérer ce pourcentage comme un "coefficient de détermination", soit le carré d'un coefficient de corrélation. La corrélation qu'il s'agit d'estimer est celle qui lie le score simple (nombre de réponses correctes) au score calculé "avec indices de certitude". On voit que les corrélations théoriques sont élevées (tableau 3.20).

Le professeur Ch. Heuchenne nous a suggéré que SCC $>$ 0 dans les cas où les conséquences (Y) des mauvaises réponses sont des fonctions plus complexes des conséquences (X) des bonnes réponses. Dans les six barèmes du tableau 3.18, on a la fonction la plus simple à imaginer : Y (i) = - X (i).

Voici d'autres barèmes (tableau 3.21).

Tableau 3.21

	Barème D BR	Barème D MR	Barème E BR	Barème E MR	Barème F BR	Barème F MR
OM	0	0	0	0	0	0
1	+ 1	0	+ 1	- 2	+ 1	- 1
2	+ 2	-1	+ 2	- 4	+ 2	- 4
3	+ 3	-2	+ 3	- 6	+ 3	- 9

Barème D	Barème E	Barème F
$Y(i) = -(X(i)-1)$	$Y(i) = -2\,X(i)$	$Y(i) = -(X(i))^2$
SCT = 17,875	SCT = 64	SCT = 101,5
SCR = 10,125 (56,6 %)	SCR = 54 (84,4 %)	SCR = 65,76 (64,8 %)
SCC = 0,375 (2,1 %)	SCC = 1 (1,6 %)	SCC = 9,5 (9,4 %)
SCRC = 7,375 (41,3 %)	SCRC = 9 (14,0 %)	SCRC = 26, 24 (25,8 %)
SCE = 0	SCE = 0	SCE = 0

Trois barêmes de conséquences (où SCC $\neq$ 0) et leur analyse correspondante

On remarque en effet que SCC $>$ 0 pour les barèmes D, E et F.

Le tableau 3.22 présente un programme FORTRAN destiné à opérer ces analyses, dont on trouvera un exemple dans le tableau 3.23.

Les corrélations entre scores simples (SS), proportionnels au nombre de bonnes réponses, et scores calculés selon la matrice de conséquences sont calculées par un programme FORTRAN (tableau 3.24) et présentées dans le tableau 3.25.

Tableau 3.22

```
C      CE PROGRAMME CALCULE LA PART DE VARIANCE DUE A LA CERTITUDE DANS UNE
C                    MATRICE DE CONSEQUENCES
       DIMENSION X(2,10)
       REAL M,MR(2),MC(10)
     1 READ(5,500,END=3) N,((X(I,J),I=1,2),J=1,N)
   500 FORMAT(4X,I1,8F5.2)
C--------- CALCUL DE LA MOYENNE GENERALE
       T=0
       DO 2 I=1,2
       DO 2 J=1,N
     2 T=T+X(I,J)
       M=T/(N*2)
C------CALCUL DE LA MOYENNE DE CHAQUE COLONNE (REPONSES)
       MR(1)=0
       MR(2)=0
       DO 13 I=1,2
       DO 13 J=1,N
    13 MR(I)=MR(I)+X(I,J)
       MR(1)=MR(1)/N
       MR(2)=MR(2)/N
C------CALCUL DE LA MOYENNE DE CHAQUE LIGNE (CERTITUDES)
       DO 4 J=1,N
     4 MC(J)=0
       DO 5 J=1,N
       DO 5 I=1,2
     5 MC(J)=MC(J)+X(I,J)
       DO 6 J=1,N
     6 MC(J)=MC(J)/2
C--------CALCUL DE LA VARIANCE TOTALE
       ST=0
       DO 7 I=1,2
       DO 7 J=1,N
     7 ST=ST+(X(I,J)-M)**2
C-----CALCUL DE LA VARIANCE DUE AUX REPONSES
       SR=0
       DO 8 I=1,2
     8 SR=SR+((MR(I)-M)**2)*N
C-------CALCUL DE LA VARIANCE DUE AUX CERTITUDES
       SC=0
       DO 9 J=1,N
     9 SC=SC+((MC(J)-M)**2)*2
C------ CALCUL DE LA VARIANCE D'INTERACTION
       SI=0
       DO 10 I=1,2
       DO 10 J=1,N
    10 SI=SI+(X(I,J)-MC(J)-MR(I)+M)**2
C-----CALCUL DES POURCENTAGES DE VARIANCE
       PR=SR/ST
       PC=SC/ST
       PI=SI/ST
C------- ECRITURE DES RESULTATS
       WRITE(6,601)
   601 FORMAT('0'///'  BR      MR')
       DO 11 J=1,N
    11 WRITE(6,600)(X(I,J),I=1,2),MC(J)
   600 FORMAT(3(F5.2,2X))
       WRITE(6,600) MR(1),MR(2)
       WRITE(6,604) M
   604 FORMAT('0','M = ',F10.4)
       WRITE(6,602)
   602 FORMAT('0',30X,'SOMME DES       POURCENTAGE'/35X,'CARRES          DE
      1VARIANCE')
       WRITE(6,603) ST,SR,PR,SC,PC,SI,PI
   603 FORMAT('OVARIANCE TOTALE ',14X,F10.4,9X,'  1'/'OVARIANCE DUE AUX R
      1EPONSES     ',F10.4,6X,F10.4/'OVARIANCE DUE AUX CERTITUDES  ',
      2F10.4,6X,F10.4/'OVARIANCE DUE A L INTERACTION ',F10.4,6X,F10.4)
       GOTO 1
     3 STOP
       END
```

Tableau 3.23

BR	MR	
0.0	0.0	0.0
1.00	-1.00	0.0
2.00	-2.00	0.0
3.00	-5.00	-1.0
1.50	-2.00	

M = -0.2500

	SOMME DES CARRES	POURCENTAGE DE VARIANCE
VARIANCE TOTALE	43.5000	1
VARIANCE DUE AUX REPONSES	24.5000	0.5632
VARIANCE DUE AUX CERTITUDES	1.5000	0.0344
VARIANCE DUE A L INTERACTION	17.5000	0.4023

Tableau 3.24

```
      DIMENSION X(73),V(7,73),Y(73),W(7,73),SW(7),SWY(7),COR(7)
      REAL MY,MW(7)
   13 K=0
      DO 1 J=1,73
      READ(5,500)X(J),(V(I,J),I=1,7),L
  500 FORMAT(8F9.2,6X,I2)
C------------NETTOYAGE POUR ABSENCES
      IF(X(J))1,1,5
    5 K=K+1
      Y(K)=X(J)
      DO 6 I=1,7
    6 W(I,K)=V(I,J)
    1 CONTINUE
      NK=K
C-------CALCUL DE M,S,COV ET COR
      MY=0
      SY=0
      DO 11 I=1,7
      MW(I)=0
      SW(I)=0
   11 SWY(I)=0
      DO 7 K=1,NK
      MY=MY+Y(K)
      DO 7 I=1,7
    7 MW(I)=MW(I)+W(I,K)
      MY=MY/NK
      DO 8 I=1,7
    8 MW(I)=MW(I)/NK
      DO 9 K=1,NK
      SY=SY+((Y(K)-MY)**2)
      DO 10 I=1,7
      SW(I)=SW(I)+(W(I,K)-MW(I))**2
   10 SWY(I)=SWY(I)+(Y(K)-MY)*(W(I,K)-MW(I))
    9 CONTINUE
      DO 15 I=1,7
   15 COR(I)=SWY(I)/SQRT(SW(I)*SY)
      WRITE(6,600)L,NK,MY,(MW(I),I=1,7)
      WRITE(6,601)SY,(SW(I),I=1,7)
      WRITE(6,602)(SWY(I),I=1,7)
      WRITE(6,602)(COR(I),I=1,7)
  600 FORMAT('0',2I3,8(F10.3,2X))
  601 FORMAT(' ',6X,8(F10.3,2X))
  602 FORMAT(' ',18X,7(F10.3,2X))
      GOTO 13
   14 STOP
      END
```

Programme FORTRAN qui calcule les corrélations (COR) entre les scores simples (X) et les scores avec certitude (V) calculés selon six matrices de conséquences différentes. L'utilisation des barèmes A,B,C,D,E et F a donné le tableau 3.25.

Tableau 3.25

1	50	14.240	18.090	25.740	33.390	31.360	14.760	-5.800
		537.116	8834.723	11159.566	14628.914	6663.473	22367.031	67169.875
			1954.215	2144.115	2334.014	1714.675	2825.874	4431.574
			0.897	0.876	0.833	0.906	0.815	0.738
2	50	12.120	7.950	13.620	19.290	22.080	-4.100	-30.260
		327.278	3456.739	5095.746	7937.609	3667.674	10556.445	37003.531
			784.996	1009.276	1233.556	956.516	1087.597	1291.555
			0.738	0.782	0.765	0.873	0.585	0.371
3	50	14.780	16.172	23.480	30.788	30.580	7.660	-14.580
		724.577	8565.289	12996.422	19153.309	9012.141	24835.125	55020.074
			2280.687	2764.275	3247.863	2375.375	3486.254	4975.586
			0.915	0.901	0.872	0.930	0.822	0.788
4	50	14.440	17.349	24.800	32.252	30.980	10.180	-9.580
		790.316	7026.641	12109.949	19147.098	9288.922	20031.328	42006.090
			2120.439	2767.395	3414.351	2533.434	3171.034	3945.754
			0.900	0.895	0.878	0.935	0.797	0.685
5	52	30.038	41.258	57.423	73.588	67.731	32.923	-0.519
		1647.922	17364.629	33958.609	57680.461	25732.152	59575.594	131156.062
			4834.766	6758.121	8681.480	6095.520	8145.109	10672.992
			0.904	0.903	0.890	0.936	0.822	0.726
6	50	12.540	11.492	17.720	23.948	24.860	0.640	-23.820
		464.417	4751.832	8244.039	13358.957	6127.988	16723.445	37781.285
			1310.911	1711.555	2112.199	1542.775	2000.715	2784.135
			0.882	0.875	0.848	0.915	0.718	0.665
7	51	11.961	12.104	17.980	23.857	24.216	3.300	-14.882
		697.918	6830.852	11720.906	18442.727	8950.574	20698.000	34045.187
			2038.002	2695.955	3353.907	2411.426	3321.995	3813.229
			0.933	0.943	0.935	0.965	0.874	0.782
8	52	11.077	10.365	15.385	20.404	21.558	-0.308	-18.385
		651.689	5966.961	9950.250	15580.590	7498.789	19568.973	36660.219
			1704.833	2209.455	2714.079	2036.763	2559.225	2901.534
			0.865	0.868	0.852	0.921	0.717	0.594
9	52	14.481	19.896	27.962	36.027	32.942	15.192	-0.212
		994.975	8948.156	17537.836	29273.437	13926.742	29655.977	49208.551
			2850.690	3981.954	5113.203	3614.437	4905.168	5963.250
			0.955	0.953	0.947	0.971	0.903	0.852
10	50	11.660	8.462	12.620	16.778	20.280	-6.440	-29.760
		293.218	4753.172	7967.715	12413.133	5044.043	20010.215	42091.027
			977.749	1227.534	1477.321	1045.755	1651.514	2177.075
			0.828	0.803	0.774	0.860	0.682	0.620
11	51	9.824	3.898	7.216	10.533	15.275	-13.020	-37.020
		1085.406	13498.094	24614.523	39581.691	17812.059	49504.887	86614.500
			3667.275	4940.914	6214.570	4263.461	6648.789	8405.770
			0.958	0.956	0.948	0.970	0.907	0.867
12	49	9.286	3.396	6.408	9.420	14.163	-12.286	-37.102
		453.998	5793.008	10027.742	15919.984	7050.660	21063.898	44698.383
			1359.352	1865.280	2371.208	1667.708	2350.993	2825.423
			0.838	0.874	0.882	0.932	0.760	0.627
13	40	9.225	4.500	7.650	10.800	14.800	-9.325	-33.025
		686.972	7238.160	13497.035	21986.902	10152.367	25626.691	50880.910
			2043.695	2827.145	3610.594	2535.795	3604.919	4166.219
			0.913	0.928	0.929	0.960	0.859	0.705
14	50	9.980	6.494	10.940	15.386	17.920	-5.500	-28.560
		484.978	5176.211	9410.777	15431.906	6721.641	20614.500	48142.215
			1279.590	1707.936	2136.281	1576.915	1952.496	2302.436
			0.808	0.799	0.781	0.873	0.618	0.477

Pour chacun des 14 tests (en ligne) : nombre de sujets, moyenne et somme des carrés des écarts (SCE) à la moyenne du score simple (colonne 1) et des scores avec certitude (6 colonnes suivantes). Dans ces six dernières colonnes, figurent en outre la somme des produits des écarts (SWY) et, enfin, la corrélation avec le score simple.

Les corrélations réelles entre les scores simples (SS) et les scores "avec certitude" (SC) ont été calculées sur les données présentées au chapitre 2 (14 tests présentés à 53 sujets).

Pour chacune des 14 épreuves, le tableau 3.2 présente successivement, d'abord :

- le nombre de sujets (les absents ont été éliminés),
- la moyenne des réponses correctes,
- la somme des carrés des écarts à cette moyenne;

ensuite, dans chacune des six colonnes (correspondant aux barèmes A, B, C, D, E et F des tableaux 3.18 et 3.21) :

- la moyenne des scores avec certitude calculés avec ce barème,
- la somme des carrés des écarts de ces scores,
- la covariance (XN) entre ces scores et le nombre de réponses correctes
- la corrélation correspondante.

Le tableau 3.25 a été établi par le programme FORTRAN présenté dans le tableau 3.24.

Les corrélations figurant au tableau 3.25 ont été rassemblées synoptiquement dans le tableau 3.26.

Tableau 3.26

	Barèmes de cotation					
	A	B	C	D	E	F
Test n° 1	.897	.876	.833	.906	.815	.738
2	.738	.782	.765	.873	.585	.371
3	.915	.901	.872	.930	.822	.788
4	.900	.895	.878	.935	.797	.665
5	.904	.903	.890	.936	.822	.726
6	.882	.875	.848	.915	.718	.665
7	.933	.943	.935	.965	.874	.782
8	.865	.868	.852	.921	.717	.594
9	.955	.953	.947	.971	.903	.852
10	.828	.803	.774	.860	.682	.620
11	.958	.956	.948	.970	.907	.867
12	.838	.874	.882	.932	.760	.627
13	.913	.928	.929	.960	.859	.705
14	.808	.799	.781	.873	.618	.477
Corrélations théoriques	.865	.802	.684	.878	.919	.805

Corrélations entre les scores simples (SS) et les scores calculés selon l'un des six barèmes de certitude (voir tableaux 3.18 et 3.21) pour chacun des 14 tests de l'expérience de l'année scolaire 1971-1972 (cf. chapitre 2).

On constate que, pour les barèmes A, B, C et D, les corrélations observées sont très souvent supérieures aux corrélations théoriques. Cependant, l'inverse se produit (9 fois sur 56 corrélations, soit dans 16 % des cas).

Le barème B a un statut différent des autres : c'est le barème qui a été effectivement appliqué lors des 14 tests qui ont fourni les données de base.

La corrélation entre les scores simples et les scores calculés selon la matrice des conséquences est influencée, d'une part, par cette "corrélation théorique" extraite de la matrice et, d'autre part, par le comportement des sujets (la répartition de leurs réponses dans les diverses cellules).

Ces deux effets devraient, logiquement, agir de façon indépendante.

CONCLUSIONS

Il est vain de recueillir des indices de certitude auprès d'étudiants sans prendre certaines précautions méthodologiques.

La plus fondamentale d'entre elles consiste à utiliser une matrice des conséquences calculée selon le principe de l'utilité attendue; le présent chapitre a tenté quelques mises au point sur ce sujet. Une seconde précaution réside dans les consignes et la préparation des sujets. Il s'agit de leur faire comprendre (par conditionnement, entraînement, simulation, etc.) les propriétés de la matrice, les modes de réponses permis et leurs conséquences.

Si ces deux précautions sont prises, on peut sérieusement espérer obtenir ce que SHUFFORD, ALBERT et MASSENGILL (1966) appellent "une mesure acceptable de la probabilité". Les principes exposés au chapitre 3 ont été appliqués lors de l'expérience 1972-1973 qui va être analysée au chapitre 4.

Un problème reste cependant sans réponse :
Quelle importance relative faut-il attribuer à la certitude (par rapport à l'exactitude de la réponse) ?

CHAPITRE 4

UNE EXPERIENCE CONTINUE, AVEC CERTITUDES ORDINALES ET CONSEQUENCES CONFORMES A LA THEORIE MODERNE DE L'UTILITE (CRITERE E.S.U.)

LE PROBLEME DE LA VALIDITE DE L'ESPER

1. PRESENTATION DE L'EXPERIENCE

A. *LE PROBLEME*.

Puisque les sujets se conforment au critère de l'espérance subjective de l'utilité (E.S.U.), on voit l'intérêt d'une expérience semblable à celle de l'année 1971-1972, mais qui lui soit supérieure par quatre aspects :

a) Une matrice de conséquences conforme à la théorie moderne de l'utilité;

b) Une facilité des tests successifs plus stable autour d'une valeur moyenne de 50 %. Il est bien difficile de réaliser ce dernier point quand les questions n'ont pas été prétestées;

c) Une consigne précisant aux sujets quelles zones de l'axe des PER sont codées par les indices de certitude;

d) Des sujets entraînés par des exercices intensifs.

A la lumière des résultats de l'année 1971-1972, le Commandant VAN ROY, directeur de la banque de questions de l'E.T.F.A., a adopté, pour l'année 1972-1973, une matrice des conséquences calculée selon les principes exposés au chapitre 3. Le premier (a) des quatre aspects ci-dessus était ainsi réalisé.

Dans la série de résultats (13 tests successifs en mécanique) qui sera examinée dans ce chapitre, le second (b) des quatre aspects souhaités a aussi été réalisé. En effet, la corrélation entre le numéro d'ordre et la facilité du test est .066 (proche de 0). De plus, la facilité moyenne des 13 tests est 56,63 % (proche de 50 %).

d'accord

Le troisième (c) des quatre aspects souhaitables n'a pas été réalisé car, à ce moment, son importance ne nous apparaissait pas encore. Quant au quatrième (d) aspect, il a été partiellement réalisé puisque les mêmes sujets ont été soumis, dans d'autres disciplines que la mécanique, à des épreuves (à choix multiple) notées à l'aide de la même matrice de conséquences, ainsi qu'à la même procédure de renforcement.

Comme pour l'expérience de 1971-1972, les résultats seront confrontés à des hypothèses formulées *a posteriori*.

Les huit premières hypothèses (H1 à H8) émises au chapitre 2 restent valables :

- <u>Cohérences en taux d'utilisation (τ)</u> :

  H1. Cohérence globale en taux d'utilisation

  H2. Cohérence inter-individuelle en taux d'utilisation

  H3. Cohérence inter-épreuve en taux d'utilisation.

- <u>Cohérences en taux d'exactitude (μ)</u> :

  H4. Cohérence globale en taux d'exactitude

  H5. Cohérence individuelle en taux d'exactitude

  H6. Stabilité inter-individuelle du taux d'exactitude d'un degré de certitude donné

  H7. Cohérence en taux d'exactitude à l'intérieur d'un test

  H8. Stabilité entre épreuves du taux d'exactitude d'un degré de certitude donné.

finalement, la certitude intervient pour très peu !

B. LA PROCEDURE DE RENFORCEMENT DE L'EXPERIENCE 1972-73

Une centaine d'étudiants répartis en plusieurs classes ont été soumis, de septembre 1972 à juin 1973, à une trentaine de tests successifs (de 25 questions environ), portant sur diverses branches (mécanique, chimie, physique, électricité et électronique).

En plus de leur réponse, les étudiants doivent indiquer l'un des trois indices de certitude suivants :

- La certitude 1 (Je ne suis pas sûr de ma réponse);
- La certitude 2 (Ma réponse me semble bonne);
- La certitude 3 (Je suis très sûr de ma réponse).

L'omission est conseillée en cas d'ignorance totale.

On remarquera que la consigne est restée ordinale. La matrice des conséquences (en points) est la suivante (tableau 4.1).

Tableau 4.1

	réponse correcte	réponse incorrecte
Omission	0	- 0
Certitude 1	1,75	- 0,5
Certitude 2	2	- 0,75
Certitude 3	2,25	- 1,50

Matrice des conséquences utilisée durant l'année scolaire 1972-1973

Selon le modèle d'analyse de la variance décrit au chapitre 3,

SCT = 13,40

SCR = 12,75 (95,2 % de variance due à la réponse)

SCC = 0,152 (1,1 % de variance due à la certitude)

SCRC = 0,498 (3,7 % de variance due à l'interaction réponse x certitude)

Ceci n'est pas naturel.

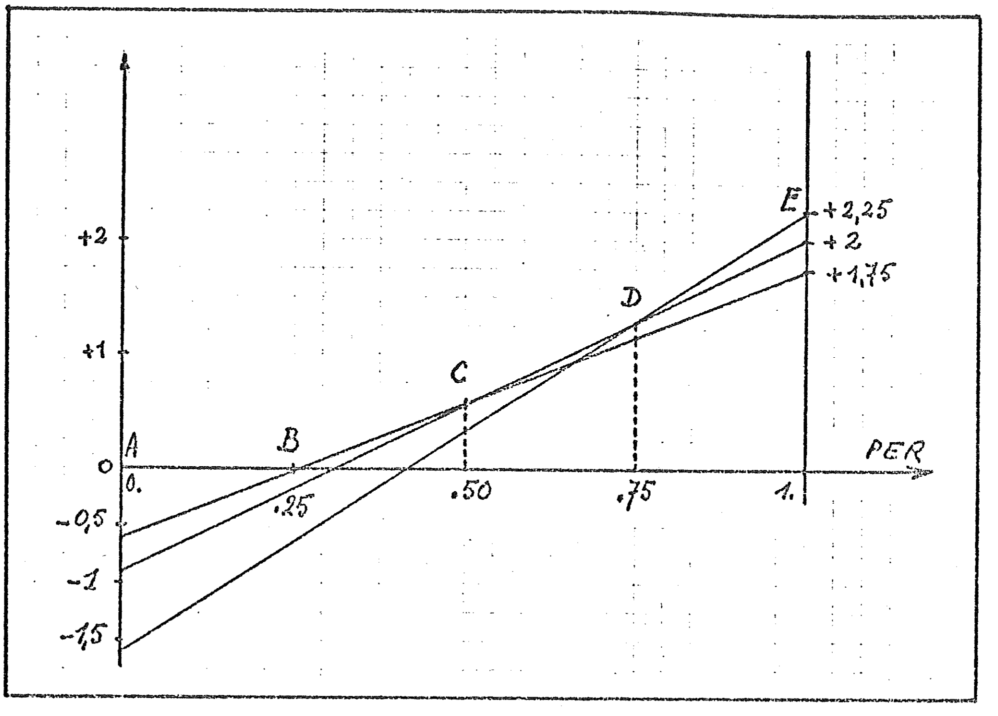
C'est aussi contraire à la philosophie que vous défendiez au départ: "il vaut mieux s'abstenir que se tromper."

En outre, il a été décidé de n'autoriser que 20 % d'omissions par test. Ainsi, dans un test de 25 questions, l'élève ne peut se permettre que cinq omissions. Au-delà de ce taux, toute omission supplémentaire est pénalisée par le retrait de 0,35 Point. Cette procédure ne correspond à aucun principe de la théorie des décisions et nous y sommes opposé. La direction de l'école a pris cette décision afin de "ne pas permettre à certains étudiants d'obtenir la moitié des points (1) sans avoir répondu à la moitié des questions." Cette consigne pourrait réduire le taux d'omissions au profit des certitudes 1. Cependant, dans l'expérience de l'année précédente, ce n'est qu'au septième test que le taux moyen d'omissions a atteint 20 % , alors que la matrice des conséquences favorisait l'emploi d'omissions. La légère entorse méthodologique que constitue cette pénalisation des omissions n'influencera donc que peu les résultats.

Les conditions operantes sont restées les mêmes qu'en 1971-72, A nouveau, le total des points détermine le passage dans une classe supérieure en fin d'année et l'obtention de permissions de sorties anticipées. Les valeurs de la matrice des conséquences sont, rappelons-le, destinées à forcer l'étudiant à conformer ses certitudes à ses seules ESPER pour maximiser son score. Le graphique suivant (tableau 4.2) en est la démonstration :

(1) Voir, en annexe 2, le système de calcul du niveau de réussite pédagogique (20/20 au bulletin).

Tableau 4.2



Fonctions d'utilité attendue ($\bar{u}$) pour chacun des quatre indices de certitude

On voit que quatre segments de droite sont optimaux :

AB : entre 0 et 25 % (stratégie optimale : omission)

BC : entre 25 et 50 % (stratégie optimale : certitude 1)

CD : entre 50 et 75 % (stratégie optimale : certitude 2)

DE : entre 75 et 100 % (stratégie optimale : certitude 3).

Les hypothèses 9 à 14 sont les mêmes que pour l'expérience de l'année scolaire 1971-1972. Elles concernent les stratégies incompatibles avec la validité des indices de certitude. L'hypothèse H18 a disparu. Les hypothèses H16, H17 et H19' diffèrent fortement par rapport à celles de l'année précédente, car elles dépendent étroitement de la matrice des conséquences. Ces diverses hypothèses seront à nouveau formulées au moment de leur vérification.

C. LES DONNEES DE BASE

Les résultats bruts qui feront l'objet de la présente étude ont été rassemblés à l'Ecole Technique de la Force Aérienne de Saffraanberg par la section "Contrôle des Etudes" dirigée par le commandant P. VAN ROY, avec l'aide, entre autres, de P. DE SCHUTTER et C. DEBATY. Nous jouions le rôle de conseiller pédagogique dans cette entreprise devenue autonome.

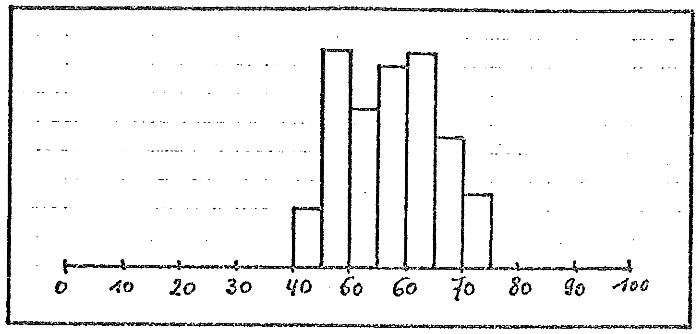
Encore une fois, seule sera considérée la plus longue série de résultats successifs portant sur une même matière (la mécanique de nouveau), présentés aux mêmes élèves. Il s'agit de 13 tests de 25 questions en moyenne, qui ont été administrés en neuf mois (de septembre 1972 à juin 1973) à six classes différentes (trois néerlandophones, trois francophones). Certaines classes ont reçu une ou deux questions en plus ou en moins, selon le souhait exprès des enseignants.

Ces six classes totalisaient 102 sujets, mais 29 d'entre eux ont soit quitté définitivement leur classe à Noël, soit rejoint la classe après Noël. Dans l'un et l'autre cas, les résultats sont trop incomplets. Seuls les résultats de 73 étudiants ont été retenus. Parmi ces étudiants, seize ont été absents lors d'un test, cinq lors de deux tests et un lors de trois tests. Par conséquent, on ne dispose que de 920 *patterns* au lieu de 949 (73 x 13).

Rappelons que le caractère operant de la procédure n'est pas parfait et que les sujets ne sont pas réellement entraînés.

Les pourcentages de réussite ($M|_i$ ER) des divers sujets sont les suivants (tableau 4.3).

Tableau 4.3

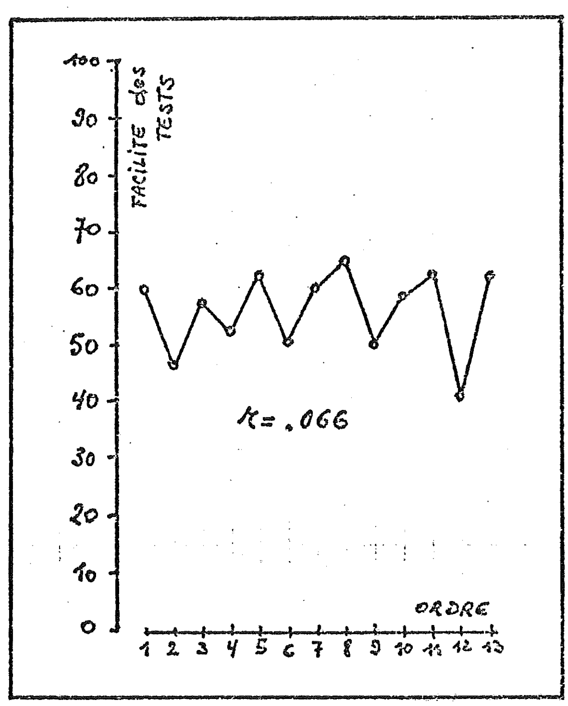


Distribution des pourcentages individuels de réussite
Année scolaire 1972-1973

La facilité moyenne des épreuves (56,63 %) est bonne (cf. les souhaits de HAMBLETON, ROBERTS et TRAUB (1970)). La dispersion faible entre les sujets (S = 8,44 %) est aussi une caractéristique favorable à notre étude.

Comme le montre le tableau 4.4, la difficulté des tests successifs varie, mais sans relation avec l'ordre de présentation des tests (r = .066).

Tableau 4.4



Facilité des 13 tests successifs
Année scolaire 1972-1973

VALIDITE ET COHERENCES DANS LE TAUX D'UTILISATION DES DIVERS DEGRES DE CERTITUDE

Cohérence en τ

1. La cohérence globale en taux d'utilisation.

H1. Le taux d'utilisation de certitudes maximales (ici certitude 3) est plus élevé et le TUC minimales (ici certitude 1) est plus faible pour les réponses correctes (ER = 1) que pour les réponses incorrectes (ER = 0).

$$(TUC_3 \mid ER = 1) > (TUC_3 \mid ER = 0)$$
$$(TUC_1 \mid ER = 1) < (TUC \mid ER = 0)$$

Sur les 19706 items, on a observé en 1972-73	Rappel 1971-72
11160 réponses correctes (soit 56,63 %)	51,32 %
6588 réponses incorrectes (soit 33,43 %)	27,37 %
1958 omissions (soit 9,93 %)	21,29 %

On voit ici l'effet de la pénalisation de l'omission lors de l'année scolaire 1972-73.

- Parmi les réponses correctes, on a observé

1340 certitudes 1 (soit 12,00 %)	6,37 %
2353 certitudes 2 (soit 21,08 %)	15,37 %
7467 certitudes 3 (soit 66,90 %)	78,24 %

Parmi les réponses incorrectes, on a observé

2139 certitudes 1 (soit 32,46 %)	18,98 %
2121 certitudes 2 (soit 32,19 %)	25,90 %
2328 certitudes 3 (soit 35,33 %)	55,11 %

L'hypothèse H1 est donc pleinement confirmée puisque

12 % < 32,46 % NB.: 1971-72 : 6,37 % < 18,98 %

et 66,90 % > 35,33 % 78,24 % > 55,11 %

3. La cohérence entre individus en taux d'utilisation.

Les individus plus compétents que d'autres ($\mathcal{M}|_i$ER élevée) devraient utiliser plus souvent les certitudes fortes, d'où l'hypothèse H2.

H2. Le taux de bonnes réponses d'un sujet ($\mathcal{M}|_i$ ER) est positivement corrélé avec le taux d'utilisation de certitudes fortes (τ_F).

Le tableau 4.5 permet de répondre à cette question.

Le taux d'utilisation de certitudes fortes (2 et 3) est corrélé avec le nombre de bonnes réponses au total de chaque élève : .489 mais bien moins qu'en 1971-72 (.816). Cela provient peut-être, en partie, de la dispersion plus faible des pourcentages de bonnes réponses (σ = 8,44 alors que σ = 10,25 en 1971-72).

Tableau 4.5

								µ				µ		
								% BR							
Code	-3	-2	-1	OM	+1	+2	+3	(1)	(2)	(3)	NR	NBR	% BR	Fortes	% fortes
501	12	51	50	14	38	58	51	43	53	80	274	147	53,64	172	62,77
502	42	45	45	7	11	38	108	19	45	72	296	157	53,04	233	78,71
503	15	53	82	6	29	59	52	26	52	77	296	140	47,24	179	60,47
504	12	41	93	8	33	47	62	26	53	83	296	142	47,92	162	54,72
505	21	32	44	33	30	39	97	40	54	82	296	166	56,08	189	63,85
507	9	36	53	57	36	61	44	40	62	83	296	141	47,63	150	50,67
508	12	55	51	16	30	76	56	37	58	82	296	162	54,72	199	67,22
509	24	40	49	42	37	35	69	43	46	74	296	141	47,63	168	56,75
510	34	43	40	26	25	49	79	38	53	69	296	173	58,44	205	69,25
511	36	37	58	29	27	37	72	31	50	66	296	136	45,94	182	61,48
513	33	46	33	35	10	38	81	23	45	71	276	129	46,73	198	71,73
514	18	75	59	11	25	57	27	29	43	60	272	109	40,07	177	65,07
515	9	12	110	7	85	29	44	43	70	83	296	158	53,37	94	31,75
516	72	12	48	24	15	10	115	23	45	61	296	140	47,29	209	70,60
518	20	69	24	29	12	70	47	33	50	70	271	129	47,60	206	76,01
603	25	38	24	29	20	57	103	45	60	80	296	180	60,81	223	75,33
604	26	44	29	24	17	40	116	36	47	81	296	173	58,44	226	76,35
605	34	34	39	41	19	28	101	32	45	74	296	148	50,00	197	66,55
606	15	27	60	56	39	34	65	39	55	81	296	138	46,62	141	47,63
607	19	67	29	39	16	69	57	35	50	75	296	142	47,97	212	71,62
609	15	15	51	42	42	11	96	45	42	86	272	149	54,77	137	50,36
610	37	42	25	50	11	35	96	30	45	72	296	142	47,97	210	70,94
611	38	44	52	42	21	30	69	28	40	64	296	120	40,54	181	61,14
612	47	62	16	36	4	46	85	20	42	64	296	135	45,00	240	81,08
613	54	37	38	26	22	17	77	36	31	58	271	116	42,80	185	68,56
614	41	31	39	25	18	18	80	31	36	66	252	116	46,03	170	67,46
701	77	14	3	16	0	11	175	0	44	69	296	186	62,83	277	93,58
702	20	18	44	41	39	30	104	46	62	83	296	173	58,44	172	58,10
703	17	47	10	41	6	63	112	37	57	86	296	181	61,14	239	80,74
704	63	23	2	42	1	29	116	33	55	64	276	146	52,89	231	83,69
708	9	41	5	38	4	69	69	44	62	88	235	142	60,42	188	80,00
709	20	45	27	38	13	44	109	32	49	84	296	166	56,08	218	73,64
711	44	7	10	41	10	20	164	50	74	78	296	194	65,54	235	79,39
712	26	17	24	37	21	32	139	46	65	84	296	192	64,86	214	72,29
713	48	8	8	31	11	5	185	57	38	79	296	201	67,90	246	83,10
714	30	28	15	46	22	34	101	59	54	77	276	157	56,88	193	69,92
715	22	25	12	58	13	31	135	52	55	85	296	179	60,47	213	71,95
1201	24	32	28	23	14	56	83	33	63	77	260	153	58,84	195	75,00
1202	15	28	46	8	24	44	95	34	61	86	260	163	62,69	182	70,00
1205	30	25	11	15	11	26	124	50	50	80	242	161	66,52	205	84,71
1206	10	19	16	33	23	61	98	58	76	90	260	182	70,00	188	72,30
1209	39	47	50	8	26	24	66	34	33	62	260	116	44,61	176	67,69
1210	36	33	8	6	4	20	153	33	37	80	260	177	68,07	242	93,07
1211	27	27	43	13	20	41	87	31	60	76	258	148	57,36	182	70,00
1214	35	29	8	0	4	21	144	33	42	80	241	169	70,12	229	95,31
1215	54	1	7	19	10	2	129	58	66	70	222	141	63,51	186	83,78
1216	13	11	27	39	28	28	96	50	71	88	242	152	62,80	148	61,15
1217	15	8	38	7	36	29	127	48	78	89	260	192	73,84	179	68,84
1301	29	21	20	25	12	21	97	37	50	76	225	130	57,77	168	74,66
1302	51	7	12	22	9	14	145	42	66	73	260	168	64,61	217	83,46
1303	62	13	20	13	7	19	109	25	59	63	243	135	55,55	203	83,53
1304	37	8	10	15	9	13	168	47	61	81	260	190	73,07	226	86,92
1305	53	1	0	21	0	4	181	0	80	77	260	185	71,15	241	92,69
1306	64	17	14	22	9	16	118	39	48	64	260	143	55,00	215	82,69
1307	45	21	11	6	9	15	131	45	41	74	240	155	64,58	212	88,33
1308	51	10	10	4	7	13	145	41	56	73	240	165	68,75	219	91,25
1310	46	7	10	26	5	9	157	33	56	77	260	171	65,75	219	84,23
1311	22	28	17	25	14	33	103	45	54	82	242	150	61,98	186	76,85
1312	52	20	11	9	4	12	113	26	37	68	221	129	58,37	197	89,14
1313	37	11	11	24	11	14	152	50	56	80	260	177	68,07	214	82,30
1314	28	12	8	31	2	19	139	20	61	83	239	160	66,94	198	82,84
1701	41	7	37	38	23	12	102	38	63	71	261	137	52,49	162	62,30
1702	36	31	14	33	6	30	110	30	49	75	260	146	56,15	207	79,61
1703	36	25	35	40	25	26	73	41	50	66	251	125	47,89	160	61,53
1704	19	12	26	23	31	32	117	54	72	86	260	180	69,23	180	69,23
1706	16	40	26	12	17	57	92	39	58	85	260	166	63,84	205	78,84
1707	31	12	24	41	22	7	105	47	36	77	242	134	55,37	218	90,08
1708	6	31	48	50	39	63	23	44	67	79	260	125	48,07	123	47,30
1709	16	46	33	14	29	34	50	46	42	75	222	113	50,90	146	65,76
1710	41	34	14	12	12	23	124	46	40	75	260	159	61,15	222	85,38
1711	35	32	28	29	14	18	104	33	36	74	260	136	52,30	189	72,69
1712	52	29	5	38	6	13	117	54	30	69	260	136	52,30	211	81,15
1713	28	25	12	29	6	28	132	33	52	82	260	166	63,84	213	81,92

Ventilation des certitudes individuelles au total
des 13 épreuves de mécanique

Année scolaire 1972-1973

3. La cohérence entre épreuves en taux d'utilisation des divers degrés de certitude.

Les tests les plus faciles devraient entraîner une utilisation plus grande des indices de certitude élevés, d'où l'hypothèse H3.

H3. Il existe une corrélation positive et élevée entre la facilité de chaque test $M|_t$ER et le taux d'utilisation des certitudes fortes ($TUC_{2\ t}$ + $TUC_{3\ t}$).

Voici les résultats test par test (tableau 4.6) :

Tableau 4.6

Test	Mauvaises réponses -3	-2	-1	Omissions 0	Bonnes réponses 1	2	3	Total	% ER	U1	U2	U3
1	168	173	174	166	91	237	656	1665	59,09	265	410	824
2	216	209	200	131	69	145	453	1423	46,87	269	354	669
3	138	183	167	159	87	188	588	1510	57,15	254	371	726
4	196	187	144	140	98	148	491	1404	52,49	242	335	687
5	259	225	233	191	185	297	1059	2449	62,92	418	522	1318
6	191	187	174	119	91	136	464	1362	50,73	265	323	655
7	137	155	155	119	115	213	538	1432	60,47	270	368	675
8	181	120	110	116	83	158	746	1514	65,19	193	278	927
9	203	156	167	168	84	168	444	1390	50,07	251	324	647
10	170	131	141	153	121	181	578	1475	59,66	262	312	748
11	151	105	128	133	112	182	593	1404	63,17	240	267	744
12	227	200	217	204	98	142	332	1420	40,28	315	342	559
13	91	90	129	159	106	158	535	1258	63,51	235	248	616
	2328	2121	2139	1958	1340	2353	7467	19706		3479	4474	9735

UTILISATION DE CHAQUE DEGRE DE CERTITUDE AU COURS DES 13 EPREUVES SUCCESSIVES DE MECANIQUE (ANNEE SCOLAIRE 1972-73)

Les données qui concernent l'hypothèse H3 sont fournies par le tableau 4.7.

Tableau 4.7

Test	Pourcentage d'utilisation				Facilité du test	Total des réponses (1)	% certitudes faibles	% certitudes fortes
	Omission	Certitude 1	Certitude 2	Certitude 3				
1	9,96	15,91	24,62	49,48	59,09	1665	25,9	74,1
2	9,20	18,90	24,87	47,01	46,87	1423	28,1	71,9
3	10,52	16,82	24,56	48,07	57,15	1510	27,3	72,7
4	9,97	17,23	23,86	48,93	52,49	1404	27,2	72,8
5	7,79	17,06	21,31	53,81	62,92	2449	24,9	75,1
6	8,73	19,45	23,71	48,09	50,73	1362	28,2	71,8
7	8,31	18,85	25,69	47,13	60,47	1432	27,2	72,8
8	7,66	12,74	18,36	61,22	65,19	1514	20,4	79,6
9	12,08	18,05	23,30	46,54	50,07	1390	30,1	69,9
10	10,37	17,76	21,15	50,71	59,66	1475	28,1	71,9
11	9,47	17,09	20,44	52,99	63,17	1404	26,6	73,4
12	14,36	22,18	24,08	39,36	40,28	1420	36,5	63,5
13	12,63	18,68	19,71	48,96	63,51	1258	31,3	68,7
Moyenne	10,08	17,75	22,74	49,22	56,63	1515,8 (2)		

(1) Omissions comprises

(2) Le total des réponses aux 13 tests est 19.706

POURCENTAGES D'UTILISATION DE CHAQUE DEGRE DE CERTITUDE AU COURS DES 13 EPREUVES SUCCESSIVES DE MECANIQUE (ANNEE SCOLAIRE 1972-73)

La corrélation entre la facilité d'un test et le pourcentage de certitudes fortes utilisées vaut .70.

L'hypothèse H3 est donc pleinement confirmée.

VALIDITE ET COHERENCES DANS LE TAUX D'EXACTITUDE DES DIVERS DEGRES DE CERTITUDE

Cohérence en μ

1. La cohérence globale en taux d'exactitude.

Plus les réponses sont accompagnées d'un degré de certitude (ESPER ou C) élevé, plus leur succès ($M|_C$ ER) doit être élevé, d'où l'hypothèse H4.

H4. Il existe, pour l'ensemble des réponses, une relation croissante entre le degré de certitude (C) utilisé et le taux d'exactitude des réponses auxquelles il a été associé M (ER | ESPER = C) ou μ_C.

Sur le total des 19706 questions, on a observé :

1958 omissions, soit 9,93 %
3479 certitudes 1, soit 17,65 %
4474 certitudes 2, soit 22,70 %
9795 certitudes 3, soit 49,70 %.

La ventilation totale figure au tableau 4.8.

Tableau 4.8

	- 3	- 2	- 1	0	1	2	3	Total
F :	2328	2121	2139	1958	1340	2353	7467	19706
P :	11,8 %	10,7 %	10,8 %	9,9 %	6,8 %	11,9 %	37,9 %	100 %

On trouvera, en annexe 4, les tableaux des résultats par test et par sujet.

Parmi les réponses fournies avec	(1972-73)	(1970-71)
- Certitude 1, 1340 sont correctes, soit	= 38,5 %	38, 6 %
- Certitude 2, 2353 sont correctes, soit	= 52,6 %	52,6 %
- Certitude 3, 7467 sont correctes, soit	= 76,2 %	72,7 %

$$\mu_1 < \mu_2 < \mu_3$$

puisque 38,5 % < 52,6 % < 76, 2 %.

L'hypothèse H4 est donc de nouveau confirmée. On remarquera combien les μ des deux années sont proches. Ici aussi, le rho de Spearman vaut 1.

2. La cohérence individuelle en taux d'exactitude (μ_{ic}).

Au niveau de chaque individu, les degrés de certitude faibles doivent être attribués à des réponses dont les taux de succès seront faibles, d'où l'hypothèse H5.

H5. Il existe, pour chaque sujet, une relation croissante entre le degré de certitude utilisé (C) et le taux d'exactitude des réponses auxquelles il a été associé ($M|_i(ER|C=c)$ ou μ_{ic}).

Le tableau 4.5 contient le μ_{ic} de chaque degré de certitude pour chaque sujet. Nous avons appelé :

- cohérence forte la situation $\mu_{i1} \leq \mu_{i2} \leq \mu_{i3}$
- cohérence simple la situation $\mu_{i1} < \mu_{i3}$
- non cohérence les autres situations (où $\mu_{i1} \geq \mu_{i3}$

Sur les 73 étudiants, on dénombre 62 cohérences fortes et 11 cohérences simples.

Ces observations sont nettement à l'avantage de l'expérience 1972-73 par rapport à celle de l'année précédente (tableau 4.9).

Tableau 4.9

	1971-72	1972-73
cohérences fortes	71,69 %	84,93 %
cohérences simples	22,64 %	15,06 %
incohérences	5,66 %	0

Cohérence individuelle
Comparaison entre les résultats des deux années successives

Si on élimine cinq étudiants qui n'ont pas utilisé dix fois l'un des degrés de certitude, on dénombre :

58 cohérences fortes (85,29 %)

10 cohérences simples (14,70 %)

3. La stabilité entre individus du taux d'exactitude (d'un degré de certitude donné).

Quelle que soit la compétence des sujets, le taux d'exactitude des réponses auxquelles un degré de certitude a été associé ne devrait pas varier, d'où l'hypothèse H6.

H6. Il existe une corrélation nulle entre le taux d'exactitude des réponses $M|i\ (ER|C = c)$ auxquelles un degré de certitude a été associé par un sujet et le taux d'exactitude des réponses de ce sujet, $M|i\ ER$.

On peut déjà raisonnablement prédire que la consigne ordinale va amener une variabilité regrettable entre les individus pour leurs moyennes de réussite pour chaque degré de certitude

Ces paramètres sont les suivants (tableau 4.10):

Tableau 4.10

	M	S^2	S
μ_1	37,58	130,24	10,3
μ_2	52,65	130,49	10,3
μ_3	76,08	59,99	7,75
% BR	56,98	71,01	8,42

Les corrélations avec le taux de bonnes réponses sont les suivantes :

(de la certitude 1) : .2794

(de la certitude 2) : .4797

(de la certitude 3) : .0489

Si la troisième corrélation est nettement favorable à l'hypothèse H6, la première et surtout la seconde lui sont légèrement défavorables. Ce problème devra donc être examiné avec attention dans les expériences ultérieures.

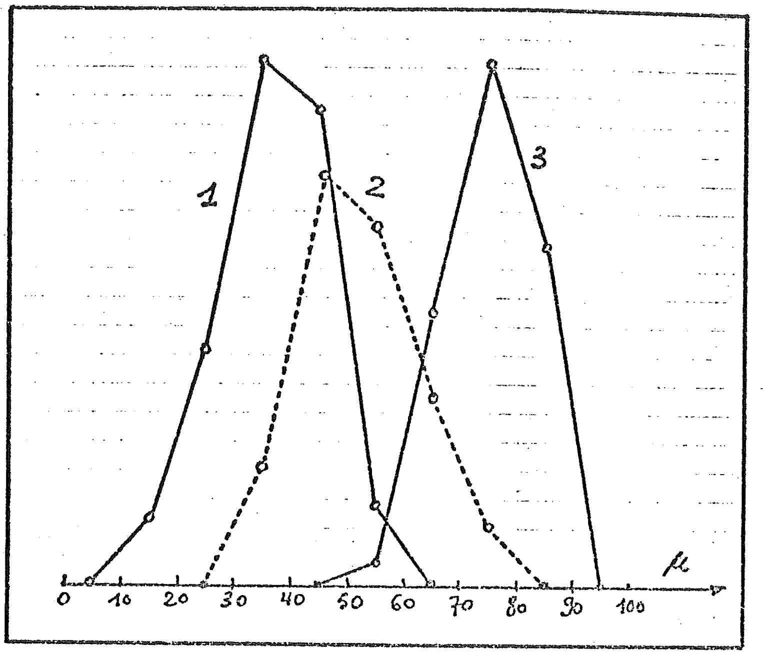
On ne peut cependant pas accorder un même poids à un μ calculé sur n = 5 et à un μ calculé sur n = 50. Dès lors, une représentation d'ensemble pourrait être basée sur le principe suivant : à chaque μ (de chaque degré) est attribuée une répétition égale au nombre de réponses sur lequel il a été calculé (tableau 4.11).

Tableau 4.11

Fiabilité	Certitude 1	Certitude 2	Certitude 3
91-100 %	0	0	0
81-90 %	0	0	.29,10
71-80 %	0	5,11	45,10
61-70 %	0	16,38	23,98
51-60 %	7,05	31,17	1,79
41-50 %	42,14	35,74	0
31-50 %	46,89	10,40	0
21-30 %	20,57	0,92	0
11-20 %	5,91	0	0
0-10 %	0,1	0	0

Les répétitions obtenues sont alors portées sur un histogramme (classes regroupées). Voici l'"histogramme" de chacun des trois degrés de certitude (tableau 4.12).

Tableau 4.12



"Histogrammes" des μ de chaque degré de certitude obtenus par multiplication des μ par les répétitions correspondantes. Année scolaire 1972-73

Les sigmas de la présente expérience sont légèrement plus faibles que ceux de 1971-1972.

4. La cohérence en taux d'exactitude à l'intérieur d'un test.

Quelles que soient les caractéristiques d'un test et notamment sa facilité $M|_t$ER, les degrés de certitude faibles doivent correspondre à des réponses dont le taux d'exactitude $M|_t$(ER | C = 0 ou 1) est faible, d'où l'hypothèse H7.

H7. Il existe, pour chaque test, une relation croissante entre C et $M|_t$ (ER|C = c) ou μ_{tc}.

Voici les pourcentages au cours des 1 tests successifs (tableau 4.13).

Tableau 4.13

Test	1	2	3
1	34	57	80
2	25	40	68
3	34	50	81
4	40	44	71
5	44	56	80
6	34	42	71
7	42	57	80
8	43	56	80
9	33	51	69
10	46	58	77
11	46	63	80
12	31	41	59
13	45	63	85
totale (1)	38,5	52,6	76,2

(1) Les moyennes totales ont été calculées sur l'ensemble des réponses.

On ne constate, à aucun test, une interversion des μt_c de deux degrés de certitude. Le rho de Spearman vaut 1. L'hypothèse H7 est donc confirmée.

5. La stabilité en taux d'exactitude entre épreuves.

Quelle que soit la facilité du test, le taux de succès des réponses correspondant à une certitude donnée devrait ne pas varier, d'où l'hypothèse H8.

H8. Il existe une corrélation nulle entre $M|_t ER$ et $M|_t (ER|C=c)$.

Les pourcentages de réussite (μt_c) sont corrélés comme suit avec la facilité du test. (Les corrélations de l'expérience précédente sont notées pour mémoire.)

		1971-72
μt_1 (de la certitude 1) :	r = .810	(.791)
μt_2 (de la certitude 2) :	r = .877	(.818)
μt_3 (de la certitude 3) :	r = .949	(.948)

On remarquera, ici encore, le parallélisme entre les résultats de 1971-1972 et ceux de 1972-1973.

H8'. La corrélation entre $M|_t (ER|C = c)$ et l'ordre du test est nulle.

Les corrélations brutes entre chacun de ces trois μt_c et le numéro d'ordre du test sont :

		1971-72
pour μt_1 :	r = .305	(-.430)
pour μt_2 :	r =-.556	(-.524)
pour μt_3 :	r =-.065	(-.808)

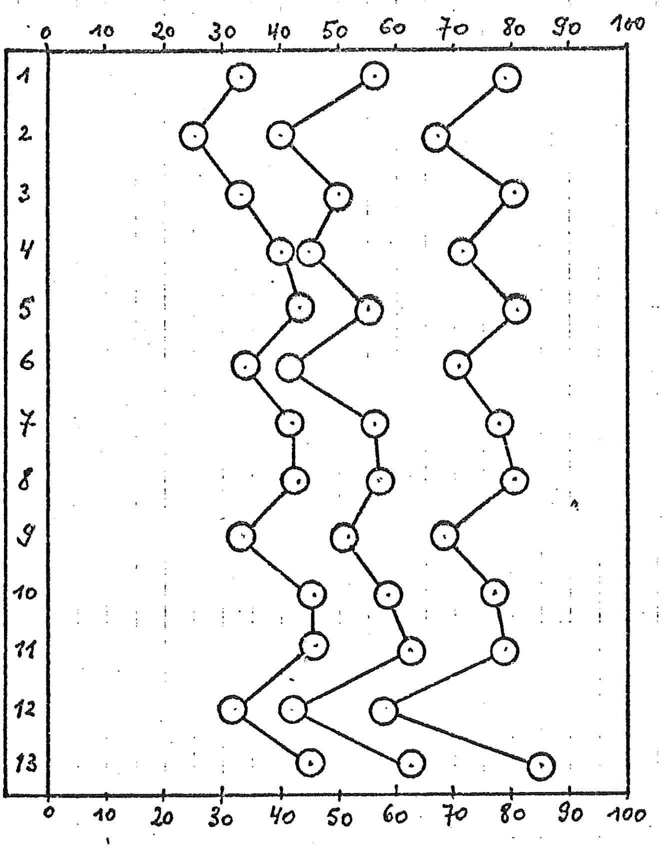
La variation d'un test à l'autre est, dans une très large mesure, expliquée par les différences de facilité des tests successifs (tableau 4.14).

Les μt_c de chaque degré ont été prédits à partir de la facilité du test (f) par les équations suivantes :

$$\mu t_1 = 0{,}728\ f - 2{,}697 \qquad (r = .810)$$
$$\mu t_2 = 0{,}949\ f - 1{,}175 \qquad (r = .877)$$
$$\mu t_3 = 0{,}927\ f + 23{,}037 \qquad (r = .949)$$

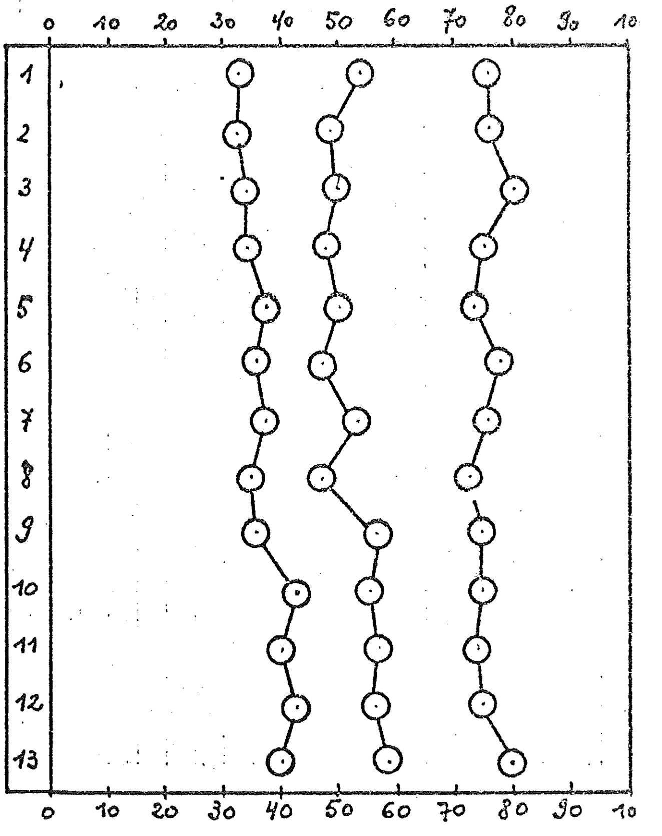
Les moyennes des μt_c respectifs ont été ajoutés aux résidus obtenus. Portées sur graphique, les nouvelles valeurs mettent en évidence la stabilité des ESPER (tableau 4.15).

Tableau 4.14



Moyennes d'exactitude des réponses fournies avec chaque indice de certitude au cours des 13 tests de l'année scolaire 1972-1973.

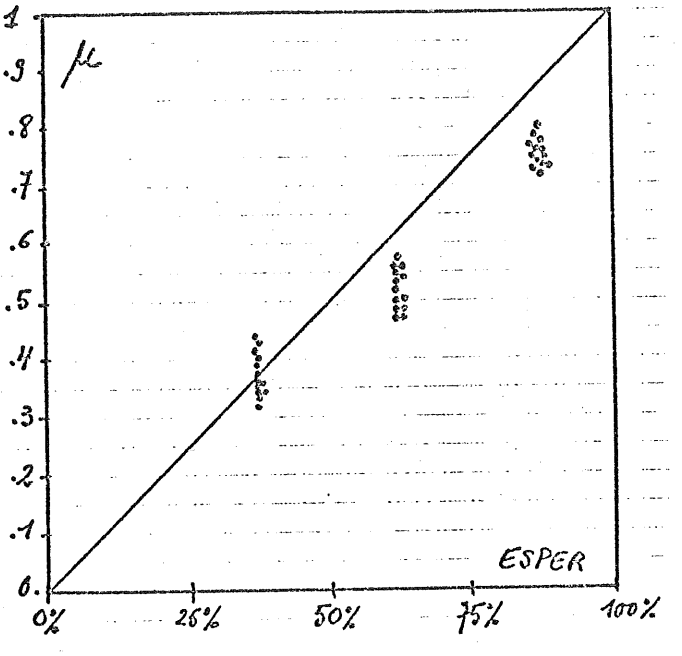
Tableau 4.15



Moyennes d'exactitude des réponses fournies avec chaque degré de certitude au cours des 13 tests de l'année scolaire 1972-1973, après élimination de la facilité du test.

Sur le nuage de points μ / ESPER (tableau 4.16), on remarque l'inclinaison de la droite de régression, légèrement inférieure à 45 degrés, ce qui confirme l'observation antérieure. Les individus ont en effet tendance à se surestimer légèrement lors de l'emploi des certitudes 2 et 3 : ils ne les emploient pas avec suffisamment de parcimonie.

Tableau 4.16



Liaison, pour chacune des 13 épreuves de l'année scolaire 1972-1973, entre l'ESPER et la moyenne d'exactitude des réponses fournies avec un indice de certitude (μ)

4. STRATEGIES INCOMPATIBLES AVEC LA VALIDITE DES INDICES DE CERTITUDE

Comme pour l'expérience précédente, on cherchera à savoir si chacun des critères de la théorie des décisions explique les *patterns* comportementaux. Par "*pattern* comportemental", on entend la configuration des nombres d'utilisation de chaque degré (omission comprise) d'un individu donné à un test donné.

Il n'y a pas 949 *patterns* (73 sujets à 13 tests), mais 920 seulement.

Les hypothèses suivantes vont être testées :
Les étudiants n'utilisent pas ...

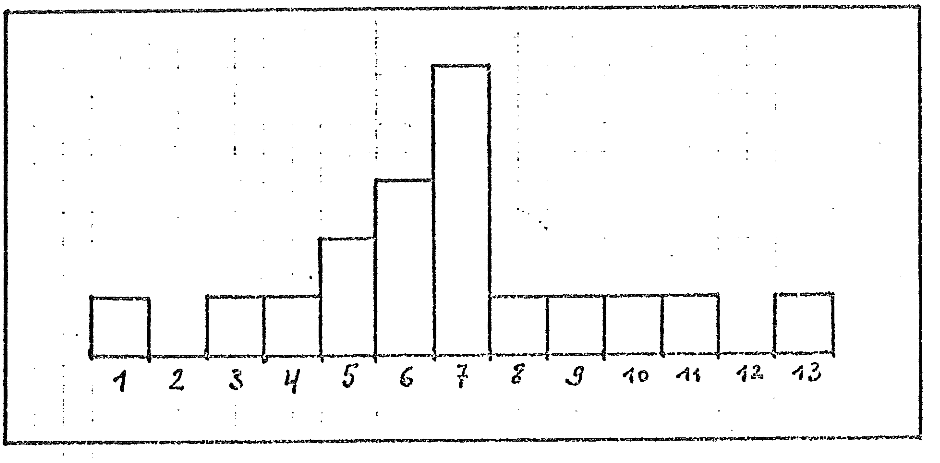
H9 ... la stratégie 1 (critère maximax)
H10 ... la stratégie 2 (critère maximin)
H11 ... la stratégie 3 (critère d'équiprobabilité)
H12 ... la stratégie 4 (critère de HURWICZ)
H13 ... la stratégie 5 (critère de SAVAGE)
H14 ... la stratégie 6 (critère de COOMBS)

H9. Les étudiants n'utilisent pas la stratégie 1 (application du critère maximax), qui consiste à ne fournir que des réponses avec la certitude 3 (qui donne le gain maximal en cas d'événement favorable).

Ce *pattern* se produit chez dix sujets sur 73. Chez un sujet, lors de quatre tests, chez un autre lors de trois tests, chez trois sujets lors de deux tests et chez les cinq autres sujets lors d'un seul test. Ainsi, on peut dire que ce pattern a été observé dans 18 tests sur 920, soit 1,95 %.

Il est remarquable cependant qu'à une seule exception près (sujet 701), les individus qui ont présenté ce *pattern* viennent tous de la même classe (sujets 1301, 1303, 1304, 1305, 1306, 1307, 1308, 1312, 1314). Ces *patterns* sont répartis dans les 13 épreuves (tableau 4.17) avec un mode lors du test 7, dont la facilité est supérieure à la moyenne (60,47 %).

Tableau 4.17

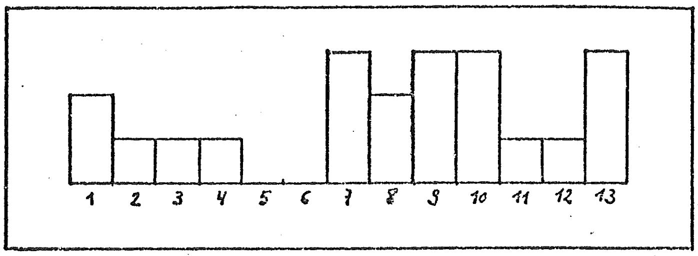


Evolution au cours des 13 tests de l'année scolaire 1972-1973 du nombre de patterns compatibles avec le critère maximax

Considérons maintenant un *pattern* moins strict, inspiré lui aussi du critère maximax : l'utilisation des seules certitudes 2 et 3. Ce *pattern* se produit chez 11 sujets (dont six de la classe 13). Deux sujets présentent ce *pattern* lors de quatre tests, un lors de trois tests, deux lors de deux tests et six lors d'un seul test.

En tout, ces 21 *patterns* constituent 2,28 % de l'ensemble. Ces *patterns* sont largement répartis dans le temps et sont en légère augmentation à partir du test 7 (tableau 4.18).

Tableau 4.18



Evolution, au cours des 13 tests de l'année scolaire 1972-1973, du nombre de patterns compatibles avec une interprétation élargie (utilisation des seules certitudes fortes) du critère maximax

Si l'on élargit encore beaucoup plus l'interprétation du critère maximax, on aboutit à considérer le *pattern* "absence d'omission", comme dans les exemples (imaginaires) 1 et 2 ci-dessous.

	-3	-2	-1	0	1	2	3
Exemple 1 :	1	1	2	0	1	5	12
Exemple 2 :	1	0	3	0	0	6	12
Exemple 3 :	8	0	2	0	1	0	5
Exemple 4 :	0	2	13	0	2	3	0

L'exemple 3 n'est pas retenu car la certitude 2 n'a pas été utilisée, contrairement à la certitude 1. Ce *pattern* est incompatible avec le critère maximax. L'exemple 4 constitue un cas semblable. Les deux types de *patterns* déjà notés au début de ce type de stratégie ne sont pas repris dans le comptage qui suit.

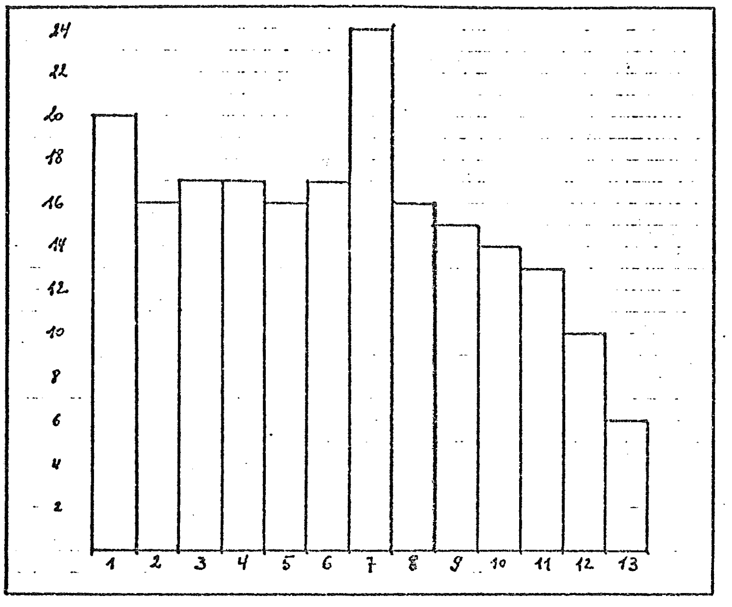
Les *patterns* "absence d'omission" ont été observés chez 53 sujets et, en tout, 167 fois (soit dans 18,15 % des cas).

Ce *pattern* s'est présenté 1 fois pour 22 sujets
2 fois pour 8 sujets
3 fois pour 6 sujets
4 fois pour 1 sujet
5 fois pour 4 sujets
6 fois pour 4 sujets
7 fois pour 1 sujet
8 fois pour 4 sujets
9 fois pour 2 sujets.

Ce taux (18,15 %) révèle que 81,85 % des *patterns* comportaient des omissions. La pénalisation des omissions excédentaires n'a donc pas supprimé l'omission.

La répartition temporelle de ces *patterns* est la suivante (tableau 4.19).

Tableau 4.19



On a observé 1958 omissions sur les 920 *patterns;* on peut ainsi calculer que le nombre moyen d'omissions par test individuel est 2,022. L'absence totale d'omission à un *pattern* n'a donc rien de surprenant.

Si l'on totalise tous les *patterns* comportementaux qui pourraient être compatibles avec une acception fort large de la stratégie 1, on ne compte que 206 *patterns*, soit 22,39 %. De plus, seuls sept sujets ont présenté de tels *patterns* dans plus de la moitié de leurs tests (sept, huit et neuf *patterns*). Le nombre de ces *patterns* diminue à partir du septième test. Ils se produisent chez des sujets différents, à peu près à n'importe quel test. Ce fait indique que le phénomène est plus lié au hasard qu'à des stratégies systématiques.

H10. Les étudiants n'utilisent pas la stratégie 2 (application du critère maximin).

L'utilisation du critère maximin consiste à omettre sans cesse puisque c'est le comportement qui donne la pénalisation la moins grave en cas d'échec (0). Aucun *pattern* semblable n'a été observé.

En considérant ce critère de la façon la plus complaisante possible, on pourrait accepter le *pattern* où l'omission est plus utilisée que tout autre type de réponse. Ce *pattern* ne s'observe que deux fois :

	-3	-2	-1	0	1	2	3
pour le sujet 715 (test 3) :	2	0	0	13	0	0	9
pour le sujet 711 (test 12) :	2	3	0	8	1	2	4

Ce petit nombre d'observations écarte le critère maximin dans l'explication des comportements.

H11. Les étudiants n'utilisent pas la stratégie 3 (application du critère de LAPLACE).

Ce critère consiste à attribuer autant (équiprobabilité) de chances au succès de la réponse fournie qu'à son échec (soit 50 % à chacun).

L'espérance mathématique de chaque action possible est :

Omission : 0
Certitude 1 : $(0,5 \times 1,75) + (0,5 \times -0,5) = 0,875 - 0,25 = 0,625$
Certitude 2 : $(0,5 \times 2) + (0,5 \times -0,75) = 1 - 0,375 = 0,625$
Certitude 3 : $(0,5 \times 2,25) + (0,5 \times -1,5) = 1,125 - 0,75 = 0,375$

La stratégie 3 consiste donc à ne fournir que des réponses avec certitudes 1 et 2, à l'exclusion de la certitude 3 et de l'omission. Ce *pattern* a été observé deux fois chez un sujet (sujet 515) et une fois chez deux autres (503, 504). Cette stratégie n'a donc été utilisée que dans 0,4 % des cas.

H12. <u>Les étudiants n'utilisent pas la stratégie 4</u> (application du critère de HURWICZ).

L'utilisation de ce critère consiste à

- omettre sans cesse si l'on est très pessimiste (c'est-à-dire si on attribue une PER commune inférieure à 25 % à toutes les réponses d'un test;
- répondre sans cesse avec la certitude 1 si l'on est légèrement pessimiste (si la PER commune est comprise entre 25 et 50 %);
- répondre sans cesse avec la certitude 2 si l'on est modérément optimiste (50 % < PER commune < 75 %);
- répondre sans cesse avec la certitude 3 si l'on est très optimiste (PER commune > 75 %).

A aucun moment, les *patterns* "très pessimiste", "pessimiste" et "optimiste" n'ont été observés. Le *pattern* "très optimiste" (répondre sans cesse avec la certitude 3) a été rencontré, rappelons-le, chez dix sujets, au cours de 18 tests au total. Il faut bien convenir que, si cette hypothèse est correcte, l'optimisme en question est très éphémère puisque il a été observé comme suit :

- Pour un sujet, lors de quatre tests
- Pour un sujet, lors de trois test
- Pour trois sujets, lors de deux tests
- Pour cinq sujets, lors d'un seul test.

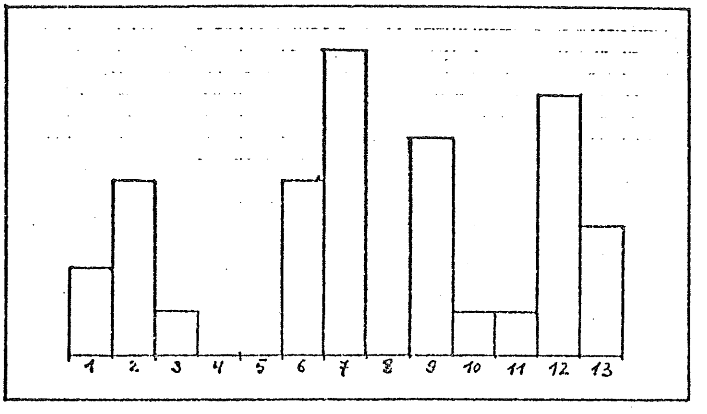
Si l'on considère le critère d'une façon beaucoup plus large, on peut classer dans les *patterns* "pessimistes" ceux qui consistent à ne pas utiliser la certitude 3. Ce *pattern* a été observé chez 11 sujets (3 fois chez 2 sujets, 2 fois chez 2 autres et une seule fois chez les 7 autres sujets).

En élargissant encore l'interprétation du critère de HURWICZ (pessimisme), on peut considérer les *patterns* où une seule réponse a été fournie avec certitude 3. Ce dernier *pattern* et le *pattern* "absence de réponse 3" ont été observés chez 19 sujets :

Pour un sujet, lors de 5 tests
Pour un sujet, lors de 4 tests
Pour deux sujets, lors de 3 tests
Pour quatre sujets, lors de 2 tests
Pour onze sujets, lors d'un seul test.

Si ces *patterns* étaient dus à une attitude pessimiste, il faudrait bien convenir que ce pessimisme fut loin d'être systématique. Ces *patterns* ont été observés lors des tests suivants (tableau 4.20).

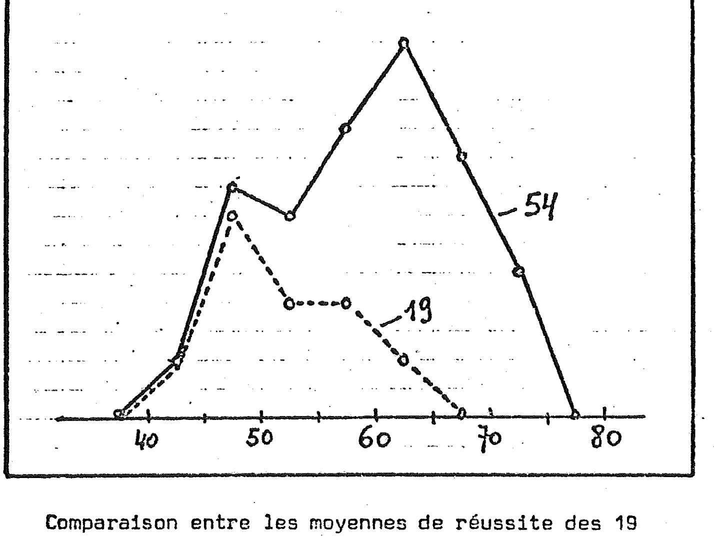
Tableau 4.20



Malgré la concentration sur certains tests (2, 6, 7, 9, 12, 13), ce *pattern* est très réparti quant à l'ordre des tests.

Cependant, il n'est nullement prouvé que ce *pattern* soit causé par une attitude "pessimiste". Il pourrait aussi bien s'agir de l'évaluation correcte, par les sujets, de leur faiblesse. Ainsi, si l'on considère le pourcentage de réponses correctes fournies par ces 19 sujets, on constate que la moyenne vaut 51,45 %, alors que la moyenne des 54 autres sujets vaut 58,93 %. Le tableau 4.21 montre bien que ces *patterns* ont été observés chez les élèves les plus faibles.

Tableau 4.21



Comparaison entre les moyennes de réussite des 19 sujets qui ont présenté des patterns "pessimiste" et les moyennes de réussite des 54 autres sujets

Pour 15 des 19 sujets, ces *patterns* se sont produits dans des tests où leur pourcentage de bonnes réponses était plus faible que dans les autres tests. Le tableau 4.22 présente, pour ces 19 sujets, l'exactitude moyenne de leurs réponses aux tests présentant ce *pattern* "pessimiste" (appelés "tests P") et aux autres tests.

Tableau 4.22

Numéro du sujet	Nombre de tests P	MliER des tests P	moyennes des autres tests	Cas aberrants
501	3	43,75	56,65	
502	3	30,30	59,56	
503	1	25,0	48,90	
504	1	10,52	50,53	
505	1	31,57	57,76	
507	4	46,06	48,30	
508	1	70,0	53,61	x
509	1	60,0	46,73	x
510	1	54,54	58,75	
514	1	65,0	38,09	x
515	2	63,63	51,57	x
518	2	45,0	47,80	
606	1	26,31	48,01	
611	1	18,18	40,77	
703	2	37,77	65,32	
708	2	40,42	65,42	
1301	1	35,0	59,44	
1303	1	42,10	56,69	
1708	5	45,76	49,69	

Comparaison entre moyennes de réussite des tests P (où pattern "pessimiste" s'est présenté) et les moyennes de réussite des autres tests pour les 19 sujets qui ont présenté des patterns compatibles avec le critère de HURWICZ (pessimisme)

On remarque que 63 % des cas se produisent dans la classe 5 (1) et que les quatre cas aberrants se produisent dans cette même classe. Il se pourrait que ces quatre sujets aient adopté cette stratégie par imitation de leurs condisciples, alors qu'elle n'était pas en concordance avec leurs ESPER personnelles. Si cela est vrai, voilà en effet cinq *patterns* (sur 920, soit 0,5 %) superstitieux, incompatibles avec la validité de l'ESPER.

(1) Le code du sujet est fait du code de la classe, suivi du numéro de l'élève proprement dit.

H13. Les étudiants n'utilisent pas la stratégie 5 (utilisation du critère du regret minimax).

Les profits maximum sont :

- en cas de réponse correcte, + 2,25
- en cas de réponse incorrecte, 0.

Dans la matrice des regrets (donc valeurs négatives ou nulles) du tableau 4.23, le regret maximum pour chaque action possible a été souligné.

Tableau 4.23

		réponse correcte	réponse incorrecte
A1	Omission	-2,25	0
A2	Certitude 1	-0,5	-0,5
A3	Certitude 2	-0,25	-0,75
A4	Certitude 3	0	-1,5

Matrice des regrets en application du critère de SAVAGE à la matrice de conséquences du tableau 4.1

Le plus petit des regrets maximum est - 0,5; il se produit pour la certitude 1. Le *pattern* de réponses découlant de la stratégie 5 consiste donc à fournir uniquement des réponses avec certitude 1; il n'est observé que pour trois sujets, chacun lors d'un seul test et jamais lors du même test (tests 2, 7 et 12).

H14. Les étudiants n'utilisent pas la stratégie 6 (application du critère "ampleur du risque").

Dans cette expérience, la matrice était la suivante :

Tableau 4.24

		Evénements possibles		Ampleur du risque (r)
		Réussite	Echec	
	Omission	0	0	0
Actions	Rép. cert. 1	+1,75	-0,5	2,25
disponibles	Rép. cert. 2	+2	-0,75	2,75
	Rép. cert. 3	+2,25	-1,5	3,75

L'étude portera non seulement sur r, mais en plus sur les échelles jointes obtenues à partir de r^2, d'une fonction "inverse" et d'une fonction logarithmique (1) de l'ampleur du risque (r).

Tableau 4.25

	r	r^2	inverse	log.10(r+1)	log. n(r+1)
Omission	0	0	0	0	0
Rép. cert. 1	2,25	5,06	0,44	0,51	1,18
Rép. cert. 2	2,75	7,56	0,80	0,57	1,32
Rép. cert. 3	3,75	14,06	1,06	0,68	1,56

Quand on dessine les échelles jointes, les mi-distances se placent dans l'ordre suivant :

Tableau 4.26

Echelles jointes							
	r	M01	M02	M03	M12	M13	M23
	r^2	M01	M02	M12	M03	M13	M23

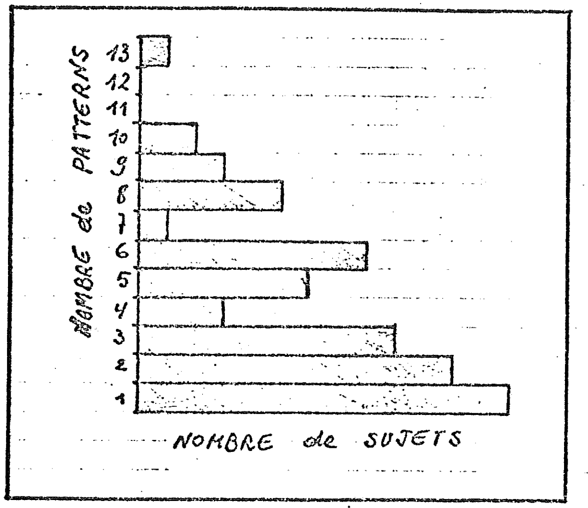
zones													
	1	2	3	4	5	6	7	8	9	10	11	12	13

(1) Voir le détail de ces fonctions dans le chapitre 2.

Les trois autres fonctions (fonction inverse et les deux fonctions logarithmiques) donnent le même ordre que la fonction r.

Sur les 920 *patterns* observés, 259 sont compatibles (soit 28,5 %) avec la fonction r. Sur 73 étudiants, 11 sujets ne présentent jamais de *pattern* compatible au cours des 13 tests. Pour les 62 autres sujets, les nombres de *patterns* présentés sont les suivants (tableau 4.27).

Tableau 4.27



Répartition du nombre de patterns compatibles avec la théorie du dépliage parmi les 62 sujets qui les ont présentés (année scolaire 1972-1973)

Dans ces 259 *patterns*, on dénombre les répétitions suivantes (tableau 4.28).

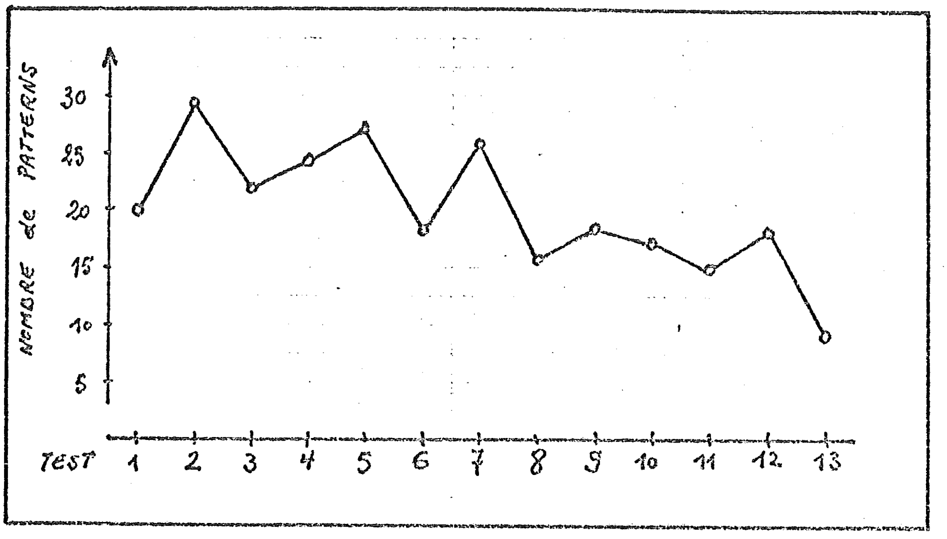
Tableau 4.28

125 patterns	13	48,2 %
9	12	3 %
31	11	12 %
19	10	7,3 %
23	9	8,8 %
6	8	2 %
24	7	9 %
4	6	1,5 %
10	5	3,8 %
6	4	2,3 %
2	3	0,75 %

Répartition des types de patterns compatibles avec la théorie du dépliage (année scolaire 1972-1973)

Si l'on observe le moment où se sont présentés les *patterns*, on obtient le tableau 4.29.

Tableau 4.29



Evolution, au cours des 13 tests successifs de l'année scolaire 1972-1973 des patterns compatibles avec la théorie du dépliage

Parmi les douze sujets qui ont présenté sept *patterns* et plus, un seul (S.1710) a présenté constamment le même *pattern* (dix fois le *pattern* 13). En dehors de cette exception, la diversité est de règle. Ainsi, le seul étudiant qui totalise à lui seul 13 *patterns* compatibles avec la stratégie 6, présente sept *patterns* différents : les *patterns* 6, 7, 8, 9, 10, 11 et 13. Les dix autres étudiants présentent chacun entre 3 et 6 *patterns* différents. La diversité dans les *patterns* est une constatation générale.

° ° °

Si l'on considère les échelles individuelles issues de la fonction r^2, elles sont semblables aux échelles issues de la fonction r, excepté pour les positions 6, 7 et 8, dénombrées 34 fois pour r et 18 fois seulement pour r^2.

Pour 53 sujets, r et r^2 sont aussi plausibles.
Pour 12 sujets, r est plus plausible que r^2.
Pour 7 sujets, r^2 est plus plausible que r.

Pour ces 19 derniers sujets, 12 ne font pencher la balance qu'avec un *pattern*. D'autres sujets (plus rares, évidemment) se présentent comme suit :

Tableau 4.30

Sujet 515		
test	r	r^2
1	-	-
2	7	-
3	7	-
4	-	-
5	-	-
6	7	-
7	6	-
8	-	-
9	-	-
10	7	-
11	-	-
12	-	-
13	7	-

Sujet 1708		
test	r	r^2
1	-	-
2	-	-
3	-	6
4	4	4
5	5	5
6	-	6
7	-	7
8	-	7
9	5	5
10	5	5
11	5	5
12	-	7
13	-	8

Deux exemples observés (année scolaire 1972-1973) de l'utilisation de patterns compatibles avec la théorie du dépliage

Ces deux sujets présentent un grand nombre de *patterns* compatibles avec une seule échelle jointe. En dehors des sujets 515 et 1708, il faut bien convenir que l'ensemble des résultats est défavorable à la théorie du dépliage, car, si 28,5 % des *patterns* sont compatibles, 71,5 % en revanche sont incompatibles. Cependant, les résultats acquis ne sont valables que pour une ampleur de risque allant de 0 à 6. Il se pourrait, - et c'est probable, - que la théorie fonctionne pour des ampleurs de risque beaucoup plus élevées.

Conclusions relatives aux stratégies non compatibles.

Aucune des six stratégies non compatibles n'a été utilisée dans des proportions importantes. Bien souvent, quand un *pattern* est compatible avec une de ces stratégies, il peut s'expliquer par d'autres variables, comme la facilité du test, ou par la stratégie E.S.U.

5. LA STRATEGIE COMPATIBLE AVEC LA VALIDITE DES INDICES DE CERTITUDE

Quand les hypothèses de l'expérience 1971-1972 diffèrent de celles de 1972-1973, ces dernières sont marquées d'un apostrophe.

<u>Rappel H16</u> (1971-72). Les taux moyens d'utilisation de l'omission et de la certitude 3 <u>augmentent</u> avec le numéro d'ordre du test, tandis que les TUC 1 et TUC 2 diminuent.

H16' (1972-73). Les taux moyens d'utilisation de chaque degré de certitude <u>restent stables</u> dans le temps, parce que la matrice des conséquences a été calculée selon le critère E.S.U.

<u>Rappel H17</u> (1971-72). Si le taux d'utilisation de certitudes fortes est très lié (cf. H2) à la facilité, alors les taux relatifs d'utilisation de l'omission (parmi les certitudes faibles) et de la certitude 3 (parmi les certitudes fortes) <u>augmentent</u> avec le temps.

H17' (1972-73). Les taux relatifs d'utilisation de l'omission (parmi les certitudes faibles) et de la certitude 3 (parmi les certitudes fortes) <u>restent stables</u> avec le temps.

<u>Rappel H19</u>. Les types d'évolution des TUC ne diffèrent pas selon les sujets : tous présentent une <u>disparition monotone</u> des certitudes 1 et 2.

H19' . Les types de stabilité des TUC ne diffèrent pas selon les sujets : les TUC sont constants chez les sujets quel que soit l'ordre du test (corrélation proche de 0).

VÉRIFICATION DES HYPOTHÈSES.

Utiliser le critère de l'E.S.U., c'est utiliser une stratégie qui ne dépend que de l'ESPER à chaque question. Dans une telle perspective, tous les niveaux de certitude sont utilisés, sauf compétence très élevée (et de façon homogène) ou compétence très faible (et de façon homogène). Il n'existe donc pas de *pattern* représentatif de cette stratégie.

Plus exactement, tous les *patterns*, à l'exception de certains de ceux que nous avons déjà analysés, sont compatibles avec la stratégie qui applique le critère de l'Espérance Subjective de l'Utilité (E.S.U.).

Dès lors, l'hypothèse H15 (voir expérience de 1971-72) n'a plus de sens

H15. Les *patterns* compatibles avec le critère E.S.U. sont plus nombreux que les *patterns* compatibles avec d'autres critères.

Il faut cependant examiner un *pattern* qui, quoique compatible avec la validité de l'ESPER, ne correspond pas à l'hypothèse H1. Il s'agit de la stratégie qui consiste à fournir la certitude 3 (critère maximax) pour des réponses dont l'ESPER dépasse un certain seuil (par exemple 50 %) et à omettre (critère maximin) pour les réponses dont l'ESPER est inférieure à ce seuil.

Il s'agit de la même dichotomie que dans l'expérience de l'année précédente (1971-72), à cette différence près qu'avec la présente matrice, ce comportement n'est plus adapté.

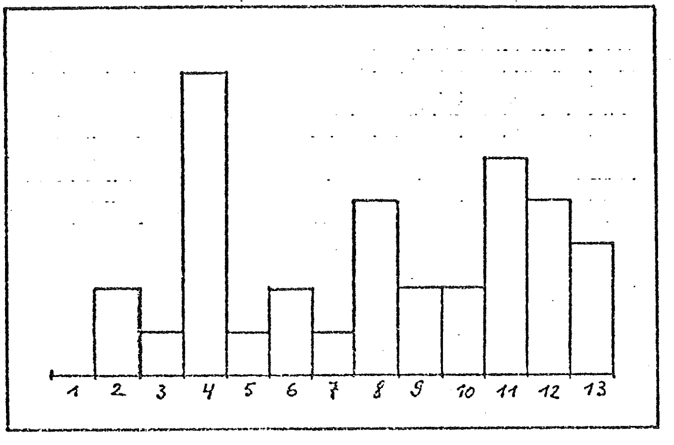
Le *pattern* "uniquement omissions et certitude 3" s'est présenté chez 17 sujets, dans 34 tests au total.

Pour un sujet (1035), dans 7 tests (5, 8, 9, 10, 11, 12, 13)
Pour un sujet (701), dans 5 tests
Pour un sujet (1312), dans 4 tests
Pour quatre sujets, dans 2 tests
Pour neuf sujets, dans un seul test.

On constate ici aussi un phénomène "groupe" puisque cinq sujets font partie de la classe 7 et neuf sujets de la classe 13.

Ces *patterns* sont apparus dans tous les tests, à l'exception du premier (tableau 4.31).

Tableau 4.31



Répartition temporelle du pattern constitué uniquement de l'omission et de certitude 3 (année scolaire 1972-1973)

Tableau 4.32

Test	Nombre d'items	Omissions %	Certitude 1 %	Certitude 2 %	Certitude 3 %
1	1665	9,96	15,91	24,62	49,48
2	1423	9,20	18,90	24,87	47,01
3	1510	10,52	16,82	24,56	48,07
4	1404	9,97	17,23	23,86	46,93
5	2449	7,79	17,06	21,31	53,81
6	1362	8,73	19,45	23,71	45,66
7	1432	8,31	18,85	25,69	47,13
8	1514	7,66	12,74	18,36	61,32
9	1390	12,08	18,05	23,30	46,54
10	1475	10,37	17,76	21,15	50,71
11	1404	9,47	17,09	20,44	52,99
12	1420	14,36	22,18	24,08	39,36
13	1258	12,63	18,68	19,71	48,96

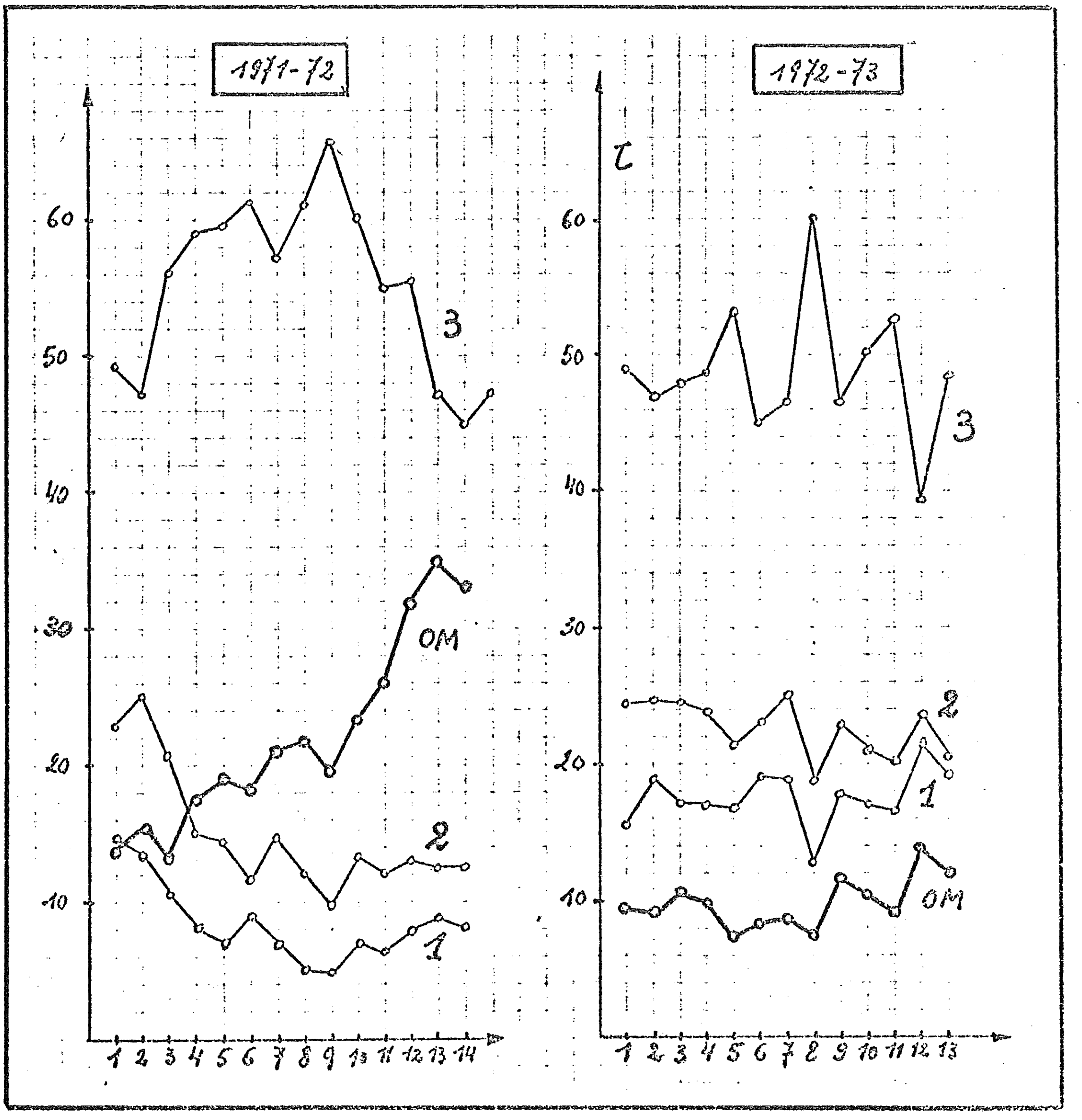
POURCENTAGES D'UTILISATION DE CHAQUE DEGRE DE CERTITUDE

AU COURS DES 13 EPREUVES SUCCESSIVES DE MECANIQUE (ANNEE 1972 - 1973)

H16'A. Les taux d'utilisation de chaque degré de certitude restent stables dans le temps (contrairement à l'expérience 1971-72), à facilité égale.

Les tableaux 4.32 et 4.33 permettent de constater que cette hypothèse se vérifie.

Tableau 4.33



Evolution comparée du taux d'utilisation de chaque degré de certitude au cours des deux années scolaires (1971-1972 et 1972-1973)

Comme indices de stabilité, on considérera non seulement la corrélation linéaire entre l'ordre des tests et le TUC (taux d'utilisation de chaque indice), mais aussi le coefficient de pente de la droite de régression (des TUC sur l'ordre des tests). Ces coefficients sont soulignés dans le tableau 4.34.

Tableau 4.34

1971-1972	1972-1973
$\tau'_0 = 1{,}583\ x + 9{,}779\ (r = .939)$	$\tau'_0 = 0{,}255\ x + 8{,}299\ (r = .505)$
$\tau'_1 = -0{,}411\ x + 11{,}64\ (r = -.624)$	$\tau'_1 = 0{,}170\ x + 16{,}557\ (r = .305)$
$\tau'_2 = -0{,}833\ x + 21{,}27\ (r = -.757)$	$\tau'_2 = -0{,}326\ x + 25{,}027\ (r = -.556)$
$\tau'_3 = -0{,}339\ x + 57{,}291\ (r = -.212)$	$\tau'_3 = -0{,}064\ x + 49{,}810\ (r = -.065)$

Equations de régression du taux d'utilisation d'un degré de certitude (τ) sur l'ordre des tests (x)

Les coefficients de corrélation et les coefficients de régression sont systématiquement plus faibles pour les données de 1972-73 que pour celles de 1971-72. La variance des TUC dans le temps est, elle aussi, systématiquement plus faible, comme le montre le tableau 4.35.

Tableau 4.35

1971-1972 Moyennes	1971-1972 Sigmas		1972-1973 Moyennes	1972-1973 Sigmas
21,65	7,05	τ_0	10,15	2,08
8,55	2,75	τ_1	17,74	2,17
15,02	4,60	τ_2	22,74	2,28
54,74	6,70	τ_3	49,22	5,07

Paramètres des taux d'utilisation de chaque degré de certitude à chaque test

Les taux d'utilisation bruts (tableau 4.7) ont été prédits à partir de la facilité de chaque test. Le tableau 4.36_a reprend ces valeurs pour l'expérience de 1971-72; le tableau 4.36_b reprend ces valeurs pour l'expérience de 1972-73.

1971-72   X = facilité

MX = 56.219 MY = 10.081 SX = 7.224 SY = 1.888
SXY = -95.674 SX2 = 678.340 SY2 = 46.327
A = -0.141 B = 18.010 R = -0.540

	X	Y	Y'	RESIDUS
1	59.0991	9.9600	9.6746	0.2854
2	46.8728	9.2000	11.3990	-2.1990
3	57.1523	10.5200	9.9492	0.5708
4	52.4929	9.9700	10.6063	-0.6363
5	62.9236	7.7900	9.1352	-1.3452
6	50.7342	8.7300	10.8544	-2.1244
7	60.4749	8.3100	9.4805	-1.1705
8	65.1915	7.6600	8.8153	-1.1553
9	50.0719	12.0800	10.9478	1.1322
10	59.6610	10.3700	9.5953	0.7747
11	63.1766	9.4700	9.0995	0.3705
12	40.2817	14.3600	12.3286	2.0314
13	62.7186	12.6300	9.1641	3.4659

Y = TAUX D'UTILISATION de l'OMISSION

MX = 56.219 MY = 22.743 SX = 7.224 SY = 2.195
SXY = -124.588 SX2 = 678.340 SY2 = 62.619
A = -0.184 B = 33.069 R = -0.605

	X	Y	Y'	RESIDUS
1	59.0991	24.6200	22.2141	2.4059
2	46.8728	24.8700	24.4597	0.4103
3	57.1523	24.5600	22.5717	1.9883
4	52.4929	23.8600	23.4274	0.4326
5	62.9236	21.3100	21.5117	-0.2017
6	50.7342	23.7100	23.7505	-0.0405
7	60.4749	25.6900	21.9614	3.7286
8	65.1915	18.3600	21.0951	-2.7351
9	50.0719	23.3000	23.8721	-0.5721
10	59.6610	21.1500	22.1109	-0.9609
11	63.1766	20.4400	21.4652	-1.0252
12	40.2817	24.0800	25.6702	-1.5902
13	62.7186	19.7100	21.5493	-1.8393

Y = TAUX d'UTILISATION de C2

MX = 56.219 MY = 17.748 SX = 7.224 SY = 2.085
SXY = -140.158 SX2 = 678.340 SY2 = 56.535
A = -0.207 B = 29.364 R = -0.716

	X	Y	Y'	RESIDUS
1	59.0991	15.9100	17.1526	-1.2426
2	46.8728	18.9000	19.6788	-0.7788
3	57.1523	16.8200	17.5549	-0.7349
4	52.4929	17.2300	18.5176	-1.2876
5	62.9236	17.0600	16.3624	0.6976
6	50.7342	19.4500	18.8810	0.5690
7	60.4749	18.8500	16.8684	1.9816
8	65.1915	12.7400	15.8938	-3.1538
9	50.0719	18.0500	19.0178	-0.9678
10	59.6610	17.7600	17.0365	0.7235
11	63.1766	17.0900	16.3101	0.7799
12	40.2817	22.1800	21.0407	1.1393
13	62.7186	18.6800	16.4048	2.2752

Y = TAUX d'UTILISATION de C1

MX = 56.219 MY = 49.221 SX = 7.224 SY = 4.878
SXY = 373.772 SX2 = 678.340 SY2 = 309.370
A = 0.551 B = 18.243 R = 0.816

	X	Y	Y'	RESIDUS
1	59.0991	49.4800	50.8075	-1.3275
2	46.8728	47.0100	44.0707	2.9393
3	57.1523	48.0700	49.7348	-1.6648
4	52.4929	48.9300	47.1674	1.7625
5	62.9236	53.8100	52.9149	0.8951
6	50.7342	45.6600	46.1984	-0.5384
7	60.4749	47.1300	51.5656	-4.4356
8	65.1915	61.2200	54.1646	7.0555
9	50.0719	46.5400	45.8335	0.7065
10	59.6610	50.7100	51.1172	-0.4072
11	63.1766	52.9900	53.0543	-0.0643
12	40.2817	39.2600	40.4390	-1.0790
13	62.7186	48.9600	52.8019	-3.8419

Y = TAUX d'UTILISATION de C3

1970-71

MX = 50.432 MY = 21.654 SX = 8.660 SY = 6.799
SXY = -734.122 SX2 = 1050.030 SY2 = 647.032
A = -0.699 B = 56.913 R = -0.891

	X	Y	Y'	RESIDUS
1	61.9130	13.6500	13.6269	0.0231
2	50.4160	14.3900	21.6650	-7.2750
3	59.1200	12.4800	15.5737	-3.0937
4	57.7600	17.5200	16.5305	0.9895
5	60.7782	18.3600	14.4203	3.9397
6	52.2500	18.0000	20.3828	-2.3828
7	52.0034	20.8800	20.5552	0.3248
8	50.3496	21.5900	21.7114	-0.1214
9	60.2365	18.9100	14.7291	4.1309
10	46.2331	23.3900	24.5834	-1.1834
11	40.9314	25.4900	28.2961	-2.8061
12	37.1429	31.8300	30.9448	0.8852
13	36.9000	34.5000	31.1146	3.3854
14	39.9200	32.1600	29.0032	3.1568

X = FACILITE Y = TAUX d' OM

X = FACILITE Y = TAUX de C1

MX = 50.432 MY = 8.559 SX = 8.660 SY = 2.652
SXY = 73.249 SX2 = 1050.030 SY2 = 98.488
A = 0.070 B = 5.040 R = 0.228

	X	Y	Y'	RESIDUS
1	61.9130	14.5200	9.3594	5.1606
2	50.4160	13.7200	8.5574	5.1626
3	59.1280	10.5600	9.1646	1.3954
4	57.7600	8.0000	9.0697	-1.0697
5	60.7782	7.6200	9.2803	-1.6603
6	52.2500	9.1600	8.6854	0.4746
7	52.0034	6.9900	8.6682	-1.6782
8	50.3496	5.3300	8.5529	-3.2229
9	60.3365	5.2900	9.2535	-3.9635
10	46.2331	7.4500	8.2656	-0.8156
11	40.9314	6.2900	7.8958	-1.6058
12	37.1429	8.0000	7.6315	0.3685
13	36.9000	8.9000	7.6146	1.2854
14	39.9200	8.0000	7.8252	0.1748

MX = 50.432 MY = 15.028 SX = 8.660 SY = 4.436
SXY = 195.413 SX2 = 1050.030 SY2 = 275.525
A = 0.186 B = 5.642 R = 0.363

	X	Y	Y'	RESIDUS
1	61.9130	22.6000	17.1644	5.4356
2	50.4160	25.2900	15.0248	10.2652
3	59.1200	20.9600	16.6446	4.3154
4	57.7600	15.2800	16.3915	-1.1115
5	60.7782	14.3100	16.9532	-2.6432
6	52.2500	11.5000	15.3661	-3.8661
7	52.0034	14.9100	15.3202	-0.4102
8	50.3496	11.9700	15.0125	-3.0425
9	60.3365	9.7700	16.8710	-7.1010
10	46.2331	13.7100	14.2464	-0.5364
11	40.9314	12.2500	13.2597	-1.0097
12	37.1429	13.3000	12.5546	0.7454
13	36.9000	12.3000	12.5094	-0.2094
14	39.9200	12.2400	13.0715	-0.8315

X = FACILITE Y = TAUX de C2

X = FACILITE Y = TAUX de C3

MX = 50.432 MY = 54.747 SX = 8.660 SY = 6.461
SXY = 465.119 SX2 = 1050.030 SY2 = 584.405
A = 0.443 B = 32.408 R = 0.594

	X	Y	Y'	RESIDUS
1	61.9130	49.2100	59.8325	-10.6225
2	50.4160	46.5500	54.7333	-8.1833
3	59.1200	56.0000	58.5953	-2.5953
4	57.7600	59.2000	57.9929	1.2071
5	60.7782	59.6800	59.3298	0.3502
6	52.2500	61.3300	55.5522	5.7778
7	52.0034	57.2000	55.4430	1.7570
8	50.3496	61.1000	54.7104	6.3896
9	60.3365	66.0200	59.1342	6.8858
10	46.2331	55.4300	52.8870	2.5430
11	40.9314	55.9600	50.5385	5.4215
12	37.1429	46.8500	48.8604	-2.0104
13	26.9000	44.3000	48.7528	-4.4528
14	39.9200	47.6000	50.0905	-2.4905

H16'B. L'hypothèse H16' ne se vérifie pleinement qu'après contrôle de la variable "facilité du test".

Rappelons que la facilité des tests a oscillé autour de la moyenne générale (56,63 %) de façon horizontale, si bien que la corrélation entre le numéro d'ordre du test et sa facilité est quasi nulle (r = 0,066), comme on le voit dans le tableau 4.4. De plus, avec la nouvelle matrice, la facilité 50 % a perdu son caractère de "charnière".

La facilité du test présente les corrélations suivantes avec les taux d'utilisation

- de l'omission	r = -.540
- de la certitude 1	r = -.716
- de la certitude 2	r = -.605
- de la certitude 3	r = .816
- des certitudes fortes	r = .706

Le taux d'utilisation de chaque degré de certitude (TUC ou τ) ont été prédits (τ') à partir de la facilité (f) de chaque test.

Les équations sont fournies dans le tableau 4.37.

Tableau 4.37

1971-72		1972-73	
$\tau'0 = -0,699\ f + 56,913$	(r= -.891)	$\tau'0 = -.141\ f + 18,01$	(r= -.540)
$\tau'1 = 0,070\ f + 5,040$	(r= .228)	$\tau'1 = -.207\ f + 29,36$	(r= -.716)
$\tau'2 = 0,168\ f + 5,642$	(r= .363)	$\tau'2 = -.184\ f + 33,06$	(r= -.605)
$\tau'3 = 0,443\ f + 32,408$	(r= .594)	$\tau'3 = .551\ f + 18,24$	(r = .816)

Equations de prédiction des taux d'utilisation de chaque degré de certitude à chaque test à partir de la facilité du test

Le tableau 4.38 contient les corrélations partielles entre l'ordre des tests et les TUC.

Tableau 4.38

	1971-1972		1972-1973	
	r de départ	r partielles	r de départ	r partielles
TU om.	.939	.810	.505	.639
TUCert.1	-.624	-.782	.305	.428
TUCert.2	-.754	-.861	-.556	-.645
TUCert.3	-.211	-.595	-.065	-.200

Estimation (par corrélation partielle) de la liaison réelle entre le taux d'utilisation des divers degrés de certitude et l'ordre des tests, après élimination de la facilité du test

Pour l'expérience 1972-73, les corrélations partielles restent inférieures à celles de 1971-72; elles sont plus élevées que les corrélations de départ. Il faut cependant considérer l'ampleur de la corrélation à la lumière du coefficient de pente de la régression des TUC sur l'ordre (facilité contrôlée).

Aux résidus de la prédiction, nous avons ajouté la moyenne (sur l'ensemble des tests) des taux d'utilisation de la certitude considérée. La dispersion des résidus de prédiction à partir des équations du tableau 4.37 est nettement plus faible en 1972-73 qu'en 1971-72, comme le montre le tableau 4.39.

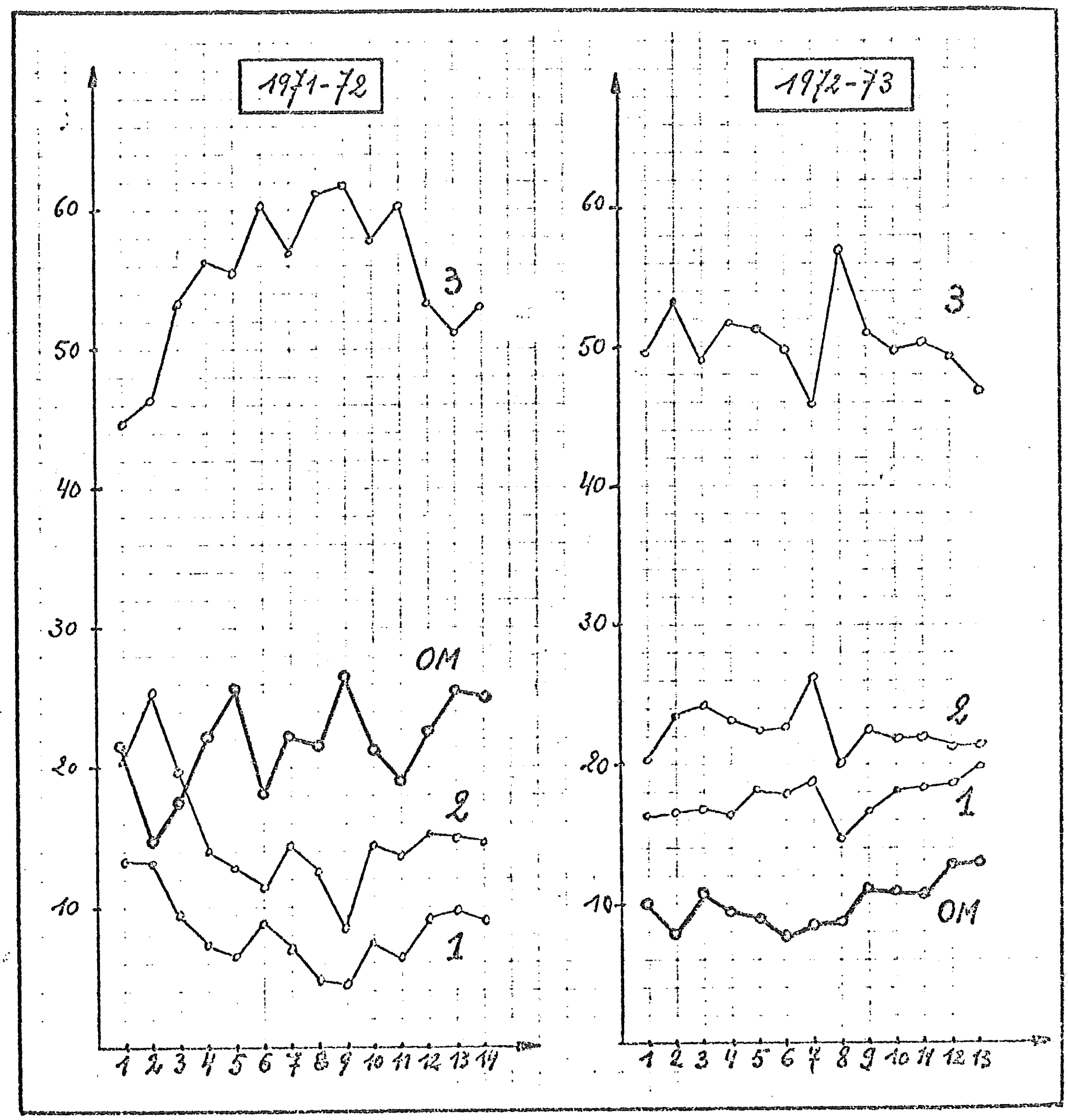
Tableau 4.39

	1971-72		1972-73		F	Signif.
	σ	σ^2	σ	σ^2		
$\tau'_0-\tau_0$	3,09	9,54	1,58	2,49	3,83	P.05
$\tau'_1-\tau_1$	2,57	6,60	1,45	2,10	3,14	P.05
$\tau'_2-\tau_2$	4,13	17,05	1,74	3,02	5,64	P.01
$\tau'_3-\tau_3$	5,19	26,93	2,82	7,95	3,38	P.05

Variances des résidus de prédiction du taux d'utilisation de chaque degré de certitude par l'ordre du test. Comparaison entre les variances de l'année scolaire 1971-1972 et celles de l'année scolaire 1972-1973

Les $\tau'-\tau$ (résidus des TUC prédits après contrôle de la facilité du test) sont assez différents des τ (TUC observés). Les τ' des deux expériences sont présentés dans le tableau 4.40.

Tableau 4.40



Graphique des résidus $\tau'-\tau$. Les τ' sont calculés par les équations du tableau 4.37. Aux résidus sont ajoutés la $M\tau$ sur l'ensemble des tests

En examinant les tableaux 4.40 et 4.41, on notera combien les valeurs initiales (celles du premier test) des taux d'utilisation de chaque indice sont proches en 1971-72 et 1972-73 : .

Tableau 4.41

	1971-72	1972-73
Cert. 1	14,52	15,91 %
Cert. 2	22,60 %	24,62 %
Cert. 3	49,21 %	49,48 %

Valeurs de τ lors du premier test : valeurs initiales de τ

Ces taux peuvent être mis en parallèle car, par chance, les premiers tests de chaque année avaient une facilité fort proche : 61,91 % en 1971-72 et 59,09 % en 1972-73. Il en va autrement pour l'omission : respectivement 13,65 % et 9,96 %. La différence est vraisemblablement attribuable à la procédure de pénalisation des omissions.

On se rappellera qu'en 1971-72, les trois premiers tests avaient constitué une période d'adaptation et qu'un comportement stable inspiré des critères E.S.U. pouvait être observé à partir du quatrième test. Dès lors, il est intéressant de considérer (tableau 4.42) la dispersion des TUC (1) des onze derniers tests de l'expérience de l'année scolaire 1971-72.

Tableau 4.42

	$\mathcal{M}$	σ^2 1971-72 (11 tests)	σ^2 1972-73 (13 tests)	F	Signif.
OM.	23,87	5,58	2,49	2,24	NS
Cert. 1	7,36	2,35	2,10	1,11	NS
Cert. 2	12,86	4,46	3,02	1,47	NS
Cert. 3	55,87	13,78	7,95	1,73	NS

Paramètres ($\mathcal{M}$, S) des 11 derniers tests de l'année scolaire 1971-1972. Comparaison avec la variance des 13 tests de l'année scolaire 1972-1973

On remarque que les dispersions sont du même ordre en 1971-72 qu'en 1972-73 (les F sont non significatifs).

(1) Ou, plus exactement, des résidus de leur prédiction par la facilité du test, additionnée de la moyenne.

Les TUC de 1972-73 (ou plutôt les résidus τ' après prédiction par la facilité) sont-ils en augmentation ou, selon l'hypothèse H16', stables ?

Pour répondre à cette question, on a calculé la régression de ces taux (τ') sur l'ordre des tests (x). Ensuite, la valeur de l'intervalle de confiance du coefficient de x (coefficient de pente de l'équation $\tau' = ax + b$) a été calculée par les formules suivantes :

Erreur standard de a : $\sqrt{\hat{\sigma}^2_{\tau x} \;/\; SCE_x}$

Intervalle de confiance de a = $a \pm 1{,}96 \sqrt{\hat{\sigma}^2_{\tau x} \;/\; SCE_x}$

où $\hat{\sigma}^2_{\tau x} = (1-r^2)\, SCE_\tau \,/\, (n-2)$ (DAGNELIE P., 1970,p. 268

Ces valeurs sont rassemblées dans le tableau 4.43.

Tableau 4.43

	Omissions	C1	C2	C3
Σx^2	182	182	182	182
Σy^2	46,327	56,535	62,619	309,370
Σxy	46,33	30,96	- 59,39	- 15,33
r	.505	.305	-.556	-.065
r^2	.255	.093	.309	.003
a	+ 0,255	+ 0,170	- 0,326	- 0,084
E S(a)	0,12	0,15	0,14	0,38
x 2,201	0,264*	0,330*	0,308	0,836*
x 3,106			0,434*	

CALCUL DE LA SIGNIFICATION DU COEFFICIENT DE PENTE

(ANNEE SCOLAIRE 1972-73)

Les signes * indiquent les valeurs de l'intervalle de confiance supérieures (en valeur absolue) au coefficient de régression a;

On constate donc que l'hypothèse H16a (H20) ne peut être rejetée à P.05 pour les certitudes 1 et 3 et à P.01 pour l'omission et la certitude 2.

La même procédure a été appliquée pour les onze derniers tests de l'expérience 1971-72 (voir tableau 4.44). Selon l'hypothèse H16', après une période d'adaptation (les trois premiers tests), les TUC se stabilisent.

Tableau 4.44

	Omissions	C1	C2	C3
Σx^2	110	110	110	110
Σy^2	61,48	25,95	49,16	151,65
Σxy	11,12	20,76	26,24	57,57
r	.1349	.394	.356	.445
r^2	.018	.155	.126	.198
a	0,101	0,188	0,238	0,523
E.S. (a)	0,25	0,15	0,21	0,35
x 2,179	0,544 *	0,326*	0,457*	0,762*
x 1,356	0,339 *	0,203*	0,284 *	0,474

CALCUL DE LA SIGNIFICATION DU COEFFICIENT DE PENTE

(ANNEE SCOLAIRE 1971-72)

Ainsi, l'hypothèse (H16) ne peut être rejetée à P.05 pour tous les degrés de certitude. On voit même qu'elle ne peut non plus être rejetée à P.20, excepté pour la certitude 3.

Une façon différente d'estimer la plausibilité de l'hypothèse nulle consisterait à diviser le coefficient de pente par sa propre erreur standard. La valeur ainsi trouvée serait reportée directement dans la table des valeurs t.

H17'. Les taux relatifs d'utilisation de l'omission et de la certitude 3 restent stables avec le temps.

Le tableau 4.45 présente ces TUC relatifs.

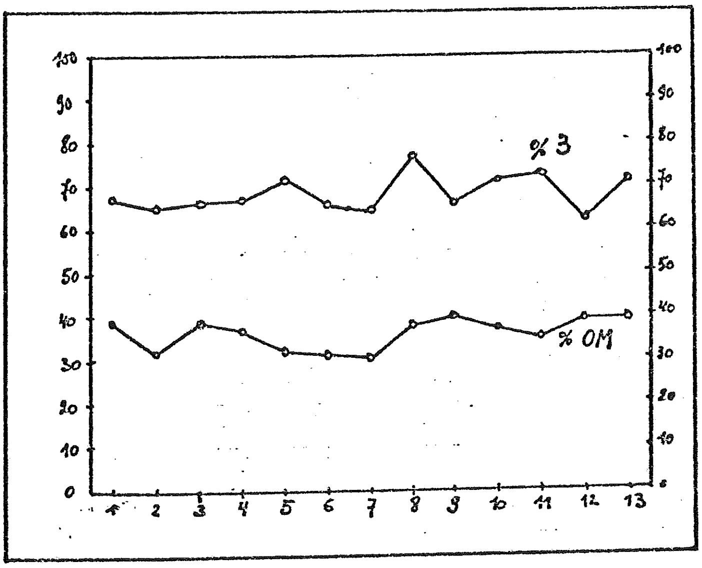
Tableau 4.45

	OM	Certitude 1	Certitudes faibles	% OM	Certitude 2	Certitude 3	Certitudes fortes	% certitude 3
1	9,96	15,91	25,87	38,50	24,62	49,48	74,10	66,77
2	9,20	18,90	28,10	32,7	24,87	47,01	71,88	65,41
3	10,52	16,82	27,34	38,47	24,56	48,07	72,63	66,18
4	9,97	17,23	27,20	36,65	23,86	48,93	72,79	67,22
5	7,79	17,06	24,85	31,34	21,31	53,81	75,12	71,63
6	8,73	19,45	28,18	30,97	23,71	48,09	71,80	66,97
7	8,31	18,85	27,16	30,59	25,69	47,13	72,82	64,72
8	7,66	12,74	20,40	37,54	18,36	61,22	79,58	76,92
9	12,08	18,05	30,13	40,09	23,30	46,54	69,84	66,63
10	10,37	17,76	28,13	36,86	21,15	50,71	71,86	70,86
11	9,47	17,09	26,56	35,65	20,44	52,99	73,43	72,16
12	14,36	22,18	36,54	39,29	24,08	39,36	63,44	62,04
13	12,63	18,68	31,31	39,47	19,71	48,96	68,67	71,29
M	10,08			36,00				68,25

POURCENTAGE D'UTILISATION DE CHAQUE DEGRE DE CERTITUDE LORS DES 13 EPREUVES SUCCESSIVES DE MECANIQUE (ANNEE SCOLAIRE 1972-73)

Les données du tableau 4.45 ont été portées sur graphique (tableau 4.46).

Tableau 4.46



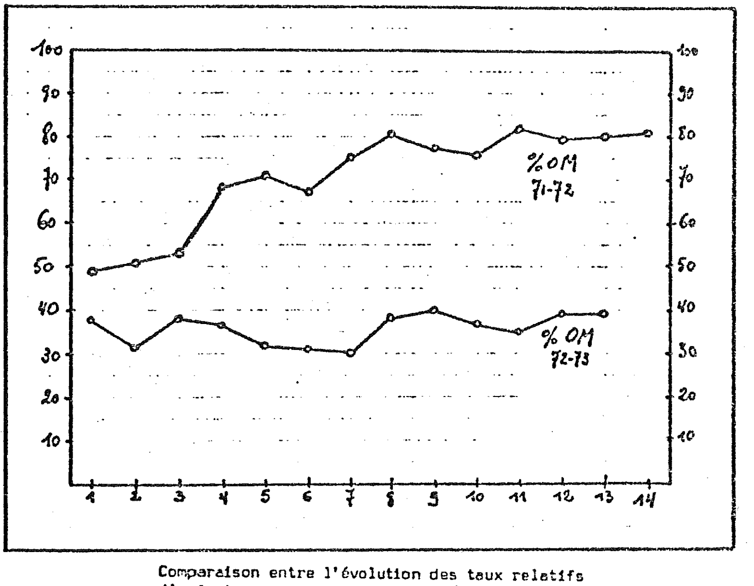
Taux relatifs d'omissions (par rapport aux certitudes faibles) (année scolaire 1972-1973)
Taux relatifs de certitudes 3 (par rapport aux certitudes fortes) (année scolaire 1972-1973)

On ne constate aucune tendance vers l'augmentation, contrairement à l'année précédente.

Rappelons que ce recours à des taux relatifs n'est qu'une manière imparfaite de neutraliser la variable "facilité du test" (qui n'est corrélée qu'à .70 avec l'utilisation de certitudes fortes).

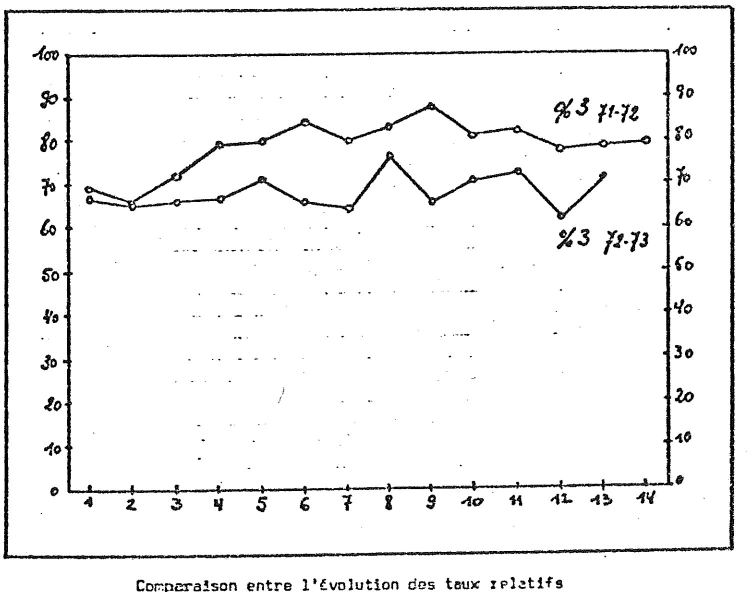
Les tableaux 4.47 et 4.48 permettent de comparer l'évolution entre les TUC relatives lors des deux expériences.

Tableau 4.47



Comparaison entre l'évolution des taux relatifs d'omission au cours de l'année scolaire 1971-1972 et cette évolution en 1972-1973

Tableau 4.48



Comparaison entre l'évolution des taux relatifs de certitude 3 au cours de l'année scolaire 1971-1972 et cette évolution en 1972-1973

H17'B. L'hypothèse H17' ne se vérifie pleinement qu'après contrôle de la facilité du test.

Les corrélations partielles (la facilité contrôlée) des TUC avec l'ordre sont les suivantes (tableau 4.50).

	1971-1972		1972-1973	
	r partielles	(r anciennes)	r partielles	(r anciennes)
ordre/% rel. d'OM	.859	(.885)	.330	(.323)
ordre/% rel. Cert.3	.803	(.556)	.314	(.242)

Tableau 4.50 Estimation (par corrélation partielle) de la liaison réelle entre les taux relatifs (d'utilisation des omissions et des certitudes 3) et l'ordre des tests, après élimination de la facilité du test

Les équations de régression des taux relatifs sur l'ordre ont de nouveau été calculées (tableau 4.51).

Tableau 4.51

	Taux rel. d'omissions		Taux rel. de certitude 3	
	1971-1972	1972-1973	1971-1972	1972-1973
$\sum x^2$	227,50	182,00	227,50	182,00
$\sum y^2$	1704,23	143,03	475,52	187,11
$\sum xy$	550,94	52,22	182,90	44,04
r	.8854	.3237	.5564	.2426
r^2	.7839	.1047	.3095	.0588
a	2,4217	0,2669	0,8039	0,241
E.S. (a)	0,367	0,253	0,129	0,295
$x\alpha$ (P.05)	0,779	0,556*	0,281	0,649*
$x\alpha$ (P.01)	1,121		0,394	

CALCUL DE LA SIGNIFICATION DES COEFFICIENTS DE PENTE

Les signes * indiquent les valeurs de l'intervalle de confiance supérieures (en valeur absolue) au coefficient de régression a.

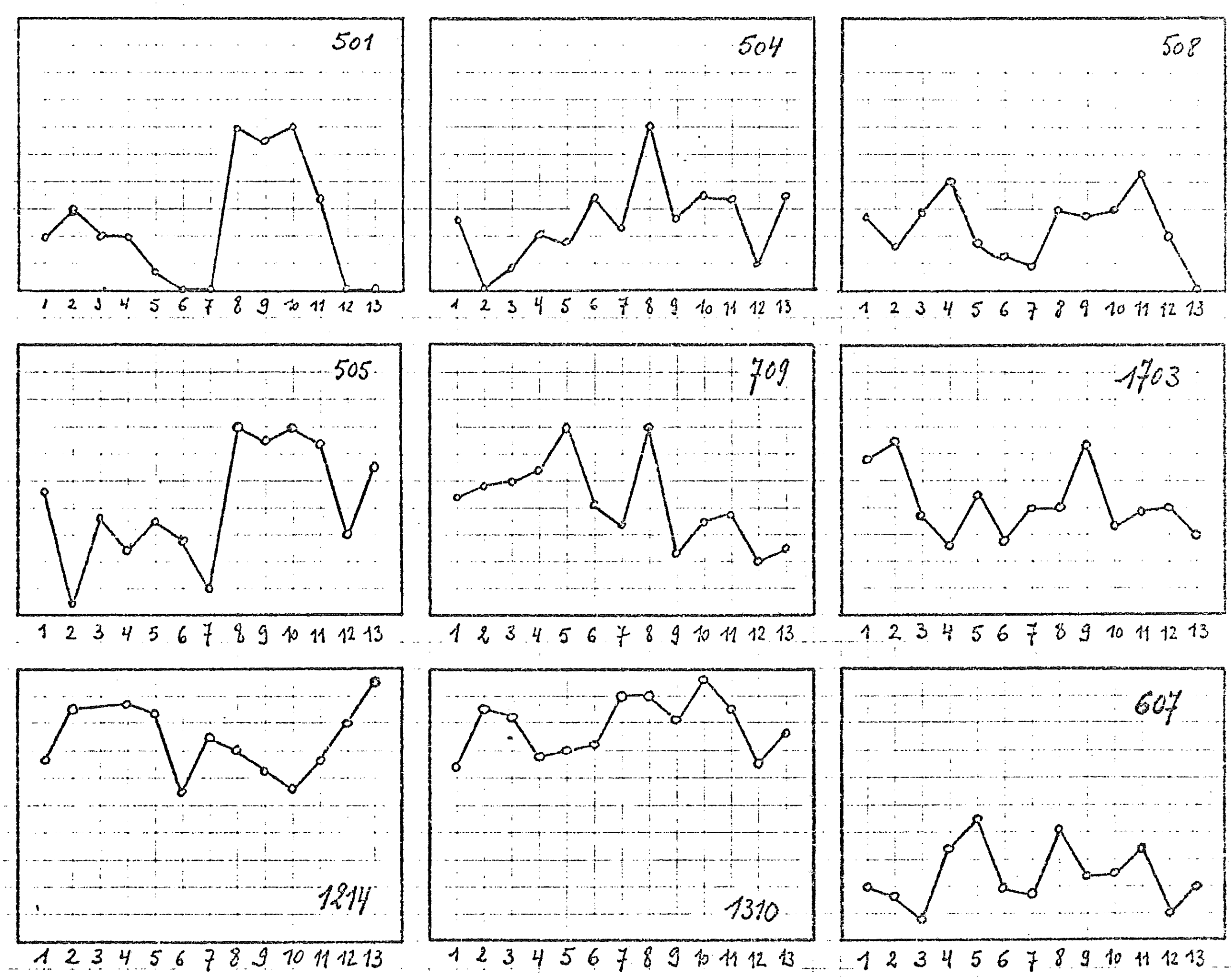
On voit donc que l'hypothèse H17' ne peut être rejetée au seuil de P.05 pour l'expérience 1972-1973. Les valeurs des coefficients de pente (2,421 et 0,803) de l'expérience 1971-1972 permettent de rejeter H17'b au seuil de P.01.

dents de scie : peut-on parler de stabilité ?

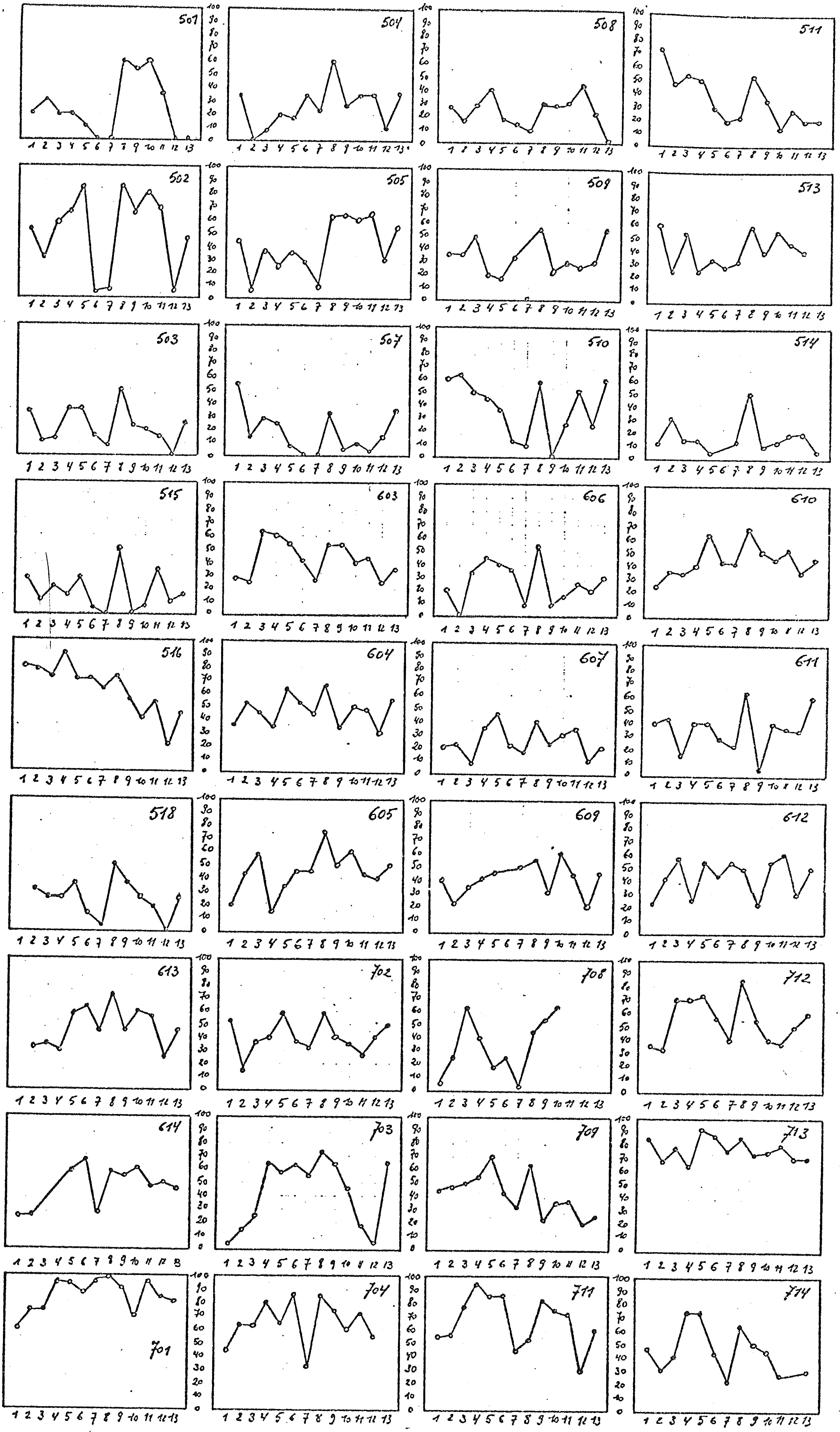
H19'. Les types de stabilité diffèrent selon les sujets : monotonie (légèrement) ascendante ou descendante, dents de scie (mais régression horizontale), etc. 

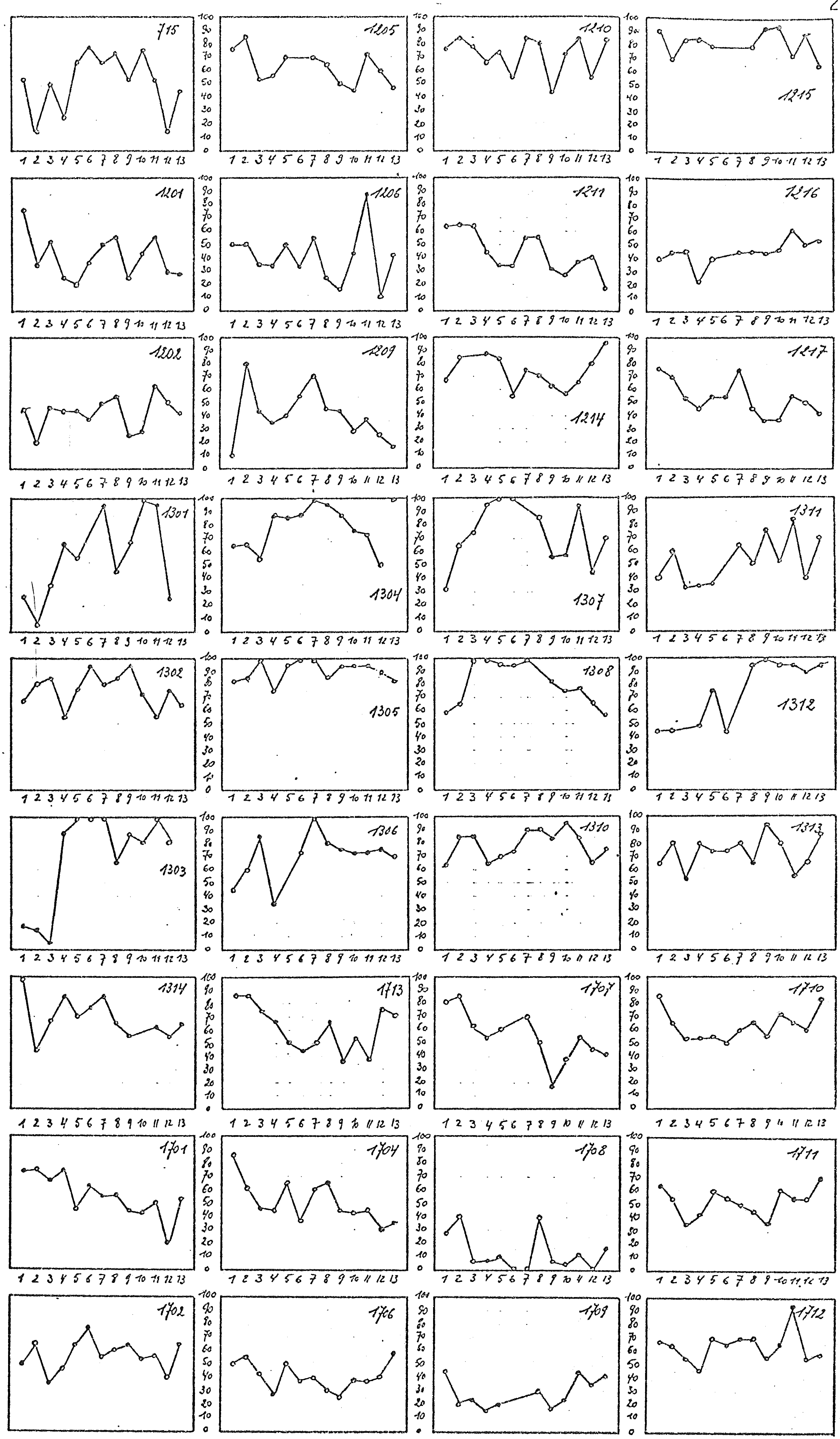
Les taux d'utilisation de chaque degré à chaque test, pour chaque sujet, sont présentés en annexe 7.

Voici, à titre d'illustration, puisés dans les six classes différentes, neuf exemples de l'évolution du taux d'utilisation de la certitude 3 (tableau 4.52). Ils sont à l'image de l'ensemble des résultats. Que le taux d'utilisation soit élevé (sujets 1214 et 1310), moyen (sujets 505 et 1703) ou faible (sujets 504 et 607), la régression est proche de l'horizontale. Certains montrent de légères diminutions au fil des tests (sujet 709), d'autres des "pointes" passagères (sujet 501).



L'hypothèse H19' se confirme donc.


501
504
508
511
502
505
509
513
503
507
510
514
515
603
606
610
516
604
607
611
518
605
609
612
613
702
708
712
614
703
709
713
701
704
711
714


715
1205
1210
1215
1201
1206
1211
1216
1202
1209
1214
1217
1301
1304
1307
1311
1302
1305
1308
1312
1303
1306
1310
1313
1314
1713
1707
1710
1701
1704
1708
1711
1702
1706
1709
1712

Comme pour l'expérience de l'année précédente (1971-72), les corrélations entre les scores simples (SS) et "avec certitude" (SC) ont été calculées (voir tableau 4.53) grâce au programme FORTRAN du tableau 3.

Les corrélations du tableau 4.53 sont rassemblées synoptiquement dans le tableau 4.54.

Tableau 4.54

	Barèmes de cotation						
	A	B	C	D	E	F	G
Test n° 1	.951	.946	.927	.955	.912	.858	.973
2	.976	.947	.881	.931	.900	.891	.973
3	.895	.909	.897	.938	.847	.737	.954
4	.960	.947	.914	.951	.917	.836	.976
5	.900	.866	.818	.898	.799	.705	.955
6	.825	.833	.796	.877	.711	.624	.907
7	.965	.951	.926	.956	.931	.877	.983
8	.931	.911	.871	.924	.863	.778	.964
9	.944	.926	.882	.934	.888	.805	.973
10	.951	.913	.855	.918	.851	.845	.963
11	.944	.881	.811	.882	.838	.853	.954
12	.921	.905	.839	.913	.782	.729	.942
13	.936	.871	.776	.869	.861	.766	.965
Corrélations théoriques	.865	.802	.684	.878	.919	.805	.976

Corrélations entre les scores simples et les scores calculés selon chacun des six barèmes de certitude (voir tableaux 3.18 et 3.21) et du barème de la présente expérience (barème G) pour chacun des 13 tests.

On constate que, pour les barèmes A, B et C, les corrélations observées sont toujours supérieures aux corrélations théoriques, tendance plus nette que dans l'expérience de l'année précédente. De nombreux cas contraires s'observent dans les barèmes E et F (un seul pour le barème D), ainsi que pour le barème G où toutes les corrélations observées sont très élevées.

Tableau 4.53

1	71	13.859	14.015	21.254	28.491	28.507	5.831	-19.873	23.106
		708.586	12514.211	20133.328	30605.059	13093.660	39795.824	94334.937	7198.645
			2832.546	3574.528	4316.492	2908.053	4841.262	7013.219	2197.799
			0.951	0.946	0.927	0.955	0.912	0.858	0.973
2	73	9.137	1.655	6.192	10.729	14.753	-11.151	-41.945	11.634
		782.624	12263.598	15961.187	22161.484	10239.484	29597.203	87755.125	6350.199
			3022.443	3346.073	3669.703	2635.458	4331.488	7384.367	2168.406
			0.976	0.947	0.881	0.931	0.900	0.891	0.973
3	70	12.329	11.471	18.286	25.100	25.257	4.757	-20.243	20.336
		579.438	10832.965	15962.172	23183.016	9953.305	32998.730	98268.000	5749.820
			2243.549	2765.420	3287.291	2252.077	3704.576	5563.547	1741.522
			0.895	0.909	0.897	0.938	0.847	0.737	0.954
4	72	10.236	6.307	10.569	14.832	17.889	-4.792	-28.958	14.806
		578.981	9246.914	12787.555	17760.238	8151.047	24633.742	60460.746	5015.000
			2222.274	2577.311	2932.349	2064.881	3464.451	5064.262	1663.549
			0.960	0.947	0.914	0.951	0.917	0.856	0.976
5	73	21.110	23.737	34.192	44.646	44.014	14.192	-21.986	35.983
		1659.115	18901.750	29289.199	44603.449	20458.895	53433.137	126739.562	12082.578
			5042.141	6039.387	7036.660	5232.836	7519.379	10216.801	4274.363
			0.900	0.866	0.818	0.898	0.799	0.705	0.955
6	64	10.797	4.900	9.906	14.912	18.531	-7.609	-38.313	15.023
		338.359	6384.145	8491.398	12127.453	5129.926	18817.219	69957.562	3133.961
			1212.693	1412.781	1612.855	1155.906	1795.078	3035.937	933.548
			0.825	0.833	0.796	0.877	0.711	0.624	0.907
7	68	12.735	12.972	18.809	24.646	25.382	5.926	-18.353	21.154
		847.230	12411.031	19988.414	30347.316	13545.973	35603.492	97090.750	7590.590
			3128.787	3911.552	4694.293	3237.875	5112.637	7951.582	2493.021
			0.965	0.951	0.926	0.956	0.931	0.877	0.983
8	72	13.708	16.822	24.222	31.622	29.931	11.819	-7.944	23.934
		398.872	5041.004	8204.367	12897.754	5788.590	16494.559	35997.652	3464.991
			1320.361	1647.660	1974.960	1404.535	2214.201	2949.159	1133.607
			0.931	0.911	0.871	0.924	0.863	0.778	0.964
9	73	9.534	5.101	9.096	13.090	16.301	-5.808	-32.452	13.384
		994.157	11200.465	17280.199	26525.395	12537.252	33667.156	68639.750	7945.414
			3149.039	3839.250	4529.449	3297.237	5139.469	6648.559	2733.782
			0.944	0.926	0.882	0.934	0.888	0.805	0.973
10	72	12.222	12.761	18.111	23.461	24.250	5.431	-17.319	20.146
		600.439	8144.629	11497.004	16563.008	7637.500	21831.547	57975.520	4758.062
			2102.315	2399.214	2696.115	1964.993	3080.103	4985.070	1627.409
			0.951	0.913	0.855	0.918	0.851	0.845	0.963
11	72	12.319	14.608	20.333	26.058	25.657	9.347	-11.167	21.181
		423.649	6647.418	10619.918	16631.367	6963.945	21452.203	51249.844	3979.647
			1584.201	1868.326	2152.452	1515.659	2525.006	3972.824	1238.841
			0.944	0.881	0.811	0.882	0.838	0.853	0.954
12	71	8.056	-1.713	1.127	3.966	10.197	-17.155	-51.197	8.500
		473.771	7245.789	8257.785	10541.934	5077.199	17971.176	62251.117	3623.750
			1705.542	1790.485	1875.427	1416.203	2282.610	3957.781	1234.743
			0.921	0.905	0.839	0.913	0.782	0.729	0.942
13	68	11.603	14.760	20.809	26.857	25.358	12.250	-6.941	20.798
		358.278	5382.430	7196.457	10412.926	4581.793	14028.750	39551.641	2921.024
			1299.421	1398.831	1498.243	1112.921	1929.743	2882.580	987.035
			0.936	0.871	0.776	0.869	0.861	0.766	0.965

Pour chacun des 13 tests (en ligne) : nombre de sujets, moyenne et SCE du score simple (col.1) et des scores avec certitude calculés à partir des six barèmes A, B, C, D, E et F (voir tableaux 3.18 et 3.21) et du barème G utilisé dans l'expérience 1972-1973. La somme des produits des écarts et la corrélation avec le score simple est chaque fois indiquée.

CONCLUSIONS

1. Les conclusions de l'expérience de 1971-72 concernant la validité de l'ESPER et les stratégies présidant au choix des indices de certitude sont confirmées dans l'expérience de 1972-73.

2. Les TUC c (taux d'utilisation des différents degrés de certitude) sont stables tout au long des treize tests successifs (H16', H17').

Cette constatation, très importante, confirme à elle seule les hypothèses générales :

- Il est possible de recueillir des ESPER valides;
- Les conditions expérimentales (situation operante, étude longitudinale) sont appropriées;
- Une matrice calculée selon le critère E.S.U. est requise pour une étude valable de l'ESPER.

3. Les recherches sur la validité de l'ESPER doivent être poursuivies et étendues.

Il serait intéressant, en premier lieu, d'étudier l'impact d'un entraînement spécifique préalable (exercices systématiques d'ESPER, explication du principe de l'E.S.U., communication des cohérences aux sujets, utilisation de ces cohérences pour déterminer les conséquences, etc.).

Il serait nécessaire de confronter les principes et les résultats expérimentaux relatifs à l'ESPER avec la théorie de la généralisabilité. Dans cette perspective, il semble pertinent de recueillir, au préalable, un minimum d'informations sur la fidélité et la sensibilité de l'ESPER. C'est l'objet des chapitres 5 (où sont décrites des procédures de testing) et 6 (où sont analysées des données).

CHAPITRE 5

PROCEDURES EXPERIMENTALES DESTINEES A L'ETUDE DE L'ESPER

A. Le jeu des prédictions de SHANNON

B. Le jeu des prédictions de SHANNON
avec indication du degré de certitude

C. Le jeu des ESPER

D. Un dispositif automatique

E. Une procédure collective "à grande
échelle"

F. Perspectives.

A. LE JEU DES PREDICTIONS DE SHANNON

Cette méthode (1) d'estimation de la redondance d'un texte anglais exploite le fait que quiconque parle une langue naturelle possède, implicitement, une connaissance importante des statistiques de cette langue. Sa familiarité avec les mots, les expressions, les clichés et la grammaire mettent le sujet à même de remplir des lettres manquantes dans un texte écrit, ou de compléter une phrase inachevée dans la conversation. Ce phénomène est systématiquement exploité dans le test de closure (2).

SHANNON (1951) demande à un sujet de prédire successivement chaque lettre d'un texte d'anglais littéraire. Vingt-sept réponses sont disponibles : les vingt-six lettres de l'alphabet plus le blanc qui inclut non seulement les espaces, mais aussi tous les signes de ponctuation.

En prédisant la première lettre du passage, le sujet ne peut s'appuyer que sur des probabilités dites "de premier ordre" (3), mais, à mesure qu'il découvre les lettres et les espaces, il lui est plus facile de prédire. Les indices qui facilitent les prédictions successives sont, non seulement les probabilités de certaines séquences "d'ordre deux" ou "d'ordre trois", etc, mais aussi le contenu sémantique. Grâce à ce dernier type d'indice, un sujet moyen réussit mieux au jeu de Shannon qu'un ordinateur qui disposerait, en mémoire, de toutes les probabilités de séquences jusqu'à "l'ordre cinq", mais qui serait incapable de comprendre le "sens" de la phrase.

---

(1) L'appellation originale bien connue est SHANNON's Guessing Game. Par "Jeu de prédictions", nous trahissons le terme guessing, mais ce dernier est lui-même impropre, car il évoque une réponse fournie au hasard, ce qui ne décrit pas exactement la technique dont il s'agit.

(2) Voir TAYLOR (1953) et G. DE LANDSHEERE (1973).

(3) Par exemple, la probabilité (d'ordre deux) que la lettre Q, en tête de mot, soit suivie d'un U, est très élevée.

Exploration du possible,
pas du réel

La procédure utilisée par SHANNON est la suivante. Quand la prédiction d'une lettre donnée est incorrecte, le sujet doit fournir d'autres réponses jusqu'à ce que l'une d'elles soit correcte. Ainsi il est possible d'enregistrer, pour chaque lettre du passage, le nombre de réponses fournies par le sujet avant de trouver cette lettre.

ATTNEAVE (1959, p. 30) montre un résultat typique obtenu par SHANNON (tableau 5.1) :

Tableau 5.1

THERE IS NO REVERSE ON A MOTORCYCLE
1 1 1 5 11211 2 1115117 1 1121 3 21 22 7 1 1 1 1 4 1 1 1 1 1

A FRIEND OF MINE FOUND THIS OUT
318613 1 1111111 1 1116 2 1 1 111 1211 1 1 1 1

RATHER DRAMATICALLY THE OTHER DAY
4 1 1 1 1 11115 1 1 1 1 1 1 1 1 1 1 6 1 1 1 1 1 1 1 1 1 1 1 1

La première ligne de chacun des trois exemples constitue le texte original à deviner, lettre après lettre. La seconde ligne présente le nombre de réponses nécessaires pour trouver chaque lettre, espace (ou "blanc") compris.

Nous avons pratiqué ce jeu comme expérimentateur et comme sujet (voir l'exemple ci-dessous). La technique est simple et les consignes souples (1). Le tableau 5.2 présente, à la manière de SHANNON, les résultats d'une expérience réalisée en présence de témoins où nous avons reconstitué une phrase de FREUD (2). La ligne supérieure est le texte à deviner, lettre après lettre; la ligne inférieure présente le nombre de réponses nécessaires pour trouver chaque lettre.

---

(1) On peut (ou non) signaler au sujet le nom de l'auteur, le titre de l'ouvrage, etc.; ces indices auront pour effet de faciliter le jeu. En outre, il est préférable de demander au sujet de noter lui-même la série des lettres qu'il a déjà énumérées: il a ainsi sous les yeux les réponses incorrectes à ne pas répéter. De même, le sujet se constituera plusieurs exemplaires de l'alphabet afin d'avoir constamment sous les yeux les 27 réponses possibles.

(2) S. FREUD, <u>Totem et tabou</u>, Paris, Ed. Payot, 1965, p. 73.

Tableau 5.2

```
L'HYPOTHESE SELON LAQUELLE
1 5 6 1 2 2 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

LES MORTS L ES PLUS CHERS
1 1 2 1 7 4 6 1 1 1 12 3 1 1 1 1 1 1 1 8 1 1 1 1 1

SE TROUVERAIENT TRANSFORMES
9 1 2 1 1 2 1 1 2 1 1 1 1 1 1 5 2 2 1 1 2 1 1 1 1 1

EN DEMONS F AIT SURGIR T OUT
1 1 1 5 2 7 1 1 1 1 13 1 8 1 1 3 1 1 1 1 1 10 1 1 1 2

NATURELLEMENT UNE AUTRE QUESTION
8 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 9 1 1 1 1 2 1 1 1 1 1 1 1
```

Résultat d'une expérience à partir du jeu de SHANNON modifié sur un texte de FREUD (Totem et Tabou, Ed. Payot, 1965, p. 73).

Sur les 141 lettres (et espaces) de ce texte, 107 ont été trouvées dès la première réponse. Les 34 autres lettres furent découvertes comme suit (tableau 5.3).

Tableau 5.3

14 en deux réponses	
2 en trois réponses	
1 en autre réponses	
3 en cinq réponses	soit 272 réponses
2 en six réponses	pour trouver 141 symboles.
2 en sept réponses	
1 en huit réponses	Le pourcentage de réussite
4 en neuf réponses	est 51,8 %.
1 en dix réponses	
1 en onze réponses	
1 en douze réponses	

Nombre de réponses nécessaires pour prédire les lettres correctes du texte repris au tableau 5.2

L'intérêt de cette technique dépasse celui d'un jeu, ce que reconnaissent divers auteurs. Le texte qui suit ne souligne pas seulement les possibilités de la technique, mais apporte, en outre, des suggestions (que nous avons soulignées dans le texte original) propres à améliorer le jeu lui-même.

"Supposons dix solutions possibles dont l'une est correcte. Le sujet fournit une réponse incorrecte. Ordinairement, l'expérimentateur enregistre cela simplement comme une erreur et ne cherche pas à savoir quelle était l'ampleur de l'incertitude du sujet... La procédure de SHANNON, qui permet au sujet de deviner jusqu'à ce qu'il fournisse la réponse correcte, permet une mesure bien plus fine de la performance... avec pour conséquence que les données ont une plus grande stabilité et qu'ainsi moins de sujets sont nécessaires." (ATTNEAVE, 1959, p. 40).

Cette procédure peut être utilisée pour le testing de groupe, grâce à un appareil conçu par ATTNEAVE et M.D. ARNOULT (1956).

"Le sujet répond en perforant, avec un stylo à pointe cônique, une des positions indiquées sur chaque rangée de sa feuille de réponse. Cette feuille est solidement attachée à une planche dans laquelle des petits trous ont été forés." Sous les solutions fausses, les trous ne permettent pas au stylet de s'enfoncer profondément, contrairement à la bonne solution, où un large trou se produit dans la feuille de papier. Le nombre de trous perforés dans chaque rangée constitue un enregistrement du nombre de réponses nécessaires au sujet pour trouver la bonne.

Cette technique est, évidemment, à rapprocher des *punchboards* popularisées par l'enseignement programmé.

La technique du Shannon's guessing-game a été utilisée par POLLACK (1954) dans une série d'expériences sur l'"assimilation d'informations encodées séquentiellement". Par exemple, on lit à un sujet une série de huit symboles, chacun pouvant être une lettre ou un chiffre (donc 36 possibilités). Si la séquence BX3G9KRF a été lue et si le sujet répond par écrit BX3C9RKF, on lui dira que ses 4e, 6e et 7e symboles sont incorrects. Il peut alors répondre autant de fois qu'il est nécessaire pour trouver les symboles corrects.

B. LE JEU DES PREDICTIONS DE SHANNON AVEC INDICATION DU DEGRE DE CERTITUDE

Nous avons modifié le jeu de SHANNON afin de l'adapter à l'étude de l'ESPER : en plus de chaque symbole prédit, le sujet doit fournir un indice de certitude. Le nombre d'indices disponibles peut évidemment varier, depuis l'accent mis (ou non) sur la réponse jusqu'à l'usage d'une échelle de probabilité allant de 0 à 1, sans aucune restriction dans la précision (*continuous confidence marking*). Dans l'expérience qui suit, les sujets sont invités à utiliser dix niveaux de certitude comme suit :

Certitude 0	$0 \leqslant$ ESPER $<$ .1
Certitude 1	$.1 <$ ESPER $<$ .2
⋮	
Certitude 9	$.9 <$ ESPER $<$ 1.

Chaque degré de certitude représente une zone sur l'échelle PER. Nous appellerons $\mathcal{G}$ PER le point central de cette zone. Ainsi, $\mathcal{G}$ PER pour la certitude 1 est .15.

C'est, en fait, conformément à cette procédure que s'était déroulée l'expérience décrite ci-dessus (la phrase de FREUD). Le protocole du début de cette expérience (1) figure dans le tableau 5.4.

---

(1) L'espace (ou blanc typographique) est représenté par un tiret : – .

Tableau 5.4

	L(1)					
L	E(2)	A(3)	D(3)	U(4)	-(5)	
-	A(1)	E(1)	I(2)	O(4)	U(4)	H(6)
H	Y(1)					
Y	S(4)	P(4)				
P	N(6)	O(2)				
O	M(2)	T(2)				
T	H(6)					
H	E(8)					
E	S(9)					
S	E(9)					
E	-(9)					

Protocole du jeu de SHANNON avec certitude lors du début de l'expérience décrite au tableau 5.2

A partir de ces données (272 réponses au total), on peut dénombrer les réponses inexactes et les réponses exactes pour chaque niveau de certitude dans le tableau 5.5.

Tableau 5.5

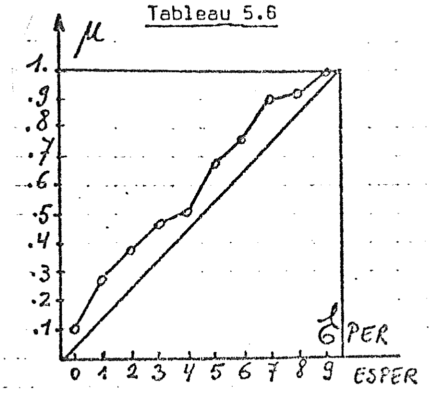
Certitude	B.R.	M.R.	TOTAL	μ	$\bar{\epsilon}$ PER
0	3	30	33	0.09	.05
1	21	52	73	.287	.15
2	16	27	43	.372	.25
3	7	8	15	.466	.35
4	5	5	10	.500	.45
5	9	4	13	.692	.55
6	7	2	9	.770	.65
7	9	1	10	.900	.75
8	11	1	12	.916	.85
9	53	1	54	.981	.95
TOTAL	141	131	272		

μ des dix degrés lors du jeu de SHANNON avec certitude (tableau 5.2)

B.R. signifie : "Bonnes réponses"
M.R. signifie : "Mauvaises réponses"
μ correspond au pourcentage de succès observé parmi toutes les réponses fournies avec un degré de certitude donné.

On peut dresser (tableau 5.6) un graphique en portant en abscisse les $\mathcal{E}$ PER et en ordonnée les μ.

Tableau 5.6



Fonction μ/ESPER ou fonction de validité de l'ESPER (les données numériques figurent au tableau 5.5)

On remarque qu'aucune interversion ne se produit entre les μ observés. Le rho de Spearman entre les deux variables vaudrait 1 et le r de Bravais-Pearson serait très élevé. Ce résultat est (par chance, vu les petits nombres de réponses) un exemple de <u>transivité stochastique</u> (1).

Un tel graphique représente la <u>fonction de probabilité subjective</u> d'une personne (COOMBS, DAWES et TVERSKY, 1970, p. 136). Nous l'appellerons f (μ| ESPER).

La corrélation entre les deux variables est élevée. Un simple examen visuel révèle que le coefficient de pente de la droite de régression et la valeur de la constante sont proches respectivement de 1 et de 0. La droite idéale est, bien entendu, $y = ax + b$, avec $a = 1$ et $b = 0$.

Des valeurs différentes <u>mais stables</u> de a et de b n'invalideraient en rien la validité de l'ESPER. Il en irait de même si l'équation était d'un autre degré, à condition qu'elle soit stable dans le temps pour un même individu, et d'un individu à l'autre, et pourvu que la fonction soit croissante.

(1) Voir F. BRESSON, 1965, p.234

L'expérience ci-dessus a été répétée dix fois (total : 3207 items). Dans le tableau 5.7, ces données sont analysées.

Tableau 5.7

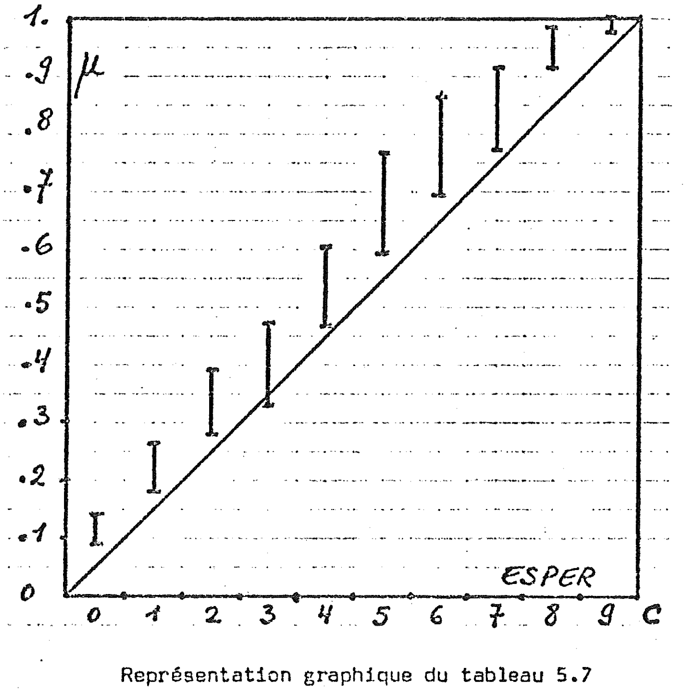
Cert.	BR	MR	Tot.	μ	$\sqrt{\frac{pq}{N}}$	SE x 1,96 (1)	Lim. de conf. de
0	101	754	855	11,8 %	1,1	$\pm$ 2,156	de 9,6 à 13,9 %
1	128	457	585	21,8 %	1,72	$\pm$ 3,37	de 18,4 à 25,1 %
2	120	226	346	34,6 %	2,5	$\pm$ 4,9	de 29,7 à 39,5 %
3	76	113	189	40,2 %	3,56	$\pm$ 6,97	de 33,23 à 47,2 %
4	117	99	216	54,1 %	3,4	$\pm$ 6,66	de 47,4 à 60,7 %
5	87	41	128	67,9 %	4,12	$\pm$ 8,07	de 59,8 à 75,9 %
6	75	22	97	77,3 %	4,25	$\pm$ 8,3	de 69 % à 85,6 %
7	93	17	110	84,5 %	3,45	$\pm$ 6,76	de 77,7 à 91,26 %
8	206	10	216	95,3 %	1,4	$\pm$ 2,75	de 92,5 à 98,05 %
9	461	4	465	99,1 %	.44	$\pm$ 0,86	de 98,2 à 99,96 %

Limites de confiance des divers μ observés pour un ensemble de 3207 items du jeu de SHANNON avec certitude présentés à un même sujet

Les trois colonnes de droite sont consacrées au calcul des limites de confiance (à P.05) des μ.

Le tableau 5.8 représente notre fonction de probabilité subjective (avec les limites de confiance à P.05).

Tableau 5.8



Représentation graphique du tableau 5.7

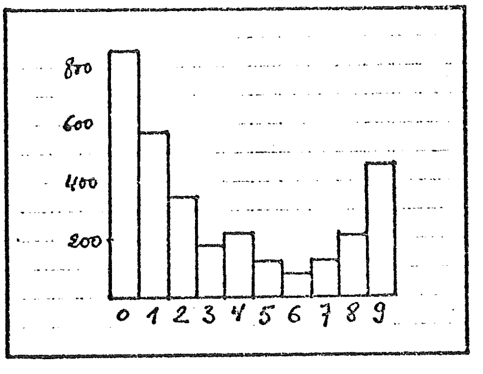
(1) En fait, il faudrait utiliser les valeurs t.

A partir du tableau 5.8, on peut faire les constatations suivantes :

1. La transitivité stochastique est respectée, ce qui indique que le sujet est cohérent avec lui-même dans ses estimations. En d'autres termes, il y a une corrélation parfaite (rho de Spearman = 1) entre les degrés de certitude et les pourcentages de réussite.
2. Tous les pourcentages sont décalés vers le haut, ce qui indique que le sujet s'est légèrement sous-estimé, puisqu'il réussit mieux que les certitudes qu'il annonce.
3. Certains pourcentages voisins sont très peu distincts (par exemple 2 et 3, 6 et 7, ainsi que 8 et 9).

On remarque aussi que le τ ou TUC (taux d'utilisation des indices de certitude) varie très fort d'un degré à l'autre, ce qui explique que les intervalles de confiance sont étroits pour les certitudes 0, 1 et 9 et nettement plus larges pour les certitudes allant de 3 à 7. Voici (tableau 5.9) les réponses pour chaque degré de certitude, nombres à partir desquels ont été calculés les dix pourcentages de réussite :

Tableau 5.9



Taux d'utilisation (τ) de chaque degré de certitude pour les 3207 items de l'expérience synthétisée dans le tableau 5.7

Ces différences dans l'utilisation des degrés de certitude s'expliquent de deux façons.

Première explication.

La matière elle-même impose l'utilisation très fréquente de la certitude 0 (quand presque toutes les lettres de l'alphabet sont possibles), de la certitude 1 (quand seules les voyelles sont possibles), des certitudes 8 et 9 (quand les possibilités sont fort restreintes). Trois degrés (0, 1 et 9) constituent à eux seuls 60 % des indices fournis (voir tableaux 5.7 et 5.9). Pour cette raison, dans les expériences à grande échelle, les lettres à deviner ne seront plus choisies au hasard.

Deuxième explication.

Le degré 0 pourrait être nuancé en deux ou trois degrés différents, car le sujet serait à même de faire la distinction entre une chance sur vingt et une chance sur dix. La sensibilité paraît plus grande aux extrêmes qu'au milieu. Pour cette raison, certains chercheurs (W. EDWARDS par exemple) utilisent plutôt des échelles d'allure logarithmique, du type :

0	1	2	3	4	5	6	7	8	9	10
$\frac{1}{10.000}$	$\frac{1}{1000}$	$\frac{1}{100}$	$\frac{1}{10}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{9}{10}$	$\frac{99}{100}$	$\frac{999}{1000}$	$\frac{9999}{10000}$

Ce sont alors des rapports et non plus directement des probabilités qui sont appliqués au problème de l'ESPER.

Les données ci-dessus sont peu généralisables puisqu'elles proviennent d'expériences où le sujet est le chercheur lui-même et que ce dernier a pu, consciemment ou non, influencer les résultats. Il est donc indispensable de procéder à de nouvelles expériences sur d'autres sujets, en suivant une procédure collective permettant de récolter un plus grand nombre de données. C'est ce type de procédure qui va être exposé.

C. LE JEU DES ESPER

Ce jeu peut être collectif. Plusieurs sujets devinent en même temps la même lettre du texte (avec indication d'un indice de certitude); quand tous ont inscrit leur réponse, l'expérimentateur révèle la réponse correcte. La même procédure se reproduit pour chaque lettre. La différence majeure entre ce jeu et le précédent est que le sujet ne peut émettre qu'une seule prédiction par question. Comme dans les autres jeux, les sujets sont invités à maximiser leur score total. Ici toutefois, à chaque réponse, des points peuvent être gagnés ou perdus selon une matrice des utilités définie sur deux événements possibles (E1 = réponse correcte; E2 = réponse incorrecte) et sur les dix actions disponibles (les dix degrés de certitude). L'omission est exclue. Dans ces expériences, la matrice (calculée selon les principes décrits au chapitre 3) était la suivante (tableau 5.10).

Tableau 5.10

	Bonne réponse	Mauvaise réponse	Idem (arrondis)
Certitude 1	21	0	0
2	22	- 0,11	- 0,1
3	23	- 0,36	- 0,33
4	24	- 0,79	- 0,8
5	25	- 1,46	- 1,5
6	26	- 2,46	- 2,5
7	27	- 3,96	- 4
8	28	- 6,29	- 6
9	29	-10,29	-10
10	30	-19,29	-20

Données de base du jeu de prédiction avec certitude
(300 items posés à 10 étudiants)

Ce jeu a été proposé à dix étudiants de vingt ans, lors de trois séances espacées d'une semaine. A chaque fois, cent lettres ont été devinées, ce qui représente environ une heure de prestation. Les réponses et degrés de certitude sont notées sur une feuille de réponses par chaque sujet.

Des techniques peuvent être nécessaires pour contrôler l'honnêteté des sujets : utilisation de dispositifs de réponses , affichage des réponses par les sujets (sur ardoises par exemple).

La phrase correcte est écrite par l'expérimentateur au tableau, lettre après lettre, au rythme du déroulement de l'expérience. Après la séance, chaque sujet établit sa matrice des résultats.

Voici le résumé des données (tableau 5.11) pour les cinq meilleurs sujets (ceux dont le pourcentage de réponses correctes sont les plus élevés) et les cinq moins bons.

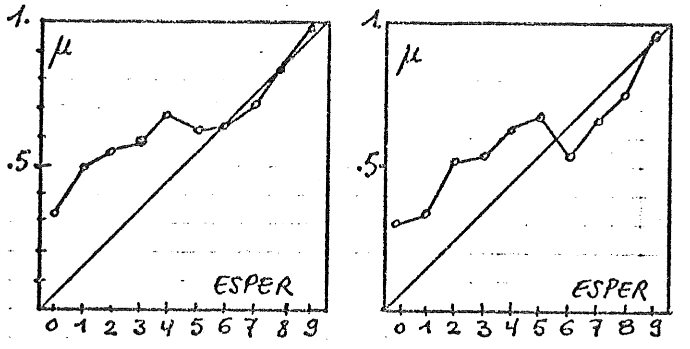
Tableau 5.11

Les cinq meilleurs					Les cinq moins bons					
Cert.	BR	MR	TOT.	μ	Cert.	BR	MR	TOT.	μ	$\mathcal{E}$ PER
0	117	256	373	.313	0	114	259	373	.305	.05
1	37	37	74	.500	1	21	40	61	.344	.15
2	44	35	79	.556	2	58	51	109	.532	.25
3	15	11	26	.576	3	12	10	22	.545	.35
4	85	40	125	.680	4	80	49	129	.620	.45
5	12	7	19	.631	5	23	11	34	.676	.55
6	24	14	38	.631	6	19	17	36	.527	.65
7	17	7	24	.708	7	24	11	35	.685	.75
8	35	6	41	.853	8	38	12	50	.760	.85
9	682	19	701	.972	9	625	26	651	.960	.95
	1068	432	1500	.712		1014	486	1500	.676	

Résultats du jeu de prédiction avec certitude
(300 items posés à 10 étudiants)

Les deux fonctions de probabilité subjective $f(\mu/\text{ESPER})$ sont les suivantes (tableau 5.12).

Tableau 5.12



Fonction μ/ESPER pour les cinq meilleurs (à gauche) et pour les cinq moins bons (à droite) sujets de l'expérience du jeu des prédictions avec certitude (tableau 5.11)

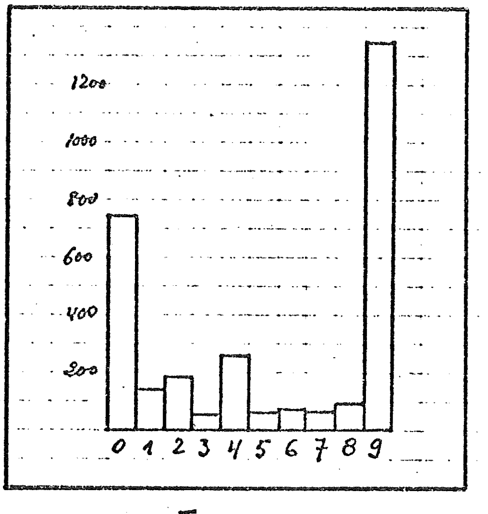
On peut calculer un indice global d'écart à la diagonale, par la formule bien connue $\sqrt{\frac{\sum d^2}{n}}$

où d = écart entre le pourcentage de réussite observé ($\mathcal{M}$(ER | ESPER)) le pourcentage de réussite attendu ($\mathcal{E}$(ESPER | C))

n = le nombre de réponses

Tableau 5.13

Le taux d'utilisation de chaque degré de certitude figure au tableau 5.13 :



Taux d'utilisation (τ) de chaque degré de certitude pour les 3000 items de l'expérience synthétisée dans le tableau 5.11

De nouveau, l'utilisation massive des degrés 0 et 9 diminue la signification des calculs relatifs aux autres degrés de certitude. Pour cette raison, dans les expériences ultérieures, nous révélerons aux sujets certaines lettres de la phrase. Il s'agit des lettres pour lesquelles la grande majorité des sujets fournirait une prédiction avec la certitude minimale et celles pour lesquelles, à l'opposé, les prédictions seraient affectées d'une certitude maximale par tous les sujets.

D. UN DISPOSITIF AUTOMATIQUE

La procédure collective qui vient d'être décrite a l'avantage de ne mobiliser qu'un expérimentateur pour une dizaine de sujets. Par contre, elle ne permet pas à ces derniers de travailler selon leur rythme individuel. De plus, elle mobilise tellement l'expérimentateur que ce dernier ne peut effectuer certaines observations (chronométrie par exemple) ou certaines tâches (affichage immédiat des scores individuels).

Pour cette raison, nous avons mis au point une technique qui consiste à écrire sur un long ruban de papier, de cinq en cinq centimètres, une même phrase, de plus en plus complète (ici à une seule lettre près : M). La phrase 94 (exemple fourni dans le tableau 5.14) offre donc la solution à la question 93, tout en sollicitant une nouvelle réponse (1).

Tableau 5.14

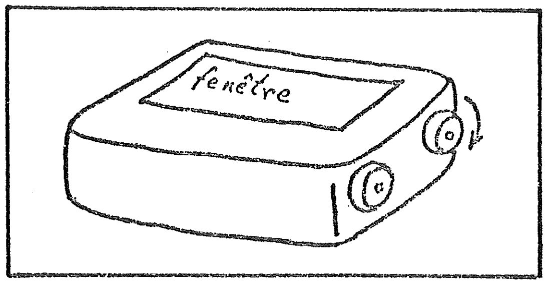
```
93
SI-L'ON-VEUT-S'AVENTURER-DANS-UN-HUMOUR-UN-PEU-MAC
ABRE,ON-IMAGINERA-QU'UN-HOMME,VENANT-DE-PERDRE-SON

-JUMEAU,FASSE-

-----A----/----B----/----C----/----D----
 0   1   2   3   4   5   6   7   8   9
|...|...|...|...|...|...|...|...|...|...|
0  10  20  30  40  50  60  70  80  90  100


94
SI-L'ON-VEUT-S'AVENTURER-DANS-UN-HUMOUR-UN-PEU-MAC
ABRE,ON-IMAGINERA-QU'UN-HOMME,VENANT-DE-PERDRE-SON

-JUMEAU,FASSE-M

-----A----/----B----/----C----/----D----
 0   1   2   3   4   5   6   7   8   9
|...|...|...|...|...|...|...|...|...|...|
0  10  20  30  40  50  60  70  80  90  100
```

Extrait du ruban présentant une phrase (de plus en plus complète) de J. ROSTAND : Aux frontières du surhumain. Trois échelles (4, 10 et 40 degrés) sont chaque fois proposés aux étudiants

(1) L'exemple du tableau 5.14 est extrait de J. ROSTAND, Aux frontières du surhumain. Le choix des textes dépend des individus qui devront les deviner. Les expériences qui vont être décrites ont été réalisées avec des adultes universitaires. Il importe de choisir un texte qui ne soit pas trop technique par le vocabulaire utilisé et dont le thème ne suppose pas de connaissance ou de prérequis spécifiques. Les séquences où trop de sigles, dates, noms propres apparaissent sont aussi écartées. La mutilation du texte porte sur cent caractères, les cent précédents étant fournis aux sujets afin qu'ils puissent disposer d'un "contexte" donnant un sens à leurs recherches.

Ce long ruban, qui commence par une seule lettre et qui se termine par la phrase complète, est introduit dans un dérouleur très simple (tableau 5.15).

Tableau 5.15



"Boîte enseignante" (Ecole Moderne - Freinet) transformée en dispositif de présentation des rubans décrits au tableau 5.14

La fenêtre permet de voir une seule phrase à la fois. Le sujet a la confirmation (ou l'infirmation) immédiate de sa prédiction; il ne peut la modifier parce qu'elle est passée sous le plexiglas.

Le ruban peut être réalisé par le programme FORTRAN suivant (tableau 5.16).

Tableau 5.16

```
C-------GENERATION DE PHRASES POUR LE JEU DES CERTITUDES
        DIMENSION K(200),A(200),R(50,4)
        READ(5,502)((R(I,J),I=1,50),J=1,4)
    502 FORMAT(50A1)
      7 READ(5,500,END=3)(K(I),I=1,200)
        READ(5,501)(A(I),I=1,200)
    500 FORMAT(80I1)
    501 FORMAT(80A1)
        L=0
        N=0
      5 N=N+1
      2 M=N+1
        IF(L.GT.100)GOTO 7
        IF(K(M).NE.1)GOTO 4
        N=N+1
        GOTO 2
      4 L=L+1
        WRITE(6,600)L
        WRITE(6,601)(A(I),I=1,N)
    600 FORMAT('0'////'0',I4)
    601 FORMAT('0',50A1/' ',50A1)
    602 FORMAT('0',50A1/' ',50A1/' ',50A1/' ',50A1)
        WRITE(6,602)((R(I,J),I=1,50),J=1,4)
        GOTO 5
      3 STOP
        END
```

Programme FORTRAN d'écriture des phrases présentées dans le tableau 5.14

Les données se présentent comme suit (tableau 5.17).

```
         1         2         3         4         5         6         7
1234567890123456789012345678901234567890123456789012345678901234567890123456789
SI-L'ON-VEUT-S'AVENTURER-DANS-UN-HUMOUR-UN-PEU-MACABRE,ON-IMAGINERA-QU'UN-HOMME
VENANT-DE-PERDRE-SON-JUMEAU,FASSE-METTRE-EN-CONSERVE-UN-JEU-D'ORGANES-FRATERNEL
-QUI-POURRAIENT,EVENTUELLEMENT,LUI-SERVIR-D'ORGANES-DE-SECOURS.
1       1   11       1111   11   1   111       1   1111   1   111 1  1       11
     11       111   1    111    11    111             1           1111     1111
1   1      111111   11111111111        111      1111       1111
```

Tableau 5.17

E. UNE PROCEDURE COLLECTIVE "A GRANDE ECHELLE"

Cette procédure a été mise au point dans le but de mettre à l'épreuve à la fois les hypothèses HF et HS relatives à la fidélité (stabilité) et à la sensibilité de l'ESPER (voir chapitre 6). Le jeu des certitudes, portant sur cent questions relatives à un texte de J. FOURASTIE (1), a été présenté à des adultes. Cet extrait a été mutilé selon les principes découlant des expériences précédentes. On trouvera, en annexe 5, le texte complet, sa version mutilée et le programme FORTRAN d'impression. Trois échelles de certitude (une à quatre degrés, une à dix degrés et une à quarante degrés) ont été utilisées, afin d'estimer la sensibilité maximale de l'ESPER.

L'expérience s'est déroulée en trois phases :

1) Les sujets répondent à une série "d'essai" de cinq questions, qui a pour but de familiariser avec le problème, le mode de réponse, la procédure de correction et les résultats.
2) Quinze jours plus tard, les sujets doivent répondre (par une lettre et l'indication du degré de certitude) à chacune des cent questions constituant l'expérience proprement dite.
3) Les sujets doivent fournir à nouveau, après un intervalle de temps (certains deux semaines, d'autres cinq semaines), les cent certitudes aux cent lettres qu'ils avaient déjà fournies (et qu'ils ne peuvent plus changer).

Le chapitre 6 est entièrement réservé à cette troisième phase : l'expérience test-retest. Dans les pages qui suivent, on décrit les documents utilisés lors des phases 1 et 2.

---

(1) J. FOURASTIE, Les quarante mille heures, Ed. Denoël, 1972, p. 155 sq.

La série d'essai.

Nous avons invité trois cents adultes (en majorité des professeurs de l'enseignement secondaire) à participer à une expérience à partir du "jeu des ESPER".

Bien que ces personnes aient été contactées directement lors de réunions, une lettre d'invitation leur a été remise. Ce document comprenait :

1° Une description brève de la recherche;

2° Des indications sur le déroulement ultérieur de l'expérience et la manière dont les informations circuleront;

3° Des documents propres à la série d'essai:
- Les consignes (tableau 5.18),
- Un exemple de démonstration (tableau 5.19),
- L'exercice d'essai proprement dit servant également de feuille de réponses (tableau 5.20).

4° Des documents relatifs au jeu des ESPER, en général :
- Les huit règles du jeu (tableau 5.21),
- Trois matrices de gains et de pertes (tableau 5.22),
- La représentation graphique des gains et des pertes (tableau 5.23).

Tableau 5.18

JEU DES CERTITUDES

Série d'essai

Tout d'abord, veuillez lire les règles du jeu (en annexe).
La présente série ne comporte que cinq questions. En moyenne, il faut entre 30 et 45 secondes pour répondre à une question. Ne consacrez donc pas plus de cinq minutes pour l'ensemble de vos réponses.

Vous faciliteriez la correction en utilisant un bic ou un marqueur de couleur vive pour entourer les points comme dans l'exemple de réponses (photocopié) joint en annexe. Vous constatez qu'il s'agit du même texte que le vôtre, mais avec une lettre en moins.
Vous connaissez donc les cinq réponses correctes et vous pouvez constater que le sujet qui a fait les cinq prédictions :

- a proposé des lettres plausibles;
- s'est trompé trois fois sur cinq, c-a-d :
    - question 1 et 5 : avec une certitutde faible.
    - question 3 : avec (hélas) une certitude forte.

Il a obtenu les points suivants :

	4 degrés	10 degrés	40 degrés
Question 1	0	0	- 0,12
Question 2	31,25	27,50	29,75
Question 3	-10,00	-11,87	-11,31
Question 4	46,25	45,50	44,75
Question 5	0	- 0,33	- 0,60
Total	67,50	60,80	62,47

Les trois totaux (différents) ne sont pas ajoutés les uns aux autres.
Il faut essayer de maximiser chacun des trois résultats.

Rappel :

Les certitudes 0, 1, 2 et 3 doivent être fournies comme suit :
Si vous estimez
vos chances de réussite entre

...0 et 25% ------ certitude 0
25 et 50% ------ certitude 1
50 et 75% ------ certitude 2
75 et 100% ----- certitude 3

Les certitudes de 0 à 9 sont elles aussi liées à des zones de pourcentages de chances de fournir la réponse correcte. Ces zones figurent sur la feuille de réponse(le listing).

Tableau 5.19

Exemple de démonstration :

```
1
POURQUOI L'ATTITUDE AFFECTIVE A L'EGARD DU SOUVERAIN COMPORTE-T-ELLE UN ELEMENT     -----0----/----1----/----2----/----3-----
SI P L                                                                                0   1   2   3   4   5   6   7   8   9
                                                                                      |...|...|...|...|...|...|...|...|...|...|
                                                                                      0  10  20  30  40  50  60  70  80  90  100

2
OLUTION DEPASSERAIT LE CADRE DE CE TRAVAIL.NOUS AVONS DEJA FAIT ALLUSION AU COMP    -----0----/----1----/----2----/----3-----
LEXE P A                                                                              0   1   2   3   4   5   6   7   8   9
                                                                                      |...|...|...|...|...|...|...|...|...|...|
                                                                                      0  10  20  30  40  50  60  70  80  90  100

3
DE LA ROYAUTE SERAIT DE NATURE A NOUS APPORTER UNE REPONSE DECISIVE A CETTE QUES     -----0----/----1----/----2----/----3-----
TION.D A                                                                              0   1   2   3   4   5   6   7   8   9
                                                                                      |...|...|...|...|...|...|...|...|...|...|
                                                                                      0  10  20  30  40  50  60  70  80  90  100

4
PROBANTES DE FRAZER,LES PREMIERS ROIS ETAIENT DES ETRANGERS QUI,APRES UNE BREVE     -----0----/----1----/----2----/----3-----
P E                                                                                   0   1   2   3   4   5   6   7   8   9
                                                                                      |...|...|...|...|...|...|...|...|...|...|
                                                                                      0  10  20  30  40  50  60  70  80  90  100
```

Tableau 5.20

```
1
POURQUOI L'ATTITUDE AFFECTIVE A L'EGARD DU SOUVERAIN COMPORTE-T-ELLE UN ELEMENT     -----0----/----1----/----2----/----3-----
SI PU                                                                                 0   1   2   3   4   5   6   7   8   9
                                                                                      |...|...|...|...|...|...|...|...|...|...|
                                                                                      0  10  20  30  40  50  60  70  80  90  100

2
OLUTION DEPASSERAIT LE CADRE DE CE TRAVAIL.NOUS AVONS DEJA FAIT ALLUSION AU COMP    -----0----/----1----/----2----/----3-----
LEXE PA                                                                               0   1   2   3   4   5   6   7   8   9
                                                                                      |...|...|...|...|...|...|...|...|...|...|
                                                                                      0  10  20  30  40  50  60  70  80  90  100

3
DE LA ROYAUTE SERAIT DE NATURE A NOUS APPORTER UNE REPONSE DECISIVE A CETTE QUES     -----0----/----1----/----2----/----3-----
TION.D'                                                                               0   1   2   3   4   5   6   7   8   9
                                                                                      |...|...|...|...|...|...|...|...|...|...|
                                                                                      0  10  20  30  40  50  60  70  80  90  100

4
PROBANTES DE FRAZER,LES PREMIERS ROIS ETAIENT DES ETRANGERS QUI,APRES UNE BREVE     -----0----/----1----/----2----/----3-----
PE                                                                                    0   1   2   3   4   5   6   7   8   9
                                                                                      |...|...|...|...|...|...|...|...|...|...|
                                                                                      0  10  20  30  40  50  60  70  80  90  100

5
NTS,AVEC ACCOMPAGNEMENT DE FETES SOLENNELLES.ON RETROUVE ENCORE L'ECHO DE CETTE     -----0----/----1----/----2----/----3-----
HISTOIRE PR                                                                           0   1   2   3   4   5   6   7   8   9
                                                                                      |...|...|...|...|...|...|...|...|...|...|
                                                                                      0  10  20  30  40  50  60  70  80  90  100
```

Tableau 5.21

REGLES DU JEU DES CERTITUDES

Les phrases que vous avez sous les yeux sont extraites d'un texte dont nous n'avons imprimé qu'une ligne sur deux. Ce texte n'a pas été tronqué selon les usages habituels, mais au bout de 80 caractères.

Pour chaque ligne, le jeu consiste à deviner la lettre qui suit immédiatement la dernière lettre imprimée. Quand le dernier caractère imprimé est un espace, il est représenté sur le listing par - (un tiret).

Règle 1 : La réponse ne peut être qu'une des 26 lettres de l'alphabet ou un tiret (qui remplace tous les signes de ponctuation, ou l'espace entre deux mots).

Règle 2 : LE MATIN,IL DIT:J'AI SOIF.
revient à LE-MATIN-IL-DIT-J-AI-SOIF-
Pour prédire le point, l'apostrophe, l'espace, la virgule, etc. , on répond par - (tiret).

Règle 3 : Parmi les mots à deviner, ne figurent aucun nom propre, sigle, abbréviation, mot étranger, ancien, argotique, ou hautement technique. Entre deux mots ne figurent qu'un espace ou un signe de ponctuation, jamais les deux.

Règle 4 : Vous devez associer à la lettre prédite un indice de certitude qui exprime le pourcentage de chances (la probabilité) d'être correcte que vous attribuez à votre réponse.

Vous avez une certaine connaissance des statistiques de votre langue, le français. Ainsi, vous savez qu'en début de mot un Q est toujours suivi d'un U, un L n'est jamais suivi d'un R, etc.

Mais le problème n'est pas seulement d'ordre statistique : le contexte amont et aval rend certains mots plus probables que d'autres. En raison de cet aspect sémantique, il est impossible de calculer objectivement les chances d'apparition d'une lettre donnée dans un contexte précis (à la fin de chaque ligne du jeu, par exemple).

Il faut recourir à des experts qui fournissent :

- la lettre qui leur semble la plus probable ;
- le pourcentage de chances d'être correcte qu'ils attribuent à leur réponse (cette estimation ,nous l'appelons probabilité subjective).

VOUS ETES L'UN DE CES EXPERTS.

Règle 5 : Le meilleur expert est celui qui obtient le plus grand nombre de points au jeu. On gagne des points en cas de réponse correcte et on perd des points en cas de réponse incorrecte. Une certitude élevée peut vous faire gagner beaucoup de points(en cas de prédiction correcte), ou vous en faire perdre beaucoup (en cas de prédiction incorrecte).
Les points à gagner et à perdre sont précisés dans un tableau en annexe.

./..

./..

Règle 6 : Il est inévitable que vous vous trompiez (sinon vous êtes un voyant extralucide) mais ce n'est pas grave si votre certitude est faible. Par contre, quand votre prédiction est correcte, il importe que ce soit avec une certitude forte. Il faut donc trouver un juste milieu.

IMPORTANT

Les points à gagner et à perdre ont été calculés de telle manière que vous ayez mathématiquement intérêt à fournir un indice de certitude le plus proche possible du pourcentage de chances réel de votre réponse.

Règle 7 : Comme nous ignorons le degré de finesse que des adultes peuvent atteindre dans l'expression des probabilités subjectives, nous utilisons trois échelles différentes :

La première, assez grossière, a 4 degrés : 0, 1, 2 et 3.

La seconde a 10 degrés : 0 1 2 3 4 5 6 7 8 9.

La troisième, très fine, a 40 degrés (de 2,5% en 2,5%) :

I . . . I . . . I . . . I . . . I . . . I . . . I . . . I . . . I . . . I . . . I
0 10 20 30 40 50 60 70 80 90 100

Sur cette dernière échelle, portez votre degré de certitude en entourant, non pas les nombres (comme sur les échelles à 4 et à 10 degrés), mais les points (ou les barres)de l'échelle elle-même.

Règle 8 : N'utilisez pas le dictionnaire et répondez assez vite (entre 30 et 45 secondes en moyenne par question).

Seule la lettre qui suit immédiatement la dernière lettre écrite nous intéresse. Néanmoins, vous pouvez indiquer, en entier, le mot auquel vous pensez : le listing vous sera rendu quand nous en aurons recopié les lettres et les certitudes.

Tableau 5.22

JEU DES CERTITUDES

Tableau des Gains et des Pertes

p	En cas de B.R.	En cas de M.R.	Cert.	En cas de B.R.	En cas de M.R.	Cert.	En cas de B.R.	En cas de M.R.
2,5%	20,75	0,0						
5 %	21,50	-0,02	0	21,50	0,0			
7,5%	22,25	-0,06						
10 %	23,00	-0,12						
12,5%	23,75	-0,20				0	23,75	0,0
15 %	24,50	-0,31	1	24,50	-0,33			
17,5%	25,25	-0,44						
20 %	26,00	-0,60						
22,5%	26,75	-0,79						
25 %	27,50	-1,01	2	27,50	-1,08			
27,5%	28,25	-1,26						
30 %	29,00	-1,54						
32,5%	29,75	-1,86						
35 %	30,50	-2,22	3	30,50	-2,37			
37,5%	31,25	-2,63				1	31,25	-2,5
40 %	32,00	-3,08						
42,5%	32,75	-3,58						
45 %	33,50	-4,13	4	33,50	-4,37			
47,5%	34,25	-4,75						
50 %	35,00	-5,42						
52,5%	35,75	-6,17						
55 %	36,50	-7,00	5	36,50	-7,37			
57,5%	37,25	-7,92						
60 %	38,00	-8,93						
62,5%	38,75	-10,06				2	38,75	-10,00
65 %	39,50	-11,31	6	39,50	-11,87			
67,5%	40,25	-12,70						
70 %	41,00	-14,26						
72,5%	41,75	-16,01						
75 %	42,50	-17,99	7	42,50	-18,87			
77,5%	43,25	-20,24						
80 %	44,00	-22,82						
82,5%	44,75	-25,82						
85 %	45,50	-29,36	8	45,50	-30,87			
87,5%	46,25	-33,61				3	46,25	-32,50
90 %	47,00	-38,86						
92,5%	47,75	-45,61						
95 %	48,50	-54,86	9	48,50	-57,87			
97,5%	49,25	-69,11						
100 %	50,00	-98,36						

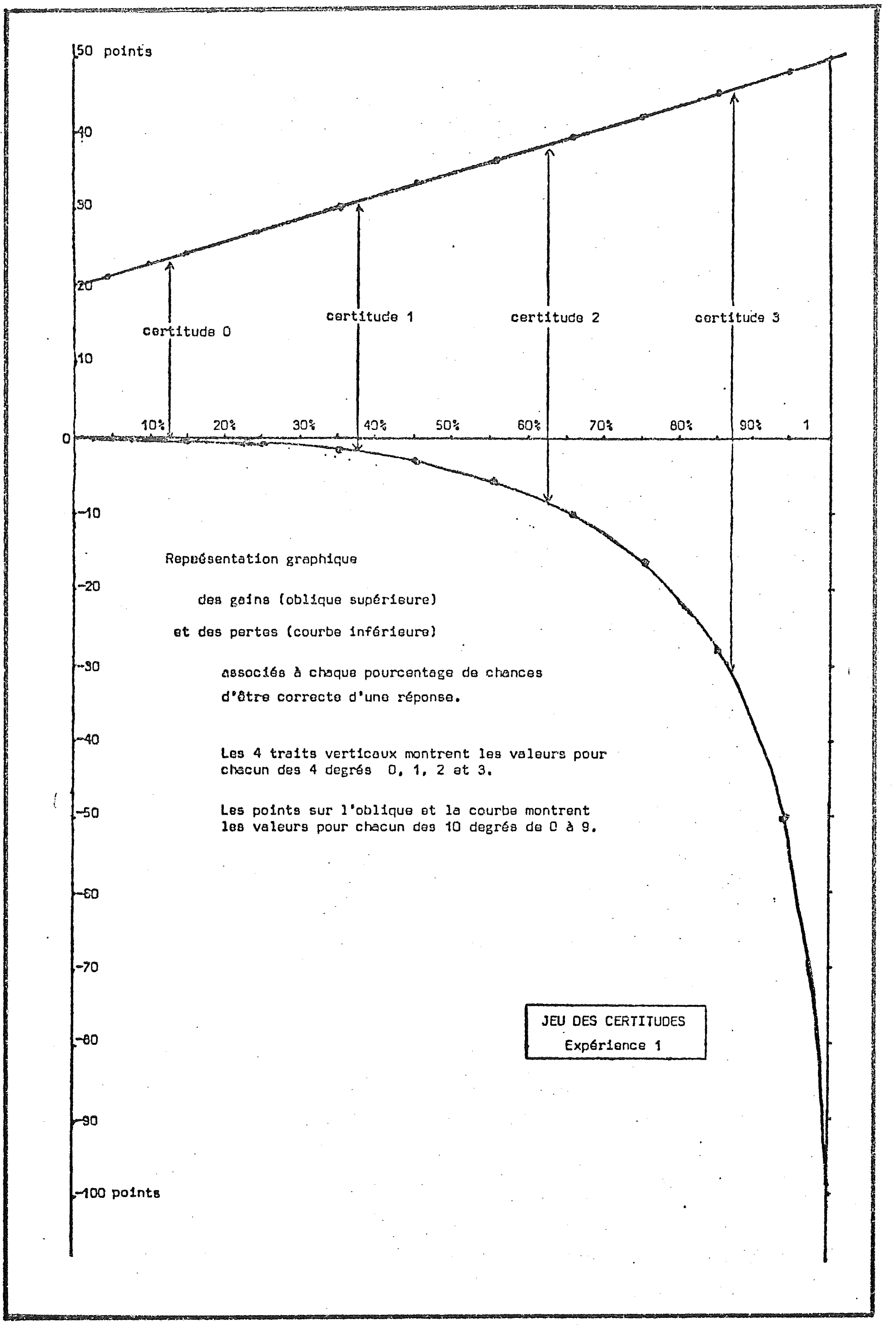
Légende : p = pourcentage de chances d'être correcte que vous attribuez à votre réponse.

B.R. = bonne réponse.

M.R. = mauvaise réponse.

Cert. = indice de certitude.

Tableau 5.23


50 points
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100 points
10%
20%
30%
40%
50%
60%
70%
80%
90%
1
certitude 0
certitude 1
certitude 2
certitude 3
Représentation graphique
des gains (oblique supérieure)
et des pertes (courbe inférieure)
associés à chaque pourcentage de chances
d'être correcte d'une réponse.
Les 4 traits verticaux montrent les valeurs pour
chacun des 4 degrés 0, 1, 2 et 3.
Les points sur l'oblique et la courbe montrent
les valeurs pour chacun des 10 degrés de 0 à 9.
JEU DES CERTITUDES
Expérience 1

L'exercice d'essai (tableau 5.20) porte sur le même texte que l'exemple de démonstration (tableau 5.19), mais avec une lettre en plus.

Pour étudier la sensibilité de l'ESPER, trois échelles (emboîtées) ont été utilisées :

- Une échelle à quatre degrés;
- Une échelle à dix degrés;
- Une échelle à quarante degrés.

Par retour du courrier (avec un délai d'une semaine), chaque sujet a reçu les réponses correctes (le texte complet) et le calcul de ses points pour chaque échelle. Voici un exemple de listing individuel (tableau 5.24).

Tableau 5.24

VOTRE CODE = 23

QUESTION NUMERO	BONNE REPONSE	VOTRE REPONSE	VOTRE CERTITUDE 4D	10D	40D	VOS POINTS 4D	10D	40D
1	I	E	1	3	40.0	-2.50	-2.37	-3.08
2	T	S	0	1	15.0	0.0	-0.33	-0.31
3	A	A	2	6	65.0	38.75	39.50	39.50
4	R	R	2	6	70.0	38.75	39.50	41.00
5	I	E	0	1	10.0	0.0	-0.33	-0.12
					TOTAUX	75.00	75.97	76.99

Résultats du sujet 23 à la série d'essai (5 items) tels qu'ils lui sont parvenus

En outre, chacun a reçu les listes ordonnées (du total le plus élevé au moins élevé) des sujets pour chacune des échelles.

Le tableau 5.25 présente la lettre d'accompagnement de cette présentation des résultats de la série d'essai (tableau 5.26).

Tableau 5.25

Madame,
Monsieur,

Vous avez accepté de vous prêter au jeu des certitudes en me renvoyant vos réponses ; je vous en remercie vivement.

A mon tour, je vous adresse divers documents.

Documents relatifs à l'expérience 1 (passée) :

A. La correction de vos réponses et le détail de vos points sur listing. Si vos calculs ne concordaient pas avec ceux de l'ordinateur, veuillez me le signaler.

B. Sur listing aussi, les résultats (ordonnés du meilleur au moins bon) des 258 personnes qui ont joué le jeu. Nous vous laissons le soin de vous situer (grâce à votre code) dans chacun des trois classements.

C. Sur une feuille stencilée, les phrases complètes et des commentaires sur le listing des 258 résultats.

Documents relatifs à l'expérience 2 (actuelle) :

D. Un exemplaire des règles du jeu (forme remaniée, mais fond inchangé; relisez cependant les règles 6, 7 et 8, s'il vous plaît) et des points à gagner et à perdre (inchangés).

E. Une enveloppe de retour.

F. Un questionnaire pour y noter vos remarques et/ou suggestions.

G. Une nouvelle feuille de jeu qui contient 100 réponses à me renvoyer dès que possible....

... avant le vendredi 29 mars 1974

H. Des consignes propres à votre groupe expérimental, car les 258 personnes vont participer à des expériences légèrement différentes.

Je vous rappelle qu'à tout moment, vous pouvez cesser de jouer; il vous suffit de ne pas renvoyer votre réponse à la date indiquée. Bien entendu, il me serait agréable de continuer à vous compter parmi les participants à la présente expérience. De mon côté, je m'engage à vous envoyer, comme pour l'expérience 1 le compte rendu détaillé des résultats.

Conscient des efforts qu'exige ce jeu, je vous remercie d'avance pour votre collaboration; elle me sera très précieuse.

Je vous prie de croire, Madame, Monsieur, à mes sentiments les meilleurs.

D. Leclercq

RESULTATS DES 258 PERSONNES DANS L'EXPERIENCE 1

AVEC 4 DEGRES *

	CODE	NBR	POINTS *
1	302	5	193.75 *
2	117	4	175.00 *
3	59	4	175.00 *
4	200	5	171.25 *
5	301	5	171.25 *
6	110	5	171.25 *
7	36	4	167.50 *
8	66	5	163.75 *
9	227	4	160.00 *
10	275	4	160.00 *
11	25	4	160.00 *
12	34	4	160.00 *
13	163	4	155.00 *
14	212	4	155.00 *
15	264	4	152.50 *
16	44	5	148.75 *
17	33	4	147.50 *
18	255	4	145.00 *
19	125	4	145.00 *
20	270	4	145.00 *
21	131	4	145.00 *
22	226	4	145.00 *
23	55	4	145.00 *
24	304	4	145.00 *
25	312	4	145.00 *
26	303	4	140.00 *
27	256	3	138.75 *
28	69	4	137.50 *
29	222	4	137.50 *
30	235	3	133.75 *
31	133	4	132.50 *
32	64	4	130.00 *
33	221	4	130.00 *
34	89	4	130.00 *
35	31	4	130.00 *
36	71	3	128.75 *
37	[illegible]	3	128.75 *
38	185	3	126.25 *
39	237	3	126.25 *
40	189	3	126.25 *
41	174	4	125.00 *
42	107	3	123.75 *
43	242	3	123.75 *
44	29	4	122.50 *
45	272	4	122.50 *
46	3	4	122.50 *
47	228	4	122.50 *
48	12	3	121.25 *
49	115	3	121.25 *
50	195	3	118.75 *
51	8	3	118.75 *
52	83	3	118.75 *
53	7	3	118.75 *
54	193	3	118.75 *
55	241	3	118.75 *
56	219	4	117.50 *
57	266	4	117.50 *
58	22	3	116.25 *
59	223	3	116.25 *
60	54	3	115.25 *
61	184	3	113.75 *
62	37	3	113.75 *
63	[illegible]	3	113.75 *
64	[illegible]	3	113.75 *
65	[illegible]	3	113.75 *
66	[illegible]	3	111.25 *
67	27	3	111.25 *
68	210	3	111.25 *
69	[illegible]	3	111.25 *
70	[illegible]	3	111.25 *
71	[illegible]	3	111.25 *
72	[illegible]	3	111.25 *
73	[illegible]	3	111.25 *
74	[illegible]	4	110.00 *
75	[illegible]	3	106.25 *
76	72	3	106.25 *
77	[illegible]	3	103.75 *
78	[illegible]	3	103.75 *
79	[illegible]	3	103.75 *
80	[illegible]	3	103.75 *
81	[illegible]	3	103.75 *
82	[illegible]	3	103.75 *
83	[illegible]	3	103.75 *
84	[illegible]	3	103.75 *
85	[illegible]	3	103.75 *
86	41	3	103.75 *
87	[illegible]	3	103.75 *
88	4	3	103.75 *
89	[illegible]	3	103.75 *
90	[illegible]	3	101.25 *
91	[illegible]	4	100.00 *
92	[illegible]	3	98.75 *
93	[illegible]	3	98.75 *
94	[illegible]	3	96.25 *
95	[illegible]	3	96.25 *
96	[illegible]	3	96.25 *
97	[illegible]	3	96.25 *
98	[illegible]	3	95.25 *
99	[illegible]	3	96.25 *
100	[illegible]	3	96.25 *

AVEC 10 DEGRES *

	CODE	NBR	POINTS *
1	302	5	203.50 *
2	301	5	194.50 *
3	117	4	179.13 *
4	110	5	173.50 *
5	200	5	170.50 *
6	36	4	169.13 *
7	275	4	167.13 *
8	59	4	165.63 *
9	34	4	165.63 *
10	227	4	165.63 *
11	25	4	163.13 *
12	264	4	162.63 *
13	212	4	157.67 *
14	44	5	152.50 *
15	312	4	151.13 *
16	66	5	149.50 *
17	163	4	149.00 *
18	33	4	148.67 *
19	255	4	147.63 *
20	55	4	146.63 *
21	270	4	144.63 *
22	256	3	142.17 *
23	226	4	141.92 *
24	303	4	141.92 *
25	131	4	141.92 *
26	304	4	141.63 *
27	69	4	138.92 *
28	235	3	137.34 *
29	125	4	137.13 *
30	64	4	134.13 *
31	133	4	133.67 *
32	128	3	133.17 *
33	222	4	131.63 *
34	221	4	130.13 *
35	237	3	129.76 *
36	228	4	129.63 *
37	185	3	128.76 *
38	189	3	128.05 *
39	29	4	125.63 *
40	3	4	125.63 *
41	107	3	125.34 *
42	242	3	125.34 *
43	115	3	124.30 *
44	174	4	121.67 *
45	223	3	121.50 *
46	9	3	121.30 *
47	76	3	120.76 *
48	10	3	118.80 *
49	71	3	118.80 *
50	12	3	118.80 *
51	266	4	118.67 *
52	22	3	118.50 *
53	89	4	118.13 *
54	3	3	116.34 *
55	83	3	116.26 *
56	195	3	116.25 *
57	31	4	115.13 *
58	60	3	115.05 *
59	54	3	114.84 *
60	193	3	114.76 *
61	7	3	114.76 *
62	218	3	114.76 *
63	197	3	114.55 *
64	196	3	113.55 *
65	179	3	113.26 *
66	241	3	113.26 *
67	210	3	112.76 *
68	253	4	112.67 *
69	219	4	112.67 *
70	272	4	111.92 *
71	181	3	110.76 *
72	121	3	110.26 *
73	309	3	109.76 *
74	164	3	109.30 *
75	27	3	108.76 *
76	296	3	108.26 *
77	300	3	108.09 *
78	160	3	106.80 *
79	37	3	106.75 *
80	233	3	106.76 *
81	50	3	105.76 *
82	4	3	104.26 *
83	41	3	102.76 *
84	205	3	102.26 *
85	72	3	102.09 *
86	258	3	101.80 *
87	259	3	101.76 *
88	262	3	101.26 *
89	86	4	100.13 *
90	190	3	99.76 *
91	119	3	97.76 *
92	287	3	97.26 *
93	94	3	96.26 *
94	194	3	96.26 *
95	132	3	96.26 *
96	18	3	94.80 *
97	244	3	94.76 *
98	78	3	94.26 *
99	214	3	94.05 *
100	139	3	93.76 *

AVEC 40 DEGRES *

	CODE	NBR	POINTS *
1	302	5	200.50 *
2	301	5	187.00 *
3	110	5	176.50 *
4	117	4	175.24 *
5	36	4	171.13 *
6	200	5	170.50 *
7	59	4	167.50 *
8	34	4	166.92 *
9	275	4	164.74 *
10	264	4	163.83 *
11	227	4	163.08 *
12	212	4	159.38 *
13	25	4	159.13 *
14	312	4	156.38 *
15	163	4	151.63 *
16	33	4	150.63 *
17	44	5	150.25 *
18	131	4	147.46 *
19	55	4	147.24 *
20	255	4	146.50 *
21	270	4	145.00 *
22	66	5	144.25 *
23	304	4	143.58 *
24	256	3	142.05 *
25	235	3	141.55 *
26	125	4	140.69 *
27	303	4	139.40 *
28	226	4	139.21 *
29	222	4	136.46 *
30	69	4	136.44 *
31	64	4	134.50 *
32	133	4	133.55 *
33	135	3	133.10 *
34	128	3	131.94 *
35	221	4	131.64 *
36	237	3	131.04 *
37	189	3	130.91 *
38	228	4	129.08 *
39	107	3	126.98 *
40	242	3	126.95 *
41	29	4	126.69 *
42	89	4	125.01 *
43	174	4	124.86 *
44	115	3	124.68 *
45	272	4	123.94 *
46	223	3	123.57 *
47	9	3	123.57 *
48	76	3	122.71 *
49	3	4	121.92 *
50	12	3	121.84 *
51	71	3	120.08 *
52	241	3	118.53 *
53	195	3	117.85 *
54	121	3	117.63 *
55	22	3	117.00 *
56	150	3	116.84 *
57	10	3	116.80 *
58	7	3	116.76 *
59	179	3	116.72 *
60	253	4	116.15 *
61	60	3	115.95 *
62	255	4	115.48 *
63	27	3	115.44 *
64	54	3	114.60 *
65	37	3	114.56 *
66	196	3	114.50 *
67	197	3	114.18 *
68	210	3	113.66 *
69	193	3	113.04 *
70	218	3	113.04 *
71	83	3	111.61 *
72	296	3	111.57 *
73	8	3	110.32 *
74	233	3	109.50 *
75	181	3	109.39 *
76	300	3	109.30 *
77	309	3	108.41 *
78	219	4	108.38 *
79	50	3	107.59 *
80	31	4	107.14 *
81	72	3	106.34 *
82	262	3	105.10 *
83	184	3	104.85 *
84	289	3	104.13 *
85	41	3	103.29 *
86	246	3	103.10 *
87	247	3	103.10 *
88	248	3	103.10 *
89	249	3	103.10 *
90	250	3	103.10 *
91	251	3	103.10 *
92	287	3	102.69 *
93	205	3	102.65 *
94	139	3	102.45 *
95	258	3	101.99 *
96	4	3	101.91 *
97	190	3	100.93 *
98	194	3	99.63 *
99	119	3	98.99 *
100	94	3	98.62 *

./..

101	94	3	96.25 *	101	313	3	93.76 *	101	18	3	97.34 *
102	119	3	96.25 *	102	13	3	93.42 *	102	132	3	97.32 *
103	191	3	96.25 *	103	266	2	91.93 *	103	78	3	96.04 *
104	244	3	95.25 *	104	191	3	91.76 *	104	244	3	95.45 *
105	13	3	93.75 *	105	192	3	91.76 *	105	19	3	94.77 *
106	118	3	91.25 *	106	90	3	91.63 *	106	217	3	94.76 *
107	61	3	91.25 *	107	21	3	91.26 *	107	118	3	94.61 *
108	79	3	91.25 *	108	158	3	91.05 *	108	120	3	94.46 *
109	168	3	91.25 *	109	120	3	90.42 *	109	243	3	94.25 *
110	138	3	91.25 *	110	19	3	90.26 *	110	13	3	93.49 *
111	90	3	91.25 *	111	295	3	90.09 *	111	192	3	93.14 *
112	18	3	91.25 *	112	20	3	89.76 *	112	191	3	93.14 *
113	199	3	91.25 *	113	247	3	89.26 *	113	38	3	93.04 *
114	248	3	88.75 *	114	248	3	89.26 *	114	236	2	92.25 *
115	243	3	88.75 *	115	249	3	89.26 *	115	90	3	92.07 *
116	282	3	88.75 *	116	250	3	89.26 *	116	214	3	91.68 *
117	250	3	88.75 *	117	251	3	89.26 *	117	86	4	91.39 *
118	251	3	88.75 *	118	217	3	89.26 *	118	313	3	89.99 *
119	293	3	88.75 *	119	246	3	89.26 *	119	295	3	89.76 *
120	175	3	88.75 *	120	118	3	89.05 *	120	149	3	89.76 *
121	65	3	88.75 *	121	166	3	88.76 *	121	20	3	89.60 *
122	160	3	88.75 *	122	261	3	88.26 *	122	168	3	87.86 *
123	257	3	88.75 *	123	65	3	86.26 *	123	175	3	87.73 *
124	23	3	88.75 *	124	175	3	84.76 *	124	146	3	87.50 *
125	57	3	88.75 *	125	41	3	84.42 *	125	225	3	85.63 *
126	19	3	88.75 *	126	79	3	84.09 *	126	21	3	85.60 *
127	246	3	86.75 *	127	2	3	83.80 *	127	65	3	85.53 *
128	247	3	86.75 *	128	57	3	83.55 *	128	199	3	85.13 *
129	313	3	86.75 *	129	225	3	83.05 *	129	307	2	82.96 *
130	1	3	85.25 *	130	199	3	82.30 *	130	57	3	82.66 *
131	285	2	85.00 *	131	149	3	82.17 *	131	261	3	82.47 *
132	108	3	83.75 *	132	1	3	82.17 *	132	293	3	82.31 *
133	76	3	81.25 *	133	243	3	81.76 *	133	146	3	82.19 *
134	243	3	81.25 *	134	293	3	81.76 *	134	1	3	81.88 *
135	17	3	81.25 *	135	254	3	81.26 *	135	97	2	81.50 *
136	225	3	81.25 *	136	307	2	81.22 *	136	2	3	81.46 *
137	33	3	81.25 *	137	105	2	79.55 *	137	51	3	81.15 *
138	102	3	81.25 *	138	38	3	79.26 *	138	106	2	80.72 *
139	186	3	81.25 *	139	265	3	77.80 *	139	254	3	80.69 *
140	141	3	81.25 *	140	171	3	77.55 *	140	79	3	80.28 *
141	217	3	81.25 *	141	202	2	77.18 *	141	282	3	79.34 *
142	146	3	81.25 *	142	262	3	76.76 *	142	45	2	78.13 *
143	254	3	81.25 *	143	97	2	76.43 *	143	305	3	77.52 *
144	21	3	81.25 *	144	23	2	75.97 *	144	202	2	77.39 *
145	307	2	80.00 *	145	114	2	75.97 *	145	114	2	77.37 *
146	202	2	77.50 *	146	45	2	74.93 *	146	17	3	77.09 *
147	5	2	77.50 *	147	305	3	74.80 *	147	23	2	75.99 *
148	209	3	76.25 *	148	146	3	74.76 *	148	265	3	75.73 *
149	265	3	76.25 *	149	127	2	74.43 *	149	306	2	75.70 *
150	2	3	76.25 *	150	17	3	74.26 *	150	127	2	75.24 *
151	305	3	76.25 *	151	306	2	73.97 *	151	141	3	74.63 *
152	52	2	75.00 *	152	102	3	73.26 *	152	297	2	74.61 *
153	114	2	75.00 *	153	180	3	73.25 *	153	292	3	73.87 *
154	6	2	75.00 *	154	5	2	71.59 *	154	70	2	72.66 *
155	23	2	75.00 *	155	62	2	71.26 *	155	62	2	70.55 *
156	95	3	73.75 *	156	177	2	71.26 *	156	102	3	69.93 *
157	234	3	73.75 *	157	209	3	70.80 *	157	177	2	69.22 *
158	261	3	73.75 *	158	70	2	70.43 *	158	188	3	69.10 *
159	106	2	72.50 *	159	266	2	69.68 *	159	5	2	68.62 *
160	170	2	72.50 *	160	51	2	66.92 *	160	170	2	68.22 *
161	45	2	72.50 *	161	297	2	66.39 *	161	209	3	68.08 *
162	97	2	72.50 *	162	126	2	66.93 *	162	155	2	67.69 *
163	155	2	72.50 *	163	155	2	66.93 *	163	198	3	67.64 *
164	70	2	72.50 *	164	292	3	66.76 *	164	126	2	67.39 *
165	127	2	72.50 *	165	116	2	66.18 *	165	6	2	66.61 *
166	177	2	70.00 *	166	6	2	65.69 *	166	51	2	66.40 *
167	297	2	70.00 *	167	170	2	64.76 *	167	268	2	66.12 *
168	80	2	70.00 *	168	257	3	63.76 *	168	28	2	65.45 *
169	268	2	70.00 *	169	156	2	63.05 *	169	80	2	64.97 *
170	11	2	67.50 *	170	80	2	62.68 *	170	150	2	64.74 *
171	306	2	67.50 *	171	285	2	61.89 *	171	116	2	63.62 *
172	51	2	67.50 *	172	40	2	61.39 *	172	40	2	62.64 *
173	125	2	65.00 *	173	62	2	61.26 *	173	56	2	62.63 *
174	155	2	65.00 *	174	176	2	60.39 *	174	259	2	61.68 *
175	206	2	65.00 *	175	150	2	60.22 *	175	156	2	60.48 *
176	92	2	65.00 *	176	188	3	59.55 *	176	285	2	59.41 *
177	150	2	65.00 *	177	23	2	59.39 *	177	203	2	59.39 *
178	229	2	62.50 *	178	56	2	58.97 *	178	176	2	59.36 *
179	235	2	62.50 *	179	229	2	58.89 *	179	11	2	59.34 *
180	56	2	62.50 *	180	213	2	58.89 *	180	229	2	59.28 *
181	175	2	62.50 *	181	11	2	58.30 *	181	231	2	59.23 *
182	252	3	59.75 *	182	141	3	57.76 *	182	122	2	58.05 *
183	58	2	57.50 *	183	108	3	57.63 *	183	206	2	57.37 *
184	122	2	57.50 *	184	206	2	56.93 *	184	213	2	56.86 *
185	167	2	57.50 *	185	224	2	56.89 *	185	178	2	56.46 *
186	213	2	55.00 *	186	96	2	56.18 *	186	98	2	56.45 *
187	46	2	55.00 *	187	167	2	56.18 *	187	157	2	55.74 *
188	116	2	55.00 *	188	122	2	55.69 *	188	39	2	54.98 *
189	248	2	55.00 *	189	259	2	54.76 *	189	233	2	54.22 *
190	98	2	55.00 *	190	203	2	54.39 *	190	32	2	53.80 *
191	249	2	55.00 *	191	259	2	54.39 *	191	280	2	53.74 *
192	40	2	55.00 *	192	45	2	54.39 *	192	299	2	53.35 *
193	253	2	55.00 *	193	299	2	53.89 *	193	46	2	53.[illegible] *
194	28	2	55.00 *	194	48	3	53.76 *	194	224	2	52.77 *
195	152	2	55.00 *	195	290	2	52.76 *	195	288	2	52.[illegible] *
196	239	2	55.00 *	196	283	2	51.43 *	196	152	2	51.[illegible] *
197	140	2	52.50 *	197	288	2	51.39 *	197	269	2	51.7[illegible] *
198	97	3	51.25 *	198	280	2	50.69 *	198	294	2	50.[illegible] *
199	48	3	51.25 *	199	96	3	50.26 *	199	284	2	48.[illegible] *
200	35	3	51.25 *	200	294	2	49.93 *	200	43	2	47.[illegible] *

./..

201	294	2	50.00 *	201	24	2	49.43 *	201	52	2	45.55 *
202	290	2	50.00 *	202	281	2	49.29 *	202	24	2	45.4[illegible] *
203	283	2	50.00 *	203	52	2	47.39 *	203	104	2	43.34 *
204	24	2	50.00 *	204	178	2	47.26 *	204	87	3	42.[illegible] *
205	240	2	47.50 *	205	43	2	46.89 *	205	239	2	42.[illegible]7 *
206	281	2	47.50 *	206	39	2	45.97 *	206	48	3	42.24 *
207	224	2	47.50 *	207	152	2	44.89 *	207	240	2	41.14 *
208	239	2	47.50 *	208	239	2	43.68 *	208	204	2	40.[illegible]7 *
209	240	2	47.50 *	209	173	2	41.43 *	209	173	2	40.53 *
210	203	2	47.50 *	210	264	2	41.39 *	210	171	3	39.74 *
211	230	2	47.50 *	211	298	2	41.18 *	211	298	2	38.13 *
212	173	2	42.50 *	212	240	2	40.83 *	212	26	2	37.[illegible]9 *
213	77	2	42.50 *	213	230	2	39.97 *	213	73	1	36.59 *
214	178	2	42.50 *	214	26	2	39.39 *	214	230	2	35.[illegible] *
215	39	2	42.50 *	215	204	2	37.43 *	215	208	2	34.[illegible]3 *
216	16	2	40.00 *	216	73	1	32.10 *	216	180	3	33.[illegible] *
217	220	2	35.00 *	217	104	2	30.89 *	217	236	1	33.51 *
218	73	1	33.75 *	218	291	1	30.81 *	218	81	1	31.[illegible]1 *
219	236	1	33.75 *	219	16	2	30.39 *	219	257	3	30.71 *
220	43	2	32.50 *	220	234	3	29.76 *	220	291	1	27.[illegible] *
221	260	2	32.50 *	221	140	2	29.47 *	221	16	2	27.[illegible] *
222	52	2	32.50 *	222	236	1	27.85 *	222	88	1	26.[illegible] *
223	104	2	32.50 *	223	77	2	27.76 *	223	172	1	24.[illegible] *
224	26	2	32.50 *	224	81	1	27.52 *	224	109	1	24.[illegible] *
225	81	1	28.75 *	225	87	3	24.76 *	225	113	1	23.[illegible]4 *
226	291	1	28.75 *	226	113	1	22.52 *	226	42	1	22.02 *
227	204	2	27.50 *	227	58	2	22.43 *	227	274	1	21.[illegible] *
228	284	2	25.00 *	228	98	1	21.02 *	228	250	2	20.[illegible] *
229	201	2	25.00 *	229	172	1	20.81 *	229	183	2	17.3[illegible] *
230	298	2	25.00 *	230	109	1	19.06 *	230	310	1	15.32 *
231	109	1	23.75 *	231	274	1	18.85 *	231	263	1	13.52 *
232	42	1	23.75 *	232	42	1	18.31 *	232	220	2	12.9[illegible] *
233	172	1	21.25 *	233	260	2	17.39 *	233	96	3	10.22 *
234	113	1	21.25 *	234	310	1	15.52 *	234	198	1	9.[illegible] *
235	208	2	20.00 *	235	208	2	15.26 *	235	267	1	9.54 *
236	274	1	18.75 *	236	220	2	9.76 *	236	252	1	6.75 *
237	88	1	13.75 *	237	101	1	8.81 *	237	215	1	6.[illegible] *
238	310	1	13.75 *	238	35	3	8.76 *	238	101	1	4.05 *
239	183	2	10.00 *	239	198	1	7.56 *	239	308	1	0.51 *
240	101	1	6.25 *	240	201	2	7.43 *	240	232	1	-2.09 *
241	148	1	3.75 *	241	263	1	7.02 *	241	207	1	-2.[illegible]5 *
242	215	1	3.75 *	242	215	1	2.60 *	242	77	2	-8.[illegible] *
243	187	2	2.50 *	243	267	1	2.52 *	243	58	2	-9.[illegible] *
244	198	1	1.25 *	244	252	1	0.81 *	244	140	2	-10.6[illegible] *
245	308	1	-1.25 *	245	308	1	-0.45 *	245	35	3	-11.22 *
246	263	1	-1.25 *	246	74	1	-7.32 *	246	290	2	-13.[illegible]1 *
247	267	1	-1.25 *	247	183	2	-9.61 *	247	216	1	-15.[illegible]2 *
248	207	1	-1.25 *	248	207	1	-9.98 *	248	74	1	-16.7[illegible] *
249	252	1	-1.25 *	249	232	1	-11.48 *	249	245	1	-20.71 *
250	74	1	-6.25 *	250	311	1	-15.98 *	250	148	1	-22.3[illegible] *
251	232	1	-8.75 *	251	245	1	-17.98 *	251	182	1	-32.7[illegible] *
252	182	1	-16.25 *	252	148	1	-24.57 *	252	201	2	-35.[illegible] *
253	95	1	-21.25 *	253	216	1	-39.44 *	253	311	1	-37.[illegible] *
254	245	1	-23.75 *	254	187	2	-44.11 *	254	95	1	-37.7[illegible] *
255	211	1	-36.25 *	255	192	1	-52.98 *	255	234	3	-46.72 *
256	311	1	-36.75 *	256	211	1	-57.94 *	256	211	1	-94.[illegible] *
257	216	1	-38.75 *	257	95	1	-59.11 *	257	238	1	-115.74 *
258	238	1	-38.75 *	258	238	1	-63.98 *	258	187	2	-124.[illegible]4 *

Chaque sujet pouvait ainsi se situer par rapport à l'ensemble des autres. Cette liste (tableau 5.26) permet de constater l'impact important de la certitude sur le total des points. Pour s'en convaincre, il suffit de comparer, dans la colonne de droite (40 degrés), le total du 114e sujet (92,25 points obtenus avec deux bonnes réponses) à celui du 117e sujet (91,39 points avec quatre bonnes réponses). Ces deux valeurs sont désignées par des flèches sur la liste.

On trouvera, en annexe 5, les programmes FORTRAN qui permettent d'établir les documents présentés dans les tableaux 5.25 et 5.26.

Les résultats ont été explicités aux participants par deux pages de "commentaires" (tableau 5.27).

Tableau 5.27

COMMENTAIRES SUR LES 258 RESULTATS DE L'EXPERIENCE 1

Voici les phrases complètes (extraites de S. FREUD, TOTEM et TABOU, Ed. Payot, Paris, 1965, p. 64).

1. POURQUOI L'ATTITUDE AFFECTIVE A L'EGARD DU SOUVERAIN COMPORTE-T-ELLE UN ELEMENT SI PUISSANT D'HOSTITLITE INCONSCIENTE?LA QUESTION EST TRES INTERESSANTE,MAIS SA S

2. OLUTION DEPASSERAIT LE CADRE DE CE TRAVAIL.NOUS AVONS DEJA FAIT ALLUSION AU COMPLEXE PATERNEL DE L'ENFANCE;AJOUTONS ENCORE QUE L'EXAMEN DE L'HISTOIRE PRIMITIVE

3. DE LA ROYAUTE SERAIT DE NATURE A NOUS APPORTER UNE REPONSE DECISIVE A CETTE QUESTION.D'APRES LES EXPLICATIONS TRES IMPRESSIONNANTES,MAIS,DE SON PROPRE AVIS,PEU

4. PROBANTES DE FRAZER,LES PREMIERS ROIS ETAIENT DES ETRANGERS QUI,APRES UNE BREVE PERIODE DE REGNE,ETAIENT SACRIFIES A LA DIVINITE DONT ILS ETAIENT LES REPRESENTA

5. NTS,AVEC ACCOMPAGNEMENT DE FETES SOLENNELLES.ON RETROUVE ENCORE L'ECHO DE CETTE HISTOIRE PRIMITIVE DE LA ROYAUTE DANS LES MYTHES DU CHRISTIANISME.

Les résultats, différents selon les trois échelles de certitude, sont classés (des points les plus élevés aux moins élevés). Il est évident que, sur cinq questions, le hasard a joué énormément (ainsi à la 2e ligne, R était au moins aussi probable que T, la bonne réponse). Avec 100 réponses, les hasards heureux et malheureux se compensant, l'évaluation précise des chances d'être correcte de chaque réponse prendra encore plus d'importance. A partir de ces résultats, on peut faire les constatations qui suivent.

./..

../.

Constatation 1 : 6 personnes (soit 2 % des personnes) ont fourni 5 bonnes réponses
36 personnes (soit 14 % des personnes) ont fourni 4 bonnes réponses
112 personnes (soit 43 % des personnes) ont fourni 3 bonnes réponses
75 personnes (soit 29 % des personnes) ont fourni 2 bonnes réponses
29 personnes (soit 11 % des personnes) ont fourni 1 bonne réponse.
Aucun sujet n'a eu la malchance de fournir 5 réponses incorrectes.

Constatation 2 : Le nombre de bonnes réponses a, bien entendu, une influence prépondérante sur le total des points de chaque sujet. Mais l'indice de certitude a, lui aussi, une grande importance. Ainsi, dans l'échelle à 40 degrés, les 112 personnes qui ont trois réponses se classent de la 24e à la 255e place (les points vont de 142 à -46). De même, on trouve en 114e position le sujet 286 qui n'a que 2 réponses correctes et en 117e position le sujet 86 qui a 4 réponses correctes, soit le double.

Constatation 3 : Les points, dans l'ensemble, sont à peu près les mêmes dans les trois échelles (voir, dans chaque échelle, les points des 1e, 50e, 100e, 150e, 200e.....sujets) sauf pour les tout derniers sujets (cfr le 258e sujet) où les pertes sont très sévères dans les échelles à 10 et surtout 40 degrés.

Constatation 4 : Dans l'ensemble, la place des individus est assez semblable d'une échelle à l'autre, mais il existe des cas de discordance nette. Ainsi, le sujet 66 est respectivement 8e, 16e et 22e dans les trois échelles.

F. PERSPECTIVES

Le jeu des ESPER, inspiré du SHANNON Guessing Game, est à la base de toutes les procédures qui ont été développées dans le présent chapitre et qui permettront des expériences décrites dans les chapitres 6 et 7.

Ce jeu présente, entre autres, les avantages suivants :

1. Il est proche de la situation pédagogique par deux aspects importants :
   a) On peut déterminer objectivement si la réponse d'un sujet est exacte ou non.
   b) Il n'existe pas de moyen de calculer objectivement la PER d'un sujet donné à une question donnée.
2. Il se prête bien à l'expérimentation, pour les raisons suivantes :
   a) Les contenus possibles sont quasi infinis, facilement disponibles, quels que soient l'âge ou les intérêts des sujets.
   b) Le niveau moyen de difficulté de la tâche peut être facilement manipulé.
   c) La consigne est uniforme et simple à comprendre.
   d) La présentation (par ordinateur) est rapide et peu coûteuse. Il en va de même pour la liste des bonnes réponses.
   e) La tâche porte en elle-même un élément de motivation pour le sujet.
   f) Il peut être utilisé dans une procédure test-retest, avec certains inconvénients (qui seront envisagés dans le chapitre 6).
   g) Il peut être modifié de manière à tester le théorème de BAYES relatif à la révision des probabilités (voir chapitre 7).

Les usages de ce jeu nous paraissent dépasser largement les applications décrites dans les chapitres 6 et 7. Cette technique pourrait s'avérer utile dans l'étude de la compréhension de textes, la rapidité de réponse, la formulation d'hypothèses, l'influence sociale sur l'ESPER (prédictions en groupe et non plus individuelles, etc.).

CHAPITRE 6

ETUDE DE LA STABILITE ET DE LA SENSIBILITE DE L'ESPER PAR UNE EXPERIENCE TEST-RETEST

A. Les objectifs et les méthodes de l'étude.

B. Les faiblesses de la technique.

C. Les histogrammes de replication.

D. Les hypothèses.

E. Présentation générale des données.

F. Vérification des hypothèses.

G. Le problème de la longueur de l'épreuve.

H. Les réponses des sujets au questionnaire.

I. Conclusions.

A. LES OBJECTIFS ET LES METHODES DE L'ETUDE

L'évaluation la plus directe de la stabilité de l'ESPER consiste à recueillir, à des intervalles temporels donnés, l'indice de certitude qu'un individu fournit pour une même réponse.

Plusieurs conditions sont requises pour assurer la qualité d'une telle mesure :

1. Les indices de certitude fournis lors d'une séance antérieure ne peuvent pas être mémorisés par le sujet.
2. Les sujets ne peuvent pas pouvoir "acquérir de l'information" sur le contenu de l'expérience entre deux séances.
3. Les sujets devraient, au cours des séances successives, réenvisager strictement les mêmes hypothèses.
4. La matrice des conséquences doit être calculée selon le critère de l'E.S.U.; la situation doit être operante et les renforcements proportionnels au score total.

OBJECTIF: Effectuer une mesure de la stabilité de l'ESPER en recueillant, à deux moments (distants de quinze jours au moins), la certitude des mêmes sujets concernant une même réponse (à une même question).

METHODE : Aux sujets qui ont répondu aux cent questions du texte mutilé de J. FOURASTIER, il a été demandé de fournir à nouveau leur certitude relative à la lettre fournie (inchangée).

Voici des extraits de la lettre qui a été adressée aux sujets.

Tableau 6.0

Madame,
Monsieur,

Vous avez accepté de répondre aux cent questions du jeu des certitudes; je vous remercie pour cet effort.

Dans la présente enveloppe, vous vous attendez à trouver les réponses correctes et les résultats. Or il n'en est rien, car je vous demande maintenant de vous prêter à une autre expérience (beaucoup moins longue que la précédente).

Le but de cette nouvelle expérience est de répondre aux deux questions suivantes :

- *Dans quelle mesure la certitude attribuée à une réponse particulière varie-t-elle d'une fois à l'autre ? Autrement dit, à une ou deux semaines d'intervalle, les certitudes sont-elles les mêmes ?*
- *La qualité dans l'évaluation des pourcentages de chance reste-t-elle constante chez ceux qui répondent ? Autrement dit : ceux qui ont obtenu les meilleurs résultats dans la première épreuve obtiennent-ils encore les meilleurs résultats dans la seconde ? (N.B.: Le nombre de réponses correctes est le même dans les deux expériences puisqu'on ne peut pas changer les réponses).*

Il ne faut rien changer aux lettres que vous aviez proposées. Vos réponses vous sont renvoyées dans cette enveloppe. Il faut simplement indiquer à nouveau les certitudes (sur la nouvelle bandelette en annexe). Pour certains d'entre vous, on propose cependant une échelle différente (voir la lettre indiquant votre groupe). Date de renvoi : lundi 8 avril 1974.

Ne renvoyez que la nouvelle bandelette (vos réponses ont été recopiées).

Dans la prochaine lettre, vous recevrez donc deux séries de résultats : ceux de la première série et ceux de la seconde.
-Les points à gagner et à perdre restent les mêmes.
-Vous avez toujours mathématiquement intérêt à estimer le plus exactement possible le pourcentage de chances d'exactitude de chacune de vos réponses.

Il n'est pas agréable derecommencer un travail. Cependant, vous comprendrez pourquoi il n'était pas possible de vous révéler la procédure lors de l'expérience précédente. La présente expérience est la dernière de ce type et de cette longueur. Les réponses aux deux questions ci-dessus seront communiquées (plus tard) à tous ceux qui auront permis, par leur participation, d'y répondre.

Des échelles différentes ont donc été proposées aux sujets :

- L'échelle A, à 4 degrés de certitude;
- L'échelle B, à 10 degrés de certitude;
- L'échelle C, à 40 degrés de certitude.

Certains sujets ont dû répondre, lors du test ou du retest, à l'aide des trois échelles, pour chacune de leurs réponses. Le type d'échelle sera noté par l'indice d (degrés).

Les données de base de l'expérience sont contenues dans une matrice $ESPER_{iqmd}$ à quatre dimensions, avec
nq = 100
nm = 2 (m = 1 pour le test et m = 2 pour le retest)
nd = 3 (d = 1 pour A, d = 2 pour B et d = 3 pour C).

L'évaluation de la stabilité de l'ESPER consiste à élaborer les données suivantes pour chaque sujet:

- L'histogramme d'utilisation de chaque degré (pour chaque échelle, au test et au retest),
  avec la moyenne de ces ESPER : $M|_{imd}$ ESPER
  et l'écart-type de ces ESPER : $S|_{imd}$ ESPER
- Le nuage de points formé par la distribution des degrés de certitude lors du test (m = 1) et du retest (m = 2), avec la corrélation r_{id} (ESPER | m = 1, ESPER | m = 2).

Le tableau 6.1 présente ces données pour deux sujets (S.262 et S.199), telles qu'elles leur ont été envoyées. Le tableau 6.2 fournit le détail des résultats individuels et le total. On trouvera, en annexe 7, le programme FORTRAN qui permet d'établir ces résultats.

Tableau 6.1

```
VOTRE CODE =  262
   BR MR          0 1 2 3 4 5 6 7 8 9
0   0  0
1   2  3          1 2 1   1
2   0  5              2 1 2
3   1  4          1   3   1
4   5 10            1 2 3 2 2       1 4
5   4  6                1 2 3 1 1     2
6   9  2          1 1   1 1 2     3   2
7   7  3                  1   1 3     5
8   3  1                      1       3
9  32  3                  1 3     1  30

BR =              1 1 2 1 7 5 2 6 038
MR =              2 3 6 5 4 5 1 2 1 8

M1 =  7.23    SIG 1 =  2.56
M2 =  7.28    SIG 2 =  2.94
R =   0.7269

VOTRE CODE =  262
(AVANT)  10  DEGRES
 0  *
 1  *****
 2  *****
 3  *****
 4  ***************
 5  **********
 6  ***********
 7  **********
 8  ****
 9  ***********************************

VOTRE CODE =  262
(APRES)  10  DEGRES
 0  ***
 1  ****
 2  ********
 3  ******
 4  ***********
 5  **********
 6  ***
 7  ********
 8  *
 9  ************************************** **
```

```
VOTRE CODE =  199
   BR MR          0 1 2 3 4 5 6 7 8 9
0   1  0          1
1   0  6              4 1 1
2   5  5          1   5 2 1 1
3   5  9              2   5 5 2
4   6  7                2 4 3 1 1 2
5   6  3                1 2 5     1
6  11  2                1   5 2 4   1
7   9  3              1 1   3 3 1 3
8   9  4                  1 3   6 3
9   8  1                    2 1   3 3

BR =              1 0 3 3 419 8 9 9 4
MR =              1 0 9 510 8 1 3 3 0

M1 =  6.12    SIG 1 =  2.45
M2 =  6.09    SIG 2 =  2.08
R =   0.6947

VOTRE CODE =  199
(AVANT)  10  DEGRES
 0  *
 1  ******
 2  **********
 3  **************
 4  *************
 5  *********
 6  *************
 7  ************
 8  *************
 9  *********

VOTRE CODE =  199
(APRES)  10  DEGRES
 0  **
 1  *
 2  ************
 3  ********
 4  **************
 5  ***************************
 6  *********
 7  ************
 8  ************
 9  ****
```

Utilisation des indices de certitude lors de l'expérience principale, au test (avant) et au retest (après). Exemples : résultats des sujets 262 et 199

Tableau 6.2

VOTRE CODE = 262

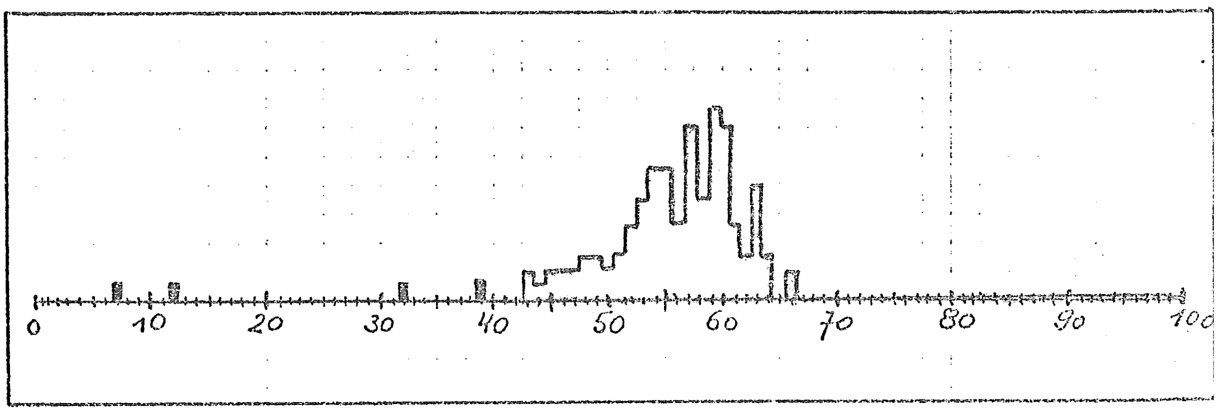
QUESTION NUMERO	BONNE REPONSE	VOTRE REPONSE	AVANT(10 DEGRES) CERT.	POINTS	APRES(10 DEGRES) CERT.	POINTS
1	–	E	6	-11.87	5	-7.37
2	–	A	1	-0.33	1	-0.33
3	I	I	9	48.50	9	48.50
4	X	–	9	-57.87	9	-57.87
5	P	P	9	48.50	9	48.50
6	U	U	5	36.50	6	39.50
7	U	U	7	42.50	9	48.50
8	U	U	8	45.50	9	48.50
9	A	A	6	39.50	1	24.50
10	R	R	9	48.50	9	48.50
11	U	U	9	48.50	9	48.50
12	E	A	3	-2.37	2	-1.08
⋮	⋮	⋮	⋮	⋮	⋮	⋮
95	S	S	5	36.50	5	36.50
96	U	U	9	48.50	9	48.50
97	I	I	7	42.50	9	48.50
98	S	I	5	-7.37	9	-57.87
99	A	A	7	42.50	9	48.50
100	T	G	4	-4.37	2	-1.08
NBR = 63				2345.88		2108.41

Extrait des résultats détaillés du sujet 262 tels qu'ils lui ont été transmis

B. FAIBLESSES DE LA TECHNIQUE

Trois sujets, sur 124, présentent un nombre de bonnes réponses anormalement bas. Voici la distribution de ces résultats totaux (tableau 6.3).

Tableau 6.3



Distribution des nombres de réponses correctes des 124 sujets de l'expérience principale

Ces résultats dénoncent une faiblesse (prévisible) des conditions de l'expérience : il n'est pas possible de contrôler le sérieux apporté à la tâche. Afin de ne pas fausser les interprétations ultérieures, les trois sujets qui ont fourni 7,12 et 32 bonnes réponses (sujets 77, 76 et 27) ont été éliminés des analyses.

Mais les faiblesses de ce jeu des certitudes ne s'arrêtent pas là. En effet, certains comportements indésirables (1) se produisent même si les sujets ont essayé de répondre sérieusement.

(1) Aucune valeur péjorative n'est attachée à ce mot. Simplement, ces comportements sont indésirables (mais difficilement évitables) dans une étude sur la stabilité des ESPER.

Premier comportement indésirable : LE CHOIX ALEATOIRE D'UN DEGRE DE CERTITUDE.

Il s'agit de l'expression de certitudes sans rapport avec les PER réelles. On peut déceler ce type de comportement au fait que ces sujets obtiennent un total de points largement inférieur à leur "total théorique". On obtient ce total théorique par le calcul suivant :

$$(NBR \times P_1(k)) + (100 - NBR) \times P_2(k))$$

où NBR = nombre de bonnes réponses

P_1 = points (positifs) attribués à une bonne réponse

P_2 = points (négatifs) attribués à une erreur

k = degré de certitude qui correspond à l'ESPER obtenue par la formule NBR/100.

Les totaux théoriques pour un nombre de réponses correctes compris entre 39 et 66 sont présentés dans le tableau 6.4.

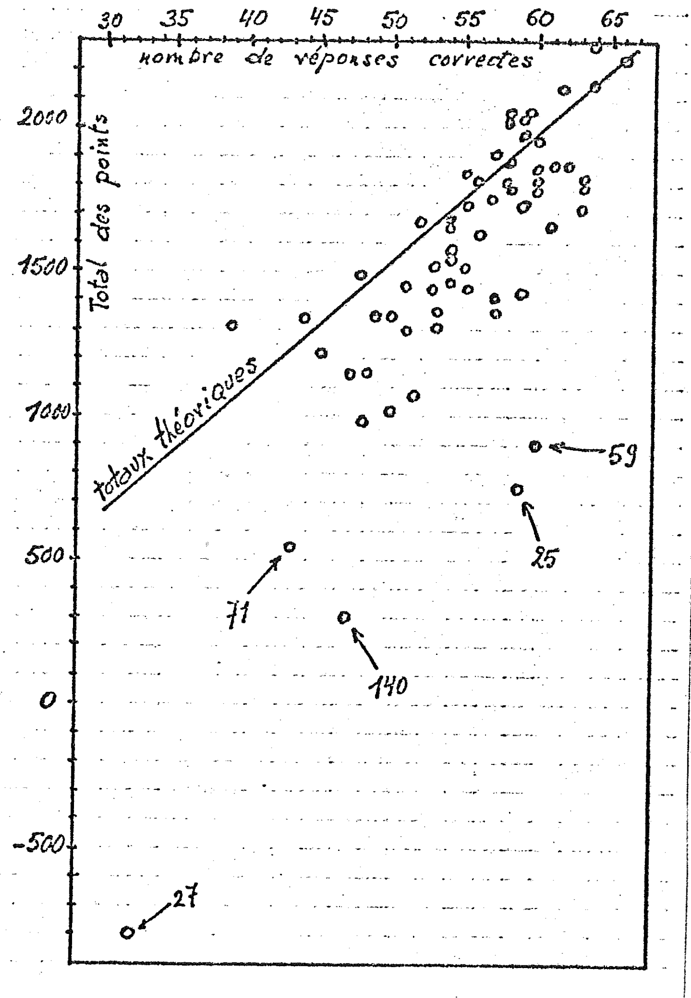
Tableau 6.4

BR	Coefficient C +	−		Coefficient B +	−		Coefficient A +	−	
66			2222,46			2203,42			2217,5
65	39,5	-11,31	2171,65			2152,05			2168,75
64			2120,84	39,5	-11,87	2100,68			2120,0
63	38,75	-10,06	2069,03			2049,31			2071,25
62			2020,22			1997,94			2022,5
61			1969,73			1939,07			1973,75
60	38,00	- 8,93	1922,8			1895,2			1925,0
59			1875,87			1851,33	38,75	-10	1876,25
58	37,25	- 7,92	1827,86			1807,46			1827,5
57			1782,69	36,5	-7,37	1763,59			1778,75
56			1736,0			1719,72			1730,0
55	36,5	- 7	1692,5			1675,85			1681,25
54			1649,0			1631,98			1632,5
53	35,75	- 6,17	1604,76			1588,11			1583,75
52			1562,84			1544,24			1535,0
51			1519,42			1494,37			1471,25
50	35	- 5,42	1479,0			1456,5			1437,5
49			1438,58			1418,63			1403,75
48	43,25	- 4,75	1397,0			1380,76			1370,0
47			1358,0	33,5	-4,37	1342,89			1336,25
46			1317,98			1305,02			1302,5
45	33,5	- 4,13	1280,35			1267,15	31,25	-2,5	1269,0
44			1242,72			1229,28			1235,0
43	32,75	- 3,58	1204,19			1191,41			1201,25
42			1167,86			1153,54			1167,5
41			1130,28			1110,67			1134,0
40	32	- 3,08	1095,2	30,5	-2,37	1077,8			1100,0
39			1060,12			1044,93			1066,25

Espérance mathématique selon le nombre de réponses correctes, pour les trois types d'échelle

Le tableau 6.5 confronte les totaux calculés aux totaux observés (nuage de points).

Tableau 6.5

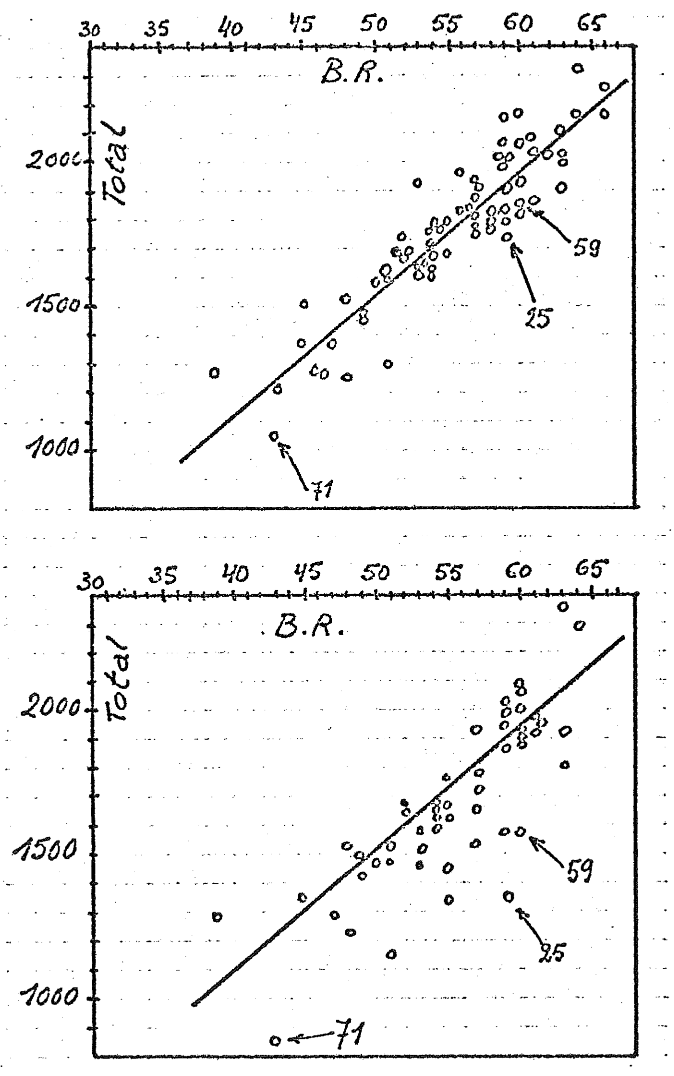


Comparaison, pour l'échelle C (40 degrés) entre les totaux théoriques (voir tableau 6.4) portés sur une droite et les totaux observés lors du test. Les numéros de code des sujets ont été indiqués dans les cas de forte discordance.

Cinq sujets (25, 27, 59, 71 et 140) ont des totaux observés nettement inférieurs aux totaux théoriques. Quant à l'ensemble des autres sujets, leurs totaux sont, le plus souvent, légèrement inférieurs à la droite théorique.

Les sujets 25, 59 et 71 étaient, lors du test, dans le groupe expérimental T. Il est donc possible de les situer sur un graphique semblable (tableau 6.6) pour les échelles A (à 4 degrés) et les échelles B (à 10 degrés).

Tableau 6.6



Comparaison entre les totaux théoriques et les totaux observés, selon les mêmes principes que le tableau 6.5

La discordance entre les totaux et les espérances mathématiques est plus importante pour les sujets 25 et 71 que pour le sujet 59. Pour les deux premiers, en effet, cette discordance se produit pour toutes les échelles. Le sujet 59 est moins discordant dans les échelles A (à 4 degrés) et B (à 10 degrés). Il a fourni des certitudes moins pertinentes dans l'échelle C (à 40 degrés).

Les sujets 27, 140, 25 et 71 seront éliminés des analyses ultérieures.

Deuxième comportement indésirable : LA MODIFICATION DES HYPOTHESES.

Lors du retest, un sujet très respectueux des consignes peut modifier complètement les hypothèses faites lors du test. Dès lors, la réponse à laquelle il avait, au départ, attribué une certitude forte va lui paraître moins probable. Si les nouvelles hypothèses confirment la réponse, celle-ci va, au contraire, paraître plus probable.

Dans les deux cas, se produira une modification de l'ESPER due à un défaut de la procédure expérimentale, et non aux sujets. En effet, ce jeu permet de reposer une même question sans qu'aucune information ne soit intervenue, après un temps assez long. Hélàs, ce n'est pas encore une situation où les hypothèses cognitives restent les mêmes d'une fois à l'autre. Une façon (imparfaite) de pallier cet inconvénient consiste à exiger (1) des sujets qu'ils écrivent lors du test le(s) mot(s) entier(s) au(x)quel(s) ils ont pensé. Lors du retest, on serait assuré que ces mots sont à nouveau envisagés. Cependant, toujours rien ne garantissait que le sujet n'envisage pas de nouveaux mots, compatibles ou non avec la réponse (2). On pourrait aussi demander au sujet de noter, sur feuilles séparées, lors du test, puis lors du post-test, les mots envisagés. On pourrait alors vérifier la concordance des hypothèses entre le test et le retest.

(1) Et non pas suggérer comme nous l'avions fait dans la règle B du règlement. En fait, personne n'a appliqué cette suggestion pour le total des réponses.

(2) Qu'ils soient compatibles ou non, ils modifient (en plus ou en moins) l'ESPER.

Dans l'expérience présente, il est impossible de déceler ces situations qui se sont inmanquablement produites et amèneront à sous-estimer la stabilité de l'ESPER.

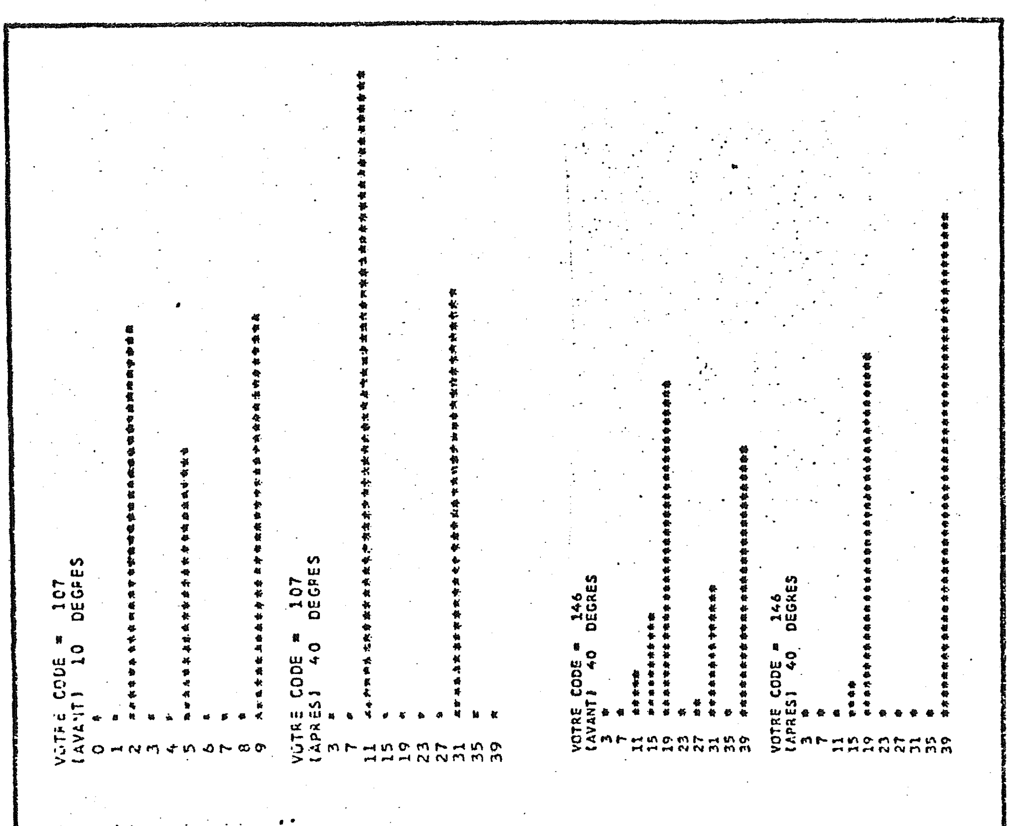
<u>Troisième comportement indésirable</u> : <u>LA MODIFICATION DE STRATEGIE</u>.

Certains sujets peuvent ne pas se conformer au seul critère de la PER, mais adopter l'une des multiples stratégies décrites dans les chapitres précédents. En fait, aucune application systématique de stratégies pures n'a été observée. Tout au plus, trois sujets ont-ils systématiquement utilisé deux des degrés sur les quarante mis à leur disposition dans l'échelle C : au retest,

- le sujet 107 a utilisé les seuls échelons 9 (22,5 %) et 31 (77,5 %),
- le sujet 110 a utilisé les seuls échelons 9 (22,5 %) et 29 (72,5 %),
- le sujet 146 a utilisé les seuls échelons 9 (22,5 %) et 29 (72,5 %).

Ils ont été retenus, quoique leur comportement soit peu propice à l'étude de la stabilité de l'ESPER. Les histogrammes au test et au retest des sujets 107 et 146 figurent au tableau 6.7.

<u>Tableau 6.7</u>



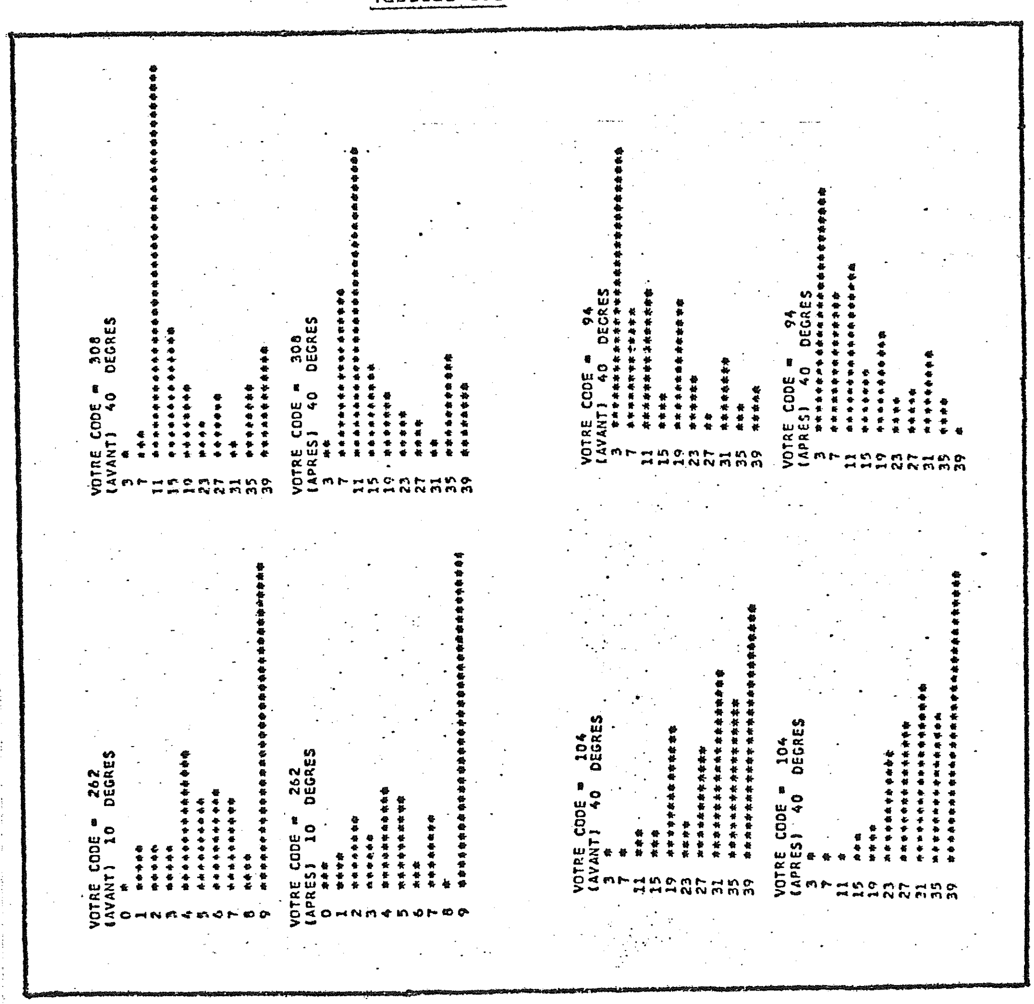
Exemples (sujets 107 et 146) d'utilisation d'un nombre très réduit de degrés d'échelles présentant beaucoup plus de possibilités

La situation n'étant pas operante, il est probable que d'autres sujets ont eu recours à des stratégies <u>différentes</u> au test et au retest, puisqu'aucune contrainte ne les obligeait à appliquer le seul critère de l'E.S.U.

La modification de stratégie peut être constatée au niveau de la distribution globale des certitudes au test et au retest.

Voici des exemples de <u>stabilité globale</u> (histogrammes semblables au test et au retest). Rappelons que n = 100 dans chacun de ces histogrammes.

<u>Tableau 6.8</u>



Exemples (sujets 308, 94, 262, 104) de stabilité globale entre le test et le retest : les histogrammes "avant" et "après" sont fort semblables

Les sujets 127, 170 et 180 constituent des exemples d'histogrammes très différents entre le pré- et le post-test, indiquant un changement de stratégie globale (instabilité globale).

Les sujets 199, 204 et 250 accusent des différences plus faibles entre le test et le retest.

La formule de Kolmogorov-Smirnov (S. SIEGEL, 1956, p. 127) a été utilisée pour quantifier ces différences.

On sait que le test de Kolmogorov-Smirnov est basé sur la différence maximale (D max) entre les fréquences cumulées des deux classes correspondantes. L'hypothèse Ho (équivalence des deux distributions) est rejetée si

$$D \max > 1.22\sqrt{\frac{n1 + n2}{n1\ n2}} \quad \text{à P.10}$$

Or, dans notre cas, n1 = n2 = 100 et $\sqrt{\frac{n1 + n2}{n1\ n2}} = 0{,}141$

Donc Ho est rejetée si D max > à 0,172 (0,141 x 1,22) à P.10
0,192 (0,141 x 1,36) à P.05
0,230 (0,141 x 1,63) à P.01
0,244 (0,141 x 1;73) à P.005
0,275 (0,141 x 1,95) à P.001

Les deux mesures ne sont pas indépendantes, comme l'exige le test K.S. (Kolmogorov-Smirnov), mais il ne s'agit pas ici de tester l'hypothèse nulle, et on ne tirera aucune conclusion au niveau individuel. Une valeur arbitraire de D max (la valeur .20) a servi à constituer deux groupes : les sujets "gardés" (D max < .20) et les sujets "rejetés" (D max ⩾ . 20).

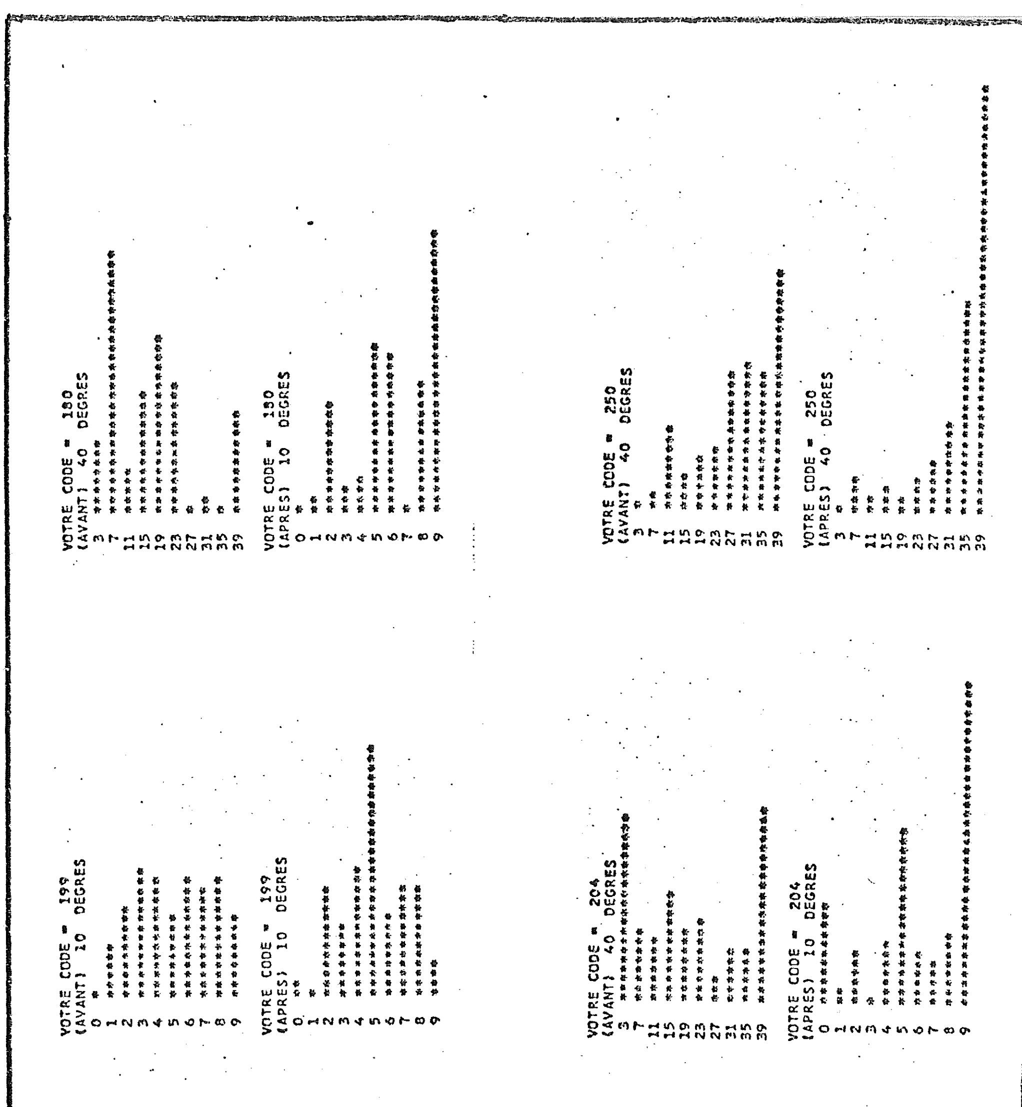
En annexe 6, on trouvera des histogrammes individuels. Une comparaison visuelle des histogrammes eut dû être guidée par quatre aspects : la similitude dans les deux histogrammes
- de la valeur modale
- de la répartition sur les divers niveaux
- de la symétrie ou de l'asymétrie (*skewness*)
- de l'aplatissement ou du leptocurtisme (*curtosis*).

Au test de Kolmogorov-Smirnov , 55 sujets seulement ont été "retenus". Ils se répartissent comme suit :

		Retest				
		A	B	C	T	
Test	A (4 degrés)	8	0	0	7	15
	B (10 degrés)	0	3	0	2	5
	C (40 degrés)	0	1	5	5	11
	T (A+B+C)	6	8	3	7	24
		14	12	8	23	55

Le tableau 6.9 présente quatre séries de résultats individuels rejetés après le test de Kolmogorov-Smirnov .

Tableau 6.9



Exemples (4 sujets) d'instabilité globale entre le test et le retest. Les histogrammes "avant" et "après" sont différents par un aspect important

On constate que, dans chacun de ces cas, un caractère important varie dans les deux distributions :

Sujet 199 : le Mode diffère;

Sujet 204 : la *Skewness* diffère;

Sujet 250 : la *Curtosis* diffère;

Sujet 180 : le Mode et la *Skewness* diffèrent, même *Curtosis*.

Retenir 55 sujets sur 124 équivaut à perdre 60 % de la population expérimentale. Cette mesure paraît inévitable pour deux raisons. Tout d'abord, la situation expérimentale n'a été operante que pour un certain nombre de sujets : ceux pour qui obtenir un total élevé de points a constitué (au test et au retest) un renforcement important. La solution élégante eût été d'exprimer les points en argent et de rétribuer les sujets au prorata du total obtenu. Encore eût-il fallu manipuler des sommes assez élevées pour que leur octroi constitue, pour la plupart des joueurs, a *satisfaying state of affairs*. Il nous a été impossible de réunir des fonds suffisants. Mais, même si nous avions pu en disposer, la situation ne serait pas devenue plus fondamentalement operante, puisqu'avant le retest, seulement deux séances de réponses suivies de renforcements auraient eu lieu. Il aurait fallu, en plus, entraîner chaque sujet par une procédure séquentielle qui consiste à fournir au sujet la solution correcte après chacune de ses réponses. Cela n'a pas été fait dans le cadre de ce travail. En effet, une telle procédure exigerait, non seulement des paiements réels, mais une application individuelle ou des procédures automatiques (coûteuses) d'enregistrement des réponses et des gains.

Nous avons considéré qu'au lieu de l'étude en situation operante, une étude sur la stabilité en situation non-operante apporterait des éléments neufs, même si une importante part de la population expérimentale doit être abandonnée. Les expériences idéales, on s'en doute, consisteraient à entraîner des sujets au moyen d'un terminal d'ordinateur qui, en même temps que les questions et les solutions correctes, afficherait les gains et les pertes.

C. LES HISTOGRAMMES DE REPLICATION

Les principes de construction des histogrammes de replication d'un degré de certitude, entre le test et le retest, sont les suivants.

1. Pour un degré de certitude donné lors du test, on établit le nombre de fois que chaque degré de l'échelle a été utilisé par le même sujet lors du retest. On dispose ainsi d'un "vecteur de replication". Considérons, par exemple, les certitudes fournies par le sujet 262 (tableau 6.10).

<u>Tableau 6.10</u>

VOTRE CODE = 262

	BR	MR	0	1	2	3	4	5	6	7	8	9
0	0	0										
1	2	3	1	2	1		1					
2	0	5			2	1	2					
3	1	4	1		3		1					
4	5	10		1	2	3	2	2			1	4
5	4	6				1	2	3	1	1		2
6	9	2	1	1		1	1	2		3		2
7	7	3					1		1	3		5
8	3	1							1			3
9	32	3					1	3		1		30
BR =			1	1	2	1	7	5	2	6	0	38
MR =			2	3	6	5	4	5	1	2	1	8

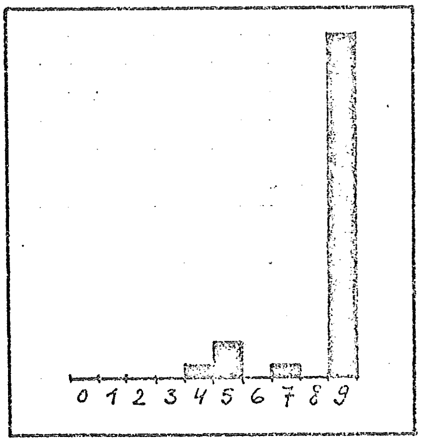
VOTRE CODE = 199

	BR	MR	0	1	2	3	4	5	6	7	8	9
0	1	0	1									
1	0	6			4	1	1					
2	5	5	1		5	2	1	1				
3	5	9			2		5	5	2			
4	6	7				2	4	3	1	1	2	
5	6	3				1	2	5			1	
6	11	2				1		5	2	4		1
7	9	3			1	1		3	3	1	3	
8	9	4					1	3		6	3	
9	8	1						2	1		3	3
BR =			1	0	3	3	4	19	8	9	9	4
MR =			1	0	9	5	10	8	1	3	3	0

Exemples (sujets 262 et 199) des nuages représentant la relation entre les certitudes lors du test et lors du retest

— Le sujet 262 a fourni, au test, 35 fois la certitude 9.
Au retest, le sujet 262 a attribué à ces 35 réponses (voir vecteur 9 horizontal) une fois la certitude 4, trois fois la certitude 5, une fois la certitude 7 et trente fois la certitude 9 (vecteur de replication).
L'histogramme de replication du degré 9 est donc le suivant (tableau 6.11).

Tableau 6.11

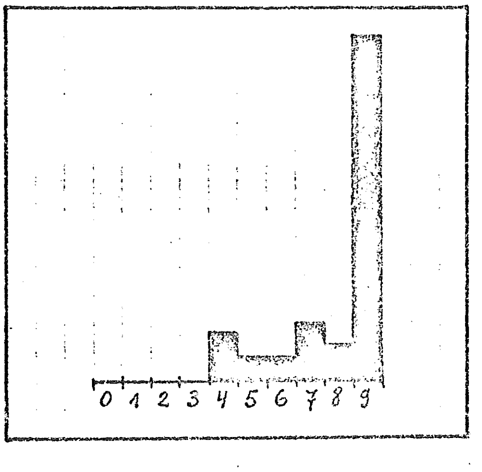


Histogramme de replication lors du retest des degrés de certitude 9 exprimés lors du test par le sujet 262

L'histogramme idéal serait un rectangle haut de 35 unités dont la base occuperait le degré 9. Ce serait l'indice d'une replication parfaite.

2. Il est possible de dessiner, pour chaque degré de l'échelle, un histogramme de replication en partant cette fois de la certitude fournie au retest.
Exemple : Le sujet 262 a fourni 46 fois la certitude 9 lors du retest. Son histogramme de replication (lors du test) de la certitude 9 (voir vecteur 9 vertical) figure au tableau 6.12.

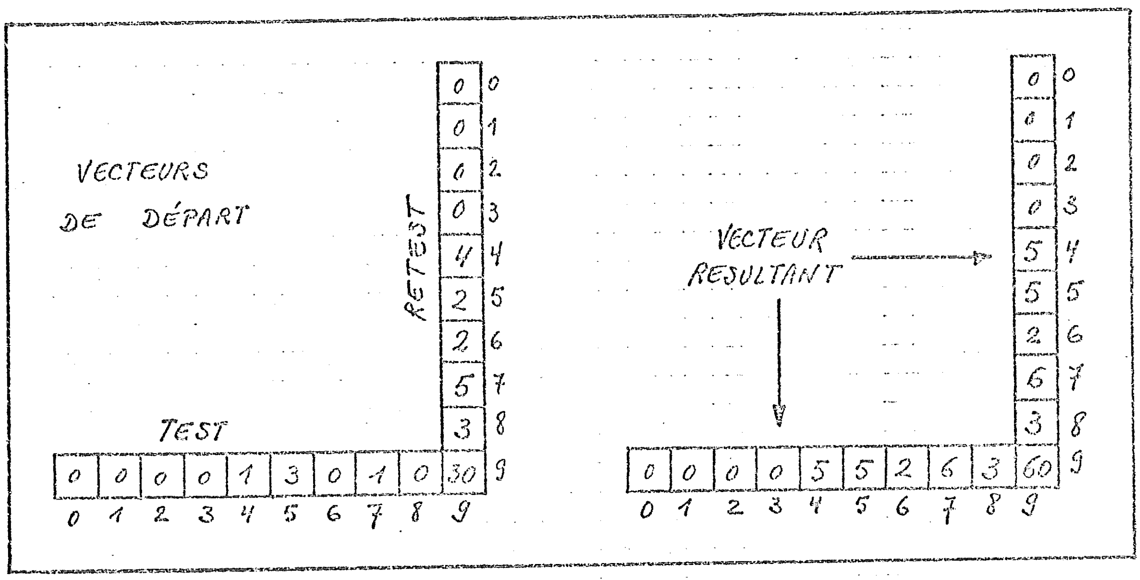
Tableau 6.12



Histogramme de replication lors du retest des degrés de certitude 9 exprimés lors du retest par le sujet 262

3. On peut combiner les deux vecteurs de replication d'un même sujet quand le nombre de degrés de certitude est le même dans les deux échelles.
   Ainsi, pour le sujet 262, on procédera à la somme des deux vecteurs (tableau 6.13).

Tableau 6.13



Somme des deux vecteurs d'un même degré de certitude
pour un même sujet (degré 9 pour le sujet 262)

Dans le tableau 6.13, on remarque que la matrice résultante (à droite) est symétrique. . Il est normal de trouver la valeur 60 (et non 30) dans le vecteur résultant, puisqu'il est présent dans les deux vecteurs de départ.

Cette procédure est, nous en convenons, quelque peu audacieuse, mais il importe de rentabiliser au maximum les données disponibles. C'est dans ce même souci que le quatrième principe a été développé.

4. On peut additionner les vecteurs de replication d'un même degré de certitude, pour différents sujets.

Ainsi, le vecteur résultant du sujet 199, pour la certitude 9, est :

0	0	0	0	0	2	2	0	3	6
0	1	2	3	4	5	6	7	8	9

La somme des vecteurs résultants des sujets 199 et 262, pour la certitude 9, est donc :

0	0	0	0	5	7	4	6	6	66
0	1	2	3	4	5	6	7	8	9

Total = 94

Il serait abusif, cependant, de considérer les valeurs absolues comme indices de replication. Ainsi, 66 replications de 9 en 9 (sur un total de 94) semblent un résultat très satisfaisant. Or il faut modérer ces observations par le taux d'utilisation de chaque indice, et ces taux varient fortement, malgré les précautions prises au niveau de la mutilation des textes.

Ainsi, sur 7000 indices de certitudes fournis (voir ci-après) par 35 sujets, on note les pourcentages d'utilisation suivants :

0	1	2	3	4	5	6	7	8	9
6,6	6,32	7,7	10,24	5,75	11,41	12,21	9,00	7,81	22,92

La certitude 9 est utilisée quatre fois plus que la certitude 4 et deux fois plus que n'importe quel autre degré. On voit que l'on ne peut comparer simplement les répétitions dans les histogrammes de replication. C'est ce qui motive le cinquième principe ci-après.

5. L'histogramme de replication d'un degré de certitude ne se dresse pas à partir des répétitions de chaque cellule du vecteur de replication, mais à partir des pourcentages. Ces pourcentages sont obtenus par division des répétitions par le nombre total des utilisations du degré de certitude correspondant.

En guise d'exemple, le tableau 6.14 présente les pourcentages calculés (matrice n° 2) à partir d'une matrice de replication symétrisée (matrice n° 1) sur les données de 40 sujets.

Tableau 6.14

	0	1	2	3	Total
0	714	355	211	67	1347
1	355	738	567	214	1874
2	211	567	934	635	2347
3	67	214	635	1516	2432
					8000

Matrice n° 1

	0	1	2	3	
0	53,0	26,3	15,6	4,9	100
1	18,9	39,3	30,2	11,4	100
2	8,9	24,1	39,7	27,0	100
3	2,7	8,7	26,1	62,3	100

Matrice n° 2

Matrice symétrisée et pourcentages horizontaux (2) sur les données de 40 sujets (1)

On remarque que chaque cellule de la matrice 1 a été divisée par le total de ligne . Désormais, les valeurs d'un même vecteur colonne peuvent être comparées entre elles. Elles n'ont plus, bien entendu, qu'une signification relative.

Bien qu'il s'agisse, dans le principe, d'histogrammes de replication, seule la ligne brisée reliant les "sommets" des rectangles sera représentée. Le but d'une telle représentation est de faciliter la lecture et de superposer deux histogrammes de replication sur un seul graphique.

D. LES HYPOTHESES DE L'ETUDE

- *Hypothèses relatives à la fidélité (stabilité) de l'ESPER.*

Thèse F : La fidélité (dans le sens de "répétabilité", "stabilité") peut être estimée à partir de la corrélation des ESPER lors d'une procédure test-retest (plutôt qu'à partir d'indices de cohérence interne).

HF1. Ces corrélations sont positives et élevées.

HF2. Ces corrélations ne varient pas significativement selon la finesse des échelles utilisées.

HF3. Pour chaque degré, le mode de son "histogramme de replication" est lui-même (ou la catégorie qui lui correspond dans une autre échelle).

- *Hypothèses relatives à la sensibilité de l'ESPER.*

Thèse S : La sensibilité peut être estimée à partir des indices suivants :

- La fréquence d'utilisation d'un niveau de certitude donné (les fréquences trop faibles indiquant une précision jugée non nécessaire ou impossible par les sujets, donc un degré trop fin de l'échelle).
- Chaque niveau de certitude doit être distinct des niveaux voisins à deux points de vue :
  - Les pourcentages de réussite (voir validité) des divers niveaux doivent être ordonnés comme les niveaux eux-mêmes;
  - Les "histogrammes de replication" de deux niveaux voisins doivent être distincts par certains paramètres (Mode, Moyenne, *Skewness*).

HS1. Le nombre de degrés de certitude distincts est supérieur à quatre (l'échelle la plus grossière); il se rapproche de dix degrés (l'échelle intermédiaire). Cette hypothèse est vraisemblable, par référence, par exemple au *magical number seven, plus or minus two* de MILLER (1956).

HS2. Certains degrés (dans les échelles fines) sont peu choisis, la préférence des sujets se marquant pour les degrés voisins (indice de superfluidité de ce degré de l'échelle).

HS3. Certains degrés voisins se confondent quant à leur validité et leur fidélité.

HS4. La sensibilité de la certitude n'est pas uniforme : elle est plus élevée aux extrêmes qu'aux valeurs intermédiaires (ESPER proche de 50 %).

E. PRESENTATION GENERALE DES DONNEES

Les consignes et les documents de l'expérience principale ont été adressées aux 258 personnes qui, sur les 313 contactées, ont répondu à la série d'essai. Quatre groupes expérimentaux ont été créés :

A = l'échelle à quatre degrés (62 personnes contactées)
B = l'échelle à dix degrés (61 personnes contactées)
C = l'échelle à quarante degrés (60 personnes contactées)
T = les trois échelles à la fois (75 personnes contactées).

Dans les consignes, aucune allusion n'est faite à la replication (test-retest) pour éviter que les sujets s'y préparent d'une quelconque façon.

Les sujets contactés ont répondu comme suit :

A = échelle à quatre degrés :	34 réponses
B = échelle à dix degrés :	29 réponses
C = échelle à quarante degrés :	32 réponses
T = les trois échelles à la fois :	54 réponses
Total	149 réponses

Voici les nombres de bonnes réponses (et leur transformation en pourcentages arrondis) pour chacune des cent questions (tableau 6.15).

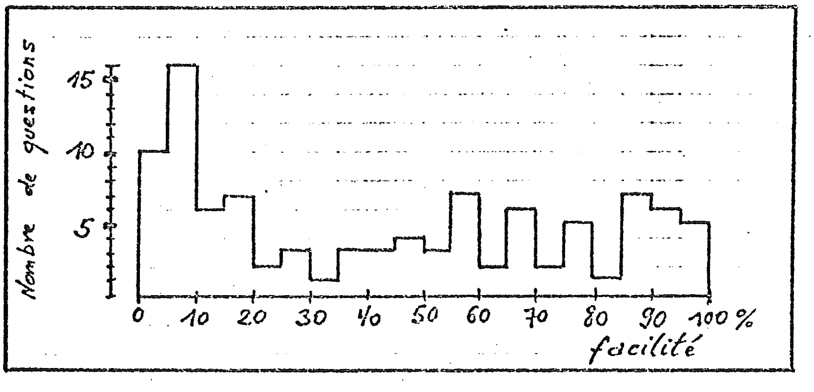
Tableau 6.15

Q	NBR	%	Q	NBR	%	Q	NBR	%	Q	NBR	%
1	64	51.	26	122	98.	51	49	39.	76	121	97.
2	35	28.	27	120	96.	52	113	91.	77	83	66.
3	3	2.	28	24	19.	53	93	75.	78	12	9.
4	112	90.	29	54	43.	54	9	7.	79	5	4.
5	6	4.	30	117	94.	55	13	10.	80	89	71.
6	11	8.	31	71	57.	56	26	20.	81	112	90.
7	62	50.	32	34	27.	57	96	77.	82	82	66.
8	27	21.	33	123	99.	58	3	2.	83	116	93.
9	68	54.	34	100	80.	59	110	88.	84	9	7.
10	8	6.	35	10	8.	60	36	29.	85	71	57.
11	15	12.	36	14	11.	61	5	4.	86	10	8.
12	70	56.	37	60	48.	62	71	57.	87	10	8.
13	116	93.	38	111	89.	63	75	60.	88	75	60.
14	14	11.	39	87	70.	64	96	77.	89	59	47.
15	79	63.	40	3	2.	65	71	57.	90	19	15.
16	10	8.	41	107	86.	66	67	54.	91	3	2.
17	63	50.	42	113	91.	67	48	38.	92	10	8.
18	10	8.	43	100	80.	68	120	96.	93	82	66.
19	104	83.	44	2	1.	69	15	12.	94	50	40.
20	9	7.	45	5	4.	70	83	66.	95	24	19.
21	12	9.	46	114	91.	71	87	70.	96	2	1.
22	81	65.	47	57	45.	72	18	14.	97	24	19.
23	61	49.	48	9	7.	73	22	17.	98	111	89.
24	8	6.	49	52	41.	74	41	33.	99	96	77.
25	109	87.	50	26	20.	75	26	20.	100	32	25.

Pourcentage de réussite (calculé sur 124 sujets) de chacune des 100 questions

La répartition des facilités est la suivante (tableau 6.16) :

Tableau 6.16



Répartition des facilités des cent questions

Les mesures prises lors de la mutilation du texte ne sont pas parvenues à rendre la distribution parfaitement horizontale. Un trop grand nombre de questions très difficiles (facilité $\leqslant$ 10 %) subsistent. Néanmoins, on peut considérer que l'étalement en facilité est satisfaisant.

Les consignes du retest ont été adressées à 149 personnes; 124 personnes ont répondu. Des échelles différentes ont été proposées à un même sujet lors du test et du retest. Seize situations ont donc été ainsi créées. Voici le nombre de réponses reçues pour chacune des seize cellules.

Tableau 6.17

		Groupe lors du re-test				
		A	B	C	T	
Groupe lors du test	A	8	6	8	8	30
	B	4	6	7	3	20
	C	6	5	11	6	28
	T	11	12	10	13	46
		29	29	36	30	124

Réponses reçues pour chacune des seize cellules de l'expérience.

Le tableau 6.18 fournit, pour chacun des 124 sujets :

1. Le nombre de réponses correctes;
2. Le total des points pour chaque échelle utilisée (échelle A = 4 degrés, échelle B = 10 degrés, échelle C = 40 degrés), au test et au retest;
3. Les corrélations entre les indices de certitude utilisés au test et les indices de certitude utilisés au retest.

Les sujets sont présentés dans l'ordre des groupes expérimentaux. Rappelons que T signifie que le sujet a dû répondre aux trois échelles A, B et C.

				AVANT			APRES			CORRELATIONS des ESPER								
		N°	BR	A	B	C	A	B	C	AA	AB	AC	BA	BB	BC	CA	CB	CC
AA	*	132	58	1797.5			1930.0			.52								
	*	178	62	2037.5			1875.0			.58								
	*	170	56	1985.0			1745.0			.56								
	*	251	43	1216.2			1146.2			.70								
	*	118	60	1830.0			1857.5			.35								
	*	185	58	1835.0			1677.0			.56								
	*	236	53	1918.7			1823.7			.73								
	*	022	54	1625.0			1630.0			.57								
BA		101	60		2072.5		1985.0						.61					
		200	60		2086.8		2075.0						.79					
		229	61		1977.0		1895.2						.52					
		160	60		1902.9		1952.5						.44					
CA		089	53			1482.6	1621.2									.50		
	O	140	47			291.9	1158.7									.07		
		163	57			1372.2	1928.7									.80		
		184	56			1621.9	1867.5									.65		
		296	54			1663.7	1575.0									.60		
		307	59			2039.3	1976.2									.57		
AB		036	59	2061.2				1839.4			.61							
		141	58	1785.0				1549.8			.38							
		173	53	1636.2				1669.5			.41							
		239	66	2187.5				2105.9			.73							
		192	66	2280.0				1968.4			.60							
		139	59	1838.7				1377.2			.57							
BB	*	262	63		2345.8			2108.4						.73				
		295	49		1496.6			1457.0						.58				
	*	199	60		2002.8			2039.6						.69				
		126	61		1966.9			1865.1						.60				
	*	128	53		1509.9			1689.3						.54				
		109	54		1644.7			1728.5						.69				
CB	*	252	56			1832.7		1931.4									.63	
		180	59			1806.6		1750.5									.47	
		204	63			1704.9		1823.7									.54	
		117	62			1886.6		1686.2									.52	
		300	55			1523.1		1458.1									.68	
AC		102	60	1927.5					1348.8			.56						
		148	54	1752.5					1245.1			.72						
		156	57	1823.7					1574.4			.67						
		176	55	1681.2					1838.9			.48						
		194	57	1883.7					1462.8			.47						
		198	45	1507.5					1346.7			.43						
		241	55	1786.2					1491.2			.48						
		302	61	2093.7					2048.2			.72						
BC		107	63		1917.9				2048.6						.48			
		131	52		1669.5				1298.7						.48			
		171	55		1663.5				1689.0						.63			
		197	55		1448.4				1055.2						.61			
		206	60		1891.4				1720.0						.55			
		258	57		1707.2				839.6						.52			
		306	59		1940.2				1995.3						.66			
CC	*	282	58			2038.3			1740.3									.51
		110	64			2147.3			2014.5									.23
	*	188	63			1838.1			1898.2									.66
		193	48			982.4												.47
		250	60			1867.4			1447.3									.45
	*	094	58			1841.6			1906.3									.71
	*	104	59			1722.4			1710.8									.59
		146	50			1032.4			474.3									.47
		210	55			1851.3			1796.0									.64
	*	308	58			2052.4			1954.1									.73
		127	57			1764.9			559.6									.46

		N°	BR	A	B	C	A	B	C	AA	AB	AC	BA	BB	BC	CA	CB	CC
TA	*	001	39	1283.7	1281.5	1310.7	1166.2			.59			.61			.63		
	*	010	59	2008.7	1983.4	1771.8	1703.7			.42			.44			.44		
	O	027	32	207.5	-210.8	-813.2	-107.5			.13			.10			.07		
		038	51	1608.7	1476.0	1312.6	1261.2			.48			.49			.43		
	*	050	59	2008.7	2015.1	2040.4	1993.7			.39			.43			.38		
	*	052	53	1636.2	1581.5	1539.3	1576.2			.70			.72			.71		
	*	069	53	1613.7	1462.1	1377.2	1675.2			.51			.54			.55		
	O	077	7	-741.2			-913.7			-.08			-.07			-.02		
		309	48	1265.0	1211.3	1175.2	1177.5			.55			.56			.34		
		061	64	2327.5	2298.6	2316.5	1985.0			.52			.54			.56		
	*	056	45	1383.7	1331.8	1237.5	1218.7			.44			.40			.38		
TB		003	55	1796.2	1751.2	1727.1		1618.3			.42			.46			.46	
	O	025	59	1721.2	1355.1	744.3		831.8			.19			.21			.20	
	*	047	57	1728.7	1733.9	1708.3		1678.2			.53			.50			.54	
		059	60	1867.5	1570.0	917.5		1617.7			.40			.43			.39	
	*	098	54	1645.0	1598.4	1443.9		1529.5			.43			.49			.53	
	*	125	59	1736.2	1581.7	1419.8		1570.8			.53			.54			.53	
	*	008	51	1508.7	1531.6	1402.1		1403.5			.37			.35			.34	
	*	006	52	1670.0	1631.0	1677.2		1558.8			.54			.58			.56	
	*	051	54	1755.0	1672.5	1686.2		1741.4			.62			.63			.62	
	*	060	57	1773.7	1649.2	1405.6		1462.0			.64			.63			.63	
	*	020	54	1715.0	1641.7	1567.0		1410.8			.62			.61			.65	
		030	63	1901.2	1800.0	1821.3		1825.1			.57			.54			.54	
TC	*	017	48	1518.7	1524.4	1501.6			1285.6			.61			.64			.63
		043	51	2013.7	1926.2	1873.9			2053.5			.52			.55			.56
	*	045	55	1793.7	1612.5	1468.2			1522.1			.39			.52			.53
		088	50	1582.5	1492.6	1364.1			1315.4			.45			.46			.45
		061	60	2065.0	1720.2	1658.5			1742.8			.53			.56			.56
		130	49	1483.7	1432.6	1364.9			1394.0			.63			.63			.63
	O	071	43	1051.2	848.8	547.0			962.3			.26			.26			.24
		058	47	1363.7	1285.1	1163.1			914.8			.35			.36			.35
	*	033	51	1295.0	1167.6	1140.9			1117.8			.36			.44			.45
		021	59	1936.2	1310.2	1833.7			1981.1			.69			.64			.63
AT		108	57	1941.2			1946.2	1918.0	1884.9	.59	.60	.61						
	*	179	61	1878.7			2078.7	1971.2	1939.0	.63	.60	.63						
	*	168	63	2013.7			2058.7	2057.5	2003.6	.63	.61	.60						
	*	212	60	2180.0			2170.0	2122.7	2067.0	.63	.66	.66						
		242	49	1453.7			1341.2	1351.3	1201.6	.58	.63	.65						
		150	63	2108.7			2013.7	1921.8	1725.1	.57	.60	.62						
	*	106	46	1295.0			1360.0	1264.9	1133.2	.59	.61	.60						
	*	218	52	1682.5			1682.5	1682.8	1741.2	.51	.53	.54						
BT	*	152	57		1536.7		1881.2	1731.1	1588.4				.46	.49	.56			
		253	57		1770.8		1883.7	1866.6	1873.8				.50	.53	.55			
	*	208	55		1337.2		1551.2	1317.7	1198.1				.52	.55	.54			
CT	*	155	53			1326.2	1666.2	1348.9	1272.0							.77	.74	.76
	*	121	62			2138.3	2130.0	2012.0	1806.3							.48	.51	.55
	*	203	54			1562.8	1665.0	1646.6	1736.6							.60	.61	.60
	*	149	60			1965.3	1880.0	1708.3	1584.1							.48	.53	.49
	*	211	44			1380.8	1322.5	1104.1	1001.2							.69	.69	.68
		283	58			1884.8	1842.5	1569.2	1110.5							.47	.48	.47
TT	*	055	56	1835.0	1773.6	1732.8	1773.0	1527.2	1512.3	.56	.54	.59	.52	.51	.52	.51	.51	.52
		133	63	2006.0	2117.0	2069.7	1901.2	1894.8	1823.8	.73	.10	.12	.45	.54	.56	.44	.58	.60
		011	57	1768.7	1697.6	1677.7	1973.7	1929.1	1938.8	.54	.56	.55	.57	.59	.58	.58	.60	.58
		044	57	1843.7	1755.9	1780.7	1883.7	1702.8	1598.6	.61	.60	.60	.63	.62	.62	.63	.62	.62
	*	057	59	1913.7	1973.8	2032.8	2088.7	2081.5	2076.2	.71	.67	.68	.81	.81	.81	.81	.81	.81
	*	019	59	2178.7	2207.4	2260.8	2168.7	2182.9	2177.1	.74	.75	.75	.73	.77	.76	.71	.76	.76
		035	52	1712.5	1733.6	1732.5	1635.5	1595.3	1545.1	.69	.70	.67	.69	.71	.70	.68	.70	.70
	*	018	56	1745.0	1848.6	1864.2	1692.5	1682.7	1682.5	.31	.24	.24	.41	.37	.38	.42	.39	.40
	*	048	46	1237.5	1063.9	674.0	1352.5	1310.4	1205.9	.53	.50	.51	.55	.51	.52	.55	.51	.52
	*	066	54	2157.5	2136.4	1999.8	2180.0	2058.4	1888.3	.66	.66	.67	.65	.66	.66	.65	.65	.66
	*	034	52	1672.5	1599.3	1534.0	1640.0	1500.8	1281.2	.52	.52	.52	.54	.53	.54	.52	.52	.53
		026	54	1642.5	1542.5	1582.1	1515.0	1155.9	794.7	.24	.27	.26	.22	.25	.15	.20	.23	.13
	O	076	12	-567.5		1518.4	-732.5			-.16	-.17	-.18	-.20	-.22	-.23	-.19	-.22	-.24

A ces 124 personnes ont été adressés, en plus de trois feuilles stencilées, leurs résultats individuels détaillés et un listing-synthèse de l'ensemble des résultats.

Pour encourager les sujets, un maximum d'informations devaient être données. Néanmoins, pour ne pas trop différer le *feedback*, certains résultats partiels et même seulement des hypothèses ont dû être communiqués. L'aspect partiel et hypothétique de certaines données a été signalé.

Tableau 6.19

Je tiens à vous remercier vivement pour l'effort que vous avez consenti; les 100 réponses constituent un travail particulièrement fastidieux.

En annexe, vous trouverez:

1. Le texte complet à deviner. Il était extrait de l'ouvrage de Jean FOURASTIER, Les 40.000 heures, Paris, Denoël, Gonthier, 1972, p.155.

2. Un listing portant le détail de vos résultats, réponse par réponse, vos totaux, votre nuage de points avant/après, la corrélation (r=) qui l'exprime mathématiquement et vos histogrammes de certitude avant et après.

3. Le tableau ordonné des points (seul, le meilleur total de chacun est pris en considération), qui indique aussi le nombre de bonnes réponses (B.R.) sur les 100 lettres.

4. Des commentaires sur les résultats généraux.

./.

Extraits de la lettre adressée aux 124 sujets qui ont participé à l'expérience jusqu'au bout.

./..

RESULTATS GENERAUX

Jeu des certitudes portant sur 100 lettres

Cent vingt quatre personnes ont participé à l'expérience complète . Pour 121 d'entre elles, le nombre de bonnes réponses (sur les 100 lettres à deviner) est compris entre 39 et 66 bonnes réponses, avec une moyenne de 56 bonnes réponses.

Afin que vous puissiez situer votre performance dans l'ensemble des résultats, voici la répartition des 121 résultats (1):

66	//
65	
64	///
63	////////
62	///
61	/////
60	////////////
59	/////////////
58	///////
57	////////////
56	/////
55	/////////
54	/////////
53	///////
52	/////
51	///
50	//
49	///
48	///
47	//
46	//
45	//
44	/
43	//
42	
41	
40	
39	/

Les corrélations entre les certitudes exprimées avant et après vont de .80 à .20, avec une valeur médiane de .56. Ces corrélations varient selon le nombre de lettres considérées. Répondre à cent lettres constitue un très long travail. Si on ne considère que les toutes premières lettres, la corrélation est souvent meilleure. Dans une étude faite sur 17 personnes, la corrélation médiane est .78, si on ne considère que les 5 premières certitudes, et de .74, si on ne considère que les dix premières, alors que pour les 100 lettres, la corrélation médiane est de .61.

§ § §

(1) Trois sujets ont eu respectivement 7, 12 et 32 bonnes réponses. Ces résultats très faibles et discordants dans l'ensemble ont été ignorés.

./..

./..

La fidélité globale avant/après est très bonne pour certains (marqués du signe F dans le tableau des points). Cela signifie qu'ils ont réparti leurs certitudes selon la même distribution avant et après.

Sur les 121 sujets considérés, 78 ont présenté un meilleur total de points avant et 42 un meilleur total après. Un seul a eu un total identique.

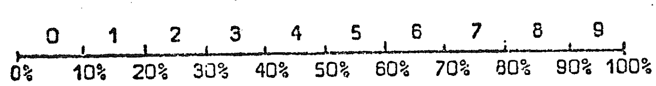
§ § §

En ce qui concerne la finesse des échelles, il apparaît que l'échelle à 40 degrés est trop fine. Vingt degrés sont amplement suffisants. La majorité des sujets préfèrent l'échelle à dix degrés.

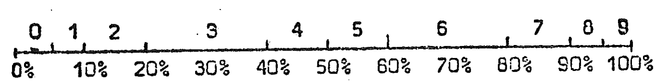
Quant à l'échelle à dix degrés, les premières analyses semblent indiquer que la sensibilité est plus forte aux extrêmes qu'au milieu. Cela signifie qu'on sait établir la différence entre 1 chance sur 100 et 5 chances sur 100, alors qu'on ne sait pas différencier 31 chances sur 100 de 35 chances sur 100.

Autour des 50% cependant, il existerait une autre zone de précision.

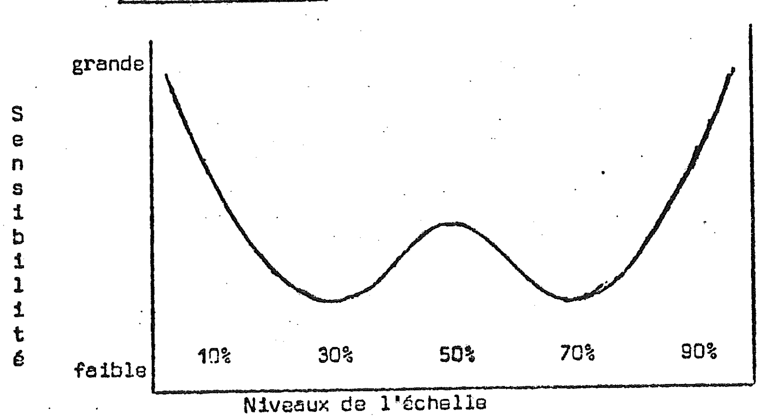
On pourrait illustrer ces hypothèses en disant que la précision n'est pas la même partout, donc que l'échelle à proposer ne doit pas être :



mais



Ceci supposerait que la sensibilité de la certitude a l'allure générale suivante :



./..

F. VERIFICATION DES HYPOTHESES

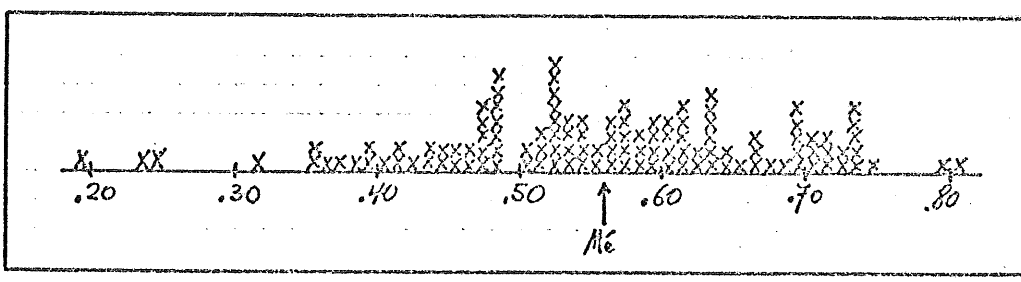
Hypothèse H F 1 :

La corrélation des ESPER au test et des ESPER au retest est positive et élevée, chez chaque sujet.

On trouvera, en annexe 7, les matrices, les histogrammes et les corrélations de chaque sujet et, en annexe 6, les données brutes, question par question.

Voici, pour les 121 sujets retenus (tableau 6.20), la distribution des corrélations entre l'ESPER au test et l'ESPER au retest. Pour chaque sujet, on a noté la corrélation entre la plus petite échelle au test et la plus petite échelle au retest.

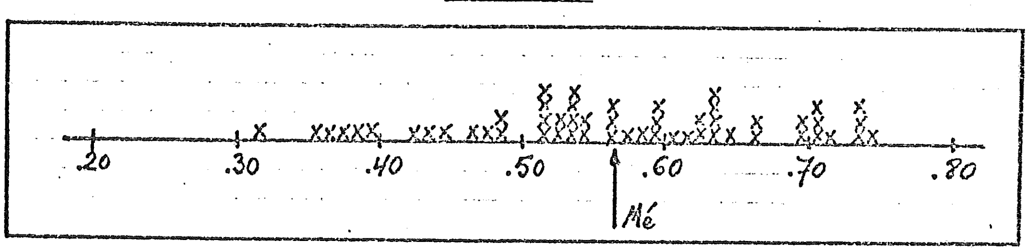
Tableau 6.20



Distribution des corrélations individuelles lors du test et du retest (la plus petite échelle a chaque fois été considérée) pour 121 sujets

L'hypothèse HF1 est donc confirmée. Si l'on ne considère que les 55 sujets retenus, la distribution est la suivante (tableau 6.21).

Tableau 6.21



Distribution des corrélations individuelles lors du test et du retest pour les 55 sujets retenus

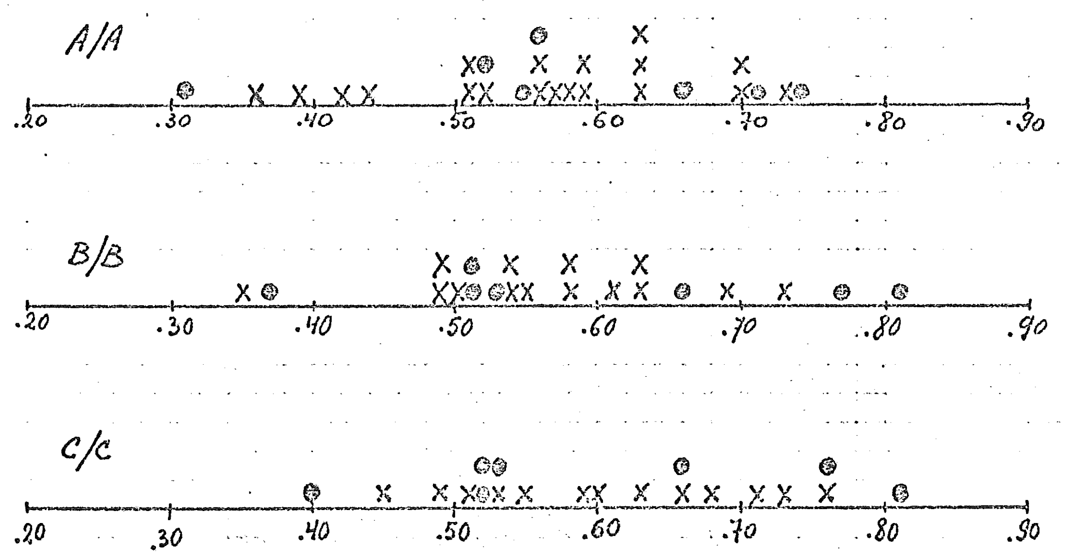
Le Médian vaut toujours .56, mais les corrélations inférieures à .32 ont disparu, ainsi que les corrélations supérieures à .75.

Hypothèse HF 2 :

Les corrélations ESPER au test - ESPER au retest ne varient pas significativement selon la finesse des échelles utilisées.

Le tableau 6.22 présente la distribution des corrélations (A/A = r entre test A et retest A; B/B = r entre test B et retest B; etc.), pour les trois types d'échelles. Seuls les sujets retenus au test de Kolmogorov-Smirnov sont pris en considération. Les sujets différents d'une situation à l'autre sont représentés par des croix. Les sept sujets représentés par des points font partie du groupe T et sont donc repris dans les trois situations.

Tableau 6.22



	N	Mé	E.S.I.	M	S
A/A	26	.56	.06	.5634	.112
B/B	21	.55	.07	.5747	.114
C/C	20	.59	.09	.5715	.116

Distribution des corrélations entre le test et le retest pour trois situations (A/A; B/B; C/C).

Les valeurs moyennes et médianes sont très proches, ce qui confirme l'hypothèse HF2.

<u>Hypothèse HF 3</u> :

Examinons maintenant, échelle par échelle, la replication des résultats au moyen de matrices globales de replication et d'histogrammes de replication de chaque degré de certitude.

L'expression "Situation 4/4 désignera la situation expérimentale dans laquelle l'échelle à quatre degrés de certitude a été utilisée au test et au retest. L'expression "Situation 4/10" désignera <u>les</u> situations expérimentales où l'échelle à quatre degrés a été utilisée <u>lors de l'un des deux tests</u> et l'échelle à dix degrés <u>lors de l'autre test</u>.

On trouvera, en annexe 7, les matrices brutes de replication.

SITUATION 4/4

Le tableau 6.23 présente la matrice de replication symétrisée (à gauche) et la matrice des pourcentages de ligne (à droite), ainsi que les μ (pourcentages de réponses correctes par indice de certitude). Ces valeurs ont été calculées à partir des réponses de 40 sujets et des réponses des 26 d'entre eux qui ont été acceptés par le test de Kolmogorov-Smirnov.

Tableau 6.23

40 SUJETS

	0	1	2	3	ToT	B.R.
0	714	355	211	67	1347	429
1	355	738	567	214	1874	775
2	211	567	934	635	2347	1052
3	67	214	635	1516	2432	1436
					8000	

	0	1	2	3	ToT	B.R.
0	53,0	26,3	15,6	4,9	100	31,8
1	18,9	39,3	30,2	11,4	100	41,3
2	8,9	24,1	39,7	27,0	100	44,8
3	2,7	8,7	26,1	62,3	100	59,0

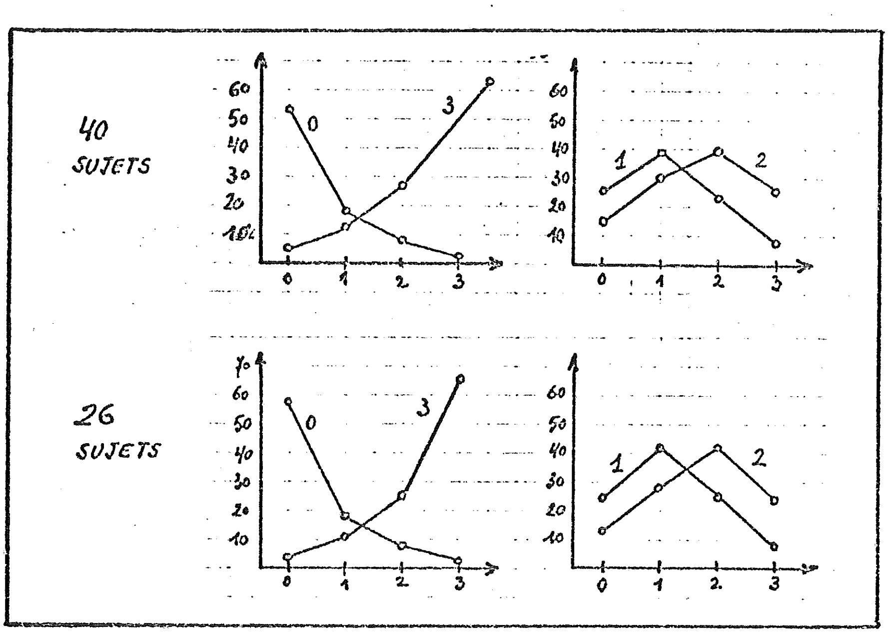
26 SUJETS

	0	1	2	3	ToT	BR
0	546	235	117	39	937	295
1	235	514	364	128	1241	526
2	117	364	590	367	1438	814
3	39	128	367	1050	1584	1217
					5200	

	0	1	2	3	ToT	BR
0	58,2	25,0	12,4	4,1	100	31,4
1	18,9	41,4	29,3	10,3	100	42,3
2	8,1	25,3	41,0	25,5	100	56,6
3	2,4	8,0	23,1	66,2	100	76,8

Les histogrammes de replication des degrés symétriquement opposés (0 et 3, 1 et 2) ont été dessinés sur un même graphique (voir tableau 6.24) pour vérifier visuellement cette symétrie.

Tableau 6.24

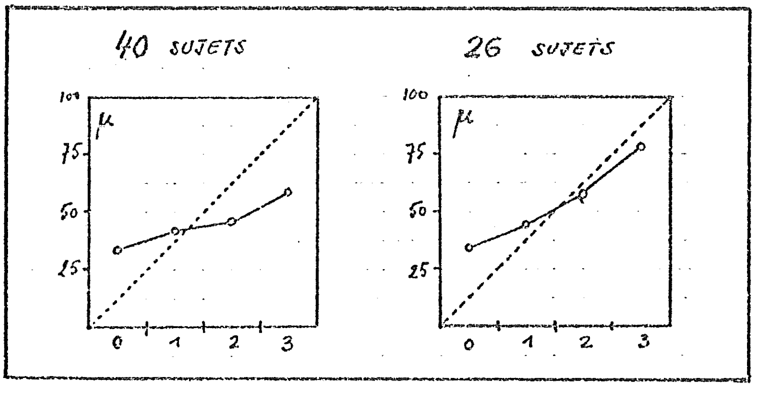


L'hypothèse H F 3 (Pour chaque degré de certitude, le Mode de son histogramme de replication est ce degré lui-même) se vérifie (tableau 6.24) pour chacun des quatre degrés (0, 1, 2 et 3) considérés.

Au seul examen visuel, la *skewness* des quatre courbes est conforme aux attentes. Par ailleurs, les "histogrammes" de replication des degrés 1 et 2 sont plus platticurtiques avec les 40 sujets qu'avec les 26 sujets retenus.

Le tableau 6.25 montre les fonctions de probabilité subjective (valeurs des différents μ portées sur graphique).

Tableau 6.25



On constate que cette fonction de probabilité subjective est une ligne brisée proche de la droite. La pente de la droite est nettement plus accusée (et plus proche de 1) dans le cas des 26 sujets "gardés" que dans le cas de l'ensemble des 40 sujets.

L'équation de régression de μ sur C peut être grossièrement évaluée, quand on a ramené les unités de l'abscisse à celles de l'ordonnée : 0 = .125, 1 = .375, 2 = .625, 3 = .875. L'équation prend alors les valeurs (approximatives) :

= 0,5 C + .15 (26 sujets)
= 0,3 C + .30 (40 sujets).

Les coefficients, estimés visuellement, sont nettement différents des valeurs idéales. Il y a une tendance à sous-estimer le succès des réponses à certitude faible et à les surestimer dans les certitudes fortes. On rejoint ici les observations de COOMBS, DAWES et TVERSKY (1970, p.147). La corrélation est, manifestement, très élevée (rho = 1) et confirme l'hypothèse H4 de validité (voir chapitres 2 et 4).

SITUATION 4/10

Dans les situations 4/10, les matrices de replication sont rectangulaires; il n'est donc pas question de les symétriser.

Le tableau 6.26 présente les matrices de replication obtenues à partir des réponses de 70 sujets et à partir des réponses des 35 sujets gardés lors du test de Kolmogorov-Smirnov.

Tableau 6.26

	0	1	2	3	4	5	6	7	8	9		BR	% BR
0	256	188	169	151	69	93	81	40	22	42	1111	372	33,4
1	119	123	201	264	139	212	215	123	66	128	1590	660	41,5
2	53	94	124	219	126	318	339	251	192	321	2037	1131	55,5
3	34	38	45	83	69	176	220	216	267	1114	2262	1618	71,52
	462	443	539	717	403	799	855	630	547	1605	7000		

	0	1	2	3	4	5	6	7	8	9			
0	147	106	90	80	21	35	30	15	9	19	552	169	30,6
1	44	63	105	171	75	105	110	55	38	41	807	328	40,6
2	31	47	52	127	61	162	197	132	112	130	1051	606	57,6
3	11	18	20	46	28	66	112	97	140	552	1090	821	75,3
	233	234	267	424	185	368	449	299	299	742	3500		

Le tableau 6.27 présente les pourcentages en ligne des matrices du tableau 6.26.

Tableau 6.27

0	1	2	3	4	5	6	7	8	9	
23,0	16,9	15,2	13,6	6,2	8,4	7,3	3,6	2,0	3,8	100
7,5	7,7	12,6	16,6	8,7	13,3	13,5	7,8	4,2	8,1	100
2,6	4,6	6,1	10,8	6,2	15,6	16,6	12,3	9,4	15,8	100
1,5	1,7	1,9	3,7	3,1	7,9	9,7	9,5	11,8	49,2	100
0	1	2	3	4	5	6	7	8	9	
26,6	19,2	16,3	14,5	3,8	6,4	5,5	2,7	1,6	3,4	100
5,5	7,8	13,0	21,2	9,3	13,0	13,6	6,8	4,7	5,1	100
2,9	4,5	4,9	12,1	5,8	15,4	18,7	12,6	10,7	12,4	100
1,0	1,7	1,8	4,2	2,6	6,1	10,3	8,9	12,8	50,6	100

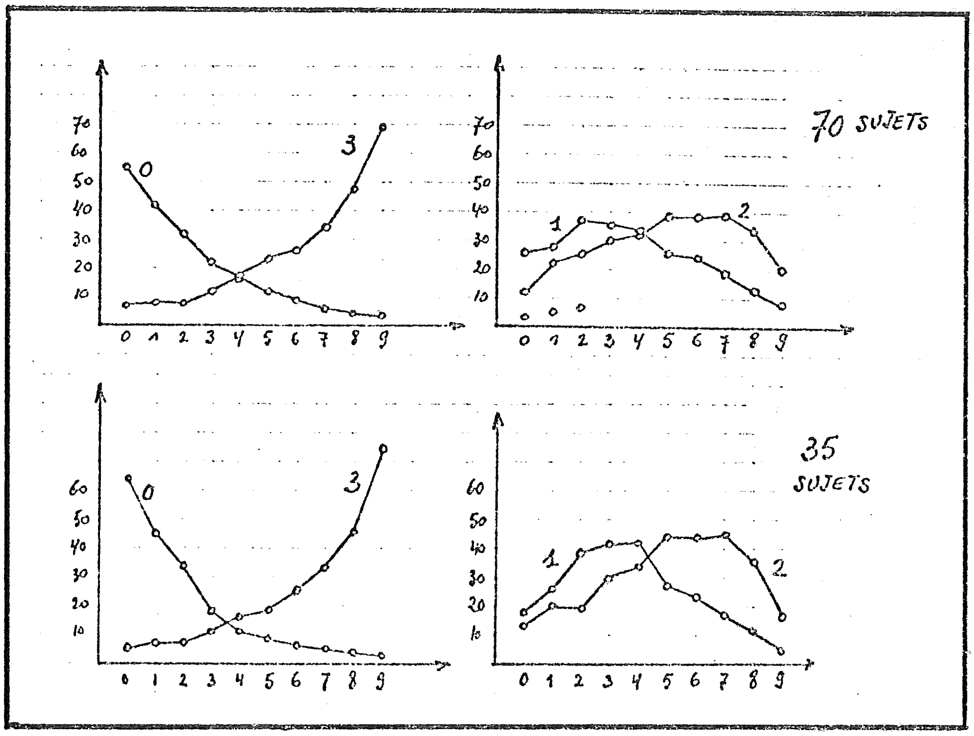
Le tableau 6.28 présente les pourcentages en colonne des matrices du tableau 6.26.

Tableau 6.28

0	1	2	3	4	5	6	7	8	9
55,4	42,4	31,4	21,1	17,1	11,6	9,5	6,4	4,0	2,6
25,8	27,8	37,3	36,8	34,5	26,5	25,2	19,5	12,1	8,0
11,5	21,2	23,0	30,5	31,3	39,8	39,6	39,8	35,1	20,0
7,3	8,6	8,3	11,6	17,1	22,1	25,7	34,3	48,8	69,4
100	100	100	100	100	100	100	100	100	100
0	1	2	3	4	5	6	7	8	9
63,1	45,3	33,7	18,9	11,4	9,5	6,7	5,0	3,0	2,6
18,9	26,9	39,3	40,3	40,5	28,6	24,5	18,4	12,7	5,5
13,3	20,1	19,5	30,0	33,0	44,0	43,9	44,2	37,5	17,5
4,7	7,7	7,5	10,8	15,1	17,9	24,9	32,4	46,8	74,4
100	100	100	100	100	100	100	100	100	100

Les histogrammes de replication du tableau 6.29 ont été tracés à partir des matrices du tableau 6.28. L'hypothèse HF3 se confirme à nouveau et le caractère leptocurtique des "histogrammes" des sujets gardés (après le test K.S.) apparaît de nouveau.

Tableau 6.29



Le tableau 6.30 fournit les fonctions de probabilité subjective pour l'échelle à 4 degrés et pour l'échelle à 10 degrés.

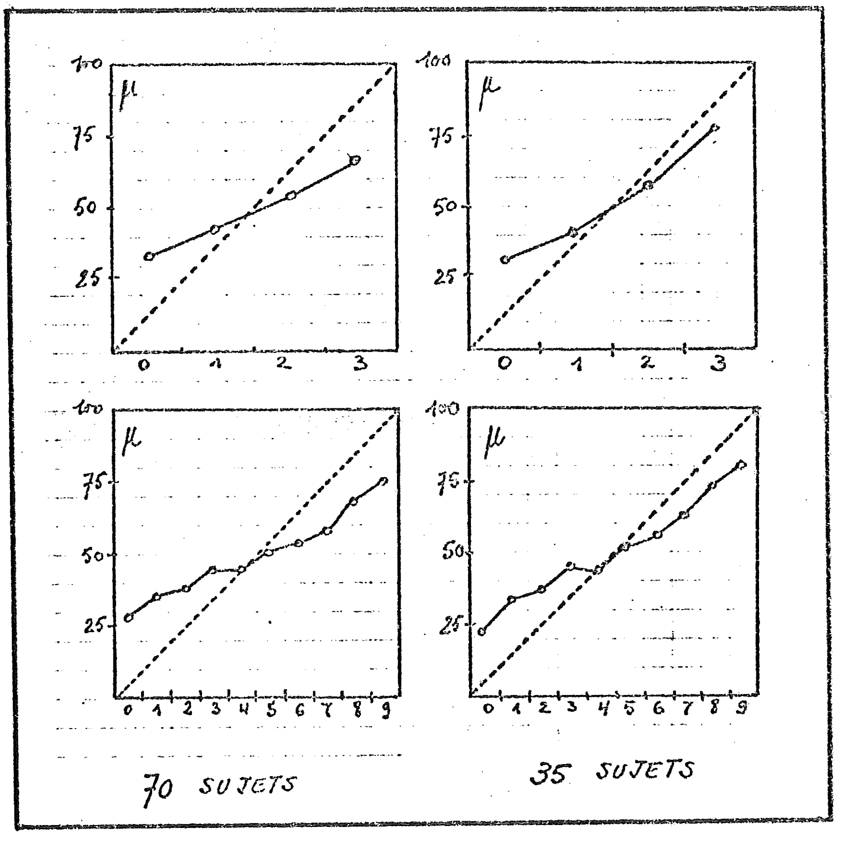
Tableau 6.30

	0	1	2	3	4	5	6	7	8	9	
	462	443	539	717	403	799	855	630	547	1605	7000
BR	123	161	209	319	174	401	447	370	377	1200	
% BR	26,6	36,3	38,7	44,4	43,1	50,1	52,2	58,7	68,9	74,7	
	233	234	267	424	185	368	449	299	299	742	3500
BR	52	77	98	189	79	186	247	185	218	593	
% BR	22,3	32,9	36,7	44,5	42,7	50,5	55,0	61,8	72,9	79,9	

		BR	% BR		BR	% BR
0	1111	372	33,4	552	169	30,6
1	1590	660	41,5	807	328	40,6
2	2037	1131	55,5	1051	606	57,6
3	2262	1618	71,52	1090	821	75,3
	7000			3500		

Le tableau 6.31 est la traduction graphique du tableau 6.30. On constate ici aussi la linéarité de ces fonctions dont la pente est proche de 1 pour les 35 sujets. On constateque, pour l'échelle à 4 degrés, le rho vaudrait 1, alors que, pour l'échelle à 10 degrés, le rho serait légèrement inférieur à 1 (interversion des μ du degré 3 et du degré 4).

Tableau 6.31



SITUATIONS 10/10

Dans les situations 10/10, les matrices de replication étant carrées, elles ont pu être symétrisées. On constatera que le total de certitudes exprimées par les 34 sujets n'est pas 3400 comme prévu, mais 3398 (6796/2). Les deux certitudes manquantes ont pour origine des erreurs de perforation.

Le tableau 6.32 présente les matrices de replication et les (pourcentages de réponses correctes par degré de certitude) obtenues à partir des réponses de 34 sujets et à partir des réponses de 20 sujets (gardés au test K.S.).

Tableau 6.32

	0	1	2	3	4	5	6	7	8	9			% BR
0	160	57	50	49	20	29	34	7	7	23	436	133	30,5
1	57	102	74	68	38	37	36	17	12	19	460	152	33,0
2	50	74	88	84	54	55	59	33	11	27	535	214	40,0
3	49	68	84	102	79	70	91	45	28	38	654	290	44,3
4	20	38	54	79	56	55	53	42	31	23	451	191	42,3
5	29	37	55	70	55	126	110	81	45	97	705	371	52,6
6	34	36	59	91	53	110	152	97	75	133	840	444	52,8
7	7	17	33	45	42	81	97	70	73	99	564	341	60,4
8	7	12	11	28	31	45	75	73	94	157	533	369	69,2
9	23	19	27	38	23	97	133	99	157	1002	1618	1237	76,4
	436	460	535	654	451	705	840	564	533	1618	6796		

	0	1	2	3	4	5	6	7	8	9			% BR
0	50	27	17	11	10	10	12	2	1	11	151	35	23,1
1	27	72	46	40	23	18	18	8	6	12	270	78	28,8
2	17	46	62	62	34	21	26	21	7	12	308	109	35,3
3	11	40	62	62	52	46	54	28	18	23	396	162	40,9
4	10	23	34	52	36	36	31	25	23	15	285	118	4,4
5	10	18	21	46	36	92	66	54	31	57	431	237	54,9
6	12	18	26	54	31	66	102	72	53	70	504	284	56,3
7	2	8	21	28	25	54	72	42	47	45	344	214	62,2
8	1	6	7	18	23	31	53	47	76	96	358	247	68,9
9	11	12	12	23	15	57	70	45	96	610	951	752	79,0
	151	270	308	396	285	431	584	344	358	951	3998		

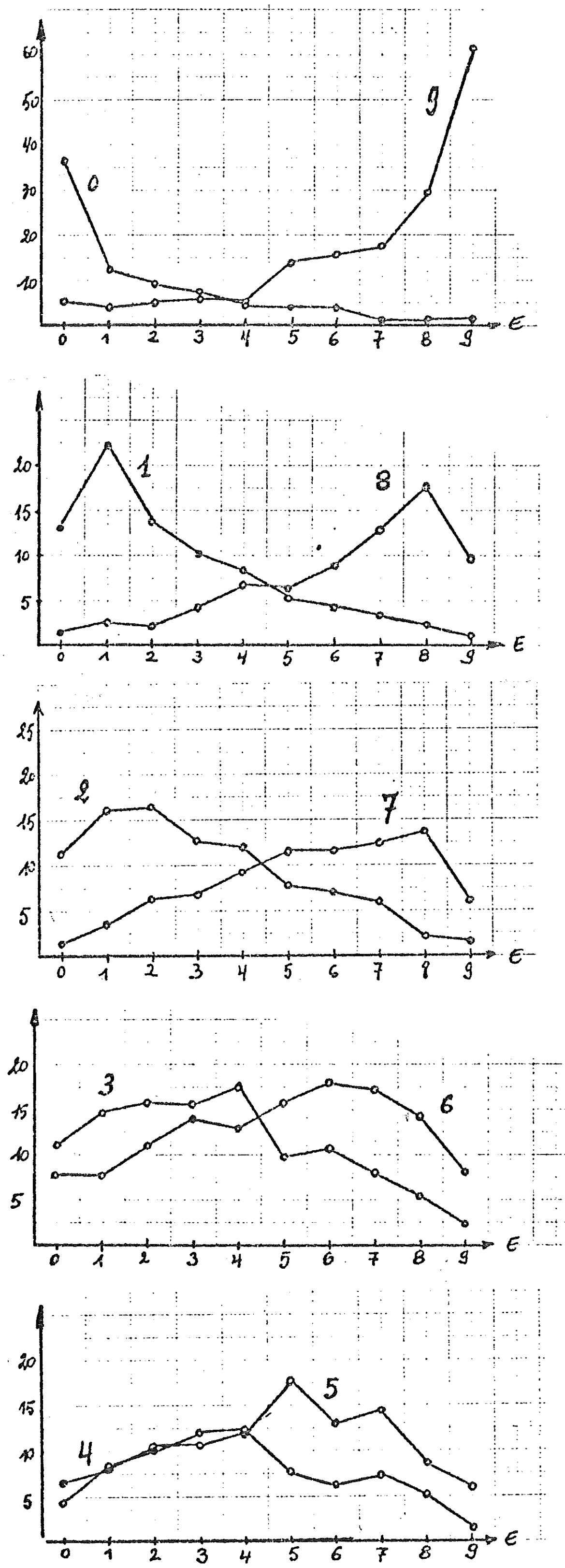
Le tableau 6.33 présente les pourcentages calculés sur les matrices du tableau 6.32.

Tableau 6.33

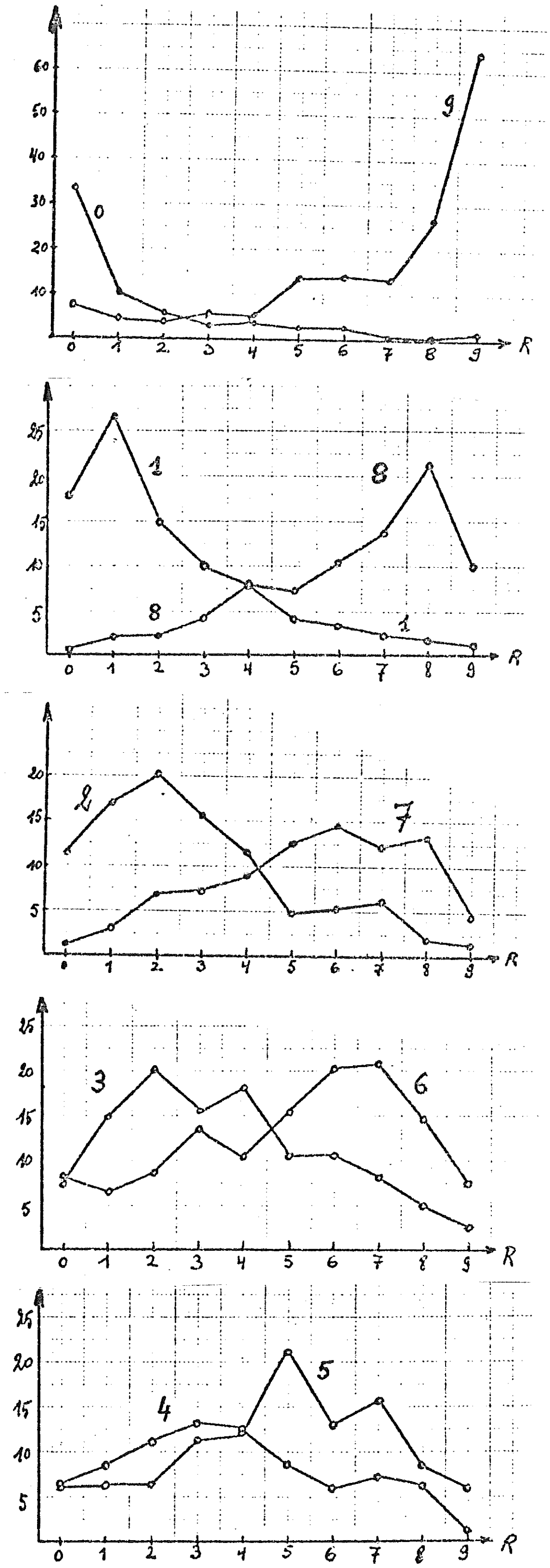
	0	1	2	3	4	5	6	7	8	9
0	36,7	12,4	9,3	7,5	4,4	4,1	4,1	1,2	1,3	1,4
1	13,1	22,2	13,8	10,4	8,4	5,2	4,3	3,0	2,2	1,2
2	11,4	16,1	16,4	12,8	12,0	7,8	7,0	5,9	2,1	1,7
3	11,2	14,8	15,7	15,6	17,5	9,9	10,8	8,0	5,3	2,3
4	4,6	8,3	10,1	12,1	12,4	7,8	6,3	7,7	5,8	1,4
5	6,7	8,0	10,3	10,7	12,2	17,9	13,1	14,4	8,4	6,0
6	7,8	7,8	11,0	13,9	11,8	15,6	18,1	17,2	14,1	8,2
7	1,6	3,7	6,2	6,9	9,3	11,5	11,6	12,4	13,7	6,1
8	1,6	2,6	2,1	4,3	6,9	6,4	8,9	12,9	17,6	9,7
9	5,3	4,1	5,1	5,8	5,1	13,8	15,8	17,6	29,5	61,9
	100	100	100	100	100	100	100	100	100	100

	0	1	2	3	4	5	6	7	8	9
0	33,1	10,0	5,5	2,8	3,5	2,3	2,4	0,6	0,3	1,2
1	17,9	26,7	14,9	10,1	8,1	4,2	3,6	2,3	1,7	1,3
2	11,3	17,0	20,1	15,7	11,9	4,9	5,2	6,1	2,0	1,3
3	7,3	14,8	20,1	15,7	18,2	10,7	10,7	8,1	5,0	2,4
4	6,6	8,5	11,1	13,1	12,6	8,4	6,1	7,3	6,4	1,6
5	6,6	6,7	6,8	11,6	12,6	21,3	13,1	15,7	8,7	6,0
6	7,9	6,7	8,5	13,6	10,9	15,3	20,2	20,9	14,8	7,4
7	1,3	3,0	6,8	7,1	8,8	12,5	14,3	12,2	13,1	4,7
8	0,7	2,2	2,3	4,5	8,1	7,2	10,5	13,7	21,2	10,0
9	7,3	4,4	3,9	5,8	5,3	13,2	13,9	13,1	21,8	64,1
	100	100	100	100	100	100	100	100	100	100

Le tableau 6.34 est la représentation graphique du tableau 6.33.



34 sujets

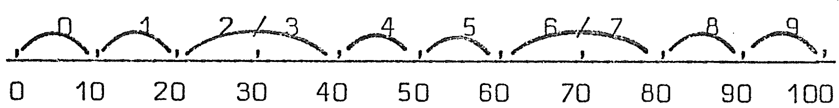


Les 20 sujets retenus

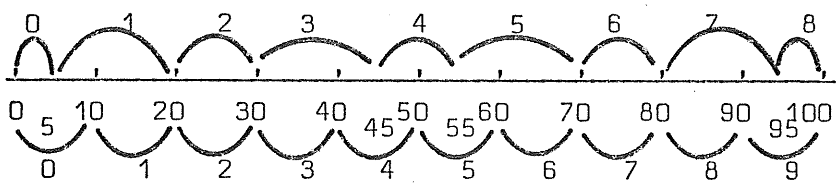
Les Modes des histogrammes de replication correspondent aux valeurs attendues, excepté pour

- le degré 7 = 8 pour les 34 sujets
  = 6 pour les 20 sujets retenus
- le degré 3 = 4 pour les 34 sujets
  = 2 pour les 20 sujets retenus
- le degré 6 = 7 pour les 20 sujets retenus
- le degré 4 = 3 pour les 20 sujets retenus.

Ces résultats contraires à l'hypothèse HF3 indiquent que certains degrés se chevauchent et qu'il faudrait opérer des regroupements. Par exemple, les degrés 6 et 7 seraient regroupés en un seul degré. Symétriquement, les degrés 2 et 3 seraient regroupés. On aurait dès lors une échelle du type :



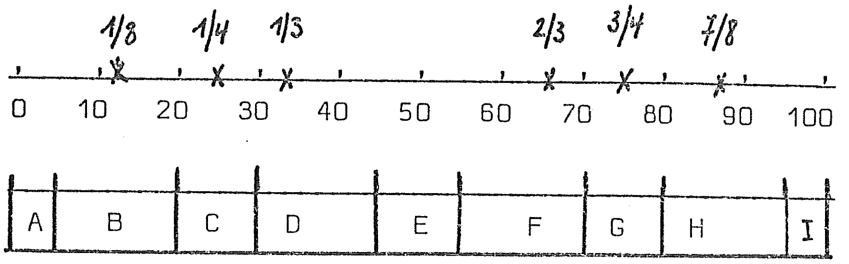
ou encore :



Cette échelle nous paraît idéale parce que :

- Elle crée des zones de certitudes extrêmes (0-5 et 95-100) d'une très petite étendue (5 %).
- Elle crée trois zones centrées sur des repères courants (25 % = zone de 20 à 30; 50 % = zone de 45 à 55; 75 % = zone de 70 à 80), d'une étendue moyenne : 10 %.
- Elle crée quatre zones intermédiaires de grande étendue (15 %) : les zones 3 et 5 contiennent les repères 1/3 et 2/3 et les zones 1 et 7 sont centrées sur les valeurs 1/8 et 7/8.

Ces marques devraient être indiquées aussi bien que les pourcentages :

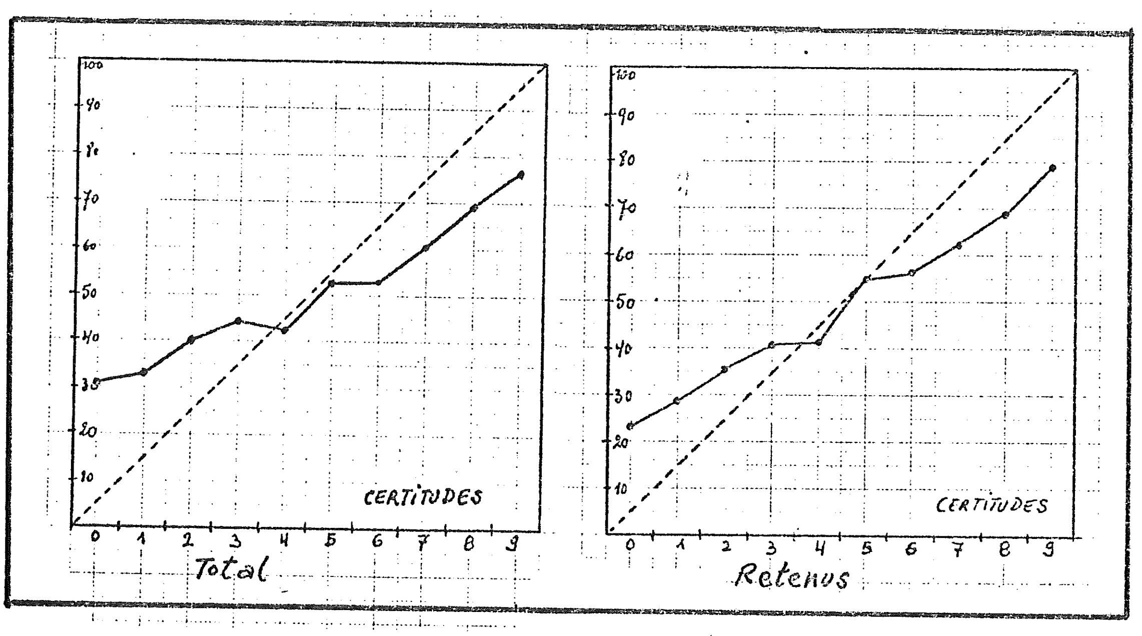


Les centres seraient, respectivement : .025 (ou 1/40)
.125 (ou 1/8)
.25 (ou 1/4)
.375 (ou 3/8)
.5 (ou 1/2)
.625 (ou 5/8)
.75 (ou 3/4)
.875 (ou 7/8)
.975 (ou 39/40)

Les *skewness** sont bonnes, même pour les degrés incriminés. Pour tous les degrés (sauf le degré 4), les histogrammes sont plus leptocurtiques* chez les 20 sujets retenus.

Les fonctions de probabilité subjective sont les suivantes (tableau 6.35).

Tableau 6.35



Le coefficient de pente est plus élevé, de nouveau, pour les 20 sujets retenus. Dans les deux cas, la ligne brisée est proche de la droite et ce sont les certitudes 4 (surtout) et 5 qui rompent la monotonie.

* Estimation grossière par inspection visuelle directe.

SITUATIONS 4/40

Dans ces situations, un problème nouveau surgit : le nombre limité de sujets. Certains degrés de certitude sont peu employés et des analyses fines deviennent impossibles quand les fréquences des matrices de replication sont trop faibles. Des études statistiques seraient d'ailleurs inutiles pour ces degrés de certitude puisque les sujets n'y ont pas recours alors qu'ils sont à leur disposition.

C'est pourquoi les quarante degrés seront regroupés en 20 degrés par la suite.

Le taux d'utilisation des 40 degrés est variable. En fait, un degré sur deux (excepté le degré 2) est très peu utilisé (hypothèse HS.2 confirmée). Deux causes pourraient expliquer ce phénomène :

1. On pourrait penser que la présentation typographique même de l'échelle défavorise l'utilisation de certains degrés. Si cela est vrai, alors on devrait s'attendre à ce que les dizaines exactes (10, 20, 30, 40, 50, 60, 70, 80, 90) soient plus attractives que les nombres tels que 15, 25, 35, 45, 55, 65, 75, 85 et 95. Or ce n'est pas du tout le cas. En effet, 95, 85 et 75 sont respectivement plus utilisés que 90, 80 et 70. Bien entendu, le degré correspondant à 50 % est très utilisé. On remarque aussi l'importance (par rapport aux degrés voisins) des repères classiques : 25, 50 et 75 %.
2. Les sujets n'utilisent pas certains degrés parce que trop précis par rapport à la sensibilité maximale de l'ESPER (hypothèse HS.2).

Le tableau 6.36 fournit les matrices de replication pour 75 sujets et pour 33 sujets (gardés au test K.S.). Les nombres totaux de degrés de certitude n'atteignent pas les valeurs théoriques. Ainsi, au lieu de 7500 certitudes attendues, on n'en observe que 7439. Les certitudes manquantes viennent du fait que quelques sujets qui devaient répondre aux trois échelles ont omis de répondre à l'échelle C. Le code est alors 0. Or ce code n'est pas prévu dans l'échelle C (contrairement aux deux autres échelles).

Tableau 6.36

	Les retenus					Total				
	0	1	2	3		0	1	2	3	
1	108	41	23	7	179	178	79	41	16	314
2	54	18	9	2	83	77	34	20	3	134
3	9	6	4	0	19	20	11	6	1	38
4	56	37	24	6	123	154	101	70	41	366
5	2	2	4	2	10	7	8	4	2	21
6	24	16	20	2	62	44	33	26	4	107
7	1	3	2	0	6	3	6	5	3	17
8	60	54	36	13	163	129	119	106	31	385
9	5	3	1	0	9	9	18	9	1	37
10	22	33	28	12	95	43	69	47	17	176
11	1	1	2	0	4	11	13	8	3	35
12	38	83	64	27	212	109	170	134	48	461
13	3	4	5	1	13	9	17	12	7	45
14	19	64	50	14	147	33	76	69	23	201
15	4	7	6	1	18	12	20	18	17	67
16	16	41	36	30	123	47	104	110	51	312
17	0	1	4	1	6	2	9	11	7	29
18	6	17	18	4	45	13	27	35	7	82
19	1	1	6	3	11	7	24	14	10	55
20	20	80	126	54	280	92	211	257	132	192
21	2	1	7	8	18	3	7	10	9	29
22	7	13	33	13	66	13	27	45	19	104
23	1	1	5	2	9	12	10	27	23	72
24	12	26	54	32	124	46	84	129	88	347
25	0	6	4	4	14	2	7	11	10	30
26	10	30	76	38	154	14	50	89	49	202
27	3	8	12	15	38	12	38	33	54	137
28	14	26	71	73	184	50	92	155	134	431
29	2	5	2	5	14	6	6	10	16	38
30	3	19	44	37	103	15	39	85	67	206
31	1	0	10	2	13	4	5	23	23	55
32	4	15	34	52	105	24	55	120	130	329
33	0	4	5	5	14	0	7	10	15	32
34	4	7	24	51	86	5	17	42	73	137
35	0	2	8	12	22	2	11	19	42	74
36	2	9	28	62	101	15	40	77	167	299
37	0	1	4	13	18	0	3	6	22	31
38	3	5	16	49	73	4	10	35	83	132
39	1	2	2	15	20	3	10	42	70	125
40	8	17	71	385	481	42	71	182	760	1055
	526	709	978	1052	3265	1271	1738	2152	2278	7439

Le tableau 6.37 fournit les pourcentages en colonne, calculés à partir de la matrice du tableau 6.36.

Tableau 6.37

0	1	2	3	0	1	2	3
20,5	5,8	2,4	0,7	14,0	4,5	1,9	0,7
10,3	2,5	0,9	0,2	6,1	2,0	0,9	0,1
1,7	0,9	0,4	0	1,6	0,6	0,3	0
10,6	5,2	2,5	0,6	12,1	5,8	3,3	1,8
0,4	0,3	0,4	0,2	0,6	0,5	0,2	0,1
4,6	2,3	2,0	0,2	3,5	1,9	1,2	0,2
0,2	0,4	0,2	0	0,2	0,3	0,2	0,1
11,4	7,6	3,7	1,2	10,1	6,8	4,9	1,4
1,0	0,4	0,1	0	0,7	1,0	0,4	0
4,2	4,7	2,9	1,1	3,4	4,0	2,2	0,8
0,2	0,1	0,2	0	0,9	0,7	0,4	0,1
7,2	11,7	6,5	2,6	8,6	10,0	6,2	2,1
0,6	0,6	0,5	0,1	0,7	1,0	0,6	0,3
3,6	9,0	5,1	1,3	2,6	4,4	3,2	1,0
0,7	1,0	0,6	0,1	0,9	1,2	0,8	0,8
3,0	5,8	3,7	2,9	3,7	6,0	5,1	2,2
0	0,1	0,4	0,1	0,2	0,5	0,5	0,3
1,1	2,4	1,8	0,4	1,0	1,6	1,6	0,3
0,2	0,1	0,6	0,3	0,6	1,4	0,7	0,4
3,8	11,3	12,9	5,1	7,2	12,1	11,9	5,8
0,4	0,1	0,7	0,8	0,2	0,4	0,4	0,4
1,3	1,8	3,4	1,2	1,0	1,6	2,1	0,8
0,2	0,1	0,5	0,2	0,9	0,5	1,3	1,0
2,3	3,7	5,5	3,0	3,6	4,8	6,0	3,9
0	0,9	0,4	0,4	0,2	0,4	0,5	0,4
1,9	4,2	7,8	3,6	1,1	2,9	4,1	2,2
0,6	1,1	1,2	1,4	0,9	2,2	1,5	2,4
2,7	3,7	7,3	6;9	3,9	5,3	7,2	5,9
0,4	0,7	0,2	0,5	0,5	0,3	0,5	0,7
0,6	2,7	4,5	3,5	1,2	2,2	3,9	2,9
0,2	0	1,0	0,2	0,3	0,3	1,1	1,0
0,7	2,1	3,5	4,9	1,9	3,2	5,6	5,7
0	0,6	0,5	0,5	0	0,4	0,4	0,7
0,7	1,0	2,5	4,9	0,4	1,0	2,0	3,2
0	0,3	0,8	1,1	0,2	0,6	0,9	1,8
0,4	1,3	2,9	5,9	1,2	2,3	3,6	7,3
0	0,1	0,4	1,2	0	0,2	0,3	1,0
0,6	0,7	1,6	4,7	0,3	0,5	1,6	3,7
0,2	0,3	0,2	1,4	0,2	0,5	2,0	3,1
1,5	2,4	7,3	36,6	3,3	4,1	8,5	33,4
100	100	100	100	100	100	100	100
Les retenus				Total			

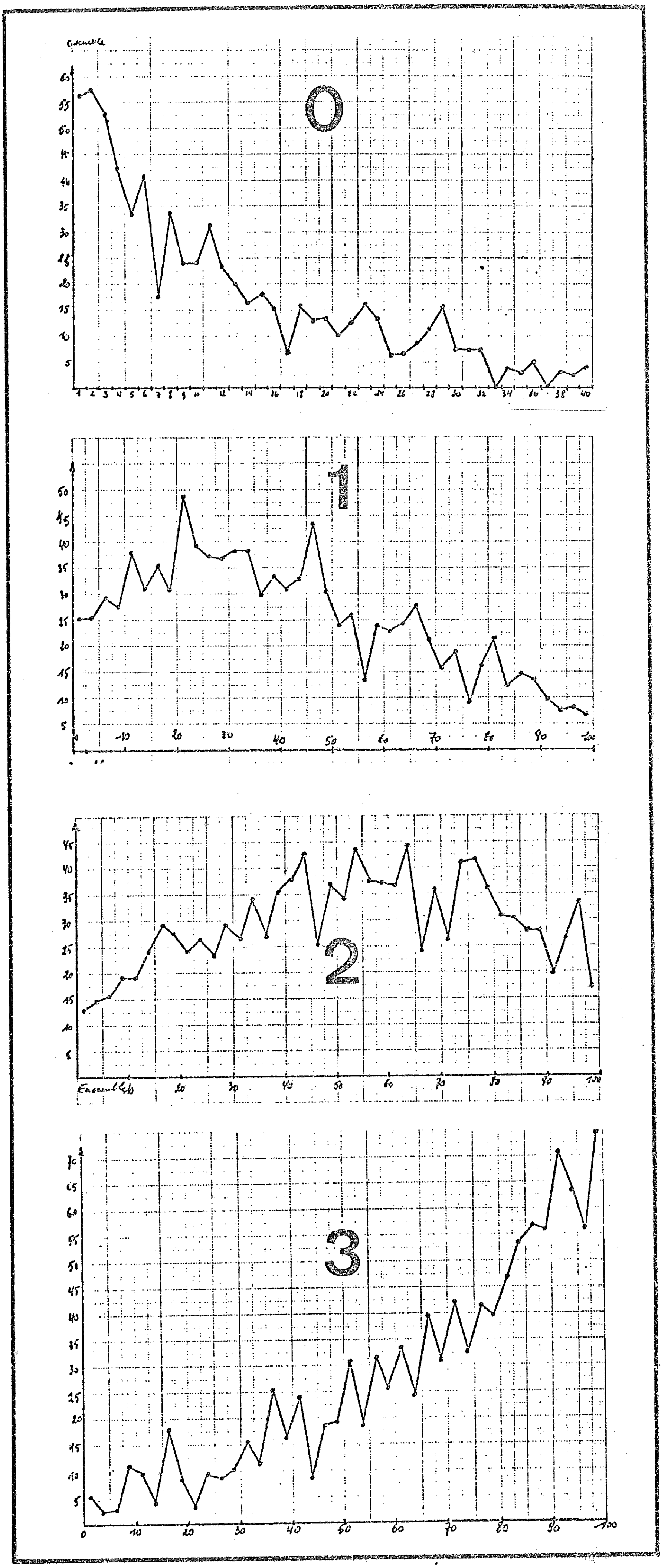
Le tableau 6.38 fournit les pourcentages en ligne calculés à partir de la matrice du tableau 6.36.

Tableau 6.38

0	1	2	3		0	1	2	3	
60,3	22,9	12,9	3,9	100	56,7	25,2	13,0	5,1	100
65,1	21,7	10,8	2,4	100	57,5	25,4	14,9	2,2	100
47,4	31,6	21,0	0	100	52,6	29,0	15,8	2,6	100
45,5	30,1	19,5	4,9	100	42,1	27,6	19,1	11,2	100
20,0	20,0	40,0	20,0	100	33,3	38,1	19,1	9,5	100
38,7	25,8	32,3	3,2	100	41,1	30,9	24,3	3,7	100
16,7	50,0	33,3	0	100	17,6	35,3	29,4	17,7	100
36,8	33,1	22,1	8,0	100	33,5	30,9	27,5	8,1	100
55,6	33,3	11,1	0	100	24,3	48,7	24,3	2,7	100
23,2	34,7	29,5	12,6	100	24,4	39,2	26,7	9,7	100
25,0	25,0	50,0	0	100	31,4	37,1	22,9	8,6	100
17,9	39,2	30,2	12,7	100	23,6	36,9	29,1	10,4	100
23,1	30,8	38,5	7,7	100	20,0	37,8	26,7	15,5	100
12,9	43,6	34,0	9,5	100	16,4	37,8	34,3	11,4	100
22,2	38,9	33,3	5,6	100	17,9	29,8	26,9	25,4	100
13,0	33,3	29,3	24,4	100	15,1	33,3	35,3	16,3	100
0	16,7	66,6	16,7	100	6,9	31,1	37,9	24,1	100
13,3	37,8	40,0	8,9	100	15,9	32,9	42,7	8,5	100
9,1	9,1	54,5	27,3	100	12,7	43,6	25,5	18,2	100
7,1	28,6	45,0	19,3	100	13,3	30,5	37,1	19,1	100
11,1	5,6	38,9	44,4	100	10,4	24,1	34,5	31,0	100
10,6	19,7	50,0	19,7	100	12,5	25,9	43,3	18,3	100
11,1	11,1	55,6	22,2	100	16,7	13,9	37,5	31,9	100
9,7	21,0	43,5	25,8	100	13,2	24,2	37,2	25,4	100
0	42,8	28,6	28,6	100	6,7	23,3	36,7	33,3	100
6,5	19,5	49,3	24,7	100	6,9	24,8	44,1	24,2	100
7,9	21,1	31,6	39,4	100	8,8	27,7	24,1	39,4	100
7,6	14,1	38,6	39,7	100	11,6	21,3	36,0	31,1	100
14,3	35,7	14,3	35,7	100	15,8	15,8	26,3	42,1	100
2,9	18,5	42,7	35,9	100	7,3	18,9	41,3	32,5	100
7,7	0	76,9	15,4	100	7,3	9,1	41,8	41,8	100
3,8	14,3	32,4	49,5	100	7,3	16,7	36,5	39,5	100
0	28,6	35,7	35,7	100	0	21,9	31,2	46,9	100
4,7	8,1	27,9	59,3	100	3,6	12,4	30,7	53,3	100
0	9,1	36,4	54,5	100	2,7	14,9	25,6	56,9	100
2,0	8,9	27,7	61,4	100	5,0	13,4	25,8	55,8	100
0	5,6	22,2	72,2	100	0	9,7	19,4	70,9	100
4,1	6,9	21,9	67,1	100	3,0	7,6	22,5	62,9	100
5,0	10,0	10,0	75,0	100	2,4	8,0	33,6	56,0	100
1,7	3,5	14,8	80,0	100	4,0	6,7	17,3	72,0	100
Les retenus					Total				

Le tableau 6.39 est la traduction graphique du tableau 6.38.

Tableau 6.39



Le tableau 6.40 présente le tableau 6.36 après regroupement des 40 échelons en 20 classes.

Tableau 6.40

0	1	2	3		0	1	2	3	
162	59	32	9	262	255	113	61	19	448
65	43	28	6	142	174	122	76	42	414
26	18	24	4	72	51	41	30	6	128
61	57	38	13	169	132	125	111	34	402
27	36	29	12	104	52	87	56	18	213
39	84	66	27	216	120	183	142	51	496
22	68	55	15	160	42	93	81	30	246
20	48	42	31	141	59	124	128	68	379
6	18	22	5	51	15	36	46	14	111
21	81	132	57	291	99	235	271	142	747
9	14	40	21	84	16	34	55	28	133
13	27	59	34	133	58	94	156	111	419
10	36	80	42	168	16	57	100	59	232
17	34	83	88	222	62	130	188	188	568
5	24	46	42	117	21	45	95	83	244
5	15	44	54	118	28	60	143	153	384
4	11	29	56	100	5	24	52	88	169
2	11	36	74	123	17	51	96	209	373
3	6	20	62	91	4	13	41	105	163
9	19	73	400	501	45	81	224	830	1180

Le tableau 6.41 fournit les pourcentages en ligne calculés à partir du tableau 6.40.

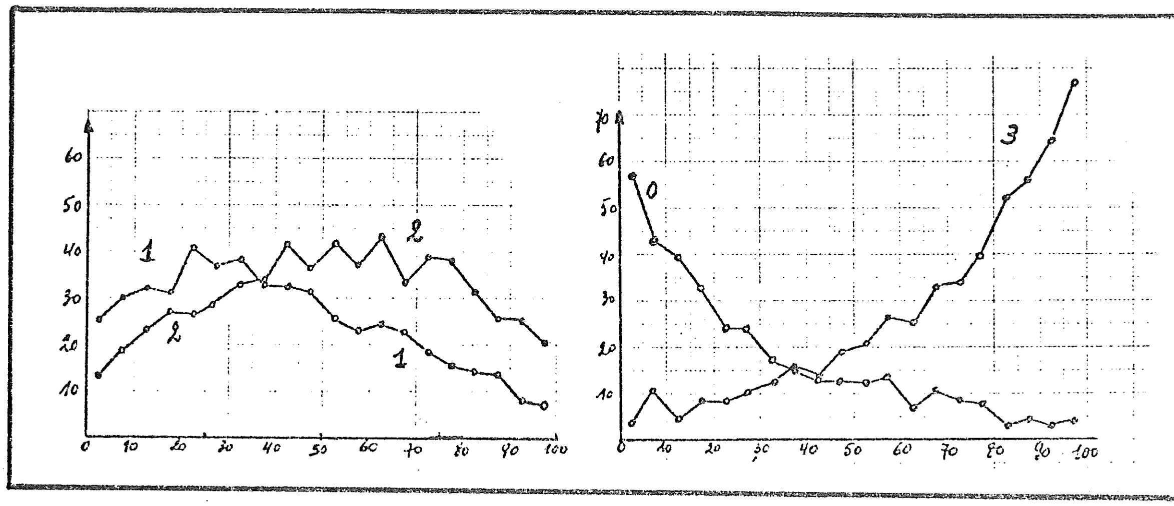
Tableau 6.41

0	1	2	3		0	1	2	3	
61,8	22,5	12,2	3,4	100	56,9	25,2	13,6	4,2	100
45,8	30,3	19,7	4,2	100	43,0	30,2	18,8	10,4	100
36,1	25,0	33,3	5,5	100	39,8	32,0	23,4	4,7	100
36,1	33,7	22,4	7,7	100	32,8	31,1	27,6	8,4	100
26,0	34,6	27,9	11,5	100	24,4	40,8	26,3	8,4	100
18,1	38,9	30,6	12,5	100	24,2	36,9	28,6	10,3	100
13,7	42,5	34,4	9,4	100	17,1	37,8	32,9	12,2	100
14,2	34,0	29,8	22,0	100	15,6	32,7	33,8	17,9	100
11,7	35,3	43,1	9,8	100	13,5	32,4	41,4	12,6	100
7,2	27,8	45,4	19,6	100	13,2	31,4	36,3	19,0	100
10,7	16,6	47,6	25,0	100	12,0	25,5	41,4	21,0	100
9,8	20,3	44,4	25,6	100	13,8	22,4	37,2	26,5	100
5,9	21,4	47,6	25,0	100	6,9	24,6	43,1	25,4	100
7,6	15,3	37,4	39,6	100	10,9	22,9	33,1	33,1	100
4,3	20,5	39,3	35,9	100	8,6	18,4	38,9	34,0	100
4,2	12,7	37,3	45,8	100	7,3	15,6	37,2	39,8	100
4,0	11,0	29,0	56,0	100	2,9	14,2	30,8	52,1	100
1,6	8,9	29,3	60,2	100	4,6	13,7	25,7	56,0	100
3,3	6,6	22,0	68,1	100	2,5	8,0	25,1	64,4	100
1,8	3,8	14,5	79,8	100	4,2	7,5	20,7	76,9	100

Les retenus   Total

Les histogrammes de replication qui traduisent graphiquement le tableau 6.41 figurent dans le tableau 6.42 pour les 75 sujets et dans le tableau 6.43 pour les 33 sujets gardés (après test K.S.).

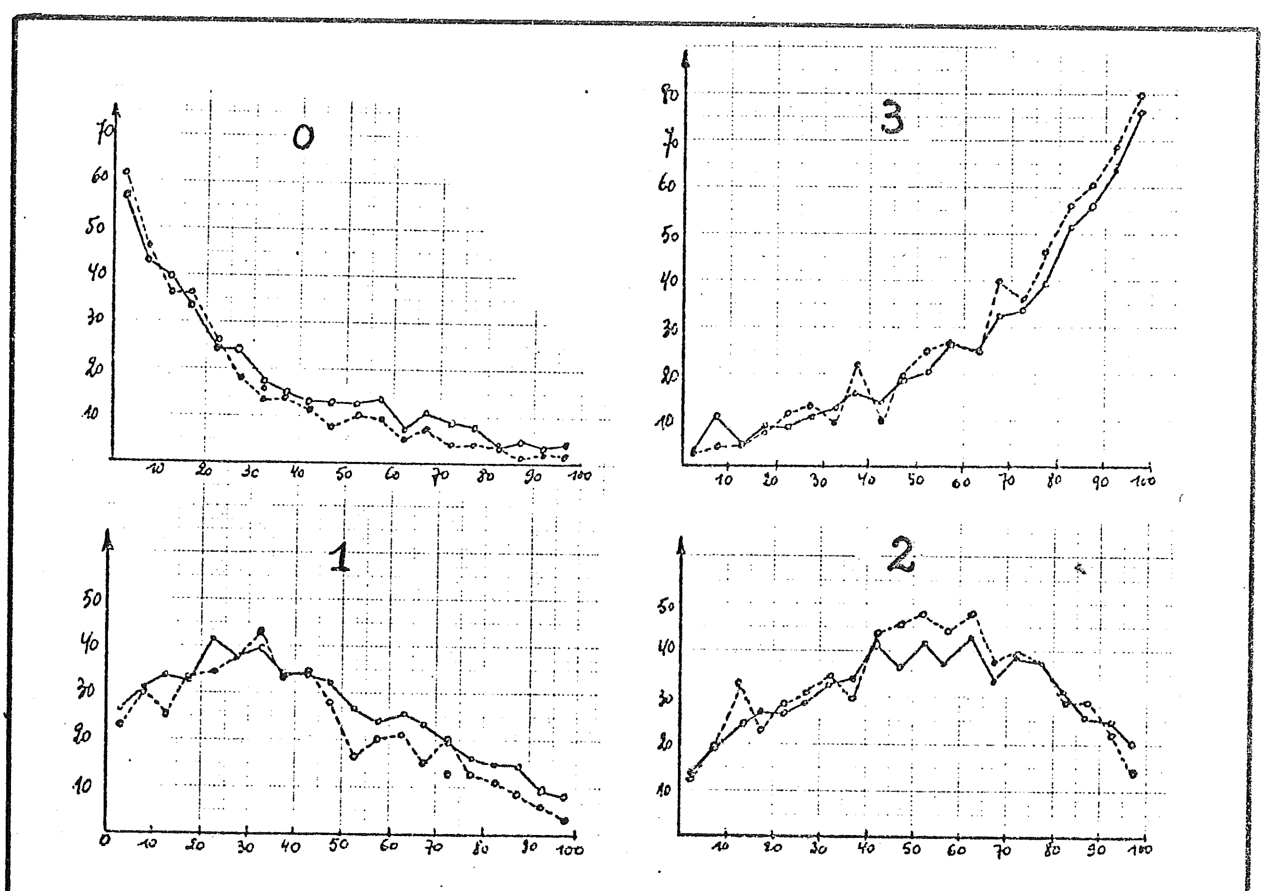
Tableau 6.42



La *skewness* de la certitude 2 est un peu faible, de même que sa *curtosis* (Les résultats de tous les sujets ont été utilisés).

Voici, portés sur le même graphique que l'ensemble des sujets (en traits pleins), les histogrammes des sujets retenus (en traits pointillés).

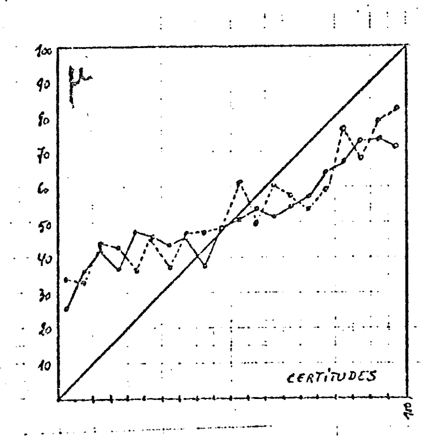
Tableau 6.43



Pour les certitudes 0 et 3, les histogrammes sont légèrement meilleurs avec les seuls sujets retenus. Pour la certitude 1, la *skewness* est plus faible, mais la *curtosis* meilleure. En ce qui concerne la certitude 2, la courbe est plus leptocurtique pour les sujets retenus que pour l'ensemble des sujets.

Les fonctions de probabilité subjective pour les deux échantillons sont les suivantes (tableau 6.44).

Tableau 6.44



En traits pleins : les 75 sujets.
En traits pointillés : les 33 sujets retenus.

Les fluctuations sont importantes. Elles sont en grande partie dues aux faibles répétitions de certains degrés.

La fonction de probabilité subjective des sujets retenus ne semble pas meilleure que celle de l'ensemble des 75 sujets.

SITUATIONS 40/40

Les situations 10/40 et plus encore 40/40 rencontrent, multiplié respectivement par 2,5 et par 10, le problème des nombres insuffisants de répétitions pour certaines classes. Le calcul des histogrammes de replication est impossible et d'ailleurs pour une large part inutile, puisqu'il est déjà difficile de discriminer entre dix degrés seulement.

Le tableau 6.45 présente la matrice de replication pour 69 sujets; le tableau 6.46 présente la matrice de replication pour les 33 sujets qui ont été gardés après le test de Kolmogorov-Smirnov.

Tableau 6.45

	0	1	2	3	4	5	6	7	8	9		BR	
1	112	35	45	21	11	17	18	6	7	10	282	84	29,8
2	21	17	13	15	5	11	5	1	0	0	88	23	26,1
3	4	12	2	6	3	2	3	1	0	1	34	11	32,3
4	65	53	45	49	23	22	26	10	10	10	313	109	34,8
5	1	4	4	2	1	2	0	0	0	2	16	6	37,5
6	9	13	13	7	5	4	6	3	2	2	64	24	37,5
7	1	4	2	3	0	1	1	1	0	4	17	9	52,9
8	53	57	56	61	30	53	31	21	14	22	398	146	37,1
9	7	5	33	4	4	24	1	3	1	14	96	46	47,9
10	17	8	16	28	9	11	16	4	2	9	120	51	42,5
11	2	5	6	5	2	3	2	3	0	2	30	13	43,3
12	37	37	56	60	38	65	46	35	14	18	406	187	46,1
13	6	5	5	10	2	4	5	6	1	3	47	23	48,9
14	8	9	19	32	10	11	22	7	4	11	133	53	39,8
15	5	9	5	4	4	8	10	2	3	19	69	29	42,0
16	16	26	30	49	38	38	42	32	30	22	323	145	44,9
17	0	0	3	3	0	3	1	1	2	2	15	3	20,0
18	4	6	2	7	3	3	10	1	2	3	41	19	46,3
19	3	1	6	6	6	6	4	2	2	7	43	17	39,5
20	43	29	52	78	29	115	81	55	36	67	585	306	52,3
21	0	4	4	2	2	11	5	4	2	10	44	21	47,7
22	5	2	4	8	3	17	14	9	5	10	77	49	63,6
23	3	2	4	5	5	9	6	5	2	20	61	36	59,0
24	18	15	30	40	24	52	74	40	37	56	386	197	51,0
25	0	1	4	3	2	1	2	4	2	8	27	16	59,2
26	4	6	14	21	5	16	40	19	5	23	153	81	52,9
27	5	4	2	23	6	6	12	2	11	32	103	52	50,5
28	14	24	31	40	28	67	71	49	73	73	470	285	60,6
29	1	4	1	2	1	3	3	5	7	4	31	17	54,8
30	4	2	6	13	2	19	19	8	11	26	110	67	60,9
31	2	2	6	4	0	14	2	5	4	28	67	47	70,1
32	2	12	16	25	18	35	45	52	48	91	344	219	63,7
33	0	1	3	4	3	2	8	3	4	10	38	24	63,1
34	0	3	3	5	2	10	8	7	17	27	82	60	73,2
35	2	1	0	8	2	3	6	5	7	22	56	34	60,7
36	8	7	4	14	12	24	33	39	62	135	338	234	69,2
37	0	1	4	6	1	0	1	3	3	12	31	17	54,8
38	3	1	0	3	3	5	6	7	8	38	74	49	66,2
39	2	4	4	10	8	24	28	13	27	116	236	173	73,3
40	12	17	11	25	15	38	72	48	79	682	999	777	77,7

Tableau 6.46

	0	1	2	3	4	5	6	7	8	9	
1	48	20	17	7	6	8	8	2	2	4	122
2	19	13	11	5	2	5	2	1	0	0	58
3	2	10	1	3	1	1	2	1	0	0	21
4	14	28	28	17	11	13	7	5	4	2	129
5	0	3	3	0	1	1	0	0	0	2	10
6	8	13	8	4	2	3	5	2	1	2	48
7	0	3	2	3	0	0	0	1	0	1	10
8	13	24	25	32	16	22	9	14	6	6	167
9	2	3	2	1	0	4	0	0	0	0	12
10	10	4	12	19	3	4	12	3	2	7	76
11	0	4	5	0	0	0	0	1	0	0	10
12	2	11	33	30	22	23	20	21	9	7	178
13	0	2	1	1	1	2	1	2	1	0	11
14	5	6	14	29	7	7	22	4	3	8	105
15	2	5	3	0	1	0	5	2	1	3	22
16	5	12	22	27	27	20	25	21	21	11	191
17	0	0	3	1	0	2	0	0	1	0	7
18	3	4	1	5	3	3	7	0	1	2	29
19	1	0	0	2	0	1	2	2	2	0	10
20	9	6	17	32	16	60	38	32	21	29	260
21	0	3	1	0	2	6	4	3	2	9	30
22	2	1	1	7	1	13	9	9	4	6	53
23	1	1	1	2	1	1	1	1	0	6	15
24	4	4	9	16	10	25	41	26	24	20	179
25	0	0	3	1	1	0	1	3	2	2	13
26	2	6	10	16	3	12	35	17	5	19	125
27	0	2	1	4	1	1	5	1	5	8	28
28	3	10	9	12	15	28	42	38	45	36	238
29	0	2	0	0	1	2	0	0	3	2	10
30	1	1	3	9	2	12	16	1	6	15	66
31	0	2	0	0	0	2	1	2	3	1	11
32	0	2	5	10	10	14	25	24	25	44	159
33	0	1	2	1	2	1	3	2	2	2	16
34	0	2	2	1	1	6	8	4	15	19	58
35	0	1	0	1	1	1	2	4	1	7	18
36	2	1	0	3	4	14	13	21	32	63	153
37	0	1	0	1	0	0	1	3	3	5	14
38	2	1	0	3	0	3	3	3	3	23	41
39	1	0	2	1	1	8	2	2	3	36	56
40	3	5	6	8	9	14	38	22	47	377	529
	164	217	263	314	184	342	415	300	305	784	3288

Les tableaux 6.47 et 6.48 présentent respectivement les pourcentages de ligne calculés sur les matrices des tableaux 6.45 et 6.46.

Tableau 6.47

	0	1	2	3	4	5	6	7	8	9	
1	39,7	12,4	16,0	7,5	3,9	6,0	6,4	2,1	2,5	3,5	100
2	23,9	19,3	14,8	17,0	5,7	12,5	5,7	1,1	0	0	100
3	11,8	35,3	5,9	17,7	8,8	5,9	8,8	2,9	0	2,9	100
4	20,8	16,9	14,4	15,7	7,3	7,0	8,3	3,2	3,2	3,2	100
5	6,2	25,0	25,0	12,5	6,3	12,5	0	0	0	17,5	100
6	14,1	20,3	20,3	10,9	7,8	6,3	9,4	4,7	3,1	3,1	100
7	5,9	23,5	11,8	17,6	0	5,9	5,9	5,9	0	23,5	100
8	13,3	14,3	14,1	15,3	7,5	13,3	7,8	5,3	3,5	5,6	100
9	7,3	5,2	34,4	4,2	4,2	25,0	1,0	3,1	1,0	14,6	100
10	14,2	6,7	13,3	23,3	7,5	9,2	13,3	3,3	1,7	7,5	100
11	6,7	16,6	20,0	16,6	6,7	10,0	6,7	10,0	0	6,7	100
12	9,1	9,1	13,8	14,8	9,4	16,0	10,3	8,6	3,5	4,4	100
13	12,8	10,6	10,6	21,3	4,3	8,5	10,6	12,8	2,1	6,4	100
14	6,0	6,8	14,3	24,1	7,5	8,3	16,5	5,2	3,0	8,3	100
15	7,3	13,0	7,3	5,8	5,8	11,6	14,5	2,9	4,3	27,5	100
16	4,9	8,0	9,3	15,2	11,8	11,8	13,0	9,9	9,3	6,8	100
17	0	0	20,0	20,0	0	20,0	6,7	6,7	13,3	13,3	100
18	9,8	14,6	4,9	17,1	7,3	7,3	24,4	2,4	4,9	7,3	100
19	7,0	2,3	14,0	14,0	14,0	14,0	9,3	4,6	4,6	16,2	100
20	7,4	4,9	8,9	13,3	4,9	19,7	13,8	9,4	6,2	11,5	100
21	0	9,1	9,1	4,5	4,5	25,0	11,4	9,1	4,5	22,8	100
22	6,5	2,6	5,2	10,4	3,9	22,1	18,2	11,7	6,5	12,9	100
23	4,9	3,3	6,6	8,2	8,2	14,7	9,8	8,2	3,3	32,7	100
24	4,6	3,9	7,8	10,3	6,2	13,5	19,2	10,4	9,6	14,5	100
25	0	3,7	14,8	11,1	7,4	3,7	7,4	14,8	7,4	29,7	100
26	2,6	3,9	9,2	13,7	3,3	10,5	26,1	12,4	3,3	15,0	100
27	4,9	3,9	1,9	22,3	5,8	5,8	11,7	1,9	10,7	31,1	100
28	3,0	5,1	6,6	8,5	6,0	14,3	15,1	10,4	15,5	15,5	100
29	3,2	12,9	3,2	6,5	3,2	9,7	9,7	16,1	22,6	12,9	100
30	3,6	1,8	5,5	11,8	1,8	17,3	17,3	7,3	10,0	23,6	100
31	2,9	2,9	9,0	4,0	0	21,0	2,9	7,5	6,0	41,8	100
32	0,6	3,5	4,6	7,2	5,3	10,2	13,1	15,1	13,9	24,5	100
33	0	2,6	7,9	10,5	7,9	5,3	21,1	7,9	10,5	26,3	100
34	0	3,7	3,7	6,1	2,4	12,2	9,8	8,5	20,7	32,9	100
35	3,6	1,8	0	14,3	3,6	5,3	10,7	8,9	12,5	39,9	100.
36	2,4	2,1	1,2	4,1	3,6	7,1	9,8	11,5	18,3	39,9	100
37	0	3,2	12,9	19,4	3,2	0	3,2	9,7	9,7	38,7	100
38	4,1	1,3	0	4,1	4,1	6,8	8,1	9,4	10,8	51,3	100
39	0,8	1,7	1,7	4,2	3,4	10,2	11,9	5,5	11,4	49,2	100
40	1,2	1,7	1,1	2,5	1,5	3,8	7,2	4,8	7,9	68,3	100

Tableau 6.48

	0	1	2	3	4	5	6	7	8	9	
1	39,4	16,4	13,9	5,7	4,9	6,6	6,6	1,6	1,6	3,3	100
2	32,8	22,4	10,0	8,	3,4	8,6	3,4	1,8	0	0	100
3	9,5	47,6	4,8	14,8	4,8	4,8	9,5	4,8	0	0	100
4	10,9	21,7	21,7	13,2	8,5	10,1	5,4	3,9	3,1	1,5	100
5	0	30,0	30,0	0	10,0	10,0	0	0	0	20,0	100
6	16,6	27,1	16,6	8,3	4,2	6,3	10,4	4,2	2,1	4,2	100
7	0	30,0	20,0	30,0	0	0	0	10,0	0	16,0	100
8	7,8	14,4	15,0	19,1	9,6	13,1	5,4	8,4	3,6	3,6	100
9	16,7	25,0	16,7	6,3	0	33,3	0	0	0	0	100
10	13,2	5,3	15,8	25,0	3,9	5,3	15,8	3,9	2,6	9,2	100
11	0	40,0	50,0	0	0	0	0	10,0	0	0	100
12	1,1	6,2	18,5	16,9	12,4	12,9	11,2	11,8	5,1	3,9	100
13	0	18,2	9,1	9,1	9,1	18,2	9,1	18,2	9,1	0	100
14	4,7	5,7	13,3	27,6	6,7	6,7	21,0	3,8	2,9	7,6	100
15	9,1	22,7	13,6	0	4,6	0	22,7	9,1	4,6	13,6	100
16	2,6	6,3	11,5	14,1	14,1	10,5	13,1	11,0	11,0	5,8	100
17	0	0	42,9	14,3	0	28,5	0	0	14,3	0	100
18	10,4	13,8	3,4	17,2	10,4	10,4	24,1	0	3,4	6,9	100
19	10,0	0	0	20,0	0	10,0	20,0	20,0	20,0	0	100
20	3,5	2,3	6,5	12,3	6,2	23,1	14,6	12,3	8,1	11,1	100
21	0	10,0	3,3	0	6,7	20,0	13,3	10,0	6,7	30,0	100
22	3,8	1,9	1,9	13,2	1,9	24,5	17,0	17,0	7,5	11,3	100
23	6,7	6,7	6,7	13,3	6,7	6,7	6,7	6,7	0	40,0	100
24	2,3	2,3	5,0	8,9	5,6	14,0	22,9	14,5	13,4	11,2	100
25	0	0	23,1	7,7	7,7	0	7,7	23,1	15,4	15,4	100
26	1,6	4,8	8,0	12,8	2,4	9,6	28,0	13,6	4,0	15,2	100
27	0	7,1	3,6	14,3	3,6	3,6	17,8	3,6	17,8	28,6	100
28	1,3	4,2	3,8	5,0	6,3	11,8	17,6	16,1	18,9	15,1	100
29	0,0	20,0	0	0	10,0	20,0	0	0	30,0	20,0	100
30	1,5	1,5	4,6	13,7	3,0	18,2	24,2	1,5	9,1	22,7	100
31	0	18,2	0	0	0	18,2	9,1	18,2	27,2	9,1	100
32	0	1,3	3,1	6,3	6,3	8,8	15,7	15,1	15,7	27,7	100
33	0	6,2	12,5	6,2	12,5	6,2	18,8	12,5	12,5	12,5	100
34	0	3,5	3,5	1,7	1,7	10,3	13,8	6,9	25,9	32,7	100
35	0	5,6	0	5,6	5,6	5,6	11,1	22,2	5,6	38,9	100
36	1,3	0,7	0	2,0	2,6	9,2	8,5	13,7	20,9	41,1	100
37	0	7,2	0	7,2	0	0	7,2	21,4	21,4	35,7	100
38	4,9	2,5	0	7,3	0	7,3	7,3	7,3	7,3	56,1	100
39	1,8	0	3,6	1,8	1,8	14,3	3,6	3,6	5,4	64,3	100
40	0,6	0,9	1,1	1,5	1,7	2,5	7,2	4,2	8,9	71,3	100

Le tableau 6.49 correspond au tableau 6.45, après regroupement des 40 échelons en 20 classes.

Tableau 6.49

0	1	2	3	4	5	6	7	8	9	
133	52	58	36	16	28	23	7	7	10	370
69	65	47	55	26	24	29	11	10	11	347
10	17	17	9	6	6	6	3	2	4	80
54	61	58	64	30	54	32	22	14	26	415
24	13	49	32	13	35	17	7	3	23	216
39	42	62	65	40	68	48	38	14	20	436
14	14	24	42	12	15	27	13	5	14	180
21	35	35	53	42	46	52	34	33	41	392
4	6	5	10	3	6	11	2	4	6	56
46	30	58	84	35	121	85	57	38	74	628
5	6	8	10	5	28	19	13	7	20	121
21	17	34	45	29	61	60	45	39	76	447
4	7	18	24	7	17	42	23	7	31	180
19	28	33	63	34	73	83	51	84	105	573
5	6	7	15	3	22	22	13	18	30	141
4	14	22	29	18	49	47	57	52	119	411
0	4	6	9	5	12	16	10	21	37	120
10	8	4	22	14	27	39	44	69	157	394
3	2	4	9	4	5	7	10	11	50	105
14	21	15	35	23	62	100	61	106	798	1235

Le tableau 6.50 présente les pourcentages de ligne calculés sur la matrice du tableau 6.49.

0	1	2	3	4	5	6	7	8	9	
35,9	14,1	15,7	9,7	4,3	7,6	6,2	1,9	1,9	2,7	100
19,9	18,7	13,5	15,8	7,5	6,9	8,3	3,2	2,9	3,2	100
12,5	21,7	21,2	11,2	7,5	7,5	7,5	3,7	2,5	5	100
13,0	14,7	14,0	15,4	7,2	13,0	7,7	5,3	3,4	6,3	100
11,1	6,0	22,6	14,8	6,0	16,2	7,9	3,2	1,4	10,6	100
8,9	9,6	14,2	14,9	9,2	15,6	11,0	8,7	3,2	4,6	100
7,8	7,8	13,3	23,3	6,7	8,3	15,0	7,2	2,7	7,8	100
5,3	8,9	8,9	13,5	10,7	11,7	13,3	8,7	8,4	10,5	100
7,1	10,7	8,9	17,8	5,3	10,7	19,6	3,6	7,1	10,7	100
7,3	4,8	9,2	13,4	5,6	19,3	13,5	9,1	6,0	11,8	100
4,1	4,9	6,6	8,3	4,1	23,1	15,7	10,7	5,8	16,5	100
4,7	3,8	7,6	10,1	6,5	13,6	17,9	10,1	8,7	17,0	100
2,2	3,9	10,0	13,3	3,9	9,4	23,3	12,8	3,9	17,2	100
3,3	4,9	5,7	10,0	5,9	12,7	14,5	8,9	14,6	18,3	100
3,5	4,2	5,0	10,6	2,1	15,6	15,6	9,2	12,8	21,3	100
0,9	3,4	5,3	7,0	4,4	11,9	11,4	13,9	12,6	28,9	100
0	3,3	5,0	7,5	4,2	10,0	13,3	8,3	17,5	30,8	100
2,5	2,0	1,0	5,6	3,6	6,8	9,9	11,1	17,5	39,8	100
2,9	1,9	3,8	8,6	3,8	4,8	6,6	9,5	10,5	47,6	100
1,1	1,7	1,2	2,8	1,9	5,0	8,1	4,9	8,6	64,6	100

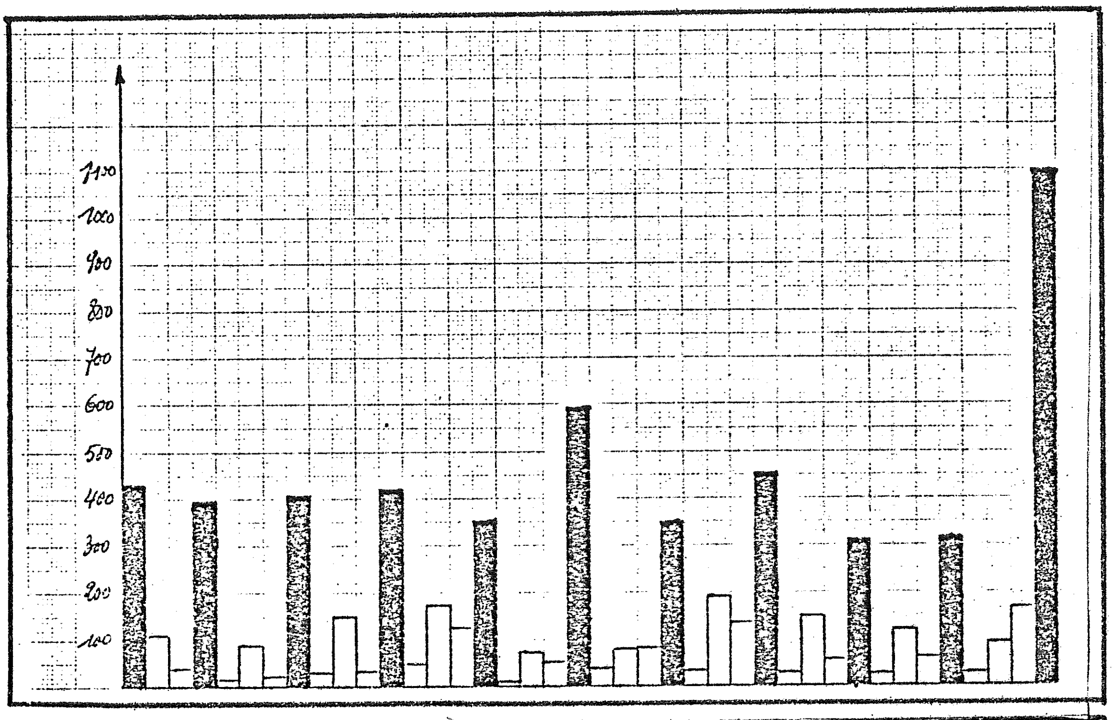
Tableau 6.50

Dans le tableau 6.51 sont présentées les valeurs numériques des μ (taux d'exactitude) de chaque degré de certitude. Les tableaux 6.52 et 6.53 présentent les τ et les μ de façon graphique. Les indices de certitude pour lesquels τ est élevé ont été représentés par des bâtons noirs et les μ correspondants par des cercles noirs. La ligne (brisée) qui joint ces cercles noirs présente nettement moins de fluctuations que la ligne (brisée) qui joint tous les μ. Cette constatation est rassurante pour la validité de l'ESPER.

Tableau 6.51

	BR	MR	Total	μ
1	118	313	431	27,37
2	27	83	110	24,5
3	11	31	42	26,2
4	146	243	389	37,53
5	5	10	15	33,3
6	30	59	89	33,7
7	9	13	22	40,9
8	149	254	403	36,97
9	10	20	30	33,33
10	59	90	149	39,59
11	13	18	33	39,39
12	190	228	418	45,45
13	19	30	49	38,77
14	79	97	176	44,88
15	54	72	126	42,85
16	168	183	351	47,86
17	4	15	19	21,05
18	30	45	75	40,0
19	21	31	52	40,38
20	309	281	590	52,37
21	16	21	37	43,24
22	44	36	80	55,0
23	46	36	82	56,09
24	189	158	347	54,46
25	22	14	36	61,11
26	104	87	191	54,45
27	73	64	137	53,28
28	291	160	451	64,52
29	16	17	33	48,48
30	86	63	149	57,71
31	34	24	58	58,6
32	204	108	312	65,38
33	23	9	32	71,87
34	91	30	121	75,20
35	46	19	65	70,76
36	233	83	316	73,73
37	18	10	28	64,28
38	72	21	98	76,26
39	125	42	167	74,85
40	879	219	1098	80,05
	4063	3337	7400	

Tableau 6.52



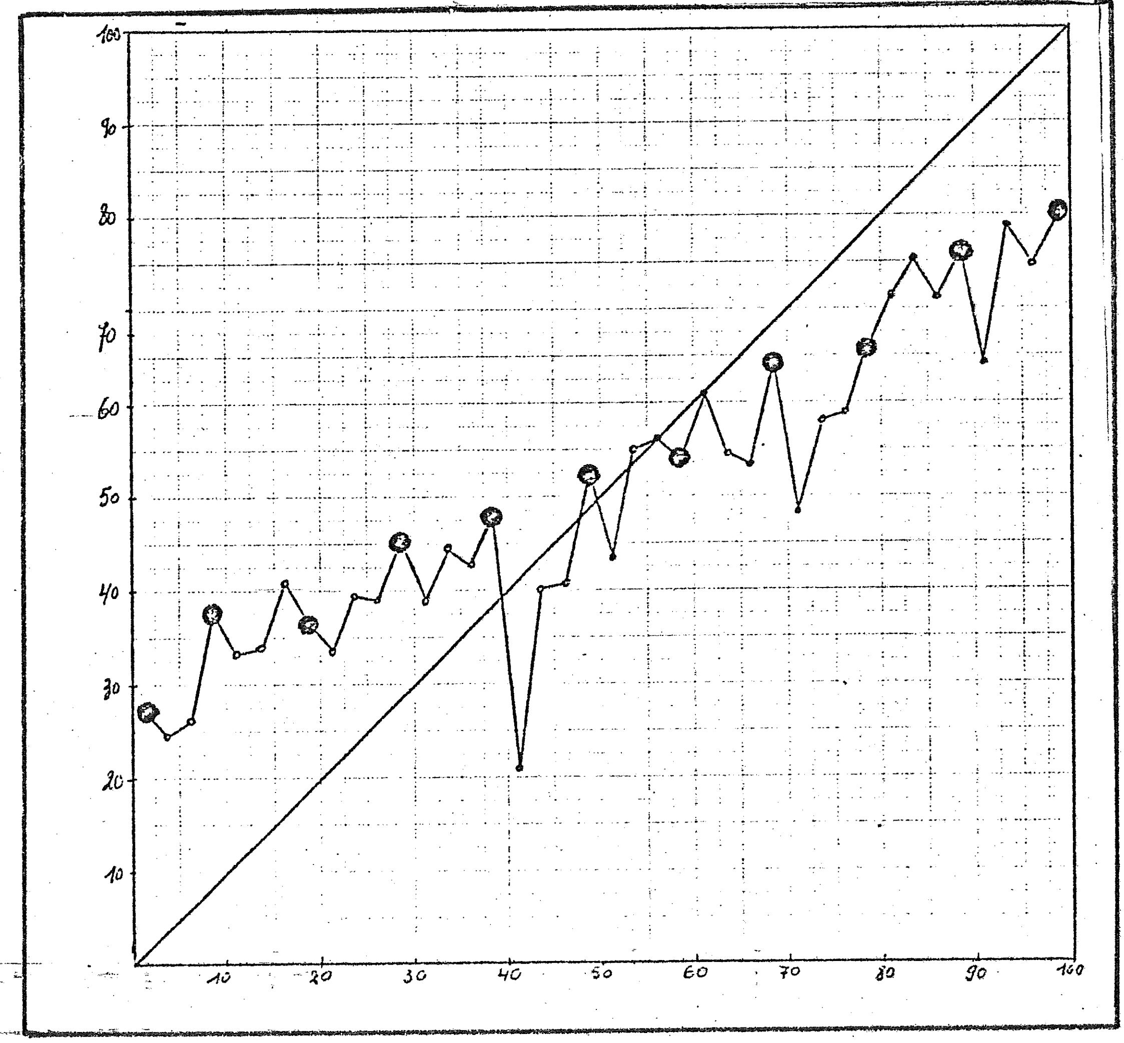


Tableau 6.53

G. LA LONGUEUR DE L'EPREUVE

En cours de travail, des réserves ont été émises, par les sujets eux-mêmes, sur la procédure : fournir cent réponses avec une certitude précise constitue un travail fastidieux. Il s'agit, en effet, de faire cent fois une série d'hypothèses, de les pondérer relativement l'une à l'autre, de choisir la meilleure et, finalement, d'attribuer une PER précise.

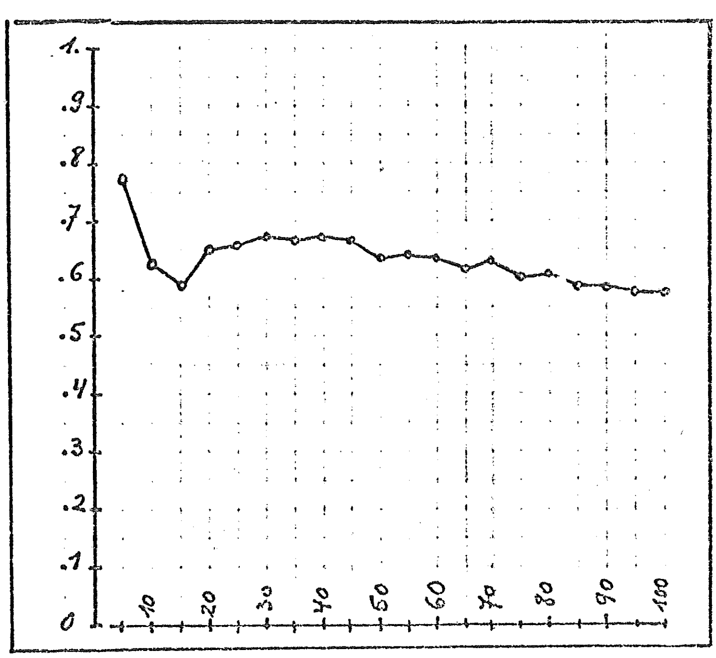
On peut penser qu'au fil des questions, les sujets se découragent et mettent moins d'application dans leur effort. La stabilité de l'ESPER devrait, en conséquence, diminuer avec le nombre de questions. Les toute premières questions étant plus mûries aussi bien avant (au test) qu'après (au retest), elles devraient présenter une corrélation (test-retest) plus forte que les questions ultérieures (hypothèse HF1). Des corrélations ont été calculées successivement pour les cinq premières, puis les dix premières, puis les quinze premières..., puis les nonante-cinq premières questions. Le calcul a été effectué à partir des données des 22 sujets suivants :

AA : 132, 251, 118, 185, 22
BA : 200, 229
CA : 163
AB : 36, 173, 239
BB : 262, 295, 199, 128
AC : 148, 156
TT : 57, 19, 18, 48, 34.

Les valeurs médianes des corrélations sont

5	.770	55	.640
10	.625	60	.645
15	.595	65	.635
20	.650	70	.620
25	.655	75	.630
30	.675	80	.600
35	.670	85	.605
40	.680	90	.580
45	.670	95	.580
50	.645	100	.575

Tableau 6.54



Valeurs des corrélations entre les degrés de certitude fournis au test et au retest pour les séries cumulées des 5, 10, 15, 20... premières lettres du jeu des ESPER (100 lettres).

L'hypothèse HF.1 est donc confirmée.

La chute aux alentours de la quinzaine de questions est difficilement interprétable. Si, comme nous le pensons, la baisse générale de corrélation provient d'une baisse dans la concentration mentale, le nombre moyen de bonnes réponses doit, lui aussi, chuter.

Les calculs qui suivent ont été effectués sur les résultats des mêmes 22 sujets.

Tableau 6.55

																							M	M'
20	58	43	60	58	54	60	61	57	59	53	66	63	49	60	53	54	57	59	59	56	46	52	56,22	65,72
19	57	42	60	57	53	60	60	56	61	53	65	63	49	60	51	54	56	60	57	56	45	51	55,72	72,1
18	55	42	58	55	53	58	60	56	60	53	64	62	48	60	51	53	57	58	56	55	43	50	54,81	76,4
17	54	41	57	52	52	56	57	55	57	52	63	62	47	57	49	54	56	56	55	54	42	50	53,54	36,3
16	55	40	57	52	53	57	60	57	58	55	65	63	47	60	51	53	57	57	56	55	43	52	54,68	50,63
15	56	37	57	52	53	57	58	60	60	56	65	64	48	60	52	53	57	56	57	54	44	53	54,95.	68,25
14	55	37	58	51	55	57	58.	60	58	54	64	62	44	57	50	50	55	57	55	54	45	52	54,0	46,98
13	56	38	58	52	55	56	6I	58	56	55	64	63	46	56	49	53	58	58	55	53	46	55	54,54	39,3
12	56	41	56	53	58	56	63	58	56	56	66	65	50	56	51	56	56	59	58	55	46	56	55,81	51,74
11	56	43	58	52	56	58	65	58	56	58	67	65	47	58	54	58	58	56	60	54	43	56	56,18	56,18
10	58	44	58	52	58	58	66	60	54	60	68	64	46	58	56	58	58	54	58	58	44	58	56,72	60,68
9	57	40	57	53	55	55	66	60	55	60	66	64	40	57	53	57	55	53	55	55	46	60	55,68	44,08
8	60	40	60	55	55	57	67	62	55	57	67	67	47	60	55	60	57	55	57	57	47	60	57,13	63,5
7	60	37	57	51	54	54	68	60	51	60	68	65	48	57	54	60	57	54	57	57	48	60	56,22	49,38
6	60	33	56	53	53	60	73	60	53	63	66	66	50	56	56	56	56	56	60	63	50	63	57,36	30,56
5	64	40	64	60	60	64	80	68	56	68	68	72	56	60	64	60	64	56	64	68	56	68	62,72	56,36
4	70	45	70	65	60	65	80	70	55	70	65	75	55	65	65	60	65	50	65	70	60	70	64,31	61,97
3	73	33	66	60	86	66	80	73	46	66	60	66	53	60	60	53	66	53	66	60	60	66	65,09	61,65
2	80	40	70	60	80	70	70	80	40	70	60	70	70	60	70	70	80	60	70	60	70	70	66,81	68,17
1	80	40	40	60	60	40	60	60	40	60	60	40	80	40	80	60	80	80	80	60	60	80	65,45	65,45

Pourcentages cumulés de réponses correctes sur 5 lettres (1), 10 lettres (2), 15 lettres (3)... 100 lettres (20) pour chacun des 22 sujets considérés. M' = la moyenne non cumulée d'une classe.

Il y a bien une baisse des bonnes réponses comme le supposait l'hypothèse. Contrairement à notre attente, une chute brusque se produit, entre les 25 premières lettres et les suivantes. Il ne faut pas perdre de vue que les classes (de 5 en 5 questions) sont cumulées. Si l'on veut connaître la facilité moyenne (M') de chaque groupe successif de 5 lettres, il suffit d'appliquer la formule :

$$M' = \frac{M_{i+1}(n+m) - n M_i}{m}$$

dérivée de $M_{i+1} = \frac{n\, M_i + m\, M'}{(n+m)}$

où M' = moyenne d'une classe donnée

M_i = moyenne de l'ensemble cumulé précédent cette classe (ou l'excluant)

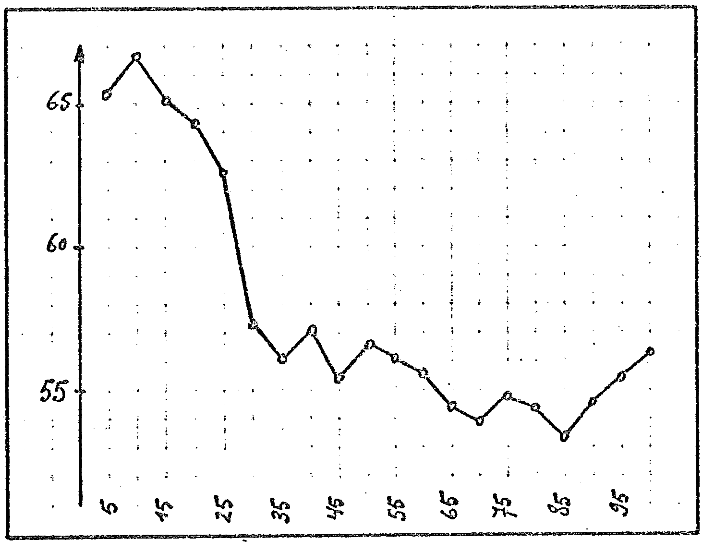
M_{i+1} = moyenne de l'ensemble (cumulé) incluant cette classe

n = effectif sur lequel M_i est calculé

m = effectif sur lequel M_{i+1} est calculé.

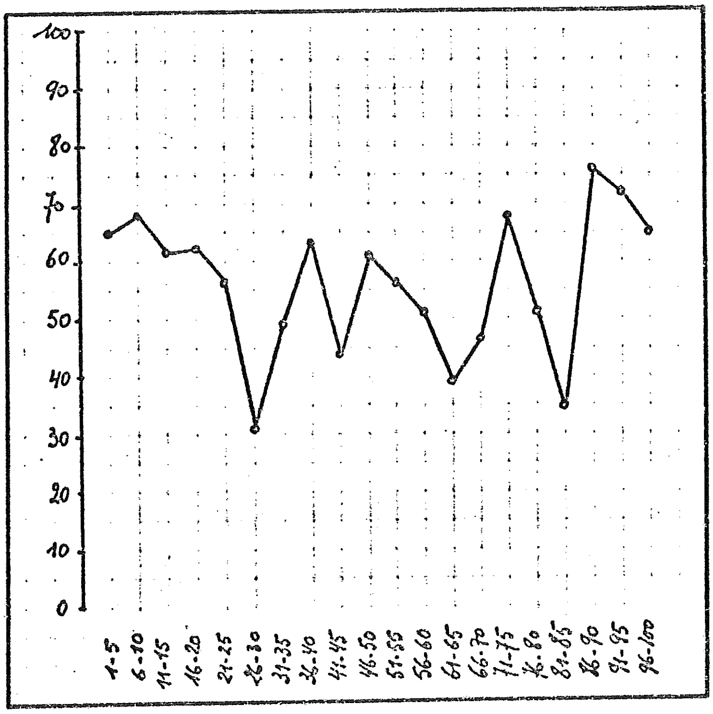
Les M' ont été portées en regard des M. Voici le graphique :

Tableau 6.56



Graphique des M du tableau 6.55

Tableau 6.57



Graphique des M' du tableau 6.55

Ainsi donc, la facilité moyenne des classes succesives de 5 questions ne décroît pas comme on aurait pu le supposer. En fait, elle dépend essentiellement du contenu : certaines questions sont très difficiles, d'autres très faciles.

La relation entre la difficulté de chaque classe de 5 questions et leur ordre est proche de 0. Ce fait renforce l'hypothèse selon laquelle la chute de corrélation (de .20 environ) entre les cinq premières questions et l'ensemble du test est liée à la fatigue ou la diminution d'application plutôt qu'à une ambiguïté croissante du texte et des questions. (Voir le taux de réussite de chacune des 100 questions.)

L'hypothèse de la diminution d'application est encore renforcée par le fait que, pour 64 % des sujets (78 sur 121), le total des points était meilleur lors du test que lors du retest.

H. LES REPONSES DES SUJETS AU QUESTIONNAIRE

Avec le test, un questionnaire avait été présenté aux sujets. Voici les quatre questions posées :

LISEZ D'ABORD TOUTES LES QUESTIONS AVANT DE REPONDRE.

QUESTION 1 : Entourez la lettre qui décrit le mieux votre comportement. Les commentaires et les précisions sont les bienvenus.

Quand j'indique le degré de ma certitude...

... Je pars mentalement des pourcentages fins (ex. 60 %), puis je remonte vers les échelles plus grossières
A - à 10 et 4 degrés indifféremment
B - à 10 puis à 4 degrés
C - à 4 puis à 10 degrés
D - autrement : ....

... Je pars mentalement de l'échelle à 10 degrés puis je réponds
E - indifféremment à 4 et 40 degrés
F - à 4 puis à 40 degrés
G - à 40 puis à 4 degrés
H - autrement : ....

... Je pars mentalement de l'échelle à 4 degrés (grossière) puis je descends vers les échelles plus fines
I - à 10 et 40 degrés indifféremment
J - à 10 puis à 40 degrés
K - à 40 puis à 10 degrés
L - autrement : ....

Je procède autrement : ....

....

....

QUESTION 2 : Entourez la lettre qui décrit le mieux votre comportement.

Le choix de mon indice de certitude dépend ...

A. ... uniquement du pourcentage exact de chances que j'attribue à ma réponse d'être correcte.

... uniquement du risque associé à la certitude :

Je préfère courir un risque...

B. ... petit : approximativement une certitude de ...... %

C. ... moyen : approximativement une certitude de ...... %

D. ... grand : approximativement une certitude de ...... %

... du pourcentage et du risque couru,

E. mais plus du pourcentage

F. mais plus du risque

G. chacun ayant une importance équivalente.

QUESTION 3 : Quelle est votre opinion sur la finesse (nombre de degrés) des trois échelles ?

L'échelle à 4 degrés est-elle	assez fine ?	OUI / NON
	trop fine ?	OUI / NON
L'échelle à 10 degrés est-elle	assez fine ?	OUI / NON
	trop fine ?	OUI / NON
L'échelle à 40 degrés est-elle	assez fine ?	OUI / NON
	trop fine ?	OUI / NON

QUESTION 4 : Selon vous, quelle serait l'échelle idéale ? ....
....
....

*

* *

Au total, 138 questionnaires nous ont été renvoyés. Des notations figuraient aux endroits réservés pour les réponses ouvertes et ailleurs. La longueur de ces réponses varie très fort. Seules les interventions les plus révélatrices seront reprises. Nous soulignerons dans le texte des sujets les éléments les plus importants.

Les questions 3 et 4 font double emploi. Une plus grande importance sera accordée à la question 4, car, dans leurs réponses à la question 3, plusieurs sujets se sont montrés incohérents.

QUESTION 1.

Quand j'indique ma certitude, je pars mentalement de l'échelle

à 40 degrés	36
à 10 degrés	34
à 4 degrés	69
	169

On pourrait interpréter ce résultat en faisant appel au principe de Decroly : "Nous pensons globalement et nous n'analysons que plus tard." Cette interprétation est renforcée par le fait que, parmi ceux qui "partent mentalement de l'échelle à 4 degrés", 81 % passent ensuite à l'échelle à 10 degrés et enfin à l'échelle à 40 degrés. Voici, d'ailleurs, la répartition des réponses :

Partent de 40 degrés	puis 10 et 4 indifféremment	A	7	(20 %)	
	puis 10, puis 4	B	18	(51 %)	36
	puis 4, puis 10	C	10	(28 %)	(25,53 %
	autrement	D	1	(3 %)	
Partent de 10 degrés	puis 40 et 4 indifféremment	E	6	(18 %)	
	puis 4, puis 40	F	11	(32 %)	34
	puis 40, puis 4	G	15	(44 %)	(24,11 %
	autrement	H	2	(6 %)	
Partent de 4 degrés	puis 10 et 40 indifféremment	I	9	(13 %)	
	puis 10, puis 40	J	56	(81 %)	69
	puis 40, puis 10	K	3	(4 %)	(48,94 %
	autrement	L	1	(1,5 %)	
Autrement		M	2	(1,42 %)	
			141		

L'interprétation avancée ci-dessus (approche globale puis analytique) doit cependant être nuancée par la très pertinente remarque du sujet 88 (S.88) : "La disposition typographique des trois échelles pousse le lecteur à aller de haut en bas, de 4 à 40 degrés."

La façon de procéder peut résulter du contenu, comme le suggère le sujet 89 (S.89) : "Le nombre de possibilités dicte parfois l'échelle, ainsi, si je vois quatre possibilités, également probables, 25 % s'imposent (donc l'échelle C)."

Certains sujets ont d'autres comportements :

- "Quand je suis tout à fait certaine, je pars de 100 %, et quand je ne le suis pas du tout, je pars de 0 %." (S.193)
- "Je pointe à un certain niveau sur un segment de droite, sans tenir compte d'une échelle quelconque: '———X——'."(S.199)
- "Pour moi, il y a trois situations : vrai, faux ou ?. Je choisis ensuite le pourcentage qui me rapportera le plus et me fera perdre le moins." (S.107)
- "Au début, je répondais sur l'échelle 40, puis 10, puis 4 degrés, puis, connaissant mieux les échelles, j'utilise l'inverse." (S.77)
- "L'évaluation du pourcentage s'effectue, pour ma part, selon un modèle triadique : réponse sûre, à moitié sûre, très peu sûre. Il reste alors un certain nombre de réponses indéterminables auxquelles j'attribue bizarrement, sans raison, environ l'indice 60 (tendance à se surestimer sans doute).
  Mes réponses sûres sont alors ventilées en - totalement sûre : 100 %
  - fort probable : 95 %
  - probable : ± 80 %
  Mes réponses à moitié sûres, je les fixe à 50 %. Pour quelques exemples décisifs où l'on a une chance sur 4, je note 25 %.
  Mes réponses peu sûres : entre 0 et 10 %."
- "J'ai d'abord mis les lettres en n'indiquant que les certitudes 100 %, puis j'ai reparcouru le listing." (S.19)
- "Dans l'échelle à 40 degrés, je considère seulement le continuum, faisant abstraction des valeurs 10, 20, etc. Toutefois, on dirait que je prends des repères tels que 0,50,100." (S.28)
- "Pour les réponses dont je suis sûr, je pars de 9; pour les autres, je pars de 0 vers 4 (ceci dans l'échelle à 10 degrés)." (S.83)

QUESTION 2.

Les réponses ci-dessous sont énumérées dans l'ordre d'importance où elles ont été fournies par les sujets.

Le choix de mon indice de certitude dépend...

39 (27,8 %)	E.	du pourcentage et du risque couru, mais plus du pourcentage.
36 (25,7 %)	A.	uniquement du pourcentage exact de chances.
26 (18,6 %)	G.	du pourcentage et du risque, chacun ayant une importance équivalente.
13 (9,3 %)	F.	du pourcentage et du risque couru, mais plus du risque.
26 (18,5 %)	B + C + D :	uniquement du risque associé à la certitude.
140		

Je préfère courir un risque...

8	B.	petit (de 10 à 30 %, Mé = 20)
13	C.	moyen (de 25 à 80 %, Mé = 50)
5	D.	grand (de 75 à 100 %, Mé =90)

Le taux élevé de réponses A, E et G (72,1 %) est rassurant quant à l'efficacité de la procédure et aux conditions expérimentales. Les réponses B, C et D, elles, sont des applications (modérées) du critère de HURWICZ. Un sujet a cependant nuancé sa réponse B (Je préfère courir un risque petit) par : "dans le doute absolu" (25 %). On ne saurait donner tort à ce sujet (S.250).

Certaines réponses ouvertes sont intéressantes :

. "Quand la réponse est une vraie devinette, je divise mes 100 % de certitude par le nombre d'éventualités (Exemple : $1 \left\{ \begin{array}{l} e \\ a \\ -\ (') \end{array} \right.$) mais j'y ajoute parfois quelques pourcents subjectifs (exemple : le est plus probable car il peut donner le ou les)." (S.204)

. "Dans ce jeu, je n'ai tenu compte que du pourcentage de chances, mais mon comportement serait différent s'il s'agissait de la vie d'un individu, de finances, de l'obtention d'une promotion. Dans ce cas, je choisirais après avoir lu le tableau de gains et pertes (je trouve le vôtre fort discutable)." (S. 171)

. "Quatre degrés ne me permettent pas de faire jouer le risque. Quarante degrés sont illusoires... A cause des gains et pertes, et surtout de leur différence, je me cantonne entre 3 et 7 (dans l'échelle à 10 degrés)." (S.306)

. "Au-delà de 70 %, le risque de perdre des points est trop grand par rapport au risque d'en gagner. Je préfère 70 % plutôt que 100 %." (S.203)

. "Je ne réponds jamais plus de 92,5 % car le gain ne compense pas la perte quand on va plus haut." (S.43)

Les trois dernières remarques posent le problème de la consigne et de la familiarité avec le tableau de gains et pertes.

QUESTION 3.

La question 3, redondante par rapport à la question 4, a été l'objet de réponses inconséquentes chez plusieurs sujets. Voici les résultats bruts :

	ECHELLE à 4 degrés		ECHELLE à 10 degrés		ECHELLE à 40 degrés	
	assez fine	trop fine	assez fine	trop fine	assez fine	trop fine
Oui	26	3 (2)	97	8	85	71
Non	92	67	17	65	4 (3)	23

La disparité des réponses vient du fait que de nombreux sujets n'ont répondu qu'à certaines des questions. La tendance générale est :

- 4 degrés ne sont pas assez fins.
- 10 degrés sont assez fins, mais pas trop fins.
- 40 degrés sont assez fins pour la majorité et trop fins pour une importante minorité.

Les réponses à la question 4 sont plus claires.

QUESTION 4.

Le nombre idéal de degrés est :

4 degrés	3	... étant donnée la grande part de subjectivité (S.1)
5 degrés	1	... les 5 degrés seraient 0, 25, 50, 75 et 100 (S.36)
6 degrés	1	
7 degrés	1	
8 degrés	1	
7 à 10 deg.	1	
4 à 10 deg.	1	
10 degrés	41	. Mais avec possibilité d'indiquer dans quelle partie de l'intervalle on se place (exemple : 7+ et 7-. On se rapproche évidemment de l'échelle à 40 degrés (S.3) . Avec ± 1 degré (S.171). . Avec possibilité de finesses (S.228) . Avec possibilité de se situer entre deux divisions, parfois. Ex. 25 %. (S.180) . Quand on doit répondre à 100 questions, 10 degrés suffisent. Si on n'avait que 5 questions, 40 degrés se justifieraient. (S. 173).

		. Dans l'application scolaire (S.194) . Pourcentages arrondis aux dizaines (S.241) . Partir de l'échelle à 10 degrés résulte sans doute de l'habitude professionnelle de la cotation sur 10 (S.262) . On ne prend pas la peine de calculer l'incidence sur le résultat d'une approximation plus détaillée. Et, dans le doute, vu l'absence de responsabilité, on ne regarde pas à un dixième près (S.258)
10 à 20 degrés	3	
20 degrés	13	. Parce que c'est celle que j'utilise moi-même couramment (S.192) . De 0 à 100 avec graduation tous les 5 % (S.309) . 1 à 20 (S.66)
40 degrés	8	
50 degrés	2	
100 degrés	4	. "Pourquoi faut-il ces trois échelles ? Ne serait-il pas plus simple de dire : ma réponse vaut 36, 42, 58,87 ou 99 ? (S.26)

Quelques remarques supplémentaires.

- "Malgré tout, ce jeu des certitudes peut s'apprendre et le sujet peut s'y perfectionner, c'est-à-dire changer d'échelle (S.148)."
- "J'ai l'impression qu'on peut affiner sa certitude." (S.163)
- "Ce test sur la langue est irritant pour moi, licencié en philologie romane : la manière de répondre que vous imposez est... frustrante ! Pour chaque cas, en effet, une analyse minutieuse (que je n'ai pas faite et qui d'ailleurs ici ne servirait pas) permettrait de chiffrer avec exactitude le nombre de réponses impossibles et, sans doute, le pourcentage de solutions possibles. Or vous exigez une seule solution et, pour cette solution unique, un pourcentage de chance. En résumé,... les processus de réflexion sont rendus inutiles, au moins en partie."(S.282)
- "Comprenez le caractère éprouvant du jeu : entre cinq solutions aussi vraisemblables les unes que les autres, vous devez en choisir une au hasard, en lui attribuant un indice de certitude." (S.282)

Les deux remarques du sujet 282 vont dans le sens d'un raisonnement qui a été développé dans l'introduction à propos du statut politique de la Malaisie en 1939. Le sujet S.282 regrette de ne pouvoir fournir l'ensemble de ses réponses (situation stigmatisée comme celle de l'expérimentateur C), mais de devoir se limiter à une seule réponse avec certitude (expérimentateur B). Qu'aurait dit le sujet S.282 si on ne lui avait même pas permis de fournir sa certitude (expérimentateur A) ?

Le sujet S.282, par ses remarques, évoque des expériences complémentaires plus complexes. Il va même jusqu'à proposer certaines procédures : "Ne pourrait-on imaginer qu'une question exige une échelle d'évaluation particulière ? Exemple : IL EST TR. Au maximum, 6 lettres (a, e, i, o, u, y) sont possibles sur 26, ce qui restreint considérablement l'échelle des probabilités de 40 degrés. On peut donc, dans ce cas, pénaliser fortement tous ceux qui répondent avec un indice de certitude inférieur à 1/6."

o

o o

Conclusions sur les réponses à la question 4 (échelle idéale).

Un certain nombre de sujets ont, à juste titre selon nous, estimé qu'il n'existe pas d'échelle idéale, mais que cela dépend :

- du nombre de questions (S.173);
- du type de jeu (S.171), du type de question (S.218), des matières (S.218)
- de dispositions d'esprit momentanées (S.148);
- de la finesse d'esprit du sujet. S'il éprouve déjà de fortes difficultés à répondre, combien plus lui sera-t-il difficile de "se doser" (S.148)
- du thème abordé par le sujet (un scientifique peut avoir une intuition particulière dans sa spécialité sans pour autant réagir aussi bien ailleurs) (S.148).

I. CONCLUSIONS

A. Conclusions sur la procédure.

La fidélité observée est fonction de la situation de testing. Une des préoccupations de notre travail est d'estimer la fidélité maximale que peuvent atteindre, en moyenne, les individus d'un âge et d'un niveau intellectuel donnés, dans une situation idéale.

Nous pensons que la présente procédure ne représente pas une telle situation idéale et que, par conséquent, les fidélités observées n'atteignent pas les fidélités maximales possibles. C'est pourquoi les conclusions porteront autant sur la procédure que sur les résultats.

La présente étude ne fournit aucune information sur les variations de fidélité inter et intra-individus pour une même épreuve et pour des épreuves différentes.

AVANTAGES :

1. L'application est collective.

   On pourrait même imaginer des réponses "en groupe". Le groupe permettrait peut-être d'enrichir les hypothèses et de nuancer les probabilités.

2. Le contenu est inépuisable et facile à manipuler, par exemple pour l'adapter aux âges, aux intérêts et à la compétence de la population visée.

3. Un très grand nombre de questions peut être posé sans que les sujets doivent se replonger dans un tout autre contexte sémantique : il s'agit toujours du même thème littéraire.

4. Le contenu est proche de la situation pédagogique la plus courante.

   D'une part, une (et une seule) réponse est exacte. D'autre part, il n'existe aucun moyen d'estimer objectivement la PER des différentes réponses possibles (ici n = 27). En effet, aux contraintes morphologiques, syntaxiques s'ajoutent des phénomènes sémantiques et stylistiques, sans parler de l'arbitraire de l'auteur.

5. La facilité moyenne de l'épreuve peut être fixée à volonté.

   Bien entendu, la facilité visée était 50 %. La facilité observée fut de 56 %. Un prétest eût sans doute permis un meilleur ajustement. L'écart-type des $\mathcal{M}_{|i}$ ER (moyennes de réussite des sujets) est faible (5,31), ce qui constitue un avantage. L'écart-type des facilités des questions est plus élevée ($\mathcal{M}_{|q}ER$), ce qui est un autre avantage.

6. Le gain d'information entre le pré et le post-test est nul, sans que des précautions particulières doivent être prises.

7. Le codage de la réponse pour le sujet comme pour l'expérimentateur, ainsi que la correction par ordinateur sont aisés.

INCONVENIENTS :

1. Cette procédure permet, suscite même, beaucoup trop de révision des hypothèses. Il serait vraisemblablement possible de recueillir des indices de fidélité plus élevés grâce à certaines modifications suggérées ci-avant.

2. La stabilité décroît avec le nombre de questions. Nous pensons cependant que la diminution est contenue dans des limites raisonnables quand on considère le nombre important de questions.

B. Conclusions sur les résultats.

1. Pour les adultes concernés et pour cette tâche, l'échelle à quatre degrés est trop peu sensible et l'échelle à quarante degrés trop sensible. Tous les échelons de l'échelle à dix degrés ne sont pas également attractifs. L'échelle idéale(1) comprendrait entre sept et dix degrés, mais quelle devrait être sa forme ? Doit-on faire choisir des espaces (entre 70 et 80 %) ou des points fixes (75 %) ? Les échelons doivent-ils rester égaux ou non ? La référence doit-elle être une échelle de probabilités exprimées en pourcentages (80 %) ou en rapports (quatre chances sur cinq) ? Autant de questions que devraient trancher des études expérimentales ultérieures. Nous espérons que la présente recherche constituera une source de réflexions pour ceux qui les entreprendront.

2. L'équation de régression (μ sur ESPER) est, à peu près, $\mu = .5$ ESPER + .25.

La corrélation est d'environ .56 (Médiane) dans des conditions assez difficiles, mais peut monter jusqu'à .78 dans des conditions plus favorables de concentration et d'attention.

(1) Pour ce type d'exercice, pour des sujets de cet âge, avec ce mode de présentation, etc.

CHAPITRE 7

LA REVISION DES ESPER

A. Le théorème de BAYES.

B. Le théorème de Bayes et les situations de la vie courante.

C. Expériences de vérification de la validité du théorème.

1. Les expériences de H. ROUANET

2. Les expériences de W. EDWARDS

D. Un jeu des ESPER propre à l'étude du théorème.

E. Résultats expérimentaux.

F. Une expérience en situation scolaire : le test relatif à l'usage du dictionnaire.

Dans les chapitres précédents, il importait de mesurer des <u>états</u> de la connaissance d'un individu sur un contenu. Dans le présent chapitre, on étudiera le problème de la <u>modification</u> de la connaissance à la suite d'un événement. Selon les cas, cet événement est appelé <u>stimulus</u>, <u>message</u>, <u>information</u>, etc.

Ici encore, les recherches menées dans le cadre de la théorie des décisions sont d'un précieux secours. Mis au point pour les sciences économiques, un outil comme le théorème de BAYES (voir ci-après) a déjà prouvé son intérêt en psychologie (1), pour rendre compte des comportements de choix.

Cette méthodologie va maintenant être explicitée; elle sera ensuite appliquée dans les expériences.

°
° °

La connaissance d'un individu sur une question précise est représentée par <u>les</u> ESPER que cet individu accorde à <u>chacune</u> des réponses possibles à la question. La connaissance est parfaite quand le sujet attribue à la solution correcte une ESPER égale à 1 et aux solutions incorrectes une ESPER égale à 0.

On distingue la connaissance du sujet mesurée <u>avant</u> l'information, de la connaissance du même sujet mesurée <u>après</u> cette information. Dans le théorème de Bayes, ces deux mesures sont appelées respectivement "probabilités *a priori*" et "probabilités *a posteriori*" (2).

---

(1) Voir les travaux de W. EDWARDS (1967), H. ROUANET (1961).

(2) Dans le cas présent, les expressions suivantes (au sens plus restreint) seraient préférables : "ESPER a priori" et "ESPER a posteriori".

Dans cette perspective, il est possible de donner une définition psychologique opérationnelle (proche de celle du terme stimulus) non pas au terme information (1) en général, mais à l'information pour un individu donné à un moment donné:

> "Tout événement qui, consciemment perçu ou non par l'individu, modifie ses ESPER."

Dans les pages qui suivent, on supposera, pour la facilité des raisonnements, que les certitudes exprimées représentent sans biais les ESPER. Les précautions méthodologiques mises en lumière dans les chapitres précédents seront appliquées dans la mesure du possible. Toutes les nuances relatives à la validité, la fidélité et la sensibilité des ESPER restent d'application, mais nous ne les soulignerons plus.

°

° °

Le problème de la modification des ESPER peut se poser de deux façons :

1) La connaissance *a priori* et la valeur informative de l'événement-stimulus sont données. On cherche à prédire quelle sera la connaissance *a posteriori*.

2) Quand la connaissance *a priori* et la connaissance *a posteriori* sont données, on cherche à déduire quelle était la valeur informative de l'événement-stimulus (ou message).

Dans ces deux situations, il s'impose de recourir au théorème de Bayes.

---

(1) PIERON H., Vocabulaire de la Psychologie, Paris, P.U.F., 1973, p. 220.

information

Les problèmes de télécommunication — téléphonie, télévision — et de construction de machines à fonctions cybernétiques ont conduit à désigner du terme général d'information un ensemble de données élémentaires fournies à un être vivant, et particulièrement à l'Homme (sensations), ou à certaines machines, par le milieu extérieur; les éléments *(logon)*, répondant à des variables de *tension*, de *position* et de *durée*, exprimables en unités quantiques (correspondant à des seuils différentiels) peuvent être codifiés suivant la théorie de Shannon où l'on appelle information ce qu'une communication apporte de nouveau (d'imprévisible) au récepteur. On mesure la quantité d'information i par l'opposé du logarithme de la probabilité p du signal transmis : $i = -\log p$. Les probabilités indépendantes se multipliant pour donner la probabilité composée, on a ainsi des quantités i qui s'ajoutent. On a choisi 2 comme base des log. afin que $i = 1$ pour $p = 1/2$ (par ex. quand on apprend que c'est pile ou face). Cette unité de i est le bit (v. ce mot). En bits, $i = \log_2 n$, n $(= 1/p)$ étant le nombre de cas (équiprobables) possibles. Communiquer 1 signal sur 32 $(= 2^5)$ possibles apporte 5 bits. Pour un code chiffré de 1 à 32, communiquer « nombre pair » ou « > 16 » n'apporte que 1 bit. D. et P.

information (dégradation d'—). La théorie de l'information, en établissant une mesure de celle-ci, permet d'apprécier la plus ou moins grande fidélité d'une transmission, par exemple au cours des communications téléphoniques. Lorsqu'il y a *dégradation de l'information*, il subsiste une incertitude, un signal ayant été reçu, quant au signal qui a été émis. On peut mesurer en bits la moyenne de cette incertitude pour l'ensemble des signaux. Elle est alors désignée par les expressions : *entropie conditionnelle*, *equivocation* (anglais), *entropie relative*, *information dégradée*. La différence entre l'information émise (entropie de la source) et l'information dégradée fournit la *quantité d'information transmise* qui s'exprime en bits ou en pourcentage de l'information émise. Cette quantité est souvent désignée par le symbole R (en français) ou T (en anglais) et elle peut être considérée comme une généralisation de la notion de *corrélation* entre deux variables. V. **entropie, canal de communication, redondance.** M. et R.

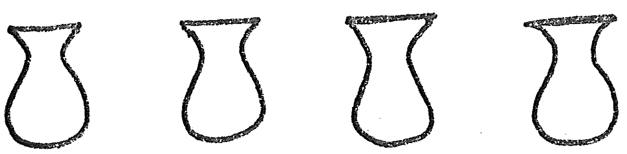
A. LE THEOREME DE BAYES

Le théorème du révérend anglais Thomas BAYES (1763) est consacré au problème de la révision des "probabilités subjectives" après information (1). Ce théorème, outil simple et puissant, a été très peu utilisé en pédagogie pour la simple raison qu'il repose sur les ESPER, jusqu'ici à peine étudiées. Les présentations françaises du théorème sont rares et les exposés non-mathématiques le sont plus encore. C'est pourquoi nous croyons devoir expliquer assez longuement ce théorème (2).

*

* *

Examinons d'abord un exemple inspiré de D. LINDLEY (1971, p. 98). On présente à un sujet quatre urnes identiques extérieurement.



A B C D

On informe le sujet qu'une de ces urnes (l'urne A) contient 2/3 de billes rouges et 1/3 de billes noires; les autres contiennent des billes dans des proportions inverses. On présente au sujet une de ces quatre urnes, choisie au hasard et on pose la question suivante : "Quelle est la

---

(1) On trouvera, dans JACKSON (1975), des commentaires sur la place de ce théorème dans l'histoire du développement de la statistique en Angleterre, jusqu'à R.A. FISCHER.

(2) Nous sommes reconnaissant au professeur BRENY et à D. DEFAYS des éclaircissements qu'ils nous ont fournis sur certains points.

probabilité que cette urne soit l'urne "rouge" (A) ?" *A priori*, cette probabilité est 1/4 puisque les trois urnes non présentées ont autant de chances d'être l'urne rouge (A). C'est cette probabilité que fourniraient en moyenne les sujets, en réponse à la question ci-dessus.

Prob. *a priori* d'être l'urne A	= p (Urne A $\mid$ H) = 1/4 = 0,25 = 1 chance contre 3.

où la barre verticale exprime la condition : "Etant donné H" et où H signifie : l'état cognitif du sujet avant information.

Une information va maintenant être introduite dans l'expérience en permettant au sujet d'extraire une bille au hasard de l'urne présentée. Si cette dernière est l'urne rouge (A), la probabilité d'extraire une bille rouge (X) est 2/3. Si le sujet est autorisé à faire plusieurs tirages successifs (avec remplacement), il va acquérir une information considérable. Cette information sera conventionnellement désignée par X.

Si, par exemple, on a extrait 33 billes dont 20 rouges et 13 noires de l'urne, quelle est la probabilité que ce soit l'urne rouge ? LINDLEY signale que la plupart des personnes interrogées évaluent cette probabilité à 9 chances contre une (.9), alors que la probabilité, calculée par le théorème de Bayes est de 99 chances contre une (exactement .992). Cet exemple montre que le théroème tire mieux parti de l'information disponible que ne le font les sujets humains.

Le théorème de Bayes est fondé sur les trois axiomes classiques (1), (2) et (3), de la théorie des probabilités et sur la définition (4) des probabilités conditionnelles (cf. COOMBS, DAWES et TVERSKY, 1970, p. 386).

Nous notons S l'ensemble des possibilités, $\cup$ la réunion, $\cap$ l'intersection et $|$ la condition.

(1) Pour tout événement E, $0 \leq p\ (E) \leq 1$

(2) $p\ (S) = 1$

(3) Pour tous événements E, F, tels que $E \cap F = \emptyset$,
$p\ (E \cup F) = p\ (E) + p\ (F)$

(4) $p\ (E | F) = p\ (E \cap F) / p\ (F)$ = p jointe / p du conditionnant.

Ainsi, si $E_1, E_2, \ldots E_n$ sont tels que $\cup\ Ei = S$ (exhaustifs) et $E_i \cap E_j = \emptyset$ (exclusifs), on peut montrer que :

(5) $$\boxed{p\ (E_i | X) = \frac{p\ (Ei)\ p\ (X | Ei)}{p\ (X)}}$$

où p (Ei) = la probabilité *a priori* d'apparition de l'événement i (Ei). Dans nos expériences, la probabilité d'apparition sera remplacée par une ESPER. Néanmoins, la notation p, très générale, sera gardée.

X = l'information ou "la situation cognitive du sujet après information"

$p\ (Ei | X)$ = la probabilité *a posteriori* de Ei (après information).

$p\ (X | Ei)$ = la probabilité de X si Ei est vrai. Cette expression est aussi appelée "vraisemblance" de Ei étant donné l'information X".

On écrit parfois le théorème de Bayes comme suit :

(6) $$\boxed{p\ (Ei | X) \propto p\ (Ei)\ p\ (X | Ei)}$$

où $\propto$ signifie proportionnel à.

Trop de je
dans l'expression

Ainsi, on peut écrire, avec D. LINDLEY (1969, p. 31) que la probabilité a posteriori d'un événement est proportionnelle au produit de la probabilité a priori et de la vraisemblance de cet événement.

Comme les Ei sont exclusifs et exhaustifs, on peut montrer que

$$(7) \quad \boxed{p\,(X) = \sum_{k=1}^{n} \left[\, p\,(X \mid Ei)\; p\,(Ei) \,\right]}$$

On rencontre souvent le théorème sous les formes suivantes, où H désigne la situation cognitive du sujet avant information :

$$(5') \quad p\,(Ei \mid X \cap H) = p\,(Ei \mid H)\; p\,(X \mid Ei \cap H) \;/\; p\,(X \mid H)$$

$$(6') \quad p\,(Ei \mid X \cap H) \propto p\,(Ei \mid H)\; p\,(X \mid Ei \cap H)$$

Nous utiliserons parfois ce type de notation.

La formule (7) montre que $p\,(X \mid H)$ peut facilement être obtenu à partir des autres données, mais que l'inverse n'est pas vrai (La somme ne permet pas de connaître chacun des termes). Par conséquent, la formule (6) décrit les données de base dont il faut disposer pour faire fonctionner le théorème.

Revenons à l'exemple des quatre urnes A, B, C et D, remplies de billes. L'urne A est rouge (c'est-à-dire composée de 2/3 de billes rouges et d'1/3 de noires). Les trois autres urnes sont noires (2/3 de billes noires, 1/3 de rouges). On présente (au hasard) une des quatre urnes à un sujet.

La probabilité *a priori* que l'urne soit l'urne rouge (A) vaut 1/4 : $p\,(A \mid H) = .25$

La probabilité *a priori* que l'urne soit noire (B, C ou D) vaut 3/4 : $p\,(B \mid H) = .25$
$p\,(C \mid H) = .25$
$p\,(D \mid H) = .25$

Ainsi, on peut écrire, avec D. LINDLEY (1965, p. 30) que la probabilité a posteriori d'un événement est proportionnelle au produit de la probabilité a priori et de la vraisemblance de cet événement.

Comme les Ei sont exclusifs et exhaustifs, on peut montrer

$$(7) \qquad \boxed{p(X) = \sum_{k=1}^{n} \left[p(X \mid E_k)\, p(E_k) \right]}$$

On rencontre souvent le théorème sous les formes suivantes, où H désigne la situation cognitive du sujet avant information :

$$(5') \quad p(E_i \mid X \cap H) = p(E_i \mid H)\, p(X \mid E_i \cap H) \,/\, p(X \mid H)$$

$$(6') \quad p(E_i \mid X \cap H) \propto p(E_i \mid H)\, p(X \mid E_i \cap H)$$

Nous utiliserons parfois ce type de notation.

La formule (7) montre que $p(X \mid H)$ peut facilement être obte à partir des autres données, mais que l'inverse n'est pas vrai (la somme ne permet pas de connaître chacun des termes). Par conséquent, la formu (6) décrit les données de base dont il faut disposer pour faire fonctionner le théorème.

Pourquoi introduit-on cette probabilité de X si on ne s'en sert pas ?

Revenons à l'exemple des quatre urnes A, B, C et D remplies de billes. L'urne A est rouge (c'est-à-dire composée de 2/3 de billes rou et d'1/3 de noires). Les trois autres (B, C et D) sont noires (2/3 de billes noires, 1/3 de rouges). On présente (au hasard) une des quatre urnes au sujet.

La probabilité a priori que l'urne soit l'urne rouge (A) vau 1/4 : $p(A \mid H) = .25$

La probabilité a priori que l'urne soit noire (B, C ou D) vau 3/4 : $p(B \mid H) = .25$
$p(C \mid H) = .25$
$p(D \mid H) = .25$

On permet au sujet d'effectuer dix tirages (avec remplacement) de l'urne qui lui est soumise. <u>Il obtient sept billes rouges et trois noires</u>. Dans cet exemple, il est possible de calculer directement $p\ (X \mid E_i \cap H)$ qui est "la probabilité de tirer au moins sept billes rouges sur dix grâce au seul hasard, si l'urne est vraiment rouge (ou urne A).

La probabilité de tirer <u>exactement</u> 7 billes rouges sur 10 dans une urne rouge (2/3 r - 1/3 n) se calcule comme suit :

$$\frac{10\,!}{3\,!\;7\,!}\ (1/3)^3\ (2/3)^7 = 120 \times .0370370 \times 058528 = .26012$$

La probabilité de tirer <u>exactement</u> 8 billes rouges sur 10 vaut :

$$\frac{10\,!}{2\,!\;8\,!}\ (1/3)^2\ (2/3)^8 = 45 \times .111111 \times .0390186 = .19509$$

La probabilité de tirer <u>exactement</u> 9 billes rouges sur 10 vaut :

$$\frac{10\,!}{1\,!\;9\,!}\ (1/3)\ \ (2/3)^9 = 10 \times .333 \times .0260124 = .08670$$

La probabilité de tirer <u>exactement</u> 10 billes rouges sur 10 vaut .01734.

La probabilité de tirer <u>au moins</u> 7 billes rouges vaut:

.26012 + .19509 + .08671 + .01734 = .55926.

La probabilité de tirer 7 billes rouges <u>au plus</u> vaut
1 - .55926 + .26012 = .70086.

Cette probabilité peut aussi être consultée directement dans une table appropriée, comme les <u>Tables of the cumulative binomial probability distribution</u> (1) dont nous reproduisons la page 216 ci-dessous (tableau 7.0).

n	r	p=0.26	p=0.27	p=0.28	p=0.29	p=0.30	p=0.31	p=5/16	p=0.32	p=0.33	p=1/3
10	0	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
	1	0.95076	0.95702	0.96256	0.96745	0.97175	0.97554	0.97641	0.97886	0.98177	0.98266
	2	0.77776	0.79807	0.81696	0.83449	0.85069	0.86564	0.86918	0.87938	0.89199	0.89595
	3	0.50422	0.53351	0.56217	0.59010	0.61722	0.64344	0.64985	0.66872	0.69300	0.70086
	4	0.24793	0.27258	0.29794	0.32392	0.35039	0.37724	0.38400	0.40436	0.43163	0.44074
	5	0.09035	0.10368	0.11812	0.13365	0.15027	0.16795	0.17253	0.18666	0.20635	0.21313
	6	0.02391	0.02872	0.03420	0.04039	0.04735	0.05511	0.05718	0.06371	0.07320	0.07656
	7	0.00446	0.00562	0.00700	0.00865	0.01059	0.01286	0.01349	0.01550	0.01855	0.01966
	8	0.00056	0.00074	0.00096	0.00124	0.00159	0.00202	0.00214	0.00254	0.00317	0.00340
	9	0.00004	0.00006	0.00008	0.00011	0.00014	0.00019	0.00020	0.00025	0.00033	0.00036
	10	0.00000	0.00000	0.00000	0.00000	0.00001	0.00001	0.00001	0.00001	0.00002	0.00002

(1) Cambridge, Massachussetts, Harvard University Press, 1955.
N.B.: Cette table ne comprenant que des proportions $0 \leqslant p \leqslant .50$, il est fréquent de devoir consulter la table à n - r au lieu de r.

Un rapide calcul permet d'obtenir chaque valeur :

p (exactement 7 billes rouges) = .70086 - .44074 = .26012
p (exactement 8 billes rouges) = .89595 - .70086 = .19509
p (exactement 9 billes rouges) = .98266 - .89595 = .08671
p (exactement 10 billes rouges) = 1. - .98266 = .01734

.55926

Ce total correspond bien à 1 - .44074.

On peut démontrer, de la même manière, que la probabilité de tirer au moins 7 billes rouges si l'urne est noire (B, C ou D) vaut .01966 (cf. table).

Revenons à présent au théorème.

La probabilité du tirage (7 rouges et 3 noires) ou probabilité de l'information vaut .70086 si l'urne est "rouge" (A)
.01966 si l'urne est "noire" (B, C ou D).

$p\ (X \mid A \cap H) = .70086$
$p\ (X \mid B \cap H) = .01966$ (idem pour C et pour D)

La probabilité *a posteriori* que l'urne considérée soit l'urne A peut s'exprimer comme suit :

$$p\ (A \mid X \cap H) = .25 \times .70086 / p\ (X \mid H)$$

Calculons maintenant la valeur de $p\ (X \mid H)$, seule inconnue du problème. Rappelons-nous que l'axiome 2 (voir plus haut) veut que $p\ (A) + p\ (B) + p\ (C) + p\ (D) = 1$; autrement dit, la somme des probabilités des divers événements vaut 1, qu'il s'agisse de probabilités *a priori* ou *a posteriori*.

Donc,

$$\begin{array}{rl} & p\ (A \mid X \cap H) = .25 \times .70086 / p\ (X \mid H) \\ + & p\ (B \mid X \cap H) = .25 \times .01966 / p\ (X \mid H) \\ + & p\ (C \mid X \cap H) = .25 \times .01966 / p\ (X \mid H) \\ + & p\ (D \mid X \cap H) = .25 \times .01966 / p\ (X \mid H) \\ \hline & 1 \qquad\qquad\qquad 1 \end{array}$$

Or, si les membres de gauche des égalités doivent satisfaire à l'axiome 2 (somme = 1), les membres de droite doivent y satisfaire également. Pour faciliter l'écriture, remplaçons $p\ (X \mid H)$ par k.

$$\begin{aligned} 1 &= (.25 \times .70086 / k) + 3\ (.25 \times .01966 / k) \\ &= .1752 / k + 3\ (.0049 / k) = (.175215 + .014745) / k \\ &= .18996 / k \end{aligned}$$

$$\boxed{k = .18996}$$

Ce facteur k peut être considéré comme un coefficient de proportionalité puisqu'il affecte également toutes les éventualités.

On peut maintenant calculer les <u>probabilités *a posteriori*</u>.

$$p\,(A \mid X \cap H) = (.25 \times .70086) / .18996 = .92237$$
$$p\,(B \mid X \cap H) = (.25 \times .01966) / .18996 = .02587$$
$$p\,(C \mid X \cap H) = \qquad .02587$$
$$p\,(D \mid X \cap H) = \qquad .02587$$

$$.99998 \simeq 1.$$

On voit que les $p\,(Ei \mid X)$ n'ont pas été calculés par la formule (5) mais la formule (8) :

$$(8) \qquad \boxed{p\,(Ei \mid X) = \frac{p\,(Ei)\; p\,(X \mid Ei)}{\sum_{k=1}^{n} p\,(E_k)\; p\,(X \mid E_k)}}$$

Cette formule est, en fait, dérivée de (5) et (7) :

$$(5) \qquad p\,(Ei \mid X) = \frac{p\,(Ei)\; p\,(X \mid Ei)}{p\,(X)}$$

$$(7) \qquad p\,(X) = \sum_{k=1}^{n} p\,(X \mid E_k)\; p\,(E_k)$$

D'autres formules permettent de calculer $p\,(Ei \mid X)$. Par exemple :

$$p\,(X \mid Ei) = \frac{p\,(X \cap Ei)}{p\,(Ei)} \qquad \text{(par (4))}$$

Par remplacement dans le numérateur de (8), on obtient :

$$p\ (Ei \mid X) = \frac{p\ (Ei)\ \dfrac{p\ (X \cap Ei)}{p\ (Ei)}}{\sum_{k=1}^{n} p\ (E_k)\ p\ (X \mid E_k)}$$

et, en simplifiant,

$$(9) \quad p\ (Ei \mid X) = \frac{p\ (X \cap Ei)}{\sum_{k=1}^{n} p\ (E_k)\ \ p\ (X \mid E_k)}$$

Les formulations (5) et (8), d'une part, et la formulation (9), d'autre part, diffèrent essentiellement sur le plan des consignes. Dans las formulations (5) et (8), on demande d'estimer des probabilités <u>conditionnelles</u>. Dans la formulation (9), on demande d'estimer des probabilités <u>conjointes</u>. Il en résulte, bien entendu, des différences importantes dans la pratique.

B. LE THEOREME DE BAYES
ET LES SITUATIONS DE LA VIE COURANTE

1. *DANS LA PRATIQUE MEDICALE ORDINAIRE.*

En France, un patient va visiter un médecin pour la première fois (exemple inspiré de D. LINDLEY, 1971, p. 104). Avant même que le patient ne soit entré dans son cabinet, le médecin a certaines idées **a priori** sur ce qui pourrait "ne pas aller" chez son visiteur. En effet, beaucoup de ses patients souffrent de maux bénins, comme le rhume; quelques-uns ont des affections plus sérieuses, comme le rhumatisme; mais très peu - du moins en Europe occidentale - souffrent d'une maladie de carence, comme le béri-béri. Ces idées peuvent être exprimées en probabilités *a priori*. Les maladies sont les événements incertains (E1, E2, E3 ou M1, M2, M3) de la situation.

Le médecin invite le patient à décrire ce qu'il ressent, pose des questions, procède à des tests et acquiert ainsi de l'information (X). Il reste au médecin à combiner ses hypothèses *a priori* et son information X pour obtenir des hypothèses *a posteriori* sur ce qui ne va pas chez son malade. Si une maladie (M1) est telle que l'ensemble des symptômes X lui est rarement associé, la vraisemblance sera faible et, par conséquent, son produit avec la probabilité *a priori* sera également faible. Aussi, le médecin pensera qu'il est peu vraisemblable que le patient souffre de cette maladie M1. De même, si la probabilité d'une maladie (M2) est faible, même si les symptômes (X) sont souvent associés à cette maladie (M2), le produit sera faible et le médecin n'attribuera pas la maladie au patient. Par exemple, il ne prendra le béri-béri en considération que si les symptômes sont très marqués. On le voit, un médecin travaille, dans une large mesure, selon les idées contenues dans le théorème de Bayes, même s'il n'en a jamais entendu parler.

2. *DANS LA PRATIQUE MEDICALE ASSISTEE PAR ORDINATEUR* (1).

Dans un proche avenir, le médecin pourrait disposer, chez lui, d'une console reliée à un ordinateur. Il introduirait, par le clavier, l'information X et l'ordinateur serait programmé pour utiliser le théorème de Bayes et dire au médecin quel est, le plus vraisemblablement, le mal du patient.

Les valeurs *a priori* $p(Mi \mid H)$ pourraient être disponibles en permanence dans l'ordinateur, ainsi que les vraisemblances de chaque Mi calculées à partir de $p(X \mid Mi \cap H)$. Les premières proviendraient de l'expérience et de la pratique, avec des valeurs élevées pour les maux courants et faibles pour les maladies rares. Les suivantes viendraient de l'ensemble de l'expérience médicale, décrivant avec quelle fréquence les patients qui souffrent de Mi présentent les symptômes X.

Les seules données qui manquent à l'ordinateur sont les informations (X) fournies par le médecin. Ce système a deux avantages : d'une part, les calculs pertinents sont effectués correctement; d'autre part, toute l'expérience médicale (et pas seulement celle du médecin consulté) est utilisée. L'ordinateur accumule plus d'informations que le médecin et il traite les données simplement par le théorème de Bayes.

Si l'on considère ceci comme un outil qui vient en aide au médecin et non comme un substitut, l'idée est potentiellement très intéressante. L'exemple numérique suivant (tableau 7.1), ramené pour des raisons de simplicité à trois maladies, peut aider à mieux comprendre.

---

(1) Le texte et l'exemple sont extraits de D. LINDLEY, 1971.

Tableau 7.1

Maladies	Probabilités a priori p (Mi | H)	Probabilités conditionnelles p (X | Mi ∩ H)	Produit	Probabilités a posteriori p (Mi | X ∩ H)
M1	0.6	0.2	0.12	0.33
M2	0.3	0.6	0.18	0.50
M3	0.1	0.6	0.06	0.17
	1.0		0.36	1.00

Fonctionnement du théorème de Bayes sur un exemple médical théorique de trois maladies M1, M2 et M3

La maladie M1 est relativement courante; M3 est rare et M2 est intermédiaire. On observe rarement les symptômes X avec M1, mais ils sont courants avec M2 et M3. Remarquons en passant que la somme des probabilités conditionnelles de X ne vaut pas un. La quatrième colonne du tableau fournit les produits désirés par le théorème de Bayes; leur somme est 0.36. Par conséquent, k vaut 0.36. En divisant par cette valeur, on obtient les probabilités *a posteriori* en colonne cinq; la somme est ici 1. La présence des symptômes (X) rend la maladie intermédiaire la plus probable, alors que la maladie d'abord suspectée (M1) est retombée à la deuxième place, tandis que M2, qui avait bien peu de chance *a priori*, a quasi doublé en probabilité.

3. *LES IMPLICATIONS DU THEOREME.*

Une conséquence simple découle du théorème de Bayes : il n'est pas recommandé d'attacher des probabilités nulles aux événements incertains, car, si la probabilité *a priori* est zéro, la probabilité *a posteriori* vaut aussi zéro, quelles que soient les informations...

En d'autres termes, si un preneur de décision attache une ESPER égale à 0 à une solution possible, il ne sera influencé par aucune information, ce qui est évidemment absurde. Une probabilité *a priori*

Ceci ne me paraît pas très convaincant.

Si on demande à 1 élève de 6° primaire $2+3 = ?$

il n'a aucun doute sur sa réponse

dire que le doute est aussi petit que l'on veut !

aussi faible qu'une chance sur un million laisse la porte ouverte à une révision de jugement.

Une probabilité de 1 est aussi dangereuse, car, alors, la probabilité de l'événement contraire sera zéro. On ne devrait donc jamais croire de façon absolue : un léger doute doit toujours subsister. Pareille prudence ne condamne pas à l'inaction. Nous continuons à rouler en voiture, à prendre l'avion, à traverser la rue, alors que nous savons très bien que la probabilité d'y perdre la vie n'est pas nulle. Agir <u>comme si</u> la probabilité était nulle n'implique pas qu'on la juge nulle.

H ?

La seule exception à ces principes est lorsque E découle logiquement de H. Alors, $p\,(E \mid H) = 1$. Ainsi, si H porte sur les règles d'arithmétique et si la base de numération est 10, $p\,(2 \times 2 = 4 \mid H) = 1$.

Une autre conséquence découlant du théorème de Bayes est la crédibilité de l'estimateur, comme le montre l'exemple de l'investissement ci-après.

On possède 1000 dollars à placer en stock. Les probabilités que le stock bonifie (E1) vaut .50. La probabilité que le stock se déprécie (E2) vaut .50. Un expert (spécialiste) estime que le stock bonifiera, opinion que nous représentons par S. Il faut trouver $p\,(E1 \mid H \cap S)$. Or, $p\,(S \mid E1 \cap H)$ est la chance que le spécialiste ait raison (chance estimée par l'investisseur). Donc, $p\,(S) \mid E1 \cap H)$ est la confiance de l'investisseur dans les prévisions <u>favorables</u> du spécialiste.

Souvent, il suffit de l'avis de quelqu'un dont la crédibilité soit supérieure à .5. On suppose souvent que

$$p\,(S \mid E1 \cap H) = 1 - p\,(S \mid E2 \cap H) = p\,(\overline{S} \mid E2)$$

Chances que la prévision du spécialiste soit correcte = 1 − Chances pour que le spécialiste se trompe.

Nous venons d'écrire "on suppose", parce que cela peut être faux. En effet, le spécialiste pourrait être un meilleur détecteur de stocks bonifiants que de stocks qui se déprécient. Dans ce cas, on dit qu'il est biaisé. Voici l'exemple (imaginaire) d'une entreprise d'aide aux automobilistes "Tourist Recours". Dans leurs renseignements téléphoniques, les employés de cette entreprise sont pessimistes, pour dissuader les usagers de prendre des risques et de devoir recourir à eux. Admettons que je téléphone à Tourist Recours pour savoir si un col n'est pas bloqué par la neige (E1). Quand les employés disent que le col n'est pas bloqué, ma confiance est de 0.8 (vraisemblance), car je sais qu'il faut un certain temps pour qu'ils soient mis au courant des avalanches ou des tombées de neige par exemple. Mais, s'ils disent que le col est bloqué, ma confiance est de 0.5 (et non de 0.8).

4. *L'ASPECT NORMATIF DU THEOREME.*

> "Le théorème de Bayes décrit comment nous apprenons ou, plus exactement, comment nous devrions apprendre. Il dit comment nos croyances, exprimées en termes de probabilités a priori, doivent être modifiées par l'information (décrite par la vraisemblance) pour donner de nouvelles croyances exprimées en termes de probabilités a posteriori."
> (LINDLEY, 1971, p. 104)

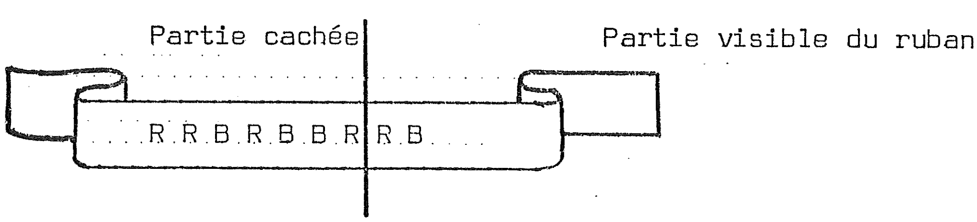
On voit combien la définition ci-dessus fait référence à une norme et ne prétend pas décrire des comportements effectifs. La question qui, de toute évidence, se pose maintenant est : dans quelle mesure les sujets humains révisent-ils leurs probabilités selon le théorème de Bayes ? C'est à cette question qu'est consacrée la deuxième partie de ce chapitre, où sont présentées des expériences testant le théorème.

C. EXPERIENCES DE VERIFICATION DE LA VALIDITE DU THEOREME

1. *LES EXPERIENCES DE H. ROUANET.*

Aux sujets qui se prêtent à ses expériences, H. ROUANET (1961) présente des sachets remplis de billes. Bien qu'uniformément constitués de papier gris, certains sachets sont dits rouges parce qu'ils contiennent 59 billes rouges et 41 bleues; les sachets dits bleus contiennent les proportions inverses : 41/59. Pour un sachet tiré de l'ensemble des sachets, le sujet doit dire s'il est bleu ou rouge. Les probabilités *a priori* (nombre de sachets) diffèrent selon l'expérience. En effet, cinq expériences différentes sont présentées à chaque sujet. Dans la première, un sachet est "rouge" et l'autre "bleu". Le rapport rouge/bleu est donc 1/1. Le rapport R/T (Rouge/Total) est .50.
Dans l'expérience suivante, deux sachets sont "rouges" et un est "bleu" (rapport 2/1). Dans les autres expériences, on présente :
5 sachets dont 4 rouges (rapport 4/1)
9 sachets dont 8 rouges (rapport 8/1)
17 sachets dont 16 rouges (rapport 16/1).

L'information est constituée par des tirages (avec remplacement) de billes, jusqu'à ce que la conviction du sujet soit faite. Malheureusement, les séries ainsi tirées sont toutes différentes et, par conséquent, les sujets se trouvent dans des situations variables. Pour remédier à cet inconvénient, ROUANET présente des séries déjà réalisées, sur ruban. Par exemple :



Le sujet arrête le déroulement dès que sa conviction est faite (avec un degré de certitude qui est propre à chaque sujet). On demande au sujet de toujours se tenir au même degré de certitude.

Les résultats montrent que les sujets obéissent à la loi de Bayes, mais en pondérant les probabilités *a priori* (le rapport R/T) par un facteur constant. Tout se passe comme si les sujets remplaçaient la probabilité R/T par la probabilité plus faible $(R/T)^{0.6}$. En fait, pour les huit sujets de l'expérience, le facteur de pondération varie de 0.4 à 0.7.

Les conclusions de H. ROUANET sont particulièrement intéressantes et nous nous permettons d'en reproduire de larges extraits.

Les décisions précédentes sont des décisions dépendant de "deux critères" : un critère a priori (les connaissances a priori sur Bleu et Rouge) et un critère 'expérimental' (l'échantillon tiré).

Lorsque les informations apportées par ces deux sources sont convergentes (Rouge plus probable a priori et différence d en faveur de Rouge), le sujet tient compte des deux sources d'information, établissant une pondération conforme (à une correction près) au théorème de Bayes.

Mais, s'il y a divergence entre les informations apportées par les deux critères (cas où la décision correcte est "le sachet est bleu"), le sujet, au lieu de devenir plus exigeant, ignore l'un des deux critères : dans le cas de l'expérience, il a ignoré le critère a priori, de nature "conceptuelle", pour ne tenir compte que du critère a posteriori, de nature plus "perceptuelle" ("J'avais beau savoir que le Bleu est rare," nous a déclaré le sujet, "je ne pouvais empêcher ma conviction de se faire dès que l'échantillon que je voyais rendait Bleu vraisemblable."). Le témoignage expérimental anéantit les connaissances qu'on pouvait avoir a priori.

Si les conclusions précédentes étaient confirmées, il serait vain d'appliquer le théorème de Bayes pour chercher à évaluer l'ESPER *a priori* d'une croyance, dans le cas où un événement est venu l'infirmer.

Supposons, par exemple, qu'une personne, devinant correctement quatre fois consécutives un chiffre compris entre 0 et 9, l'expérimentateur estime a posteriori qu'il existe plus d'une chance sur deux pour que cette personne ait une perception extra-sensorielle. On n'en pourrait pas conclure que la probabilité a priori de l'expérimentateur dépassait 1/10000 mais seulement qu'il s'est décidé à cause de la trop faible vraisemblance (1/10000) que le hasard soit la cause des résultats expérimentaux.

Un tel point de vue, s'il était confirmé, jetterait peut-être quelques lueurs sur la facilité humaine bien connue à expliquer les phénomènes rares par des causes hautement improbables; en outre, il fournirait une justification psychologique aux notions statistiques de "seuils de signification", risques de première et de seconde espèce, etc.

Il serait intéressant de construire de nouvelles expériences destinées à éprouver la validité des conclusions ci-dessus.

2. *LES EXPERIENCES DE Ward EDWARDS.*

Les expériences qui vont maintenant être décrites diffèrent des expériences de H. ROUANET sur un point essentiel : les probabilités ne sont plus objectives, mais "estimées" par des experts. La justification de cette procédure est fournie par EDWARDS (1967) :

> "L'ennui est que les probabilités (ou rapports de vraisemblance), données nécessaires pour le théorème de Bayes, ne sont pas fixables objectivement dans la plupart des situations de traitement d'informations de la vie courante. La solution consiste, évidemment,à avoir recours à des experts pour estimer les nombres nécessaires." (p.140).

A. LE CONTEXTE DU JEU.

Dans l'expérience menée en 1964, on présente au sujet six hypothèses décrivant la situation politique du monde dix ans plus tard. Par exemple, l'hypothèse 1 est : "La Russie et la Chine sont sur le point d'attaquer l'Amérique du Nord"; l'hypothèse 4 est : "La Chine est sur le point d'attaquer le Japon"; l'hypothèse 5 (ou hypothèse fourre-tout) est "Un autre conflit important est sur le point de se déclarer". Enfin, l'hypothèse 6 est "Etat de paix".

Les probabilités *a priori* et *a posteriori* de l'expérience porteront sur ces six hypothèses. Les "informations" du jeu consistent en des dépêches d'agence du type suivant :

> "Le gâteau servi lors du déjeuner des présidents japonais et chinois était surmonté du nouveau drapeau de l'alliance sino-japonaise."

Autre exemple :

> "Ce matin, deux escadres de sous-marins (non nucléaires) ont quitté VLADIVOSTOK... Il s'agit sans doute d'exercices de routine, quoi qu'il s'agisse là d'un déploiement de force tout à fait exceptionnel."

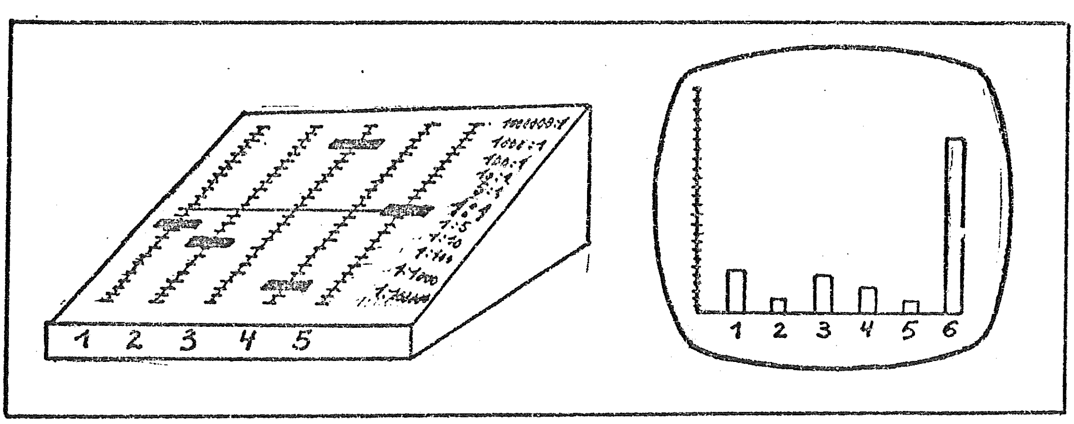
Aucune de ces "informations" ne pouvait être interprétée comme confirmant ou infirmant totalement l'une des hypothèses.

Pour cette expérience, W. EDWARDS a développé quatre systèmes différents : PIP (*Probabilistic Information Processing*) et trois systèmes concurrents : POP, PEP, PUP. Dans chacun de ces systèmes, le sujet doit considérer les hypothèses deux à deux, ce qui peut se résumer à cinq situations (1). Au début du jeu, chacune des cinq premières hypothèses est dans un rapport (2) d'une chance contre cinq (1:5), avec l'hypothèse 6, la paix. Par conséquent, les probabilités *a priori* de départ sont respectivement .1, .1, .1, .1, .1 et .5.

Par diapositives, on fournit au sujet une information (dépêche d'agence) à partir de laquelle il doit estimer la vraisemblance ou *likelihood ratio* de chaque hypothèse, et non la probabilité de l'information (3). Pour chaque information, le sujet doit donc donner cinq réponses, les rapports de vraisemblance de chaque hypothèse opposée à l'hypothèse 6.

Dans les systèmes PIP et POP, le sujet dispose d'un terminal d'ordinateur muni d'une console de réponses et d'un écran cathodique (C.R.T.). La console présente cinq curseurs (un par réponse) que l'on peut positionner sur les rapports de vraisemblance suivants : 1:1, 1:10, 1:100... jusqu'à 1:1000000, et les inverses 10:1, 100:1, jusque 1000000:1 (tableau 7.2).

Cette technique permet d'enregistrer automatiquement les probabilités (le sujet presse un bouton quand sa conviction est faite).



(1) En combinant les six hypothèses, il y a, bien entendu, quinze situations distinctes, mais non indépendantes. En fait, toutes les probabilités peuvent être calculées à partir de cinq situations seulement; dans l'expérience, il s'agissait chaque fois d'opposer les cinq hypothèses à l'Hypothèse 6 ("la paix").

(2) L'expérience utilise les rapports (odds) et non les probabilités.

(3) Ces deux notions, pourtant intimement liées, ont été dissociées pour les besoins de la recherche. On en trouvera la justification dans EDWARDS (1971, p. 146).

L'écran cathodique présente les probabilités *a posteriori* sous forme graphique, les hauteurs des rectangles variant dynamiquement avec la manipulation des curseurs par le sujet.

B. LES QUATRE GROUPES EXPERIMENTAUX.

Dans le système PIP, après chaque série de cinq comparaisons, le sujet remet ses curseurs au niveau 1:1, et les rectangles de l'écran reviennent automatiquement au *pattern* de départ (.1, .1, .1, .1, .1, .5). Par conséquent, le travail de l'ordinateur consiste uniquement à transformer les valeurs des vraisemblances en probabilités *a posteriori* (les probabilités *a priori* étant la résultante de l'opération antérieure), selon le théorème de Bayes. De plus, l'ordinateur combine les probabilités successives (comme dans le système POP), mais à l'insu du sujet.

Dans le système POP, le sujet laisse les curseurs là où il les a placés et ne modifie que les rapports qu'il lui paraît nécessaire de changer. Il fait donc des estimations basées sur l'ensemble des informations reçues jusque là. Les rectangles affichés sur l'écran sont les probabilités *a posteriori* engendrées par les réponses qui viennent d'être fournies.

Un problème commun à la formation des sujets pour les groupes PIP et POP consiste à leur faire comprendre qu'un événement suffisant pour changer un rapport 1:1 en un rapport 2:1 est aussi suffisant pour changer un rapport 100:1 en 200:1, en d'autres termes, que les informations affectent les rapports de façon multiplicative.

Dans les systèmes PEP et PUP, les sujets ne disposent pas d'ordinateur : ils écrivent leurs réponses sur des feuilles de papier. Dans le système PEP, on dit au sujet qu'en cas de guerre, il sera pénalisé de 100 points, mais qu'en cas de paix, aucune pénalisation n'intervient. Le sujet doit alors estimer le coût raisonnable (*the fair price*) d'une assurance qui payerait 100 points en cas de guerre et rien en cas de paix. La réponse PEP n'est que 100 fois la probabilité d'une guerre particulière. Dans le groupe PUP, les sujets estiment directement cette probabilité, au lieu de la dissimuler sous l'aspect du coût d'une assurance.

C. LES RESULTATS.

Avant d'aborder l'interprétation des résultats, il importe de répondre à la question : "Qu'est-ce qu'une bonne performance pour un système diagnostique ?". La réponse évidente est qu'un système diagnostique fonctionne bien si ses diagnostics sont corrects. Malheureusement, il ne manque pas de situations où la vérité est inconnue. Le critère de vérité devient alors l'accord avec un autre système diagnostique dont la validité est reconnue.

Par conséquent, on peut penser qu'un système diagnostique est meilleur si, dans des situations où divers systèmes fournissent le même diagnostic, il atteint ce diagnostic sur la base du minimum d'information. A égalité d'information, le meilleur système tirerait, par conséquent, des conclusions plus hardies que ses concurrents.

D'un tel système, on doit évidemment exiger qu'il modifie son opinion dans le sens de données nouvelles aussi bien que le feraient ses compétiteurs.

Les quatre systèmes (PIP, POP, PUP, PEP) aboutissent aux mêmes diagnostics, comme le montrent les corrélations qui suivent. Elles ont été obtenues en comparant les rapports finals pour chacune des cinq situations de guerre, dans les dix-huit scénarios présentés aux étudiants. Elles sont donc calculées sur nonante observations.

Tableau 7.3

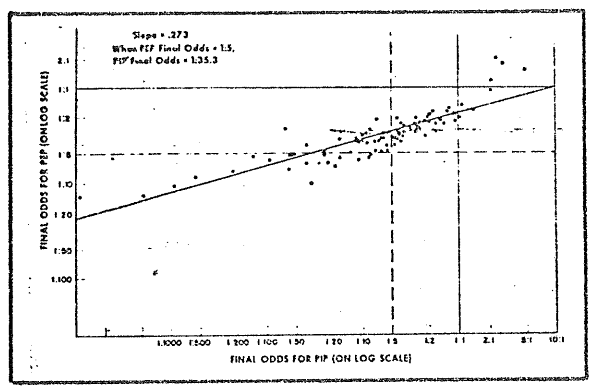
	PIP	POP	PUP	PEP
PIP	-	.89	.88	.86
POP	-	-	.91	.93
PUP	-	-	-	.93
N. de sujets	8	10	10	6

Corrélations des diagnostiques des quatre systèmes concurrents de l'expérience de W. EDWARDS (1967)

La valeur élevée des corrélations apparaît également dans les six nuages de points que l'on pourrait établir à partir des résultats de deux des quatre systèmes. Ces nuages de points sont riches d'enseignement, comme nous allons le voir à partir du *scatter plot* des rapports finals fournis par le système PIP et les rapports finals fournis par le système PEP (tableau 7.4).

Les axes sont tracés en échelles logarithmiques des rapports. La moyenne des rapports de PEP est 1:5, tandis que la moyenne des rapports de PIP est 1:35 (tableau 7.4). Rappelons que les rapports de départ (notés en pointillés sur le graphique) étaient 1:5.

Tableau 7.4



Comparaison des prédictions des systèmes PIP et PEP (in W. EDWARDS 1967)

On remarquera que la pente de la droite de régression est loin de valoir 1, comme on pourrait s'y attendre. La situation idéale se caractérise non seulement par une corrélation très élevée, mais aussi par une valeur proche de 1 de la pente de la droite de régression de la variable prédite sur celle observée. En fait, elle vaut .273.

Les pentes des droites des autres nuages de points sont présentés dans le tableau 7.5.

Tableau 7.5

		Critères (ou Y)		
		POP	PUP	PEP
Prédicteurs (ou X)	PIP	.422	.388	.273
	POP	-	.84	.63
	PUP	-	-	.67

On remarque que les pentes des nuages entre POP, PEP et PUP sont bien plus proches de 1 que les pentes des nuages où PIP intervient.

L'ampleur de la différence entre PIP et les trois autres systèmes surprit EDWARDS et son équipe d'expérimentateurs. Ils avaient espéré que PIP serait plus efficace à 10 ou 20 pour cent. Or, il est plus efficace de plus de 180 % que POP, son plus proche compétiteur, et près de 400 % plus efficace que PEP. En effet, pour obtenir des conclusions aussi hardies que PIP, il faut multiplier celles de POP par 2,37 (1/.422), celles de PUP par 2,58 (1/.388) et celles de PEP par 3,66 (1/.273). En échelle logarithmique, de telles valeurs prennent une importance énorme.

W. EDWARDS a également mené des recherches du type de celles de ROUANET (tirage de jetons de couleurs dans un sac qui contient soit 70 % de jetons rouges et 30 % de jetons bleus, soit les proportions inverses). Dans le graphique des résultats (tableau 7.6), on a porté en abscisse la différence entre le nombre de jetons rouges tirés (r) et nombre de jetons bleus tirés (b). En ordonnée, les logarithmes des rapports entre les deux hypothèses (Hr = le sac contient 70 % r; HB = le sac contient 70 % B).

Tableau 7.6

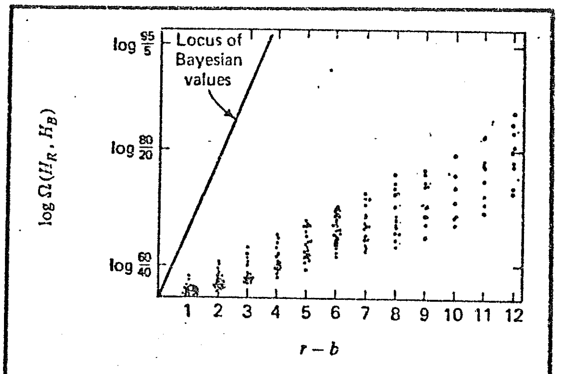


Fig. 5.15 Probability estimates of a subject compared with Bayesian estimates.

Présentés sous une forme différente, nous retrouvons des résultats convergents avec ceux de Rouanet, ce qui est encourageant.

Laissons à EDWARDS, LINDMAN et PHILLIPS (1965) le soin de commenter le graphique ci-dessus :

> "L'homme est incapable d'extraire de l'information toute la certitude dont le théorème de Bayes révèle la présence dans cette information. Autrement dit, l'homme traite l'information avec beaucoup de prudence*.."(cité par COOMBS, DAWES et TVERSKY, 1970, p. 147).

D. CONCLUSIONS.

L'aspect non réaliste de l'expérience, l'arbitraire des hypothèses et la nature non séquentielle des informations peuvent être critiqués. On peut notamment avancer que PIP est un système spécialement conçu pour traiter des informations non séquentielles et que, par conséquent, les résultats sont biaisés en sa faveur. Cependant, d'autres expériences (SCHUM et al., 1966) montrent la supériorité (moins forte cependant) de PIP sur des systèmes concurrents.

Les expériences de W. EDWARDS et de H. ROUANET - tellement différentes pourtant - aboutissent aux mêmes conclusions : les systèmes humains (chez ROUANET) et homme-machine (chez EDWARDS) sont sous bayésiens dans la révision des probabilités. Chez le premier, le coefficient de correction à utiliser va de .04 à .07. Chez le second, de .27 à .42. Cette convergence des résultats obtenus par ROUANET et EDWARDS est encourageante. Elle indique que, moyennant certaines conditions (par exemple, constantes de "correction"), le théorème de Bayes pourrait être utilisé dans les sciences de l'éducation. Bien entendu, d'autres expériences doivent être menées qui indiqueraient, outre les valeurs des constantes, les conditions et les domaines d'application du théorème, et, bien sûr, l'intérêt exact de toute l'approche.

C'est dans cette perspective que nous avons entrepris des recherches qui, on le comprend aisément, ne peuvent être qu'"exploratoires" dans un aussi vaste champ d'investigation. Dans les situations qui vont être étudiées, il n'existe aucune méthode objective de calcul de la probabilité; le recours à des experts est le seul paliatif, comme les expériences d'EDWARDS. Par contre, le sujet peut prendre connaissance de l'"état de la nature" après avoir fourni ses estimations, comme

---

* "Man is a conservative information processor."

dans les expériences de ROUANET. Cette possibilité est exclue chez EDWARDS, puisqu'on n'y manipule que des situations fictives, projetées dans le futur, par définition impossible à connaître. Il est intéressant de pouvoir confronter les probabilités subjectives à la réalité puisqu'on peut ainsi envisager d'améliorer le processus d'estimation des probabilités (par rapport à un critère objectif).

La situation pédagogique qui rassemble des éléments des expériences de ROUANET et EDWARDS est assez proche de la situation médicale. Les probabilités *a priori* et les vraisemblances des symptômes ne permettent pas de calculer de probabilité subjective : seuls les experts peuvent donner leur avis (forcément subjectif). Néanmoins, on peut prendre, *a posteriori* (lors de l'autopsie, par exemple), connaissance de l'"état de la nature". En pédagogie, au moment où les connaissances sont testées, les élèves, en état de "connaissance partielle", estiment (subjectivement) l'exactitude de leur réponse. La solution correcte fournie ensuite par le professeur (ou le document de référence) indiquera alors au sujet l'état de la nature. C'est selon ces principes de base que nous avons créé des "jeux de certitude" propres à tester le théorème de Bayes. Nous allons maintenant les décrire.

D. UN JEU DES ESPER PROPRE A L'ETUDE DU THEOREME

Nous avons proposé un jeu des révisions des ESPER pour plusieurs raisons. Tout d'abord, pour constituer une situation plus proche de la situation pédagogique que celles de Rouanet et Edwards. Cependant, il fallait que cette situation reste suffisamment manipulable, afin de vérifier expérimentalement le théorème. Cet exemple de situation et les premiers résultats que ce jeu a permis de recueillir font apparaître plus clairement les possibilités et les limites du théorème.

Dans les expériences qui suivent, nous avons remplacé, dans le théorème de Bayes, la PUR (probabilité d'utilisation ou d'apparition de la réponse) par PER. Afin de satisfaire aux quatre axiomes décrits page 372, certaines précautions ont été prises. En particulier, afin que $\sum_{i}^{n}$ PER i = 1, les consignes préciseront que les solutions proposées sont exhaustives.

1. La première version du jeu.

Nous avons d'abord conçu un jeu où l'on demande successivement au sujet :

1.- ses ESPER *a priori;*

2.- les probabilités conditionnelles qu'il attache à l'information (X);

3.- ses ESPER *a posteriori;*

L'ensemble des éléments ainsi recueilli permet de faire fonctionner l'équation dans les deux sens.

La procédure expérimentale est divisée en trois étapes.

Qu'est ce qui nous guide dans le choix des lettres proposées ?

ETAPE 1.

On présente au sujet un texte mutilé et on lui demande de choisir, pour chacune des dix lettres à prédire, parmi six solutions proposées. On garantit aux sujets que la réponse correcte figure parmi les six solutions (axiome 2). L'exemple ci-dessous ne présente que les trois premières questions (tableau 7.7).

Tableau 7.7

JEU DES CERTITUDES (THEOREME DE BAYES)

1
AINSI L'ABONDANCE DE L'HCMME ENGENDRE LE RATIONNEMENT DE LA PERSCNNALITE.MEME SI L'HOMME M

LA LETTRE QUI SUIT EST	A	E	I	O	U	N
POURCENTAGES DE CHANCES :	...	...	...	...	...	...

2
FRIVEMENT,MCINS SEFIEUSEMENT.LA CONTFEPARTIE DE L'EXTENSION DE NOS INFORMATIONS ,ET DE LA MULTIPLICATICN D

LA LETTRE QUI SUIT EST	U	E	A	I	R	O
PCURCENTAGES DE CHANCES :	...	...	...	...	...	...

3
DE NOS PROCHES.CES FAITS SONT EN RELATION AVEC LE CHANGEMENT PROFOND,QUI S'ACCO MPLIT SOUS NOS YEUX,DANS LES C

LA LETTRE QUI SUIT EST	R	U	L	O	A	E
PCURCENTAGES DE CHANCES :	...	...	...	...	...	...

Feuille de réponses de la phase 1 du jeu "révision des ESPER"(première version)

La consigne est la suivante (tableau 7.8).

Tableau 7.8

> Les vingt phrases (groupées deux par deux) reproduites sur le listing constituent un texte continu que nous avons emprunté à un auteur connu.
>
> Dans les dix séries de deux phrases, nous avons chaque fois mutilé la fin de la seconde phrase. Le jeu consiste à trouver la lettre qui suit immédiatement la dernière lettre imprimée.
>
> Juste en dessous de la phrase mutilée, figurent six lettres. La lettre à trouver est toujours l'une de ces six lettres.
>
> Veuillez chaque fois,
>
> 1) indiquer la lettre qui vous semble être la suivante
>
> 2) indiquer, pour chacune des six lettres, quel pourcentage de chances d'être correcte vous lui accordez.
>
> N.B. : le total de vos pourcentages doit être 100 %.
>
> Le but du jeu consiste à obtenir le plus grand nombre de points au total. Les points sont gagnés et perdus selon la matrice en annexe.
>
> (la même matrice que celle du jeu des certitudes).

Consignes de l'étape 1 du jeu "révision des ESPER"
(première version)

ETAPE 2.

Cette étape est consacrée au recueil des probabilités conditionnelles de l'information. L'élève doit ignorer que la lettre présentée (y à la question 1) est, en fait, la lettre qui suit dans le texte original. C'est pourquoi un texte en bas de page précise (mais c'est un mensonge) que ces lettres ont été tirées au hasard. Les trois premières situations figurent, en exemple, dans le tableau 7.9.

Tableau 7.9

```
1
AINSI L'ABONCANCE DE L'HCMME ENGENDRE LE RATIONNEMENT DE LA PERSCNNALITE.MEME SI
 L'HOMME M

QUESTION 1 : QUEL EST LE POURCENTAGE DE CHANCES PCUR QUE, DANS CETTE PHRASE-CI,
                                   A       E       I       O       U       N

                                  ...     ...     ...     ...     ...     ...
                     SOIT SUIVI DE Y ?


2
 PR.VEMENT,MOINS SERIFUSEMENT.LA CONTREPARTIE DE L'EXTENSION DE NOS INFCRMATIONS
,ET DE LA MULTIPLICATION D
QUESTION 1 : QUEL EST LE POURCENTAGE DE CHANCES POUR QUE, DANS CETTE PHRASE-CI,
                                   U       E       A       I       R       O

                                  ...     ...     ...     ...     ...     ...
                     SOIT SUIVI D'UN ESPACE ?


3
 DE NOS PROCHES.CES FAITS SONT EN RELATION AVEC LE CHANGEMENT PROFOND,QUI S'ACCO
MPLIT SOUS NOS YEUX,DANS LES C

QUESTION 1 : QUEL EST LE POURCENTAGE DE CHANCES POUR QUE, DANS CETTE PHRASE-CI,
                                   R       U       L       C       A       E

                                  ...     ...     ...     ...     ...     ...
                     SOIT SUIVI DE U ?
```

Feuille de réponses de l'étape 2 du jeu "révision des ESPER"
(première version)

Voici la consigne qui est associée à cette tâche (tableau 7.10).

Tableau 7.10

Vous retrouvez les mêmes phrases que dans le jeu n° 1, et les mêmes lettres possibles.
Ici, vous devez fournir la probabilité (pourcentage de chances) que chacune des 6 lettres proposées soit suivie d'une autre lettre, tirée au hasard.

Ainsi, à la question 1, on vous demande d'évaluer le pourcentage de chances pour que A soit suivi de Y (lettre tirée au hasard (1)), dans le cas où la lettre A est la lettre correcte.

Il faut ensuite estimer la probabilité que E soit suivi de Y, dans le cas où E est la lettre correcte.

Et ainsi de suite jusque N.

On passe ensuite à la question 2 où la lettre tirée au hasard est l'espace (considéré dans ce jeu comme la 27° lettre de l'alphabet).

On procède ainsi pour les dix questions.

(1) Ces lettres ont été tirées au hasard comme suit :

- On a constitué un ensemble de n lettres parmi lesquelles figurait la lettre qui suit réellement celle qu'il faut choisir (et qui est présentée parmi 5 autres). Le nombre de lettres de cet ensemble diffère à chaque fois, et il arrive même que la lettre correcte ait été introduite plusieurs fois dans l'ensemble.
- Quand n=2, la lettre tirée a 1 chance sur deux d'être celle qui suit rélement ; mais quand il y a 10 lettres, la lettre tirée n'a plus que 1 chance sur 10 de l'être.

Consignes de l'étape 3 du jeu "révision des ESPER"
(première version)

ETAPE 3.

On révèle au sujet que la lettre prétendûment tirée au hasard est, en réalité, la lettre suivante. Il s'agit là de l'information introduite dans le système. On présente alors une nouvelle fois les phrases de départ, mais avec l'information supplémentaire. Le sujet doit alors attribuer ses ESPER *a posteriori* aux six lettres proposées.

LES RESULTATS.

Lors de l'expérimentation, il est apparu que les sujets ne parvenaient pas, dans l'étape 2, à faire abstraction des probabilités jointes pour évaluer les probabilités conditionnelles. Par exemple, dans la question 1 :

L'homme M_Y Pourcentage de chances : A E I O U N

De nombreux sujets ont tendance à évaluer la probabilité jointe p $(AY \mid \text{texte})$ et non la probabilité conditionnelle p $(Y \mid \text{texte} \cap A)$.

Les consignes ont été modifiées mais, de toute évidence, la difficulté ne pouvait être surmontée que par une modification radicale de l'ordre de présentation des étapes. Les étapes 1 et 2 peuvent, en effet, être inversées à volonté.

Cette première expérimentation a aussi montré que le nombre de solutions proposées (six) est trop grand, car il est impossible de trouver autant de réponses plausibles. Par conséquent, on recueille beaucoup de réponses où p vaut 0. (quand p désigne les probabilités conditionnelles de l'information ou quand p désigne les probabilités *a priori*). Dans les deux cas, cela annihile l'impact de l'événement X dans le théorème. C'est pourquoi le nombre de solutions proposées a été réduit à trois dans la version actuelle du jeu.

2. La version actuelle du jeu.

L'originalité de cette version réside dans le fait que l'on recueille les probabilités conditionnelles de l'information avant les ESPER *a priori*. L'expérience se déroule en cinq étapes.

ETAPE 1.

On présente au sujet les dix questions synoptiquement (tableau 7.11).

Tableau 7.11

Quel est le pourcentage de chances pour que, dans cette phrase-ci, la lettre qui suit soit : *

%

1
AINSI L'ABONDANCE DE L'HOMME ENGENDRE LE RATIONNEMENT DE LA PERSONNALITE.MEME SI L'HOMME M**U** 1......
EO

2
PRIVEMENT,MOINS SERIEUSEMENT.LA CONTREPARTIE DE L'EXTENSION DE NOS INFORMATIONS ,ET DE LA MULTIPLICATION D**R**. 2......
UE

3
DE NOS PROCHES.CES FAITS SONT EN RELATION AVEC LE CHANGEMENT PROFOND,QUI S'ACCOMPLIT SOUS NOS YEUX,DANS LES C**R** 3......
OA

4
ERAILLES,AU DEUIL.LORSQUE CES NOUVELLES COUTUMES ONT COMMENCE DE SE FAIRE JOUR AUX ETATS-UNIS,ELLES ONT CHOQUE LES EUROPEENS ET F**O** 4......
IU

5
BARIE YANKEE.MAIS NOUS LES VOYONS AUJOURD'HUI NON PAS GAGNER L'EUROPE(LE MOT GAGNER IMPLIQUERAIT QU'ELLES SO**I** 5......
RN

6
L'EUROPE),MAIS SE FORMER,SE CONSTITUER EN EUROPE,COMME ELLES L'ONT FAIT AUX ETATS-UNIS,POUR LES MEMES CAUSES ET PAR SUITE DE L'A**V** 6......
GM

7
FAIT,CE NE SONT PAS QUELQUES USAGES ANECDOTIQUES QUI SONT EN QUESTION;CE N'EST PAS SEULEMENT L'A**T** 7......
UN

8
UMANITE MILLENAIRE QUI ACHEVE DE SE DISSIPER.LE CHANGEMENT EST DEJA SI TOTAL QUE LA PLUPART DE MES L**I** 8......
YE

9
CI QUE S'ILS FONT L'EFFORT DE LIRE LES RARES LIVRES OU LA MENTALITE PRIMITIVE ET TRADITIONNELLE EST C**O** 9......
EI

10
S UNE CONCEPTION MAGIQUE ET RITUELLE DES GRANDS EVENEMENTS DE LA VIE:NAISSANCE,MARIAGE,MORT;LES FETES,LE-**P** 10......
TG

*= lettre tirée au hasard. Dans le tableau ci-dessus, les dix lettres terminales des phrases mutilées ont été grossies pour le lecteur. Ces lettres seront modifiées lors des étapes 2 et 3 (voir ci-après). Voici les trente lettres utilisées pour les dix questions (les lettres correctes ont été soulignées).

question	1	2	3	4	5	6	7	8	9	10
étape 1	U	R	R	O	I	<u>V</u>	<u>T</u>	I	<u>O</u>	P
étape 2	E	U	<u>O</u>	I	R	G	U	Y	E	<u>T</u>
étape 3	<u>O</u>	<u>E</u>	A	<u>U</u>	<u>N</u>	M	N	<u>E</u>	I	G

L'expérimentateur montre alors au sujet (rapidement) une table de lettres au hasard (tableau 7.12), lui en explique la structure et lui demande, pour chaque question, de fournir un nombre compris entre 1 et 27. Le sujet a ainsi l'impression de désigner entièrement au hasard la solution proposée.

Chaque fois, l'expérimentateur simule la consultation de la table. En fait, il fournit les lettres qui suivent réellement la lettre à deviner. Ces lettres sont indiquées tout en haut de la table. Tous les sujets testés ont reconnu avoir été persuadés (d'emblée et jusqu'au bout de l'expérience) du caractère aléatoire de la lettre considérée. A la fin du "tirage au sort", l'expérimentateur signale que "cela va entraîner des combinaisons bizarres de lettres".

Le sujet est alors invité à répondre par écrit aux dix questions.

Tableau 7.12

	1	2	3	4	5	6	7	8	9	10
	y	-	u	r	t	e	t	c	r	r
1	p	x	e	x	q	j	n	s	u	l
2	e	s	v	y	z	a	-	z	p	c
3	l	v	q	h	j	n	j	w	v	k
4	a	g	w	n	n	s	b	y	j	m
5	o	w	x	j	d	b	w	-	l	b
6	i	q	b	s	w	-	a	a	o	y
7	b	e	n	q	g	q	k	d	k	a
8	r	l	u	d	y	c	i	c	x	q
9	g	b	l	m	e	k	q	j	b	n
10	z	n	a	g	t	g	d	q	a	f
11	n	c	s	v	k	z	f	k	q	r
12	f	d	m	e	-	i	c	h	c	j
13	s	a	g	k	v	d	k	p	i	g
14	v	r	z	s	h	f	x	e	-	z
15	d	f	k	c	s	w	h	t	g	t
16	u	f	d	-	i	e	p	f	d	i
17	c	s	j	l	p	l	v	b	n	o
18	q	i	y	b	c	r	g	o	f	-
19	w	t	f	p	l	t	z	r	t	d
20	k	h	c	i	a	p	s	l	n	p
21	-	u	-	t	r	y	e	i	e	u
22	h	j	i	u	b	h	t	m	r	h
23	x	k	t	f	o	u	r	c	y	v
24	j	m	o	a	u	x	l	u	h	s
25	y	p	h	w	m	v	m	g	w	x
26	m	z	p	r	f	o	u	x	z	e
27	t	-	r	z	x	m	o	v	s	w

ETAPE 2.

L'expérimentateur avoue au sujet que la dernière lettre imprimée dans chaque question (par exemple le U de la question 1) n'est pas correcte. Il explique qu'il s'agit d'un artifice nécessaire pour tester la probabilité relative de la lettre tirée au sort, probabilité requise par le théorème.

Le sujet reçoit alors une feuille n° 2, en tous points identique à la première, exceptées les dernières lettres imprimées. L'expérimentateur garantit que ce sont les lettres exactes et invite le sujet à recopier (avec un stylo de couleur) les lettres tirées au hasard. Le sujet répond alors aux dix lettres. A la fin de l'expérience, la majorité des sujets avouent ne pas avoir mis en doute l'authenticité des lettres terminales.

ETAPE 3.

Selon une procédure identique à l'étape 2, l'expérimentateur avoue qu'il a menti, mais que c'est la dernière fois et que, cette fois-ci, les lettres sont correctes. Le sujet reçoit alors une feuille 3 et y répond après avoir recopié les lettres tirées au hasard.

Lors de l'interview final, la majorité des sujets avouent avoir gardé des doutes quant à l'authenticité des lettres terminales.

ETAPE 4.

L'expérimentateur avoue alors au sujet que la lettre correcte lui a bien été présentée, mais pas forcément dans la feuille 3. On lui présente alors une quatrième feuille où les trois lettres déjà utilisées figurent comme solutions possibles. Le sujet doit attribuer une ESPER (*a priori*) à chacune d'elles.

ETAPE 5.

L'expérimentateur révèle enfin le dernier mensonge : la lettre que le sujet a prétendûment tirée au hasard lors de l'étape 1 est en fait la lettre qui suit la lettre à deviner; ceci constitue l'information introduite dans le système. Le sujet doit alors attribuer des ESPER *a posteriori* à chaque solution proposée.

On dispose donc de tous les éléments pour utiliser le théorème de Bayes.

E. RESULTATS EXPERIMENTAUX

Voici, confrontées, les deux probabilités *a posteriori* p $(Ei \mid X \cap H)$ sur les nuages de points :

EXPER = probabilité *a posteriori* calculées par le théorème à partir des ESPER *a priori* et des vraisemblances fournies par le sujet.

ESPER = probabilité *a posteriori* fournies par le sujet.

Sur les tableaux 7.13 à 7.18,

p $(E \mid H)$ représente la probabilité *a priori* de chaque lettre proposée;

p $(X \mid E \cap H)$ représente la probabilité de l'information pour chaque lettre proposée (ou la vraisemblance, conditionnelle à l'information, de chaque lettre).

Le produit (en colonne 5) est le résultat de la multiplication des deux premières colonnes. On remarquera que les valeurs inférieures à .0001 ont été, artificiellement, ramenées à .0001.

p $(E \mid X \cap H)$ est la probabilité *a posteriori*. La première est observée (c'est la réponse effectivement fournie par le sujet); la seconde est calculée au moyen du théorème.

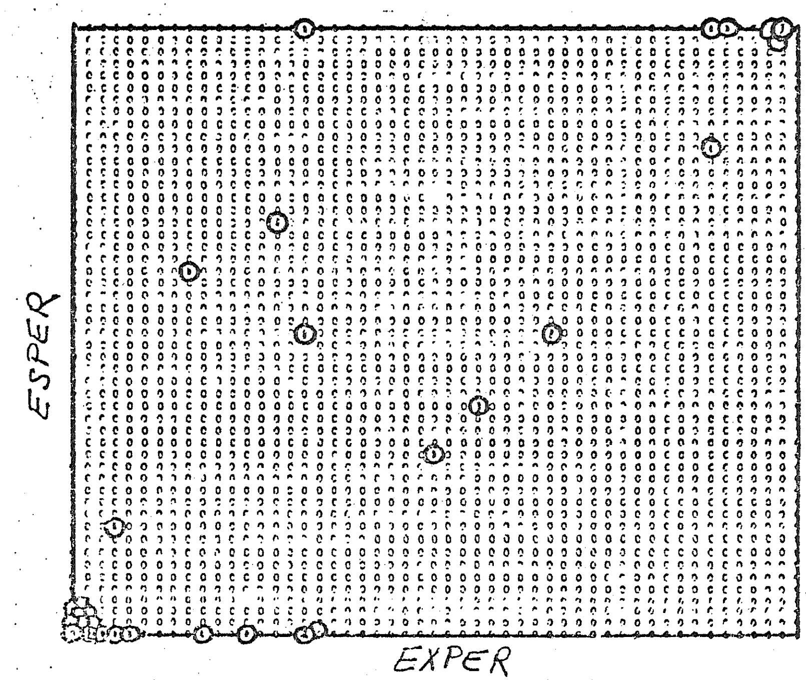
L'hypothèse générale relative à la révision des probabilités est la suivante :

HR - Le théorème de BAYES s'applique (éventuellement à une constante près qui resterait à calculer) dans les situations cognitives proches des situations pédagogiques. Plus précisément, la corrélation entre les valeurs prédites en application du théorème (EXPER) et les valeurs observées (ESPER)est élevée.

Voici l'ensemble des données pour chacun des six sujets.

	P(E|H) (2)	P(X|E+H) (3)	INF.REL.	P(E|X+H)	PRODUIT	P(E|X+H)
1	0.0001	0.0001	0.0002	0.0001	0.0001	0.0002
2	0.2000	0.0001	0.0002	0.0001	0.0001	0.0002
3	0.8000	0.5000	0.9996	1.0000	0.4000	0.9995
1	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
2	0.5000	1.0000	0.6666	0.5000	0.5000	0.6666
3	0.5000	0.5000	0.3333	0.5000	0.2500	0.3333
1	0.5000	0.0001	0.0001	0.0001	0.0001	0.0002
2	0.5000	0.8000	0.9998	1.0000	0.4000	0.9995
3	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002
1	0.4000	0.2500	0.4545	0.4000	0.1000	0.5714
2	0.4500	0.1000	0.1818	0.0001	0.0450	0.2571
3	0.1500	0.2000	0.3636	0.6000	0.0300	0.1714
1	1.0000	0.0001	0.0001	1.0000	0.0001	0.3333
2	0.0001	0.4500	0.3600	0.0001	0.0001	0.3333
3	0.0001	0.8000	0.6399	0.0001	0.0001	0.3333
1	0.8000	0.8000	0.5333	1.0000	0.6400	0.9014
2	0.1000	0.1000	0.0667	0.0001	0.0100	0.0141
3	0.1000	0.6000	0.4000	0.0001	0.0600	0.0845
1	0.7000	0.6000	0.6316	1.0000	0.4200	0.8889
2	0.1500	0.1000	0.1053	0.0001	0.0150	0.0317
3	0.1500	0.2500	0.2632	0.0001	0.0375	0.0794
1	0.5000	0.0001	0.0001	0.0001	0.0001	0.0003
2	0.0001	0.6000	0.4444	0.0001	0.0001	0.0003
3	0.5000	0.7500	0.5555	1.0000	0.3750	0.9995
1	0.3300	0.4500	0.2903	0.7000	0.1485	0.2903
2	0.3300	0.8000	0.5161	0.3000	0.2640	0.5161
3	0.3300	0.3000	0.1935	0.0001	0.0990	0.1935
1	0.1500	0.2000	0.1538	0.0001	0.0300	0.0472
2	0.7000	0.8000	0.6154	0.8000	0.5600	0.8819
3	0.1500	0.3000	0.2308	0.2000	0.0450	0.0709

Tableau 7.13

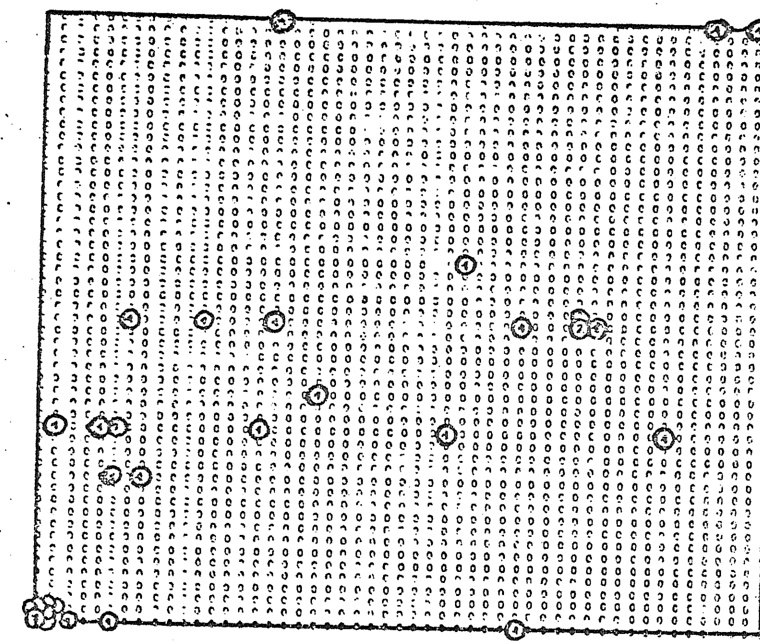


r = .8568

Résultats expérimentaux au jeu "révision des ESPER" pour le sujet n°1. En ordonnée, les ESPER a posteriori. En abscisse : les EXPER a posteriori

	P(E|H) (2)	P(X|E+H) (3)	INF.REL.	P(E|X+H)	PRODUIT	P(E|X+H)
1	0.0010	0.0010	0.0009	0.0001	0.0001	0.0001
2	0.0010	0.1000	0.0909	0.0001	0.0001	0.0001
3	0.9990	0.9990	0.9082	1.0000	0.9980	0.9998
1	0.0010	0.0010	0.0007	0.0001	0.0001	0.0001
2	0.5000	0.9990	0.6660	0.5000	0.4995	0.6664
3	0.5000	0.5000	0.3333	0.5000	0.2500	0.3335
1	0.4000	0.0010	0.0008	0.0001	0.0004	0.0011
2	0.3000	0.7500	0.5995	0.6000	0.2250	0.5994
3	0.3000	0.5000	0.3997	0.4000	0.1500	0.3996
1	0.4000	0.5000	0.6667	0.5000	0.2000	0.7407
2	0.2000	0.1500	0.2000	0.2500	0.0300	0.1111
3	0.4000	0.1000	0.1333	0.2500	0.0400	0.1481
1	0.5000	0.0010	0.0006	0.0001	0.0005	0.0011
2	0.2000	0.7500	0.4286	1.0000	0.1500	0.3332
3	0.3000	0.9990	0.5709	0.0001	0.2997	0.6657
1	0.4500	0.9000	0.4186	0.3300	0.4050	0.5745
2	0.1000	0.7500	0.3488	0.3300	0.0750	0.1064
3	0.4500	0.5000	0.2326	0.3300	0.2250	0.3191
1	0.6000	0.7500	0.6818	0.3300	0.4500	0.8654
2	0.2000	0.2500	0.2273	0.3300	0.0500	0.0962
3	0.2000	0.1000	0.0909	0.3300	0.0200	0.0385
1	0.5000	0.2500	0.2174	0.5000	0.1250	0.2380
2	0.0001	0.1000	0.0870	0.0001	0.0001	0.0002
3	0.5000	0.8000	0.6957	0.5000	0.4000	0.7618
1	0.2000	0.2500	0.2083	0.0001	0.0500	0.1163
2	0.4000	0.1500	0.1250	0.5000	0.0600	0.1395
3	0.4000	0.8000	0.6667	0.5000	0.3200	0.7442
1	0.1000	0.2500	0.3125	0.0001	0.0250	0.0581
2	0.8000	0.5000	0.6250	1.0000	0.4000	0.9382
3	0.1000	0.0500	0.0625	0.0001	0.0050	0.0116

Tableau 7.14

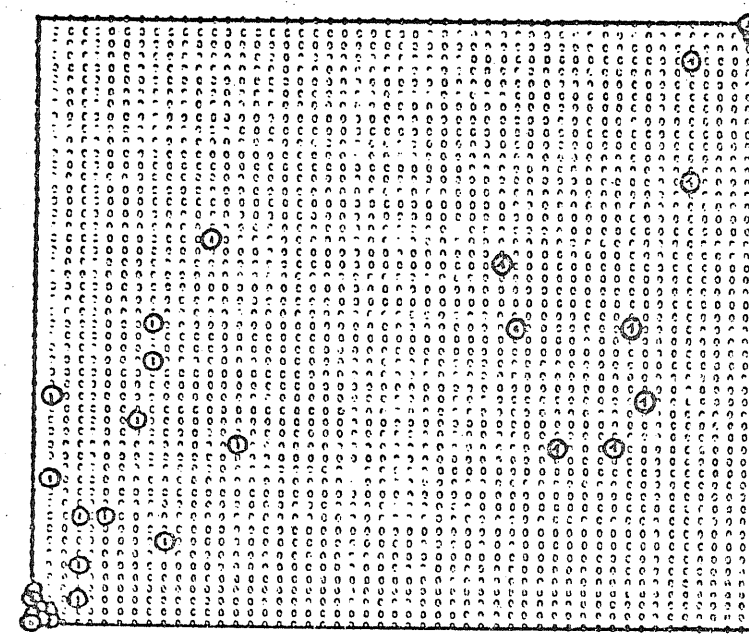


r = .6771

Résultats expérimentaux au jeu des "révision des ESPER"
pour le sujet n°2

	P(E|H) (2)	P(X|E+H) (3)	INF.REL.	P(E|X+H)	PRODUIT	P(E|X+H)
1	0.2000	0.0001	0.0001	0.0001	0.0001	0.0004
2	0.4500	0.0001	0.0001	0.0001	0.0001	0.0004
3	0.3500	0.7500	0.9997	1.0000	0.2625	0.9992
1	0.1000	0.0001	0.0001	0.0001	0.0001	0.0002
2	0.4500	0.9500	0.8260	0.5000	0.4275	0.8259
3	0.4500	0.2000	0.1739	0.5000	0.0900	0.1739
1	0.0500	0.0500	0.1111	0.0500	0.0025	0.0122
2	0.6000	0.2500	0.5556	0.3000	0.1500	0.7317
3	0.3500	0.1500	0.3333	0.6500	0.0525	0.2561
1	0.1000	0.6500	0.3939	0.2000	0.0650	0.1161
2	0.4000	0.0500	0.0303	0.4000	0.0200	0.0357
3	0.5000	0.9500	0.5758	0.4000	0.4750	0.8482
1	0.9000	0.0001	0.0002	0.0001	0.0001	0.0015
2	0.1000	0.6500	0.9997	1.0000	0.0650	0.9969
3	0.0001	0.0001	0.0002	0.0001	0.0001	0.0015
1	0.5500	0.8000	0.4103	0.5000	0.4400	0.6617
2	0.1500	0.8000	0.4103	0.1500	0.1200	0.1805
3	0.3000	0.3500	0.1795	0.3500	0.1050	0.1579
1	0.4500	0.8500	0.5484	0.6000	0.3825	0.6538
2	0.3000	0.5500	0.3548	0.3000	0.1650	0.2821
3	0.2500	0.1500	0.0968	0.1000	0.0375	0.0641
1	0.1000	0.1000	0.0588	0.0001	0.0100	0.0164
2	0.0500	0.9500	0.5588	0.0500	0.0475	0.0779
3	0.8500	0.6500	0.3824	0.9500	0.5525	0.9057
1	0.8000	0.5000	0.3333	0.3000	0.4000	0.8000
2	0.1000	0.9000	0.6000	0.4500	0.0900	0.1800
3	0.1000	0.1000	0.0667	0.2500	0.0100	0.0200
1	0.2000	0.2500	0.1786	0.2000	0.0500	0.0678
2	0.7500	0.9000	0.6429	0.7500	0.6750	0.9153
3	0.0500	0.2500	0.1786	0.0500	0.0125	0.0169

Tableau 7.15

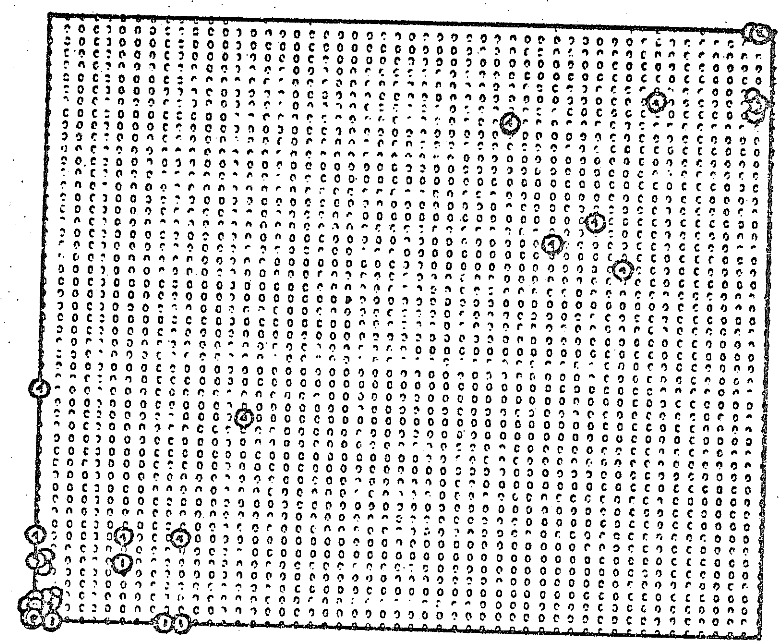


$r = .8109$

Résultats expérimentaux au jeu "révision des ESPER"
pour le sujet n°3

	P(E|H) (2)	P(X|E+H) (3)	INF.REL.	P(E|X+H)	PRODUIT	P(E|X+H
1	0.1000	0.0001	0.0003	0.0001	0.0001	0.0004
2	0.2000	0.0001	0.0003	0.0001	0.0001	0.0004
3	0.7000	0.3500	0.9994	1.0000	0.2450	0.9992
1	0.0001	0.1000	0.1666	0.0001	0.0001	0.0002
2	0.2000	0.0001	0.0002	0.0001	0.0001	0.0002
3	0.8000	0.5000	0.8332	1.0000	0.4000	0.9995
1	0.1500	0.1000	0.0952	0.0001	0.0150	0.0302
2	0.6500	0.6500	0.6190	0.9000	0.4225	0.8492
3	0.2000	0.3000	0.2857	0.1000	0.0600	0.1206
1	0.5000	0.6500	0.8666	0.9000	0.3250	0.9994
2	0.5000	0.0001	0.0001	0.0001	0.0001	0.0003
3	0.0001	0.1000	0.1333	0.1000	0.0001	0.0003
1	0.8500	0.0001	0.0001	0.0001	0.0001	0.0007
2	0.1500	0.9000	0.9998	0.9000	0.1350	0.9985
3	0.0001	0.0001	0.0001	0.1000	0.0001	0.0007
1	0.0001	0.3000	0.2069	0.4000	0.0001	0.0002
2	0.3000	0.4000	0.2759	0.0001	0.1200	0.1860
3	0.7000	0.7500	0.5172	0.6000	0.5250	0.8138
1	0.7500	0.6500	0.4333	0.7000	0.4875	0.7769
2	0.2000	0.6500	0.4333	0.1500	0.1300	0.2072
3	0.0500	0.2000	0.1333	0.1500	0.0100	0.0159
1	0.3000	0.0001	0.0001	0.1000	0.0001	0.0002
2	0.0001	0.4000	0.3636	0.0001	0.0001	0.0002
3	0.7000	0.7000	0.6363	0.9000	0.4900	0.9996
1	0.0001	0.1000	0.0606	0.0001	0.0001	0.0001
2	0.7000	0.8000	0.4848	0.6500	0.5600	0.7133
3	0.3000	0.7500	0.4545	0.3500	0.2250	0.2866
1	0.2500	0.5000	0.3226	0.0001	0.1250	0.2174
2	0.5000	0.7500	0.4839	0.8500	0.3750	0.6522
3	0.2500	0.3000	0.1935	0.1500	0.0750	0.1304

Tableau 7.16

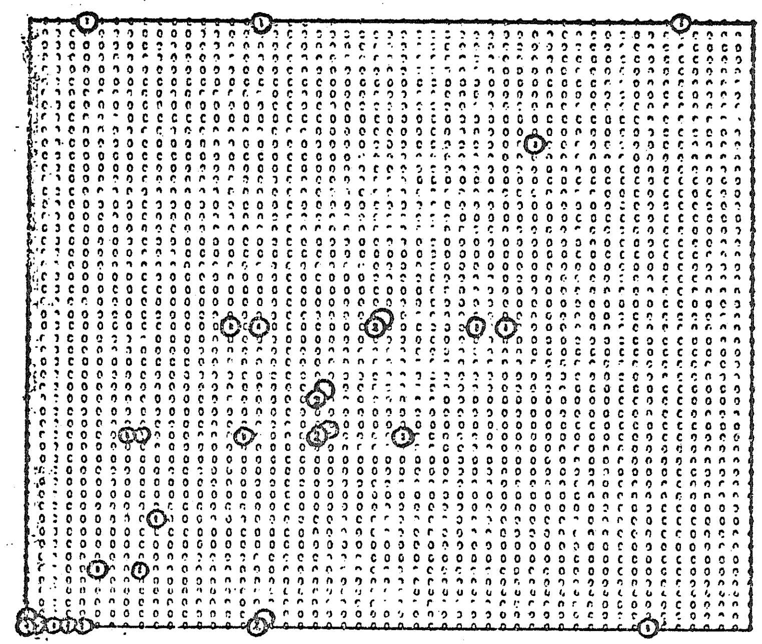


$r = .9549$

Résultats expérimentaux au jeu "révision des ESPER" pour le sujet n°4

	P(E|H) (2)	P(X|E+H) (3)	INF.REL.	P(E|X+H)	PRODUIT	P(E|X+H)
1	0.0001	0.0001	0.0002	0.0001	0.0001	0.3333
2	0.0001	0.0001	0.0002	0.0001	0.0001	0.3333
3	0.0001	0.5000	0.9996	1.0000	0.0001	0.3333
1	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002
2	0.5000	0.5000	0.5000	0.5000	0.2500	0.4999
3	0.5000	0.5000	0.5000	0.5000	0.2500	0.4999
1	0.1000	0.1500	0.3333	1.0000	0.0150	0.0811
2	0.8000	0.2000	0.4444	0.0001	0.1600	0.8649
3	0.1000	0.1000	0.2222	0.0001	0.0100	0.0541
1	0.2000	0.3000	0.2727	0.3300	0.0600	0.1579
2	0.4000	0.3000	0.2727	0.3300	0.1200	0.3158
3	0.4000	0.5000	0.4545	0.3300	0.2000	0.5263
1	0.8000	0.0001	0.0001	0.0001	0.0001	0.0007
2	0.1000	0.5000	0.3333	0.5000	0.0500	0.3331
3	0.1000	1.0000	0.6666	0.5000	0.1000	0.6662
1	0.5000	0.5000	0.4545	0.5000	0.2500	0.6250
2	0.1000	0.3000	0.2727	0.0001	0.0300	0.0750
3	0.4000	0.3000	0.2727	0.5000	0.1200	0.3000
1	0.3300	0.5000	0.4167	0.3300	0.1650	0.4167
2	0.3300	0.5000	0.4167	0.3300	0.1650	0.4167
3	0.3300	0.2000	0.1667	0.3300	0.0660	0.1667
1	0.5000	0.1000	0.0625	0.0001	0.0500	0.0909
2	0.0001	0.5000	0.3125	0.0001	0.0001	0.0002
3	0.5000	1.0000	0.6250	1.0000	0.5000	0.9089
1	0.2000	0.2000	0.3333	0.2000	0.0400	0.2000
2	0.4000	0.2000	0.3333	0.4000	0.0800	0.4000
3	0.4000	0.2000	0.3333	0.4000	0.0800	0.4000
1	0.2500	0.1500	0.2727	0.1000	0.0375	0.1765
2	0.5000	0.3000	0.5455	0.8000	0.1500	0.7059
3	0.2500	0.1000	0.1818	0.1000	0.0250	0.1176

Tableau 7.17

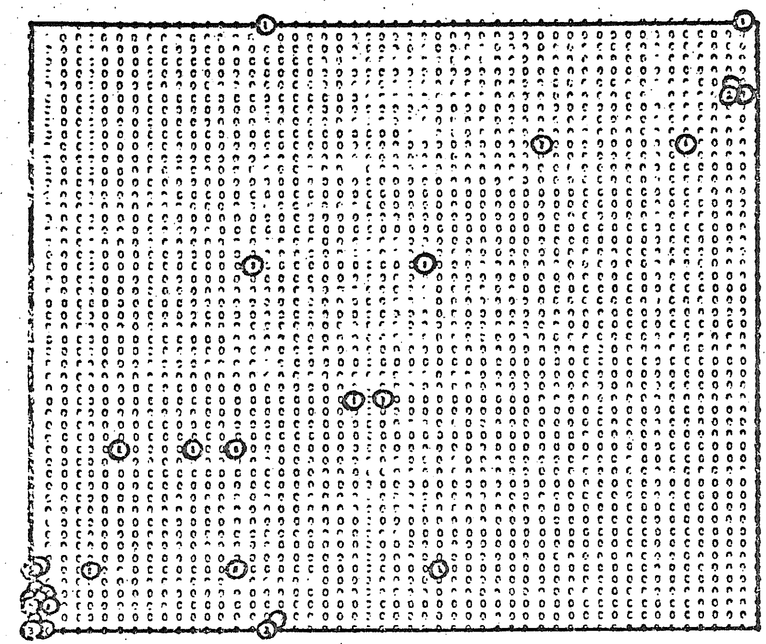


r = .4717

Résultats expérimentaux au jeu "révision des ESPER" pour le sujet n°5

Tableau 7.18

	P(E|H) (2)	P(X|E+H) (3)	INF.REL.	P(E|X+H)	PRODUIT	P(E|X+H)
1	0.0500	0.0001	0.3333	0.0001	0.0001	0.3333
2	0.0500	0.0001	0.3333	0.0001	0.0001	0.3333
3	0.9000	0.0001	0.3333	1.0000	0.0001	0.3333
1	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
2	0.5000	0.6000	0.4444	0.4000	0.3000	0.4444
3	0.5000	0.7500	0.5555	0.6000	0.3750	0.5555
1	0.0500	0.0500	0.1000	0.1000	0.0025	0.0103
2	0.5500	0.4000	0.8000	0.8000	0.2210	0.9072
3	0.4000	0.0500	0.1000	0.1000	0.0200	0.0825
1	0.3200	0.2500	0.2941	0.3000	0.0600	0.2857
2	0.3200	0.2000	0.2353	0.3000	0.0640	0.2286
3	0.3400	0.4000	0.4706	0.4000	0.1360	0.4857
1	0.4800	0.0001	0.0001	0.0500	0.0001	0.0003
2	0.0400	0.0500	0.0583	0.0500	0.0020	0.0052
3	0.4800	0.8000	0.9411	0.9000	0.3840	0.9946
1	0.6000	0.3000	0.300	0.1000	0.1800	0.5625
2	0.2000	0.5000	0.5000	0.6000	0.1000	0.3125
3	0.2000	0.2000	0.2000	0.3000	0.0400	0.1250
1	0.8000	0.5000	0.7692	0.9000	0.4000	0.9639
2	0.1000	0.0500	0.0769	0.0500	0.0050	0.0120
3	0.1000	0.1000	0.1538	0.0500	0.0100	0.0241
1	0.0500	0.2000	0.2105	0.0001	0.0100	0.0146
2	0.0500	0.0001	0.0001	0.0001	0.0001	0.0001
3	0.9000	0.7500	0.7894	1.0000	0.6750	0.9853
1	0.7000	0.0500	0.1163	0.1000	0.0350	0.2823
2	0.0500	0.0300	0.0698	0.1000	0.0015	0.0121
3	0.2500	0.3500	0.8140	0.8000	0.0875	0.7056
1	0.0500	0.1000	0.1818	0.0500	0.0050	0.0154
2	0.9000	0.3500	0.6364	0.9000	0.3150	0.9692
3	0.0500	0.1000	0.1818	0.0500	0.0050	0.0154

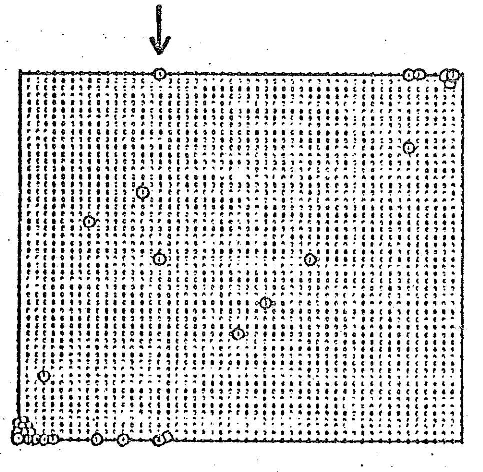


r = .8487

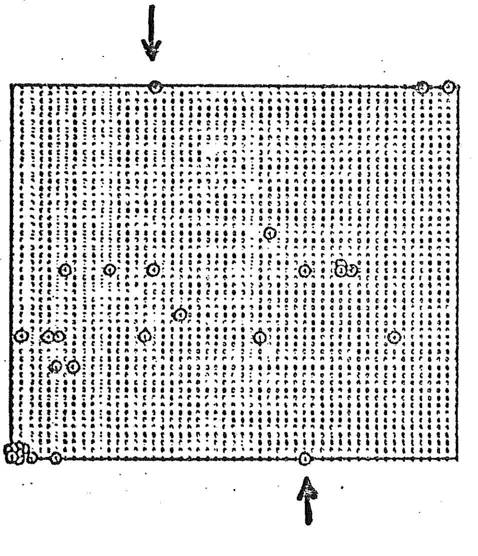
Résultats expérimentaux au jeu "révision des ESPER"
pour le sujet n° 6

Commentaires.

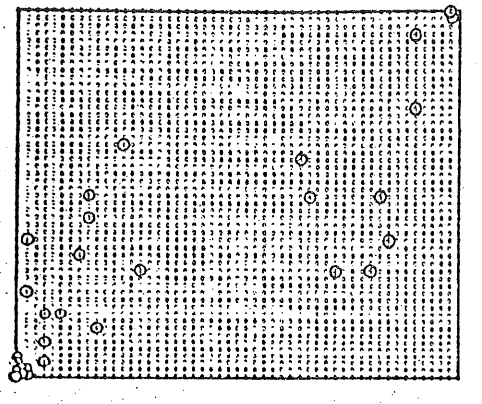
Le sujet 1 obtient une corrélation d'ensemble élevée (r=.8568). Les concordances entre le théorème et les observations sont parfaites aux extrêmes (probabilités nulles et totales), mais nettement moins bonnes dans les valeurs moyennes. La question 5 a amené une discordance particulièrement flagrante (voir flèche).



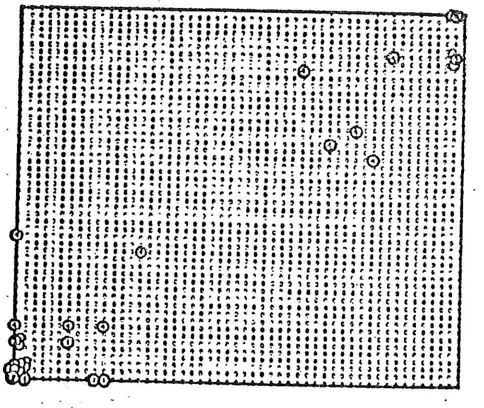
Pour le sujet 2, on observe une corrélation d'ensemble plus faible (r=.6771). Dans les valeurs moyennes, le sujet se cantonne entre .25 et .60). En deça, il répond 0; au-delà, il répond 1. La question 5 a amené deux discordances flagrantes (voir flèche).



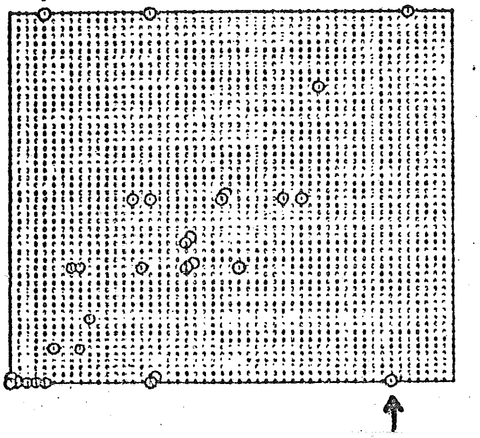
Le sujet 3 a une corrélation d'ensemble élevée (r=.8109). Le sujet utilise toutes les possibilités du continuum. Le nuage de points prend la forme générale d'une cubique; néanmoins, le nombre d'observations est trop faible pour confirmer une telle hypothèse.



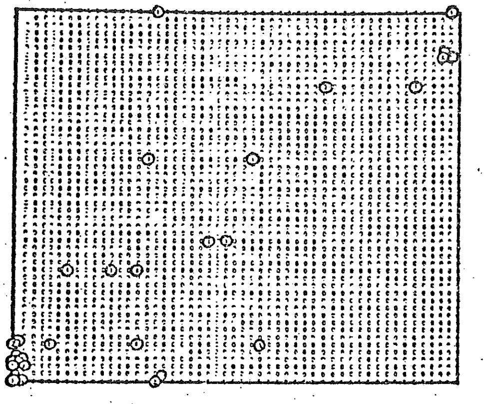
Le sujet 4 présente la corrélation la plus forte (r=.9549). Les probabilités, observées et calculées, sont fort dichotomisées : faibles ($<$.40) ou fortes ($>$.60).



Le sujet 5 présente la corrélation la plus faible (r=.4717). Les questions 1 et 3 provoquent des discordances importantes. Si on calcule la corrélation sur les huit autres questions, r=.9121.



Le sujet 6 présente une corrélation élevée (r=.8487). La question 1 fait apparaître une discordance importante. Si on calcule la corrélation sur les neuf autres questions, r=.9357.



Discussion.

Les discordances se produisent dans les questions 1, 5 et 3. En dehors de cinq cas (sur soixante questions) de discordance flagrante, la corrélation entre les valeurs observées et les valeurs fournies par le théorème est élevée (aux environs de .90).

Les cas de discordance constituent, au premier abord, une déception. Cependant, on pourrait tirer parti de tels phénomènes, par exemple pour déceler les situations où l'information n'a pas été traitée de façon algébrique, mais où une restructuration totale a eu lieu, modification de *gestalt*, *insight*, à un moment du processus de traitement de l'information.

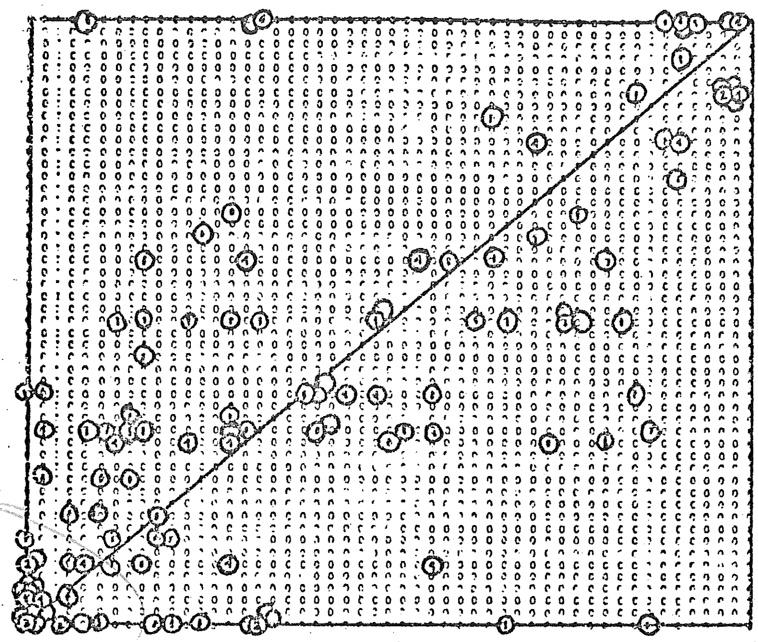
Le théorème de Bayes considère le traitement de l'information dans le cas où le nombre de possibilités envisagées par le sujet reste invariable. Les discordances entre les valeurs prédites par le théorème et les valeurs observées pourraient être révélatrices de situation où l'individu envisage, grâce à l'information, de nouvelles possibilités.

On peut regretter que le théorème de Bayes ne s'applique pas à tous les cas. Il est possible de tirer parti de cette observation car on détient, par là même, un "indicateur d'insight", outil qui pourrait se révéler fécond en recherche.

Si on découpait un rectangle de la moitié de l'aire du graphique centré sur le centre de celui-ci, la corrélation serait nulle.

Sur le tableau 7.19 sont regroupés les 180 points des tableaux 7.13 à 7.18.

Tableau 7.19



La corrélation linéaire vaut .796. A la vue du nuage, on peut se demander si les points ne se répartissent pas selon une cubique plutôt que selon une droite. La variance totale, sur les 180 observations y (ESPER) est 21,16852.

Le tableau 7.20 présente les régressions polynomiales des trois premiers degrés : x = EXPER.

Tableau 7.20

degrés	r		
1	.7960	$y = 0,7929\ x + 0,0691$	
2	.7962	$y = 0,8559\ x^2 - 0,0688\ x + 0,0639$	
3	.8020	$y = 0,1004\ x^3 + 2,2808\ x^2 - 1,6517\ x + 0,0820$	
4	.8049	$y = 0,3901\ x^4 + 0,6689\ x^3 + 1,090\ x^2 - 1,4179\ x + 0,0775$	
	r^2	% de variance expliquée par la régression	
1	.6335	63,35 %	d
2	.6339	63,39 %	0,04 %
3	.6432	64,32 %	0,93 %
4	.6479	64,79 %	0,47 %

expliquez.

Je ne comprends pas

On constate que le deuxième degré polynomial n'ajoute qu'une très faible explication. Le troisième degré explique une part supplémentaire de la variance (près de 1%). Cependant, cette part est tellement mince qu'on peut se contenter du premier degré. Le quatrième degré ajoute peu et le cinquième degré n'ajoute pas d'explication supplémentaire significative.

Conclusions.

Bien mieux que n'importe quelle démonstration verbale, le théorème de Bayes convainc qu'on ne peut évaluer la valeur informative d'un médium que par rapport à l'état cognitif d'un individu (ou d'un groupe) donné à un moment donné. Une telle constatation n'a rien de décourageant, car on peut commencer à raisonner selon des analogies psychophysiques ou physiques. On peut définir l'unité d'information comme l'événement nécessaire pour faire évoluer l'ESPER d'une certaine quantité. L'information parfaite est celle qui amène le sujet à la certitude totale. On voit combien est négligée (sauf pour le calcul) l'ESPER de départ et combien une telle mesure de "l'information" est relative. Serait-ce parce que l'information présente des analogies avec la tension électrique qui se définit par la différence de potentiel entre deux points ?

Les humains ne sont pas toujours bayésiens dans la révision de leursprobabilités, alors qu'ils devraient l'être. Dans ce sens, le théorème de Bayes est normatif. Il décrit ce qui devrait être et non ce qui est.

Dans les expériences ci-dessus, les sujets se sont montrés fort bayésiens et on peut se réjouir de ce phénomène. On ne constate même pas le "conservatisme" humain (tendance à une plus grande modération) observé par de nombreux chercheurs (ROUANET, EDWARDS, etc.), mais critiqué par MANZ (1970).

En général, le théorème s'appplique bien dans neuf cas sur dix. On peut se demander si les exceptions ne constituent pas des cas où les sujets ont totalement remodelé leur point de vue, indépendamment des ESPER *a priori*.

Un espoir est ainsi permis : les cas (10 %) qui ne répondent pas au théorème sont peut-être indicatifs d'un processus non-bayésien : un *insight* remodèle complètement l'état cognitif du sujet et néglige (souvent même contredit) les ESPER *a priori*. Ainsi, ce qui paraît une faiblesse pourrait s'avérer utile.

Quelles auraient été les observations si le théorème et ses implications avaient été exposés aux sujets ou, mieux, si les sujets y avaient été exposés, par une procédure operante ? Dans quelle mesure des sujets plus jeunes diffèrent-ils dans la révision de leurs probabilités ? Quelles sont les tâches qui se prêtent le mieux au traitement continu de l'information, et celles qui y sont plus résistantes ? Autant de questions qui mériteraient des études spécifiques, où se combineraient l'approche clinique et la procédure expérimentale.

F. UNE EXPERIENCE EN SITUATION SCOLAIRE : le test relatif à l'usage du dictionnaire

1. Prise d'information dans la vie courante.

Avant d'aborder un problème scolaire, prenons, dans la vie courante, un exemple des ESPER dans la recherche d'informations supplémentaires.

Bien des voyageurs, au moment de monter dans le train, interpellent tout porteur de képi pour demander confirmation de la destination, surtout quand la ligne de chemin de fer leur est peu familière (à l'étranger, par exemple) et quand les renseignements manquent ou sont douteux. Il s'agit là d'un comportement adapté, car les chemins de fer sont, en général, inexplicablement avares de précisions spontanées.

Les horaires de chemins de fer comportent des subtilités dont se jouent les habitués, mais qui font problème aux néophytes; ceux-ci augmentent leurs ESPER en interrogeant le plus de monde possible. Interroger plus d'une personne signifie que l'on doute non pas de la personne elle-même, mais du comportement qui découlerait de sa réponse. Cette dernière, en effet, peut être biaisée par des erreurs à bien des niveaux : dans la question du voyageur, dans l'interprétation de cette question, dans la formulation de la réponse et dans l'interprétation de la réponse. D'ailleurs, un voyageur qui pose des questions à plusieurs employés de la gare, pose des questions légèrement différentes à chaque fois, intégrant bien souvent des éléments des réponses précédentes.

Dans un tout autre domaine (1), le policier en fonction, dans son premier contact avec quelqu'un, a une attitude neutre, également

(1) Comme nous l'a fait remarquer Torcuato Perez de Gusman (Université de Séville).

distante des extrêmes (le secours et la répression), recueillant des indices qui vont le porter vers l'un de ces deux pôles, mais avec des réserves, car les probabilités restent mitigées.

Il serait peut-être instructif de développer un "jeu" du voyageur de chemins de fer ou du policier pour voir fonctionner le théorème de Bayes. Ce théorème, en effet, peut dépasser le seul domaine cognitif. Il pourrait sans doute rendre compte de bien des comportements sociaux (prise de parole en groupes, etc.).

Des recherches ont été entreprises en psychiatrie par BEENEN (1970), en planification de la production par KIDD (1970), en psychologie sociale par Mc NEEL et MESSICK (1970) et par LOVIE et DAVIES (1970), en météorologie par MURPHY (1967-1970), EPSTEIN (1967, 1969) et WINKLER (1967-1970).

2. Présentation de l'expérience scolaire.

Nous désirions terminer ce chapitre par une expérience de révision des ESPER menée dans une situation scolaire. Le théorème de Bayes n'y sera pas utilisé, mais on verra à quel point il sert l'interprétation des résultats.

La consultation du dictionnaire a été choisie; il s'agit en effet d'une prise d'information répandue, standardisée, facilement contrôlable. C'est aussi le prototype de la "manipulation des ouvrages de référence", à laquelle l'école contemporaine attache beaucoup d'importance.

Dans un premier temps, les élèves (de deuxième année dans l'enseignement secondaire rénové) ont dû répondre à 17 questions (1), sans disposer de dictionnaire. C'est le prétest. Les feuilles de réponses ont alors été reprises. Les élèves ont ensuite répondu aux mêmes questions, mais, cette fois-ci, en disposant de trois dictionnaires (2): c'est le post-test.

---

(1) En fait, il y avait vingt questions, mais trois d'entre elles n'exigeaient pas l'usage du dictionnaire. Les résultats qui s'y rapportent ont été éliminés;

(2) Le Dictionnaire du Français contemporain, le Petit Larousse (illustré), le Petit Robert.

Chaque fois, les étudiants ont dû noter leur indice de certitude (trois degrés possibles, plus l'omission), selon une consigne probabiliste et des conséquences E.S.U. Lors du post-test, les élèves devaient noter, en plus, quel dictionnaire ils avaient utilisé et à quelle page exactement. Il était ainsi possible de savoir quand il y a effectivement eu recours au dictionnaire ou non (1).

3. Le problème des effets entre prétest et post-test.

Les termes "prétest" et "post-test" remplaceront ici les expressions "avant information" et "après information". Les notations proposées par Y. TOURNEUR (1974, vol. 2, p. 239 et sq) et J. CARDINET ( ) seront utilisées (avec quelques modifications).

Voici ces notations :

R = réussite
E = échec
RE = réussite au prétest et échec au post-test
N = nombre total d'observations (N = EE + RE + ER + RR)
E. = échec au prétest (E. = EE + ER)
.E = échec au post-test (.E = EE + RE)

Habituellement, on présente les résultats comme suit (tableau 7.21).

Tableau 7.21

		Avant (pré-test)		
		E.	R.	
Après	.E	EE	RE	.E
(post-test)	.R	ER	RR	.R
		E.	R.	N

(1) Le recueil des données et le codage des réponses ont été effectués par B. LUMINGU dans le cadre de son mémoire : Etude préalable à la construction d'un test diagnostique sur la consultation du dictionnaire (1974). Cette étude portant surtout sur la vérification de la qualité des questions et de hiérarchies entre objectifs, B. Lumingu n'a pu mener l'analyse qui suit. Nous reprenons donc les données qu'il a bien voulu recueillir sous notre direction et nous l'en remercions.

Désormais, EE, RE, ER, RR, E., R., .E, .R, N, représentent des répétitions d'observations. On peut calculer les indices proposés par Y. TOURNEUR (1974) :

(1) Indice de superfluidité : I.S. = (R./N) x 100

(2) Indice d'aberrance : I.B. = (RE/R.) x 100

(3) Indice d'adéquation spécifique : I.Ad.s = (ER/E.) x 100

(4) Coefficient d'Ad.S. = (ER-EE)/E.

(5) $= (ER-RE)^2/ (ER+RE)$

(6) Indice d'adéquation générale = ER/N.

Nous proposons maintenant une présentation inspirée de la première, mais plus complexe puisque les étudiants peuvent utiliser trois degrés de certitude et l'omission (tableau 7.22).

Tableau 7.22

		avant (pré-test) E.				R.		
après (post-test)		-3	-2	-1	0	1	2	3
.E	-3	E3E3	E2E3	E1E3	E0E3	R1E3	R2E3	R3E3
	-2	E3E2	E2E2	E1E2	E0E2	R1E2	R2E2	R3E2
	-1	E3E1	E2E1	E1E1	E0E1	R1E1	R2E1	R3E1
	0	E3E0	E2E0	E1E0	E0E0	R1E0	R1E0	R3E0
.R	1	E3R1	E2R1	E1R1	E0R1	R1R1	R2R1	R3R1
	2	E3R2	E2R2	E1R2	E0R2	R1R2	R2R2	R3R2
	3	E3R3	E2R3	E1R3	E0R3	R1R3	R2R3	R3R3

Matrice A = matrice des situations quand trois degrés de certitude et l'omission sont utilisés

Dans la matrice suivante, les cellules marquées d'un + représentent une amélioration, celles marquées d'un - une aggravation et celles marquées d'un 0 un *statut quo*.

Tableau 7.23

		E.				R.		
		-3	-2	-1	0	1	2	3
.E	-3	0	-	-	-	-	-	-
	-2	+	0	-	-	-	-	-
	-1	+	+	0	-	-	-	-
	0	+	+	+	0	-	-	-
.R	1	+	+	+	+	0	-	-
	2	+	+	+	+	+	0	
	3	+	+	+	+	+	+	0

Matrice B = matrice des effets de la matrice A, ou "matrice des effets mineurs"

On remarque déjà les nuances introduites dans les grandes cellules EE et RR. On appellera "effets mineurs" les effets compris à l'intérieur de chaque cellule du carré classique (à quatre cellules).

4. Les résultats bruts.

Quatre formes (non parallèles) de dix-sept questions ont été présentées à 32 sujets chacune. On dispose donc au total de 2176 réponses, puisque 128 sujets ont répondu à dix-sept questions. Présentés dans le carré classique, les résultats sont les suivants (tableau 7.24).

Tableau 7.24

		Prétest E.	Prétest R.	
Post-test	.E	848	169	1017
	.R	684	475	1159
		1532	644	2176

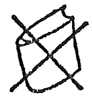
Matrice D

Indice de superfluité = .295
Indice d'aberrance = .449
Indice d'adéquation spécifique = -.107
Coefficient d'adéquation spécifique = 311.29
Indice d'adéquation générale = .3143

Cette matrice D ne permet pas de révéler deux sources importantes de variation des résultats : la première est l'utilisation effective du dictionnaire, la seconde la certitude dans la réponse.

Voici les deux matrices d'effets selon que le dictionnaire n'a pas été utilisé (tableau 7.25) ou qu'il a été utilisé (tableau 7.26).

Tableau 7.25



Post-test		Prétest E.	Prétest R.	
.E		EE 726	RE 139	865 .E
.R		ER 173	RR 244	417 .R
		E. 899	R. 383	1282

Matrice E

Tableau 7.26

Post-test		Prétest E.	Prétest R.	
.E		EE 122	RE 30	152 .E
.R		ER 511	RR 231	742 .R
		E. 633	R. 261	894

Matrice F

	Matrice E	Matrice F
IS = R./N	=.298	.291
IB = RE/R	=.362	.114
Ind. ad. spéc. : ER/E.	= .192	.807
Coef.ad. spéc.= (ER-EE)/E.	=-.615	.614
G(5) = $(ER-RE)^2/(ER+RE)$	=3,705 → table!	427,65 → table !
Ind. ad. gén.= ER/N	= .134	.571

On constate combien ces matrices sont différentes. Tous les coefficients (excepté I.S.) confirment l'importance (prévisible) de l'utilisation effective du dictionnaire.

Le tableau 7.27 tient compte des certitudes fournies avec les réponses.

Tableau 7.27

PRESTEST

	-3	-2	-1	0	1	2	3	
-3	100	106	65	16	24	15	15	341
-2	34	113	82	12	18	11	9	279
-1	12	59	126	4	22	8	8	239
0	12	26	46	35	13	13	13	158
+1	3	9	27	8	23	12	3	85
+2	9	25	43	14	23	52	19	185
+3	72	154	242	78	104	113	126	889
	252	492	631	167	226	224	193	2176

Matrice G : matrice des situations avec certitude (voir tableau 7.22) de l'expérience "dictionnaire"

Ici aussi il importe de créer deux matrices, selon que le dictionnaire n'a pas été utilisé (tableau 7.28) ou utilisé (tableau 7.29).

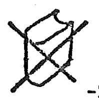


Tableau 7.28

	-3	-2	-1	0	1	2	3	
-3	84	77	38	12	16	9	11	247
-2	29	105	65	12	17	11	9	248
-1	11	53	121	4	20	8	7	224
0	11	26	43	35	10	9	12	146
1	1	4	17	6	19	10	3	60
2	4	13	25	2	17	33	13	107
3	20	26	38	17	26	47	76	250
	160	304	347	88	125	127	131	1282

MATRICE H

Le dictionnaire n'a pas été utilisé.

Tableau 7.29

	-3	-2	-1	0	3	2	3	
	16	29	27	4	8	6	4	94
	5	8	17	0	1	0	0	31
	1	6	5	0	2	0	1	15
	1	0	3	0	3	4	1	12
	2	5	10	2	4	2	0	25
	5	12	18	12	6	19	6	78
	52	128	204	61	78	66	50	639
	92	188	284	79	101	97	62	894

MATRICE I

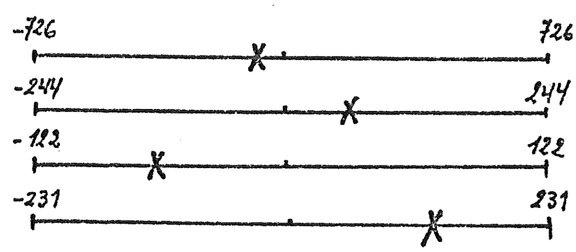
Le dictionnaire a été utilisé.

Du point de vue de l'évaluation globale des effets, qu'apporte l'indication du degré de certitude ? Elle permet d'évaluer les effets mineurs (+ , 0, -) décrits dans la matrice B ci-avant, pour les grandes cellules EE et RR. Ainsi, à partir des seules matrices E et F, les effets mineurs peuvent osciller.

- Pour la matrice E, pour EE : entre -726 et 726
  pour RR : entre -244 et 244
- Pour la matrice F, pour EE : entre -122 et 122
  pour RR : entre -231 et 231

Les moyennes des effets mineurs (tableau 7.30) ont été calculées par la matrice B :

- Pour la matrice H, pour EE : -35
  pour RR : 64
- Pour la matrice I, pour EE : -61
  pour RR : 142



Ce que les matrices E et F ne peuvent révéler tient dans les chiffres ci-dessus : la consultation du dictionnaire (matrice I) entraîne une augmentation de la certitude dans les réponses, ce qui donne des effets mineurs très positifs en RR, mais négatifs en EE.

Il est cependant possible de nuancer plus avant cette matrice des effets. Appelons θ l'effet d'une cellule. On peut poser les trois séries d'options suivantes (N.B.: les θ sont pris en valeur absolue.).

Tableau 7.31

Première série d'options

(1) $\theta(E_3E_3) = \theta(E_2E_2) = \theta(E_1E_1) = \theta(E_0E_0) = \theta(R_1R_1) = \theta(R_2R_2) = \theta(R_3R_3) = 0$

Deuxième série d'options

(2) $\theta(E_3E_2) = \theta(E_2E_3) > 0$

(3) $\theta(E_2E_1) = \theta(E_1E_2) > 0$

(4) $\theta(E_oE_1) = \theta(E_1E_0) > 0$

(5) $\theta(R_1E_o) = \theta(E_oR_1) > 0$

(6) $\theta(R_2R_1) = \theta(R_1R_2) > 0$

(7) $\theta(R_3R_2) = \theta(R_2R_3) > 0$

(8) $\theta(E_1E_3) = \theta(E_3E_1) >$ (2) , et > (3)

(9) $\theta(E_oE_2) = \theta(E_2E_o) >$ (3) , et > (4)

* (10) $\theta(R_1E_1) = \theta(E_1R_1) >$ (4) , et > (5)

* (11) $\theta(R_2E_o) = \theta(E_oR_2) >$ (5) , et > (6)

(12) $\theta(R_3R_1) = \theta(R_1R_3) >$ (6) , et > (7)

(13) $\theta(E_0E_3) = \theta(E_3E_o) >$ (8) , et > (9)

* (14) $\theta(R_1E_2) = \theta(E_2R_1) >$ (9) , et > (10)

* (15) $\theta(R_2E_1) = \theta(E_1R_2) >$ (10) , et > (11)

* (16) $\theta(R_3E_0) = \theta(E_oR_3) >$ (11) , et > (12)

* (17) $\theta(R_1E_3) = \theta(E_3R_1) >$ (13) , et > (14)

* (18) $\theta(R_2E_2) = \theta(E_2R_2) >$ (14) , et > (15)

* (19) $\theta(R_3E_1) = \theta(E_1R_3) >$ (15) , et > (16)

* (20) $\theta(R_2E_3) = \theta(E_3R_2) >$ (17) , et > (18)

* (21) $\theta(R_3E_2) = \theta(E_2R_3) >$ (18) , et > (19)

* (22) $\theta(R_3E_3) = \theta(E_3R_3) >$ (20) , et > (21)

Troisième série d'options

Les astérisques marquent des situations où il y a passage d'une réponse correcte à une réponse incorrecte et vice versa. Les autres situations ne reflètent que des changements de certitude. On peut décider que les premières *( ) reflètent un effet plus important que les secondes ( ). Cette option peut s'exprimer par

(23) $\theta(ER) + \theta(RE) > \theta(EE) + \theta(RR)$

Il existe une façon simple et systématique de donner des valeurs numériques aux effets, en partant de la matrice de conséquences. Dans notre exemple, cette matrice est la suivante (tableau 7.32).

Tableau 7.32

	E.	R.
OM	0	0
C1	-0,5	1,75
C2	-0,75	2
C3	-1,5	2,25

La matrice des effets (X) est reprise dans le tableau 7.33.

Tableau 7.33

pourquoi ce choix ?

Post-test \ Prétest	-3	-2	-1	0	+1	+2	+3
-3	0	-0,75	-1	-1,5	-3,25	-3,50	-3,75
-2	+0,75	0	-0,25	-0,75	-2,5	-2,75	-3
-1	+1	+0,25	0	-0,5	-2,25	-2,5	-2,75
0	+1,5	+0,75	+0,5	0	-1,75	-2	-2,25
+1	+3,25	+2,50	+2,25	+1,75	0	-0,25	-0,5
+2	+3,50	+2,75	+2,5	+2	+0,25	0	-0,25
+3	+3,75	+3	+2,75	+2,25	+0,5	+0,25	0

MATRICE C

Cette matrice correspond aux 23 options décrites plus haut. Elle n'est qu'une des innombrables matrices possibles et dépend entièrement des options prises lors de la création de la matrice des conséquences. On a vu, au chapitre 3, qu'il n'existe pas de moyen objectif de déterminer de telles options.

Si on applique la matrice C (tableau 7.33) sur la matrice H (tableau 7.28 : non utilisation du dictionnaire) :

Totaux

-37,5	-371,25
474	27,5

Moyennes

-0,051	-2,67
2,739	0,112

Tot = 92,75

Moy = 0,072

Tableau 7.34

		Prétest -3	-2	-1	0	+1	+2	+3
Post-test	-3	0	-57,75	-38	-18	-52	-31,5	-41,25
	-2	21,75	0	-16,25	- 9	-42,5	-30,25	-27
	-1	11	13,25	0	- 2	-45	-20	-19,25
	0	16,5	19,5	21,5	0	-17,5	-18	-27
	1	3,25	10	38,25	10,5	0	- 4	- 1,5
	2	14	35,75	62,5	4	4,25	0	- 3,25
	3	75	78	104,5	38,25	13	19	0

Matrice CxH obtenue par combinaison des matrices C (tableau 7.33) et de la matrice H (tableau 7.28) : items pour lesquels le dictionnaire n'a pas été utilisé

Si on applique la matrice C sur la matrice I (tableau 7.29 : utilisation du dictionnaire) :

Totaux

-49,75	-87,25
1441,75	55

Moyennes

-0,407	-2,908
2,821	0,238

Tot. = 1359,75

Moy. = 1,521

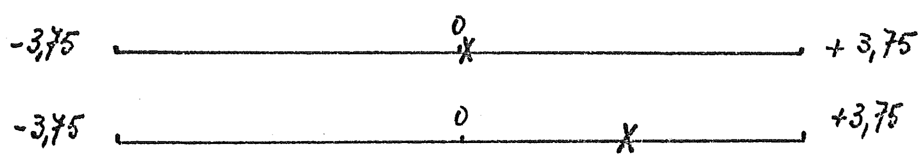
Tableau 7.35

		Prétest -3	-2	-1	0	+1	+2	+3
Post-test	-3	0	-21,75	-27	-6	-26	-21	-15
	-2	3,75	0	- 4,25	0	- 2,5	0	0
	-1	1	1,5	0	0	- 4,5	0	- 2,75
	0	1,5	0	1,5	0	- 5,25	-8	- 2,25
	+1	6,5	12,5	22,5	3,5	0	-0,5	0
	+2	17,5	33	45	24	1,5	0	- 1,5
	+3	195	384	561	137,25	39	16,5	0

Matrice CxI obtenue par combinaison des matrices C (tableau 7.33) et de la matrice I (tableau 7.29) : items pour lesquels le dictionnaire a été utilisé

Les moyennes 0,072 et 1,521 nous paraissent être le reflet le plus exact de la situation, car elles tiennent compte des effets mineurs en plus des effets principaux. Les différences principales entre l'utilisation (matrice I) et la non-utilisation (matrice H) du dictionnaire se produisent en EE et RR.

Ces moyennes peuvent être rapportées aux valeurs maximales possibles (ou valeurs extrêmes) :



Elles peuvent donc être exprimées par une valeur allant de -1 à + 1: 0,072 / 3,75 = .0192 et 1,521 / 3,75 = .4056.

Il ne faut pas surestimer l'importance de l'indication de la certitude. Elle reste un effet mineur, et la matrice de cotation est la preuve de notre volonté de lui donner cette importance secondaire (cf. au chapitre 3, où D $>$ d). Le travail effectué ci-dessus pourrait être repris au niveau de chaque item et constituerait un enrichissement possible pour l'analyse d'items.

°

° °

L'indice de certitude, cependant, permet bien plus que cette approche globale. Ainsi, si on s'intéresse aux mécanismes de production de la réponse, le degré de certitude occupe une part non négligeable dans les schémas explicatifs.

Tout d'abord, il faut scinder les matrices H et I en deux matrices chacune, selon que le sujet a changé (H" et I") ou non (H' et I") sa réponse entre le pré- et le post-test (tableaux 7.37 à 7.40).

$R_x \rightarrow R_x$

	-3	-2	-1	0	+1	+2	+3	
-3	65	50	21					136
-2	20	72	45					137
-1	7	30	70					107
0				35				35
+1					19	10	3	32
+2					17	33	13	63
+3					26	47	76	149
	92	152	136	35	62	90	92	659

MATRICE H'

	-3	-2	-1	0	+1	+2	+3	
-3	10	14	8					32
-2	4	4	9					17
-1	0	3	3					6
0								
+1					4	2	0	6
+2					6	19	6	31
+3					78	66	50	194
	14	21	20		88	87	56	286

MATRICE I'

$R_x \rightarrow R_y$

	-3	-2	-1	0	+1	+2	+3	
-3	19	27	17	12	16	9	11	111
-2	9	33	20	12	17	11	9	111
-1	4	23	51	4	20	8	7	117
0	11	26	43		10	9	12	111
+1	1	4	17	6				28
+2	4	13	25	2				44
+3	20	26	38	17				101
	68	152	211	53	63	37	39	623

MATRICE H"

	-3	-2	-1	0	+1	+2	+3	
-3	6	15	19	4	8	6	4	62
-2	1	4	8	0	1	0	0	14
-1	1	3	2	0	2	0	1	9
0	1	0	3		3	4	1	12
+1	2	5	10	2				19
+2	5	12	18	12				47
+3	52	128	204	61				445
	68	167	264	79	13	10	6	608

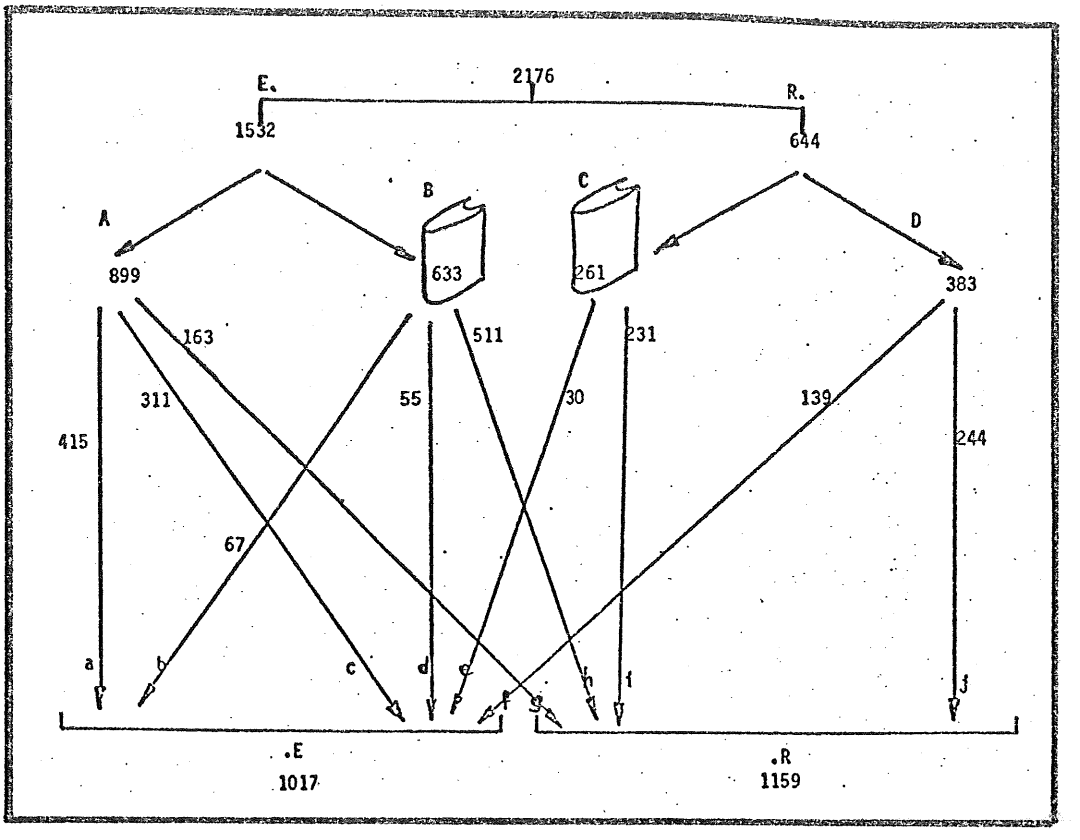
MATRICE I"

Il est possible de présenter en un seul schéma l'ensemble des données disponibles concernant le processus de décision dans la consultation du dictionnaire.

Dans les tableaux 7.4 à 7.50, les conventions graphiques sont les suivantes :

- Les flèches représentent des décisions prises par le sujet.
- Les embranchements non flèchés représentent des états de la situation non décidés par le sujet.
- BR = bonne réponse (correcte); MR = mauvaise réponse (incorrecte).
- Le haut du graphique est réservé au prétest, le bas au post-test.
- Dans le milieu du graphique, les flèches verticales symbolisent une réponse semblable au pré et au post-test, tandis que les flèches obliques symbolisent une modification de la réponse. Une telle modification entraîne toujours une .e (erreur au post-test) quand la réponse est R. (réussite au prétest). Mais quand la réponse est E. (au prétest), une modification peut amener une réponse R. ou une autre E.
- Les calculs de pourcentages sont effectués par référence aux nombres juste supérieurs.

Tableau 7.41



Ensemble des situations schématisé à partir des données des matrices H', H", I' et I" (tableau 7.37 à 7.40). Aux cinq niveaux horizontaux, le total vaut toujours 2176

Tableau 7.42

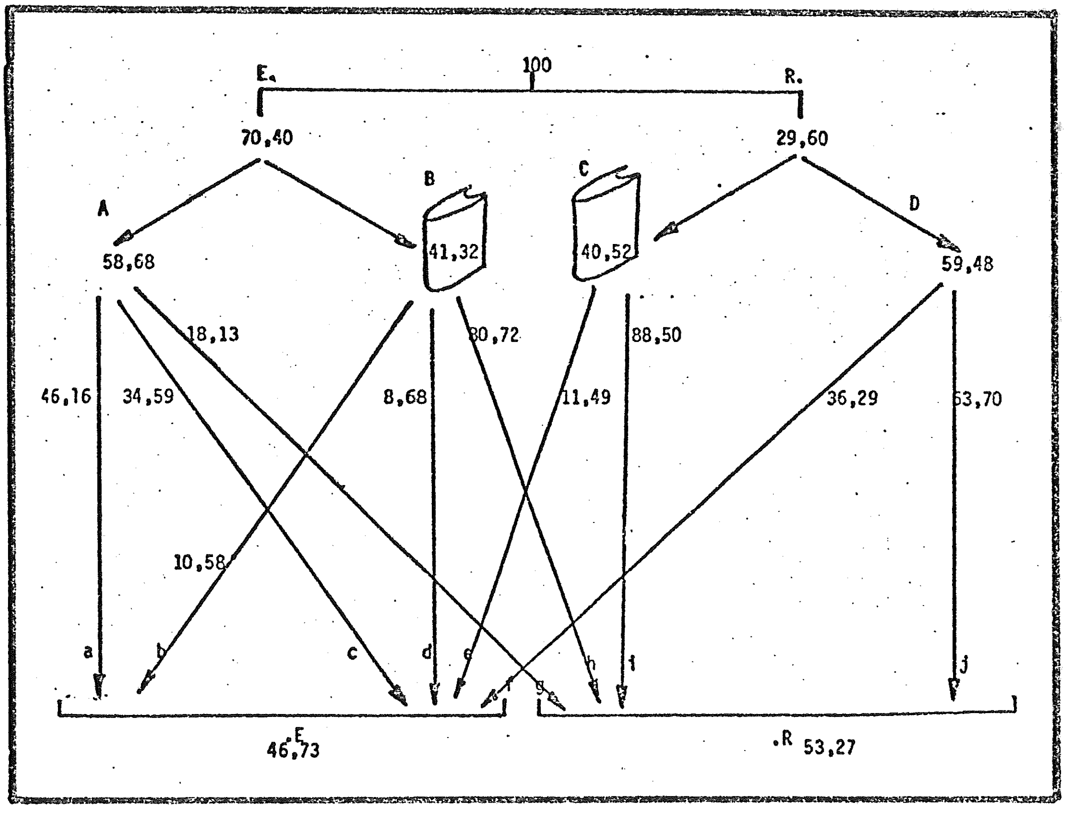


Tableau 7.41 exprimé en pourcentages. Les pourcentages sont calculés sur les répétitions au niveau horizontal juste supérieur, sauf pour le niveau inférieur

Tableau 7.43

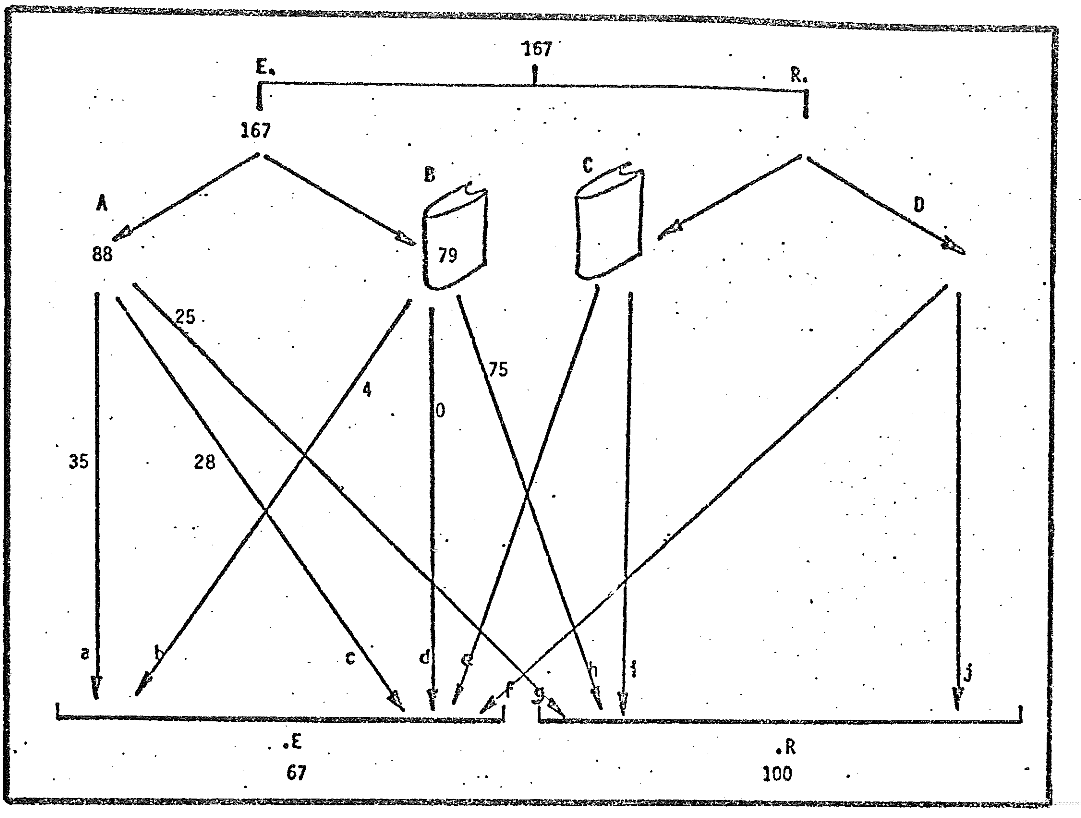


Schéma des situations pour les items omis lors du prétest
(sous ensemble du tableau 7.41)

Tableau 7.44

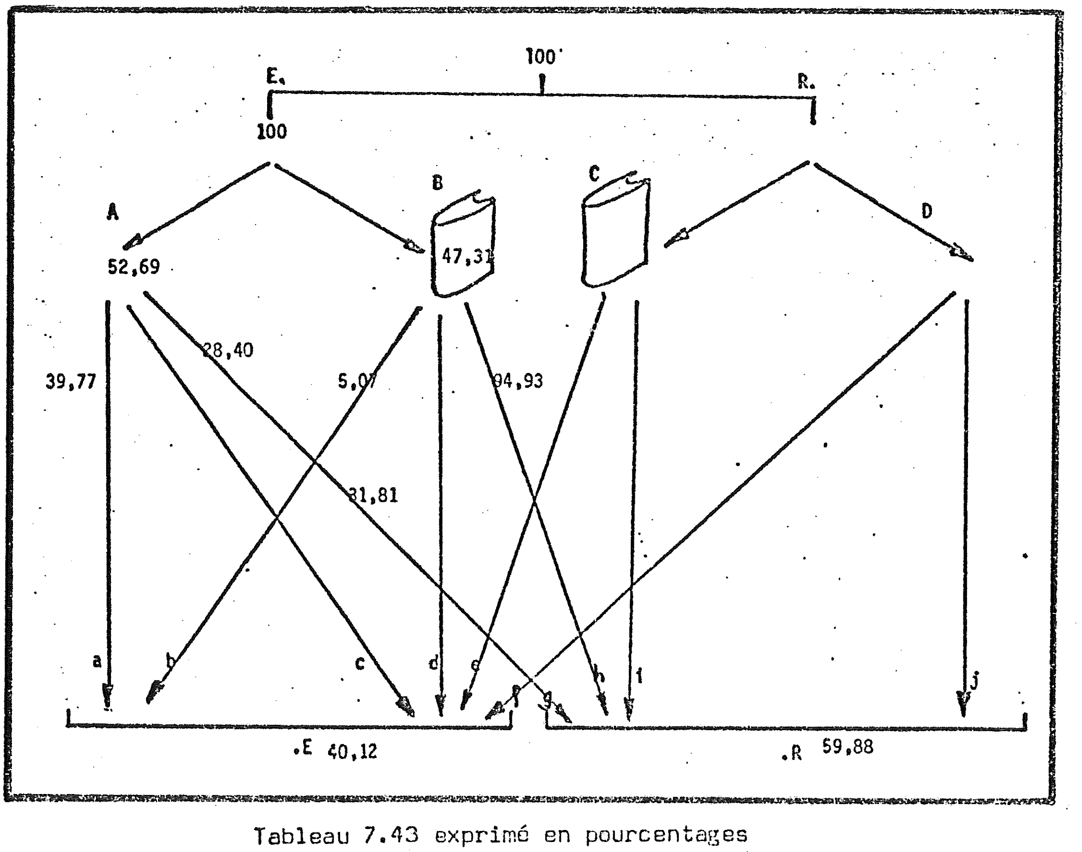


Tableau 7.43 exprimé en pourcentages

Tableau 7.45

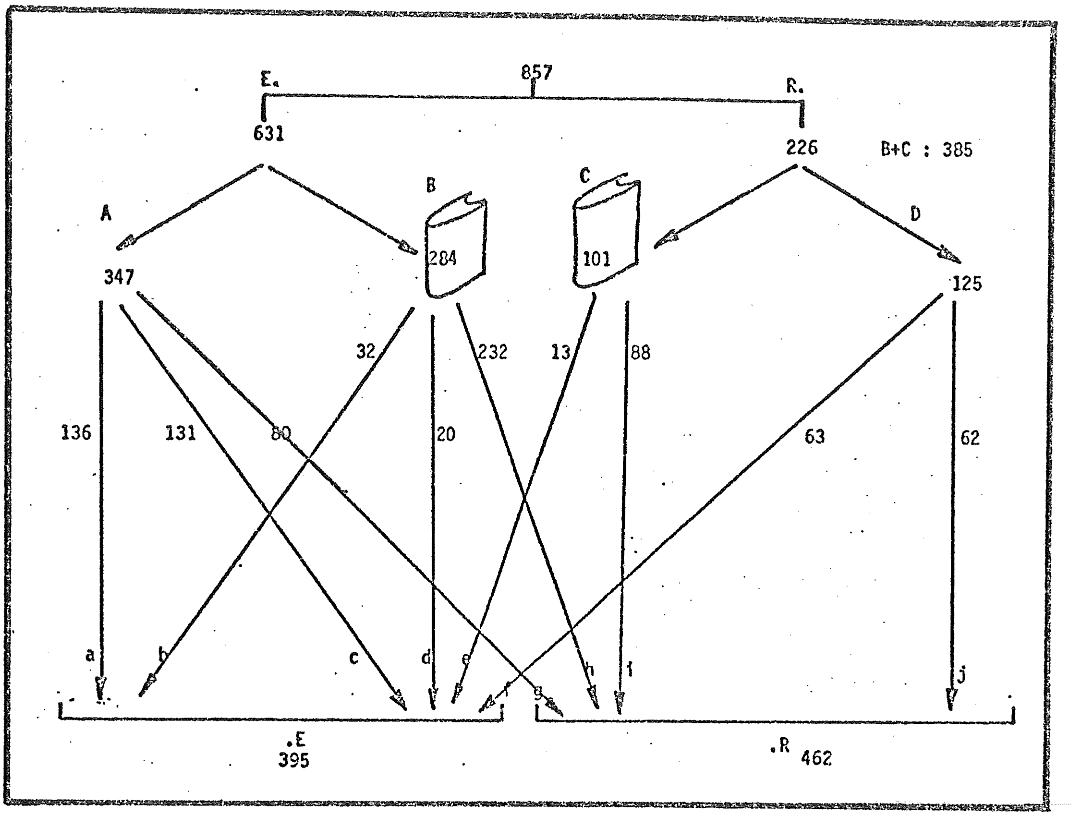


Schéma des situations pour les items dont la réponse, au prétest, était accompagnée de la certitude 1

Tableau 7.46

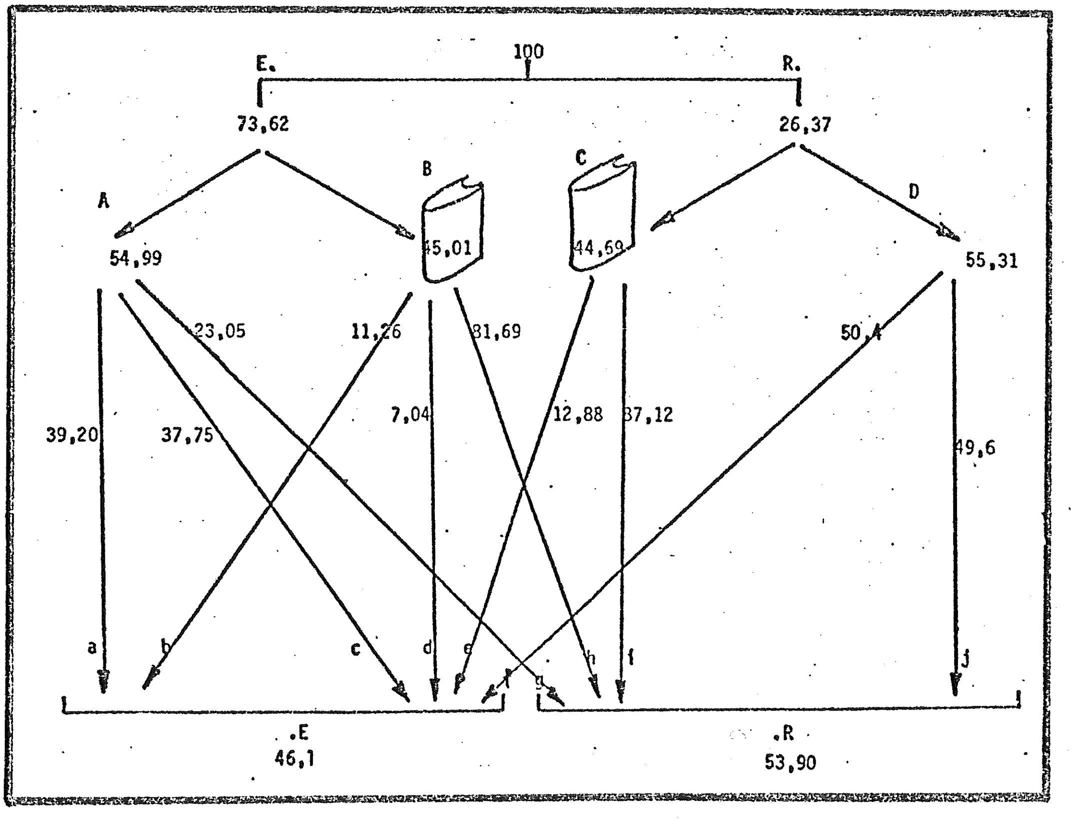


Tableau 7.45 exprimé en pourcentages

Tableau 7.47

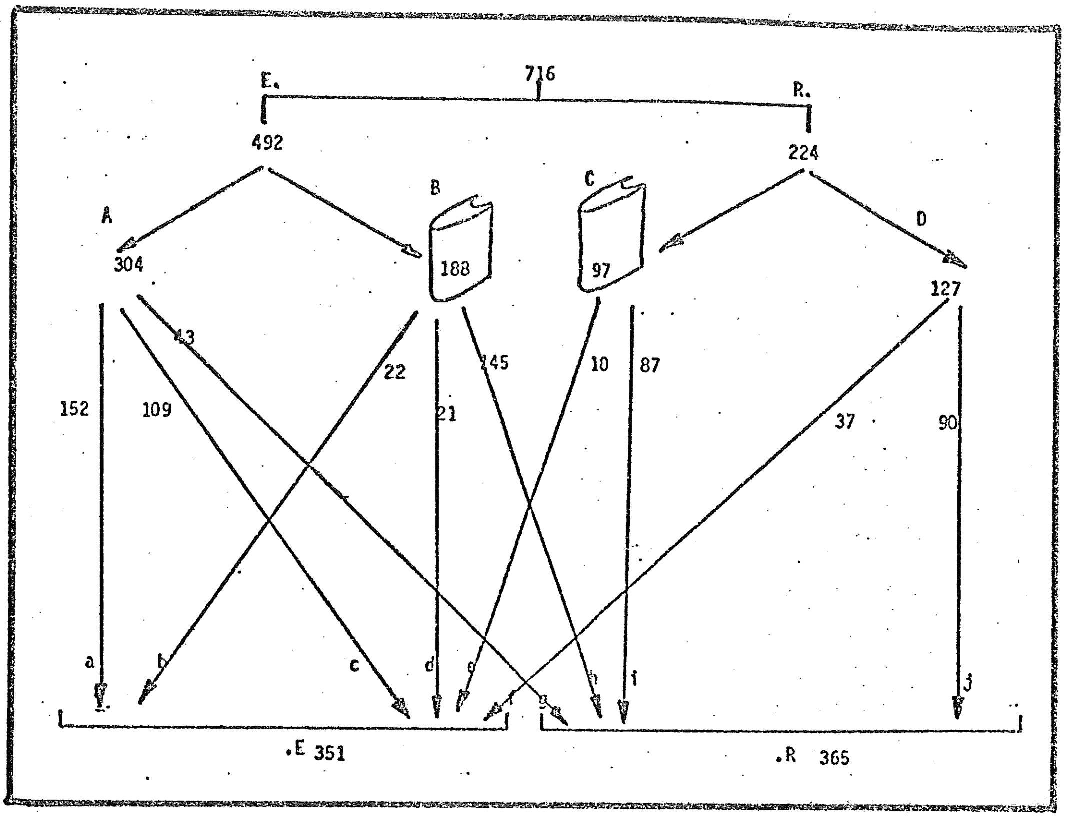


Schéma des situations pour les items dont la réponse, au prétest était accompagnée de la certitude 2

Tableau 7.48

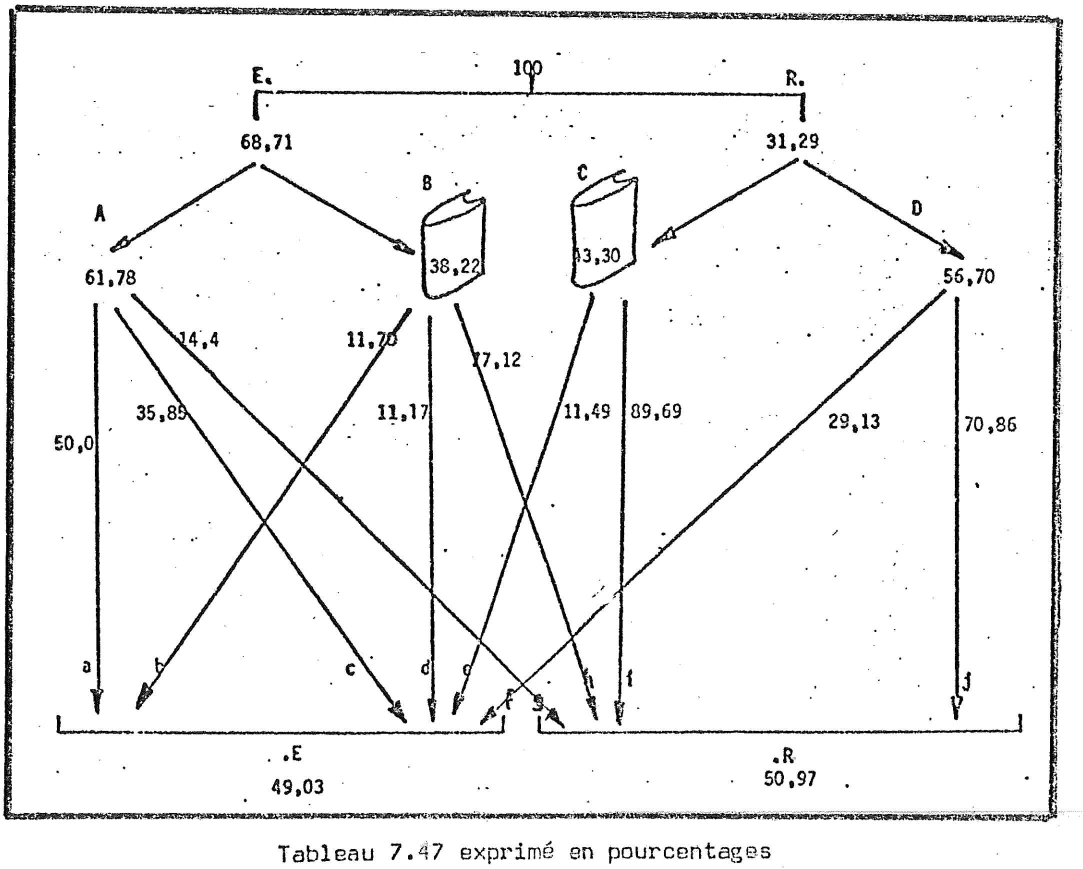


Tableau 7.47 exprimé en pourcentages

Tableau 7.49

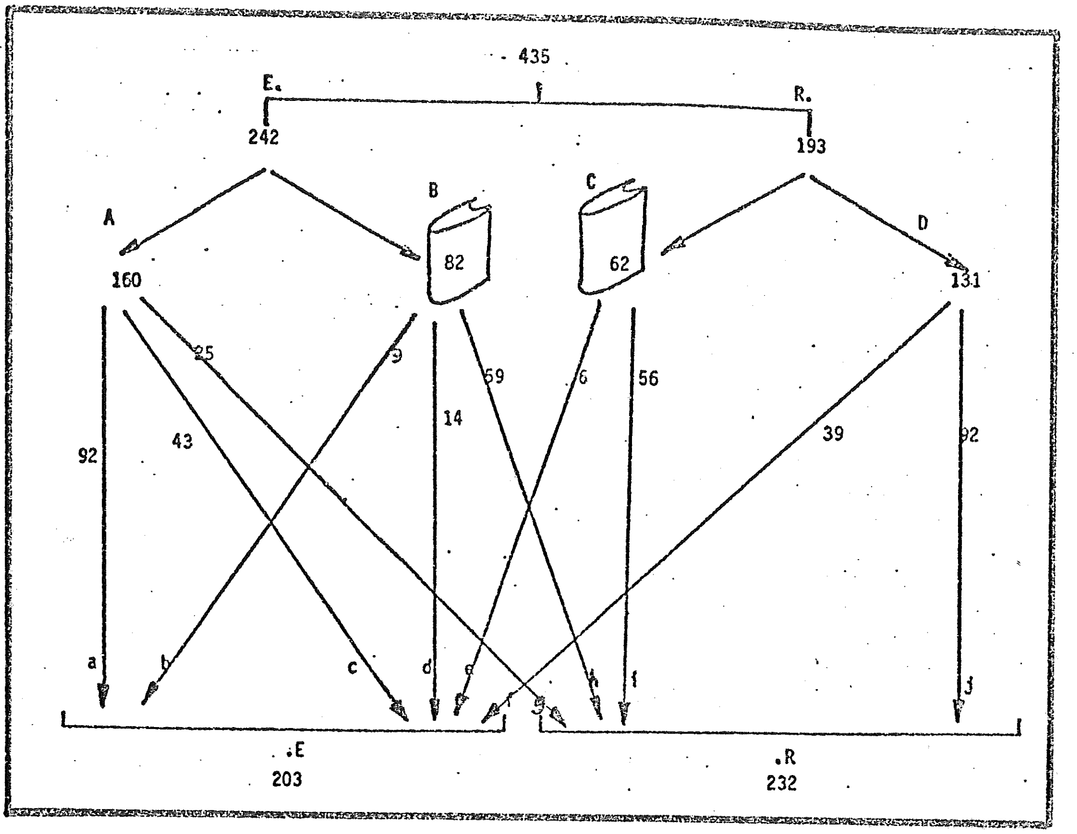


Schéma des situations pour les items dont la réponse, au prétest était accompagnée de la certitude 3

Tableau 7.50

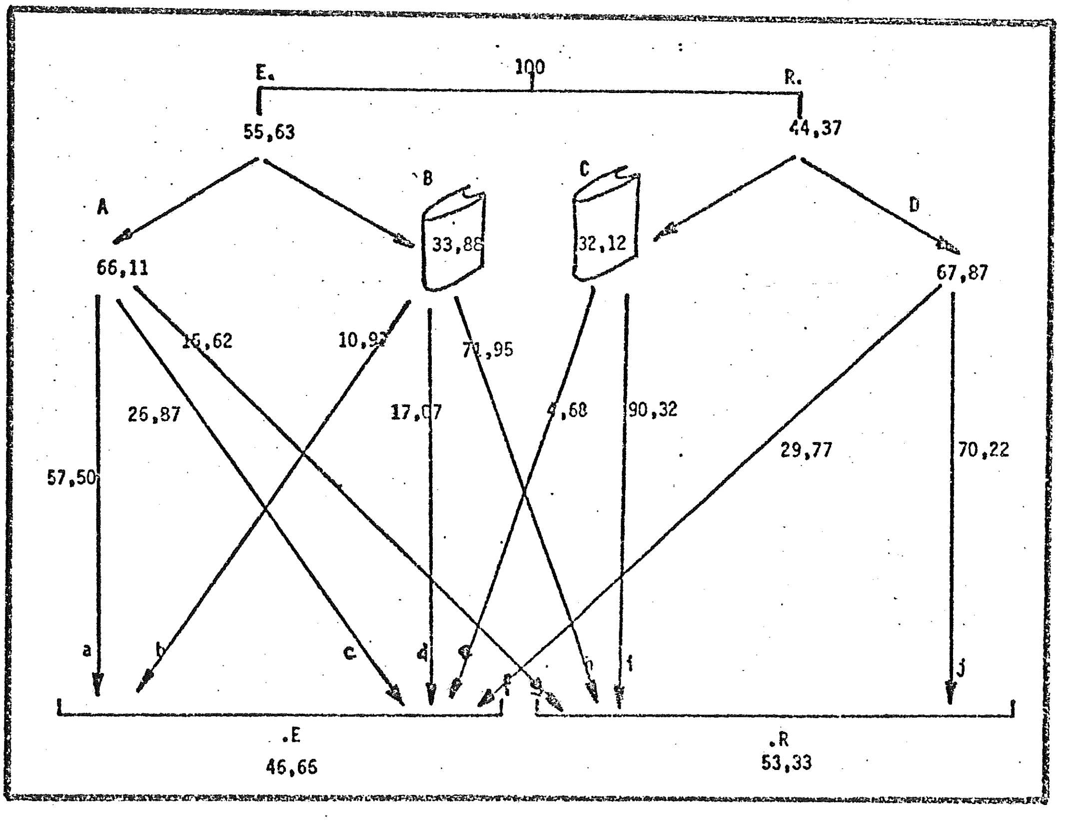


Tableau 7.49 exprimé en pourcentages

Dans l'expérience présente, trois conditions ne correspondent pas aux exigences d'une utilisation optimale des degrés de certitude.

1.- Les élèves n'ont pas été entraînés à exprimer leurs ESPER; l'expérience actuelle est, pour eux, la première du genre.

2.- Les sujets sont jeunes (13-14 ans).

3.- La situation n'est pas operante : les élèves ne seront pas exposés aux conséquences de leurs actes.

Les deux premières conditions ne peuvent qu'entraîner une sous-estimation de l'influence de l'expression du degré de certitude sur les comportements. Paradoxalement, la troisième condition est plutôt favorable à l'étude des mécanismes. En effet, les élèves ont eu recours au dictionnaire dans 41,08 % des cas. Cette valeur est relativement proche de 50 %. Si la situation avait été operante, on peut penser que le taux d'utilisation du dictionnaire eût été beaucoup plus élevé et peut-être trop proche de 100 % pour permettre de disposer de données diversifiées.

Quoi qu'il en soit, les hypothèses qui suivent restent très générales (elles portent la mention HD : hypothèses relatives au dictionnaire).

HD1. Quand on a fourni une bonne réponse au prétest (R.),on recourt moins au dictionnaire que lorsqu'on s'est trompé au prétest (E.).

Le tableau 7.51 permet de tester cette hypothèse.

Tableau 7.51

	E.(prétest)	R.(prétest)	
Non recours	899	383	1282
Recours	633	261	894
Total	1532	644	2176
% recours	41,32 %	40,52 %	

χ^2= 0,112 (N.S.)

Cette infirmation de l'hypothèse HD1 revient à dire que <u>ce n'est pas la qualité objective</u> de la réponse qui détermine le recours au dictionnaire.

HD2. Le recours au dictionnaire augmente le taux de réponses correctes au post-test.

Le tableau 7.52 fournit des données relatives à cette hypothèse.

Tableau 7.52

	Non recours	Recours	
.E (post-test)	865	152	1017
.R (post-test)	407	742	1149
Total (post-test)	1272	894	2176
% .R	31,99%	82,99%	

χ^2= 550,83 (différence très significative)

L'hypothèse HD2 est pleinement confirmée.

HD3. Plus la certitude au prétest est forte, plus le recours au dictionnaire est faible.

Le tableau 7.53 fournit des données relatives à cette hypothèse.

Tableau 7.53

Om. :	79/167 = 47,31 % d'utilisation	
C1 :	385/857 = 44,93 % d'utilisation	Moyenne = 41,08 %
C2 :	285/716 = 39,81 % d'utilisation	
C3 :	144/435 = 33,11 % d'utilisation	

L'hypothèse HD3 est entièrement confirmée.

HD4. Quand le dictionnaire n'a pas été utilisé, plus la certitude au prêtest est forte, moins la réponse change entre le pré- et le post-test.

Le tableau 7.54 fournit des données relatives à cette hypothèse.

Tableau 7.54

Om. :	(sur 88)	60,23 %	de changements de réponse
C1 :	(sur 472)	58,06 %	de changements de réponse
C2 :	(sur 431)	43,86 %	de changements de réponse
C3 :	(sur 291)	36,77 %	de changements de réponse

L'hypothèse HD4 est entièrement confirmée.

HD5. Quand le dictionnaire a été utilisé, plus la certitude au prêtest est forte, moins la réponse change entre le pré- et le post-test.

Le tableau 7.55 fournit des données relatives à cette hypothèse.

Tableau 7.55

Om. :	(sur 79)	100 %	de changements de réponse
C1 :	(sur 385)	71,94 %	de changements de réponse
C2 :	(sur 285)	65,61 %	de changements de réponse
C3 :	(sur 144)	51,38 %	de changements de réponse

L'hypothèse HD5 est entièrement confirmée.

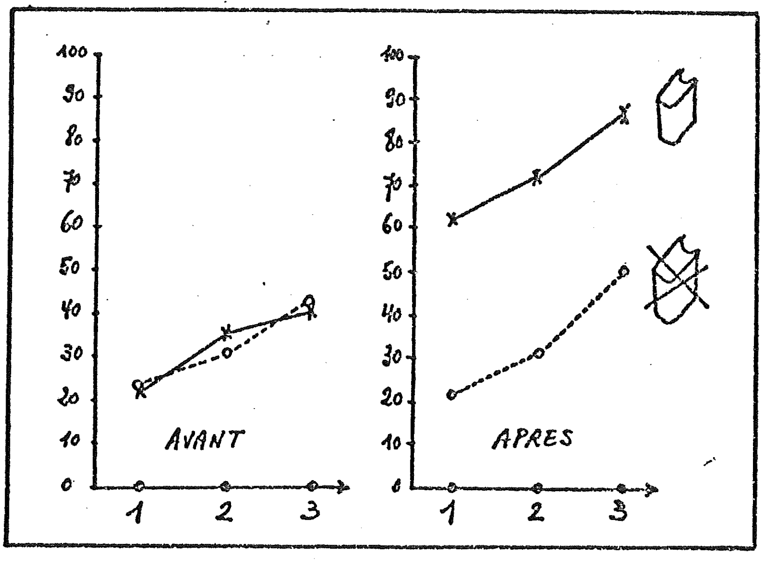
HD6. La certitude augmente avec le pourcentage de réponses correctes (corrélation positive et élevée entre l'ESPER et la moyenne de réponses exactes (μ).

Tableau 7.56

		non utilisation du dictionnaire	utilisation du dictionnaire
HD6 a			
avant	μ 3	131/291 = 45,01	62/144 = 43,05
	μ 2	127/431 = 29,46	97/285 = 34,03
	μ 1	125/472 = 26,48	101/385 = 26,23
HD6 b			
après	μ 3	250/497 = 50,30	639/733 = 87,17
	μ 2	107/355 = 30,14	78/109 = 71,55
	μ 1	60/284 = 21,12	25/40 = 62,50

Le tableau 7.57 présente sous forme graphique les valeurs du tableau 7.56.

Tableau 7.57



L'hypothèse HD6 est entièrement confirmée. Le décalage vers le haut des μ après utilisation rappelle le phénomène déjà observé aux chapitres 2 et 4 : contrairement à l'hypothèse H8, la moyenne de réponses exactes (μ) d'un degré donné varie avec la facilité de l'épreuve.

HD7. Les certitudes fortes au post-test sont plus nombreuses quand le dictionnaire a été utilisé.

Le tableau 7.58 fournit les données relatives à cette hypothèse.

Tableau 7.58

	OM	C1	C2	C3	Cert. fortes (C2 + C3)
Non utilisation	11,38	22,15	27,69	38,76	66,45
Utilisation	2,34	4,47	12,19	81,99	94,18

L'hypothèse HD7 est donc confirmée.

*

* *

La confirmation de six de ces hypothèses élémentaires permet de considérer les résultats comme très logiques et d'en approfondir l'analyse, par des hypothèses plus particulières.

QUAND LE DICTIONNAIRE N'EST PAS UTILISE (matrices H' et H").

HD8. Quand la réponse ne varie pas entre le prétest et le post-test, la certitude elle aussi reste identique (problème de la stabilité de la certitude.

HD8 a. Quand les réponses sont erronées.

Le tableau 7.59 fournit les données relatives à cette hypothèse.

Tableau 7.59

		Prétest - 3	- 2	- 1	
Post-test	- 3	65	50	21	136
	- 2	20	72	45	137
	- 1	7	30	70	107
		92	152	136	380

25% de relation seulement.

L'hypothèse HD8 a se vérifie assez bien : la diagonale contient les plus grandes répétitions. La corrélation vaut .510. On observe une augmentation des certitudes fortes. Ce phénomène pourrait s'expliquer comme suit : le seul réexamen de la question a confirmé à l'élève sa réponse (fausse) antérieure). L'utilisation du dictionnaire pour les autres questions peut apporter une impression de certitude qui s'étend même aux questions pour lesquelles il n'y a pas eu recours au dictionnaire.

HD8 b. Quand les réponses sont correctes.

Le tableau 7.60 fournit les données relatives à cette hypothèse.

Tableau 7.60

		Prétest + 1	+ 2	+ 3	
Post-test	+ 1	19	10	3	32
	+ 2	17	33	13	63
	+ 3	26	47	76	149
		62	90	92	244

La corrélation vaut seulement .225. L'hypothèse HD8b est donc moins bien vérifiée.

La diagonale présente encore des répétitions importantes, mais c'est la certitude 3 au post-test qui est le comportement le plus fréquent. Cette observation pourrait s'expliquer par les deux raisons déjà évoquées ci-dessus et par les trois autres raisons que voici :

1° Lorsqu'il aborde le post-test, l'élève a déjà parcouru toutes les questions une fois au moins (lors du prétest), ce qui lui donne plus d'assurance et augmente le temps de réflexion total.

2° L'utilisation du dictionnaire pour certaines questions peut apporter des informations relatives à d'autres questions.

3° Malgré la consigne et la surveillance stricte, des communications entre élèves ont pu se produire.

QUAND LE DICTIONNAIRE EST UTILISE (matrices I' et I").

HD9. Quand la réponse est inchangée (matrice I'), c'est la certitude 3 qui est modale au post-test.

HD9 a . Quand les réponses sont correctes.

Le tableau 7.61 fournit les données relatives à cette hypothèse.

Tableau 7.61

	+ 1	+ 2	+ 3	
+ 1	4	2	0	6
+ 2	6	19	6	31
+ 3	78	66	50	194
	88	87	56	231

L'hypothèse HD9 a se vérifie totalement.

HD9 b . Quand les réponses sont incorrectes.

Le tableau 7.62 fournit les données relatives à cette hypothèse.

Tableau 7.62

		Prétest			
		- 3	-2	- 1	
Post-test	-3	10	14	8	32
	-2	4	0	9	17
	-1	0	3	3	6
		14	21	20	55

L'hypothèse HD9 b se vérifie, excepté pour la certitude 1 en prétest, mais à une seule unité près (8 contre 9). Cette hypothèse se vérifie beaucoup moins bien que la précédente, ce qui est tout à fait normal : le dictionnaire ne peut confirmer des réponses incorrectes de la même façon que les réponses correctes.

HD10. Quand la réponse est modifiée (matrice I"), la performance est nettement meilleure au post-test.

Sur 608 observations, on observe : 67 (soit 11,01 %) EE
511 (soit 84,04 %) ER
30 (soit 4,92 %) RE.

Des analyses qui précèdent, il ressort que l'ESPER a, aux endroits de décision une influence sur les comportements : recours ou non au dictionnaire, modification ou non de la réponse. Elle fonctionne comme un relais amplificateur ou réducteur selon les cas. Ainsi, après recours au dictionnaire et réponse correcte au prétest, la réponse reste inchangée (i) dans 87,12 % pour la certitude 1,
89,69 % pour la certitude 2,
90,32 % pour la certitude 3.

Les différences sont favorables, mais faibles et peu significatives.

De même, après recours au dictionnaire et réponse correcte au prétest, la réponse change en mal (e) dans 12,88 % pour la certitude 1,
11,49 % pour la certitude 2,
9,68 % pour la certitude 3.

De nouveau, les différences sont faibles et peu significatifes. Si elles se confirmaient, cela signifierait qu la certitude intervient pour freiner le changement selon un mécanisme vraisemblablement proche du théorème de Bayes.

Mais l'ESPER ne présente pas que des aspects positifs. Ainsi, après recours au dictionnaire et réponse incorrecte au prétest, la réponse change en bien (h) dans 94,33 % pour l'omission,
81,69 % pour la certitude 1,
77,12 % pour la certitude 2,
71,95 % pour la certitude 3.

L'ESPER constitue ici un frein plus puissant, fonctionnant toujours selon le principe du théorème de Bayes.

CONCLUSIONS SUR L'EXPERIENCE "DICTIONNAIRE"

A. *CONCLUSIONS SUR LES HYPOTHESES.*

Les constatations ci-dessus, effectuées à partir d'effectifs fort restreints, soutiennent les hypothèses suivantes :

1. Ce n'est pas la qualité objective de la réponse (correcte ou incorrecte) qui explique le recours au dictionnaire.
2. La qualité subjective (ESPER) explique, en partie, le recours au dictionnaire.
3. Le recours au dictionnaire augmente la qualité objective de la réponse
4. Une certitude forte au prétest freine la modification de la réponse, que le dictionnaire ait été ou non utilisé.
5. L'utilisation du dictionnaire augmente la certitude dans les réponses au post-test.
6. Quand le dictionnaire n'est pas utilisé et si la réponse reste identique au post-test, la certitude reste identique - en général - en cas de réponse incorrecte, mais augmente nettement - en général - en cas de réponse correcte.

B. *CONCLUSIONS SUR LA METHODE.*

Avec des sujets jeunes, inexpérimentés, placés dans une situation non operante, on peut observer une expression valide des degrés de certitude. Ces certitudes prennent place à divers endroits dans des schémas explicatifs des mécanismes de recherche et de traitement des informations. Quel serait l'influence d'un entraînement au recours au dictionnaire ? Dans quelle mesure la modification des items même du dictionnaire au niveau de la lisibilité, de l'illustration, entraînerait-elle des ESPER *a posteriori* plus pertinentes ? Quel rôle jouent l'âge, le type de document de référence (atlas de géographie, manuel d'entretien d'une voiture, modes d'emploi divers), etc.? Autant de questions auxquelles pourraient s'attacher des recherches ultérieures avec le recours éventuel à la "réflexion parlée" directement enregistrée, l'enregistrement des mouvements oculaires, l'enregistrement des comportements au magnétoscope...

CONCLUSIONS GENERALES

Vous opposez doute à connaissance.
Peut être ?

A. RESUME

L'étude de l'Evaluation Subjective de la Probabilité d'Exactitude de la Réponse (ESPER) soulève, en sciences de l'Education, des problèmes méthodologiques semblables à ceux qui se sont posés en psychologie lors du développement de la psychophysique. Cette dernière a, entre autres, imposé des définitions strictes aux concepts cruciaux de "stimulus" et de "réponse". De même, à la lumière du concept d'ESPER, les notions de "connaissance", "doute", "information" prennent une signification opérationnelle.

Dans notre thèse, nous soutenions que l'ESPER est propre à rendre compte des états de connaissance partielle et qu'il est possible d'en recueillir, par des méthodes appropriées, des mesures valides, fidèles et sensibles.

Après avoir distingué la probabilité d'utilisation de la réponse (P.U.R.) de la probabilité d'exactitude de la réponse (P.E.R.), nous avons étudié la validité de l'ESPER au moyen de deux expériences.

Dans la première expérience (1971-72), on a recueilli les réponses et les indices de certitudes que 53 sujets (âge modal : 17 ans) ont fourni lors de 14 tests à choix multiple portant sur une même matière (la mécanique) et répartis sur une année scolaire. La matrice des conséquences - pour trois degrés de certitude et pour l'omission de réponse - a été déterminée pour faciliter la mémorisation et les calculs, comme la plupart des matrices utilisées jusqu'ici. Lors de cette expérience, on n'a pas tenu compte du critère de l'Espérance Subjective de l'Utilité, dont nous ne soupçonnions pas l'importance. Les étudiants ont été exposés, selon une procédure opérante, aux conséquences favorables et défavorables de leurs actes.

La seconde expérience (1972-1973) a porté sur 73 sujets dont on a recueilli les réponses et les indices de certitude pour 13 tests de mécanique. Elle est en tous points semblable à la précédente, à l'exception de la matrice de conséquences, cette fois-ci conforme au critère de l'E.S.U. (théorie des décisions).

Les données de ces deux expériences ne peuvent constituer des démonstrations, car plusieurs hypothèses ont été établies *a posteriori*. Néanmoins, il apparaît nettement que des "mesures acceptables de l'ESPER" (selon l'expression bien connue de SHUFORD, ALBERT et MASSENGILL, 1966) requièrent les conditions optimales suivantes :

1.- La procédure de testing doit être operante : les sujets doivent être exposés aux conséquences de leurs actes (des points ou toute autre valeur jugée utile par les sujets doivent dépendre des réponses fournies).

2.- Les sujets doivent avoir été entraînés à estimer leur certitude (par exemple, en étant exposés de façon répétée au programme de renforcement).

3.- Les conséquences doivent être calculées selon la théorie des décisions afin que la stratégie optimale consiste à exprimer le plus exactement possible l'ESPER.

4.- La consigne doit préciser que les indices de certitude utilisés par les sujets correspondent à une échelle des probabilités d'exactitude (PER).

La vérification expérimentale de la validité de l'ESPER a pu être menée à l'aide de trois grands critères :

1. Les cohérences en taux d'utilisation, ou cohérences en μ (trois hypothèses).
2. Les cohérences en taux d'exactitude, ou cohérences en τ (cinq hypothèses).

Il y a interaction
taux d'utili cont. fortes x familité

3. L'incompatibilité des *patterns* observés avec des stratégies contraires au principe de l'E.S.U.

   3a. Les cinq stratégies décrites par la théorie des décisions en sciences économiques.

   3b. Les stratégies décrites par des théories psychologiques : préférences de probabilités, préférences d'ampleur de risque (théorie du dépliage de Coombs).

Face à chacun de ces critères, les données ont mis en évidence la validité de l'ESPER, après une période d'exposition au programme de renforcement.

Dans les conditions particulières des expériences décrites ci-dessus (âge: 17 ans, tests de 15 en 15 jours, résultats fournis dans un délai de 2 jours), trois tests successifs ont été nécessaires pour éliminer, sous l'effet de la procédure operante, la très grande majorité des stratégies incompatibles avec la validité de l'ESPER.

Un seul critère de validité (hypothèse H8) n'est pas respecté par les données : le taux d'exactitude des divers degrés de certitude est (fortement) corrélé avec la facilité de l'épreuve. Cette constatation n'empêche en rien l'utilisation d'indices de certitude, mais impose un approfondissement du problème pour développer des indices de correction. Une augmentation du nombre de degrés de certitude disponibles supprimerait peut-être ce phénomène. Nous avons entrepris des recherches dans ce sens.

La fidélité (stabilité) des certitudes peut être mesurée dans le cadre d'expériences de test-retest, à partir de "jeux des ESPER", inspirés du *SHANNON guessing game*. Ces procédures expérimentales présentent d'incontestables avantages d'utilisation, mais aussi des défauts tels que l'on peut raisonnablement penser que, sans eux, la fidélité observée serait plus élevée.

Les trois types de comportements indésirables (choix aléatoire du degré de certitude, modification des hypothèses lors du retest, modification de la stratégie) peuvent être détectés par trois techniques : l'écart entre cotes observées et cotes théoriques, la comparaison des histogrammes d'utilisation des certitudes lors du test et du retest (par la formule de Kolmogorov-Smirnov), la comparaison des paramètres (valeur centrale, dispersion, *skewness*, *curtosis*) des distributions ci-dessus.

Une expérience test-retest a été menée, au moyen d'un jeu des ESPER, sur cent réponses, auprès d'adultes. On a pu recueillir des résultats complets pour 124 sujets. Les réponses de soixante pour cent d'entre eux présentent des caractéristiques indésirables évoquées ci-dessus. Malgré ces inconvéients méthodologiques, la stabilité moyenne des ESPER vaut .56 (corrélation entre le test et le retest), alors que ces adultes n'avaient pas été entraînés et n'étaient pas dans une situation operante. Dans les circonstances expérimentales les plus favorables, la fidélité atteint .78.

30%

B

La <u>sensibilité</u> de l'ESPER peut être étudiée à partir de trois critères :

1. Les taux d'utilisation spontanée des divers degrés de certitude disponibles;
2. Les mode, *skewness* et *curtosis* des histogrammes de replication.
3. La fonction de validité (μ/ESPER).

Les adultes testés recourent le plus fréquemment à onze degrés différents (critère 1), alors que quarante sont permis. Toutefois, sept ou huit degrés seulement se distinguent de façon nette les uns des autres (critères 2 et 3). Ainsi se confirment les observations faites par G.A. MILLER (1956) dans d'autres domaines de la perception (*The magical number seven*).

On a qd m l'impression que l'élève a le sentiment de "savoir" ou "ne pas savoir"

→ //

Les résultats permettent, en plus, de formuler l'hypothèse que certaines zones de l'échelle des PER (les extrêmes notamment) offrent une sensibilité plus grande qu'ailleurs.

. . .

Les expériences concernant la révision des hypothèses et le recours à des ouvrages de référence montrent la part importante que joue la connaissance partielle dans le traitement de l'information (problème de la cohérence) et dans la recherche de l'information (problème de la conséquence). L'information peut être simplement reçue par le sujet ou, au contraire, faire l'objet d'une recherche délibérée.

Dans la mise en oeuvre de comportements de recherche, l'ESPER *a priori* joue un rôle : moins on est sûr, plus on devrait chercher (la corrélation n'est évidemment pas parfaite). Une fois la décision prise, l'ESPER de départ continue à jouer, conformément au théorème de Bayes.

Rappelons que, par son caractère normatif, le théorème est un guide pour l'action et devrait peut-être faire l'objet d'un apprentissage à l'école.

B. PERSPECTIVES

A la fin de ce travail, il est raisonnable de considérer que les ESPER sont, dans une très large mesure, valides, fidèles et sensibles. Ils pourraient l'être plus encore avec des sujets entraînés.

Sur cette base, on pourrait donner à l'ESPER un rôle plus important dans l'évaluation sommative comme dans l'évaluation formative en éducation, d'autant plus que la certitude éprouvée conditionne la recherche d'information et constitue un "rouage" essentiel dans les processus de traitement de l'information. Elle devrait dès lors prendre une place importante à l'école et dans les laboratoires de recherche. Permettant notamment l'utilisation du théorème de Bayes en pédagogie (et pas seulement en statistique), l'ESPER donne un intérêt nouveau à diverses recherches.

Les valeurs numériques de la présente étude donnent des ordres de grandeur, des tendances, mais ne prétendent pas à une précision absolue, pour diverses raisons. D'abord, aucune des expériences n'a porté sur des échantillons représentatifs, notre recherche n'étant qu'exploratoire. En outre, les résultats obtenus sont très sensibles aux conditions expérimentales. Enfin, diverses hypothèses ont été formulées *a posteriori*. En toute rigueur, il faudrait répéter les expériences avec d'autres sujets. Cependant, dans l'immédiat, il nous paraît plus urgent de mener de nouvelles recherches exploratoires telles que celles qui sont suggérées ci-après.

corrél. ?

A. Etudes portant sur la certitude elle-même.

1. *RECHERCHE DE PARAMETRES.*

Il importerait de replacer les diverses hypothèses relatives à la validité (cf. chapitres 2 et 4) dans une axiomatique globale et de veiller à l'indépendance entre les diverses hypothèses.

En ce qui concerne la fidélité, il est peut-être possible de concevoir d'autres indices que la valeur centrale, la dispersion, la *skewness* et la *curtosis* des histogrammes de replication, ces indices pouvant eux-mêmes faire l'objet d'un approfondissement systématique.

Quant à la sensibilité, elle pourrait faire l'objet d'approches plus directes (1) que celles qui ont été décrites ci-dessus. Des indices nouveaux pourraient dès lors être envisagés.

2. *L'EFFET DE LA SITUATION.*

Dans ces études, les quatre grands principes méthodologiques (aspect operant, entraînement, matrice de conséquences, consigne) seraient suivis. Les développements de la technologie permettent cependant d'envisager des améliorations méthodologiques d'importance (2).

On construirait diverses procédures dont on mesurerait l'influence sur les paramètres. On pourrait ainsi faire apparaître dans quel type de situation la fidélité est maximale et quantifier la distance qui sépare une réalité scolaire donnée de cette situation "de référence".

---

(1) On pourrait, par exemple, demander aux sujets de répartir des étoiles (cf. le Five Stars System de DE FINETTI) sur l'échelle des probabilités. Ex/: 0 10 20 30 40 50 60 70 80 90 100
Une variante plus commode pour l'étudiant comme pour le chercheur (calculs ultérieurs) consisterait à imposer un choix entre diverses courbes en cloche différant par leur curtosis. Le sujet devrait placer l'une d'elles (plus exactement son axe de symétrie, ou moyenne) sur le point désiré de l'axe.

(2) Ainsi, le "jeu des ESPER" (deviner les lettres d'un texte mutilé, ou une variante plus proche encore de la situation scolaire) pourrait être présenté au sujet à l'aide d'un terminal d'ordinateur. Les conséquences seraient ainsi communiquées immédiatement après chaque réponse, ce qui, logiquement, devrait accélérer l'élimination des stratégies indésirables, processus qui prend trois séances successives lorsque les données sont traitées par lots (batch processing) par l'ordinateur. L'affichage, non seulement des gains et des pertes ponctuels, mais du total en évolution constante, permet de rendre plus perceptible la procédure operante.

L'importance relative du degré de certitude dans la matrice de conséquences peut faire l'objet d'enquêtes auprès d'enseignants, d'étudiants,... On peut aussi observer l'effet d'importances différentes sur le temps de réponse, le choix des degrés, l'anxiété résultante, la vitesse d'apprentissage, etc.

Enfin, l'influence des questions (contenus, niveaux taxonomiques) devrait elle aussi être prise en considération.

3. *LES DIFFERENCES ENTRE INDIVIDUS.*

A ces questions liées à la procédure s'ajoutent des hypothèses relatives aux variables individuelles : l'âge, le niveau scolaire, et intellectuel, la fatigue, la personnalité (anxiété, besoin d'accomplissement, personnalité à tendance obsessionnelle ou hystérique), la compétence, l'entraînement, etc.

Certaines de ces caractéristiques sont-elles liées à des aspects (validité, fidélité, sensibilité,...) de l'ESPER ? Sur ces préoccupations se grefferaient des expériences concernant la modification des paramètres individuels. La procédure operante qui, dans nos expériences, s'est montrée si efficace en ce qui concerne la validité globale des ESPER, a-t-elle un impact aussi direct sur la fidélité et la sensibilité ? Des critères plus précis de la validité (ex. : hypothèse H8) peuvent-ils être atteints, et à quel prix ?

4. *ETUDE D'AUTRES VARIABLES.*

Par chronométrie automatique (1), on pourra distinguer le temps de réaction pour la réponse, puis le temps qui sépare l'émission de la réponse de l'émission de la certitude. *A priori*, cependant, le temps total de réaction (réponse + ESPER) devrait être l'indice le plus significatif. Nos tentatives pour réaliser ces mesures à l'aide d'un chronomètre manuel (ou avec un compteur à impression) ont, jusqu'à présent, échoué. Le temps de réaction augmente-t-il avec le nombre de degrés de certitude disponibles, varie-t-il avec le type d'échelles (probabilités, rapports), selon que la certitude est faible ou élevée ?

(1) Par terminal d'ordinateur.

Ces questions, et bien d'autres, offrent un intérêt pratique et théorique, en rapport avec la théorie du dépliage (cf. GREENBERG, 1963), la révision des probabilités (cf. GELLER, WHITMAN et POST, 1973), la psychophysique (cf. BUJAS, KOVACIC et ROHACEK, 1975).

Les recherches sur la cohérence devraient s'étendre à des études sur la conséquence, c'est-à-dire le déclenchement (ou non) d'une action R2 consécutive à l'ESPER relative à une réponse R1.

Il conviendrait d'étudier d'abord la "conséquence" dans des situations où une seule réponse R2 est disponible (par exemple, le dictionnaire que le sujet choisit d'utiliser ou non).

Trois étapes permettraient de recueillir trois types de données complémentaires :

Etape 1. On observe l'UR2 (utilisation de la réponse R2 spontanée : par exemple, le recours spontané au dictionnaire), en rapport avec l'ER1 (exactitude de la réponse R1) et l'ESPER 1. On prendra soin, dans les consignes, de ne pas attirer l'attention sur la réponse R2, tout en rendant celle-ci très aisée (par exemple, chaque sujet disposera d'un dictionnaire individuel).

Etape 2. On recueille les ESPUR2 (évaluation subjective de la probabilité d'utilisation de la réponse R2) en demandant au sujet d'imaginer des situations où ses ESPER1 prendraient des valeurs différentes.

Etape 3. On observe l'UR2 (comme dans l'étape 1) en rapport avec l'ER1 et l'ESPER 1, en attirant l'attention du sujet sur la réponse R2 (l'application de l'étape 2 suffit à attirer l'attention).

On dispose ainsi de données de base (UR2 et ESPUR 2) semblables à celles qui sont utilisées dans l'étude de l'ESPER, avec, néanmoins, deuxmesures (étapes 1 et 3) d'UR2 là où une seule mesure de l'ER1 s'imposait. Le problème est en effet plus complexe avec l'UR puisque le sujet peut l'influencer après avoir exprimé son ESPUR, ce qui n'est pas le cas avec l'ER.

L'étape 2 ci-dessus présente des liens avec la mesure des attitudes traduites en intentions comportementales.

B. Etudes utilisant la certitude.

1. *QUANTIFICATION DE L'INFORMATION TRANSMISE.*

L'application du théorème de Bayes pourrait permettre de quantifier la valeur informative d'unités d'enseignement (texte, séquence TV, diapositive,...). Van Naerssen (1965, p. 164) fait la suggestion suivante : il fixe x (les points obtenus en cas de réponse correcte) et y (les points obtenus en cas de réponse incorrecte) en fonction de p (la PER) et q (1-p) comme suit :

$$x = C + D \log p$$

$$y = C + D \log q$$

où C et D sont des constantes arbitraires.
Une échelle qui utilise cette formule présente une particularité intéressante : la valeur attendue (VA).

$$VA = C - D\,(-p \log p - q \log q)$$

Dans l'expression entre parenthèses, on reconnaît l'incertitude H de la théorie de l'information. Si, par exemple, on construit un test à alternatives avec une échelle logarithmique , si C = D = 1, et si la base des logarithmes est 2, le score moyen du test est égal à l'information moyenne (en bits) fournie par le sujet.

Le problème mérite évidemment d'être approfondi, en le reformulant à partir des ESPER.

A la suite de nos expériences, on pourrait comparer la répartition des ESPER dans les diverses solutions possibles à une question avant l'événement X puis après l'événement X. La comparaison des entropies des deux répartitions n'est peut-être pas la meilleure option. Des indices basés sur les différences d entre les ESPER d'une même solution avant et après X seraient peut-être préférables ($\sum d$, $\sum d2$, $\sum |d|$, $\sum d/N$, etc.). Quel que soit l'indice pertinent, il nous paraît que le problème est d'importance et que sa solution est à notre portée. A nouveau, la technologie pourrait être un précieux adjuvant (1).

(1) Certains terminaux graphiques interactifs permettent, en outre, d'exprimer la certitude sous la forme d'un segment de droite dont la longueur est proportionnelle à la probabilité exprimée (comme dans l'expérience de J. BAKER, 1969). Ainsi, si n solutions sont proposées, par curseur extérieur ou par crayon lumineux stimulant directement le tube cathodique, on peut "tracer" le segment de droite. L'intérêt primordial d'un tel système réside dans le fait que l'on peut programmer la machine pour que la somme de n segments fasse 1 : toute augmentation de l'un des segments entraîne automatiquement la diminution (proportionnelle ou non à chaque longueur) des autres segments.

2. *LA PSYCHOMETRIE*.

Dans l'introduction (p. 33 à 35), nous signalions que l'ESPER pouvait jouer un rôle positif dans l'estimation du "vrai" score d'un sujet à une question. Malheureusement, et on s'en doutait, la variable "certitude", destinée à réduire l'erreur de mesure, est elle-même entachée d'une erreur de mesure.

Parallèlement aux études expérimentales qui cerneraient de plus près les fluctuations de ces erreurs de mesure selon les situations et les individus, une approche théorique s'impose. Il nous paraît que cette approche doit être menée dans le cadre général de la théorie de la généralisabilité (CRONBACH *et al.*, 1972; CRONBACH, 1975; CARDINET et TOURNEUR, 1975).

Une question s'impose particulièrement à notre attention : quelle fidélité devrait atteindre l'ESPER, dans des conditions données, pour permettre de réduire l'erreur de mesure au niveau d'un individu, d'une question, d'un test, d'une série de tests, etc. ?

On sait que, depuis quelque temps, une statistique bayésienne s'est développée et un débat est actuellement ouvert à son sujet.

Dans la mesure où elle consiste, par exemple, à estimer le "vrai" score (VS), étant donné le score observé (SO), c'est-à-dire (VS | SO), mise à part l'écriture (1), elle rencontre la même préoccupation que la statistique classique. Le fait de combiner une "probabilité *a priori*" à une "information" n'est pas neuf non plus. Mais utiliser purement et simplement les ESPER du chercheur comme probabilités *a priori* d'une hypothèse pourrait paraître abusif : peut-on donner le même poids aux résultats expérimentaux et aux probabilités qu'un individu attribue à un événement ? Ce problème dépasse notre compétence, mais nous avons voulu le mentionner, car la "statistique bayésienne" est aujourd'hui l'utilisation du théorème de Bayes la plus développée en psychologie et en sciences de l'éducation.

---

(1) Par exemple, JACKSON (1975) note $x|\tau$ le score observé (x), étant donné le vrai score (τ). Il considère ces $x|\tau$ comme des distributions a priori. Il s'efforce d'estimer $\tau|x$ (ou distributions a posteriori).

2. *LA PSYCHOPHYSIQUE*.

En raison même de la similitude de problématique, il devrait être fructueux de présenter les expériences psychophysiques classiques (méthode ascendante, descendante, des stimuli constants) avec indices de certitude. Dans le même ordre d'idées, les recherches classiques de la théorie de l'information (et les formulations mathématiques de ces problèmes) pourraient être envisagées avec utilisation d'indices de certitude. Il en va de même, évidemment, pour les modèles stochastiques d'apprentissage (ROUANET, 1967).

3. *LES ATTITUDES*.

En se limitant aux actions projetées (et aux actions aisément vérifiables), on pourrait recueillir les ESPUR (voir ci-dessus). Ce type d'attitude se définirait alors comme une tendance à l'action (tendance évaluée subjectivement) dans une situation précise (réelle ou imaginaire).

Il serait peut-être possible, comme pour l'ESPER, de développer des critères de cohérence (en τ , en μ). Les opinions pourraient être mises en corrélation avec les actions projetées. Le problème devient plus complexe quand l'UR (utilisation de la réponse) ne dépend pas du seul sujet ou quand la situation donne lieu à des interprétations diverses (par exemple pour des questions où il n'existe pas de solution unanimement admise).

°
° °

Nous voudrions que le terme de ces réflexions constitue le départ d'une série d'autres, en envisageant la relativité de ce qui, dans toutes nos recherches, a été considéré comme absolu. Toujours, en effet, nous avons utilisé des réponses qui peuvent être objectivement classées comme correctes ou incorrectes (ER = 1 ou 0). Le recours exclusif à cette dichotomie peut, lui aussi, être contesté.

De même que les ESPER sont exprimées sous forme de probabilités, on pourrait imaginer que les ER prennent des valeurs comprises entre 0 et 1. Ces valeurs pourraient fluctuer selon les temps, les lieux, les individus (les groupes d'experts, par exemple) et être déterminées grâce, entre autre , à des ESPER.

L'échaffaudage psychométrique serait plus complexe puisqu'il s'agirait d'évaluer les connaissances "floues" de données elles-mêmes "floues". N'est-ce pas dans ce cadre très général que devrait être posé tout problème de mesure en éducation ?

Disposant d'une ESPER pour une réponse dont l'ER est relative, un individu est amené à utiliser ou non cette réponse (UR = 1 ou 0) et, parfois, à modifier ER (définie, rappelons-le, à un moment donné, en un lieu donné, pour un groupe d'individus donné, etc.). Ainsi, un individu peut finir par imposer une idée alors que ses contemporains y accordent une ER faible, et *vice versa*.

C'est peut-être au prix du développement d'un réseau de relations entre les ESPER, les ER, les ESPUR et les UR que pourront être pris en compte les problèmes les plus déterminants de la réflexion et de l'action humaines.

B I B L I O G R A P H I E

ADAMS, G.S. (1964)
Measurement and Evaluation in Education, Psychology and Guidance. New York, Holt, Rinehart and Winston.

AHLGREN, A. (1967)
Confidence on Achievement Tests and the Prediction of Retention. Harvard University, thèse de doctorat inédite.

AIKEN, L.R. Jr. (1966)
Another Look at Weighting Test Items, *Journal of Educational Measurement, 3,* 2.

ALLAIS, M. (1953)
Le comportement de l'homme rationnel devant le risque : critique des postulats et axiomes de l'école américaine, *Econometrika, 21,* 503-546. (A)

ALLAIS, M. (1953)
La psychologie de l'homme rationnel devant le risque : la théorie et l'expérience, *Journ. Soc. Statist., 94,* 43-73. (B)

ASRATIAN et SIMONOV (s.d.)
Fiabilité du cerveau, Editions de Moscou.

ATKINSON, J.W. et LITWIN, G.H. (1960)
Achievement Motive and Test Anxiety Conceived as Motive to Approach Success and Motive to Avoid Failure, *Journ. Abnorm. Soc. Psychol., 60,* 52-63.

ATKINSON, J.W. (1964)
An Introduction to Motivation, Princeton, Van Nostrand.

ATTNEAVE, F. et ARNOULD, M.D. (1956)
The Quantitative Study of Shape and Pattern Perception, *Psychol. Bull., 53,* 452-471.

ATTNEAVE, F. (1959)
Application of Information Theory to Psychology, New York, Holt, Rinehart and Winston.

BAKER, J.D. (1969)
The Uncertain Student and the Understanding Computer, *La recherche en enseignement programmé, Tendances actuelles.* Paris, Dunod, 303-319.

BAYES, T. (1763)
An Essay Toward Solving a Problem in the Doctrine of Chance, *Philosophical Transactions of the Royal Society*, Londres.

BEAUJOT, A., DIDELEZ, M., FONTAINE, O. et LECLERCQ, D. (1966)
Etude d'une nouvelle technique d'évitement sans signal avertisseur chez le rat, *Psychologica Belgica, VII.*

BEENEN, F. (1970)
Psychiatric Diagnosis and Subjective Probabilities, *Acta Psychologica, 34,*

BERNOUILLI, D. (1954)
Exposition of a new theory on the measurement of risk (English translation of "Specimen theorial novea de mensura sortis", Commentarii academiae scientiarum imperialis Petropolitanae, 1. O et 1973, 5, pp. 175-192), by Louise SOMMER), *Economica,* 22 23-26.

BLALOCK, H.M. (Ed.) (1971)
Causal Models in the Social Sciences, New York, MacMillan.

BOCK, R.D. (1960)
Methods and Applications of Optimal Scaling, *Psychometric Laboratory, Report 25,* Chapel Hill, N.C.

BOCK, R.D., JONES, L.V., (1968)
The Measurement and Prediction of Judgement and Choice, San Francisco, Holden Day.

BOCK, R.D. et WOOD, R. (1971)
Test Theory, *Annual Review of Psychology, 22,* 193-224.

BOCK, R.D. (1975)
Basic Issues in the Measurement of Change, *Second International Symposium on Educ. Testing,* Montreux.

BORMUTH, J.R. (1970)
On the Theory of Achievement Test Items, Chicago, The University of Chicago Press.

BRESSON, F. (1965)
Les décisions, in FRAISSE et PIAGET, *Traité de Psychologie expérimentale, Vol. VIII,* Paris, P.U.F.,

BROWNLESS, V.T. et KEATS, J.A. (1958)
A retest Method of Studying Partial Knowledge, *Psychometrika, 23,* 1, 67-73.

BUJAS, Z., KOVACIC, M., et ROHACEK, A. (1975)
Psychophysical function based on confidence rating, in *Acta Instituti Psychologici Universitatis Zagrabiensis,* 74.

CARDINET, J. et TOURNEUR, Y. (1975)
The Generalization of Surveys of Educational Outcomes, *Second Intern. Symp. on Educ. Testing,* Montreux.

CATTELL, R.B. (1964)
Validity and Reliability, A Proposed More Basic Sets of Concepts, *Journal of Educ. Psychology, 55,* 1-22.

CHAUVIN, R. (1967)
Paradoxe dans les résultats du conditionnement, *Journ. Psych. Normale et Pathol.*, 129-141

CHERNOFF, H. (1954)
Rational Selection of Decision Functions, *Econometrica*, 422-443.

CHERNOFF et MOSES (1954)
Elementary Decision Theory, New York, Wiley.

CHERNOFF, H. (1962)
The Scoring of Multiple Choice Questionnaires, *Annals of Mathematical Statistics*, *33*, 375-393.

CHERRY, E.C. (1957)
On the Validity of Applying Communication Theory to Experimental Psychology, *Brit. Journ. Psychol.*, *48*, 176-188.

CHOPPIN, B. (1970)
An IEA Study of Guessing. A Proposal, Stockholm, International Association for the Evaluation of Educational Achievement. Unpublished Memorandum, IEA/TR/9.

CHOPPIN, B. (1974)
The Correction for Guessing on Objective Tests, Stockholm, International Association for the Evaluation of Educational Achievement.

CHOPPIN, B. (1975)
Guessing the Answer on Objective Tests, *Brit. Journ. Educ. Psych.*, *45*, 206-213.

CHOPPIN, B. (1975)
Recent Developments in Item Banking. A Review, Montreux, Second Intern. Symposium on Educational Testing.

CLARK, R.A., TEEVAN, R., RICCIUTI, H.N. (1956)
Hope of Success and Fear of Failure as Aspects of Need for Achievement, *Journ. Abnorm. Soc. Psych.*, *53*, 182-186.

C.N.R.S. (Colloque du) (1967)
Les modèles et la formalisation du comportement, Paris.

COOMBS, C.H. (1950)
Psychological Scaling Without a Unit of Measurement, *Psychol. Rev.*, *57*, 145-158.

COOMBS, C.H. (1953)
On the Use of Objective Examinations, *Educ. and Psychol. Measurement*, *13*, 308-130.

COOMBS, C.H., MILHOLLAND, J.E. et WOMER, F.B. (1956)
The Assessment of Partial Knowledge, *Educ. and Psychol. Measurement, 16,* 13-37.

COOMBS, C.H. et PRUITT (1960)
Components of Risk in Decision Making: Probability and Variance Preferences, *Journ. Exp. Psych.,* 265-277.

COOMBS, C.H. (1960)
A Theory of Data, *Psych. Review, 67,* 143-159.

COOMBS, C.H., GREENBERG, M.G. et ZINNES, J.A. (1961)
A Double Law of Comparative Judgment for the Analysis of Preferential Choice and Similarities Data, *Psychometrika, 26,* 165-171.

COOMBS, C.H. (1964)
A Theory of Data, New York, Wiley.

COOMBS, C.H., DAWES, R.M., TVERSKY, A. (1970)
Mathematical Psychology, Englewood Cliffs (N.J.), Prentice Hall.

COOMBS, C.H., et BOWEN, J.N. (1971)
A Test of VE Theories of Risk and the Effect of the Central Limit Theorem, *Acta Psychologica, 35,* 15-28.

CRAWFORD, W.R. et LEWY, A. (1965)
A rapid and Efficient Method for Scoring and Analyzing Complex Multiple Choice Examinations, Chicago, National Council on Measurement in Education (document inédit).

CRONBACH, L.J. (1950)
Further Evidence on Response Sets and Test Design, *Psychol. Measurement, 10,* 3-31.

CRONBACH, L.J. et MEEHL, P.E. (1955)
Construct Validity in Psycholocical Tests, *Psychological Bulletin, 52,* 281-302.

CRONBACH, L.J. et GLESER, G.C. (1965)
Psychological Tests and Personnel Decisions, Univ. of Illinois Press, 2e éd.

CRONBACH, L.J. (1960)
Essentials of Psychological Testing, New York, Harper and Row.

CRONBACH, L.J. (1963)
Improvement Through Evaluation, *Teachers College Record, 64,* 672-683.

CRONBACH, L.J., GLESER, G.C., NANDA, H. et RAJARATNAM, N. (1972)
The Dependability of Behavioral Measurements, New York, Wiley.

CRONBACH, L.J. (1975)
Generalizability Theory from the Perspective of Educational Evaluation, Montreux, Second International Symposium on Educational Testing.

CRONHOLM (1963)
A General Method of Obtaining Exact Sampling Probabilities of the Shannon-Wiener Measure of Information H, *Psychometrika, 28*, 4.

CROSSMAN, E.R.F.W. (1955)
The Measurement of Discriminability, *Quarterly Journ. of Exper. Psychol., 7*, 176-195.

DAGNELIE, P. (1970)
Théorie et méthodes statistiques, 2 vol., Gembloux, Duculot.

DAVIS, F.B. (1959)
Use of Correction for Chance Success in Test Scoring, *Journal of Educ. Research, 52*, 179-180.

DAVIS, F.B. et FIFER, G.,(1959)
The Effect on Test Reliability and Validity of Scoring Aptitude and Achievement Tests with Weights for Every Choice, *Educ. and Psychol. Measurement, 19*, 159-170.

DAVIS, F.B. (1964)
Educational Measurements and their Interpretation, Belmont (Calif.) Woodsworth.

DAVIS, F.B. (1966)
Analyse des items, Louvain, Ed. Nauwelaerts,

DAWES, R.M. (1972)
Fundamentals of Attitude Measurement, New York, Wiley.

DE BRUYN (1973)
Méthodes quantitatives de gestion, Presses Universitaires de Liège.

DE FINETTI, B. (1937)
La prévision : ses lois logiques, ses sources subjectives, *Annales de l'Institut Henri Poincaré, 7*.

DE FINETTI, B. (1959)
Dans quel sens la théorie de la décision est-elle et doit-elle être normative ? *in* F.N.R.S. (colloque du), *La décision*, Paris, 159-170.

DE FINETTI, B.,(1962)
Does it make sense to speak of good probability appraisers ? in GOOD, I.J. (Ed.), *The Scientist Speculates*, New York, Basic Books, 357-364.

DE FINETTI, B. (1965)
Methods for Discriminating Levels of Partial Knowledge Concerning a Test Item, *Brit. Journ. of Mathem. and Statist. Psych.*, *18*, 87-123.

DE FINETTI, B. (1963)
La décision et les probabilités, *Revue des Mathématiques pures et appliquées*, Bucarest, 405-413.

DE FINETTI, B. (1970)
Logical Foundations and Measurement of Subjective Probability, *Acta Psychologica*, *34*, 129-145.

D'HAINAUT, L. (1971)
L'enseignement des concepts scientifiques et techniques à l'aide de cours programmés, Univ. Libre de Bruxelles, thèse de doctorat.

D'HAINAUT, L. (1974)
Une méthode de compensation statistique des choix heureux par ignorance dans les questions fermées d'épreuves d'acquisition, *Les sciences de l'éducation*, *7*, 1, 57-83.

DE LANDSHEERE, G. (1970)
Introduction à la recherche en éducation, Paris, A. Colin-Bourrelier; Liège, G. Thone, 3e éd.

DE LANDSHEERE, G. (1971)
Evaluation continue et examens. Précis de docimologie, Paris, F. Nathan; Bruxelles, Labor.

DE LANDSHEERE, G. (1973)
Le test de closure, Mesure de la lisibilité et de la compréhension, Paris, F. Nathan; Bruxelles, Labor.

DIAMOND, J. et EVANS, W. (1973)
The Correction for Guessing, *Rev. of Educ. Research*, *43*, 2.

DRESSEL, P.L. et SCHMID, J. (1953)
Some Modifications of the Multiple-Choice Item, *Educ. and Psych. Measurement*, *13*, 574-595.

EBEL, R.L. (1965)
Measuring Educational Achievement, Englewood Cliffs, N.J., Prentice Hall. A.

EBEL, R.L. (1965)
Confidence-Weighting and Test Reliability, *Journ. of Educ. Measur.*, 2, 49-57 B.

EBEL, R.L. (1968)
Review of Valid Confidence Testing Demonstration Kit, *Journal of Educ. Measurement*, *5*, 353-354.

EBEL, R.L. (1969)
Expected Reliability as a Function of Choices Itr Item, *Educ. and Psych. Measurement, 29,* 565-570.

EBEL, R.L. (1971)
Criterion Referenced Measurements Limitations, *School Review, 79,* 282-288.

EBERT, R..(1971)
Sequential Decision Making: An Aggregate Scheduling Methodology, *Psychometrica, 36,*

ECHTERNACHT, G.J. (1972)
The Use of Confidence Testing in Objective Tests, *Review of Educational Research, 42,* 217-236.

EDGINGTON, E.S. (1965)
Scoring Formulas that Correct for Guessing, *Journal of Exper. Educ., 33,* 345-346.

EDWARDS, A.L. (1960)
Experimental Design in Psychological Research, New York, Holt, Rinehart and Winston,

EDWARDS, A.L. (1967)
Statistical Methods, New York, Holt, Rinehart and Winston, 2e éd.

EDWARDS, E. (1969)
Information Transmission, Londres, Chapman & Hall.

EDWARDS, W. (1953)
Probability Preferences in Gambling, *The American Journal of Psych., 66,* 349-364.

EDWARDS, W.,(1954)
Variance Preferences in Gambling, *Amer. Journ. Psychol., 67,* 441-452.

EDWARDS, W. (1954)
Probability Preferences among Bets with Differing Expected Values, *Amer. Journ. Psychol., 67,* 56-57.

EDWARDS, W. (1954)
The Reliability of Probability Preferences, *Amer. Journ. Psych., 67,* 68-95.

EDWARDS, W. (1954)
The Theory of Decision Making, *Psychol. Bull., 51,* 380-417.

EDWARDS, W. (1954)
Methods for Computing Uncertainties, *Amer. Journ. Psychol., 67,* 164-170.

EDWARDS, W. (1955)
The Prediction of Decisions Among Bets, *J. of Exp. Psych.*, *59*, 201-214.

EDWARDS, W. (1960)
Measurement of Utility and Subjective Probability, in GULLIKSEN, H.,et MESSICKS, Ed., *Psychological Scaling: Theory and Applications*, New York, Wiley.

EDWARDS, W. (1961)
Probability Learning in 1000 Trials, *Journ. of Exp. Psychol.*, *62*, 4, 385-394.

EDWARDS, W. (1961)
Behavioral Decision Theory, *Ann. Rev. Psychol.*, *12*, 473-498.

EDWARDS, W. (1962)
Subjective Probabilities Inferred from Decisions, *Psychol. Rev.*, *69*, 109-135.
Id., in LUCE, BUSH and GALANTER *(vol. 2*, p. 465, 1965).

EDWARDS, W. (1962)
Utility, Subjective Probability, Their Interaction, and Variance Preferences, *J. Conf. Res.*, *6*, 42-51.

EDWARDS, W., LINDMAN, H., PHILLIPS, L.D. (1965)
Emerging Technologies for making decisions. In *New Directions in Psychology*, Vol. II, New York, Holt, Rinehart and Winston, 261-325.

EDWARDS, W. (1967)
Probabilistic Information Processing by Men and Man-Machine Systems, in *La simulation du comportement humain*, Paris, Dunod, p. 187.

EPSTEIN, E.S. (1969)
A Scoring System for Probability Forecasts of Ranked Categories, *J. Appl. Meteorol.*, *8*, 985-987.

ESCALONA, S.K. (1940)
The Effect of Success and Failure upon the Level of Aspiration and Behavior in Manic-Depressive Psychoses, *Univ. of Iowa Stud. Child Welf.*, *16*, 199-302.

FESTINGER, L. (1942)
A Theoretical Interpretation of Shifts in Level of Aspiration, *Psychol. Rev.*, *49*, 235-250.

FISCHER, G.H. (1975)
Some Probabilistic Models for Measuring Change, Montreux, Second International Symposium on Educational Testing.

FISCHER, I. (1906)
The Nature of Capital and Income, New York, Wiley.

FLANAGAN, J.C. (1955)
The Development of an Index of Examinee Motivation, *Educ. Psych. Measur.*, *15*, 144-151.

FOURASTIE, J. (1972)
Les quarante mille heures, Paris, Ed. Denoël.

FREUD, S. (1965)
Totem et Tabou, Paris, Ed. Payot.

GELLER, E.S., WHITMAN, P., et POST, D.S. (1973)
Stimulus probability and prediction outcome as determinants of choice reaction time: some procedural considerations, *Acta Psychologica*, *37*, 1-14.

GEORGE, Ch. (1971)
Choix et apprentissage en situation aléatoire, Paris, C.N.R.S., Monographies Françaises de Psychologie.

GHISELLI, E.E. (1964)
Theory of Psychological Measurement, New York, McGraw-Hill.

GLASS, G.V. et WILEY, D.E. (1964)
Formula Scoring and Test Reliability, *Journal of Educational Measurement*, 43-47.

GOSLIN, D.A. (1966)
The Search for Ability, Standardized Testing in Social Perspective, New York, Wiley.

GREENBERG, M. (1963)
J. Scale Models for Preference Behavior, *Psychometrika*, *28*, 3, 265-271.

GREENE, F.P. (1964)
A Modified Cloze Procedure for Assessing Adult Reading Comprehension, Ann Arbor, Univ. of Michigan, thèse de doctorat.

GUILFORD, J.P. (1954)
Psychometric Methods, New York, McGraw-Hill.

GUILFORD, J.P. (1965)
Fundamental Statistics in Psychology and Education, New York McGraw-Hill.

GUTTMAN, L. et SCHLESINGER, I.M. (1967)
Systematic Construction of Distractors for Ability and Achievement Test Items, *Educ. and Psych. Measurement*, *27*, 159-170.

HAMBLETON, R.K., ROBERTS, D.M. et TRAUB, R.E. (1970)
A Comparison of the Reliability and Validity of Two Methods for Assessing Partial Knowledge on a Multiple-Choice Test, *Journ. of Educ. Measurement*, *7*, 75-82.

HAMILTON, C.H. (1950)
Bias and Error in Multiple Choice Tests, *Psychometrika, 15,* 151-168.

HAMMERTON, M. (1965)
The Guessing Correction in Vocabulary Tests, *Brit. Journ. of Educ. Psych., 35,* 249-251.

HANCOCK, J.G. et TEEVAN, R.C. (1964)
Fear of Failure and Risk-Taking Behavior, *J. Pers., 32,* 200-209.

HENMON, V.A.C. (1911)
The Relation of the Time of a Judgment to its Accuracy, *Psych. Review, 18,* 186-201.

HEVNER, K.A. (1932)
A Method of Correcting for Guessing in True-False Tests and Empirical Evidence in Support of it, *Journ. of Soc. Psych., 3,* 359-362.

HOFFMAN, B. (1964)
The Tyranny of Testing, New York, Collier.

HOLLINGWORTH, R.L. (1913)
Experimental Studies in Judgment, *Archives of Psychology, 29,* 1-119.

HOPKINS, K.D. (1964)
Extrinsic Reliability, Estimating and Attenuating Variance from Response Styles, Chance and Other Irrelevant Sources, *Educ. and Psych. Meas., 24,* 271-281.

HOPKINS, K.D., HAKSTIAN, R.A., HOPKINS, B.R. (1973)
Validity and Reliability Consequences of Confidence Weighting, *Educ. and Psych. Meas., 33,* 135-141.

HORST, P. (1933)
The Difficulty of a Multiple-Choice Test Item, *Journ. of Educ. Psych., 24,* 229-232.

HOYT, C. (1941)
Test Reliability Estimated by Analysis of Variance, *Psychometrika, 6,* 153-160.

HURWICZ, L. (1951)
Optimality Criteria for Decision Making under Ignorance,

IRWIN, W.S. et SMITH, W.A. (1957)
Value, Cost and Information as Determiners of Decision, *J. Exp. Psych., 54,* 229-232.

ISAACSON, R.L.
Relation between Achievement, Test Anxiety and Curricular Choices, *J. Abnorm. Soc. Psych.*, *68*, 447-452.

JACKSON, P.H. (1975)
The Philosophy and Methodology of Bayesian Inference, Montreux, Second Intern. Sympos. on Educ. Testing.

JACOBS, P.I. et VANDEVENTER, M. (1968)
Information in Wrong Responses, *Research Bulletin*, 68, 25.

JACOBS, S.S. (1968)
An Empirical Investigation of the Relationship Between Selected Aspects of Personality and Confidence-Weighting Behaviors, Université de Maryland, thèse de doctorat, in *University Micro-Films*, 68, 16, 676.

JACOBS, S.S. (1971)
Correlates of Unwarranted Confidence in Responses to Objective Test Items, *Journ. of Educ. Meas.*, *8*, 1.

KARRAKER, R.J. (1967)
Knowledge of Results and Incorrect Recall of Plausible Multiple-Choice Alternatives, *Journ. of Educ. Psych.*, *58*, 11-14.

KAUFMANN, A., CULMANN, G., DENIS-PAPIN, M. (1960)
Eléments de calcul informationnel, Paris, Ed. A. Michel.

KENDALL, M.G. et STUART, A. (1961)
The Advanced Theory of Statistics, 3 vol., Londres, Griffin.

KUHLMAN, C. (1962)
Thought in the Young Child, Society for Research in Child Development.

KIDD, J.B. (1970)
The Utilization of Subjective Probabilities in Production Planning, *Acta Psychologica*, *34*, 338-347.

KINTSCH, W. (1963)
A Response Time Model for Choice Behavior, *Psychometrika*, *28*, 1.

KOEHLER, R.A. (1971)
A Comparison of the Validities of Conventional Choice Testing and Various Confidence Marking Procedures, *Journ. of Educ. Meas.*, *8*, 4.

KUDER, G.F. (1950)
Identifying the Faker, *Personnal Psychology*, *3*, 155-167.

LECLERCQ, D. (1971)
Une banque de questions pour l'enseignement, *Education, 132.*

LECLERCQ, D. (1973)
Critique des méthodes d'application,de correction et de cotation des questions à choix multiple, *Scientia Paedagogica Experim., X, 1.*

LECLERCQ, D. (1975)
Banque de questions et indices de certitude; options docimologiques adaptées à l'enseignement secondaire, *Education, 149-150.*

LEONARD, J.A. (1958)
Partial Advance Information in a Choice Reaction Time, *Brit. Journ. Psych., 49,* 89-96.

LEONARD, J.A. (1961)
Choice Reaction Time Experiments and Information Theory, in CHERRY, E.C. (Ed.).

LEWIS, D.G. (1968)
Experimental Design in Education, Univ. of London Press.

LEWY, A. et McGUIRE, C. (1966)
A Study of Alternative Approaches in Estimating the Reliability of Conventional Tests, American Educational Research Association, Chicago, document inédit.

LIEBLICH, A. (1968)
The Effect of Stress and the Motivation to Succeed on Test Risk, *Journ. of Personality, 36,* 608-615.

LINDER, D., WORTMAN, C. et BREHM, J.W. (1971)
Temporal Changes in Predecision Preferences among Choice Alternatives, *Journ. of Personality and Social Psych., 19,* 3,282-284.

LINDLEY, D.V. (1969)
Introduction to Probability and Statistics from a Bayesian Viewpoint, Part 1: Probability, Londres, Cambridge Univ. Press.

LINDLEY, D.V. (1970)
Introduction to Probability... Part 2: Inference, Londres, Cambridge Univ. Press.

LINDLEY, D.V. (1971)
Making Decisions, Londres, Wiley.

LITTIG, L.W. (1964)
The Effect of Motivation on Probability Preference and Subjective Personality, Ann Arbor, Univ. de Michigan (1959), cité par ATKINSON, J.W.

LITTLE, E. et CREASER, J. (1966)
Uncertain Responses on Multiple-Choice Examinations, *Psychol. Reports, 18,* 801-802.

LORD, F.M. (1963)
Formula Scoring and Validity, *Educ. and Psych. Meas., 23,* 663-672.

LORD, F.M. (1964)
The Effect of Random Guessing on Test Validity, *Educ. Psych. Meas., 24,* 745-747.

LORD, F.M. et NOVICK, M.R. (1968)
Statistical Theories of Mental Test Scores, Reading (Mass.), Addison-Wesley.

LORD, F.M. (1970)
Some Test Theory for Tailored Testing, in HOLTZMAN, W.H.

LORD, F.M. (1970)
The Self-Scoring Flexilevel Test, *E.T.S. Res. Bull., 70,* 43.

LOVIE, A.D. et DAVIES, D.M. (1970)
The Effect of Rate of Revision and Initial Revision on the Perception of Another's Age, *Acta Psych., 34,* 322-327.

LUCE, R.D. (1959)
Individual Choice Behavior, New York, Wiley.

LUCE, R.D., BUSH, R.R. et GALANTER, E. (1963)
Readings in Mathematical Psychology, New York, Wiley (2 vol.).

LUCE, R.D., RAIFFA, H. (1966)
Games and Decision, New York, Wiley.

LUCE, R.D., BUSH, R.R. et GALANTER, E. (1963)
Handbook of Mathematical Psychology, New York, Wiley (3 vol.).

LUMINGU, B. (1974)
Etude préalable à la construction d'un test diagnostique sur la consultation du dictionnaire, Université de Liège, mémoire de licence ronéotypé.

LYERLY, S.B. (1951)
A Note for Correcting for Chance Success in Objective Tests, *Psychometrika, 16,* 21-30.

Mc CLELLAND (1955)
Studies in Motivation, New York, Appleton.

Mc NEEL, S.P. et MESSICK, D.M. (1970)
A Bayesian Analysis of Subjective Probabilities of Interpersonal Relationships, *Acta Psych., 34,* 311-321.

MANZ, W. (1970)
Experiments on Probabilistic Information Processing, *Acta Psych.*, *34*, 184-200.

MARTIN, J.J. (1967)
Bayesian Decision Problems and Markov Chains, New York, Wiley.

MASSENGILL, H.E. et SHUFORD, E.H. (1967)
What Pupils and Teachers Should Know about Guessing, Lexington, Mass., Shuford-Massengill Corporation.

MASSENGILL, H.E. et SHUFORD, E.H. (1968)
A Report on the Effect of Degree of Confidence in Student Teaching, Air Force Office of Scientific Research, United States Air Force.

MEDLEY, D.M. (1966)
The Effects of Heterogenity of Content and Guessing on the Accuracy of Scores in Multiple-Choice Tests, *Amer. Educ. Research Journ.*, *3*, 27-33.

MELLENBERGH, G.J. (1967)
Nieuwe Ervaringen met een Zekerheidsaanduiding, *Ned. T. Psychol.*, *22*, 168-181.

MEUWESE, W., BARENDREGT, J.T. et VASTENHOUT, J. (1960)
Een Onderzoek naar de Relatie tussen de Juistheid van Oordelen en het Begeleidend Gevoel van Zekerheid, *Ned. Tijdschrift voor de Psychol.*, *15*, 529-541.

MICHAEL, J.J. (1968)
The Reliability of a Multiple-Choice Examination under Various Test-Making Instructions, *Journ. of Educ. Meas.*, *5*, 307-314.

MILLER, G.A. (1956)
The Magical Number Seven, Plus or Minus Two, *Psychol. Review*, *63*, 81-97.

MURPHY, A.H. (1966)
A note on the Utility of Probabilistic Predictions and the Probability Score in the Cost-Loss Ratio Decision Situation, *J. Appl. Meteorol.*, *5*, 534-537.

MURPHY, A.H. (1969a)
The Evaluation of Probabilistic Predictions in Meteorology, Ann Arbor, Univ. of Michigan, thèse de doctorat inédite.

MURPHY, A.H. (1969b)
Measures of the Utility of Probabilistic Predictions in Cost-Loss Ratio Decision Situations in Which Knowledge of the Cost-Loss Ratio is incomplete, *J. Appl. Meteorol.*, *8*, 863-873.

MURPHY, A.H. (1969c)
On Expected-Utility Measures in Cost-Loss Ratio Decision Situations, *J. Appl. Meteorol.*, *8*, 989-991.

MURPHY, A.H. (1970)
The Ranked Probability Score and the Probability Score: A Comparison, *Monthly Weather Review*, *98*.

MURPHY, A.H. et EPSTEIN, E.S. (1967)
Verification of Probabilistic Predictions: A Brief Review, *Journ. Appl. Meteorol*, *6*, 748-755.

MYERS, A.E.(1965)
Risk Taking and Academic Success and their relation to an Objective Measure of Achievement Motivation, *Educ. Psych. Meas.*, *25* 355-363.

NEWELL et SIMON (1961)
Computer Simulation of Human Thinking, *Science*, *134*.

NEWELL et SIMON (1965)
Computer in Psychology, in LUCE, BUSH, GALANTER, *Handbook of Mathematical Psychology*, p. 361.

NOVICK, M.R. (1969)
Bayesian Methods in Psychological Testing, *E.T.S. Res. Bulletin*, *69*, 31.

NOVICK, M.R. (1975)
Bayesian *Methods to Educational Testing. A Survey*, Montreux, Second Intern. Sympos. on Educational Testing.

OWENS, R.E., HANNA, G.S. et COPPEDGE, F.L. (1970)
Comparison of Multiple-Choice Tests Using Different Types of Distractor Selection Techniques, *Journ. of Educ. Meas.*, *7*, 2.

PHILIPS, A.J. (1943)
Further Evidence Regarding Weighted Versus Unweighted Scoring of Examinations, *Educ. and Psych. Meas.*, *3*, 151-155.

PIERON, H. (1973)
Vocabulaire de la Psychologie, Paris, P.U.F., p. 220.

POLLACK, I. (1954)
The Assimilation of Sequentially-Encoded Information, HFORL Report, TR 54-5.

POSTMAN, L. (1950)
Choice Behavior and the Process of Recognition, *Amer. Journ. of Psych.*, *63*, 576-583.

POTTHOFF, E.F. et BARRETT, N.E. (1932)
Comparison of Marks Based upon Weighted and Unweighted Items in New-Type Examinations, *Journ. of Educ. Psych.*, *23*, 92-98.

PROGNEAUX, Th. (1974)
Besoin d'accomplissement, Univ. de Liège, mémoire de licence inédit.

RAIFFA, H. (1970)
Decision Analysis, Introductory Lectures on Choice Under Uncertainty, New York, Addison-Wesley.

RICHELLE, M. (1966)
Le conditionnement operant, Neuchâtel, Delachaux et Niestlé.

RICHELLE, M. (1970)
Malentendus sur les apports du conditionnement, *Rev. Comp. Animal*, *4*, 1, 22-31.

RIPPEY, R.M. (1968)
A Fortran Program for Scoring and Analyzing Probabilistic Tests, *Behavioral Science*, *13*, 424.

RIPPEY, R.M. (1968)
Probabilistic Testing, *Journ. of Educ. Meas.*, *5*, 211-215.

RIPPEY, R.M. (1970)
A comparison of Five Different Scoring Functions for Confidence Tests, *Journ. of Educ. Meas.*, *7*, 3.

ROMBERG, T. et SHEPLER, J. (1968)
An Experiment Involving a Probability Measurement Procedure, Chicago, Annual Meeting of the American Educ. Res. Association.

ROSTAND, J. (1962)
Aux frontières du surhumain, Paris, Payot.

ROSTAND, J. (1959)
Carnet d'un biologiste, Paris, Ed. Stock.

ROUANET, H. (1961)
Etudes de décisions expérimentales et calcul de probabilités, in *La décision*, Paris, C.N.R.S., 33-43.

ROUANET, H. (1967)
Les modèles stochastiques d'apprentissage, Paris, Gauthier-Villars.

ROUANET, H. (1975)
Bayesian-Fiducial Methodology for the Analysis of Psychometric Data, Montreux, Second Intern. Symposium on Educational Testing.

RUCH, G.M. et STODDARD, G.D. (1925)
Comparative Reliabilities of Five Types of Objective Examinations, *Journ. of Educ. Psych., 16*, 89-103.

RUCH, G.M. et DEGRAAFF, M.H. (1926)
Corrections for Chance and Guess vs. Do Not Guess. Instructions in Multiple-Choice Tests, *Journ. of Educ. Psych., 17*, 368-375.

SANDBERGEN, S. (1968)
Test Strategie/Test Strategy, *Ned. T. Psychol., 23*, 16-38.

SANDBERGEN, S. (1972)
Meningen van Studenten over Zekerheidscoring / Students Opinions about Confidence Marking, R.I.T.P. memorandum (inédit).

SANDBERGEN, S. (1972)
Guessing and Confidence in Testing Educational Achievement, in CHOPPIN, B. (A/106 IEA Memorandum).

SAVAGE, L.J. (1951)
The Foundations of Statistics, New York, Wiley.

SCHUM, D.A., GOLDSTEIN, I.L., HOWELL, W.C. et SOUTHARD, J.F. (1967)
Subjective Probability under Several Cost Payoff Arrangements, *Org. Behav. Hum. Perform., 2*, 84-104.

SHANNON, C.E. (1948)
A Mathematical Theory of Communication, *Bell System Technical Journal, 27*.

SHANNON, C.E. et WEAVER, W. (1949)
The Mathematical Theory of Communication, Univ. of Illinois Press.

SHANNON, C.E. (1951)
Prediction and Entropy of Printed English, *Bell Syst. Techn. Journ., 30*, 50-64.

SHERRIFS, A.C. et BOOMER, D.S. (1954)
Who is Penalized by the Penalty for Guessing, *Journ. of Educ. Psych., 45*, 81-90.

SHOEMAKER, D.M. (1973)
Principles and Procedures of Multiple Matrix Sampling, Cambridge (Mass.), Ballinger.

SHUBIK, M. (1967)
Game Theory and Related Approaches to Social Behaviors, New York, Wiley.

SHUFORD, E.H. (1967)
How to Shorten a Test and Increase its Reliability and Validity, Lexington, Shuford-Massengill Corporation, technical report SCM R-8;

SHUFORD, E. (1969)
Systems of Confidence Weighting, Theory and Practice, Los Angeles, Annual Meeting of the American Educational Research Association.

SHUFORD, E., ALBERT, A. et MASSENGILL, N.E. (1966)
Admissible Probability Measurement Procedures, *Psychometrika, 31*, 125-145.

SIDMAN, M. (1953)
Avoidance Conditionning with Brief Schock and no Exteroceptive Warning Signal, *Science*.

SIEGEL, S. (1956)
Non parametric Statistics for the Behavioral Sciences, New York, Mc Graw-Hill.

SIEGEL, S., SIEGEL, A.E. et McMICHAEL, A.J. (1961)
Choice, Strategy and Utility, New York, Mc Graw-Hill.

SIMON, H.A. (1957)
Models of Man, New York, Wiley.

SIMON, H.A. (1962)
An Information Processing Theory of Intellectual Development, in KESEN, W. et KUHLMAN, C.

SIMON, H.A. (1967)
The Use of Information Processing Languages in Psychology, in *Colloque du C.N.R.S.*, 303-326.

SKINNER, B.F. (1971)
Analyse expérimentale du comportement (trad. M. RICHELLE), Paris, Dessart.

SLAKTER, M.J. (1967)
Risk-Taking on Objective Examinations, *American Educ. Journ., 4*, 31-43.

SLAKTER, M.J. (1968)
The Penalty for not Guessing, *Journ. of Educ. Meas., 5*, 141-144.

SLOVIC, P. (1962)
Convergent Validation of Risk Taking Measures, *Journ. of Abnorm. and Social Psych., 65*, 68-71.

SLOVIC, P. et LICHTENSTEIN, S. (1968)
Relative Importance of Probabilities and Payoffs in Risk Taking, *Journ. of Exper. Psych., 78.*

SLOVIC, P., LICHTENSTEIN, S. et EDWARDS, W. (1968)
Boredon Induced Changes in Preferences among Bets, *Amer. Journ. of Psych., 78,* 208-217.

SMITH, C.P. (1964)
Relationship Between Achievement-Related Motives and Intelligence, Performance Level, and Persistence, *J. Abnorm. Soc. Psych., 68,* 523-533.

SMITH, R.B. (1970)
An Empirical Investigation of Complexity and Process in Multiple-Choice Items, *Journ. of Educ. Meas., 7,* 1.

SODERQUIST, H.L. (1936)
A new Method of Weighting Scores in a True-False Test, *Journ. of Educ. Res., 30,* 290-292.

SOLOMON, H. (1961)
Studies in Item Analysis and Prediction, Stanford Univ. Press.

STANLEY, J.C. et WANG, M.C. (1968)
Differential Weighting. A Survey of Methods and Empirical Studies, New York, College Entrance Exam. Board.

STANLEY, J.C. et WANG, M.D. (1970)
Weighting Test Items and Test-Item Opinions, *Educ. and Psych. Measurement, 30,* 21-35.

STEVENS, S.S. (1951)
Handbook of Experimental Psychology, New York, Wiley.

STEVENS, S.S. (1957)
On the Psychophysical Law, *Psychol. Rev., 64,* 153-181.

STEVENS, S.S. (1959)
Measurement, Psychophysics and Utility, in C.W. CHURCHMAN et P. RATOOSH (Eds.), *Measurement, Definitions and Theories,* New York, Wiley.

STEVENS, S.S. (1962)
The Surprising Simplicity of Sensory Metrics, *Amer. Psychol., 17,* 29-39.

SWINEFORD, F. (1938)
The Measurement of a Personality Trait, *Journ. of Educ. Psych., 29,* 289-292.

SWINEFORD, F. (1941)
Analysis of a Personality Trait, *Journ. of Educ. Psych., 32,* 438-444.

SWINEFORD, F. et MILLER, P.M. (1953)
Effects of Directions Regarding Guessing on Item Statistics of a Multiple-Choice Vocabulary Test, *Journ. of Educ. Psych., 44*, 129-133.

Tables of the Cumulative Binomial Probability Distribution. (1955)
Cambridge, Mass., Harvard Univ. Press.

TAYLOR, W.L. (1953)
Cloze Procedure: A New Tool for Measuring Readability, *Journalism Quarterly, 30*, 415-433.

TAYLOR, W.L. (1956)
Recent Developments in the Use of Cloze Procedure, *Journalism Quarterly, 33*, 42-46.

TAYLOR, W.L. (1957)
Cloze Readability Scores as Indices of Individual Differences in Comprehension and Aptitude, *J. Appl. Psych., 41*, 19-26.

TERRACE, H.S. (1963)
Discrimination Learning with and without Errors, *Journ. Exper. Anal. Behav., 6*, 1-27.

TERRACE, H.S. (1963)
Errorless Transfert of a Discrimination Accross Two Continua, *Journ. Exp. Anal. Behav., 6*, 223-232.

THORNDIKE, R.L. et HAGEN, E. (1969)
Measurement and Evaluation in Psychology and Education, New York, Wiley, 3e éd.

THORNDIKE, R.L. (1971)
Educational Measurement, Amer. Council on Education, 2e éd.

THRALL, R.M., COOMBS, C.H. et DAVIES, R.L. (1954)
Decision Processes, New York, Wiley.

TIBERGHIEN, G. (1968)
Etude de la certitude du rappel au cours d'un apprentissage verbal, *Année psychol., 18*, 32-39.

TORGERSON, W. (1967)
Theory and Methods of Scaling, New York, Wiley, 1967.

TROW, W.C. (1923)
The Psychology of Confidence, an experimental Inquiry, *Archives of Psychology, 67*, 1-47.

TOURNEUR, Y. (1973)
Les objectifs du domaine cognitif (3 vol.), Univ. de Mons, thèse de doctorat inédite.

VAN NAERSSEN, R.F. (1962)
A Scale for the Measurement of Subjective Probability, *Acta Psychologica, 20,* 2, 159-166.

VAN NAERSSEN, R.F. et VAN BEAUMONT, R. (1965)
Ervaringen met een Zekerheidsaanduiding bij objektieve Tentamens, *Ned. T. Psychol., 20,* 308-315.

VAN NAERSSEN, R.D., SANDBERGEN, S. et BRUYNIS, E. (1966)
Is de Utiliteitscruve van Examenscores een Ogief ?, *Ned. T. Psychol., 21,* 6, 358-363.

VON NEUMANN, J. et MORGENSTERN, O. (1947)
Theory of Games and Economic Behavior, Princeton Univ. Press.

VOTAW, D.F. (1936)
The Effect of Do-Not-Guess Directions upon the Validity of True-False or Multiple-Choice Tests, *Journ. of Educ. Psychol., 28,* 698-703.

WATERS, C.W. et WATERS, L.K. (1971)
Validity and Likeability Ratings for three Scoring Instructions for a Multiple-Choice Vocabulary Test, *Educ. and Psych. Measur., 31,* 935-938.

WILEY, L.N. et TRIMBLE, O.C. (1936)
The Ordinary Objective Test as a Possible Criterion of Certain Personality Traits, *School and Society, 43,* 446-448.

WILLIAMSON, J. (1964)
Assessing Clinical Judgment, *J. of Medical Educ., 39,* 893.

WINKLER, R.L. (1967)
The Quantification of Judgment: Some Methodological Suggestions, *J. Amer. Statist. Ass., 62,* 1105-1120.

WINKLER, R.L. (1969)
Scoring Rules and the Evaluation of Probability Assessors, *Amer. Statist. Ass.,64* , 1073-1078.

WINKLER, R.L. et MURPHY, A.H. (1968)
"Good" Probability Assessors, *J. Appl. Meteorol., 7,* 751-758.

WINKLER, R.L. (1970)
Nonlinear Utility and the Probability Score, *J. Appl. Meteorol., 9,*143-148.

WOOD, R. (1971)
Computerized Adaptive Sequential Testing, Chicago University, Dissertation abstract.

ZILLER, R.C. (1957)
A Measure of the Gambling Response Set in Objective Tests, *Psychometrica, 22,* 289-292.

ANNEXES

ANNEXE 1. - Le programme FORTRAN "EVAL".

ANNEXE 2. - L'article "Banque de questions et indices de certitude" paru dans les numéros 149 (décembre 1974) et 150 (janvier 1975) de la revue Education.

ANNEXE 3. - Transformation d'une équation de VAN NAERSSEN.

ANNEXE 4. - Patterns individuels de l'expérience 1971-72
Patterns individuels de l'expérience 1972-73.

ANNEXE 5. - Texte utilisé dans l'expérience décrite aux chapitres 5 et 6.

ANNEXE 6. - Programmes FORTRAN utilisés au chapitre 6.

ANNEXE 7. - Les histogrammes individuels de distribution des degrés de certitude au(x) prétest(s) et post-test(s) lors de l'expérience du chapitre 6.

ANNEXE 1 (Cf p. 54)

Le programme FORTRAN "EVAL"

Ce programme, fréquemment modifié, présente, dans sa version de novembre 1975 les caractéristiques suivantes :

- Lecture et traitement optionnels en ce qui concerne :
  - le recours ou non aux indices de certitude;
  - le codage des réponses (alphabétiques ou numériques);
  - le nom complet (alphanumérique) des sujets;
  - l'ordre (nom-réponses ou réponses-nom);
  - le système "perfostyle" (cartes pré-perforées 40 colonnes).

- Un programme principal appelant les subroutines suivantes :
  - la SUBROUTINE CALCS (calcul simple) qui calcule le score individuel quand les indices de certitude n'ont pas été utilisés;
  - la SUBROUTINE CALCC (calcul avec certitude) dont on trouvera un exemple de sortie en p. 483;
  - la SUBROUTINE RPBIS (corrélation point bisériale) qui analyse les questions et dont on trouvera un exemple de sortie en p. 484;
  - la SUBROUTINE ESCAL (escalier) qui dresse l'escalier des résultats et dont on trouvera un exemple de sortie en p. 485;
  - la SUBROUTINE BULTIN (bulletin) qui ventile les résultats individuels selon n catégories (9 au maximum) dont on trouvera un exemple de sortie en p. 486;
  - la SUBROUTINE ORDRE, appelée par la SUBROUTINE ESCAL.

Actuellement, nos études portent sur les deux réalisations suivantes :

- Une version du programme qui traiterait les réponses de mêmes élèves aux mêmes questions, à deux moments différents (par exemple une même épreuve présentée en prétest puis en post-test d'un cours programmé);
- Une version de la SUBROUTINE CALCC qui traiterait jusqu'à dix degrés de certitude.

POST-TESTDU COURS PROGRAMME QU'EST-CE QU'UNE QUESTION A CHOIX MULTIPLE ?
SEMINAIRE EVALUATION NOVEMBRE 1975 LABOR. PEDAG. EXPER. UNIV. LIEGE

NUM.	MR	OM	BR	SS	SG	SC	-3	-2	-1	0	+1	+2	+3	%1	%2	%3
1	2	0	23	18.40	17.73	19.80	1	0	1	0	0	1	22	0	100	95
2	6	0	19	15.20	13.07	13.70	4	1	1	0	1	3	15	50	75	78
3	3	1	21	16.80	16.00	17.20	1	2	0	1	0	5	16	0	71	94
4	1	0	24	19.20	18.30	20.20	1	0	0	0	2	4	18	100	100	94
5	3	0	22	17.60	16.93	18.10	1	1	1	0	1	4	17	50	90	94
6	0	0	25	20.00	20.00	21.00	0	0	0	0	1	4	20	100	100	100
7	11	0	14	11.20	8.50	7.50	5	6	0	0	0	3	11	0	33	68
8	4	0	21	16.80	15.73	16.70	3	1	0	0	0	1	20	0	50	86
9	8	0	17	13.60	11.47	11.40	0	8	0	0	0	15	2	0	65	100
10	5	0	20	16.00	14.53	15.70	3	1	1	0	0	0	20	0	0	86
11	3	1	21	16.80	15.80	16.90	1	1	1	1	3	3	15	75	75	93
12	3	1	21	16.80	16.00	17.50	0	1	2	1	1	5	15	33	83	100
14	3	0	22	17.60	16.80	18.10	1	0	2	0	1	5	16	33	100	94
15	1	1	23	18.40	18.30	20.00	1	0	0	1	0	1	22	0	100	95
16	3	0	22	17.60	16.73	18.10	2	1	0	0	0	2	20	0	66	90
17	3	0	22	17.60	16.60	17.50	2	1	0	0	4	0	18	100	0	90
18	6	0	19	15.20	13.93	13.80	3	1	2	0	2	4	13	50	80	31
19	3	0	22	17.60	17.10	18.10	1	1	1	0	3	0	19	75	0	95
20	3	0	22	17.60	16.73	18.90	0	0	3	0	1	1	20	25	100	100
21	6	2	17	13.60	12.40	12.30	2	0	4	2	3	4	10	42	100	83
22	4	0	21	16.80	15.87	16.70	2	1	1	0	0	5	16	0	83	33
23	1	1	23	18.40	18.00	20.30	0	0	1	1	0	2	21	0	100	100
24	3	0	22	17.60	16.67	18.60	0	3	0	0	0	3	19	0	50	100
25	4	0	21	16.80	15.50	16.80	3	1	0	0	0	0	21	0	0	87
28	6	0	19	15.20	13.07	14.00	4	2	0	0	0	1	18	0	33	81
29	5	0	20	16.00	14.93	15.30	3	0	2	0	2	1	17	50	100	85
30	2	0	23	18.40	17.73	19.50	1	0	1	0	1	2	20	50	100	95
31	3	1	21	16.80	15.80	17.30	1	1	1	1	1	3	17	50	75	94
32	7	0	18	14.40	12.50	11.80	5	2	0	0	1	6	11	100	75	68
33	0	0	25	20.00	20.00	22.10	0	0	0	0	0	4	21	0	100	100
34	2	0	23	18.40	18.00	19.70	1	0	1	0	1	0	22	50	0	95
36	1	0	24	19.20	18.30	20.00	1	0	0	0	4	2	18	100	100	94
37	6	0	19	15.20	13.70	14.30	0	2	4	0	3	8	8	42	80	100
38	5	0	20	16.00	14.80	15.30	3	2	0	0	0	3	17	0	60	85
39	9	0	16	12.80	9.87	10.60	2	7	0	0	0	5	11	0	41	84
40	7	1	17	13.60	12.13	11.50	5	0	2	1	1	2	14	33	100	73
41	3	1	21	16.80	16.00	16.50	1	1	1	1	4	5	12	80	83	92
42	6	1	18	14.40	12.53	13.10	3	2	1	1	1	3	14	50	60	82
43	7	5	13	10.40	8.27	8.80	2	4	1	5	1	1	11	50	20	84
44	3	5	17	13.60	12.53	13.30	2	0	1	5	3	0	14	75	0	87
45	9	1	15	12.00	9.93	9.30	4	3	2	1	0	5	10	0	62	71
46	1	0	24	19.20	19.10	20.60	1	0	0	0	0	4	20	0	100	95
47	4	0	21	16.80	15.60	16.80	2	1	1	0	0	4	17	0	80	89
50	4	1	20	16.00	14.67	16.00	1	3	0	1	0	5	15	0	62	93
51	10	3	12	9.60	7.43	5.10	2	6	2	3	11	1	0	84	14	0
52	5	0	20	16.00	14.67	14.80	3	2	0	0	3	2	15	100	50	83
53	1	0	24	19.20	18.93	21.00	0	0	1	0	2	0	22	66	0	100
54	7	0	18	14.40	12.73	13.20	1	2	4	0	3	4	11	42	66	91
55	5	2	18	14.40	13.33	14.30	0	3	2	2	1	4	13	33	57	100
57	4	1	20	16.00	14.87	16.10	1	1	2	1	1	4	15	33	80	93
58	4	0	21	16.80	15.47	16.30	1	3	0	0	0	11	10	0	78	90
59	4	1	20	16.00	14.83	16.20	1	0	3	1	3	0	17	50	0	94
60	4	0	21	16.80	15.80	17.00	2	1	1	0	0	2	19	0	66	90
62	10	0	15	12.00	9.13	10.10	2	2	6	0	1	2	12	14	50	85
63	8	0	17	13.60	11.73	11.10	5	1	2	0	0	5	12	0	83	70
64	3	0	22	17.60	17.07	18.30	2	0	1	0	0	1	21	0	100	91
65	0	0	25	20.00	20.00	21.90	0	0	0	0	3	0	22	100	0	100
66	8	1	16	12.80	10.47	11.40	2	1	5	1	1	3	12	16	75	85
67	7	0	18	14.40	12.97	12.60	5	2	0	0	0	0	18	0	0	78
68	2	0	23	18.40	17.60	19.70	1	1	0	0	0	1	22	0	50	95
69	3	0	22	17.60	17.00	17.30	2	1	0	0	3	4	15	100	80	33
70	3	3	19	15.20	14.30	15.20	3	0	0	3	0	1	18	0	100	85
71	7	1	17	13.60	11.77	12.30	2	2	3	1	1	4	12	25	66	85
72	9	0	16	12.80	10.50	8.60	9	0	0	0	0	4	12	0	100	57
73	5	0	20	16.00	14.60	16.00	1	2	2	0	1	2	17	33	50	94
74	4	0	21	16.80	15.90	16.00	3	1	0	0	0	2	19	0	66	86
75	5	0	20	16.00	14.53	15.40	4	0	1	0	0	0	20	0	0	33
76	8	0	17	13.60	11.70	10.40	5	2	1	0	3	5	9	75	71	64
77	6	0	19	15.20	13.67	14.40	1	5	0	0	0	6	13	0	54	92
78	4	0	21	16.80	16.00	17.50	1	1	2	0	0	1	20	0	50	95
79	1	0	24	19.20	18.30	20.30	1	0	0	0	1	0	23	100	0	95
80	7	4	14	11.20	9.13	9.10	5	1	1	4	0	0	14	0	0	73
81	9	1	15	12.00	9.47	9.20	4	4	1	1	2	1	12	66	20	75
92	2	0	23	18.40	18.00	19.30	1	1	0	0	0	0	23	0	0	95
TOT.	334	40	1476	15.96	14.75	15.59	149	106	79	40	86	209	1181	52	66	88
LES SIGMAS =				2.40	3.04	3.82	-3	-2	-1	OM	+1	+2	+3			

POST-TESTDU COURS PROGRAMME QU'EST-CE QU'UNE QUESTION A CHOIX MULTIPLE ?
SEMINAIRE EVALUATION NOVEMBRE 1975 LABOR. PEDAG. EXPER. UNIV. LIEGE

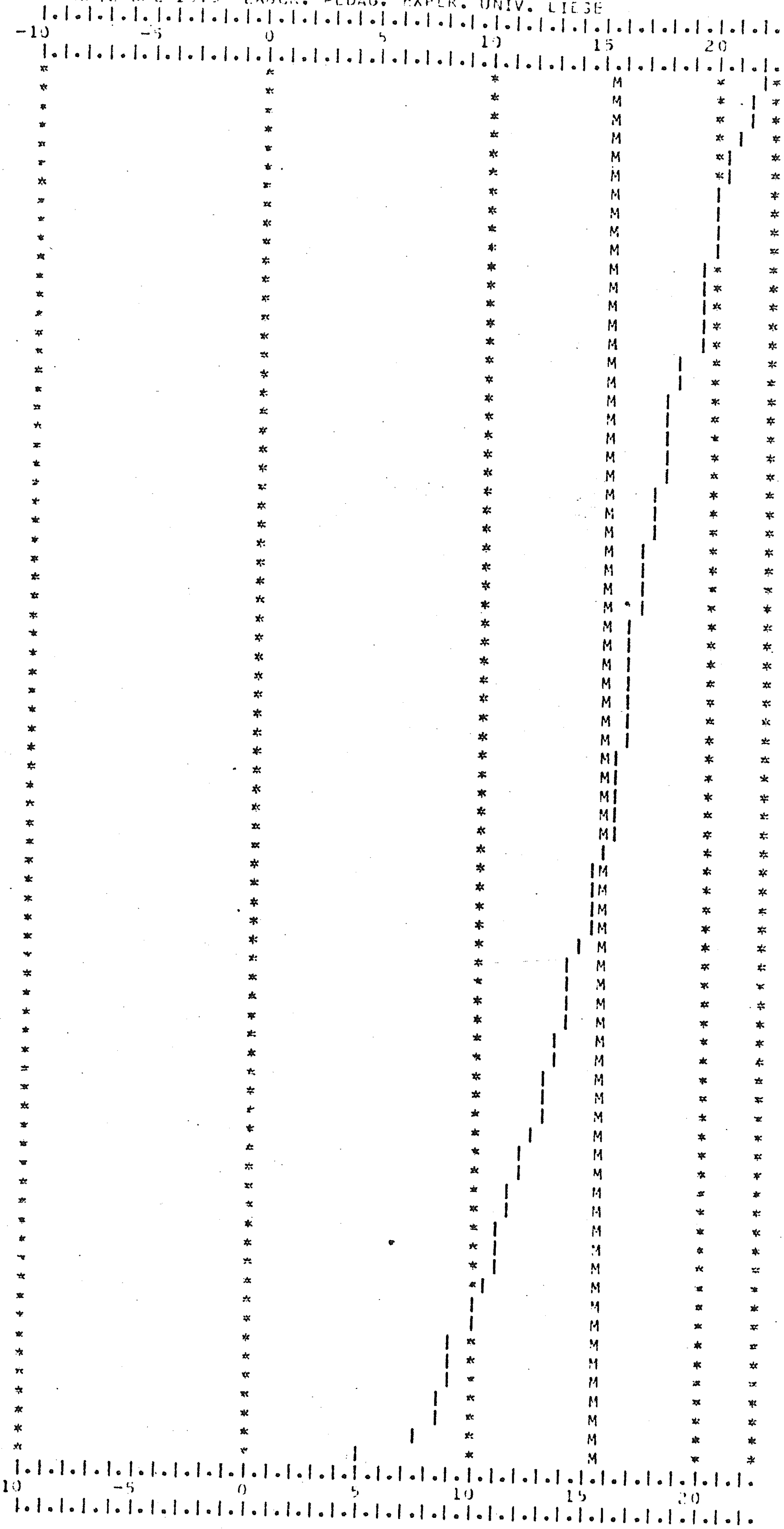
RESULTATS DE L'ANALYSE D'ITEMS

ITEM	BONNE REPONSE			DISTRIBUTION (EN %) DES REPONSES										CORRELATIONS									
		%	RPBIS BR	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
1	3	81	0.39	0	5	10	81	0	2	0	0	0	0	0.0	-0.39	-0.12	0.39	0.0	-0.10	0.0	0.0	0.0	0.0
		60		0	4	8	60	0	2	0	0	0	0										
2	1	85	0.55	1	85	2	4	1	4	1	0	0	0	-0.18	0.55	-0.04	-0.43	-0.12	-0.25	-0.11	0.0	0.0	0.0
		63		1	63	2	3	1	3	1	0	0	0										
3	4	97	0.06	0	1	1	0	97	0	0	0	0	0	0.0	-0.12	0.04	0.0	0.06	0.0	0.0	0.0	0.0	0.0
		72		0	1	1	0	72	0	0	0	0	0										
4	7	52	0.48	4	5	1	2	0	13	20	52	0	0	-0.10	-0.25	-0.08	-0.16	0.0	-0.27	-0.07	0.48	0.0	0.0
		39		3	4	1	2	0	10	15	39	0	0										
5	1	77	0.39	0	77	9	5	0	1	6	0	0	0	0.0	0.39	-0.27	-0.17	0.0	-0.11	-0.12	0.0	0.0	0.0
		57		0	57	7	4	0	1	5	0	0	0										
6	2	86	0.26	1	8	86	2	0	0	0	1	0	0	0.05	-0.30	0.26	-0.09	0.0	0.0	0.0	0.04	0.0	0.0
		64		1	6	64	2	0	0	0	1	0	0										
7	4	89	0.24	0	1	9	0	89	0	0	0	0	0	0.0	-0.05	-0.23	0.0	0.24	0.0	0.0	0.0	0.0	0.0
		66		0	1	7	0	66	0	0	0	0	0										
8	2	71	0.29	4	24	71	0	0	0	0	0	0	0	-0.16	-0.24	0.29	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		53		3	18	53	0	0	0	0	0	0	0										
9	3	85	0.26	2	8	1	85	0	2	0	0	0	0	-0.16	-0.09	-0.14	0.26	0.0	-0.09	0.0	0.0	0.0	0.0
		63		2	6	1	63	0	2	0	0	0	0										
10	2	75	0.12	1	22	75	0	0	0	0	0	0	0	-0.18	-0.06	0.12	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		56		1	17	56	0	0	0	0	0	0	0										
11	5	82	0.28	4	4	0	2	0	82	6	0	0	0	-0.20	-0.26	0.0	-0.09	0.0	0.28	-0.00	0.0	0.0	0.0
		61		3	3	0	2	0	61	5	0	0	0										
12	3	66	0.35	1	6	5	66	20	0	0	0	0	0	-0.18	-0.19	-0.22	0.35	-0.10	0.0	0.0	0.0	0.0	0.0
		49		1	5	4	49	15	0	0	0	0	0										
13	2	60	0.23	5	32	60	1	0	0	0	0	0	0	0.13	-0.30	0.23	-0.02	0.0	0.0	0.0	0.0	0.0	0.0
		45		4	24	45	1	0	0	0	0	0	0										
14	1	78	0.45	13	78	8	0	0	0	0	0	0	0	-0.45	0.45	-0.11	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		58		10	58	6	0	0	0	0	0	0	0										
15	4	78	0.47	2	10	4	0	78	4	0	0	0	0	-0.05	-0.28	-0.26	0.0	0.47	-0.21	0.0	0.0	0.0	0.0
		58		2	8	3	0	58	3	0	0	0	0										
16	2	94	0.26	0	4	94	1	0	0	0	0	0	0	0.0	-0.24	0.26	-0.10	0.0	0.0	0.0	0.0	0.0	0.0
		70		0	3	70	1	0	0	0	0	0	0										
17	2	90	0.40	0	9	90	0	0	0	0	0	0	0	0.0	-0.40	0.40	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		67		0	7	67	0	0	0	0	0	0	0										
18	2	33	0.25	8	13	33	43	1	0	0	0	0	0	-0.23	-0.23	0.25	0.09	-0.17	0.0	0.0	0.0	0.0	0.0
		25		6	10	25	32	1	0	0	0	0	0										
19	3	77	0.45	1	5	14	77	0	1	0	0	0	0	0.02	-0.11	-0.44	0.45	0.0	-0.02	0.0	0.0	0.0	0.0
		57		1	4	11	57	0	1	0	0	0	0										
20	1	89	0.12	0	89	10	0	0	0	0	0	0	0	0.0	0.12	-0.12	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		66		0	66	8	0	0	0	0	0	0	0										
21	2	85	0.21	0	12	85	2	0	0	0	0	0	0	0.0	-0.22	0.21	-0.01	0.0	0.0	0.0	0.0	0.0	0.0
		63		0	9	63	2	0	0	0	0	0	0										
22	3	98	0.14	0	0	1	98	0	0	0	0	0	0	0.0	0.0	-0.14	0.14	0.0	0.0	0.0	0.0	0.0	0.0
		73		0	0	1	73	0	0	0	0	0	0										
23	4	94	0.10	0	0	4	1	94	0	0	0	0	0	0.0	0.0	-0.01	-0.17	0.10	0.0	0.0	0.0	0.0	0.0
		70		0	0	3	1	70	0	0	0	0	0										
24	2	78	0.37	0	16	78	5	0	0	0	0	0	0	0.0	-0.35	0.37	-0.09	0.0	0.0	0.0	0.0	0.0	0.0
		58		0	12	58	4	0	0	0	0	0	0										
25	1	85	0.47	2	85	4	8	0	0	0	0	0	0	-0.24	0.47	-0.25	-0.27	0.0	0.0	0.0	0.0	0.0	0.0
		63		2	63	3	6	0	0	0	0	0	0										

MOYENNE DES R P BIS = 0.3035
FIDELITE ESTIMEE A PARTIR DES RPBIS = 0.7173

POST-TESTDU COURS PROGRAMME   QU'EST-CE QU'UNE QUESTION A CHOIX MULTIPLE ?
SEMINAIRE EVALUATION   NOVEMBRE 1975 LABOR. PEDAG. EXPER. UNIV. LIEGE

1	33	22.10
2	6	21.90
3	65	21.90
4	53	21.00
5	79	20.80
6	46	20.60
7	23	20.30
8	4	20.20
9	15	20.00
10	36	20.00
11	1	19.80
12	82	19.80
13	34	19.70
14	68	19.70
15	30	19.50
16	20	18.90
17	24	18.60
18	64	18.30
19	16	18.10
20	19	18.10
21	14	18.10
22	5	18.10
23	78	17.50
24	17	17.50
25	12	17.50
26	69	17.30
27	31	17.30
28	3	17.20
29	60	17.00
30	11	16.90
31	47	16.80
32	25	16.80
33	22	16.70
34	8	16.70
35	74	16.60
36	41	16.50
37	58	16.30
38	59	16.20
39	57	16.10
40	50	16.00
41	73	16.00
42	10	15.70
43	75	15.40
44	38	15.30
45	29	15.30
46	70	15.20
47	52	14.80
48	77	14.40
49	37	14.30
50	55	14.30
51	28	14.00
52	18	13.80
53	2	13.70
54	44	13.30
55	54	13.20
56	42	13.10
57	67	12.60
58	71	12.30
59	21	12.30
60	32	11.80
61	40	11.50
62	9	11.40
63	66	11.40
64	63	11.10
65	39	10.60
66	76	10.40
67	62	10.10
68	45	9.30
69	81	9.20
70	80	9.10
71	43	8.80
72	72	8.60
73	7	7.50
74	51	5.10



BANQUE DE QUESTIONS ET INDICES DE CERTITUDE (1)

OPTIONS DOCIMOLOGIQUES ADAPTEES A L'ENSEIGNEMENT SECONDAIRE

INTRODUCTION

L'évaluation (...) a toujours, directement ou indirectement, rapport avec le progrès, en extension ou en qualité, de l'apprentissage.

G.DE LANDSHEERE, 1971, p. 14.

La thèse ci-dessus est de plus en plus souvent démontrée dans les actions expérimentales récentes (1). Il faut cependant que l'évaluation, processus si important, puisse s'appuyer sur une technologie adéquate. En effet, la définition des objectifs puis la formulation des questions constituent des tâches suffisamment laborieuses pour que leurs résultats soient partagés par le plus grand nombre. C'est ce but que vise la création d'une banque de questions (2). La correction des épreuves, l'établissement des graphiques, d'indices diagnostiques, de bulletins, ...,.forment un volume de travail routinier tel qu'il doit être mécanisé. C'est la raison pour laquelle on gère la banque de questions par ordinateur.

Le présent article décrit les options générales et circonstancielles prises dans deux banques de questions (3) à la création desquelles nous avons activement collaboré. Ces banques servent non seulement l'évaluation, mais aussi la formation permanente des enseignants.

Les options techniques portent sur les types de questions, les modes de réponses, les procédures de correction et de cotation, l'utilisation d'un indice de certitude dans la réponse. Cette dernière option mérite un mot d'explication. On sait que, dans les questions à choix multiple, l'élève peut répondre au hasard. Par conséquent, des **corrections for guessing** ont été mises au point. La plus classique d'entre elles consiste à accorder

(1) Voir, par exemple, M. DETHEUX et E. LECLERCQ, 1972.
(2) Voir D. LECLERCQ, 1971.
(3) La première à l'Ecole Technique de la Force aérienne belge (Saffraanberg), la seconde dans certains instituts d'enseignement secondaire de la province de Liège.

un point en cas de choix correct et à retirer 0,25 point en cas de choix erroné (lorsque la question présente cinq solutions possibles). Certains auteurs (4) proposent de demander à l'élève, en plus de sa réponse, un indice de certitude qui serait l'expression de sa « probabilité de fournir la réponse correcte ». Un tel indice permet non seulement d'éviter, dans une certaine mesure (5), le guessing, mais également d'évaluer ce que Coombs et al. (6) appellent la « connaissance partielle ».

L'indice de certitude a été utilisé à titre expérimental, et les résultats de deux ans d'expérience sont particulièrement satisfaisants. Dans le présent article, nous en montrons l'apport pédagogique. En ce qui concerne sa justification statistique, nous nous permettons de renvoyer à B. De Finetti, 1965, et D. Leclercq, 1973 et 1975.

1. PRINCIPES DE CREATION ET D'UTILISATION D'UNE BANQUE DE QUESTIONS

Les banques de questions décrites ci-après résultent de six grandes options.

Option 1 : *Le choix de la question comme unité d'évaluation*

Pendant longtemps, les tests standardisés ont été considérés comme l'unité d'information dans l'évaluation pédagogique. Malheureusement, les tests sont beaucoup trop rigides : l'addition ou le retrait d'items rendent les résultats finaux ininterprétables. L'expérience montre pourtant qu'il est rare que tous les items d'un test pédagogique conviennent à la situation d'évaluation particulière où se trouve un enseignant à un moment donné. Dans de nombreux cas, on observe une discordance entre les objectifs de l'apprentissage scolaire et l'épreuve standardisée.

La perspective envisagée ici est bien plus souple ; elle rejette la standardisation à l'arrière-plan en faisant coller l'évaluation à l'enseignement.

Option 2 : *La prise en charge de l'évaluation par l'enseignant*

Jusqu'à ces derniers temps, il n'était pas exceptionnel qu'une autorité impose l'application d'une épreuve pédagogique et qu'un service technique se charge de l'administration et de l'interprétation de cette épreuve, en dehors de toute intervention des enseignants. Or, dissocier l'enseignement de l'évaluation équivaut à ignorer un des principes les plus fondamentaux de la pédagogie. Pour accomplir efficacement leur mission, les maîtres doivent donc savoir à la fois susciter les apprentissages et les évaluer.

(4) Surtout DE FINETTI, 1965.
(5) Voir D. LECLERCQ, 1973.
(6) COOMBS, MILHOLLAND et WOMER (1956).

Cette évaluation, comme on la comprend aujourd'hui, est lourde. Aussi convient-il d'en réserver les aspects les plus nobles (décision et exploitation) aux enseignants et d'en confier les aspects routiniers (reproduction, calcul, présentation) à des machines ou à des services techniques.

Option 3 : *Le recours aux techniques les plus performantes*

L'ordinateur prend en charge le traitement de l'information. Pour la reproduction de cette information, on recourt tantôt à la photocopie, tantôt au listage de cartes par l'ordinateur.

Les résultats d'un test de vingt questions passé par 60 étudiants sont traités en cinq secondes par le programme EVAL exécuté sur l'IBM 360/65, alors qu'il faudrait près d'une centaine d'heures à un homme (équipé d'une machine à calculer électronique) pour exécuter le même travail (calcul des moyennes et sigmas par classe, des notes brutes et pondérées de chaque sujet, des indices de chaque item, établissement de bulletins et de courbes cumulatives, ...).

Option 4 : *La spécialisation des tâches et la multi-disciplinarité de l'équipe*

Les compétences et les talents doivent être utilisés là où ils sont. Le spécialiste de la matière crée les questions et les tests. L'informaticien rédige le programme d'exploitation que le pédagogue a conçu.

Ceci implique cependant une large communauté de préoccupations parmi les membres d'une telle équipe. L'enseignant et le pédagogue doivent prendre des options méthodologiques semblables et le pédagogue et l'informaticien doivent se retrouver sur le terrain du programme (FORTRAN ou PL1). Des petits ouvrages illustrés (7) ont été élaborés dans le but de créer une unité de méthode, d'options et de langage entre tous.

Option 5 : *La mise en commun des ressources*

En faisant éclater les limites des établissements et des réseaux scolaires, on peut réaliser ce que P.H. COOMBS (1968, p. 197) appelle des « économies d'échelle » par une capitalisation de biens intellectuels. C'est une grande satisfaction pour nous d'avoir pu intéresser à une entreprise commune des établissements scolaires relevant de la Force aérienne, de la province de Liège, de l'Etat, de l'Enseignement libre.

Option 6 : *La sensibilité aux conditions circonstancielles*

Un système pédagogique ne peut accepter une trop grande quantité d'innovations simultanées ; il faut donc définir des priorités. Presque toujours, il importe de ménager des périodes de transition pour ne pas heurter les enseignants, les parents, et les élèves eux-mêmes.

(7) D. LECLERCQ, P. VAN ROY et P. MOONEN, *Tests objectifs*. Saint-Trond, Ecole Technique de la Force aérienne, 1972 (3 fascicules).

L'enseignement est imprégné de principes tenaces, souvent tenus pour évidents. Il ne parait, par exemple, guère pensable d'attribuer des cotes négatives, de permettre de gagner la moitié des points en répondant à moins de la moitié des questions, ou encore de plus sanctionner une erreur qu'une omission. Très liés à un moment de l'histoire pédagogique, ces principes sont plus ou moins explicites, plus ou moins « évidents », diversement fondés et controversés (voir le débat actuel entre la notation chiffrée et l'utilisation d'échelles ordinales ou descriptives).

Ces considérations nous ont amené à prendre une série d'options. Nous y consacrons le reste de cet article.

2. OPTIONS

A. Types d'items utilisés

La liste ci-dessous est inspirée de divers auteurs (8) ; elle ne contient que les types d'items permettant une **correction objective**.

1. *LES ITEMS DE PRODUCTION* (ou à « réponse ouverte »), à **réponse courte**. Ces items, le plus souvent écrits, doivent être corrigés par le professeur qui utilise, par exemple, les codes

0 (ou O) en cas d'omission,
1 (ou R) en cas de réussite,
2 (ou E) en cas d'échec,
3 (ou P) en cas de réussite partielle, etc.

2. *LES ITEMS DE SELECTION* (ou « à réponse fermée ») :

— *à alternative :*

- opinion/fait (9)
- vrai/faux (également : oui/non, est/n'est pas, même/différent, etc.)

— *à choix multiple :*

- avec 3 solutions proposées, par exemple :
  vrai/faux/opinion
  féminin/masculin/neutre
  négatif/nul/positif
- avec 4 ou 5 solutions, ménageant éventuellement la *possibilité d'omission (le* code est 0 ou O dans ce cas) *et de rejet* (le code est alors 6 ou R) de toutes les solutions proposées.

3. *LES ITEMS COMBINES* peuvent être considérés comme plusieurs items, aussi bien au point de vue de la correction que de la pondération.

a) *Les questions à appariements* sont considérées comme n questions à choix multiple dont les solutions proposées seraient identiques ;

(8) Notamment G. DE LANDSHEERE, 1971 b, pp. 56 et 85; L. VANDE VELDE, 1970.
(9) L. VANDE VELDE, 1972.

b) *Les questions « double face »*, où une seconde question de justification suit la première, permettent des combinaisons très diverses : une question à choix multiple justifiée par une question à réponse courte, une question à choix multiple justifiée par une autre question à choix multiple, etc. Ces questions sont aussi appelées « questions en cascade » ou « questions emboîtées ».

B. Le coefficient de certitude

1. LA CERTITUDE « ORDINALE »

Pour chaque réponse, l'élève doit, en outre, proposer un coefficient de certitude selon le code suivant :

1 = Je ne suis pas certain de ma réponse
2 = Certitude normale
3 = Je suis très certain de ma réponse.

Les élèves sont informés qu'en cas de bonne réponse, ils gagneront d'autant plus de points que leur certitude est élevée ; en cas de mauvaise réponse, ils perdront d'autant plus de points que leur certitude est élevée.

On remarque que la consigne est ordinale et qu'il n'est pas possible de définir les intervalles qui séparent les degrés de certitude. On remarquera aussi que les points attribués sont très simples, faciles à retenir et faciles à calculer. Malheureusement, ils ne correspondent pas à la théorie des décisions. (D. LECLERCQ, 1975.)

L'expérimentateur communique aux élèves le tableau des points obtenus pour chaque réponse :

	Certitude	*Points*
	3	+ 3
En cas de bonne réponse	2	+ 3
	1	+ 1
Omission		0
	1	— 1
En cas de mauvaise réponse	2	— 2
	3	— 3

A partir de ce tableau, on montre aux élèves qu'il est préférable d'obtenir 0 (omission) plutôt que — 3 points. Cependant, l'expérience a montré que les élèves tiennent peu compte de la consigne lors du premier test : ils utilisent presque uniquement la certitude 3, ce qui amène des scores très faibles. Dès le second test, la loi de l'effet ayant joué, les élèves ajustent bien mieux leur certitude.

2. LA FIABILITE SUBJECTIVE

Afin de définir les intervalles qui séparent les différents degrés de certitude, on fournit une consigne qui fait référence à la *probabilité de succès* (ou fiabilité) de la réponse.

Cette consigne est, évidemment, préférable à la consigne ordinale.

Les points, plus compliqués, sont conformes à la théorie des décisions. Nous présentons ceux qui ont été choisis par le commandant P. VAN ROY à la banque de questions de l'E.T.F.A. à Saffraanberg.

Voici les consignes d'utilisation.

Comment utiliser l'indice de certitude

Vous devez inscrire à côté de chaque réponse (et entre parenthèses) le degré de certitude avec lequel vous répondez à la question.

Si vous êtes *très certain* (entre 75 et 100 % de chance de fournir la bonne réponse), utilisez la certitude 3.

Si vous êtes *certain* (entre 50 et 75 % de chance de fournir la réponse correcte), utilisez la certitude 2.

Si vous êtes *peu certain* (entre 25 et 50 % de chance de fournir une réponse correcte), utilisez la certitude 1.

Si votre certitude est inférieure à 25 %, ne répondez pas. Il n'y a évidemment aucun indice de certitude.

EN RESUME :

omettre	certitude 1	certitude 2	certitude 3	
0 %	25 %	50 %	75 %	100 %

chances de fournir la réponse correcte.

VOUS AVEZ MATHEMATIQUEMENT INTERET A CONFORMER VOTE INDICE DE CERTITUDE A VOTRE CERTITUDE REELLE.

Les points à gagner ou à perdre ont été calculés **dans ce but** :

Si votre réponse est exacte	*vous obtiendrez les points suivants :*
avec la certitude 3	2,25
la certitude 2	2
la certitude 1	1,75
Si vous omettez	0
Si votre réponse est fausse	
avec la certitude 1	— 0,5
la certitude 2	— 0,75
la certitude 3	— 1,5

Vous voyez qu'il est préférable d'obtenir 0 que — 1,5.

ATTENTION : Dans les questions à choix multiples,

— la bonne réponse ne figure pas toujours parmi les solutions proposées. Il faut alors les rejeter toutes ;

— si vous oubliez d'indiquer votre certitude pour une réponse, nous indiquerons automatiquement la certitude 2.

Il est possible de présenter des points à gagner et à perdre qui tiennent compte du nombre de solutions proposées dans la question. Cette méthode présente cependant des inconvénients au niveau de la consigne.

C. Les feuilles de réponses

La feuille de réponse de l'élève est une bandelette :

NOM :							
QUESTION	1	2	3	4	5	37	38
REPONSE							
CERTITUDE							

Voici les réponses d'un élève à une épreuve de 9 questions :

NOM		Date :					
QUESTION	1	2	3	4	5	37	38
REPONSE	3	4	1	1	2		
CERTITUDE	1	3	2		1		

Commentaires

A la question 6, l'élève n'a pas répondu.

A la question 4, l'élève n'a pas noté sa certitude (la perforation par défaut est 2).

Voici les réponses-modèles (clés) fournies par le professeur, ainsi que l'importance de chaque question :

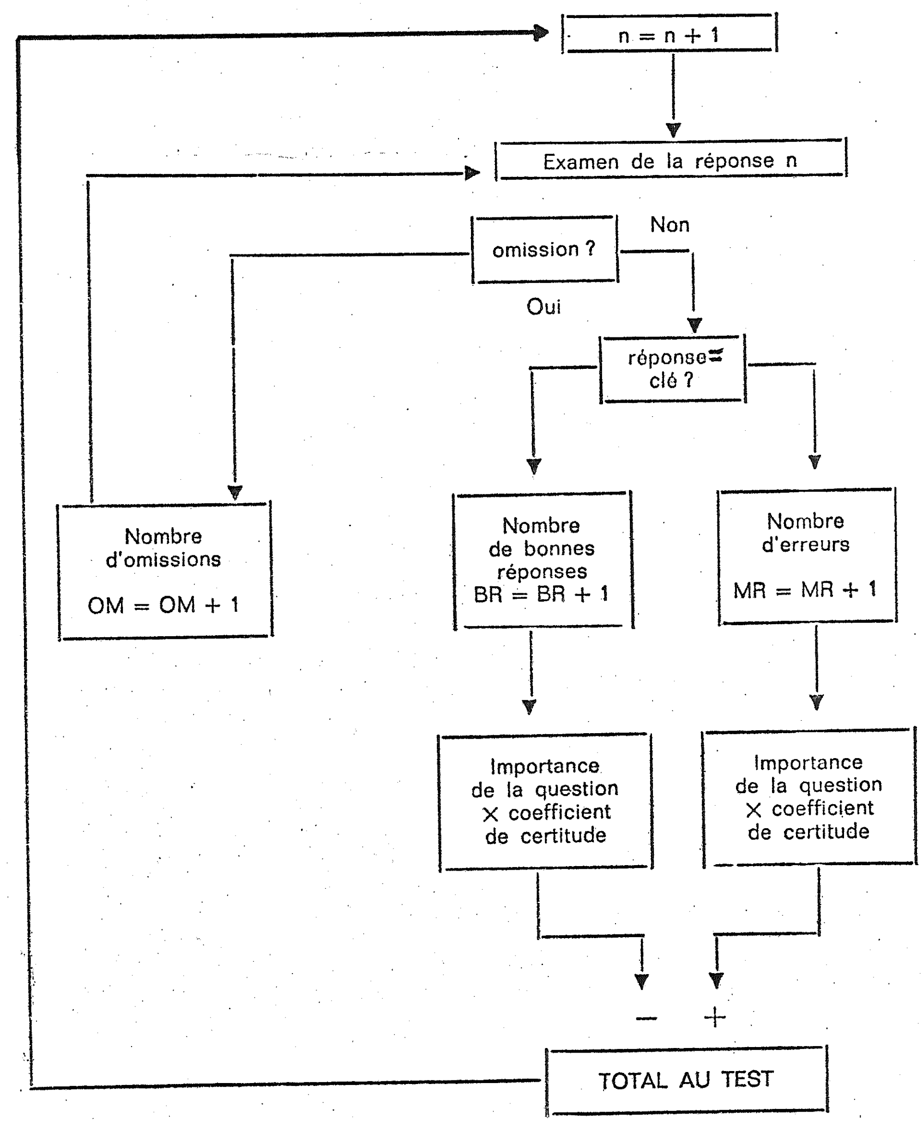
TEST : Electricité		Date : 8-2-72					
QUESTION	1	2	3	4	5	37	38
CLE	3	1	1	2	4		
IMPORTANCE	6	7	5	7	7		

Commentaires

La question 3 (importance 5) est $\frac{5}{7}$ fois plus importante que la question 2 (importance 7 ; cette nuance est légère. Si l'importance de la question 2 avait été 1, la question 3 aurait été 5 fois plus importante.

D. La correction

On porte les réponses, les certitudes et l'identification de l'élève sur une carte perforée. Le programme confronte chaque réponse avec la clé, selon l'algorithme suivant :



Sur le listing figureront les nombres de mauvaises réponses (MR), d'omissions (OM) et de bonnes réponses (BR). On y lira, en outre :

— le nombre de mauvaises réponses avec la certitude 3 (—3)
2 (—2)
1 (—1)

— le nombre d'omissions 0

— le nombre de bonnes réponses avec la certitude 1 (+1)
2 (+2)
3 (+3)

Voici, par exemple, les résultats obtenus pour un test [10] de trente questions, passé le 21 octobre 1971 par seize élèves.

	NOM	MR	OM	BR	Mauvaises réponses			CM	Bonnes réponses			Cote totale
					—3	—2	—1		+1	+2	+3	
1	TOM	9	0	21	8	1	0	0	0	2	19	11.63
2	BUL	4	6	20	4	0	0	6	0	0	20	13.47
3	VOS	9	2	19	9	0	0	2	2	1	16	7.96
4	FLO	4	0	26	4	0	0	0	0	2	24	21.43
5	DAN	6	4	20	5	0	1	4	1	2	17	10.61
6	STI	4	0	26	2	1	1	0	0	0	26	20.00
7	WAE	4	2	24	4	0	0	2	0	0	24	16.53
8	KAM	7	1	22	7	0	0	1	0	0	22	12.86
9	BAR	11	0	19	10	1	0	0	0	0	19	9.39
10	LEC	6	0	24	4	2	0	0	0	3	21	17.14
11	PLA	3	0	27	3	0	0	0	0	2	25	23.27
12	VER	6	0	24	2	2	2	0	1	9	14	15.10
13	LAZ	7	1	22	5	1	1	1	3	0	19	10.41
14	FNN	11	0	19	3	5	3	0	1	1	17	10.61
15	KER	4	0	26	3	0	1	0	3	3	20	20.20
16	VER	4	3	23	2	2	0	3	4	8	11	12.24
TOT.	480	99	19	362	75	15	9	19	15	33	314	M = 14.55 0 = 4.7

E. L'exploitation des données du listing

• Les élèves 3 (VOS) et 9 (BAR) répondent presque toujours avec la certitude 3, ... mais se trompent souvent et ont, par conséquent, les cotes les plus basses de la série. La procédure de cotation défavorise donc ceux qui sont toujours sûrs d'eux-mêmes alors qu'ils se trompent.

(10) La matière du test portait sur les conventions de langage dans les transmissions aériennes.

Dans un atelier, un hôpital, une administration, les élèves 3 (VOS) et 9 (BAR) seraient de véritables catastrophes. La procédure de cotation les pénalise (voir leur cote).

• Les élèves 2 (BUL) et 16 (VER) ont omis plus de réponses que les élèves 1 (TOM) et 14 (ENN), et ont pourtant des totaux supérieurs. La procédure de cotation ne défavorise pas les élèves qui omettent (voir leur cote).

Ce type de procédure peut amener le professeur à une analyse plus fine de la personnalité et des attitudes de l'élève vis-à-vis de la connaissance. Voici deux attitudes peu souhaitables et les remèdes à y apporter :

1. *L'élève beaucoup trop sûr de lui.* Il y a lieu de l'entrainer à mieux estimer l'exactitude de sa réponse et de lui faire prendre conscience des conséquences des erreurs de jugement dans la vie.

2. *L'élève très peu sûr de lui,* alors qu'il connait bien la matière. Il y a lieu de l'entrainer à mieux estimer l'exactitude de sa réponse et de l'aider à prendre confiance en lui.

Dans tous les cas, le professeur agit en conseiller ; il intervient auprès de l'élève pour lui faire prendre conscience de la situation, le motiver, augmenter ou diminuer la confiance en soi.

D. LECLERCQ.

Bibliographie

La bibliographie complète paraîtra avec la seconde partie de cet article.

BANQUE DE QUESTIONS ET INDICES DE CERTITUDE (II)

G. Maximum et minimum

Dans l'enseignement secondaire, l'évaluation a un rôle essentiellement formatif [1] par l'information en retour (*feedback*) qu'elle fournit aux étudiants et aux enseignants, en ce qui concerne la maîtrise d'un contenu (*mastery learning*). Dans la suite de cet article, nous n'envisagerons que les situations où tous les aspects d'un apprentissage doivent être vérifiés chez tous les élèves. Par conséquent, les étudiants n'ont pas la possibilité de répondre aux questions de leur convenance, comme cela peut être le cas dans une perspective d'évaluation d'aptitude.

Un second problème soulevé est celui du maximum des points. Il est courant d'exiger 75 % des points pour considérer qu'il y a réussite pédagogique. Cette pratique, qui remonte à la mise au point des premières épreuves d'intelligence (BINET), est proposée par plusieurs auteurs (par ex. : G. GOOSENS et S. ROLLER) [2]. Dans les exemples qui suivent, nous seront également amenés à fixer *arbitrairement* un « niveau de réussite pédagogique » qui est, évidemment, inférieur au maximum des points. Tant que ce niveau pédagogique n'est pas atteint, l'élève poursuivra son apprentissage.

a) Maximum

Dans une épreuve classique, on demande à l'élève de connaître les réponses, sans plus. Ici, on lui demande, en outre, d'être certain de sa réponse, ce qui est plus exigeant. C'est la raison pour laquelle on considère qu'un étudiant qui répond correctement à toutes les questions avec la certitude 2 (ou certitude normale) doit obtenir le *niveau de réussite pédagogique* (qui se traduira par le maximum des points *au bulletin*).

Niveau de réussite pédagogique = $(\Sigma_1^n \text{ importances}_i) \times 2$.

Dans nos exemples, le niveau de réussite pédagogique est toujours ramené à 20 points (par règle de trois), et le maximum à 30 points.

Par conséquent, un étudiant peut obtenir une cote plus élevée que le niveau de réussite pédagogique s'il donne des bonnes réponses avec l'assurance 3. Bien entendu, on n'écrira pas 24/20 dans le bulletin de l'élève, mais simplement 20/20. Cependant, le professeur ne manquera pas

(1) Voir G. DE LANDSHEERE, 1971 b, p. 15.

(2) G. GOOSENS et S. ROLLER, *Recherches expérimentales sur la connaissance de l'accord du participe passé à Genève et à Bruxelles*, Neuchâtel et Paris, Delachaux et Niestlé, 1954.

de s'intéresser aux cotes qui dépassent le niveau de réussite pédagogique [3].

Pourquoi ne ménagerait-on pas la possibilité d'atteindre le niveau de réussite pédagogique, donc d'obtenir 20/20 au bulletin, de diverses façons ? Cette procédure pourrait diminuer le stress chez les élèves puisqu'une erreur ou une omission n'entraînent plus la perte irrévocable du maximum (20/20) au bulletin.

Voici, à titre d'exemple, les résultats d'un test de 25 questions d'électronique, posé à neuf élèves, le 29 septembre 1971 :

	Score réel	Bulletin	—3	—2	—1	0	+1	+2	+3
1. DO...	27.33	20	0	0	0	2	0	0	23
2. LE...	20.44	20	0	0	0	0	4	11	9
3. NE...	28.00	20	0	0	1	0	0	6	19
4. CA...	21.22	20	0	1	0	1	0	14	9
5. HA...	26.00	20	0	0	0	1	0	10	14
6. LO...	21.56	20	1	0	0	2	1	6	15
7. MO...	27.11	20	0	0	0	0	0	10	15
8. VA...	22.22	20	0	0	0	1	0	17	7
9. WI...	26.44	20	0	1	0	0	0	4	20
Moyenne	24.47	**Total :**	1	2	1	7	5	78	131
Niveau de réussite pédagogique	20.00								
Maximum	30.00								

Pourquoi serait-il aberrant qu'après 82 heures de cours massés sur deux mois, ces neuf élèves adultes atteignent tous le niveau de réussite pédagogique (20/20 au bulletin) ? Si les questions sont représentatives des objectifs et si ceux-ci sont adéquats, le professeur peut être satisfait de son enseignement et les élèves de leur apprentissage. On pourrait, bien entendu, craindre une trop grande facilité des questions. Le dossier de chaque question doit permettre de vérifier cette hypothèse.

Les étudiants sont docimologiquement discriminés entre eux, mais la distinction ne subsiste plus au point de vue social (bulletin).

b) Minimum

L'apparition de scores totaux négatifs n'est pas rare dans les épreuves, hélas ! Dans un grand nombre de cas, l'élève s'est trompé à un point tel qu'il a obtenu moins de points que celui qui aurait omis sans cesse (total des points : 0) On peut caricaturer la situation en disant qu'il « en sait moins que quelqu'un qui ne sait rien ».

Sur le bulletin, les scores négatifs sont ramenés à 0.

(3) Des cotes négatives ou supérieures à 20/20 ne doivent pas étonner car l'échelle utilisée est une échelle d'intervalles égaux, mais pas encore de quotient, dans ce sens qu'elle n'a pas de zéro absolu (un autre exemple en est l'échelle des températures où le zéro varie d'une échelle à l'autre).

H. Les omissions

A un certain moment de l'expérience, les omissions ont considérablement augmenté. Or, un principe largement répandu dans le système scolaire est qu'un élève ne devrait pas pouvoir obtenir plus de la moitié des points (4) en répondant à moins de la moitié des questions.

Pour donner satisfaction aux utilisateurs, chaque professeur peut fixer un taux d'omissions autorisées (OA) pour chaque test. Ainsi, si on autorise 20 % d'omissions à un test de vingt questions, aucune pénalisation n'interviendra pour quatre omissions au moins. Par contre, pour cinq omissions ou plus, une pénalisation sera appliquée. Le commandant P. VAN ROY, qui a proposé cette technique, a calculé que la pénalisation devrait être de — 0,35 point pour que se réalise le principe : « Il faut avoir répondu à la moitié des questions pour obtenir la moitié des points. »

Par exemple, à un test de 25 questions, pour lequel on autorise 40 % d'omissions (soit 10 omissions permises), un élève omet 12 questions (deux omissions seront donc pénalisées) (OP). On calcule alors la moyenne d'importance des douze questions omises et on effectue le calcul :

Pénalisation : — 0,35 × nombre d'OP × Moyenne des importances des 12 questions.

Ici : — 0,35 × 2 × 2,7.

Cette pénalisation a, évidemment, pour effet de ramener le taux des omissions en-deçà du seuil autorisé, chez la grande majorité des étudiants. Cette pratique ne concorde cependant pas avec la théorie des décisions.

I. Les points au bulletin

Dans tous les barèmes classiques de cotation, on ne peut obtenir ni plus du maximum, ni des scores négatifs ; de plus, certaines grandeurs ont une signification.

- 2/20 est toujours « très mauvais », qu'il s'agisse d'enseignement primaire, secondaire ou universitaire.
- En dessous de 8/20, il y a toujours « échec ».
- Le seuil d'acceptabilité d'un score varie entre 8/20 et 12/20, avec une moyenne de 10/20.
- Au-delà de 12/20, les scores deviennent bons.

Pour gagner sur d'autres terrains, plus importants à nos yeux, nous n'avons pas tenté, dans l'immédiat, de faire disparaître cette échelle traditionnelle. Pour rendre les résultats conformes aux habitudes, on disposait de deux possibilités : une normalisation ou une série de fonctions plus empiriques de transformation.

(4) Dans le cas où toutes les questions ont le même poids.

a) La normalisation

Ce premier moyen consiste à calculer la moyenne et le sigma des cotes obtenues, puis à transformer ces cotes en notes Z afin de les exprimer, par exemple, dans une échelle à moyenne = 12 et à sigma = 3.

Cette façon de faire a été rejetée, car elle est uniquement centrée sur les notes Z, c'est-à-dire la place de l'individu dans son groupe. Aucune référence n'est faite à la matière, et on assiste à des phénomènes regrettables, comme le montre l'exemple ci-dessous.

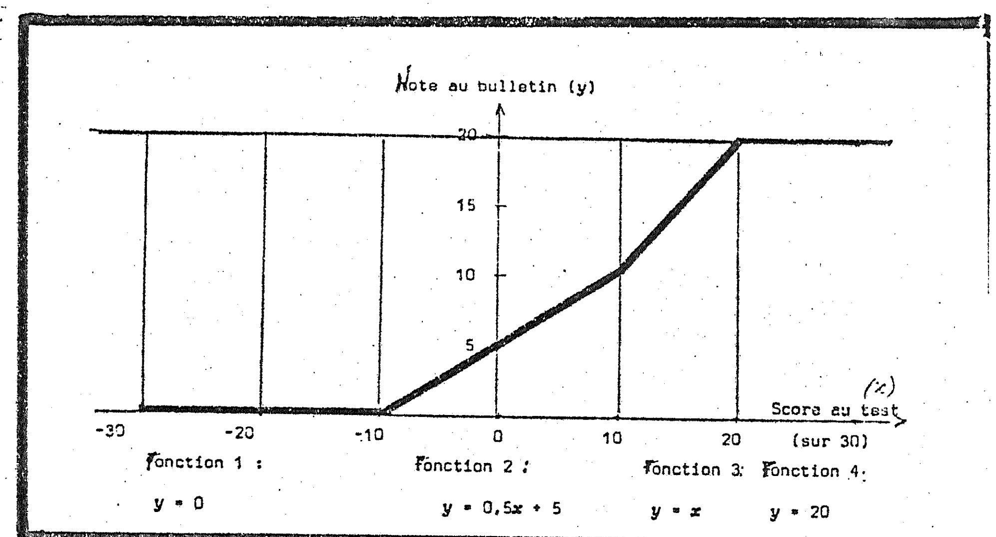
Une classe obtient, à un test de physique, une moyenne de 17/20 et un sigma de 1, ce qui se traduit sur le bulletin par une moyenne de 12/20 et un sigma de 3. On voit déjà le mensonge introduit par la normalisation, mensonge particulièrement préjudiciable aux élèves qui avaient obtenu 14/20 au test, qui vont obtenir trois points au bulletin.

Mais, si la classe obtient, à un test de chimie, une moyenne de 3/20 et un sigma de 1, les notes au bulletin auront une moyenne de 12/20 avec un sigma de 3, si bien que certains élèves pourront avoir 20/20 au bulletin, alors qu'aucun n'a eu plus de 6/20 lors du test.

Les lecteurs du bulletin seront là aussi abusés. Mais l'abus le plus flagrant est la comparaison des notes de physique et de chimie, semblables dans le bulletin et tellement dissemblables en réalité.

b) Les fonctions de transformation

La note Z suffit à situer l'enfant dans le groupe. Pour la note au bulletin, nous avons — très empiriquement, et d'ailleurs provisoirement — opté pour les quatre fonctions de transformation ci-dessous.



En abscisse sont portés les scores réels (minimum = — 30 ; maximum = + 30). En ordonnée, les notes au bulletin sont comprises entre 0 (nivellement par le bas des scores négatifs) et 20 points (écrétage par le haut des scores supérieurs à 20).

On voit qu'entre — 30 et — 10, les scores sont ramenés à 0 point au bulletin. Entre — 10 et + 10, ils sont ramenés à une note au bulletin comprise entre 0 et 10 points. Entre 10 et 20, les scores sont traduits tels quels au bulletin. Enfin, au-delà de 20, les scores sont ramenés à 20 points au bulletin.

Cette procédure, beaucoup plus souple que la précédente, permet, à tout moment, de modifier une des quatre fonctions indépendamment des autres, ce qui ne serait pas possible si, par perfectionnisme, on essayait de représenter ces fonctions par des courbes (équations non linéaires).

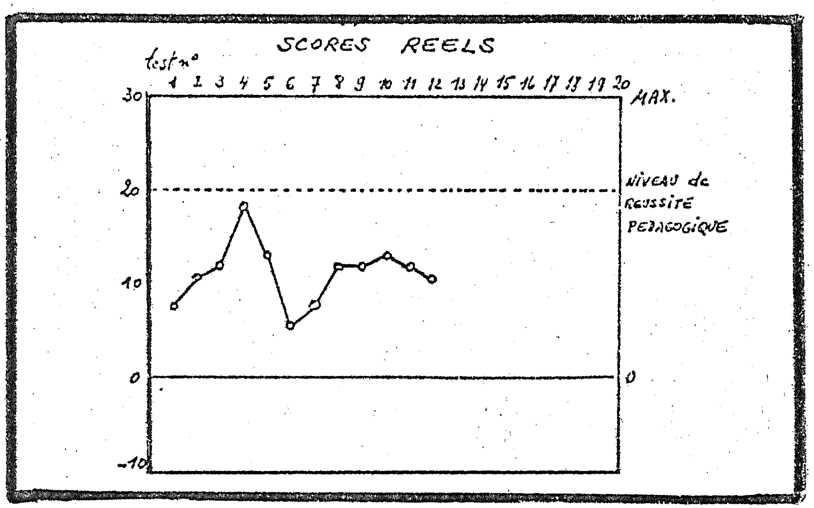
J. L'évolution des étudiants

Plusieurs tests ont été subis par les élèves au cours d'une année scolaire. Il était donc possible de noter l'évolution de chaque étudiant pour chaque discipline.

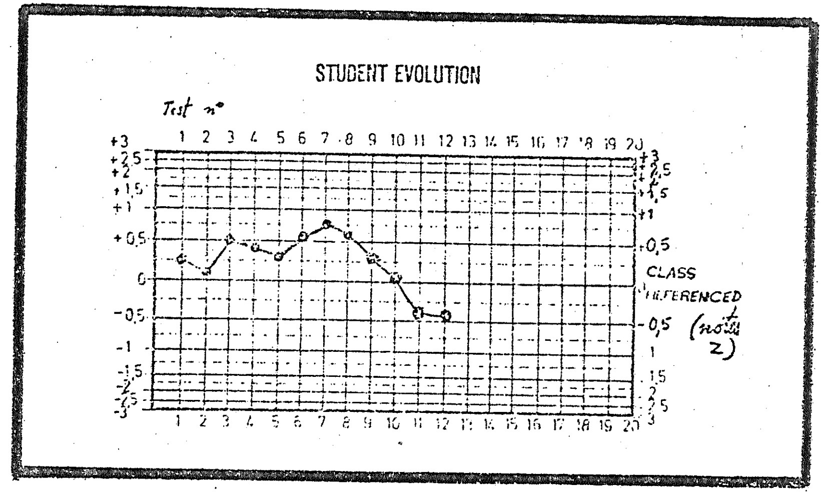
Afin d'éviter définitivement l'ambiguïté dénoncée ci-dessus, deux graphiques ont été conçus : le premier pour les scores réels (et non pour les scores du bulletin, qui ne sont qu'une traduction adaptée), le second pour les notes Z. Ces deux graphiques peuvent être mis à jour par l'élève lui-même.

Voici un exemple de ces deux graphiques :

A. **Evolution par rapport à la matière (scores réels)**



B. **Evolution par rapport à la classe (notes Z)**



L'intérêt de ce type de donnée pour l'évaluation et la guidance continues, ainsi que pour les délibérations finales, est trop évident pour que nous nous y attardions.

K. La courbe des résultats

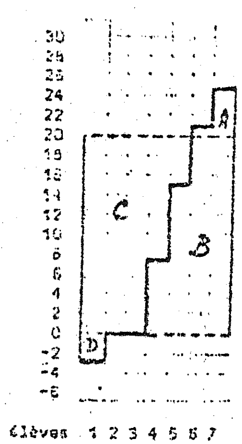
Dans une publication sur l'enseignement programmé (LECLERCQ et al., 1972), nous avons montré l'intérêt qu'il y a à présenter les résultats sous forme de courbe, à partir du vecteur ordonnée des cotes. Dans ces graphiques, on porte en abscisse les élèves et en ordonnées les scores.

Imaginons une classe de sept élèves où les cotes (ordonnées) seraient : — 2,5 - 0 - 0 - 7,5 - 15 - 21 - 24,5, ce qui donne la ligne brisée suivante :

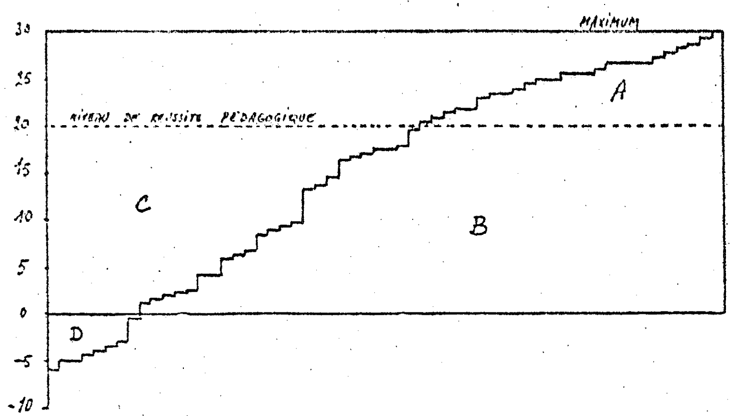
Considérons les quatre surfaces ainsi délimitées :

— La surface B représente ce qui est connu.
— La surface A représente les connaissances supérieures au maximum souhaité.
— La surface C montre ce qui n'est pas connu.
— La surface D reflète un état de connaissances plus déplorable encore que l'ignorance.

Pour l'exemple proposé, l'importance de la surface C montre que pour beaucoup d'élèves, les objectifs de l'enseignement ne sont pas atteints.



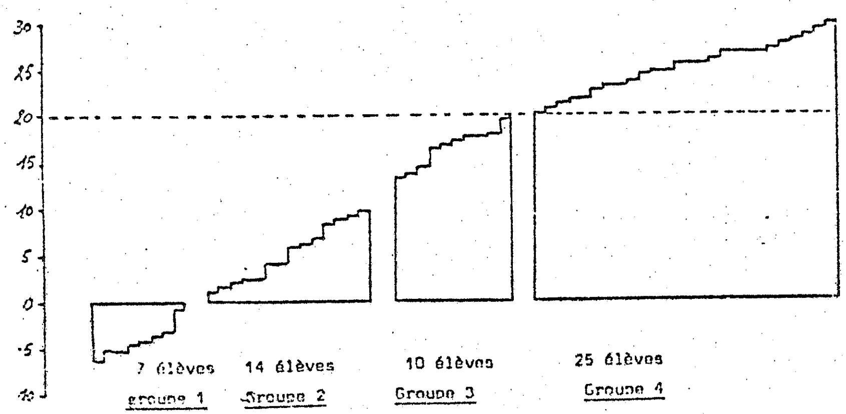
Voici la ligne brisée obtenue quand on ordonne les cotes sur 20 de 57 élèves qui ont subi une épreuve de géographie le 4 novembre 1971 (5).



Discussion

En examinant ce graphique et en connaissant les intentions du professeur, on peut se féliciter de l'importance de la zone A et déplorer la présence d'une zone D. Cette dernière ne devrait jamais exister dans une situation pédagogique. En effet, un élève totalement ignorant, mais sage, refuserait de répondre et obtiendrait 0, mais pas une cote négative.

Pour clarifier la situation, nous proposons de découper le graphique ci-dessus en quatre graphiques :



(5) Vu le nombre élevé de sujets, nous avons réduit nos dimensions de plus de moitié sur l'abscisse.

Pour la matière considérée, le groupe 4 (près de la moitié de la classe) n'a plus besoin de l'intervention du professeur. Quant aux trois autres groupes, leurs besoins sont divers : le groupe 3 est le groupe des élèves qui ont obtenu la moitié des points, mais qu'un traitement approprié devrait amener au maximum ; le groupe 2, faible, nécessite presque un réapprentissage ; quant au groupe 1, il faut non seulement le traiter comme le groupe 2, mais, en outre, modifier ses attitudes vis-à-vis de la situation de contrôle et de l'auto-évaluation des connaissances.

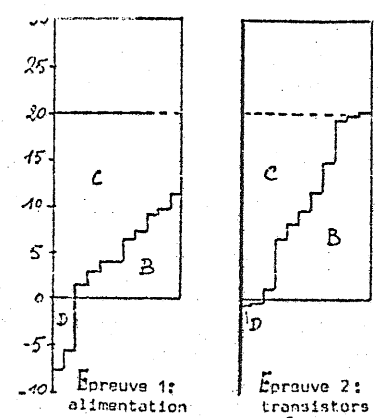
L. Le nombre de questions d'un test

L'enseignant qui a bien déterminé ses objectifs, qui sait lesquels il veut contrôler et qui dispose d'un répertoire suffisant de questions, est en mesure de construire des épreuves homogènes et très valides.

Dans cette perspective, il est souhaitable de présenter à l'élève n tests homogènes très courts, plutôt qu'un test hétérogène long. Cette position semble aller à l'encontre des théories classiques de construction de tests, selon lesquelles l'augmentation des items entraîne une augmentation de la fidélité. Mais n'oublions pas que la fidélité est subordonnée, en pédagogie, à la validité qui, elle, doit être maximisée par d'autres moyens (tests critères, analyse des objectifs, ...).

Voici un exemple illustrant ce problème.

Deux épreuves d'électronique (une sur les transistors et l'autre sur l'alimentation et la stabilisation, comptant chacune 25 questions) ont été appliquées le même jour à onze élèves. Il serait tentant, pour satisfaire à certaines exigences statistiques (Guilford, 1956, p. 458), de regrouper ces questions en un seul test de 50 questions. Et pourtant, quelle perte d'informations cela eût été, quand on examine les graphiques !



En regroupant ces deux épreuves, on n'aurait pas vu que les rapports entre les surfaces C et B sont bien différents et que la zone D n'existe pratiquement que dans l'épreuve 1.

CONCLUSIONS

Dans les pages qui précèdent, nous n'avons décrit que les options docimologiques qui nous paraissaient les moins répandues dans l'enseignement secondaire actuel. Nous avons voulu montrer les possibilités de ces options et non les présenter comme les seules possibles. Il est bien évident que le système compte bien d'autres principes, tels que « le feed-back aux élèves et au professeur doit se faire le plus rapidement possible », etc. Ces derniers principes sont suffisamment acceptés comme allant de soi pour que nous les discutions ici.

On aura remarqué combien les divers principes s'interpénètrent et comme leur combinaison constitue un système cohérent. Ce système est le résultat d'ajustements successifs et de compromis entre les enseignants, les directeurs et les chercheurs. Pour cette raison, le système est en constante évolution et diffère d'une école à l'autre. La participation des enseignants eux-mêmes à l'élaboration du système a été très fructueuse et encourageante. Nous pensons avoir réalisé un vœu exprimé par G. De Landsheere (1971) dans son ouvrage sur la docimologie (p. 195) :

> « Un vaste effort d'information sera aussi entrepris auprès des enseignants appelés à changer profondément leurs habitudes d'évaluation. Le moyen le plus efficace semble résider dans l'organisation d'expériences locales portant sur la préparation des examens, la notation et la modération... Progressivement, toutes les écoles y participeraient. »

D. LECLERCQ.

Bibliographie

COOMBS, C. H., MILHOLLAND, J. E., WOMER, F. B., The Assessment of Partial Knowledge, *Educational and Psychological Measurement*, 16, 1956, 13-37.

COOMBS, P. H., *La crise mondiale de l'éducation*, Paris, P.U.F., 1968.

DE FINETTI, B., Methods for discriminating levels of partial knowledge concerning a test item, in *British Journal of Mathematical and Statistical Psychology*, 18, 123, 1965, p. 87.

DE LANDSHEERE, G., *Introduction à la recherche en éducation*, Paris, A. Colin; Liège, G. Thone, 1971 a, 3e éd.

DE LANDSHEERE, G., *Evaluation continue et examens, Précis de docimologie*, Paris, Nathan; Bruxelles, Labor, 1971 b, 2e éd.

DETHEUX, M. et LECLERCQ, E., Essai d'évaluation continue en éducation compensatoire par contrôles successifs du rendement dans l'apprentissage de la lecture en première année primaire, *Actes du Colloque sur la compensation des handicaps socio-culturels*, Esneux, octobre 1972.

GUILFORD, J. P., *Fundamental Statistics in Psychology and Education*, New York, Mc Graw Hill, 1956.

LECLERCQ, D., VAN ROY, P., MOONEN, P., Test objectifs (3 fascicules), Saint-Trond, Ecole Technique de la Force Aérienne, 1972.

LECLERCQ, D., Une banque de questions pour l'enseignement, in *Education*, n° 132, nov.-déc. 1971.

LECLERCQ, D., Critique des méthodes d'application de correction et de cotation des questions à choix multiple, in *Scientic Paedagogica Experimentalis*, 1973, X.

LECLERCQ, D., L'utilisation de la fiabilité subjective dans l'évaluation pédagogique. Thèse de doctorat. A paraître en 1975.

PAQUAY, J., Essai d'évaluation continue en éducation compensatoire, par contrôles successifs du rendement de l'apprentissage de la mathématique en première année primaire, *Actes du Colloque sur la compensation des handicaps d'origine socio-culturelle*, Esneux, octobre 1972.

VAN DE VELDE, L., *Adaptation des questions du développement d'activités intellectuelles définies*, Bruxelles, janvier 1972, document ronéotypé.

EDUCATION

tribune libre

N° 150

Revue bimestrielle

ANNEE 1975

Annexe 3 (Cf chap. 3, p. 165)

Pour arriver à (4) $\frac{p_i}{1-p_i}(x_j - x_i) < y_i - y_j < \frac{p_j}{1-p_j}(x_j - x_i)$,

on part de (1) : $ESU = p_i x_i + (1-p_i) y_i$

et de (2) : $p_i x_j + (1-p_i) y_j < p_i x_i + (1-p_i) y_i$

On divise (2) par $1-p_i$ et on a (2') : $\frac{p_i x_j}{1-p_i} + y_j < \frac{p_i x_i}{1-p_i} + y_i$

On retire y_j à chaque membre et on a (2'') :

$$(2'') : \frac{p_i x_j}{1-p_i} < \frac{p_i x_i}{1-p_i} + y_i - y_j$$

On retire $\frac{p_i x_i}{1-p_i}$ de chaque membre et on a (2''') :

$$(2''') : \frac{p_i x_j}{1-p_i} - \frac{p_i x_i}{1-p_i} < y_i - y_j$$

Dans le membre de gauche, on met en évidence $\frac{p_i}{1-p_i}$, et on a (2'''')

$$(2'''') : \frac{p_i}{1-p_i}(x_j - x_i) < y_i - x_j$$

De même, (3) : $p_j x_i + (1-p_j) y_i < p_j x_j + (1-p_j) y_j$

$$(3') : \frac{p_j x_i}{1-p_j} + y_i < \frac{p_j x_j}{1-p_j} + y_j$$

$$(3'') : y_i < \frac{p_j x_j}{1-p_j} + y_j - \frac{p_j x_i}{1-p_j}$$

$$(3''') : y_i - y_j < \frac{p_j x_j}{1-p_j} - \frac{p_j x_i}{1-p_j}$$

$$(3'''') : y_i - y_j < \frac{p_j}{1-p_j}(x_j - x_i)$$

Annexe 4 (Cf Chap. 2)

1

TEST CODE	-3	-2	-1	0	+1	+2	+3
1602	2	0	1	2	0	8	10
1603	1	2	0	12	0	1	7
1604	4	3	0	0	0	6	10
1606	2	0	2	5	1	1	12
1607	2	4	3	4	0	1	9
1608	1	1	3	6	0	7	5
1610	6	1	1	0	1	3	11
1611	1	5	2	5	3	1	6
1612	1	5	0	3	0	4	10
1613	4	0	0	2	1	11	5
1701	0	3	6	4	2	2	6
1702	3	3	3	1	2	3	8
1703	3	6	3	0	0	4	7
1704	4	5	6	0	1	1	6
1705	2	2	7	0	4	4	4
1707	0	0	6	0	7	5	5
1708	2	5	3	0	2	5	6
1709	1	8	1	0	0	6	7
1710	2	2	5	0	2	2	10
1711	2	2	6	0	3	4	6
1712	2	8	2	0	3	1	7
1713	0	8	8	0	1	2	4
1714	2	2	8	0	2	5	4
1715	0	0	0	0	0	0	0
1801	1	1	0	8	0	0	13
1802	1	1	1	0	5	6	9
1803	2	0	0	2	0	5	14
1804	0	0	1	6	0	3	13
1805	1	2	1	3	1	5	10
1807	0	4	0	6	2	6	5
1810	0	1	0	2	0	7	13
1812	1	1	1	5	2	1	12
1813	1	3	3	3	1	2	10
1814	0	1	3	2	1	8	8
1815	0	0	0	9	2	2	10
1816	0	0	0	0	0	0	23
1817	3	3	1	2	0	1	13
1818	0	0	0	0	0	0	0
1901	1	1	2	2	4	4	9
1902	2	0	0	5	0	0	16
1903	3	0	0	5	1	2	12
1904	0	0	0	0	0	0	0
1905	2	1	0	5	1	1	13
1906	3	0	0	3	4	2	11
1908	1	0	1	3	2	4	12
1909	1	0	0	6	1	2	13
1910	1	0	1	9	2	0	10
1911	3	2	0	4	0	0	14
1913	1	0	2	9	0	3	8
1914	1	3	3	3	2	4	7
1915	4	1	0	0	0	0	18
1916	0	0	0	5	3	2	13
1918	4	0	1	6	1	3	8

84100 97157 70160482

2

TEST CODE	-3	-2	-1	0	+1	+2	+3
1602	4	6	0	0	2	6	6
1603	7	2	0	3	0	0	12
1604	4	5	0	0	0	4	11
1606	2	6	0	0	0	2	14
1607	4	3	1	0	0	3	13
1608	0	3	0	5	0	11	5
1610	3	5	0	1	0	5	10
1611	0	5	4	0	4	5	6
1612	6	5	0	0	0	1	12
1613	0	0	0	0	0	0	0
1701	3	1	6	0	2	2	10
1702	4	2	6	0	0	2	10
1703	4	6	3	0	0	4	7
1704	1	5	3	3	1	5	6
1705	0	2	10	0	2	5	5
1707	1	4	4	4	2	3	6
1708	0	0	12	0	7	2	3
1709	3	4	4	0	3	2	8
1710	4	5	3	0	0	1	11
1711	2	4	6	0	3	2	7
1712	2	7	3	0	0	9	3
1713	6	5	3	0	3	2	5
1714	1	6	1	1	1	8	6
1715	2	4	5	5	2	0	6
1801	3	3	0	9	0	0	9
1802	3	2	3	0	0	3	13
1803	7	4	0	3	1	2	7
1804	4	6	1	6	0	1	6
1805	3	5	0	6	0	3	7
1807	0	5	1	9	1	7	1
1810	4	5	0	0	0	2	13
1812	1	6	3	5	2	3	4
1813	3	3	2	3	1	3	9
1814	3	2	3	3	3	7	5
1815	2	0	0	10	2	0	10
1816	4	0	0	0	0	0	20
1817	1	5	0	9	2	4	3
1818	0	0	0	0	0	0	0
1901	2	4	1	6	0	4	7
1902	2	1	1	9	0	0	11
1903	0	0	0	0	0	0	0
1904	6	1	0	5	0	0	12
1905	4	2	4	0	0	1	13
1906	2	2	3	6	2	3	6
1908	5	4	0	6	0	1	8
1909	3	2	3	7	0	1	8
1910	2	1	1	7	2	1	10
1911	2	1	2	10	0	1	8
1913	2	2	1	7	1	1	10
1914	1	4	1	8	4	3	3
1915	9	1	0	2	0	0	12
1916	2	0	0	11	3	1	7
1918	5	1	4	4	1	1	8

148167108173 57137412

3

TEST CODE	-3	-2	-1	0	+1	+2	+3
1602	8	0	0	0	0	3	14
1603	5	2	0	3	0	2	13
1604	4	5	0	0	0	1	15
1606	3	2	0	0	0	4	16
1607	6	3	2	0	0	1	13
1608	0	4	1	0	1	14	5
1610	3	2	0	0	1	1	18
1611	1	2	3	4	3	4	8
1612	3	3	0	1	0	4	14
1613	5	3	1	0	0	7	9
1701	6	3	2	0	1	1	12
1702	5	2	3	0	0	4	11
1703	6	6	2	0	1	4	6
1704	4	0	9	4	1	1	6
1705	1	4	4	1	6	4	5
1707	1	0	7	5	2	4	6
1708	1	6	4	0	0	8	6
1709	7	2	2	0	0	1	13
1710	2	8	4	2	2	3	4
1711	1	2	4	2	2	1	13
1712	7	3	0	0	0	5	10
1713	3	5	2	5	0	4	6
1714	0	3	4	4	0	9	5
1715	4	7	1	5	1	2	5
1801	6	0	0	7	1	0	11
1802	4	1	1	0	1	4	14
1803	5	2	1	2	1	3	11
1804	2	4	0	5	2	2	10
1805	3	4	0	3	0	3	12
1807	0	0	0	0	0	0	0
1810	3	0	0	1	6	2	13
1812	1	2	1	7	1	10	3
1813	4	4	3	4	0	4	6
1814	2	0	1	2	2	5	13
1815	3	0	0	7	0	0	15
1816	0	0	0	0	0	0	25
1817	4	6	0	5	1	3	6
1818	0	0	0	0	0	0	0
1901	1	4	3	4	2	4	7
1902	3	0	0	10	0	0	12
1903	0	0	0	0	0	0	0
1904	2	0	0	1	0	0	22
1905	5	1	0	5	0	0	14
1906	2	0	1	12	1	0	9
1908	4	0	2	6	2	3	8
1909	4	1	1	1	0	4	14
1910	1	0	0	10	0	0	14
1911	6	3	0	5	0	0	11
1913	0	2	2	1	6	1	13
1914	1	3	2	9	0	4	6
1915	2	2	0	1	0	2	18
1916	1	0	3	6	6	0	9
1918	5	0	3	6	0	0	11

160116 79156 53146540

4

TEST CODE	-3	-2	-1	0	+1	+2	+3
1602	6	3	0	0	0	2	14
1603	3	0	0	5	0	0	17
1604	6	2	1	0	0	1	15
1606	6	0	0	0	0	0	19
1607	5	2	1	0	0	3	14
1608	5	0	0	1	0	0	19
1610	5	0	0	1	0	0	19
1611	6	1	0	5	0	3	10
1612	5	0	0	3	0	0	17
1613	3	1	0	1	0	1	19
1701	6	2	0	0	0	0	17
1702	4	2	1	0	0	1	17
1703	4	4	0	0	1	8	8
1704	3	1	7	3	5	4	2
1705	3	1	5	4	2	0	10
1707	1	5	0	5	1	4	9
1708	1	1	4	1	2	11	5
1709	5	1	1	0	1	3	14
1710	1	4	0	5	1	2	12
1711	4	0	1	2	2	2	14
1712	1	3	1	0	4	8	8
1713	7	0	0	3	0	0	15
1714	1	2	3	4	3	6	6
1715	6	2	0	3	0	4	10
1801	5	0	2	10	1	0	7
1802	4	0	0	4	1	0	15
1803	4	1	4	5	3	1	7
1804	4	4	2	4	0	6	5
1805	7	4	0	6	0	3	5
1807	0	5	3	5	2	6	4
1810	0	0	3	1	1	2	18
1812	3	2	1	8	0	5	6
1813	4	8	5	0	2	0	6
1814	0	2	4	4	1	4	10
1815	1	0	0	10	1	0	13
1816	3	0	0	2	0	0	20
1817	2	3	0	11	0	3	6
1818	0	0	0	0	0	0	0
1901	3	2	4	8	0	4	4
1902	3	0	0	10	0	0	12
1903	0	0	0	0	0	0	0
1904	0	0	0	0	0	0	0
1905	4	1	0	4	0	2	14
1906	4	0	1	10	0	0	10
1908	3	2	2	8	2	0	8
1909	3	1	0	12	0	2	7
1910	4	0	0	12	0	0	9
1911	2	0	0	7	0	0	16
1913	2	0	2	2	5	1	13
1914	3	3	0	9	0	9	1
1915	2	0	0	1	0	1	21
1916	1	1	1	11	0	1	10
1918	4	2	0	9	0	0	10

172 78 59219 41113568

5.

TEST CODE	-3	-2	-1	0	+1	+2	+3
1602	20	0	0	0	0	0	30
1603	4	5	0	10	0	5	26
1604	14	1	7	0	2	0	26
1606	10	0	0	4	0	0	36
1607	10	0	0	14	0	1	25
1608	2	7	3	12	0	22	4
1610	7	0	0	14	0	0	29
1611	3	4	2	17	7	1	16
1612	9	0	0	4	0	0	37
1613	9	1	0	6	0	2	32
1701	3	8	5	9	2	4	19
1702	3	1	0	14	0	2	29
1703	14	2	0	9	0	4	21
1704	7	6	3	14	4	7	9
1705	4	3	5	14	5	6	13
1707	2	4	3	6	3	11	21
1708	7	3	10	9	3	4	9
1709	5	3	1	14	1	6	20
1710	6	2	4	7	5	6	20
1711	4	2	1	18	1	2	22
1712	3	6	4	5	2	10	20
1713	8	0	0	17	0	0	25
1714	0	8	0	9	0	33	0
1715	5	3	1	19	2	3	17
1801	17	0	0	8	0	0	24
1802	6	2	5	0	2	2	32
1803	5	5	5	3	3	8	20
1804	2	6	7	7	5	8	14
1805	8	5	0	7	1	2	26
1807	0	6	6	8	5	13	11
1810	3	0	1	1	5	0	39
1812	4	1	2	12	7	6	17
1813	4	5	6	3	2	9	20
1814	3	4	2	3	4	9	24
1815	8	1	0	10	0	3	27
1816	6	0	0	0	0	0	43
1817	3	3	2	15	3	5	18
1818	0	0	0	0	0	0	0
1901	11	0	0	9	0	1	28
1902	6	1	0	12	0	1	29
1903	3	0	0	12	0	0	34
1904	3	0	0	5	0	0	41
1905	3	0	0	15	0	0	31
1906	6	0	0	10	0	2	31
1908	7	4	1	17	3	6	11
1909	4	6	0	3	0	7	27
1910	1	0	3	14	2	1	28
1911	9	0	2	10	0	1	27
1913	9	0	3	12	2	0	23
1914	2	8	1	8	2	16	13
1915	2	1	0	1	0	3	42
1916	4	1	5	11	4	6	18
1918	7	0	3	11	1	2	25

305128103472 93240***

TEST CODE	6 -3	-2	-1	0	+1	+2	+3
1602	10	0	0	1	0	0	13
1603	10	1	1	1	0	0	11
1604	6	2	1	0	0	1	14
1606	8	0	0	0	0	0	16
1607	7	0	0	3	0	0	14
1608	2	1	0	2	0	4	15
1610	8	0	0	4	0	0	12
1611	5	2	0	2	0	2	13
1612	6	0	0	1	0	0	17
1613	7	0	0	0	0	0	17
1701	6	2	1	5	0	1	9
1702	2	0	1	11	0	2	8
1703	1	4	1	4	2	4	8
1704	0	4	6	7	0	1	6
1705	3	0	7	6	2	1	5
1707	3	3	5	4	0	3	6
1708	0	0	6	7	9	0	2
1709	1	3	3	7	0	1	9
1710	4	2	4	1	2	2	9
1711	9	0	0	7	0	0	8
1712	0	5	3	3	1	4	8
1713	4	0	0	8	0	0	12
1714	1	5	0	6	1	3	8
1715	4	0	0	9	0	2	9
1801	10	0	0	6	0	0	8
1802	4	2	3	0	1	1	13
1803	5	3	2	3	0	4	7
1804	1	2	5	2	3	2	9
1805	4	4	0	3	1	3	9
1807	0	5	2	7	1	2	7
1810	3	0	1	1	0	0	19
1812	0	0	0	0	0	0	0
1813	5	2	1	4	2	2	8
1814	0	1	6	3	3	4	7
1815	9	2	0	3	0	1	9
1816	3	0	0	1	0	0	20
1817	7	4	0	4	0	1	8
1818	1	3	3	6	0	2	9
1901	1	2	1	4	0	4	12
1902	4	1	0	4	0	0	15
1903	5	0	0	7	0	0	12
1904	0	0	0	0	0	0	0
1905	5	0	0	5	0	0	14
1906	5	0	0	7	0	0	12
1908	6	0	0	9	0	0	9
1909	0	0	0	0	0	0	0
1910	3	0	0	9	0	0	12
1911	4	2	0	6	0	1	11
1913	4	1	6	0	2	1	10
1914	0	3	1	8	1	6	5
1915	6	0	0	0	1	1	16
1916	1	0	3	10	3	1	6
1918	10	0	0	5	2	0	7

213 71 73216 37 67523

TEST CODE	7 -3	-2	-1	0	+1	+2	+3
1602	3	1	3	0	2	1	13
1603	3	2	0	3	0	4	11
1604	3	1	1	0	0	3	15
1606	4	0	0	0	0	0	19
1607	7	1	0	1	0	0	14
1608	4	4	0	3	0	5	7
1610	7	0	0	1	0	1	14
1611	7	0	1	4	0	0	11
1612	3	0	0	1	0	0	19
1613	8	0	0	2	0	0	13
1701	4	6	0	3	0	4	6
1702	3	3	0	6	0	2	9
1703	5	4	0	2	0	2	10
1704	6	0	0	10	0	0	7
1705	1	8	5	6	0	3	0
1707	3	3	0	8	0	4	5
1708	3	4	3	1	2	1	9
1709	2	3	0	8	1	1	8
1710	3	3	1	4	3	0	9
1711	4	0	4	5	1	1	8
1712	1	7	2	1	1	3	8
1713	8	0	0	8	0	0	7
1714	1	3	1	7	2	6	3
1715	4	3	0	9	0	1	6
1801	3	2	1	7	2	0	8
1802	3	0	2	0	1	2	15
1803	3	2	2	4	2	1	9
1804	6	4	0	6	0	2	5
1805	2	6	1	6	1	3	4
1807	0	4	2	6	2	7	2
1810	1	0	4	0	2	1	15
1812	0	0	0	0	0	0	0
1813	2	5	1	8	2	1	4
1814	0	3	4	3	0	3	10
1815	10	0	0	3	0	0	10
1816	2	0	0	2	0	0	19
1817	1	2	3	6	1	4	6
1818	5	1	1	6	0	1	9
1901	2	1	0	7	2	1	10
1902	4	0	0	7	2	1	9
1903	4	0	0	8	0	0	11
1904	1	0	0	6	0	0	16
1905	4	0	0	6	0	0	13
1906	3	0	0	8	0	0	12
1908	2	1	2	10	0	1	7
1909	0	2	0	2	0	3	16
1910	4	0	0	12	0	0	7
1911	5	0	0	7	0	0	11
1913	5	0	0	5	0	0	13
1914	1	1	0	5	2	7	7
1915	0	0	0	0	0	0	0
1916	1	2	2	6	3	1	8
1918	5	2	2	6	0	0	8

176 94 48245 34 81495

TEST CODE	8 -3	-2	-1	0	+1	+2	+3
1602	3	1	7	0	0	2	9
1603	4	0	0	7	0	0	11
1604	5	0	0	0	0	3	14
1606	9	0	0	0	0	0	13
1607	8	0	0	1	0	0	13
1608	5	1	0	0	0	1	15
1610	5	0	0	3	0	0	14
1611	4	0	2	0	0	3	13
1612	8	0	0	1	0	0	13
1613	4	0	0	1	0	0	17
1701	6	3	1	3	0	1	8
1702	6	2	0	7	1	1	5
1703	8	3	1	0	0	0	10
1704	2	1	2	9	0	2	6
1705	1	8	0	2	0	3	8
1707	9	1	0	3	0	1	8
1708	8	3	1	0	0	0	10
1709	5	4	2	2	1	3	5
1710	4	2	1	2	0	2	11
1711	8	1	1	3	0	0	9
1712	3	2	0	0	0	4	13
1713	3	0	0	8	0	0	11
1714	0	2	1	6	2	9	2
1715	6	2	1	5	0	0	8
1801	5	0	1	5	1	0	10
1802	5	1	0	1	0	1	14
1803	4	1	2	8	0	1	6
1804	1	2	1	6	2	7	3
1805	2	5	1	5	1	2	6
1807	0	8	1	6	2	1	4
1810	5	0	0	2	0	1	14
1812	0	0	0	0	0	0	0
1813	3	4	2	5	0	2	6
1814	1	1	2	1	7	3	7
1815	9	0	0	10	0	0	3
1816	3	0	0	0	0	0	19
1817	2	5	0	10	0	1	4
1818	3	0	0	8	0	1	10
1901	2	0	1	9	1	0	9
1902	8	0	0	7	1	0	6
1903	3	0	0	8	0	0	11
1904	3	0	0	2	0	0	17
1905	2	0	0	9	0	0	11
1906	6	0	0	8	0	0	8
1908	1	1	0	10	2	3	5
1909	2	2	0	9	0	1	8
1910	3	0	1	10	1	0	7
1911	3	0	0	6	0	0	13
1913	2	0	0	8	0	0	12
1914	4	1	0	8	0	4	5
1915	4	2	2	0	1	1	12
1916	2	2	1	11	0	1	5
1918	0	0	3	12	0	1	6

212 71 38247 23 66487

TEST CODE	9 -3	-2	-1	0	+1	+2	+3
1602	1	0	0	0	0	0	23
1603	7	0	0	1	0	1	15
1604	2	0	1	0	0	3	18
1606	4	0	0	0	0	0	20
1607	4	0	1	1	0	1	17
1608	1	0	0	3	0	1	19
1610	2	0	0	3	0	1	18
1611	0	1	0	2	0	1	20
1612	4	0	0	0	0	0	20
1613	3	0	0	0	0	0	21
1701	6	1	0	3	0	4	10
1702	5	0	0	5	0	0	14
1703	7	0	0	2	0	0	15
1704	5	0	0	9	0	0	10
1705	2	1	2	6	0	2	11
1707	3	1	0	5	0	2	13
1708	3	1	0	5	0	2	13
1709	9	0	2	3	1	0	9
1710	4	0	1	1	2	0	16
1711	5	1	0	3	0	2	13
1712	4	2	0	0	0	0	18
1713	5	0	0	9	0	0	10
1714	5	0	1	3	0	5	10
1715	5	1	0	11	0	0	7
1801	6	0	3	8	2	0	5
1802	3	2	0	0	1	2	16
1803	3	3	1	3	3	1	10
1804	2	2	2	5	2	4	7
1805	2	5	1	5	0	5	6
1807	0	2	3	8	1	7	3
1810	2	0	2	3	1	0	16
1812	0	0	0	0	0	0	0
1813	3	4	0	4	4	4	5
1814	1	0	3	2	4	5	9
1815	9	3	0	4	0	0	8
1816	3	0	0	1	0	0	20
1817	2	4	1	7	0	6	4
1818	5	0	0	9	1	1	8
1901	4	0	2	6	1	0	11
1902	1	0	0	6	1	0	16
1903	1	0	0	7	0	0	16
1904	1	0	0	4	0	0	19
1905	6	0	0	8	0	0	10
1906	5	0	0	10	1	0	8
1908	1	2	3	7	3	2	6
1909	2	4	0	4	0	4	10
1910	6	1	0	13	0	0	4
1911	0	0	0	8	0	1	15
1913	5	0	1	5	0	0	13
1914	1	5	1	4	0	5	8
1915	2	0	1	1	2	2	16
1916	2	0	0	11	3	1	7
1918	5	1	1	8	0	0	9

179 47 33236 33 75645

TEST CODE	10 -3	-2	-1	0	+1	+2	+3
1602	6	6	2	3	0	2	6
1603	3	2	3	5	0	3	9
1604	10	1	2	1	0	2	9
1606	11	0	0	0	0	0	14
1607	10	1	0	5	0	2	7
1608	6	0	0	10	0	0	9
1610	8	0	0	3	0	2	12
1611	6	0	0	7	0	1	11
1612	8	0	0	3	0	0	14
1613	5	0	0	5	0	0	15
1701	10	1	1	4	0	1	8
1702	9	0	0	8	0	0	8
1703	11	1	0	4	0	1	8
1704	6	1	0	6	3	0	9
1705	6	4	1	7	1	1	5
1707	3	6	1	4	0	2	9
1708	5	5	4	2	1	3	5
1709	1	1	3	7	1	1	11
1710	5	0	3	5	2	0	10
1711	4	1	1	6	3	0	10
1712	7	4	1	3	0	2	8
1713	9	0	0	5	1	0	10
1714	3	3	0	10	0	0	9
1715	6	2	0	5	1	0	11
1801	9	0	0	5	0	0	11
1802	3	2	0	0	2	3	15
1803	3	4	2	1	3	4	8
1804	2	2	6	5	4	4	2
1805	0	11	0	5	0	9	0
1807	0	4	2	9	1	7	2
1810	3	1	4	1	5	1	10
1812	0	0	0	0	0	0	0
1813	3	5	1	4	2	4	6
1814	1	3	6	5	0	4	6
1815	2	0	0	9	0	3	11
1816	0	0	0	0	0	0	0
1817	1	2	2	8	2	4	6
1818	3	2	1	6	2	0	11
1901	6	0	0	8	0	2	9
1902	3	0	0	9	0	0	13
1903	0	0	0	12	0	0	13
1904	5	0	0	19	0	0	11
1905	6	0	0	7	0	1	11
1906	0	0	0	0	0	0	0
1908	4	0	2	9	3	1	6
1909	3	3	0	4	0	5	10
1910	3	0	1	8	1	0	12
1911	6	0	1	7	0	1	11
1913	10	0	0	5	0	0	10
1914	0	4	0	8	0	6	7
1915	3	3	0	4	1	1	13
1916	2	1	4	9	0	4	5
1918	3	0	1	10	0	0	11

242 86 55295 39 87457

TEST CODE	11 −3	−2	−1	0	+1	+2	+3
1602	13	0	0	6	0	0	5
1603	1	4	0	7	0	2	10
1604	8	3	0	1	0	0	12
1606	4	0	0	2	0	0	18
1607	6	0	0	4	0	0	14
1608	1	3	0	5	0	3	12
1610	4	0	0	3	0	0	17
1611	7	1	0	7	0	0	9
1612	17	0	0	5	0	0	2
1613	6	0	0	2	0	0	16
1701	8	1	1	5	0	0	9
1702	6	7	2	5	1	2	1
1703	5	2	1	4	1	5	6
1704	7	0	4	3	0	0	10
1705	3	8	0	6	0	2	5
1707	4	7	2	7	0	1	3
1708	2	5	1	5	0	6	5
1709	0	0	0	0	0	0	0
1710	11	2	0	4	0	2	5
1711	7	0	2	5	0	0	10
1712	2	5	1	4	0	2	10
1713	8	1	0	5	0	0	10
1714	0	3	0	8	0	5	8
1715	10	0	0	5	0	0	9
1801	6	0	3	9	1	0	5
1802	4	0	2	1	2	2	13
1803	8	0	2	4	1	2	7
1804	4	6	2	5	2	1	4
1805	4	8	0	9	1	2	0
1807	2	10	3	6	1	1	1
1810	1	1	4	4	0	0	14
1812	1	1	5	13	2	1	1
1813	2	2	1	4	0	3	12
1814	1	0	4	4	5	2	8
1815	6	0	0	5	0	0	8
1816	2	0	0	1	0	0	21
1817	3	3	4	7	1	2	4
1818	9	1	3	7	0	2	2
1901	0	1	1	10	0	2	10
1902	5	2	0	9	0	0	8
1903	10	0	0	11	0	0	3
1904	3	0	0	5	0	0	16
1905	8	0	0	10	0	0	6
1906	7	0	0	10	0	0	7
1908	3	1	1	10	2	0	7
1909	2	3	0	7	0	1	11
1910	8	0	2	6	0	0	8
1911	11	1	0	8	0	0	4
1913	6	0	1	11	0	0	6
1914	3	0	0	10	0	4	7
1915	2	0	0	4	0	0	18
1916	3	1	2	9	3	2	4
1918	0	0	0	0	0	0	0

264 93 54312 23 57421

TEST CODE	12 −3	−2	−1	0	+1	+2	+3
1602	1	2	5	6	1	5	5
1603	6	6	0	7	0	1	5
1604	9	5	5	1	2	0	3
1606	5	0	0	6	0	0	14
1607	4	2	0	10	0	1	8
1608	4	1	0	13	0	5	2
1610	7	2	0	7	0	0	9
1611	8	5	0	3	0	0	9
1612	12	1	0	3	0	1	8
1613	5	2	0	12	0	0	6
1701	5	3	0	10	0	2	5
1702	5	0	0	10	0	0	10
1703	6	2	1	12	0	0	4
1704	7	2	1	10	0	1	4
1705	6	3	2	7	1	4	2
1707	4	4	2	10	1	2	2
1708	5	1	5	7	1	2	4
1709	0	0	0	0	0	0	0
1710	3	1	2	9	1	0	9
1711	4	0	2	9	3	0	7
1712	2	4	2	5	1	5	6
1713	5	0	1	13	0	2	4
1714	2	4	0	9	2	5	3
1715	6	1	0	10	0	1	7
1801	7	0	0	9	0	0	9
1802	4	2	4	0	1	0	14
1803	8	1	1	8	1	2	4
1804	2	4	5	8	1	1	4
1805	4	8	0	4	0	5	4
1807	1	3	6	5	0	7	3
1810	6	0	5	3	3	0	8
1812	2	3	4	8	1	3	4
1813	4	1	3	7	1	0	9
1814	2	1	4	6	1	3	8
1815	5	0	0	9	0	0	11
1816	2	0	0	2	0	0	21
1817	4	0	1	10	2	1	7
1818	4	2	3	7	1	3	5
1901	3	0	0	14	0	0	6
1902	0	0	0	0	0	0	0
1903	4	0	0	8	0	0	13
1904	1	0	0	14	0	0	10
1905	8	0	0	8	0	0	9
1906	6	0	0	10	0	0	9
1908	3	0	2	11	2	0	7
1909	0	0	0	0	0	0	0
1910	6	0	0	13	0	0	6
1911	6	0	0	8	0	0	11
1913	2	1	0	12	3	0	7
1914	3	6	0	5	0	6	5
1915	5	2	0	2	0	0	16
1916	1	3	2	10	0	7	2
1918	0	0	0	0	0	0	0

224 88 68390 30 75350

TEST CODE	13 −3	−2	−1	0	+1	+2	+3
1602	3	0	3	6	2	0	11
1603	9	4	0	11	1	0	0
1604	12	0	1	3	0	0	9
1606	8	0	0	4	0	0	13
1607	5	0	0	9	1	0	10
1608	2	0	0	14	0	1	8
1610	0	0	0	0	0	0	0
1611	8	0	0	9	0	1	7
1612	0	0	0	0	0	0	0
1613	5	1	0	11	0	3	5
1701	9	1	0	8	0	0	7
1702	1	0	3	11	3	0	7
1703	7	2	2	9	0	0	5
1704	4	3	1	12	0	1	4
1705	2	8	3	5	0	3	4
1707	4	3	2	11	0	4	1
1708	2	2	9	6	4	0	2
1709	0	0	0	0	0	0	0
1710	5	0	3	11	2	1	3
1711	2	1	2	14	0	0	6
1712	4	5	0	7	0	4	5
1713	3	2	0	13	1	1	5
1714	5	4	0	10	0	4	2
1715	4	2	1	14	0	0	4
1801	5	0	3	5	2	0	10
1802	1	3	2	3	1	1	14
1803	0	0	0	0	0	0	0
1804	3	3	1	8	3	3	4
1805	1	8	0	6	0	8	2
1807	0	6	1	9	3	3	3
1810	1	0	4	4	3	0	13
1812	2	1	9	1	4	4	4
1813	4	3	0	9	0	2	7
1814	0	0	0	0	0	0	0
1815	8	0	0	11	0	0	6
1816	1	0	0	0	0	0	24
1817	5	0	4	10	2	2	2
1818	0	0	0	0	0	0	0
1901	0	0	0	0	0	0	0
1902	3	0	0	11	0	1	10
1903	2	0	0	14	0	0	9
1904	3	0	0	7	0	0	15
1905	6	0	0	7	0	0	12
1906	0	0	0	0	0	0	0
1908	0	0	0	0	0	0	0
1909	4	0	3	7	0	5	6
1910	9	0	0	12	0	0	4
1911	1	0	0	10	0	0	14
1913	0	0	0	0	0	0	0
1914	1	3	0	13	0	6	2
1915	0	0	0	0	0	0	0
1916	0	0	0	0	0	0	0
1918	0	0	0	0	0	0	0

164 65 57345 32 58279

TEST CODE	14 −3	−2	−1	0	+1	+2	+3
1602	3	5	0	7	1	0	9
1603	4	4	0	8	0	0	9
1604	12	0	1	0	1	0	11
1606	10	0	0	4	0	0	11
1607	5	1	1	5	1	0	12
1608	3	6	0	7	0	4	5
1610	8	0	0	6	0	0	11
1611	9	0	1	4	0	1	10
1612	12	0	0	2	0	1	10
1613	11	0	0	7	0	1	6
1701	7	1	0	10	0	0	7
1702	3	2	2	5	1	4	8
1703	4	5	0	10	1	0	5
1704	0	0	0	0	0	0	0
1705	3	3	1	10	0	4	4
1707	0	6	0	13	0	3	3
1708	5	0	5	11	0	0	4
1709	0	0	0	0	0	0	0
1710	4	1	3	10	0	0	7
1711	2	0	2	13	0	1	7
1712	2	5	0	5	0	3	10
1713	2	3	2	12	0	2	4
1714	3	1	0	12	0	2	7
1715	2	4	0	12	0	2	5
1801	4	0	5	5	3	0	6
1802	3	5	0	0	2	0	15
1803	5	1	5	2	0	2	10
1804	2	1	5	8	0	4	5
1805	4	8	0	7	0	4	2
1807	1	2	3	7	2	8	2
1810	4	0	4	4	5	0	8
1812	5	1	5	7	2	4	1
1813	5	5	0	8	0	3	4
1814	1	1	8	4	0	3	8
1815	2	0	0	11	0	0	12
1816	4	0	0	1	0	0	20
1817	3	1	5	8	3	1	4
1818	2	0	3	11	1	4	4
1901	4	3	2	8	1	0	7
1902	3	0	0	11	0	0	11
1903	4	0	0	14	0	0	7
1904	0	0	0	9	0	0	16
1905	7	0	0	6	0	0	12
1906	4	0	0	14	0	0	7
1908	3	3	1	11	2	2	3
1909	2	3	0	6	0	2	12
1910	2	0	1	10	0	0	12
1911	2	0	2	10	0	1	10
1913	2	0	3	10	0	0	10
1914	1	1	0	12	0	3	8
1915	1	1	0	10	0	1	12
1916	1	0	1	15	3	0	5
1918	0	0	0	0	0	0	0

195 83 71402 29 70400

1602

TEST	-3	-2	-1	0	+1	+2	+3
1	2	0	1	2	0	8	10
2	4	6	0	0	2	6	6
3	8	0	0	0	0	3	14
4	6	3	0	0	0	2	14
5	20	0	0	0	0	0	30
6	10	0	0	1	0	0	13
7	3	1	3	0	2	1	13
8	3	1	7	0	0	2	9
9	1	0	0	0	0	0	23
10	6	6	2	3	0	2	6
11	13	0	0	6	0	0	5
12	1	2	5	6	1	8	5
13	3	0	3	6	2	0	11
14	3	5	0	7	1	0	9

83 24 21 31 8 29160

1603

TEST	-3	-2	-1	0	+1	+2	+3
1	1	2	0	12	0	1	7
2	7	2	0	3	0	0	12
3	5	2	0	3	0	2	13
4	3	0	0	5	0	0	17
5	4	5	0	10	0	5	26
6	10	1	1	1	0	0	11
7	3	2	0	3	0	4	11
8	4	0	0	7	0	0	11
9	7	0	0	1	0	1	15
10	3	2	3	5	0	3	9
11	1	4	0	7	0	2	10
12	6	6	0	7	0	1	5
13	9	4	0	11	1	0	0
14	4	4	0	8	0	0	9

67 34 4 83 1 19156

1604

TEST	-3	-2	-1	0	+1	+2	+3
1	4	3	0	0	0	6	10
2	4	5	0	0	0	4	11
3	4	5	0	0	0	1	15
4	6	2	1	0	0	1	15
5	14	1	7	0	2	0	23
6	6	2	1	0	0	1	14
7	3	1	1	0	0	3	15
8	5	0	0	0	0	3	14
9	2	0	1	0	0	3	18
10	10	1	2	1	0	2	9
11	0	1	0	1	0	0	12
12	9	5	5	1	2	0	3
13	12	0	1	3	0	0	9
14	12	0	1	0	1	0	11

99 23 20 6 5 24182

1606

TEST	-3	-2	-1	0	+1	+2	+3
1	2	0	2	5	1	1	12
2	2	6	0	0	0	2	14
3	3	2	0	0	0	4	16
4	6	0	0	0	0	0	19
5	10	0	0	4	0	0	36
6	8	0	0	0	0	0	16
7	4	0	0	0	0	0	19
8	9	0	0	0	0	0	13
9	4	0	0	0	0	0	20
10	11	0	0	0	0	0	14
11	4	0	0	2	0	0	18
12	5	0	0	6	0	0	14
13	8	0	0	4	0	0	13
14	10	0	0	4	0	0	11

86 8 2 25 1 7235

1607

TEST	-3	-2	-1	0	+1	+2	+3
1	2	4	3	4	0	1	9
2	4	3	1	0	0	3	13
3	6	3	2	0	0	1	13
4	5	2	1	0	0	3	14
5	10	0	0	14	0	1	25
6	7	0	0	3	0	0	14
7	7	1	0	1	0	0	14
8	6	0	0	1	0	0	13
9	4	0	1	1	0	1	17
10	10	1	0	5	0	2	7
11	6	0	0	4	0	0	14
12	4	2	0	10	0	1	8
13	5	0	0	9	1	0	10
14	5	1	1	5	1	0	12

83 17 9 57 2 13183

1608

TEST	-3	-2	-1	0	+1	+2	+3
1	1	1	3	6	0	7	5
2	0	3	0	5	0	11	5
3	0	4	1	0	1	14	5
4	5	0	0	1	0	0	19
5	2	7	3	12	0	22	4
6	2	1	0	2	0	4	15
7	4	4	0	3	0	5	7
8	5	1	0	0	0	1	15
9	1	0	0	3	0	1	19
10	6	0	0	10	0	0	9
11	1	3	0	3	0	3	12
12	4	1	0	13	0	5	2
13	2	0	0	14	0	1	8
14	3	6	0	7	0	4	5

36 31 7 81 1 76130

1610

TEST	-3	-2	-1	0	+1	+2	+3
1	6	1	1	0	1	5	11
2	3	5	0	1	0	5	10
3	3	2	0	0	1	1	18
4	5	0	0	1	0	0	19
5	7	0	0	14	0	0	29
6	8	0	0	4	0	0	12
7	7	0	0	1	0	1	14
8	5	0	0	3	0	0	14
9	2	0	0	3	0	1	16
10	8	0	0	3	0	2	12
11	4	0	0	3	0	0	17
12	7	2	0	7	0	0	9
13	0	0	0	0	0	0	0
14	8	0	0	6	0	0	11

73 10 1 46 2 13194

1611

TEST	-3	-2	-1	0	+1	+2	+3
1	1	5	2	5	3	1	6
2	0	5	4	0	4	5	6
3	1	2	3	4	3	4	8
4	6	1	0	5	0	3	10
5	3	4	2	17	7	1	16
6	5	2	0	2	0	2	13
7	7	0	1	4	0	0	11
8	4	0	2	0	0	3	15
9	0	1	0	2	0	1	20
10	6	0	0	7	0	1	11
11	7	1	0	7	0	0	9
12	3	5	0	3	0	0	9
13	8	0	0	9	0	1	7
14	9	0	1	4	0	1	10

65 26 15 69 17 23149

1612

TEST	-3	-2	-1	0	+1	+2	+3
1	1	5	0	3	0	4	10
2	6	5	0	0	0	1	12
3	3	3	0	1	0	4	14
4	5	0	0	3	0	0	17
5	9	0	0	4	0	0	37
6	6	0	0	1	0	0	17
7	3	0	0	1	0	0	19
8	8	0	0	1	0	0	13
9	4	0	0	0	0	0	20
10	8	0	0	3	0	0	14
11	17	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
12	12	1	0	3	0	1	0
13	0	0	0	0	0	0	0
14	17	0	0	7	0	1	10

94 14 0 27 0 11153

1613

TEST	-3	-2	-1	0	+1	+2	+3
1	4	0	0	2	1	11	[illegible]
2	0	0	0	0	0	0	0
3	5	3	1	0	0	7	9
4	3	1	0	1	0	1	19
5	9	1	0	6	0	2	32
6	7	0	0	0	0	0	17
7	6	0	0	2	0	0	13
8	4	0	0	1	0	0	17
9	3	0	0	0	0	0	21
10	5	0	0	5	0	0	15
11	6	0	0	2	0	0	16
12	5	2	0	12	0	0	6
13	5	1	0	11	0	3	9
14	11	0	0	7	0	1	6

75 8 1 45 1 25131

1701

TEST	-3	-2	-1	0	+1	+2	+3
1	0	3	6	4	2	2	6
2	3	1	6	0	2	2	10
3	6	3	2	0	1	1	12
4	6	2	0	0	0	0	17
5	3	8	5	9	2	4	19
6	6	2	1	5	0	1	9
7	4	6	5	8	0	2	6
8	6	3	1	3	0	1	[illegible]
9	6	1	0	3	0	4	10
10	10	1	1	4	0	1	9
11	8	1	1	5	0	0	9
12	5	3	0	10	0	2	5
13	[illegible]	1	0	[illegible]	0	[illegible]	7
14	7	1	0	10	0	0	7

79 36 23 64 7 22133

1702

TEST	-3	-2	-1	0	+1	+2	+3
1	3	3	3	1	2	3	8
2	4	2	6	0	0	2	10
3	5	2	3	0	0	4	11
4	4	2	1	0	0	1	17
5	3	1	0	14	0	2	29
6	2	0	1	11	0	2	0
7	3	3	0	4	0	2	7
8	6	2	0	7	1	1	5
9	5	0	0	5	0	0	14
10	9	0	0	8	0	0	8
11	4	7	7	[illegible]	1	2	1
12	5	0	0	10	0	0	[illegible]
13	1	0	3	11	5	0	7
14	3	2	2	5	1	4	[illegible]

59 24 21 5 1 25145

1703

TEST	-3	-2	-1	0	+1	+2	+3
1	3	6	3	0	0	4	7
2	4	6	3	0	0	4	7
3	6	6	2	0	1	2	6
4	4	4	0	0	1	[illegible]	[illegible]
5	14	2	0	9	0	4	21
6	1	4	1	4	2	4	7
7	5	4	0	2	0	2	10
8	8	3	1	0	0	0	10
9	7	0	0	2	0	0	15
10	11	1	0	4	0	1	3
11	5	2	1	4	1	5	6
12	6	2	1	12	0	0	4
13	7	2	2	9	0	0	5
14	4	5	0	10	1	0	5

85 47 14 56 6 36120

1704

TEST	-3	-2	-1	0	+1	+2	+3
1	4	5	6	0	1	1	5
2	1	5	3	3	1	5	6
3	4	0	9	4	1	1	6
4	3	1	7	3	5	4	2
5	7	6	3	14	4	7	9
6	0	4	6	7	0	1	6
7	6	0	0	10	0	0	7
8	2	1	2	9	0	2	6
9	5	0	0	0	0	0	10
10	6	1	0	6	3	0	9
11	7	0	4	3	0	0	10
12	7	2	1	10	0	1	4
13	4	3	1	12	0	1	4
14	0	0	0	0	0	0	0

56 28 42 90 15 23 85

1705

TEST	-3	-2	-1	0	+1	+2	+3
1	2	2	7	0	4	4	4
2	0	2	10	0	2	5	5
3	1	4	4	1	6	4	5
4	3	1	5	4	2	0	10
5	4	3	5	14	5	6	13
6	3	0	7	6	2	1	5
7	1	8	5	6	0	3	0
8	1	0	0	2	0	3	[illegible]
9	2	1	2	6	0	2	11
10	6	4	1	7	1	1	5
11	3	8	0	6	0	2	5
12	6	3	2	7	1	4	2
13	2	0	3	[illegible]	0	5	4
14	3	3	1	10	0	4	4

37 55 52 74 23 42 81

1707

TEST	-3	-2	-1	0	+1	+2	+3
1	0	0	6	0	7	5	5
2	1	4	4	4	2	3	6
3	1	0	7	5	2	4	6
4	1	5	0	5	1	4	9
5	2	4	3	6	3	11	21
6	3	3	5	4	0	3	6
7	5	3	0	8	0	4	5
8	9	1	0	3	0	1	8
9	3	1	0	5	0	2	13
10	3	6	1	4	0	2	9
11	4	7	2	7	0	1	3
12	4	4	2	10	1	2	2
13	4	3	2	11	0	4	1
14	0	6	0	13	0	3	3

38 47 32 85 16 49 97

1708

TEST	-3	-2	-1	0	+1	+2	+3
1	2	5	3	0	2	5	6
2	0	0	12	0	7	2	3
3	1	6	4	0	0	8	6
4	1	1	4	1	2	11	5
5	7	3	10	9	[illegible]	4	9
6	0	0	6	7	9	0	2
7	3	4	3	1	2	1	9
8	8	3	1	0	0	0	10
9	3	1	0	[illegible]	0	2	13
10	5	5	4	2	1	[illegible]	[illegible]
11	2	5	1	5	0	6	[illegible]
12	5	1	5	7	1	2	4
13	2	2	9	6	4	0	2
14	5	0	5	11	0	0	4

44 36 67 54 36 44 83

1709

TEST	-3	-2	-1	0	+1	+2	+3
1	1	8	1	0	0	6	7
2	3	4	4	0	3	2	8
3	7	2	2	0	0	1	13
4	5	1	1	0	1	3	14
5	5	3	1	14	1	6	20
6	1	3	3	7	0	1	9
7	2	3	0	8	1	1	8
8	5	4	2	2	1	3	5
9	9	0	2	3	1	0	9
10	1	1	3	7	1	1	11
11	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0

39 29 19 41 9 24104

1710

TEST	-3	-2	-1	0	+1	+2	+3
1	2	2	5	0	2	2	10
2	4	5	3	0	0	1	11
3	2	0	4	2	2	3	4
4	1	4	0	5	1	2	12
5	6	2	4	7	5	6	20
6	4	2	4	1	2	2	9
7	3	3	1	4	3	0	9
8	4	2	1	2	0	2	11
9	4	0	1	1	2	0	16
10	5	0	3	5	2	0	10
11	11	2	0	4	0	2	5
12	3	1	2	9	1	0	9
13	5	0	3	11	2	1	3
14	4	1	3	10	0	0	7

58 32 34 61 22 21136

1711

TEST	-3	-2	-1	0	+1	+2	+3
1	2	2	6	0	3	4	6
2	2	4	6	0	3	2	7
3	1	2	4	2	2	1	13
4	4	0	1	2	2	2	14
5	4	2	1	10	1	2	22
6	9	0	0	7	0	0	8
7	4	0	4	5	1	1	3
8	2	1	1	3	0	0	9
9	5	1	0	3	0	2	15
10	4	1	1	6	5	0	10
11	7	0	2	5	0	0	10
12	4	0	2	9	3	0	7
13	2	1	2	14	0	0	6
14	2	0	2	13	0	1	7

58 14 32 87 18 15140

1712

TEST	-3	-2	-1	0	+1	+2	+3
1	2	3	2	0	3	1	7
2	2	7	3	0	0	9	3
3	7	3	0	0	0	5	10
4	1	3	1	0	4	8	8
5	3	6	4	5	2	10	20
6	0	5	3	3	1	4	3
7	1	7	2	1	1	3	8
8	3	2	0	0	0	4	13
9	4	2	0	0	0	0	13
10	7	4	1	3	0	2	8
11	2	5	1	4	0	2	10
12	2	4	2	5	1	5	6
13	4	5	0	7	0	4	5
14	2	3	0	5	0	3	10

41 66 19 33 12 60134

1713

TEST	-3	-2	-1	0	+1	+2	+3
1	0	8	0	0	1	2	4
2	6	5	3	0	3	2	5
3	3	5	2	5	0	2	6
4	7	0	0	3	0	0	13
5	8	0	0	17	0	0	25
6	4	0	0	8	0	0	12
7	8	0	0	8	0	0	7
8	3	0	0	8	0	0	11
9	5	0	0	9	0	0	10
10	9	0	0	5	1	0	10
11	8	1	0	5	0	0	10
12	5	0	1	13	0	2	4
13	3	2	0	13	1	1	5
14	2	3	2	12	0	2	4

71 24 16106 6 13128

1714

TEST	-3	-2	-1	0	+1	+2	+3
1	2	2	8	0	2	5	4
2	1	6	1	1	1	3	6
3	0	3	4	4	0	9	5
4	1	2	3	4	3	6	6
5	0	8	0	9	0	33	0
6	1	5	0	6	1	3	8
7	1	3	1	7	2	6	3
8	0	2	1	6	2	9	2
9	5	0	1	3	0	5	10
10	3	3	0	10	0	0	9
11	0	3	0	8	0	5	8
12	2	4	0	9	2	5	3
13	5	4	0	10	0	4	2
14	3	1	0	12	0	2	7

24 46 19 89 13160 73

1715

TEST	-3	-2	-1	0	+1	+2	+3
1	0	0	0	0	0	0	0
2	2	4	8	8	2	0	6
3	4	7	1	5	1	2	6
4	6	2	0	3	0	4	10
5	5	3	1	19	2	3	17
6	4	0	0	9	0	2	9
7	4	3	0	9	0	1	6
8	6	2	1	5	0	0	8
9	5	1	0	11	0	0	7
10	6	2	0	5	1	0	11
11	10	0	0	5	0	0	9
12	6	1	0	10	0	1	7
13	4	2	1	14	0	0	4
14	2	4	0	12	0	2	5

64 31 9112 6 15104

1801

TEST	-3	-2	-1	0	+1	+2	+3
1	1	1	0	0	0	0	13
2	3	3	0	9	0	0	9
3	6	0	0	7	1	0	11
4	5	0	2	10	1	0	7
5	17	0	0	8	0	0	24
6	10	0	0	6	0	0	8
7	3	2	1	7	2	0	8
8	5	0	1	5	1	0	10
9	6	0	3	8	2	0	5
10	9	0	0	5	0	0	11
11	6	0	3	9	1	0	5
12	7	0	0	9	0	0	9
13	5	0	3	5	2	0	10
14	4	0	5	5	3	0	8

87 6 18101 13 0138

1802

TEST	-3	-2	-1	0	+1	+2	+3
1	1	1	1	0	5	6	9
2	3	2	3	0	0	3	13
3	4	1	1	0	1	4	14
4	4	0	0	4	1	0	16
5	6	2	5	0	2	2	32
6	4	2	3	0	1	1	13
7	3	0	2	0	1	2	15
8	5	1	0	1	0	1	14
9	3	2	0	0	1	2	16
10	3	2	0	0	2	3	15
11	4	0	2	1	2	2	13
12	4	2	4	0	1	0	14
13	1	3	2	3	1	1	14
14	3	5	0	0	2	0	15

48 23 23 9 20 27213

1803

TEST	-3	-2	-1	0	+1	+2	+3
1	2	0	0	2	0	5	14
2	7	4	0	3	1	2	7
3	5	2	1	2	1	3	11
4	4	1	4	5	3	1	7
5	5	5	5	3	3	8	20
6	5	3	2	3	0	4	7
7	3	2	2	4	2	1	9
8	4	1	2	8	0	1	6
9	3	3	1	3	3	1	10
10	3	4	2	1	3	4	8
11	8	0	2	4	1	2	7
12	8	1	1	8	1	2	4
13	0	0	0	0	0	0	0
14	5	1	5	2	0	2	10

62 27 27 48 18 36120

1804

TEST	-3	-2	-1	0	+1	+2	+3
1	0	0	1	6	0	3	13
2	4	6	1	6	0	1	6
3	2	4	0	5	2	2	10
4	4	4	2	4	0	6	5
5	2	6	7	7	5	8	14
6	1	2	5	2	3	2	9
7	6	4	0	6	0	2	5
8	1	2	1	6	2	7	3
9	2	2	2	5	2	4	7
10	2	2	6	5	4	4	2
11	4	6	7	5	2	1	4
12	2	4	5	8	1	1	4
13	3	3	1	8	3	3	4
14	2	1	5	8	0	4	5

35 46 38 81 24 48 91

1805

TEST	-3	-2	-1	0	+1	+2	+3
1	1	2	1	3	1	5	10
2	3	5	0	6	0	3	7
3	3	4	0	3	0	3	12
4	7	4	0	6	0	3	5
5	8	5	0	7	1	2	26
6	4	4	0	3	1	3	9
7	2	6	1	6	1	3	4
8	2	5	1	5	1	2	6
9	2	5	1	5	0	5	6
10	0	11	0	5	0	9	0
11	4	8	0	9	1	2	0
12	4	8	0	4	0	5	4
13	1	8	0	6	0	8	2
14	4	8	0	7	0	4	2

45 83 4 75 6 57 93

1807

TEST	-3	-2	-1	0	+1	+2	+3
1	0	4	0	6	2	6	5
2	0	5	1	9	1	7	1
3	0	0	0	0	0	0	0
4	0	5	3	5	2	6	4
5	0	6	6	8	5	13	11
6	0	5	2	7	1	2	7
7	0	4	2	6	2	7	2
8	0	8	1	6	2	1	4
9	0	2	3	8	1	7	3
10	0	4	2	9	1	7	2
11	2	10	3	6	1	1	1
12	1	3	6	5	0	7	3
13	0	6	1	9	3	3	3
14	1	2	3	7	2	8	2

4 64 33 91 23 75 48

1810

TEST	-3	-2	-1	0	+1	+2	+3
1	0	1	0	2	0	7	14
2	4	5	0	0	0	2	13
3	3	0	0	1	6	2	13
4	0	0	3	1	1	2	18
5	3	0	1	1	5	0	39
6	3	0	1	1	0	0	19
7	1	0	4	0	2	1	15
8	5	0	0	2	0	1	14
9	2	0	2	3	1	0	16
10	3	1	4	1	5	1	10
11	1	1	4	4	0	0	14
12	6	0	5	3	3	0	8
13	1	0	4	4	3	0	13
14	4	0	4	4	5	0	8

36 8 32 27 31 16213

1812

TEST	-3	-2	-1	0	+1	+2	+3
1	1	1	1	5	2	1	12
2	1	6	3	5	2	3	4
3	1	2	1	7	1	10	3
4	3	2	1	8	0	5	6
5	4	1	2	12	7	6	17
6	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0
11	1	1	5	13	2	1	1
12	2	3	4	8	1	3	4
13	2	1	9	1	4	4	4
14	5	1	5	7	2	4	1

20 18 31 66 21 37 52

1813

TEST	-3	-2	-1	0	+1	+2	+3
1	1	3	3	3	1	2	10
2	3	3	2	3	1	3	9
3	4	4	3	4	0	4	6
4	4	8	5	0	2	0	6
5	4	5	6	3	2	9	20
6	5	2	1	4	2	2	8
7	2	5	1	8	2	1	4
8	3	4	2	5	0	2	6
9	3	4	0	4	4	4	5
10	3	5	1	4	2	4	6
11	2	2	1	4	0	3	12
12	4	1	3	7	1	0	9
13	4	3	0	9	0	2	7
14	5	5	0	8	0	3	4

47 54 28 66 17 39112

1814

TEST	-3	-2	-1	0	+1	+2	+3
1	0	1	3	2	1	0	0
2	3	2	3	3	3	7	5
3	2	0	1	2	2	5	13
4	0	2	4	4	1	4	10
5	3	4	2	3	4	9	24
6	0	1	6	3	3	4	7
7	0	3	4	3	0	3	10
8	1	1	2	1	7	3	7
9	1	0	3	2	4	5	9
10	1	3	6	5	0	4	6
11	1	0	4	4	5	2	8
12	2	1	4	6	1	3	0
13	0	0	0	0	0	0	0
14	1	1	8	4	0	3	8

15 19 50 42 31 60123

1815

TEST	-3	-2	-1	0	+1	+2	+3
1	0	0	0	9	2	2	10
2	2	0	0	10	2	0	10
3	3	0	0	7	0	0	15
4	1	0	0	10	1	0	13
5	8	1	0	10	0	3	27
6	9	2	0	3	0	1	9
7	10	0	0	3	0	0	10
8	9	0	0	10	0	0	3
9	9	3	0	4	0	0	8
10	2	0	0	9	0	3	11
11	6	0	0	10	0	0	8
12	5	0	0	9	0	0	11
13	8	0	0	11	0	0	6
14	2	0	0	11	0	0	12

74 6 0116 5 9153

1816

TEST	-3	-2	-1	0	+1	+2	+3
1	0	0	0	0	0	0	23
2	4	0	0	0	0	0	20
3	0	0	0	0	0	0	25
4	3	0	0	2	0	0	20
5	6	0	0	0	0	0	43
6	3	0	0	1	0	0	20
7	2	0	0	7	0	0	19
8	3	0	0	0	0	0	19
9	3	0	0	1	0	0	20
10	0	0	0	0	0	0	0
11	2	0	0	1	0	0	21
12	2	0	0	2	0	0	21
13	1	0	0	0	0	0	24
14	4	0	0	1	0	0	20

33 0 0 10 0 0295

1817

TEST	-3	-2	-1	0	+1	+2	+3
1	3	3	1	2	0	1	11
2	1	5	0	9	2	4	3
3	4	6	0	5	1	3	6
4	2	3	0	11	0	3	6
5	3	3	2	15	3	5	18
6	7	4	0	4	0	1	0
7	1	2	3	6	1	4	6
8	2	5	0	10	0	1	4
9	2	4	1	7	0	6	4
10	1	2	2	8	2	4	6
11	3	3	4	7	1	2	4
12	4	0	1	10	2	1	7
13	5	0	4	10	2	2	2
14	3	1	5	8	3	1	4

41 41 23112 17 38 91

1818

TEST	-3	-2	-1	0	+1	+2	+3
1	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0
6	1	3	3	6	0	2	9
7	5	1	1	6	0	1	9
8	3	0	0	8	0	1	10
9	5	0	0	9	1	1	8
10	3	2	1	6	2	0	11
11	9	1	3	7	0	2	2
12	4	2	3	7	1	3	5
13	0	0	0	0	0	0	0
14	2	0	3	11	1	4	4

32 9 14 60 5 14 58

1901

TEST	-3	-2	-1	0	+1	+2	+3
1	1	1	2	2	4	4	9
2	2	4	1	6	0	4	7
3	1	4	3	4	2	4	7
4	3	2	4	8	0	4	4
5	11	0	0	9	0	1	28
6	1	2	1	4	0	4	12
7	2	1	0	7	2	1	10
8	2	0	1	9	1	0	9
9	4	0	2	6	1	0	11
10	6	0	0	8	0	2	9
11	0	1	1	10	0	2	10
12	3	0	0	14	0	0	8
13	0	0	0	0	0	0	0
14	4	3	2	8	1	0	7

40 18 17 95 11 26131

1902

TEST	-3	-2	-1	0	+1	+2	+3
1	2	0	0	5	0	0	16
2	2	1	1	9	0	0	11
3	3	0	0	10	0	0	12
4	3	0	0	10	0	0	12
5	6	1	0	12	0	1	29
6	4	1	0	4	0	0	15
7	4	0	0	7	2	1	9
8	8	0	0	7	1	0	6
9	1	0	0	6	1	0	16
10	3	0	0	9	0	0	13
11	5	2	0	9	0	0	8
12	0	0	0	0	0	0	0
13	3	0	0	11	0	1	10
14	3	0	0	11	0	0	11

47 5 1110 4 3168

1903

TEST	-3	-2	-1	0	+1	+2	+3
1	3	0	0	5	1	2	12
2	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0
5	3	0	0	12	0	0	34
6	5	0	0	7	0	0	12
7	4	0	0	8	0	0	11
8	3	0	0	8	0	0	11
9	1	0	0	7	0	0	16
10	0	0	0	12	0	0	13
11	10	0	0	11	0	0	3
12	4	0	0	8	0	0	13
13	2	0	0	14	0	0	9
14	4	0	0	14	0	0	7

39 0 0106 1 2141

1904

TEST	-3	-2	-1	0	+1	+2	+3
1	0	0	0	0	0	0	0
2	6	1	0	5	0	0	12
3	2	0	0	1	0	0	22
4	0	0	0	0	0	0	0
5	3	0	0	5	0	0	41
6	0	0	0	0	0	0	0
7	1	0	0	6	0	0	16
8	3	0	0	2	0	0	17
9	1	0	0	4	0	0	19
10	5	0	0	19	0	0	11
11	3	0	0	5	0	0	16
12	1	0	0	14	0	0	10
13	3	0	0	7	0	0	15
14	0	0	0	9	0	0	16

28 1 0 77 0 0195

1905

TEST	-3	-2	-1	0	+1	+2	+3
1	2	1	0	5	1	1	11
2	4	2	4	0	0	1	13
3	5	1	0	5	0	0	14
4	4	1	0	4	0	2	14
5	3	0	0	15	0	0	31
6	5	0	0	5	0	0	14
7	4	0	0	6	0	0	13
8	2	0	0	9	0	0	11
9	6	0	0	8	0	0	10
10	6	0	0	7	0	1	11
11	8	0	0	10	0	0	6
12	8	0	0	8	0	0	9
13	6	0	0	7	0	0	12
14	7	0	0	6	0	0	12

70 5 4 95 1 5183

1906

TEST	-3	-2	-1	0	+1	+2	+3
1	3	0	0	3	4	2	11
2	2	2	3	6	2	3	6
3	2	0	1	12	1	0	9
4	4	0	1	10	0	0	10
5	6	0	0	10	0	2	31
6	5	0	0	7	0	0	12
7	3	0	0	8	0	0	12
8	6	0	0	8	0	0	8
9	5	0	0	10	1	0	8
10	0	0	0	0	0	0	0
11	7	0	0	10	0	0	7
12	6	0	0	10	0	0	9
13	0	0	0	0	0	0	0
14	4	0	0	14	0	0	7

53 2 5108 8 7130

1908

TEST	-3	-2	-1	0	+1	+2	+3
1	1	0	1	3	2	4	12
2	5	4	0	6	0	1	8
3	4	0	2	6	2	3	8
4	3	2	2	8	2	0	8
5	7	4	1	17	3	6	11
6	6	0	0	9	0	0	9
7	2	1	2	10	0	1	7
8	1	1	0	10	2	3	5
9	1	2	3	7	3	2	6
10	4	0	2	9	3	1	6
11	3	1	1	10	2	0	7
12	3	0	2	11	2	0	7
13	0	0	0	0	0	0	0
14	3	3	1	11	2	2	3

43 18 17117 23 23 97

1909

TEST	-3	-2	-1	0	+1	+2	+3
1	1	0	0	6	1	2	13
2	3	2	3	7	0	1	8
3	4	1	1	1	0	4	14
4	3	1	0	12	0	2	7
5	4	6	0	3	0	7	27
6	0	0	0	0	0	0	0
7	0	2	0	2	0	3	16
8	2	2	0	9	0	1	8
9	2	4	0	4	0	4	10
10	3	3	0	4	0	5	10
11	2	3	0	7	0	1	11
12	0	0	0	0	0	0	0
13	4	0	3	7	0	5	6
14	2	3	0	6	0	2	12

30 27 7 68 1 37142

1910

TEST	-3	-2	-1	0	+1	+2	+3
1	1	0	1	9	2	0	10
2	2	1	1	7	2	1	10
3	1	0	0	10	0	0	14
4	4	0	0	12	0	0	9
5	1	0	3	14	2	1	28
6	3	0	0	9	0	0	12
7	4	0	0	12	0	0	7
8	3	0	1	10	1	0	7
9	6	1	0	13	0	0	4
10	3	0	1	8	1	0	12
11	8	0	2	6	0	0	8
12	6	0	0	13	0	0	6
13	9	0	0	12	0	0	4
14	2	0	1	10	0	0	12

53 2 10145 8 2143

1911

TEST	-3	-2	-1	0	+1	+2	+3
1	3	2	0	4	0	0	14
2	2	1	2	10	0	1	8
3	6	3	0	5	0	0	11
4	2	0	0	7	0	0	16
5	9	0	2	10	0	1	27
6	4	2	0	6	0	1	11
7	5	0	0	7	0	0	11
8	3	0	0	6	0	0	13
9	0	0	0	8	0	1	15
10	6	0	1	7	0	1	11
11	11	1	0	8	0	0	4
12	6	0	0	8	0	0	11
13	1	0	0	10	0	0	14
14	2	0	2	10	0	1	10

60 9 7106 0 6176

1971-72

1913

TEST	-3	-2	-1	0	+1	+2	+3
1	1	0	2	9	0	3	8
2	2	2	1	7	1	1	10
3	0	2	2	1	6	1	13
4	2	0	2	2	5	1	13
5	9	0	3	12	2	0	23
6	4	1	6	0	2	1	10
7	5	0	0	5	0	0	13
8	2	0	0	8	0	0	12
9	5	0	1	5	0	0	13
10	10	0	0	5	0	0	10
11	6	0	1	11	0	0	6
12	2	1	0	12	3	0	7
13	0	0	0	0	0	0	0
14	2	0	3	10	0	0	10
	50	6	21	87	19	7	148

1916

TEST	-3	-2	-1	0	+1	+2	+3
1	0	0	0	5	3	2	13
2	2	0	0	11	3	1	7
3	1	0	3	6	6	0	9
4	1	1	1	11	0	1	10
5	4	1	5	11	4	6	18
6	1	0	3	10	3	1	6
7	1	2	2	6	3	1	8
8	2	2	1	11	0	1	5
9	2	0	0	11	3	1	7
10	2	1	4	9	0	4	5
11	3	1	2	9	3	2	4
12	1	3	2	10	0	7	2
13	0	0	0	0	0	0	0
14	1	0	1	15	3	0	5
	21	11	24	125	31	27	99

1914

TEST	-3	-2	-1	0	+1	+2	+3
1	1	3	3	3	2	4	7
2	1	4	1	8	4	3	3
3	1	3	2	9	0	4	6
4	3	3	0	9	0	9	1
5	2	8	1	8	2	16	13
6	0	3	1	8	1	6	5
7	1	1	0	5	2	7	7
8	4	1	0	8	0	4	5
9	1	5	1	4	0	5	8
10	0	4	0	8	0	6	7
11	3	0	0	10	0	4	7
12	3	6	0	5	0	6	5
13	1	3	0	13	0	6	2
14	1	1	0	12	0	3	8
	22	45	9	110	11	83	84

1918

TEST	-3	-2	-1	0	+1	+2	+3
1	4	0	1	6	1	3	8
2	5	1	4	4	1	1	8
3	5	0	3	6	0	0	11
4	4	2	0	9	0	0	10
5	7	0	3	11	1	2	25
6	10	0	0	5	2	0	7
7	5	2	2	6	0	0	8
8	0	0	3	12	0	1	6
9	5	1	1	8	0	0	9
10	3	0	1	10	0	0	11
11	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0
	48	6	18	77	5	7	103

1915

TEST	-3	-2	-1	0	+1	+2	+3
1	4	1	0	0	0	0	18
2	9	1	0	2	0	0	12
3	2	2	0	1	0	2	18
4	2	0	0	1	0	1	21
5	2	1	0	1	0	3	42
6	6	0	0	0	1	1	16
7	0	0	0	0	0	0	0
8	4	2	2	0	1	1	12
9	2	0	1	1	2	2	16
10	3	3	0	4	1	1	14
11	2	0	0	4	0	0	18
12	5	2	0	2	0	0	16
13	0	0	0	0	0	0	0
14	1	1	0	10	0	1	12
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

1713 → 1972-73

TEST	-3	-2	-1	0	+1	+2	+3
1	2	2	0	0	0	1	17
2	6	1	0	2	0	0	11
3	1	1	2	0	1	1	13
4	1	2	0	0	0	4	12
5	2	4	2	0	3	6	13
6	0	4	1	2	0	3	8
7	2	3	0	4	1	2	8
8	2	1	0	3	1	2	11
9	1	1	2	4	0	3	5
10	1	4	2	3	0	1	10
11	0	0	3	6	0	2	7
12	8	1	0	2	0	2	7
13	2	1	0	3	0	1	10
	28	25	12	29	6	28	132

Annexe 4 (Cf chap. 4)

1

TEST CODE	-3	-2	-1	0	+1	+2	+3
501	2	4	7	2	3	4	3
502	2	3	4	0	1	4	11
503	3	4	4	0	4	5	5
504	0	1	9	3	2	3	7
505	2	3	6	1	1	3	9
507	4	4	1	4	0	2	10
508	1	8	3	0	2	5	6
509	6	4	4	2	2	4	3
510	4	2	4	3	0	1	11
511	4	2	1	0	0	4	14
513	3	3	2	5	0	0	12
514	2	6	9	0	2	5	1
515	0	2	8	0	3	5	7
516	7	2	2	0	0	1	13
518	0	0	0	0	0	0	0
603	2	3	3	2	4	6	5
604	1	3	3	5	0	5	8
605	5	6	2	4	0	2	6
606	2	1	6	5	5	3	3
607	3	5	3	4	2	6	2
609	2	4	3	4	1	3	8
610	1	3	5	5	2	4	5
611	2	4	3	5	1	2	8
612	3	5	3	5	1	5	3
613	0	0	0	0	0	0	0
614	3	4	5	4	3	3	3
701	4	3	0	4	0	2	11
702	6	0	2	3	5	2	7
703	1	5	3	5	1	10	0
704	5	7	1	3	1	2	6
706	0	11	0	4	0	9	1
709	4	4	0	5	0	5	7
711	6	2	1	4	1	3	8
712	1	4	1	3	0	6	8
713	5	1	1	1	1	0	16
714	3	1	1	6	3	2	9
715	5	1	2	4	1	4	8
1201	1	0	2	0	1	2	16
1202	1	2	4	0	1	5	9
1205	1	0	2	1	2	0	16
1206	0	0	1	3	2	5	11
1209	1	5	5	2	5	3	1
1210	0	1	0	1	0	3	17
1211	1	0	0	4	0	2	13
1214	0	2	0	0	1	4	15
1215	2	0	0	1	1	0	18
1216	2	0	1	5	3	4	7
1217	0	1	0	0	1	3	17
1301	0	1	5	2	1	7	6
1302	1	1	0	3	0	3	14
1303	1	3	3	3	0	9	3
1304	2	2	0	2	0	4	12
1305	1	0	0	3	0	1	17
1306	3	4	1	5	0	2	7
1307	1	1	5	1	4	4	6
1308	1	1	2	0	1	5	12
1310	4	1	3	3	0	1	10
1311	0	2	4	1	3	3	9
1312	1	3	3	2	1	3	9
1313	2	0	3	3	0	2	12
1314	6	0	0	0	0	0	16
1701	1	0	4	0	0	2	15
1702	1	2	0	2	1	6	10
1703	2	0	1	6	2	0	11
1704	5	0	1	0	1	1	14
1706	1	1	3	0	2	5	10
1707	2	0	0	3	0	1	16
1708	0	2	2	4	1	7	6
1709	2	3	3	0	3	3	8
1710	10	1	0	0	0	2	9
1711	2	3	3	1	1	0	12
1712	4	4	1	0	1	1	11
1713	2	2	0	0	0	1	17

168173174166 91237656

2

TEST CODE	-3	-2	-1	0	+1	+2	+3
501	1	3	3	1	5	2	4
502	0	3	8	0	0	2	6
503	0	5	10	0	1	1	2
504	0	3	14	0	1	1	0
505	0	7	4	2	2	3	1
507	2	5	1	6	0	4	1
508	0	3	9	0	3	1	3
509	6	4	2	3	1	2	1
510	8	3	1	1	2	0	4
511	5	4	3	0	0	3	4
513	1	8	5	1	0	0	4
514	5	7	2	0	3	1	1
515	1	4	8	1	1	3	1
516	6	0	2	1	0	1	9
518	2	5	3	2	2	1	4
603	1	2	1	2	3	6	4
604	3	6	1	0	1	1	7
605	2	5	3	2	1	0	6
606	0	1	11	2	4	1	0
607	2	6	2	0	2	5	2
609	0	5	7	1	0	2	4
610	1	4	1	3	0	4	6
611	4	2	3	3	0	3	4
612	3	7	0	1	1	2	5
613	4	5	3	0	3	2	2
614	3	6	5	2	0	1	2
701	7	3	0	0	0	2	7
702	0	2	8	2	1	3	3
703	1	7	2	2	0	5	2
704	4	3	0	4	0	0	8
708	1	4	1	3	0	6	4
709	1	2	1	0	2	5	8
711	4	0	0	1	2	5	7
712	1	2	4	2	1	4	5
713	1	2	2	1	0	1	12
714	4	2	3	2	5	1	2
715	1	8	1	4	0	3	2
1201	5	3	5	0	0	5	2
1202	0	3	9	0	0	4	4
1205	6	1	0	1	0	1	11
1206	1	1	0	2	2	5	9
1209	8	0	1	1	2	0	8
1210	5	1	0	0	0	2	12
1211	5	2	2	1	0	2	8
1214	4	1	1	0	1	0	13
1215	5	1	0	3	2	0	9
1216	2	2	3	2	2	2	7
1217	1	1	3	0	0	2	13
1301	0	3	5	5	4	2	1
1302	6	0	0	4	0	0	10
1303	3	4	6	2	1	4	0
1304	3	1	0	4	0	2	10
1305	3	0	0	2	0	1	14
1306	5	3	1	4	0	0	7
1307	2	1	2	3	1	0	11
1308	4	1	2	2	0	2	9
1310	6	1	1	1	0	0	11
1311	2	0	2	3	2	1	10
1312	5	6	3	0	1	1	4
1313	5	0	0	4	0	0	11
1314	2	1	1	4	0	5	7
1701	5	1	0	4	0	0	10
1702	3	2	1	2	0	2	10
1703	3	0	1	4	1	1	10
1704	2	2	2	0	1	3	10
1706	2	1	4	0	0	4	9
1707	6	1	0	2	0	0	11
1708	5	3	1	4	0	4	3
1709	4	6	3	2	2	2	1
1710	3	3	2	1	0	1	10
1711	0	3	3	3	0	0	11
1712	4	1	2	4	0	0	9
1713	6	1	0	2	0	0	11

216209200131 69145453

3

TEST CODE	-3	-2	-1	0	+1	+2	+3
501	1	6	5	0	2	6	4
502	8	3	4	0	0	3	6
503	1	2	10	0	4	5	2
504	0	7	11	0	4	0	2
505	0	1	11	3	0	3	7
507	0	5	3	7	0	2	7
508	2	2	8	0	2	5	5
509	6	3	3	1	1	4	6
510	3	7	1	2	1	1	9
511	5	7	2	0	1	1	8
513	6	2	3	5	0	2	6
514	0	6	3	3	3	5	4
515	1	3	10	0	3	3	4
516	7	3	2	0	0	2	10
518	1	1	6	1	5	5	5
603	1	1	3	1	1	3	14
604	2	4	3	1	1	4	9
605	2	3	0	4	0	3	12
606	1	2	6	4	3	1	7
607	0	8	4	3	0	7	2
609	0	2	5	5	3	1	8
610	1	5	3	4	2	2	7
611	0	3	3	5	4	5	4
612	7	2	2	4	1	1	7
613	2	3	4	4	1	4	6
614	0	0	0	0	0	0	0
701	5	4	1	0	0	1	13
702	1	2	2	4	3	4	8
703	1	5	0	5	0	9	5
704	4	4	0	3	0	2	11
708	1	2	0	3	0	4	14
709	1	5	0	4	0	3	11
711	1	0	1	4	0	0	18
712	3	3	0	2	0	2	14
713	3	0	0	4	1	0	16
714	1	1	2	6	2	3	9
715	2	0	0	13	0	0	9
1201	2	2	2	2	2	1	8
1202	1	1	4	0	2	3	8
1205	1	5	1	1	0	2	9
1206	1	2	0	3	0	7	6
1209	3	3	2	1	1	4	5
1210	5	3	0	1	0	0	10
1211	2	0	3	1	2	1	10
1214	0	0	0	0	0	0	0
1215	1	0	2	1	0	0	15
1216	1	0	1	2	4	3	8
1217	1	0	3	0	3	3	9
1301	2	6	0	1	1	4	5
1302	2	0	0	2	0	1	14
1303	0	1	8	2	4	3	1
1304	1	1	1	4	0	3	9
1305	1	0	0	0	0	0	18
1306	7	1	1	0	0	1	0
1307	0	5	0	0	0	0	14
1308	1	0	0	0	0	0	18
1310	3	0	1	2	0	0	13
1311	0	6	0	2	0	5	6
1312	0	0	0	0	0	0	0
1313	0	0	0	3	0	6	10
1314	0	1	0	2	0	3	13
1701	4	1	1	2	1	1	9
1702	3	6	1	4	0	1	4
1703	1	1	1	4	2	4	6
1704	1	0	3	1	2	4	8
1706	0	4	1	1	1	4	8
1707	4	3	0	2	0	2	8
1708	0	3	3	4	4	4	1
1709	2	4	2	2	2	5	2
1710	3	1	1	0	2	5	7
1711	2	3	4	0	4	1	5
1712	2	2	1	4	1	1	8
1713	1	1	2	0	1	1	13

138132167159 87198508

4

TEST CODE	-3	-2	-1	0	+1	+2	+3
501	1	7	5	0	1	3	2
502	5	1	4	0	2	1	7
503	3	4	2	0	3	4	4
504	1	3	5	0	4	4	3
505	3	2	4	1	1	5	2
507	2	5	1	3	1	5	3
508	2	2	3	0	1	6	6
509	2	5	3	3	1	4	2
510	2	5	3	1	1	1	7
511	6	4	3	0	0	3	4
513	3	8	1	5	1	0	2
514	1	6	6	2	2	1	2
515	1	1	10	0	5	1	2
516	8	0	2	0	0	0	10
518	3	5	5	0	1	4	2
603	4	2	1	1	1	3	8
604	1	7	0	2	2	2	6
605	1	4	5	5	2	1	2
606	3	1	3	5	1	1	6
607	1	8	1	0	1	3	6
609	2	4	2	3	1	2	6
610	2	2	3	5	0	2	6
611	5	5	3	2	1	1	3
612	2	5	0	3	0	7	3
613	1	3	4	2	4	1	5
614	0	0	0	0	0	0	0
701	8	0	0	1	0	0	11
702	1	0	1	4	5	2	7
703	2	1	1	4	0	1	11
704	5	0	0	4	0	0	11
708	0	3	0	2	1	6	8
709	4	3	1	4	0	1	7
711	4	0	0	1	0	0	15
712	2	1	2	3	0	0	12
713	4	0	1	3	3	0	9
714	4	2	0	2	1	0	11
715	0	0	3	4	2	6	5
1201	2	4	3	4	0	3	3
1202	1	3	5	0	2	1	7
1205	4	3	0	3	0	2	7
1206	1	1	2	2	3	5	5
1209	2	4	3	2	1	3	4
1210	4	1	1	2	1	1	9
1211	3	4	3	1	2	1	5
1214	2	1	0	0	1	0	15
1215	6	0	0	3	0	0	10
1216	1	2	3	4	3	2	4
1217	2	0	3	1	3	3	7
1301	4	0	0	3	0	3	9
1302	1	3	4	0	1	1	9
1303	8	0	0	2	0	0	9
1304	3	0	0	1	1	0	14
1305	6	1	0	2	0	2	8
1306	2	3	3	3	3	1	4
1307	4	0	0	1	0	0	14
1308	3	0	0	0	0	0	16
1310	2	2	1	3	1	0	10
1311	1	4	3	2	0	4	5
1312	3	5	1	0	2	2	6
1313	3	1	0	2	0	1	12
1314	2	0	0	3	0	0	14
1701	6	0	1	3	1	0	8
1702	3	7	2	1	0	0	6
1703	4	5	2	0	1	6	1
1704	1	1	2	0	6	2	7
1706	1	4	2	3	1	4	4
1707	4	2	1	4	2	0	6
1708	0	1	3	4	7	3	1
1709	1	5	5	0	2	4	2
1710	4	3	1	1	1	3	6
1711	0	3	2	1	2	3	8
1712	2	3	0	4	0	3	7
1713	1	2	0	0	0	4	12

196167144140 98148491

5

TEST CODE	-3	-2	-1	0	+1	+2	+3
501	1	7	4	0	7	17	1
502	14	2	1	0	0	3	17
503	3	4	4	0	6	10	10
504	1	6	12	0	4	8	6
505	5	2	4	5	7	6	8
507	0	9	4	7	1	13	3
508	0	8	11	0	2	9	7
509	0	5	6	3	5	9	7
510	4	6	5	1	5	6	10
511	3	3	8	4	3	8	9
513	2	1	4	4	3	13	10
514	0	10	5	2	4	14	2
515	1	2	9	0	14	1	10
516	10	4	5	1	1	0	15
518	6	7	3	6	0	8	7
603	0	3	4	4	1	5	22
604	1	2	1	1	2	8	22
605	1	2	7	7	4	5	11
606	2	3	3	8	4	4	13
607	2	6	2	1	3	8	15
609	2	0	3	4	12	1	15
610	5	1	2	4	1	6	15
611	3	3	9	2	2	6	12
612	5	6	3	1	0	7	15
613	12	2	4	1	5	3	17
614	6	6	6	0	2	1	16
701	7	1	0	1	0	0	25
702	1	0	2	3	5	5	21
703	2	3	3	4	2	3	22
704	6	0	0	6	0	7	19
708	0	2	2	7	1	10	7
709	2	5	2	0	1	3	24
711	4	0	0	4	1	0	28
712	1	0	4	2	4	0	26
713	5	2	1	0	0	0	29
714	5	4	0	0	3	3	22
715	3	1	1	3	5	2	22
1201	1	8	5	4	2	5	5
1202	2	2	2	3	2	8	11
1205	6	5	1	2	1	0	15
1206	2	3	3	3	2	4	13
1209	6	11	2	0	1	4	6
1210	4	4	2	1	1	0	18
1211	2	8	6	2	2	2	8
1214	6	2	0	0	1	2	19
1215	8	0	2	3	1	0	16
1216	1	2	7	6	3	0	11
1217	2	2	5	0	4	2	15
1301	5	2	3	5	2	1	12
1302	3	0	3	4	0	0	20
1303	8	0	0	0	0	0	22
1304	2	0	2	0	2	0	24
1305	7	0	0	1	0	0	22
1306	6	0	3	4	1	1	15
1307	7	0	0	0	0	0	23
1308	6	0	0	0	1	0	23
1310	5	2	0	4	0	3	16
1311	2	3	3	3	3	7	9
1312	8	2	1	2	0	2	15
1313	1	4	0	4	0	0	21
1314	3	3	0	4	0	2	19
1701	4	0	5	4	6	1	10
1702	4	2	2	6	0	1	15
1703	4	0	5	6	5	1	9
1704	1	0	5	1	1	3	19
1706	2	6	2	1	0	6	13
1707	4	3	1	6	1	1	14
1708	0	3	6	7	6	5	3
1709	1	7	4	0	7	6	5
1710	0	5	1	5	1	1	17
1711	4	2	3	2	3	2	14
1712	5	2	1	2	3	1	16
1713	2	4	2	0	3	0	13

259225233191185297***

TEST CODE	6: -3	-2	-1	0	+1	+2	+3
501	0	6	5	1	5	7	6
502	1	6	8	0	4	5	6
503	1	7	8	0	1	4	3
504	1	2	10	0	2	2	7
505	1	3	5	6	2	1	5
507	0	2	8	6	3	5	0
508	1	2	4	0	2	4	2
509	1	5	7	3	1	0	7
510	1	4	9	2	3	3	2
511	0	7	5	2	3	2	5
513	4	5	4	4	1	3	3
514	0	0	0	0	0	0	0
515	1	0	14	0	5	4	0
516	11	1	3	2	1	0	6
518	0	9	1	3	0	8	3
603	1	5	3	1	2	3	9
604	4	6	1	2	0	2	9
605	3	3	4	2	3	1	8
606	3	1	4	4	5	1	6
607	2	3	1	2	1	6	3
609	0	0	0	0	0	0	0
610	4	7	1	3	1	2	6
611	1	5	7	2	1	2	6
612	3	5	4	0	1	3	8
613	6	2	2	2	2	1	9
614	6	2	2	2	1	1	10
701	7	0	0	1	0	2	14
702	0	2	5	4	2	2	9
703	2	4	1	2	0	2	13
704	9	0	0	2	0	1	12
708	2	5	0	3	0	10	4
709	0	2	4	4	2	2	10
711	8	0	0	3	0	0	13
712	4	0	3	4	2	1	10
713	11	0	0	2	0	1	10
714	1	4	2	4	2	2	9
715	7	0	1	3	0	1	12
1201	1	4	0	3	1	3	6
1202	0	2	5	0	4	0	7
1205	0	0	0	0	0	0	0
1206	1	1	2	4	2	3	5
1209	2	6	0	0	1	1	8
1210	1	5	1	0	1	1	9
1211	0	3	3	0	2	4	6
1214	1	4	2	0	0	2	9
1215	0	0	0	0	0	0	0
1216	0	0	0	0	0	0	0
1217	2	0	2	2	3	1	8
1301	0	0	0	0	0	0	0
1302	6	0	0	1	0	0	9
1303	8	0	0	0	0	0	10
1304	4	0	2	0	0	0	12
1305	6	0	0	0	0	0	12
1306	4	2	0	0	1	2	9
1307	7	0	0	0	0	0	11
1308	7	0	1	0	0	0	10
1310	0	1	0	0	3	1	13
1311	0	0	0	0	0	0	0
1312	1	3	3	0	0	4	7
1313	4	3	0	2	0	0	9
1314	1	1	1	2	0	0	13
1701	6	1	1	4	1	2	5
1702	2	1	0	3	0	0	12
1703	2	2	2	2	4	2	3
1704	2	2	2	2	3	2	5
1706	2	4	1	2	2	2	5
1707	0	0	0	0	0	0	0
1708	0	3	6	2	2	5	0
1709	0	0	0	0	0	0	0
1710	2	2	0	1	3	3	7
1711	5	3	1	3	0	1	5
1712	5	2	0	3	0	1	7
1713	0	4	1	2	0	3	8

101167174110 01134444

TEST CODE	7: -3	-2	-1	0	+1	+2	+3
501	0	0	0	0	0	0	0
502	0	9	4	0	1	7	1
503	0	3	9	0	4	4	2
504	0	2	5	0	3	7	5
505	0	2	8	3	6	1	2
507	0	1	8	7	3	3	0
508	0	7	0	3	0	10	2
509	0	1	7	2	12	0	0
510	1	3	5	1	2	9	1
511	2	2	5	6	2	2	3
513	2	4	2	3	2	4	5
514	1	8	4	1	1	5	2
515	0	0	10	0	11	1	0
516	7	0	4	3	1	0	7
518	0	8	0	2	0	11	1
603	1	3	1	3	3	6	5
604	1	1	2	0	3	6	9
605	3	3	4	0	1	4	7
606	0	3	6	1	5	5	2
607	1	5	2	3	2	6	3
609	2	0	6	1	3	1	9
610	1	2	1	4	2	4	8
611	0	8	3	2	0	4	5
612	4	5	1	0	0	4	8
613	6	6	2	1	2	1	4
614	0	5	4	1	2	4	6
701	4	1	0	0	0	0	17
702	4	0	5	3	4	3	3
703	1	4	0	1	2	3	11
704	3	4	1	3	0	7	4
708	0	5	1	7	1	7	1
709	0	6	1	3	2	3	7
711	3	0	2	4	2	4	7
712	1	1	1	0	6	5	8
713	1	1	0	2	1	1	16
714	2	6	1	2	0	8	3
715	1	2	0	2	2	1	14
1201	1	2	3	1	2	2	9
1202	1	2	4	0	3	1	9
1205	1	2	1	0	0	3	13
1206	1	1	0	0	1	7	10
1209	7	4	1	0	0	1	7
1210	3	0	1	0	0	2	14
1211	4	1	5	0	1	2	7
1214	6	3	0	0	0	2	9
1215	0	0	0	0	0	0	0
1216	1	0	2	4	2	3	8
1217	1	0	3	0	1	1	14
1301	2	0	0	0	1	0	17
1302	2	1	0	0	0	3	14
1303	3	0	0	0	0	0	17
1304	2	0	0	0	0	0	18
1305	3	0	0	0	0	0	17
1306	6	0	0	0	0	0	14
1307	0	0	0	0	0	0	0
1308	4	0	0	0	0	0	16
1310	1	0	0	1	0	1	17
1311	1	2	0	2	1	2	12
1312	0	0	0	0	0	0	0
1313	0	0	2	0	0	2	16
1314	1	0	0	3	0	0	16
1701	6	0	4	4	1	0	5
1702	3	2	0	3	0	4	8
1703	3	2	2	3	2	3	5
1704	1	0	0	2	1	5	11
1706	0	3	1	3	1	4	8
1707	3	1	2	2	1	0	11
1708	0	0	6	4	1	5	0
1709	0	0	0	0	0	0	0
1710	2	3	0	1	3	1	10
1711	7	1	3	4	0	2	3
1712	7	1	0	4	1	0	7
1713	2	3	0	4	1	2	8

137155155119115213538

TEST CODE	8: -3	-2	-1	0	+1	+2	+3
501	1	2	3	0	4	2	10
502	3	0	3	0	0	0	16
503	2	4	4	0	0	1	9
504	4	3	2	0	2	3	8
505	2	2	0	2	2	2	12
507	0	0	6	0	3	6	7
508	0	4	2	1	2	7	6
509	0	2	3	1	2	2	12
510	3	1	2	1	3	2	10
511	2	3	2	0	2	3	10
513	3	3	1	1	0	4	10
514	3	6	1	0	1	3	8
515	1	0	5	0	4	2	10
516	3	0	2	2	2	0	13
518	2	5	0	2	0	4	9
603	2	4	0	0	2	4	10
604	1	2	2	1	1	1	14
605	3	2	1	2	1	0	13
606	1	2	2	2	1	3	11
607	3	0	3	5	1	4	6
609	0	0	2	1	6	1	12
610	1	2	0	3	1	1	14
611	7	4	1	0	0	3	7
612	2	2	1	3	0	5	9
613	3	1	2	2	0	1	13
614	3	3	1	2	2	1	10
701	4	0	0	0	0	0	18
702	2	1	3	4	1	0	11
703	0	2	0	3	0	1	16
704	3	0	0	2	0	1	16
708	0	4	0	2	1	5	10
709	2	1	1	3	0	3	12
711	2	0	1	3	0	2	14
712	5	1	1	1	0	0	14
713	4	0	0	3	0	0	15
714	1	2	0	2	0	4	13
715	0	2	0	3	1	0	16
1201	3	2	1	0	1	5	8
1202	3	2	3	0	1	3	8
1205	4	0	1	1	2	3	9
1206	0	1	2	3	5	4	5
1209	3	3	4	2	2	0	6
1210	4	2	0	0	1	1	12
1211	4	1	4	0	1	3	7
1214	3	3	1	0	0	2	11
1215	6	0	1	2	0	1	10
1216	0	1	2	4	1	3	9
1217	2	0	5	0	2	4	7
1301	4	4	1	4	0	2	5
1302	7	0	0	1	0	2	13
1303	5	2	2	2	0	1	8
1304	3	0	0	1	0	0	16
1305	5	0	0	3	0	0	12
1306	4	1	2	0	1	0	12
1307	6	1	0	0	0	2	11
1308	0	0	0	0	0	0	0
1310	5	0	1	1	0	0	13
1311	3	3	1	2	0	4	7
1312	5	1	0	0	0	0	14
1313	4	1	0	1	3	2	9
1314	3	1	0	2	0	4	10
1701	0	1	2	4	1	1	11
1702	3	1	0	4	1	2	9
1703	1	5	1	4	0	2	7
1704	2	1	1	2	0	3	11
1706	0	1	2	0	4	7	6
1707	3	0	3	4	3	0	7
1708	1	3	2	4	4	4	2
1709	0	2	5	0	4	1	6
1710	2	2	2	1	0	2	13
1711	1	1	4	2	0	4	8
1712	2	3	0	2	0	1	12
1713	2	1	0	3	1	2	11

131120110116 83156746

TEST CODE	9: -3	-2	-1	0	+1	+2	+3
501	1	2	2	2	1	3	11
502	0	2	7	1	1	2	14
503	0	4	6	1	2	4	5
504	2	3	4	1	0	8	4
505	2	1	0	2	0	5	12
507	0	1	3	4	8	5	1
508	0	4	1	3	1	7	6
509	0	3	3	5	4	2	5
510	0	2	3	5	3	9	0
511	4	0	9	0	2	3	4
513	0	2	4	3	1	3	9
514	0	9	7	0	1	3	2
515	0	0	6	0	16	0	0
516	3	0	3	4	1	2	9
518	2	4	1	4	0	5	6
603	1	3	3	2	0	2	11
604	3	5	4	3	2	1	4
605	5	0	2	7	0	2	6
606	1	4	4	7	2	3	1
607	2	4	3	3	1	6	3
609	1	0	6	6	3	0	6
610	6	3	1	4	0	3	5
611	1	7	5	5	1	3	0
612	3	11	0	3	0	3	2
613	5	3	3	4	0	2	5
614	4	2	5	1	1	1	8
701	3	0	0	1	0	1	17
702	0	1	4	4	3	1	9
703	0	1	0	2	0	5	14
704	5	2	0	2	0	2	11
708	1	4	1	4	0	1	11
709	1	5	2	3	2	5	4
711	3	0	1	3	0	0	15
712	0	1	1	5	2	1	12
713	1	0	2	3	1	0	15
714	4	1	1	8	0	1	7
715	0	0	1	5	0	4	12
1201	2	2	2	3	0	5	2
1202	0	3	2	2	1	4	4
1205	2	4	1	1	0	2	6
1206	0	3	1	2	1	6	3
1209	1	1	8	0	0	0	6
1210	2	4	2	1	0	2	5
1211	2	2	4	0	2	3	3
1214	3	3	2	0	0	1	7
1215	9	0	0	0	1	0	6
1216	2	3	3	0	1	2	5
1217	1	1	3	1	3	2	5
1301	5	3	1	0	0	1	6
1302	8	0	0	0	0	1	7
1303	6	1	0	0	0	1	8
1304	6	1	0	0	0	1	8
1305	6	0	0	3	0	0	7
1306	7	0	0	0	1	3	5
1307	4	4	1	1	0	1	5
1308	8	0	2	0	1	0	5
1310	6	0	1	2	0	0	7
1311	6	1	0	1	1	1	6
1312	10	0	0	0	0	0	6
1313	8	0	0	1	0	0	7
1314	3	1	1	3	0	2	7
1701	3	0	5	2	1	1	4
1702	5	1	1	1	0	3	5
1703	7	0	2	2	2	0	3
1704	1	1	2	2	2	2	6
1706	2	3	1	0	2	6	2
1707	1	1	4	4	3	1	2
1708	0	2	6	3	1	3	1
1709	2	5	4	3	1	0	1
1710	4	3	2	0	0	2	5
1711	2	5	1	2	1	1	4
1712	5	3	0	4	0	0	4
1713	1	1	2	4	0	3	5

203156147143 84148444

TEST CODE	10: -3	-2	-1	0	+1	+2	+3
501	2	2	3	2	0	1	10
502	4	0	1	2	0	1	12
503	1	6	3	0	1	6	3
504	0	4	3	1	4	1	7
505	2	1	2	1	4	0	11
507	0	1	6	3	2	6	2
508	1	1	2	3	2	6	8
509	1	5	2	4	1	2	5
510	0	1	1	4	1	8	5
511	0	3	4	4	5	1	3
513	1	1	5	1	1	1	10
514	1	6	5	0	2	5	1
515	1	0	11	0	6	2	0
516	3	1	4	3	3	1	5
518	1	9	0	1	0	5	4
603	4	1	2	3	1	5	4
604	1	0	3	2	2	3	9
605	2	1	0	2	2	3	10
606	1	5	4	4	0	4	2
607	2	4	1	4	1	5	3
609	2	0	1	4	3	0	10
610	3	2	5	3	1	0	6
611	3	3	4	3	2	0	5
612	7	4	1	1	0	3	4
613	5	3	0	3	2	0	7
614	4	1	2	3	1	1	8
701	6	2	2	0	0	2	8
702	1	2	3	3	2	3	6
703	1	1	0	2	0	8	8
704	4	1	0	5	0	2	8
708	4	1	0	3	0	3	9
709	1	1	5	3	1	3	6
711	1	1	0	2	2	0	14
712	2	2	1	4	1	4	6
713	4	0	0	3	2	0	11
714	1	1	2	6	0	2	8
715	1	2	0	2	0	1	14
1201	3	2	0	0	2	8	5
1202	2	1	1	1	5	7	4
1205	0	1	1	2	4	3	10
1206	1	2	2	1	3	4	8
1209	1	1	7	0	5	2	5
1210	2	3	0	0	0	3	13
1211	2	2	2	1	4	6	4
1214	3	5	0	0	0	4	9
1215	7	0	0	0	1	0	13
1216	0	0	1	3	3	4	13
1217	1	1	3	0	8	1	7
1301	3	0	0	0	0	0	18
1302	3	0	3	3	0	0	12
1303	8	1	0	0	2	1	9
1304	1	0	1	1	2	1	15
1305	4	0	0	1	0	0	16
1306	7	1	2	1	0	2	8
1307	3	4	0	0	0	5	9
1308	7	2	0	0	2	1	9
1310	5	0	0	1	0	0	15
1311	2	2	2	3	1	2	9
1312	5	0	0	1	0	0	15
1313	2	0	0	2	2	0	15
1314	0	0	0	0	0	0	0
1701	2	0	3	5	3	1	7
1702	3	1	1	2	2	4	8
1703	1	1	5	4	3	1	6
1704	1	1	2	4	3	2	8
1706	1	3	3	0	1	6	7
1707	1	1	4	5	3	0	7
1708	0	3	4	4	5	4	1
1709	1	2	2	4	2	6	4
1710	4	4	2	0	0	0	11
1711	3	3	0	3	0	2	10
1712	2	5	0	4	0	2	12
1713	1	4	2	3	0	3	10

170131141153121181573

502	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
503	1	4	3	2	1	8	2
504	3	1	2	1	3	5	4
505	2	2	1	0	1	4	11
507	0	3	1	4	7	5	1
508	3	3	1	1	4	3	6
509	1	2	3	4	2	4	5
510	3	1	1	0	2	6	8
511	3	2	4	3	2	4	3
513	3	2	0	2	0	7	7
514	2	8	4	2	0	3	2
515	2	0	6	1	5	2	5
516	4	0	4	5	1	0	7
518	3	6	0	1	0	10	1
603	2	3	1	3	1	4	7
604	3	2	3	1	1	4	7
605	2	2	1	2	1	6	7
606	0	0	4	4	6	2	5
607	1	4	1	4	2	3	6
609	1	0	4	4	4	0	8
610	3	4	0	4	0	2	8
611	4	0	5	4	3	1	4
612	3	3	0	4	0	1	10
613	6	2	3	3	0	1	6
614	3	1	4	2	1	3	7
701	9	0	0	1	0	0	11
702	0	2	5	2	5	1	6
703	1	3	0	4	1	9	3
704	8	0	0	3	0	3	7
708	0	0	0	0	0	0	0
709	3	1	2	3	0	7	5
711	4	1	0	2	1	2	11
712	2	1	1	3	2	6	6
713	2	0	1	1	2	0	15
714	3	4	0	4	4	3	3
715	1	3	0	4	1	2	10
1201	2	0	1	1	0	6	8
1202	2	2	1	1	1	2	9
1205	0	1	2	1	0	1	13
1206	1	0	0	2	0	0	15
1209	1	1	4	0	4	2	6
1210	1	2	0	0	0	1	14
1211	0	0	3	0	3	5	7
1214	1	2	2	0	0	2	11
1215	1	0	1	1	3	0	12
1216	0	0	2	3	1	1	11
1217	1	0	4	0	0	4	9
1301	1	0	0	0	1	0	16
1302	1	1	0	1	5	1	9
1303	5	0	0	0	0	0	13
1304	1	0	2	0	3	0	12
1305	3	0	0	1	0	0	14
1306	2	0	1	0	1	3	11
1307	2	0	0	0	1	0	15
1308	1	2	0	0	1	1	13
1310	3	0	0	3	0	0	12
1311	2	1	0	1	1	0	13
1312	3	0	0	1	0	0	14
1313	0	0	2	0	5	1	10
1314	1	2	0	2	0	3	10
1701	2	0	4	2	2	1	7
1702	2	1	3	2	1	1	8
1703	3	1	5	2	0	3	4
1704	1	2	2	2	4	0	7
1706	1	1	4	0	2	4	6
1707	0	0	3	3	1	1	10
1708	0	2	4	3	3	4	2
1709	1	3	3	0	3	1	7
1710	3	4	2	0	0	0	9
1711	0	1	2	2	1	2	10
1712	6	0	0	1	0	0	11
1713	0	0	3	6	0	2	7

151105128133112182593

502	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
503	0	[illegible]	13	0	2	3	0
504	0	2	13	0	1	2	2
505	1	5	1	4	0	4	5
507	1	0	7	5	2	3	2
508	2	0	4	4	5	3	2
509	1	1	1	9	2	1	5
510	3	7	3	2	1	2	2
511	1	0	6	8	2	0	3
513	5	7	2	1	1	1	3
514	3	2	8	0	2	4	1
515	0	0	10	3	5	0	2
516	2	1	11	0	2	2	2
518	0	4	3	4	3	6	0
603	2	7	1	5	0	2	3
604	2	5	4	3	0	2	4
605	5	2	6	1	2	1	3
606	0	3	4	6	2	1	4
607	0	7	2	5	0	4	2
609	2	0	8	5	3	0	2
610	6	3	3	4	1	2	1
611	4	0	4	5	4	0	3
612	3	5	1	5	0	3	3
613	2	3	6	3	2	1	3
614	7	1	3	4	2	0	3
701	9	0	0	3	0	0	8
702	3	5	2	3	1	1	5
703	1	9	0	4	0	6	0
704	7	2	0	5	0	2	4
708	0	0	0	0	0	0	0
709	1	5	5	3	1	2	3
711	2	3	0	8	1	2	4
712	3	1	3	5	1	0	7
713	6	1	0	4	0	1	8
714	0	0	0	0	0	0	0
715	1	5	2	7	0	3	2
1201	1	2	2	3	1	6	5
1202	2	3	3	1	1	2	8
1205	5	3	0	1	0	4	7
1206	0	4	3	3	1	7	2
1209	3	4	7	0	1	3	2
1210	4	6	1	0	0	2	7
1211	1	1	4	2	0	5	7
1214	5	2	0	0	0	2	11
1215	8	0	0	1	1	0	10
1216	2	1	2	3	2	2	8
1217	1	2	3	1	3	1	9
1301	3	2	5	5	2	1	2
1302	6	1	1	2	1	0	9
1303	7	1	1	2	0	0	9
1304	4	3	2	2	1	2	6
1305	7	0	0	2	0	0	11
1306	9	0	0	4	0	0	7
1307	6	4	3	1	1	2	3
1308	7	2	3	1	0	1	6
1310	5	0	2	2	1	2	8
1311	2	4	1	3	1	3	6
1312	8	0	0	2	0	0	10
1313	4	0	4	2	1	0	9
1314	6	2	2	4	1	0	5
1701	1	1	5	2	6	2	3
1702	3	4	0	2	1	5	5
1703	3	5	3	0	2	2	5
1704	1	1	3	4	3	3	5
1706	2	6	2	1	1	2	6
1707	3	0	3	3	5	0	6
1708	0	2	3	4	3	8	0
1709	2	8	1	0	2	2	5
1710	3	2	1	0	2	3	9
1711	5	3	1	3	2	0	6
1712	4	3	0	4	0	2	7
1713	8	1	0	2	0	2	7

227200217204 98142332

502	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
503	0	4	6	3	0	2	5
504	0	2	3	2	3	[illegible]	7
505	1	1	1	3	2	2	[illegible]
507	0	0	4	1	6	2	7
508	0	4	1	1	4	10	0
509	0	0	3	2	3	1	11
510	2	1	2	3	1	1	[illegible]
511	1	0	6	2	5	3	3
513	0	0	0	0	0	0	[illegible]
514	0	1	5	1	4	8	1
515	0	0	3	2	7	5	3
516	1	0	4	3	3	1	8
518	0	6	2	3	1	3	5
603	4	1	1	2	1	8	3
604	3	1	2	3	2	1	[illegible]
605	0	1	4	3	2	0	[illegible]
606	1	1	3	4	1	5	5
607	0	1	4	5	0	6	4
609	1	0	4	4	3	0	9
610	3	4	0	4	0	3	6
611	4	0	2	4	2	0	8
612	2	2	0	6	0	2	8
613	2	4	5	1	1	0	7
614	2	0	2	4	3	2	7
701	4	0	0	4	0	0	12
702	1	1	2	2	2	3	9
703	4	2	0	3	0	2	9
704	0	0	0	0	0	0	0
708	0	0	0	0	0	0	0
709	0	5	3	3	2	2	5
711	2	0	4	2	0	2	[illegible]
712	1	0	2	3	2	1	11
713	1	1	0	4	0	1	13
714	1	0	3	4	2	5	5
715	0	1	1	4	1	4	9
1201	0	1	2	2	2	5	5
1202	0	2	3	0	1	4	7
1205	0	0	1	1	2	5	8
1206	1	0	0	5	1	4	6
1209	1	4	6	0	3	1	2
1210	1	1	0	0	0	2	13
1211	1	3	4	1	1	5	7
1214	1	1	0	0	0	0	15
1215	1	0	1	4	0	1	[illegible]
1216	1	0	0	3	3	2	8
1217	0	0	1	2	5	2	7
1301	0	0	0	0	0	0	0
1302	3	0	1	1	2	2	8
1303	0	0	0	0	0	0	0
1304	5	0	0	0	0	0	12
1305	1	0	0	3	0	0	13
1306	2	2	0	1	1	1	[illegible]
1307	3	1	0	1	2	1	9
1308	2	2	0	1	1	3	8
1310	1	0	0	3	0	1	12
1311	1	0	1	2	1	1	11
1312	3	0	0	1	0	0	13
1313	4	2	0	0	0	0	11
1314	1	0	3	2	1	0	10
1701	1	2	2	2	0	2	8
1702	1	1	3	1	0	1	10
1703	2	3	5	3	1	0	3
1704	0	1	1	3	4	2	[illegible]
1706	2	3	0	1	0	3	8
1707	0	0	3	3	3	1	7
1708	0	4	2	3	2	3	3
1709	0	1	1	3	1	4	7
1710	1	1	0	2	0	0	13
1711	4	1	1	3	0	0	8
1712	4	4	0	2	0	1	6
1713	2	1	0	3	0	1	10

91 90129159106158525

1972-73

501

TEST	−3	−2	−1	0	+1	+2	+3
1	2	4	7	2	3	4	3
2	1	3	3	1	5	2	4
3	1	6	5	0	2	6	4
4	1	7	5	0	1	3	3
5	1	7	4	0	7	17	1
6	0	6	5	1	5	7	0
7	0	0	0	0	0	0	0
8	1	2	3	0	4	2	10
9	1	2	2	2	1	3	11
10	2	2	3	2	0	1	10
11	2	4	2	1	2	5	5
12	0	6	6	4	3	1	0
13	0	2	5	1	5	7	0
	12	51	50	14	38	58	51

502

TEST	−3	−2	−1	0	+1	+2	+3
1	2	3	4	0	1	4	11
2	0	3	8	0	0	2	4
3	8	3	4	0	0	3	6
4	5	1	4	0	2	1	7
5	14	2	1	0	0	3	17
6	1	6	8	0	4	5	[illegible]
7	0	9	4	0	1	7	1
8	5	0	3	0	0	0	14
9	0	2	2	1	1	2	14
10	4	0	1	2	0	1	12
11	2	0	1	3	0	4	9
12	1	13	4	0	1	1	[illegible]
13	0	3	1	1	1	5	9
	42	45	45	7	11	38	108

503

TEST	−3	−2	−1	0	+1	+2	+3
1	3	4	4	0	4	5	5
2	0	5	10	0	1	1	2
3	1	2	10	0	4	5	2
4	3	4	2	0	3	4	4
5	3	4	4	0	6	10	10
6	1	7	8	0	1	4	3
7	0	3	9	0	4	4	2
8	2	4	4	0	0	3	9
9	0	4	6	1	2	4	5
10	1	6	3	0	1	6	3
11	1	4	3	2	1	8	2
12	0	2	13	0	2	3	0
13	0	4	6	3	0	2	5
	15	53	82	6	29	59	52

504

TEST	−3	−2	−1	0	+1	+2	+3
1	0	1	9	3	2	3	7
2	0	3	14	0	1	1	0
3	0	7	11	0	4	0	2
4	1	3	5	0	4	4	3
5	1	6	12	0	4	8	6
6	1	2	10	0	2	2	7
7	0	2	5	0	3	7	5
8	4	3	2	0	2	3	8
9	2	3	4	1	0	8	4
10	0	4	3	1	4	1	7
11	3	3	2	1	3	5	4
12	0	2	13	0	1	2	2
13	0	2	3	2	3	3	7
	12	41	93	8	33	47	62

505

TEST	−3	−2	−1	0	+1	+2	+3
1	2	3	6	1	1	3	9
2	0	7	4	2	2	3	1
3	0	1	8	3	0	3	9
4	3	2	4	1	3	5	2
5	5	2	4	5	7	6	8
6	1	3	5	6	2	1	6
7	0	2	8	3	5	1	2
8	2	2	0	2	2	2	12
9	2	1	0	2	1	5	12
10	2	1	2	1	4	0	10
11	2	2	1	0	1	4	11
12	1	5	1	4	0	4	5
13	1	1	1	3	2	2	10
	21	32	44	33	30	39	97

507

TEST	−3	−2	−1	0	+1	+2	+3
1	4	4	1	4	0	2	10
2	2	5	1	6	0	4	1
3	0	5	3	7	0	2	7
4	2	5	1	3	1	5	3
5	0	9	4	7	1	13	3
6	0	2	8	6	3	5	0
7	0	1	8	7	3	3	0
8	0	0	6	0	3	6	7
9	0	1	3	4	8	5	1
10	0	1	6	3	2	6	2
11	0	3	1	4	7	5	1
12	1	0	7	5	2	3	2
13	0	0	4	1	6	2	7
	9	36	53	57	36	61	44

508

TEST	−3	−2	−1	0	+1	+2	+3
1	1	8	3	0	2	5	6
2	0	3	9	0	3	1	3
3	2	2	8	0	2	5	5
4	2	2	3	0	1	6	6
5	0	8	11	0	2	9	7
6	1	9	6	0	2	4	2
7	0	7	0	3	0	10	2
8	0	4	2	1	2	7	6
9	0	4	1	3	1	7	6
10	1	1	2	3	2	6	5
11	3	3	1	1	4	3	6
12	2	0	4	4	5	3	2
13	0	4	1	1	4	10	0
	12	55	51	16	30	76	56

509

TEST	−3	−2	−1	0	+1	+2	+3
1	6	4	4	2	2	4	3
2	6	4	2	3	1	2	1
3	6	3	3	1	1	4	6
4	2	5	3	3	1	4	2
5	0	5	8	3	5	9	7
6	1	5	7	3	1	0	7
7	0	1	7	2	12	0	0
8	0	2	3	1	2	2	12
9	0	3	3	5	4	2	5
10	1	5	2	4	1	2	5
11	1	2	3	4	2	4	5
12	1	1	1	9	2	1	5
13	0	0	3	2	3	1	11
	24	40	49	42	37	35	69

510

TEST	−3	−2	−1	0	+1	+2	+3
1	4	2	4	3	0	1	11
2	8	3	1	1	2	0	4
3	3	7	1	2	1	1	9
4	2	5	3	1	1	1	7
5	4	6	5	1	5	6	10
6	1	4	9	2	3	3	2
7	1	3	5	1	2	9	1
8	3	1	2	1	3	2	10
9	0	2	3	5	3	9	0
10	0	1	1	4	1	8	5
11	3	1	1	0	2	6	8
12	3	7	3	2	1	2	2
13	2	1	2	3	1	1	10
	34	43	40	26	25	49	79

511

TEST	−3	−2	−1	0	+1	+2	+3
1	4	2	1	0	0	4	14
2	5	4	3	0	0	3	4
3	5	7	2	0	1	1	8
4	6	4	3	0	0	3	4
5	3	3	8	4	3	8	8
6	0	7	5	2	3	2	5
7	2	2	5	6	2	2	3
8	2	3	2	0	2	3	10
9	4	0	9	0	2	3	4
10	0	3	4	4	5	1	3
11	3	2	4	3	2	4	3
12	1	0	6	8	2	0	3
13	1	0	6	2	5	3	3
	36	37	58	29	27	37	72

513

TEST	−3	−2	−1	0	+1	+2	+3
1	3	3	2	5	0	0	12
2	1	8	5	1	0	0	4
3	6	2	3	5	0	2	6
4	3	8	1	5	1	0	2
5	2	1	4	4	3	13	10
6	4	5	4	4	1	3	3
7	2	4	2	3	2	4	5
8	3	3	1	1	0	4	10
9	0	2	4	3	1	3	9
10	1	1	5	1	1	1	10
11	3	2	0	2	0	7	7
12	5	7	2	1	1	1	3
13	0	0	0	0	0	0	0
	33	46	33	35	10	38	81

514

TEST	−3	−2	−1	0	+1	+2	+3
1	2	6	9	0	2	5	1
2	5	7	2	0	3	1	1
3	0	6	3	3	3	5	4
4	1	6	6	2	2	1	2
5	0	10	5	2	4	14	2
6	0	0	0	0	0	0	0
7	1	8	4	1	1	5	2
8	3	6	1	0	1	3	8
9	0	9	7	0	1	3	2
10	1	6	5	0	2	5	1
11	2	8	4	2	0	3	2
12	3	2	8	0	2	4	1
13	0	1	5	1	4	8	1
	18	75	59	11	25	57	27

515

TEST	−3	−2	−1	0	+1	+2	+3
1	0	2	9	0	3	5	7
2	1	4	8	1	1	3	1
3	1	3	10	0	3	3	4
4	1	1	10	0	5	1	2
5	1	2	9	0	14	1	10
6	1	0	14	0	5	4	0
7	0	0	10	0	11	1	0
8	1	0	5	0	4	2	10
9	0	0	6	0	16	0	0
10	1	0	11	0	6	2	0
11	2	0	6	1	5	2	5
12	0	0	10	3	5	0	2
13	0	0	3	2	7	5	3
	9	12	110	7	85	29	44

516

TEST	−3	−2	−1	0	+1	+2	+3
1	7	2	2	0	0	1	13
2	6	0	2	1	0	1	9
3	7	3	2	0	0	2	10
4	8	0	2	0	0	0	10
5	10	4	5	1	1	0	16
6	11	1	3	2	1	0	6
7	7	0	4	3	1	0	7
8	3	0	2	2	2	0	13
9	3	0	3	4	1	2	9
10	3	1	4	3	3	1	5
11	4	0	4	5	1	0	7
12	2	1	11	0	2	2	2
13	1	0	4	3	3	1	8
	72	12	48	24	15	10	115

518

TEST	−3	−2	−1	0	+1	+2	+3
1	0	0	0	0	0	0	0
2	2	5	3	2	2	1	4
3	1	1	6	1	5	5	5
4	3	5	5	0	1	4	2
5	6	7	3	6	0	8	7
6	0	9	1	3	0	8	3
7	0	8	0	2	0	11	1
8	2	5	0	2	0	4	9
9	2	4	1	4	0	5	6
10	1	9	0	1	0	5	4
11	3	6	0	1	0	10	1
12	0	4	3	4	3	6	0
13	0	6	2	3	1	3	5
	20	69	24	29	12	70	47

603

TEST	−3	−2	−1	0	+1	+2	+3
1	2	3	3	2	4	6	5
2	1	2	1	2	3	6	4
3	1	1	3	1	1	3	14
4	4	2	1	1	1	3	8
5	0	3	4	4	1	5	20
6	1	5	3	1	2	3	9
7	1	3	1	3	3	6	5
8	2	4	0	0	2	4	10
9	1	3	3	2	0	2	11
10	4	1	2	3	1	5	4
11	2	3	1	3	1	4	7
12	2	7	1	5	0	2	3
13	4	1	1	2	1	8	3
	25	38	24	29	20	57	103

604

TEST	−3	−2	−1	0	+1	+2	+3
1	1	3	3	5	0	5	8
2	3	6	1	0	1	1	7
3	2	4	3	1	1	4	9
4	1	7	0	2	2	2	6
5	1	2	1	1	2	8	22
6	4	6	1	2	0	2	9
7	1	1	2	0	3	6	9
8	1	2	2	1	1	1	14
9	3	5	4	3	2	1	4
10	1	0	3	2	2	3	9
11	3	2	3	1	1	4	7
12	2	5	4	3	0	2	4
13	3	1	2	3	2	1	8
	26	44	29	24	17	40	116

605

TEST	−3	−2	−1	0	+1	+2	+3
1	5	6	2	4	0	2	6
2	2	5	3	2	1	0	6
3	2	3	0	4	0	3	12
4	1	4	5	5	2	1	2
5	1	2	7	7	4	5	11
6	3	3	4	2	3	1	8
7	3	3	4	0	1	4	7
8	3	2	1	2	1	0	13
9	5	0	2	7	0	2	6
10	2	1	0	2	2	3	10
11	2	2	1	2	1	6	7
12	5	2	6	1	2	1	3
13	0	1	4	3	2	0	10
	34	34	39	41	19	28	101

606

TEST	−3	−2	−1	0	+1	+2	+3
1	2	1	6	5	5	3	3
2	0	1	11	2	4	1	0
3	1	2	6	4	3	1	7
4	3	1	3	5	1	1	6
5	2	3	3	8	4	4	13
6	3	1	4	4	5	1	6
7	0	3	6	1	5	5	2
8	1	2	2	2	1	3	11
9	1	4	4	7	2	3	1
10	1	5	4	4	0	4	2
11	0	0	4	4	6	2	5
12	0	3	4	6	2	1	4
13	1	1	3	4	1	5	5
	15	27	60	56	39	34	65

607

TEST	−3	−2	−1	0	+1	+2	+3
1	3	5	3	4	2	6	2
2	2	6	2	0	2	5	2
3	0	8	4	3	0	7	2
4	1	8	1	0	1	3	6
5	2	6	2	1	3	8	15
6	2	9	1	2	1	6	3
7	1	5	2	3	2	6	3
8	3	0	3	5	1	4	6
9	2	4	3	3	1	6	3
10	2	4	1	4	1	5	3
11	1	4	1	4	2	3	6
12	0	7	2	5	0	4	2
13	0	1	4	5	0	6	4
	19	67	29	39	16	69	57

609

TEST	−3	−2	−1	0	+1	+2	+3
1	2	4	3	4	1	3	8
2	0	5	7	1	0	2	4
3	0	2	5	5	3	1	8
4	2	4	2	3	1	2	6
5	2	0	3	4	12	1	15
6	0	0	0	0	0	0	0
7	2	0	6	1	3	1	9
8	0	0	2	1	6	1	12
9	1	0	6	6	3	0	6
10	2	0	1	4	3	0	10
11	1	0	4	4	4	0	8
12	2	0	8	5	3	0	2
13	1	0	4	4	3	0	8
	15	15	51	42	42	11	96

610

TEST	−3	−2	−1	0	+1	+2	+3
1	1	3	5	5	2	4	5
2	1	4	1	3	0	4	6
3	1	5	3	4	2	2	7
4	2	2	3	5	0	2	6
5	5	1	2	4	1	6	18
6	4	7	1	3	1	2	6
7	1	2	1	4	2	4	8
8	1	2	0	3	1	1	14
9	6	3	1	4	0	3	5
10	3	2	5	3	1	0	6
11	3	4	0	4	0	2	8
12	6	3	3	4	1	2	1
13	3	4	0	4	0	3	6
	37	42	25	50	11	35	96

611

TEST	−3	−2	−1	0	+1	+2	+3
1	2	4	3	5	1	2	8
2	4	2	3	3	0	3	4
3	0	3	3	5	4	5	4
4	5	5	3	2	1	1	3
5	3	3	9	2	2	6	12
6	1	5	7	2	1	2	6
7	0	8	3	2	0	4	5
8	7	4	1	0	0	3	7
9	1	7	5	5	1	3	0
10	3	3	4	3	2	0	5
11	4	0	5	4	3	1	4
12	4	0	4	5	4	0	3
13	4	0	2	4	2	0	8
	38	44	52	42	21	30	69

612

TEST	−3	−2	−1	0	+1	+2	+3
1	3	5	3	5	1	5	3
2	3	7	0	1	1	2	5
3	7	2	2	4	1	1	7
4	2	5	0	3	0	7	3
5	5	6	3	1	0	7	15
6	3	5	4	0	1	3	8
7	4	5	1	0	0	4	8
8	2	2	1	3	0	5	9
9	3	11	0	3	0	3	2
10	7	4	1	1	0	3	4
11	3	3	0	4	0	1	10
12	3	5	1	5	0	3	3
13	2	2	0	6	0	2	8
	47	62	16	36	4	46	85

1972-73

613

TEST	-3	-2	-1	0	+1	+2	+3
1	0	0	0	0	0	[illegible]	0
2	4	5	3	0	3	[illegible]	2
3	2	3	4	4	1	[illegible]	5
4	1	3	4	2	4	[illegible]	5
5	12	2	4	1	5	[illegible]	10
6	6	2	2	2	2	[illegible]	9
7	6	6	2	1	2	[illegible]	4
8	3	1	2	2	0	[illegible]	13
9	5	3	3	4	0	2	5
10	5	3	[illegible]	3	2	[illegible]	7
11	6	2	3	3	0	[illegible]	6
12	2	3	6	3	2	[illegible]	3
13	2	4	5	1	1	[illegible]	7
	54	37	38	26	22	[illegible]	77

702

TEST	-3	-2	-1	0	+1	+2	+3
1	6	0	2	3	5	2	7
2	0	2	8	2	1	3	3
3	1	2	2	4	3	4	6
4	1	0	1	4	5	2	7
5	1	0	2	3	5	5	21
6	0	2	5	4	2	2	9
7	4	0	5	3	4	3	3
8	2	1	3	4	1	0	11
9	0	1	4	4	3	1	9
10	1	2	3	3	2	3	6
11	0	2	5	2	5	1	6
12	3	5	2	3	1	1	5
13	1	1	2	2	2	3	9
	20	18	44	41	39	30	104

708

TEST	-3	-2	-1	0	+1	+2	+3
1	0	11	0	4	0	9	1
2	1	4	1	3	0	6	4
3	1	2	0	3	0	4	14
4	0	3	0	2	1	6	8
5	0	2	2	7	1	18	7
6	2	5	0	3	0	10	4
7	0	5	1	7	1	7	1
8	0	4	0	2	1	5	10
9	1	4	1	4	0	1	11
10	4	1	0	3	0	3	9
11	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0
	9	41	5	38	4	65	69

712

TEST	-3	-2	-1	0	+1	+2	+3
1	1	4	1	3	0	8	[illegible]
2	1	2	4	2	1	4	5
3	3	3	0	2	0	2	14
4	2	1	2	3	0	0	12
5	1	0	4	2	4	0	26
6	4	0	3	4	2	1	10
7	1	1	1	0	6	5	6
8	5	1	1	1	0	0	14
9	0	1	1	5	2	1	12
10	2	2	1	4	1	4	6
11	2	1	1	3	2	6	6
12	3	1	3	5	1	0	7
13	1	0	2	3	2	1	11
	26	17	24	37	21	32	139

614

TEST	-3	-2	-1	0	+1	+2	+3
1	3	4	5	4	3	[illegible]	3
2	3	6	5	2	0	1	2
3	0	0	[illegible]	0	0	[illegible]	0
4	0	0	[illegible]	0	0	[illegible]	0
5	6	6	6	0	2	[illegible]	16
6	6	2	2	2	1	[illegible]	10
7	0	5	4	1	2	[illegible]	6
8	3	3	1	2	2	[illegible]	10
9	4	2	5	1	1	[illegible]	8
10	4	1	2	3	1	[illegible]	8
11	3	1	4	2	1	[illegible]	7
12	7	1	3	4	2	[illegible]	3
13	2	0	2	4	3	2	7
	41	31	39	25	18	[illegible]	80

703

TEST	-3	-2	-1	0	+1	+2	+3
1	1	5	3	5	1	10	0
2	1	7	2	2	0	5	2
3	1	5	0	5	0	8	5
4	2	1	1	4	0	1	11
5	2	3	3	4	2	3	20
6	2	4	1	2	0	2	13
7	1	4	0	1	2	3	11
8	0	2	0	3	0	1	16
9	0	1	0	2	0	[illegible]	14
10	1	1	0	2	0	8	[illegible]
11	1	3	0	4	1	9	3
12	1	5	0	4	0	6	0
13	4	2	0	3	0	2	9
	17	47	10	41	6	63	112

709

TEST	-3	-2	-1	0	+1	+2	+3
1	4	4	0	5	0	5	7
2	1	2	1	0	2	5	8
3	1	5	0	4	0	3	11
4	4	3	1	4	0	1	7
5	2	5	2	0	1	3	24
6	0	2	4	4	2	2	10
7	0	6	1	3	2	3	7
8	2	1	1	3	0	3	12
9	1	5	2	3	2	5	4
10	1	1	5	3	1	3	6
11	3	1	2	3	0	7	5
12	1	5	5	3	1	2	3
13	0	5	3	3	2	2	5
	20	45	27	38	13	44	109

713

TEST	-3	-2	-1	0	+1	+2	+3
1	5	1	1	1	1	0	16
2	1	2	2	1	0	1	12
3	3	0	0	4	1	0	16
4	4	0	1	3	3	0	[illegible]
5	5	2	1	0	0	0	29
6	11	0	0	2	0	1	10
7	1	1	0	2	1	1	16
8	4	0	0	3	0	0	15
9	1	0	2	3	1	0	15
10	4	0	0	3	2	0	11
11	2	0	1	1	2	0	15
12	6	1	0	4	0	1	8
13	1	1	0	4	0	1	13
	48	8	8	31	11	5	165

701

TEST	-3	-2	-1	0	+1	+2	+3
1	4	3	0	4	0	3	11
2	7	3	0	0	0	2	7
3	5	4	1	0	0	1	13
4	8	0	0	1	0	0	11
5	7	1	[illegible]	1	0	[illegible]	28
6	7	0	[illegible]	1	0	[illegible]	14
7	4	1	0	0	0	[illegible]	17
8	4	0	0	0	0	[illegible]	18
9	3	0	0	1	0	1	17
10	6	2	2	0	0	[illegible]	8
11	9	0	0	1	0	[illegible]	11
12	9	0	0	3	0	[illegible]	8
13	4	0	0	4	0	[illegible]	12
	77	14	3	16	0	11	175

704

TEST	-3	-2	-1	0	+1	+2	+3
1	5	7	1	3	1	2	6
2	4	3	0	4	0	0	6
3	4	4	0	3	0	2	11
4	5	0	0	4	0	0	11
5	6	0	0	6	0	7	19
6	9	0	0	2	0	1	12
7	3	4	1	3	0	7	4
8	3	0	0	2	0	1	16
9	5	2	0	2	0	2	11
10	4	1	0	5	0	2	8
11	6	0	0	3	0	3	7
12	7	2	0	5	0	2	4
13	0	0	0	0	0	0	0
	63	23	2	42	1	29	116

711

TEST	-3	-2	-1	0	+1	+2	+3
1	6	2	1	4	1	3	8
2	4	0	0	1	2	5	7
3	1	0	1	4	0	0	18
4	4	0	0	1	0	0	15
5	4	0	0	4	1	0	28
6	8	0	0	3	0	0	13
7	3	0	2	4	2	4	7
8	2	0	1	3	0	2	14
9	5	0	1	3	0	0	15
10	1	1	0	2	2	0	14
11	4	1	0	2	1	2	11
12	2	3	0	8	1	2	4
13	2	0	4	2	0	2	10
	44	7	10	41	10	20	164

714

TEST	-3	-2	-1	0	+1	+2	+3
1	3	1	1	6	3	2	9
2	4	2	3	2	5	1	2
3	1	1	2	6	2	3	9
4	4	2	0	2	1	0	11
5	5	4	0	0	3	3	22
6	1	4	2	4	2	2	9
7	2	6	1	2	0	8	3
8	1	2	0	2	0	4	13
9	4	1	1	8	0	1	7
10	1	1	2	6	0	2	8
11	3	4	0	4	4	3	3
12	0	0	0	0	0	0	0
13	1	0	3	4	2	5	5
	30	28	15	46	22	34	101

715

TEST	-3	-2	-1	0	+1	+2	+3
1	5	1	2	4	1	4	8
2	1	8	1	4	0	3	2
3	2	0	0	13	0	0	9
4	0	0	3	4	2	6	5
5	3	1	1	3	5	2	22
6	7	0	1	3	0	1	12
7	1	2	0	2	2	1	14
8	0	2	0	3	1	0	16
9	0	0	1	5	0	4	[illegible]
10	1	2	0	2	0	1	14
11	1	3	0	4	1	2	10
12	1	5	2	7	0	3	2
13	0	1	1	4	1	4	[illegible]
	22	25	12	58	13	31	135

1205

TEST	-3	-2	-1	0	+1	+2	+3
1	1	0	2	1	2	0	16
2	6	1	0	1	0	1	11
3	1	5	1	1	0	2	9
4	4	3	0	3	0	2	7
5	6	5	1	2	1	0	15
6	0	0	0	0	0	0	[illegible]
7	1	2	1	0	0	3	13
8	4	0	1	1	2	3	9
9	2	4	1	1	0	2	6
10	0	1	1	2	4	3	10
11	0	1	2	1	0	1	13
12	5	3	0	1	0	4	7
13	0	0	1	1	2	5	8
	30	25	11	15	11	26	124

1210

TEST	-3	-2	-1	0	+1	+2	+3
1	0	1	0	1	0	3	17
2	5	1	0	0	0	2	12
3	5	3	0	1	0	0	10
4	4	1	1	2	1	1	9
5	4	4	2	1	1	0	18
6	1	5	1	0	1	1	9
7	3	0	1	0	0	2	14
8	4	2	0	0	1	1	12
9	2	4	2	1	0	2	5
10	2	3	0	0	0	3	13
11	1	2	0	0	0	1	14
12	4	6	1	0	0	2	7
13	1	1	0	0	0	2	13
	36	33	8	6	4	20	153

1215

TEST	-3	-2	-1	0	+1	+2	+3
1	2	0	0	1	1	0	18
2	5	1	0	3	2	0	9
3	1	0	2	1	0	0	15
4	6	0	0	3	0	0	[illegible]
5	8	0	2	3	1	0	16
6	0	0	0	0	0	0	[illegible]
7	0	0	0	0	0	0	[illegible]
8	6	0	1	2	0	1	10
9	9	0	0	0	1	0	6
10	7	0	0	0	1	[illegible]	13
11	1	0	1	1	3	0	12
12	8	0	0	1	1	0	10
13	1	0	1	4	0	1	10
	54	1	7	19	10	2	129

1201

TEST	-3	-2	-1	0	+1	+2	+3
1	1	0	2	0	1	2	16
2	5	3	5	0	0	5	2
3	2	2	2	2	2	1	8
4	2	4	3	4	0	3	3
5	1	6	5	4	2	5	5
6	1	4	0	3	1	3	6
7	1	2	3	1	2	2	9
8	3	2	1	0	1	5	8
9	2	2	2	3	0	5	2
10	3	2	0	0	2	8	[illegible]
11	2	0	1	1	0	6	8
12	1	2	2	3	1	6	5
13	0	1	2	2	2	5	5
	24	32	26	23	14	56	83

1206

TEST	-3	-2	-1	0	+1	+2	+3
1	0	0	1	3	2	5	11
2	1	1	0	2	2	5	9
3	1	2	0	3	0	7	6
4	1	1	2	2	3	5	5
5	2	3	3	3	2	4	13
6	1	1	2	4	2	3	5
7	1	1	0	0	1	7	10
8	0	1	2	3	5	4	5
9	0	3	1	2	1	6	3
10	1	2	2	1	3	4	8
11	1	0	0	2	0	0	15
12	0	4	3	3	1	7	2
13	1	0	0	5	1	4	6
	10	19	16	33	23	61	98

1211

TEST	-3	-2	-1	0	+1	+2	+3
1	1	0	0	4	0	2	13
2	5	2	2	1	0	2	8
3	2	0	3	1	2	1	10
4	3	4	3	1	2	1	5
5	2	8	6	2	2	2	8
6	0	3	3	0	2	4	6
7	4	1	5	0	1	2	7
8	4	1	4	0	1	3	7
9	2	2	4	0	2	3	3
10	2	2	2	1	4	6	4
11	0	0	3	0	3	5	7
12	1	1	4	2	0	5	7
13	1	3	4	1	1	5	2
	27	27	43	13	20	41	87

1216

TEST	-3	-2	-1	0	+1	+2	+3
1	2	0	1	5	3	4	7
2	2	2	3	2	2	2	7
3	1	0	1	2	4	3	8
4	1	2	3	4	3	2	4
5	1	2	7	6	3	0	11
6	0	0	0	0	0	0	0
7	1	0	2	4	2	3	6
8	0	1	2	4	1	3	9
9	2	3	3	0	1	2	5
10	0	0	1	3	3	4	10
11	0	0	2	3	1	1	11
12	2	1	2	3	2	2	6
13	1	0	0	3	3	2	8
	13	11	27	39	28	28	56

1202

TEST	-3	-2	-1	0	+1	+2	+3
1	1	2	4	0	1	5	9
2	0	3	0	0	0	4	4
3	1	1	4	0	2	3	8
4	1	3	5	0	2	1	7
5	2	2	2	3	2	8	11
6	0	2	5	0	4	[illegible]	7
7	1	[illegible]	[illegible]	0	1	1	[illegible]
8	3	2	3	0	1	3	6
9	0	3	2	2	1	4	4
10	2	1	1	1	5	7	4
11	2	2	1	1	1	2	9
12	2	1	1	1	1	2	8
13	0	2	1	0	1	4	7
	15	28	46	8	24	44	95

1209

TEST	-3	-2	-1	0	+1	+2	+3
1	1	5	5	2	5	3	1
2	8	0	1	1	2	0	8
3	3	3	2	1	1	4	5
4	2	4	3	2	1	3	4
5	6	11	2	0	1	4	6
6	2	6	0	0	1	1	8
7	[illegible]	4	1	0	0	1	7
8	3	3	4	2	2	0	4
9	1	1	8	0	0	0	6
10	1	1	7	0	5	2	5
11	1	1	4	0	4	2	6
12	3	4	7	0	1	3	2
13	1	4	0	0	1	1	2
	39	47	50	8	26	24	66

1214

TEST	-3	-2	-1	0	+1	+2	+3
1	0	2	0	0	1	4	15
2	4	1	1	0	1	0	13
3	0	0	0	0	0	0	0
4	2	1	0	0	1	0	15
5	6	2	0	0	1	2	19
6	1	4	2	0	0	2	9
7	[illegible]	3	0	0	0	2	9
8	3	3	1	0	0	2	11
9	3	3	2	0	0	1	7
10	3	5	0	0	0	4	9
11	1	2	2	0	0	2	11
12	5	2	0	0	0	2	11
13	1	1	0	0	0	0	15
	35	29	8	0	4	21	144

1217

TEST	-3	-2	-1	0	+1	+2	+3
1	0	1	0	0	1	3	17
2	1	1	3	0	0	2	13
3	1	0	3	0	3	3	9
4	2	0	3	1	3	3	7
5	2	2	5	0	4	2	15
6	2	0	2	2	3	1	8
7	1	0	3	0	1	1	14
8	2	0	5	6	2	4	7
9	1	1	3	1	3	2	5
10	1	1	3	0	8	1	7
11	1	0	4	0	0	4	9
12	1	2	3	1	3	1	9
13	0	0	1	2	5	2	7
	15	8	38	7	36	29	127

1972-73

1301

TEST	-3	-2	-1	0	+1	+2	+3
1	0	1	5	2	1	7	6
2	0	3	5	5	4	7	1
3	2	6	0	1	1	4	5
4	4	0	0	3	0	3	9
5	5	2	3	5	2	1	12
6	6	0	0	0	0	0	0
7	2	0	0	0	1	0	17
8	4	4	1	4	0	2	5
9	5	3	1	0	0	1	6
10	3	0	0	0	0	0	18
11	1	0	0	0	1	0	16
12	3	2	5	5	2	1	2
13	0	0	0	0	0	0	0
	29	21	20	25	12	21	57

1304

TEST	-3	-2	-1	0	+1	+2	+3
1	2	2	0	2	0	4	12
2	3	1	0	4	0	2	10
3	1	1	1	4	0	3	9
4	3	0	0	1	1	0	14
5	2	0	2	0	2	0	24
6	4	0	2	0	0	0	12
7	2	0	0	0	0	0	18
8	3	0	0	1	0	0	16
9	6	1	0	0	0	1	8
10	1	0	1	1	2	1	15
11	1	0	2	0	3	0	12
12	4	3	2	2	1	2	6
13	5	0	0	0	0	0	12
	37	8	10	15	9	13	169

1307

TEST	-3	-2	-1	0	+1	+2	+3
1	1	1	5	1	4	4	6
2	2	1	2	3	1	0	11
3	0	5	0	0	0	0	14
4	4	0	0	1	0	0	14
5	7	0	0	0	0	0	23
6	7	0	0	0	0	0	11
7	0	0	0	0	0	0	0
8	6	1	0	0	0	2	11
9	4	4	1	1	0	1	5
10	3	4	0	0	0	5	9
11	2	0	0	0	1	0	15
12	6	4	3	1	1	2	3
13	3	1	0	1	2	1	9
	45	21	11	8	9	15	131

1311

TEST	-3	-2	-1	0	+1	+2	+3
1	0	2	4	1	3	3	9
2	2	0	2	3	2	1	10
3	0	6	0	2	0	5	6
4	1	4	3	2	0	4	5
5	2	3	3	3	3	7	9
6	0	0	0	0	0	0	0
7	1	2	0	2	1	2	12
8	3	3	1	2	0	4	7
9	6	1	0	1	1	1	6
10	2	2	2	3	1	2	9
11	2	1	0	1	1	0	13
12	2	4	1	3	1	3	6
13	1	0	1	2	1	1	11
	22	28	17	25	14	33	103

1302

TEST	-3	-2	-1	0	+1	+2	+3
1	1	1	0	3	0	3	14
2	6	0	0	4	0	1	10
3	2	0	0	2	0	1	14
4	1	3	4	0	1	1	9
5	3	0	3	4	0	0	22
6	8	0	0	1	0	0	9
7	2	1	0	0	0	0	14
8	7	0	0	1	0	2	10
9	8	0	0	0	0	1	7
10	3	0	3	3	0	0	12
11	1	1	0	1	5	1	9
12	6	1	1	2	1	0	9
13	3	0	1	1	2	2	8
	51	7	12	22	9	14	145

1305

TEST	-3	-2	-1	0	+1	+2	+3
1	1	0	0	3	0	1	17
2	3	0	0	2	0	1	14
3	1	0	0	0	0	0	18
4	6	1	0	2	0	2	8
5	7	0	0	1	0	0	22
6	6	0	0	0	0	0	12
7	3	0	0	0	0	0	17
8	5	0	0	3	0	0	12
9	6	0	0	3	0	0	7
10	4	0	0	1	0	0	16
11	5	0	0	1	0	0	14
12	7	0	0	2	0	0	11
13	1	0	0	3	0	0	13
	53	1	0	21	0	4	181

1308

TEST	-3	-2	-1	0	+1	+2	+3
1	1	1	2	0	1	5	12
2	4	1	2	2	0	2	9
3	1	0	0	0	0	0	18
4	3	0	0	0	0	0	16
5	6	0	0	0	1	0	23
6	7	0	1	0	0	0	10
7	4	0	0	0	0	0	16
8	0	0	0	0	0	0	0
9	8	0	2	0	1	0	5
10	7	2	0	0	2	1	9
11	1	2	0	0	1	1	13
12	7	2	3	1	0	1	6
13	2	2	0	1	1	3	8
	51	10	10	4	7	13	145

1312

TEST	-3	-2	-1	0	+1	+2	+3
1	1	3	3	2	1	3	9
2	5	6	3	0	1	1	4
3	0	0	0	0	0	0	0
4	3	5	1	0	2	2	6
5	8	2	1	2	0	2	15
6	1	3	3	0	0	4	7
7	0	0	0	0	0	0	0
8	5	1	0	0	0	0	14
9	10	0	0	0	0	0	6
10	5	0	0	1	0	0	15
11	3	0	0	1	0	0	14
12	8	0	0	2	0	0	10
13	3	0	0	1	0	0	13
	52	20	11	9	4	12	113

1303

TEST	-3	-2	-1	0	+1	+2	+3
1	1	3	3	3	0	9	3
2	3	4	6	2	1	4	0
3	0	1	8	2	4	3	1
4	6	0	0	2	0	0	0
5	8	0	0	0	0	0	22
6	8	0	0	0	0	0	17
7	3	0	0	0	0	0	17
8	5	2	2	2	0	1	8
9	6	1	0	0	0	1	8
10	8	1	0	0	2	1	9
11	5	0	0	[illegible]	0	0	13
12	7	1	1	[illegible]	0	0	[illegible]
13	0	0	0	0	0	0	0
	62	13	[illegible]	13	7	15	[illegible]

1306

TEST	-3	-2	-1	0	+1	+2	+3
1	3	4	1	5	0	2	7
2	5	3	1	4	0	0	7
3	7	1	1	0	0	1	9
4	2	3	3	3	3	1	4
5	6	0	3	4	1	1	15
6	4	2	0	0	1	2	9
7	6	0	0	0	0	0	14
8	4	1	2	0	1	0	12
9	7	0	0	0	1	3	5
10	7	1	2	1	0	2	8
11	2	0	1	0	1	3	11
12	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
13	2	2	0	1	1	1	10
	64	17	14	27	9	16	118

1310

TEST	-3	-2	-1	0	+1	+2	+3
1	4	1	3	3	0	1	10
2	6	1	1	1	0	0	11
3	3	0	1	2	0	0	13
4	2	2	1	3	1	0	10
5	5	2	0	4	0	3	16
6	0	1	0	0	3	1	13
7	1	0	0	1	0	1	17
8	5	0	1	1	0	0	13
9	6	0	1	2	0	0	7
10	5	0	0	1	0	0	15
11	3	0	0	3	0	0	12
12	5	0	2	2	1	[illegible]	6
13	1	0	0	3	0	1	12
	46	7	10	26	5	5	157

1313

TEST	-3	-2	-1	0	+1	+2	+3
1	2	0	3	3	0	2	12
2	5	0	0	4	0	0	11
3	0	0	0	3	0	6	10
4	3	1	0	2	0	1	12
5	1	4	0	4	0	0	21
6	4	3	0	2	0	0	9
7	0	0	2	0	0	2	16
8	4	1	0	1	3	2	9
9	8	0	0	1	0	0	7
10	2	0	0	2	2	0	15
11	0	0	2	0	5	1	10
12	4	0	4	2	1	0	9
13	4	2	0	0	0	0	11
	37	11	11	24	11	14	152

1314

TEST	-3	-2	-1	0	+1	+2	+3
1	6	0	0	0	0	0	16
2	2	1	1	4	0	5	7
3	0	1	0	2	0	3	13
4	2	0	0	3	0	0	14
5	3	3	0	4	0	2	18
6	1	1	1	2	0	0	13
7	1	0	0	3	0	0	16
8	3	1	0	2	0	4	10
9	2	1	1	3	0	2	7
10	0	0	0	0	0	0	0
11	1	2	0	2	0	3	10
12	6	2	2	4	1	0	5
13	1	0	3	2	1	0	10
	28	12	8	31	2	15	139

1703

TEST	-3	-2	-1	0	+1	+2	+3
1	2	0	1	6	2	0	11
2	3	0	1	4	1	1	10
3	1	1	1	4	2	4	6
4	4	5	2	0	1	6	1
5	4	0	5	6	5	1	9
6	2	2	2	2	4	3	3
7	3	2	2	3	2	3	5
8	1	5	1	4	0	2	7
9	7	0	2	2	2	0	3
10	1	1	5	4	3	1	6
11	3	1	5	2	0	3	4
12	3	5	3	0	2	2	5
13	2	3	5	3	1	0	3
	36	25	35	40	25	26	73

1707

TEST	-3	-2	-1	0	+1	+2	+3
1	2	0	0	3	0	1	16
2	6	1	0	2	0	0	11
3	4	3	0	2	0	2	8
4	4	2	1	4	2	0	6
5	4	3	1	6	1	1	14
6	0	0	0	0	0	0	0
7	3	1	2	2	1	0	11
8	3	0	3	4	3	0	7
9	1	1	4	4	3	1	2
10	1	1	4	5	3	0	7
11	0	0	3	3	1	1	10
12	3	0	3	3	5	0	6
13	0	0	3	3	3	1	7
	31	12	24	41	22	7	105

1710

TEST	-3	-2	-1	0	+1	+2	+3
1	10	1	0	0	0	2	9
2	3	3	2	1	0	1	10
3	3	1	1	0	2	5	7
4	4	3	1	1	1	3	6
5	0	5	1	5	1	1	17
6	2	2	0	1	3	3	7
7	2	3	0	1	3	1	10
8	2	2	2	1	0	2	11
9	4	3	2	0	0	2	5
10	4	4	2	0	0	0	11
11	3	4	2	0	0	0	9
12	3	2	1	0	2	3	9
13	1	1	0	2	0	0	13
	41	34	14	12	12	23	124

1701

TEST	-3	-2	-1	0	+1	+2	+3
1	1	0	4	0	0	2	15
2	5	1	0	4	0	0	10
3	4	1	1	2	1	1	9
4	6	0	1	3	1	0	8
5	4	0	5	4	0	1	10
6	6	1	1	4	1	0	5
7	6	0	4	4	1	0	5
8	0	1	2	4	1	1	11
9	3	0	5	2	1	1	4
10	2	0	3	5	3	1	7
11	2	0	4	2	2	1	7
12	1	1	5	2	6	2	3
13	1	2	2	2	0	2	8
	41	7	37	38	23	12	102

1704

TEST	-3	-2	-1	0	+1	+2	+3
1	5	0	1	0	1	1	14
2	2	2	2	0	1	3	10
3	1	0	3	1	2	4	8
4	1	1	2	0	6	2	7
5	1	0	5	1	1	3	14
6	2	2	2	2	3	2	5
7	1	0	0	2	1	5	11
8	2	1	1	2	0	3	11
9	1	1	2	2	2	2	6
10	1	1	2	4	3	2	8
11	1	2	2	2	4	0	7
12	1	1	3	4	3	3	5
13	0	1	1	3	4	2	6
	19	12	26	23	31	32	117

1708

TEST	-3	-2	-1	0	+1	+2	+3
1	0	2	2	4	1	7	6
2	5	3	1	4	0	4	3
3	0	3	3	4	4	4	1
4	0	1	3	4	7	3	1
5	0	3	6	7	6	5	3
6	0	3	6	2	2	5	0
7	0	0	6	4	1	9	0
8	1	3	2	4	4	4	2
9	0	2	6	3	1	3	1
10	0	3	4	4	5	4	1
11	0	2	4	3	3	4	2
12	0	2	3	4	3	8	0
13	0	4	2	3	2	3	3
	6	31	48	50	39	63	23

1711

TEST	-3	-2	-1	0	+1	+2	+3
1	2	3	3	1	1	0	12
2	0	3	3	3	0	0	11
3	2	3	4	0	4	1	5
4	0	3	2	1	2	3	8
5	4	2	3	2	3	2	14
6	5	3	1	3	0	1	5
7	7	1	3	4	0	2	3
8	1	1	4	2	0	4	8
9	2	5	1	2	1	1	4
10	3	3	0	3	0	2	10
11	0	1	2	2	1	2	10
12	5	3	1	3	2	0	6
13	4	1	1	3	0	0	8
	35	32	28	29	14	18	104

1702

TEST	-3	-2	-1	0	+1	+2	+3
1	1	2	0	2	1	6	10
2	3	2	1	2	0	2	10
3	3	6	1	4	0	1	4
4	3	7	2	1	0	0	6
5	4	2	2	6	0	1	15
6	2	1	0	3	0	4	12
7	3	2	0	3	0	4	8
8	3	1	0	4	1	2	9
9	5	1	1	1	0	3	5
10	3	1	1	2	2	4	8
11	2	1	1	2	1	1	[illegible]
12	3	4	0	2	1	[illegible]	[illegible]
13	1	1	3	1	0	1	10
	36	31	14	33	6	30	110

1706

TEST	-3	-2	-1	0	+1	+2	+3
1	1	1	3	0	2	5	10
2	2	1	4	0	0	4	9
3	0	4	1	1	1	4	8
4	1	4	2	3	1	4	4
5	2	6	2	1	0	6	13
6	2	4	1	2	2	2	[illegible]
7	0	3	1	3	1	4	8
8	0	1	2	0	4	7	6
9	2	3	1	0	2	6	2
10	1	3	3	0	1	6	7
11	1	1	4	0	2	4	6
12	2	6	2	1	1	2	[illegible]
13	2	3	0	1	0	3	8
	16	40	26	12	17	57	92

1709

TEST	-3	-2	-1	0	+1	+2	+3
1	2	3	3	0	3	3	8
2	4	6	3	2	2	2	1
3	2	4	2	2	2	5	2
4	1	5	5	0	2	4	2
5	1	7	4	0	7	6	5
6	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0
8	0	2	5	0	4	1	8
9	2	5	4	3	1	0	1
10	1	2	2	4	2	6	4
11	1	3	3	0	3	1	7
12	2	8	1	0	2	2	[illegible]
13	0	1	1	3	1	4	7
	16	46	33	14	29	34	50

1712

TEST	-3	-2	-1	0	+1	+2	+3
1	4	4	1	0	1	1	11
2	4	1	2	4	0	0	9
3	2	2	1	4	1	1	8
4	2	3	0	4	0	3	7
5	5	2	1	2	3	1	16
6	5	2	0	3	0	1	7
7	7	1	0	4	1	0	7
8	2	3	0	2	0	1	12
9	5	3	0	4	0	0	4
10	2	1	0	4	0	2	12
11	6	0	0	1	0	0	11
12	4	3	0	4	0	2	7
13	4	4	0	2	0	1	6
	52	29	5	38	6	13	117

Annexe 5 (Cf chap. 5, p. 291)

AINSI L'ABONDANCE DE L'HOMME ENGENDRE LE RATIONNEMENT DE LA PERSONNALITE.MEME SI
L'HOMME MOYEN EST CONNU D'UN GRAND NOMBRE,IL EST CONNU MOINS PROFONDEMENT,MOINS
PRIVEMENT,MOINS SERIEUSEMENT.LA CONTREPARTIE DE L'EXTENSION DE NOS INFORMATIONS
,ET DE LA MULTIPLICATION DE NOS"RELATIONS",C'EST LA REDUCTION DE LA CONNAISSANCE
DE NOS PROCHES.CES FAITS SONT EN RELATION AVEC LE CHANGEMENT PROFOND,QUI S'ACCO
MPLIT SOUS NOS YEUX,DANS LES COUTUMES RELATIVES A LA MORT,AUX FAIRE-PART,AUX FUN
ERAILLES,AU DEUIL.LORSQUE CES NOUVELLES COUTUMES ONT COMMENCE DE SE FAIRE JOUR A
UX ETATS-UNIS,ELLES ONT CHOQUE LES EUROPEENS ET FURENT MISES AU COMPTE DE LA BAR
BARIE YANKEE.MAIS NOUS LES VOYONS AUJOURD'HUI NON PAS GAGNER L'EUROPE(LE MOT GAG
NER IMPLIQUERAIT QU'ELLES SONT AUJOURD'HUI IMPORTEES,IMITEES DES ETATS-UNIS PAR
L'EUROPE),MAIS SE FORMER,SE CONSTITUER EN EUROPE,COMME ELLES L'ONT FAIT AUX ETAT
S-UNIS,POUR LES MEMES CAUSES ET PAR SUITE DE L'AVENEMENT DE CES MEMES CAUSES.EN
FAIT,CE NE SONT PAS QUELQUES USAGES ANECDOTIQUES QUI SONT EN QUESTION;CE N'EST P
AS SEULEMENT L'ATTITUDE DEVANT LA MORT;C'EST TOUTE L'ATMOSPHERE DE PENSEE DE L'H
UMANITE MILLENAIRE QUI ACHEVE DE SE DISSIPER.LE CHANGEMENT EST DEJA SI TOTAL QUE
LA PLUPART DE MES LECTEURS NE POURRONT REELLEMENT COMPRENDRE DE QUOI JE PARLE I
CI QUE S'ILS FONT L'EFFORT DE LIRE LES RARES LIVRES OU LA MENTALITE PRIMITIVE ET
TRADITIONNELLE EST CORRECTEMENT DECRITE.L'HUMANITE A VECU JUSQU'A NOS JOURS DAN
S UNE CONCEPTION MAGIQUE ET RITUELLE DES GRANDS EVENEMENTS DE LA VIE:NAISSANCE,M
ARIAGE,MORT;LES FETES,LE TRAVAIL ETAIENT DES CEREMONIES FIXEES PAR LES ANCETRES;
LA SITUATION SOCIALE ET MORALE DE CHAQUE PERSONNE ETAIT REGLEE PAR UN SACREMENT
.PLUS ENCORE QUE PAR SA RELIGION,L'HOMME MOYEN ETAIT SOUTENU,GUIDE,ENTRETENU DAN
S SON ARDEUR DE VIVRE,PAR SA PAROISSE,SA COMMUNAUTE,LE PETIT GROUPE HUMAIN QUI L
'ENTOURAIT COMME LA COUVEE FAIT LE POUSSIN.L'HOMME MOYEN D'AUJOURD'HUI EST SEUL
DANS LA FOULE.CE NE SONT PLUS LES FACTEURS AFFECTIFS,SENTIMENTAUX ET RITUELS QUI
SOUTIENNENET SA VIE;CE N'EST QUE PARTIELLEMENT D'EUX QU'IL ATTEND SON EQUILIBRE
.IL L'ATTEND DE PLUS,ET PEUT-ETRE DAVANTAGE ENCORE,DE LA SATISFACTION DE SES BES
OINS ECONOMIQUES ET DE LA SATISFACTION DE SES BESOINS INTELLECTUELS.C'EST SUR LU
I-MEME,ET PLUS PRECISEMENT SUR SA PROPRE ACTION,SUR SA PROPRE PENSEE,QU'IL DOIT
FONDER SA PROPRE ACTION,SUR SA PROPRE PENSEE,QU'IL DOIT FONDER SA SITUATION.LES
PARAGRAPHES QUI PRECEDENT MONTRENT L'IMPORTANCE POUR LA SITUATION DE L'HOMME DE
LA CONNAISSANCE QU'IL A SUR LES REALITES QUI L'ENTOURENT,DE LA CONCEPTION QU'IL
SE FAIT DE LUI-MEME ET DE L'UNIVERS.DEPUIS DES MILLENAIRES,L'HUMANITE A VECU PAR
INSTINCT ET PAR CROYANCE.DANS SA FAIBLE FRACTION CONSCIENTE,ELLE A BEAUCOUP CRU
,UN PEU IMAGINE,TRES PEU RAISONNE,PRATIQUEMENT PAS OBSERVE(L'HOMME MOYEN EST EFF
ECTIVEMENT,A L'ETAT NATUREL,FORT PEU APTE A L'OBSERVATION).C'EST PARCE QUE L'HUM
ANITE TRADITIONNELLE N'A PRATIQUEMENT PAS SU OBSERVER QUE L'ON A PU DIRE QUE LES
NEUF DIXIEMES DES CHERCHEURS QU'ELLE A PRODUITS SONT AUJOURD'HUI VIVANTS.LE RES
ULTAT EST QUE L'HUMANITE TRADITIONNELLE SAVAIT TRES PEU DE CHOSES SUR ELLE-MEME
ET SUR L'UNIVERS QUI L'ENTOURAIT;LA QUASI-TOTALITE DE CE PEU,ELLE LE TENAIT OU C

ROYAIT LE TENIR DE CE QUE LES RELIGIONS APPELLENT LA REVELATION,ET ELLE LE TRANS
METTAIT PAR CROYANCE.CE QUE CONNAITRA L'HOMME DU 21E SIECLE,NOUS NE POUVONS ENCO
RE LE DIRE,PUISQUE NOUS SAVONS QUE NOTRE CONNAISSANCE CROIT TRES RAPIDEMENT ET Q
UE PREVOIR UNE DECOUVERTE SERAIT LA FAIRE.MAIS CE QUE NOUS SAVONS DEJA NOUS-MEME
S AUJOURD'HUI SUFFIT A NOUS PLACER DANS UNE SITUATION BIEN DIFFERENTE DE NOS AIE
UX.IL NE PEUT ETRE EVIDEMMENT QUESTION DE DRESSER UN INVENTAIRE DE NOS CONNAISSA
NCES S'OPPOSANT A L'IGNORANCE RECENTE.NOTRE OBJET EST D'IDENTIFIER LES GRANDS OR
DRES DE FAITS QUI COMMANDENT LA CONDITION HUMAINE.EN CETTE MATIERE,NOUS PENSONS
POUVOIR RAMENER A QUATRE LES PRISES DE CONSCIENCE QUI ONT MODIFIE NOTRE CONCEPTI
ON DU MONDE:LES TROIS PREMIERES CONCERNENT TROIS IMMENSITES;LE QUATRIEME EST L'I
DENTIFICATION D'UNE METHODE.1-NOS ANCETRES CROYAIENT L'UNIVERS BORNE A LA TERRE
ET A DES ACCESSOIRES DESTINES A SON SERVICE.NOUS SAVONS QUE NOUS VIVONS SUR UNE
PLANETE QUI FAIT PARTIE D'UN UNIVERS DE DIX MILLIARDS D'ANNEES-LUMIERE,ET QUE CE
T INDEFINIMENT GRAND EST CONSTITUE D'UN NOMBRE INDEFINI DE PARTICULES INDEFINIME
NT PETITES...2-NOS ANCETRES CROYAIENT QUE LA TERRE ET L'HUMANITE REMONTAIENT A C
INQ OU DIX MILLE ANS ET QUE LA FIN DU MONDE ETAIT PROCHE.NOUS SAVONS QUE LA TERR
E EXISTE DEPUIS PLUSIEURS MILLIARDS D'ANNEES,DES HOMMES,OU DU MOINS DES HOMINIEN
S,DEPUIS AU MOINS 500 MILLE ANS;L'ASTRONOMIE NE NOUS DONNE AUCUNE RAISON DE PENS
ER QUE LA TERRE NI L'HUMANITE SOIENT MENACEES DE DISPARAITRE AVANT 100 OU 500 MI
LLE ANS.3-CET UNIVERS IMMENSE EST INDEFINIMENT COMPLEXE,CES DEUX IMMENSITES COMP

LEXES FONT L'IMMENSITE DE NOTRE IGNORANCE.SI GRANDS QUE SOIENT LES PROGRES DE NO
TRE CONNAISSANCE,L'UNIVERS EN DEVENIR NOUS APPARAIT COMME UNE SOURCE INEPUISABLE
DE NOUVELLES CONNAISSANCES;L'HOMME DE SCIENCE SE TROUVE PERDU DANS L'IMMENSITE
DES CHOSES A CONNAITRE,COMME L'ETRE VIVANT DANS L'IMMENSITE DU TEMPS ET DE L'ESP
ACE.4-CEPENDANT,POUR CETTE TACHE DE SISYPHE,L'HOMME A AUJOURD'HUI UNE METHODE EP
ROUVEE,QUI LUI A DONNE EN QUATRE CENTS ANS PLUS DE RESULTATS QUE TOUTES LES AUTR
ES EN CENT MILLE ANS;LA METHODE EXPERIMENTALE.AUJOURD'HUI,L'HOMME A UNE CONFIANC
E TELLE EN CETTE METHODE QUE TOUTE CONNAISSANCE REVELEE,IMAGINEE OU RAISONNEE,LU
I EST STRICTEMENT SUBORDONNEE.ALORS QUE L'HUMANITE MILLENAIRE EUT SANS HESITATIO
N DONNE LE PAS AU RAISONNEMENT OU A LA REFLEXION SUR L'EXPERIENCE,NOUS INTERPRET
ONS LA REVELATION ET NOUS PLIONS LA RATIONALITE EN FONCTION DE L'OBSERVATION.CES
PROGRES DANS NOTRE CONNAISSANCE DE L'UNIVERS ET CES CHANGEMENTS DANS NOTRE CONC
EPTION MEME DE LA CONNAISSANCE ONT EVIDEMMENT DE GRANDES CONSEQUENCES;IL NE SEMB
LE PAS DOUTEUX QU'ILS EUSSENT FORTEMENT MODIFIE,S'ILS EUSSENT ETE CONNUS D'EUX,L
'OEUVRE DES GRANDS HOMMES QUI ONT,AU COURS DES SIECLES,ELABORE LES RELIGIONS,LES
MORALES ET LES PHILOSOPHIES QUI CONSTITUENT AUJOURD'HUI NOTRE CONCEPTION DU MON
DE:ILS SONT DONC CAPABLES,DES MAINTENANT,MAIS SURTOUT D'ICI A UNE CINQUANTAINE D
'ANNEES,PAR REFLEXION ET CRITIQUE CONSTRUCTIVE,DE RENOUVELER LE CONTENU DU CERVE
AU HUMAIN, LE STOCK D'IDEES,OU COMME DIT TRES BIEN M.PIERRE AUGER,LA POPULATION
D'IDEES ENGENDREE PAR L'HUMANITE.EN CE SENS ENCORE,ON PEUT PENSER QUE L'HUMANITE

CHANGE QUALITATIVEMENT PLUS ENCORE QUE QUANTITATIVEMENT.MAIS LES FACTEURS BIOLO
GIQUES METTENT PLUS ENCORE EN QUESTION LA NATURE DE L'HOMME.EN MATIERE DE BIOLOG
IE,L'ACCROISSEMENT,DES POUVOIRS DE L'HOMME A D'ABORD ETE UTILISE A LA LUTTE CONT
RE LES MALADIES ET A LA PROTECTION DES FAIBLES;ELLE ABOUTIT AUJOURD'HUI A ENVISA
GER L'ASSURANCE ET LE PERFECTIONNEMENT MEME DES HOMMES NORMAUX.CEPENDANT LA SCIE
NCE BIOLOGIQUE MET EN EVIDENCE LA FRAGILITE,LA COMPLEXITE ET L'ORIGINALITE DU PH
ENOMENE HUMAIN.NOUS N'AVONS PAS A REVENIR ICI,MALGRE LEUR IMPORTANCE,SUR LE FAIT
DE L'ALLONGEMENT DE LA VIE MOYENNE DANS LA FIXITE DE LA VIE MAXIMALE;SUR LE FAI
T QUE L'HOMME MOYEN EST MAINTENANT APPELE A VIVRE UNE VIE BIOLOGIQUEMENT COMPLET
E;NI SUR LES AUTRES FAITS RELATIFS A LA MORTALITE ET A LA MORBIDITE.CE SONT DES
REALITES DONT NOUS AVONS DEJA DIT UN MOT AU CHAPITRE 2,ET QUI SONT EXPOSEES CORR
ECTEMENT DANS DE NOMBREUX LIVRES.IL NOUS FAUT PAR CONTRE EVOQUER ICI,EN NOUS PLA
CANT DANS LA PERSPECTIVE DU 21E SIECLE,LES CONSEQUENCES INATTENDUES ET AUJOURD'H
UI ENCORE PEU APPARENTES ET PEU CLAIRES,DE NOTRE EFFORT CONSCIENT ET FRUCTUEUX P
OUR ALLONGER LA VIE MOYENNE ET REDUIRE LA MALADIE.LES BIOLOGISTES SE SONT AVISES
DEPUIS LONGTEMPS QUE LA PROLONGATION DE LA VIE DES FAIBLES AVAIT POUR CONSEQUEN
CE UN AFFAISSEMENT PROGRESSIF DE L'ESPECE;ELLE PERMET EN EFFET LA REPRODUCTION D
ES ETRES QUE LA"SELECTION NATURELLE"ELIMINAIT DUREMENT.CE N'EST QUE PLUS RECEMME
NT QUE FUT CONSTATEE LA FREQUENCE DES EVENEMENTS DEFAVORABLES,DANS LA MUTATION D
ES GENES,ET LA RARETE DES EVOLUTIONS FAVORABLES,DE TELLE SORTE QUE LE PROFESSEUR
L'HERITIER PUT ECRIRE"LA MUTATION EST LA SOURCE DU MAL BIOLOGIQUE".M.JEAN ROSTAND
A ATTIRE L'ATTENTION NON SEULEMENT DU GRAND PUBLIC,MAIS DU MONDE SAVANT,SUR CE
DOUBLE DANGER,PAR DES ECRITS CELEBRES:"LA THERAPEUTIQUE EST POURVOYEUSE DE TARES
;ELLE SE RECRUTE DES CLIENTS,ET ELLE CREE LES HOMMES QUI AURONT BESOIN D'AVOIR R
ECOURS A ELLE.QUE L'ESPECE HUMAINE SOIT DIRECTEMENT MENACEE DANS LA QUALITE DE S
ON PATRIMOINE HEREDITAIRE,C'EST LA LA CONCLUSION A LAQUELLE IL EST DIFFICILE D'E
CHAPPER...IL Y A UNE CONTRADICTION ESSENTIELLE ENTRE LE MIEUX-ETRE INDIVIDUEL ET
LE MAL GENETIQUE.LA CIVILISATION,CLEMENTE AUX MAUVAIS GENES,ACCROIT LA SOURCE B
IOLOGIQUE DU MAL."BIEN ENTENDU,ROSTAND N'EN CONCLUT PAS LA DEGRADATION IRR
EMEDIABLE DE L'ESPECE,MAIS LA NECESSITE D'UN ENORME EFFORT DE CONTREPARTIE.CETTE
CONTREPARTIE EXISTE DES MAINTENANT,C'EST NOTAMMENT LA PHARMACOGENETIQUE.MAIS EL
LE EST ENCORE INFANTILE,ET IL EST CLAIR QU'ELLE IMPLIQUE,POUR ETRE APPLIQUEE A L
'ECHELLE DES GRANDS NOMBRES,D'ENORMES QUANTITES DE TRAVAIL HAUTEMENT QUALIFIE.IL
EST DONC PREVISIBLE QUE,D'ICI AU 21E SIECLE,LA PARTIE NE SERA PAS GAGNEE.ET MEME
SI ELLE L'EST DU POINT DE VUE DE LA CONNAISSANCE SCIENTIFIQUE,ELLE N'EVITERA PAS
DANS LA POPULATION VIVANTE,AU COURS DES 100 PROCHAINES ANNEES,L'ACCROISSEMENT D
ES DESORDRES MENTAUX ET PHYSIOLOGIQUES.NOUS PARLERONS DE CES FAITS,QUI PEUVENT N
'ETRE QUE TRANSITOIRE,AU CHAPITRE SUIVANT.LA LUTTE CONTRE LA MALADIE A D'AUTRES
CONSEQUENCES INDIRECTES,SINON FACHEUSES,DU MOINS AMBIGUES.LA TENDANCE ACTUELLE E
ST EN EFFET DANS CETTE LUTTE,DE RECOURIR BEAUCOUP MOINS A LA STIMULATION DES DE
FENSES NATURELLES DU CORPS HUMAIN QU'A LA SUPPRESSION DES AGENTS D'AGRESSION.SAN
S QUE L'ON PUISSE JUGER DEFINITIVEMENT AVANT L'EXPERIENCE A TRES LONG TERME,L'ON
DOIT DES MAINTENANT CRAINDRE QUE CE FAIT N'ABOUTISSE A LA LONGUE A UNE REDUCTIO
N DE L'ENSEMBLE DES REFLEXES VITAUX.D'AUTRE PART,BEAUCOUP D'HOMMES EMINENTS CRAI
GNENT AUSSI QUE LA DESTRUCTION DES AGENTS NATURELS JUGES AGRESSIFS,PAR SON ARBI
TRAIRE ET SON EFFICACITE PLANETAIRE,NE PRIVE EN REALITE L'HOMME DU MILIEU NATURE
L,ANIMAL ET VEGETAL,QUI LUI EST INDISPENSABLE.NOUS VOYONS AINSI QU'EN MATIERE DE
BIOLOGIE,NON SEULEMENT LA SITUATION DE L'HUMANITE DU 21E SIECLE ET LA SITUATION
MILLENAIRE DIFFERENT RADICALEMENT,MAIS ENCORE CERTAINS ELEMENTS DE CES DIFFEREN
CES SONT PREOCCUPANTS.CEPENDANT,NOUS N'AVONS ENCORE RIEN DIT DE CE QUI EST LE PL
US EXTRAORDINAIRE:LA GRANDE PROBABILITE DE VOIR L'HOMME DU 21E SIECLE EN MESURE
D'INTERVENIR DANS LES PROCESSUS GENETIQUES.LES THEMES TOUCHENT ICI A LA SCIENCE-
FICTION;MAIS LES PROBABILITES DE REALISATION SONT ASSEZ FORTES;ET L'IMPORTANCE P
OUR L'HOMME TELLE QUE NOUS DEVONS Y REFLECHIR,MEME EN ADMETTANT QUE LES MODALITE
S EFFECTIVES DE L'EVENEMENT DIFFERENT SENSIBLEMENT DE CELLES QUI SONT AUJOURD'HU
I VOLONTE ET NAGUERE DE LA CONSCIENCE DES PARENTS.DEJA,IN VITRO,CERTAINS CHERCHEU
RS SONT CAPABLES DE L'INFLUENCER;LE SUCCES IN VIVO EST TRES PROBABLE,A L'ECHELLE
DU LABORATOIRE SINON A L'ECHELLE DE LA CLINIQUE POUR L'HOMME OU LA FEMME MOYENS
.SANS DOUTE EST-IL UTOPIQUE D'ENVISAGER DES LE 21E SIECLE QUE LES FUTURS PARENTS
VIENNENT COMMANDER AU GENETICIEN-GYNECOLOGUE D'ASSEMBLER LES GENES DE LEUR SEME
NCE DE MANIERE A OBTENIR UNE FILLE AYANT LA COULEUR DE CHEVEUX DE TELLE AIEULE,L
'APTITUDE AUX MATHEMATIQUES DE TEL AUTRE,LA VIGUEUR DU LION ET LA DOCILITE DE L'
AGNEAU(JE TOMBE,EN ECRIVANT CES VOEUX,DANS L'ORNIERE DE L'HUMANITE TRADITIONNELL
E;CAR LE BESOIN DU 21E SIECLE EXCLUERA UNE COULEUR DE CHEVEUX DEFINIE POUR S'ATT
ACHER A UNE NEUTRALITE AUTORISANT LE MAXIMUM DE VARIETE DES TEINTES ARTIFICIELLE
S).IL EST PAR CONTRE A PEU PRES CERTAIN QUE LE PROCESSUS DEJA EN COURS SERA CONF
IRME QUI CONDUIT A DESOLIDARISER L'ACTE SEXUEL DE LA PROCREATION ET A FAIRE DE L
A PROCREATION UN ACTE TECHNICO-SCIENTIFIQUE.LES CONSEQUENCES DIRECTES ET INDIREC
TES DE CES FAITS SONT GRANDES ET NOMBREUSES.NOUS PARLERONS DE CERTAINES D'ENTRE
ELLES AUX CHAPITRES SUIVANTS.NOUS NOTERONS SEULEMENT ICI QUE L'UNE DE CES CONSEQ
UENCES SEMBLE ETRE LA TENDANCE A LA REDUCTION DE LA SPECIFICITE DES ROLES SEXUEL
S QUI EST CONSTATEE DEPUIS QUELQUES ANNEES.MAIS LE FACTEUR LE PLUS BOULEVERSANT
EST LA DECOUVERTE DE LA FRAGILITE DU PATRIMOINE GENETIQUE DE L'HUMANITE ET DE SA
SENSIBILITE A DES TROUBLES INFIMES.IL SUFFIT QUE L'ENERGIE MECANIQUE REVETE LA
FORME DE CERTAINES RADIATIONS POUR QUE QUELQUES CENTIEMES OU MILLIEMES DE WATT-S
ECONDE DETRUISENT CE PATRIMOINE OU LUI CAUSENT D'IRREPARABLES DEGRADATIONS.C'EST
POURQUOI LE DR MAROIS A PU ECRIRE:"SI L'ON ME DEMANDAIT CE QUE L'HOMME DOIT SAU
VEGARDER D'ABORD,JE NE DIRAIS PAS SEULEMENT LES MONUMENTS DU DESERT DE NUBIE,LE
PARTHENON ET LA CHAPELLE SIXTINE,MAIS CES QUELQUES ACIDES NUCLEIQUES QUI DANS NO
S CELLULES GERMINALES ASSURENT D'AGE EN AGE LA PROPAGATION DE NOTRE ESPECE."EN
DEFINITIVE,AU MOMENT DE CLORE CE CHAPITRE QUI,TOUT EN ETANT TRES INCOMPLET,NE P

```
EUT LAISSER AUCUN DOUTE SUR L'ABIME QUANTITATIF ET QUALITATIF QUI SEPARE LA SITU
ATION MILLENAIRE ET TRADITIONNELLE DE L'HOMME DE SA SITUATION PROCHAINE,QUANT A
1111111 11111111 1  11111111111        11111    11 1111111   1 111111      11
 LA DEMOGRAPHIE,QUANT A L'ECONOMIE,QUANT AU GENRE DE VIE,QUANT AU LOISIR,QUANT A
UX MILIEUX TECHNIQUES,QUANT AUX CONNAISSANCES,QUANT A L'ESPRIT SCIENTIFIQUE...,C
 11    1111   1111111 1 1 1   11     11111111111 1 1            1111 1111111111111
'EST ENCORE LES PROCHAINS POUVOIRS DE L'HOMME SUR SA PROPRE PERSONNALITE BIOLOGI
QUE QUI PARAISSENT LES PLUS BOULEVERSANTS.CE SONT EUX QUI ONT CONDUIT M.JEAN ROS
111111 11   1 111111  11  111 111111111111  1  11  11 1 11  11   1111111111111
TAND A S'ECRIER:"OU DONC APPRENDRE LE METIER DE DIEU?"NOUS DECOUVRONS QUE CHAQUE
 ETRE HUMAIN EST UNE MOLECULE ORGANIQUE,MOLECULE GEANTE ET COMPLEXE AU SENS QUE
11  111 111111 11   11   111111  111 1111  1111111  11111  1       111  1   11111
LA CHIMIE DONNE A CES MOTS,C'EST-A-DIRE:ASSEMBLAGE D'ATOMES LIES ENSEMBLE DE TEL
LE SORTE QUE CHACUN D'EUX EST NECESSAIRE A L'EXISTENCE NORMALE DE L'ENSEMBLE:LA
1111  11111 11   1111 1111111 11111111          1111111  111111        111111111
MOLECULE HUMAINE PEUT ETRE REPRESENTEE,COMME TOUTE MOLECULE ORGANIQUE,PAR UNE"FO
RMULE DEVELOPPEE",ICI IMMENSE,OU CHAQUE ATOME CONSTITUANT EST NECESSAIRE A L'EQU
  1 11    11111111  11    1111  1  11111  1111      11  111  11   11111111111111
ILIBRE DE L'ENSEMBLE.CHAQUE HOMME A UNE FORMULE DISTINCTE DE TOUS LES AUTRES.LA
COMPLEXITE,L'ORIGINALITE ET LA FRAGILITE DE CHAQUE ETRE HUMAIN RESULTE DE CES FA
1       11111       111 111111  1 11    11 1111  1 111111  111 111111     1111  111111
ITS.CES CARACTERES IMPLIQUENT QUE LORSQUE NON SEULEMENT L'UN DES MEMBRES,NON SEU
LEMENT L'UN DES ORGANES,MAIS SEULEMENT L'UN DES ATOMES D'UN CORPS VIVANT EST DET
11111111   . 1  11  111111  111  11111111     1  11 111111     1  1111 11111111111
RUIT OU MENACE,C'EST L'ENSEMBLE DES ATOMES QUI SOUFFRE ET QUI EST MENACE.LES CON
SEQUENCES DE CES FAITS SONT ENORMES:ELLES CONDUISENT A PERSONNALISER,A INDIVIDUA
1111111111  1  11    111 1111    11111  1111    111 1111 1      11 1111111 1     1111
LISER NON SEULEMENT TOUT ACTE MEDICAL MAIS TOUTE RELATION SOCIALE.NOUS Y REVIEND
RONS PLUS LOIN.NE RETENONS ICI QU'UN EXEMPLE NOTABLE:L'UNITE ORGANIQUE DE L'ETRE
111111  111  111  1  11 1111  111  1 1  111111     1111       111  11  1111           1
 VIVANT EXPLIQUE LES LIENS JUSQU'ICI MAL COMPRIS,ET SOUVENT MEME MECONNUS,QUI EX
ISTENT ENTRE LE PSYCHIQUE ET LE PHYSIQUE.EN FAIT,LES EMOTIONS,MEME SI LEUR ELECT
 111111    111  1  11  1111 11   1    1 1111   1    11  11    111111     11111  111  11
ION N'EST QUE MICROPHYSIQUE,AGISSENT SUR LE CORPS MACROPHYSIQUE:LES MALADIES PSY
CHOSOMATIQUES SONT PROVOQUEES PAR LEUR IMPACT SOMATIQUE,C'EST-A-DIRE A L'ECHELLE
11  11111111111  111     1111111  11  111    1111    1111111    11111111111     11 1
 DU CORPS MACROPHYSIQUE.CE FAIT ANNONCE QUE LES EMOTIONS,LA VIE AFFECTIVE,LA SEN
SIBILITE,PHENOMENES INDIVIDUELS ET INDIVIDUALISANTS,ACQUERRONT DANS LA CITE FUTU
111111111  11111111      111111111  1         111 111  1111  1 1111111  111 11  111  1
RE UNE PLACE QU'ELLES SONT LOIN D'AVOIR DANS LES REVENDICATIONS SOCIALES ET DANS
 LES PREOCCUPATIONS POLITIQUES D'AUJOURD'HUI.PLUS PRECISEMENT,LA RECONNAISSANCE
11  11   111111111111 1  1111111    111111111     1      1   11111 11      111111111
CONNEXE DE LA FRAGILITE ET DE LA RELATIVE PLASTICITE DES GRANDES MOLECULES ORGAN
IQUES DONNE A L'HOMME DES RESPONSABILITES QUI SONT DE L'ORDRE DE CE QUE L'ON APP
1 11111    1 1        1111    1    1111111111111 11  111          111   1  11       111111
ELAIT JUSQU'ICI"LA CREATION".CELA NE MODIFIE PAS LE PROBLEME DE L'EXISTENCE DE D
IEU,MAIS SEMBLE AU MINIMUM IMPLIQUER UNE DELEGATION DE POUVOIRS A LAQUELLE NOUS
11111  111  11  1  1    1 111    111111    11    11111111      11111111 1 11111111111
NE SOMMES PAS PREPARES.LE SOCIALISME TRADITIONNEL S'EST BORNE EN EFFET A RENDRE
CLASSIQUES LES SEULS PROBLEMES ECONOMIQUES DU NIVEAU DE VIE ET DE LA PROPRIETE D
11  11111111  11  1111       11 111     111111111 11   1 111   1 111  1      11       11111
ES MOYENS DE PRODUCTION.NOUS DECOUVRONS AUJOURD'HUI SEULEMENT L'AMPLEUR DE LA ME
TAMORPHOSE DE L'HUMANITE.DANS LE CHAPITRE PRECEDENT NOUS AVONS TENTE DE METTRE E
 111111111          11111111  1 1  1  111111    11111 11  111  1111     111  1     1111
N EVIDENCE LES FACTEURS QUI,SOUS NOS YEUX,TRANSFORMENT CE QUE NOUS AVONS APPELE
LA SITUATION DE L'HOMME SUR LA TERRE,C'EST-A-DIRE SON CADRE DE VIE,LA NATURE,LE
1111    1111111          11111        11  1111  111  1 1111   1  1111  1  11 11   111111
```

Annexe 6 (Cf chap. 6, p. 297)

```
C          CALCUL DES SCORES TOTAUX ET DES CORRELATIONS INTERECHELLES
C            DANS L'EXPERIENCE DE LA FIDELITE (STABILITE) DES CERTITUDES
      DIMENSION CSQ(2,3,40)
      INTEGER * 2 C(2,3,100),RC(100),L(100),KEY(100),KZ(3),H(2,3)
      KZ(1)=1
      KZ(2)=2
      KZ(3)=2
      READ(5,504)((CSQ(I,1,K),K=1,4),I=1,2),((CSQ(I,2,K),K=1,10),I=1,2),
     1((CSQ(I,3,K),K=1,40),I=1,2)
  504 FORMAT(10(10F8.3))
C     LECTURE INDIVIDUELLE
    1 READ(3,303,END=3)NUM,((H(L1,L2),L2=1,3),L1=1,2),CRI
      READ(3,502)(L(I),I=1,100)
  502 FORMAT(100A1)
      READ(3,501)(RC(I),I=1,100)
  501 FORMAT(100I2)
  303 FORMAT(I3,1X,6I1,1X,I1)
      DO 2 L1=1,2
      DO 2 L2=1,3
      READ(3,300)(C(L1,L2,I),I=1,100)
  300 FORMAT(100I2)
    2 CONTINUE
      IF(H(1,3).GT.0.AND.H(2,3).GT.0)GOTO 27
      GOTO 1
   27 CONTINUE
      WRITE(6,601)NUM
  601 FORMAT('0----------',I4)
      DO 8 L2=1,3
      DO 8 L1=1,2
      T=0
      DO 7 I=1,100
      KI=RC(I)
      K=KZ(KI)
    7 T=T+CSQ(L1,L2,K)
    8 WRITE(6,600)T
  600 FORMAT(' ',F10.4)
      DO 5 LX=1,3
      DO 5 LY=1,3
      SX=0
      SY=0
      SXY=0
      DO 28 I=1,100
      SX=SX+C(1,LX,I)
   28 SY=SY+C(2,LY,I)
      XM=SX/100
      YM=SY/100
      SX=0
      SY=0
      SXY=0
      DO 29 I=1,100
      X=C(1,LX,I)
      Y=C(2,LY,I)
      SX=SX+(X-XM)**2
      SY=SY+(Y-YM)**2
   29 SXY=SXY+(X-XM)*(Y-YM)
      R=SXY/SQRT(SX*SY)
      WRITE(6,600)R
    5 CONTINUE
      GOTO 1
    3 STOP
      DEBUG SUBCHK
```

```
C        PROGRAMME FXL
       INTEGER * 2 C,NU,TT,H,M1(40),M2(40),M3(40),TR(200,20),TB(6,20),
      1TBR(100),RC,L,KEY
       COMMON C(2,3,100),NU(3,3,45,45),L(100),KEY(100),RC(100),TT(3,3),
      1H(2,3)
       REAL NU1,NU2
       REAL MA(3,3),MB(3,3)
       DIMENSION SA(3,3),SB(3,3),RAB(3,3)
       DIMENSION LX(5),KEYX(5),CX(2,5)
       DIMENSION SAB(3,3),KZ(3)
       KZ(1)=1
       KZ(2)=0
       KZ(3)=0
    94 CONTINUE
       DO 6 K1=1,3
       DO 6 K2=1,3
       TT(K1,K2)=0
       DO 6 K3=1,45
       DO 6 K4=1,45
     6 NU(K1,K2,K3,K4)=0
       DO 42 I=1,100
    42 TBR(I)=0
       DO 40 LAR=1,20
       DO 41 MR=1,200
    41 TR(MR,LAR)=0
       DO 40 NBR=1,6
    40 TB(NBR,LAR)=0
       NSUJ=0
       READ(5,500)(KEY(I),I=1,100)
   500 FORMAT(50A1)
C      LECTURE INDIVIDUELLE
     1 READ(3,303,END=3)NUM,((H(L1,L2),L2=1,3),L1=1,2),CRI
       NSUJ=NSUJ+1
       READ(3,502)(L(I),I=1,100)
   502 FORMAT(100A1)
       READ(3,501)(RC(I),I=1,100)
   501 FORMAT(100I2)
   303 FORMAT(I3,1X,6I1,1X,I1)
       DO 43 I=1,100
       KI=RC(I)
    43 TBR(I)=TBR(I)+KZ(KI)
       DO 2 L1=1,2
       DO 2 L2=1,3
       READ(3,300)(C(L1,L2,I),I=1,100)
   300 FORMAT(100I2)
     2 CONTINUE
C--------- DESSIN DES HISTOGRAMMES
       DO 5 L1=1,2
       DO 5 L2=1,3
       IF(H(L1,L2).GE.1)CALL HIST(L1,L2,NUM)
     5 CONTINUE
       GOTO 1
     3 STOP
       DEBUG SUBCHK
       END
```

./.

./..

```
      SUBROUTINE HIST(L1,L2,NUM)
C      KC=INDICE DE CLASSE DE L'HISTOGRAMME
C     V(KC,2)= VALEURS SIMPLES PUIS CUMULEES DE CHAQUE CLASSE DE L'HISTOGRAMME
C     L1=INDICE  AVANT(1)   APRES(2)
C     L2= INDICE D'ECHELLE 4 DEGRES(1)   10 DEGRES (2)   40 DEGRES(40)
      COMMON C(2,3,100),NU(3,3,45,45),L(100),KEY(100),RC(100),TT(3,3),
     1H(2,3)
      DIMENSION X(100)
      DIMENSION NT(3)
      INTEGER * 2 C,NU,TT,V(10,2),H,RC,L,KEY
      DATA B/' '/,X/100*'*'/
      WRITE(6,601)NUM
  601 FORMAT('0','CODE = ',I3//)
      NT(1)=4
      NT(2)=10
      NT(3)=10
      DO 2 KC=1,10
    2 V(KC,1)=0
      DO 5 I=1,100
      IF(L2.EQ.3)C(L1,L2,I)=C(L1,L2,I)/4
      KC=C(L1,L2,I)
      IF(L2.LT.3)KC=KC+1
      IF (KC.EQ.0) KC=1
    5 V(KC,1)=V(KC,1)+1
C    LA PREMIERE VALEUR CUMULEE VAUT LA PREMIERE VALEUR SIMPLE
      V(1,2)=V(1,1)
      DO 4 KC=2,10
      KB=KC-1
    4 V(KC,2)=V(KB,2)+V(KC,1)
      NX=NT(L2)
      DO 3 KC=1,NX
      N=V(KC,1)
      IF(N)7,7,8
    8 WRITE (6,600)KC,V(KC,1),V(KC,2),(X(I),I=1,N)
  600 FORMAT(' ',3I4,100A1)
      GOTO 3
    7 WRITE(6,600)KC,V(KC,1),V(KC,2),B
    3 CONTINUE
      RETURN
      DEBUG SUBCHK
      END
```

ANNEXE 7 (Cf. p. 297)

```
CODE = 118

  1   2    2**
  2  13   15*************
  3  34   49**********************************
  4  51  100***************************************************

CODE = 118

  1   5    5*****
  2   9   14*********
  3  24   38************************
  4  62  100**************************************************************

CODE = 185

  1   0    0
  2  20   20********************
  3  37   57*************************************
  4  43  100*******************************************

CODE = 185

  1   0    0
  2  13   13*************
  3  32   45********************************
  4  55  100*******************************************************

CODE = 236

  1  22   22**********************
  2  21   43*********************
  3  12   55************
  4  45  100*********************************************

CODE = 236

  1  17   17*****************
  2  22   39**********************
  3  23   62***********************
  4  38  100**************************************

CODE =  22

  1  54   54******************************************************
  2  25   79*************************
  3  13   92*************
  4   8  100********

CODE =  22

  1  45   45*********************************************
  2  37   82*************************************
  3   8   90********
  4  10  100**********

CODE = 132

  1  16   16****************
  2  37   53*************************************
  3  22   75**********************
  4  25  100*************************

CODE = 132

  1  13   13*************
  2  46   59**********************************************
  3  30   89******************************
  4  11  100***********

CODE = 178

  1  27   27***************************
  2  21   48*********************
  3  15   63***************
  4  37  100*************************************

CODE = 178

  1  22   22**********************
  2  13   35*************
  3  29   64*****************************
  4  36  100************************************

CODE = 170

  1  23   23***********************
  2  17   40*****************
  3  17   57*****************
  4  43  100*******************************************

CODE = 170

  1   9    9*********
  2  16   25****************
  3  23   48***********************
  4  52  100****************************************************

CODE = 251

  1  20   20********************
  2  41   61*****************************************
  3  14   75**************
  4  25  100*************************

CODE = 251

  1  19   19*******************
  2  34   53**********************************
  3  23   76***********************
  4  24  100************************
```

```
CODE = 179

  1   2    2**
  2  25   27*************************
  3  34   61**********************************
  4  39  100***************************************

CODE = 179

  1   3    3***
  2  26   29**************************
  3  43   72*******************************************
  4  28  100****************************

CODE = 179

  1   0    0
  2   8    8********
  3  12   20************
  4   7   27*******
  5   2   29**
  6   1   30*
  7   5   35*****
  8  34   69**********************************
  9   9   78*********
 10  22  100**********************

CODE = 179

  1   9    9*********
  2  11   20***********
  3   7   27*******
  4   2   29**
  5   1   30*
  6   4   34****
  7  24   58************************
  8  16   74****************
  9  20   94********************
 10   6  100******

CODE = 108

  1  26   26**************************
  2  23   49***********************
  3  19   68*******************
  4  32  100********************************

CODE = 108

  1   5    5*****
  2  25   30*************************
  3  36   66************************************
  4  34  100**********************************

CODE = 108

  1  10   10**********
  2   7   17*******
  3  12   29************
  4   5   34*****
  5   6   40******
  6  13   53*************
  7  12   65************
  8   8   73********
  9   4   77****
 10  23  100***********************

CODE = 108

  1  17   17*****************
  2   9   26*********
  3   8   34********
  4   6   40******
  5  13   53*************
  6  11   64***********
  7  10   74**********
  8   3   77***
  9   0   77
 10  23  100***********************

CODE =  21

  1  39   39***************************************
  2  33   72*********************************
  3  15   87***************
  4  13  100*************

CODE =  21

  1  30   30******************************
  2   4   34****
  3   9   43*********
  4  22   65**********************
  5   5   70*****
  6   3   73***
  7   4   77****
  8   8   85********
  9   3   88***
 10  12  100************

CODE =  21

  1  31   31*******************************
  2   8   39********
  3   7   46*******
  4  19   65*******************
  5   4   69****
  6   4   73****
  7   6   79******
  8   7   86*******
  9   2   88**
 10  12  100************

CODE =  21

  1   6    6******
  2  20   26********************
  3  19   45*******************
  4   7   52*******
  5  11   63***********
  6   8   71********
  7   7   78*******
  8   6   84******
  9  14   98**************
 10   2  100**
```

```
CODE = 206

  1   2   2**
  2   4   6****
  3   3   9***
  4   8  17********
  5  10  27**********
  6  10  37**********
  7  13  50*************
  8  11  61***********
  9  15  76***************
 10  24 100************************
CODE = 206

  1   7   7*******
  2  11  18***********
  3   7  25*******
  4   4  29****
  5   1  30*
  6  10  40**********
  7  10  50**********
  8  12  62************
  9  11  73***********
 10  27 100***************************
CODE = 258

  1   1   1*
  2   2   3**
  3  12  15************
  4  15  30***************
  5  16  46****************
  6  10  56**********
  7   8  64********
  8  11  75***********
  9   1  76*
 10  24 100************************
CODE = 258

  1   3   3***
  2   7  10*******
  3   6  16******
  4  11  27***********
  5  10  37**********
  6   7  44*******
  7   5  49*****
  8   3  52***
  9  12  64************
 10  36 100************************************
CODE = 306

  1   0   0
  2   1   1*
  3   6   7******
  4  14  21**************
  5   2  23**
  6  34  57**********************************
```

```
CODE = 171

  1   8   8********
  2  10  18**********
  3  12  30************
  4   4  34****
  5   4  38****
  6  21  59*********************
  7  10  69**********
  8   9  78*********
  9   5  83*****
 10  17 100*****************
CODE = 171

  1  10  10**********
  2   5  15*****
  3   6  21******
  4   4  25****
  5  13  38*************
  6   8  46********
  7  21  67*********************
  8   5  72*****
  9  22  94**********************
 10   6 100******
CODE = 197

  1  13  13*************
  2  10  23**********
  3  14  37**************
  4  16  53****************
  5   7  60*******
  6   1  61*
  7   6  67******
  8   6  73******
  9   3  76***
 10  24 100************************
CODE = 197

  1  11  11***********
  2  10  21**********
  3   6  27******
  4   4  31****
  5   9  40*********
  6   8  48********
  7   8  56********
  8  12  68************
  9  32 100********************************
 10   0 100
```

```
CODE = 107

  1   1   1*
  2   0   1
  3  36  37************************************
  4   0  37
  5   0  37
  6  25  62*************************
  7   0  62
  8   1  63*
  9   0  63
 10  37 100*************************************
CODE = 107

  1   1   1*
  2  60  61************************************************************
  3   0  61
  4   0  61
  5   0  61
  6   0  61
  7  38  99**************************************
  8   1 100*
  9   0 100
 10   0 100
CODE = 131

  1   0   0
  2  16  16****************
  3   6  22******
  4  14  36**************
  5   2  38**
  6  22  60**********************
  7   1  61*
  8   7  68*******
  9  13  81*************
 10  19 100*******************
CODE = 131

  1   4   4****
  2   8  12********
  3   1  13*
  4   2  15**
  5  21  36*********************
  6   4  40****
  7   7  47*******
  8  11  58***********
  9  41  99*****************************************
 10   1 100*
```

VOTRE CODE = 104

	BR	MR	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
1	0	0																																								
2	0	0																																								
3	0	0																																								
4	0	0																																								
5	0	0																																								
6	0	0																																								
7	0	0																																								
8	0	0																																								
9	0	0																																								
10	0	0																																								
11	0	0																																								
12	2	1																1				1				1																
13	0	0																																								
14	0	0																																								
15	0	0																																								
16	0	3																				2					1															
17	0	0																																								
18	0	0																																								
19	0	0																																								
20	3	11																1							1	1	1	1	1	1		2	1	2				1		1		
21	0	0																																								
22	0	0																																								
23	0	0																																								
24	2	2																								1				2				1								
25	0	0																																								
26	2	1																						1						1							1					
27	1	0																																						1		
28	5	3																								2				2								2		2		
29	0	0																																								
30	0	1																								1																
31	0	0																																								
32	10	9																								3				3	1	4				1		2		2	1	2
33	0	0																																								
34	1	1																1																				1				
35	0	0																																								
36	11	4																				1						1		1			1	3		1		2		1	3	1
37	0	0																																								
38	8	2																																3		1	1		1	1	1	2
39	7	1																								1							1					2		2	2	
40	7	2																																				1			1	7
BR =			0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	5	0	1	0	3	0	2	3	6	0	3	2	8	1	9	6	9
MR =			0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	3	0	1	1	5	2	1	1	7	1	4	0	3	0	0	0	3	0	1	2	3

M1 = 30.87 SIG 1 = 7.72
M2 = 31.89 SIG 2 = 6.56
R = 0.5942

VOTRE CODE = 104
(AVANT) 40 DEGRES

```
 3  *
 7  *
11  ***
15  ***
19  **************
23  ****
27  ************
31  ********************
35  *****************
39  ***************************
```

VOTRE CODE = 104
(APRES) 40 DEGRES

```
 3  *
 7  *
11  *
15  ***
19  ****
23  ************
27  ***************
31  *******************
35  ****************
39  *******************************
```

VOTRE CODE = 308

	BR	MR	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
1	0	0																																								
2	0	0																																								
3	0	0																																								
4	0	0																																								
5	0	0																																								
6	0	1										1																														
7	1	1						1								1																										
8	0	0																																								
9	1	2							1		1	1																														
10	13	23	1					9	1		1	11	4				3			1	3						1										1					
11	0	3														1				1				1																		
12	0	0																																								
13	0	0																																								
14	3	4						2				2	1			1	1																									
15	5	2						2				2	1			1													1													
16	0	0																																								
17	2	1										1	1																								1					
18	1	3										1				1				1					1																	
19	1	0						1																																		
20	0	0																																								
21	0	0																																								
22	3	0										1													1											1						
23	1	0										1																														
24	0	0																																								
25	5	0						1				1				1																1					1					
26	2	0											1								1																					
27	0	0																																								
28	0	0																																								
29	1	0																																		1						
30	0	0																																								
31	0	0																																								
32	0	1	1																																							
33	1	1																							1											1						
34	4	0																									1	1									1					1
35	2	0										1																					1									
36	0	0																																								
37	2	0																																			1					1
38	4	0																							1												3					
39	0	0																																								
40	6	0																																					1			5
BR =			0	0	0	0	0	7	1	0	1	10	5	0	0	4	3	0	0	0	1	0	0	0	2	0	2	1	1	0	0	1	1	0	0	3	7	0	1	0	0	7
MR =			2	0	0	0	0	9	1	0	1	13	3	0	0	2	1	0	0	3	3	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0

M1 = 18.50 SIG 1 = 10.62
M2 = 16.99 SIG 2 = 11.26
R = 0.7314

```
VOTRE CODE =  308
(AVANT)  40  DEGRES
 3  *
 7  ***
11  ******************************************
15  **************
19  ********
23  ****
27  *******
31  **
35  ********
39  ************

VOTRE CODE =  308
(APRES)  40  DEGRES
 3  **
 7  ******************
11  ********************************
15  **********
19  *******
23  *****
27  ****
31  **
35  ***********
39  ********
```

TABLE DES MATIERES