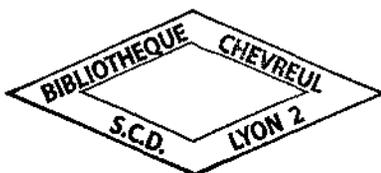
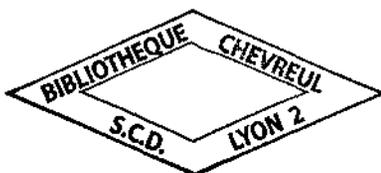


S O M M A I R E



INTRODUCTION.		5
PREMIERE PARTIE:	APPROCHE PREALABLE DE L'EVALUATION EN ARTS PLASTIQUES: UN CERTAIN TYPE D'ERRANCE ET SES FONDEMENTS SUPPOSES.	11
■ 1.	<u>ENONCE DE QUELQUES CONSTATS INITIAUX.</u>	13
1.1.	DE L'ECTOPLASME A L'ICONOCLASTE.	14
	111 - L'évidence des désaccords.	14
	112 - Difficulté et illégitimité de la notation.	17
	113 - De la question du (bon) goût.	19
1.2.	PETITE HISTOIRE D'UNE GESTION EMPIRIQUE.	22
	121 - La notation en A.P.: Evolution des données institutionnelles.	23
	122 - Comment notent aujourd'hui les P.A.P.?	27
	123 - Nécessité d'une formation à l'évaluation.	32
■ 2.	<u>ARCHITECTURE NAISSANTE D'UNE INTERROGATION SUR LE PHENOMENE DE L'EVALUATION EN ARTS PLASTIQUES.</u>	35
2.1.	DESENFOUISSEMENT D'UN PREMIER DEGRE D'INTELLIGIBILITE.	37
	211 - Le concept d'évaluation.	37
	212 - Objet et fonction de l'évaluation en arts plastiques.	42
	213 - Attentes et représentations.	49
2.2.	DE L'IDENTIFICATION DE QUELQUES MODELES AU CIBLAGE DE NOS INTERROGATIONS.	53
	221 - "Dessin", "arts plastiques" ou "éducation plastique".	53
	222 - Diverses modélisations plus implicites.	57
	223 - Problématisation de la recherche.	62

S O M M A I R E



INTRODUCTION.		5
PREMIERE PARTIE:	APPROCHE PREALABLE DE L'EVALUATION EN ARTS PLASTIQUES: UN CERTAIN TYPE D'ERRANCE ET SES FONDEMENTS SUPPOSES.	11
■ 1.	<u>ENONCE DE QUELQUES CONSTATS INITIAUX.</u>	13
1.1.	DE L'ECTOPLASME A L'ICONOCLASTE.	14
	111 - L'évidence des désaccords.	14
	112 - Difficulté et illégitimité de la notation.	17
	113 - De la question du (bon) goût.	19
1.2.	PETITE HISTOIRE D'UNE GESTION EMPIRIQUE.	22
	121 - La notation en A.P.: Evolution des données institutionnelles.	23
	122 - Comment notent aujourd'hui les P.A.P.?	27
	123 - Nécessité d'une formation à l'évaluation.	32
■ 2.	<u>ARCHITECTURE NAISSANTE D'UNE INTERROGATION SUR LE PHENOMENE DE L'EVALUATION EN ARTS PLASTIQUES.</u>	35
2.1.	DESENFOUISSEMENT D'UN PREMIER DEGRE D'INTELLIGIBILITE.	37
	211 - Le concept d'évaluation.	37
	212 - Objet et fonction de l'évaluation en arts plastiques.	42
	213 - Attentes et représentations.	49
2.2.	DE L'IDENTIFICATION DE QUELQUES MODELES AU CIBLAGE DE NOS INTERROGATIONS.	53
	221 - "Dessin", "arts plastiques" ou "éducation plastique".	53
	222 - Diverses modélisations plus implicites.	57
	223 - Problématisation de la recherche.	62

TROISIEME PARTIE :	EXPLOITATIONS ET PROPOSITIONS: QUELQUES IMPLICATIONS PEDAGOGIQUES.	178
■ 6.	<u>VERS UNE MEILLEURE EVALUATION DE L'EDUCATION PLASTIQUE ?</u>	179
6.1.	ASPERITES NOUVELLES SUR LA VOIE DU BIEN EVALUER.	180
611 -	Peut-on espérer harmoniser les jugements en A.P. ?	180
612 -	Notation critériée: Que peut-on attendre de la concertation?	185
613 -	La question de la relation d'ordre.	200
6.2.	JUGER EN ARTS PLASTIQUES: FIABILITE DES INDICATEURS.	209
621 -	Légitimité des examens et concours.	211
622 -	Evaluer une production ou évaluer son auteur?	221
623 -	Les objectifs de l'éducation artistique sont-ils opérationnalisables?	226
6.3.	ORIENTATIONS POUR UNE EVALUATION PLUS PERFORMANTE.	233
631 -	Evaluation collective et auto-évaluation.	233
632 -	Recherche d'indicateurs comportementaux dans le champ scolaire élargi.	237
633 -	Proposition de dispositifs d'évaluation.	242
■ 7.	<u>DE L'EVALUATION COMME OUTIL AIDANT A LA FORMULATION D'AUTRES PROCEDURES PEDAGOGIQUES.</u>	250
7.1.	PLAIDOYER POUR UNE NON-FONDAMENTALISATION DES ACTIVITES PLASTIQUES.	252
711 -	La P.P.O. au temps de la post-créativité.	252
712 -	La création plastique: Désir et initiative.	258
713 -	Une voie tracée au travers de la pensée éducative.	264
7.2.	POUR UNE GESTION HEUREUSE DE LA DIMENSION CREATRICE.	269
721 -	Comment concilier apprentissage et création?	269
722 -	Quelques lieux de l'expression plastique, hors l'institution scolaire.	276
723 -	Analyse des situations nouvelles introduites en collège.	280
7.3.	UNE PETITE UTOPIE ET SES CONSEQUENCES INSTITUTIONNELLES.	285
731 -	Le concept de "classe-atelier".	286
732 -	Les "Activités Artistiques d'Accompagnement" (3A).	295
733 -	L'évaluation repensée dans le cadre de l'espace-atelier.	299
	CONCLUSION.	306
	ANNEXES.	317
	BIBLIOGRAPHIE.	344

DEUXIEME PARTIE (EXPERIMENTALE).

2

**RECHERCHE
DE QUELQUES UNS DES DETERMINANTS
DE L'EVALUATION
DES PRODUCTIONS PLASTIQUES.**

ANCRAGE THEORIQUE ET OUTILLAGE METHODOLOGIQUE.

3

3.1. CONNAISSANCES PRELIMINAIRES: CHAMPS D'INVESTIGATION.

- 311 - Apport de la docimologie générale.
- 312 - Lecture del'oeuvre: Sémiotique des faits visuels.
- 313 - Psychologie de l'esthétique.

3.2. METHODOLOGIE DE LA RECHERCHE.

- 321 - Problèmes relatifs au recueil des données.
- 322 - La notation "en laboratoire": Protocoles expérimentaux.
- 323 - La question du traitement statistique des résultats.

Nous disposons, pour sous-tendre nos protocoles expérimentaux, de connaissances émanant de plusieurs corpus référentiels.

Nous nous placerons de prime abord sur le terrain de la psychologie de l'évaluation scolaire dont les recherches récentes sont de la plus haute importance: La connaissance du schéma décrivant les étapes du mécanisme de l'évaluation fournira, dans un premier temps, les outils initiaux nécessaires à l'appréhension du comportement du professeur évaluateur d'objets plastiques.

Mais l'on sait déjà qu'une partie des disparités de jugement entre correcteurs, lors de l'activité cognitive de la reconnaissance des indices utiles à l'évaluation, provient d'effets systématiques liés au fonctionnement de mécanismes perceptifs.

Il apparaît alors que pour mieux comprendre le désarroi du P.A.P., confronté à un objet visuel qui se perçoit dans une sorte de globalité où de multiples niveaux de sens peuvent être présents, face à la "co-occurrence" d'une infinité d'indices, seul le secours des connaissances actuelles en matière de sémiotique visuelle pourrait permettre, associées aux observations déjà anciennes de la psychologie de l'esthétique, d'énoncer quelques hypothèses quant à d'éventuels déterminants spécifiques.

3.1. CONNAISSANCES PRELIMINAIRES: CHAMPS D'INVESTIGATION.

3. 1. 1.

APPORT DE LA DOCIMOLOGIE GENERALE.

Du grec dokimé (épreuve), la docimologie est l'étude systématique des examens. C'est H. PIERON, un des premiers français à entreprendre, en 1922, avec son collègue H. LAUGIER, une critique détaillée de la notation des examens, qui lui donna son nom.

Toute étude docimologique s'inaugure par le constat des divergences d'évaluation observées et la citation d'exemples frappants et caractéristiques, aussi ne ferons-nous que rappeler très brièvement quelques étapes.

Les premières recherches (PIERON, LAUGIER, WEINBERG) furent entreprises en 1935. En 1936, paraissaient sous le titre "La correction des épreuves écrites dans les examens" les conclusions de la commission CARNEGIE française, à la suite d'une enquête internationale suggérée par les américains. Ces investigations mirent en évidence les différences d'appréciation d'un correcteur à l'autre et conclurent par une mise en doute de la fiabilité des examens.

Les chiffres recueillis à cette occasion sont impressionnants pour qui les lit pour la première fois: La commission, s'appuyant sur des copies de Baccalauréat, put constater en effet des écarts allant jusqu'à 13 points dans la correction de la composition française. Cette information provoqua une certaine émotion mais n'eût, aux dires de PIERON (1), que bien peu d'échos. Depuis, d'autres expériences, notamment celle menée en 1967 par le Service de la Recherche du C.R.D.P. de Lyon, portant sur la multicorrection de rédactions effectuées par 150 professeurs de français, ont révélé des marges de variation tout aussi importantes.

(1) PIERON (H.), Examens et docimologie, op. cit., page 26.

On sait que le niveau d'exigence varie considérablement selon les examinateurs. G. de LANDSHEERE cite une expérience de R. DUQUENNE (1967) démontrant, même, que les correcteurs n'ont pas une conscience objective de leur degré de sévérité ou d'indulgence (1)!

" Pour prédire la note d'un candidat, il vaut mieux connaître son examinateur que lui-même" s'était exclamé PIERON (2). En serait-il de même en arts plastiques? L'objet de cette étude est évidemment de tenter d'apporter quelques éléments de réponse. S'agissant d'une expression plastique, tout porte à croire que la référence en matière de fiabilité de l'évaluation risque de se rapprocher davantage de la copie de philosophie ou de la rédaction (autre matière d'expression) plutôt que des écarts moindres enregistrés en mathématiques ou physique (où le produit-norme est beaucoup plus facile à cerner).

Quels sont les résultats actuellement en notre possession qui nous permettent de préciser à quoi tiennent les désaccords entre correcteurs?

Depuis les premiers travaux de LAUGIER, WEINBERG et PIERON, de nombreuses expériences de multicorrection ont permis de comparer les distributions de notes de plusieurs évaluateurs selon un certain nombre de facteurs:

Chaque distribution de note est caractérisée par un indice de position (la moyenne, autrement dit une échelle de sévérité/indulgence) et un indice de dispersion (éventail des notes réellement utilisées entre le 0 et le 20). Non seulement ces données varieront d'un évaluateur à l'autre, mais aussi l'ordre de classement: Chacun ordonnera les travaux de manière différente.

Ensuite, c'est le grand mérite du Groupe de Recherche sur l'évaluation de l'Université de Provence animé par J.P. CAVERNI et G. NOIZET et auquel ont participé J.J.BONNIOL, M. PIOLAT, R. AMIGUES et J.M. FABRE, d'avoir pu faire apparaître

- " l'hypothèse que les divergences d'évaluation - constat de la docimologie
- " classique - ne sont pas dues au hasard, ni même à des contrastes dans
- " la personnalité des évaluateurs. La conclusion générale qui se dessine
- " est que l'évaluation d'une production scolaire relève de déterminants

(1) LANDSHEERE (G. de), Evaluation continue et examens. Précis de docimologie, Bruxelles-Paris, Labor-Nathan, 1980 (5^e éd.), page 30.

(2) PIERON (H.) op. cit., page 24.

" systématiques, mais qui ne réfèrent pas tous à des caractéristiques
" proprement scolaires" (1).

1) Le mécanisme de l'évaluation.

Rappelons brièvement le mécanisme général de l'évaluation tel que ces chercheurs l'ont exposé:

La tâche d'évaluation est une tâche de lecture: A ce titre, elle dépend de mécanismes généraux relatifs à la perception (et plus précisément, pour ce qui nous concerne, de mécanismes relatifs à la perception des faits visuels).

C'est aussi une activité de comparaison, ce qui la présente également comme activité cognitive. Cette activité s'établit relativement à un modèle de référence construit à partir de l'estimation initiale de ce qu'il semblerait concevable d'attendre, estimation qui se précise ensuite peu à peu dans l'esprit de l'évaluateur compte tenu des premières observations réellement recueillies.

Ainsi, suivant un certain nombre de points de vue correspondant à des critères, l'évaluateur va recueillir des indices qu'il mettra en relation à la fois avec les informations émanant des autres travaux et avec l'image intérieure de ce qu'il attendait.

L'évaluation repose donc sur la succession de deux ordres d'activité: D'une part, le recueil des indices, tributaire d'une problématique de la perception, d'autre part, la constitution du modèle de référence par rapport auquel s'exercera un comportement cognitif.

La perception des indices est tributaire de notre sensibilité et de nos capacités de repérage: On conçoit qu'il puisse à avoir dénaturation des indices recueillis, interférence d'informations extérieures à la tâche, exploration polarisée des travaux, mauvais repérage des indices pertinents.

J.J. BONNIOL, dans son oeuvre de synthèse (2), a précisé les processus conduisant à la constitution du modèle de référence. Nous retiendrons de son étude trois phases successives:

- La phase d'encadrement, correspondant à la constitution du produit norme issu de l'ensemble des produits possibles et porteur des caractéristiques qui attestent de la réussite de la tâche.
- Le dosage des exigences, qui amène à préciser la nature du produit

(1) NOIZET (G.) et CAVERNI (J.P.), Psychologie de l'évaluation scolaire, op. cit. p. 192.

(2) BONNIOL (J.J.), Déterminants et mécanismes des comportements d'évaluation d'épreuves scolaires, op. cit., pages 75 à 110.

attendu en fonction de ce que le professeur pense obtenir de ses élèves compte tenu de ce qu'il sait sur ces derniers et sur les conditions de la production.

- Le codage et la définition du modèle de référence: L'évaluation ne peut en effet avoir lieu sans repère correspondant à une échelle de mesure.

A chaque niveau interviennent des facteurs subjectifs. Ce schéma général a fait l'objet, durant ces quinze dernières années, d'expérimentations qui ont conduit à faire émerger de nombreux déterminants systématiques que nous présenterons ici succinctement:

2) Déterminants de l'évaluation scolaire:

a) - Les effets "mécaniques".

Nous placerons ici les effets déterminés par les conditions matérielles de l'acte d'évaluation, indépendamment du contenu des objets à évaluer et de toute référence aux personnes à juger.

- Nous avons déjà évoqué:

* Différence de sévérité entre correcteurs.

* Utilisation différente de l'échelle des notes.

- Il faut y ajouter:

* Infidélité d'un même correcteur dans le temps.

Il a été démontré à plusieurs reprises dès 1935 (LAUGIER et WEINBERG, HARTOG et RHODES) que le même examinateur note différemment les mêmes travaux lorsqu'il les examine à nouveau dix mois ou a fortiori plusieurs années après: Le taux de corrélation est souvent inférieur à 0.60 et près de la moitié des travaux reçoivent une note sensiblement différente.

* Un effet séquentiel: L'effet d'ordre.

J.J. BONNIOL, en confrontant lors d'une expérience datant de 1965 (1) la notation de 26 versions anglaises selon qu'elles étaient présentées dans un ordre ou dans l'ordre inverse, a pu montrer une tendance à surévaluer les copies corrigées en premier et à sous-évaluer celles corrigées en dernier.

* L'effet de contraste dans la succession des copies.

L'activité d'évaluation est une activité de comparaison. Il est facile d'imaginer qu'un devoir "moyen" situé au milieu d'excellentes

(1) BONNIOL (J.J.), "Les différences de notation tenant aux effets d'ordre de correction", Cahiers de Psychologie, 1965, 8, pages 181 à 188.

copies paraîtra plus modeste alors qu'il risquera d'être mis en valeur par des travaux insuffisants. Les expériences de BONNIOL et PIOLAT (1) ont contribué à identifier ce qu'ils ont appelé des "effets d'ancrage".

b) - Informations a priori relatives au producteur.

Ces expériences sont à situer parallèlement à l'esprit qui a inspiré le travail de ROSENTHAL et JACOBSON (1968) dont nous avons déjà fait état: L'évaluateur possède souvent un certain nombre d'informations, soit sur les conditions de production du travail, soit sur la personnalité des producteurs, et qui vont peser dans la détermination du produit attendu. Ces effets "d'information" composent le deuxième lot de déterminants systématiques repérés.

THORNDIKE (1920) nommait "effet de halo" le phénomène de brouillard produit par la réputation d'un élève, par son regard, son vêtement, sa prestance physique ou sa diction en cas d'examen oral. On parle aujourd'hui, selon les chercheurs, d'effets d'assimilation (BONNIOL, CAVERNI, NOIZET), d'effet-source (FABRE) ou d'effet de stéréotypie (de LANDSHEERE).

Ces informations ont trait principalement au niveau scolaire de l'élève ainsi qu'à diverses données d'ordre social:

- Statut scolaire
- Performances antérieures
- Statut socio-économique
- Origine ethnique.

Suivant un grand nombre d'expériences destinées à interroger ces différents facteurs, il ressort globalement que les informations qui s'y rapportent induisent chez l'enseignant une tendance à juger des aptitudes présentes et futures en fonction du caractère plutôt favorable ou plutôt défavorable qui s'en dégage:

NOIZET et CAVERNI, proposant une interprétation générale de l'effet des informations a priori (2), expliquent que les informations provoquent des attentes et contribuent à fixer la représentation que l'enseignant se fait de l'élève. Le terme de "stéréotypie" utilisé

(1) BONNIOL (J.J.) et PIOLAT (M.), "Comparaison des effets d'ancrage obtenus dans une tâche d'évaluation", Actes du XVII^e Congrès International de Psychologie appliquée, Liège, 1971, pages 1179 à 1189.

(2) NOIZET (G.) et CAVERNI (J.P.), Psychologie de l'évaluation scolaire, op. cit., page 97.

par G. de LANDSHEERE (1) semble particulièrement heureux: C'est effectivement un cliché que d'imaginer qu'un élève issu d'un milieu défavorisé aura nécessairement des performances moindres, une pesanteur de l'esprit que de continuer de juger un élève conformément à la première impression reçue. On notera (malgré, il est vrai, de fréquents rappels à la prudence) que le livret scolaire peut apparaître, selon l'usage qui en est fait, suivant la formule de P. MARC (2), comme l'instrument d'une "quasi-institutionnalisation des attentes".

Nous venons d'évoquer les principaux acquis susceptibles de nourrir nos recherches. Certes, les chercheurs ont soigneusement écarté du champ de leurs investigations les matières artistiques. Quand elles sont mentionnées, c'est le plus souvent dans le cadre de considérations pessimistes:

A. BONBOIR (3) : " ... Il en est qui, s'ils se laissent difficilement analyser
" de façon complète (dessin, rédaction...), présentent encore
" l'inconvénient de ne pouvoir offrir le modèle..."

G. de LANDSHEERE (4), s'ancrant dans quelque certitude innée:

" La composition française, considérée comme oeuvre
" d'art, échappe à l'évaluation parcellaire "

Le mythe de l'acte créateur aurait-il fait reculer les docimologues? ...A moins que ce ne soit plutôt le manque d'intérêt des enseignants plasticiens pour la docimologie qui ait privé la recherche des collaborateurs qui lui étaient nécessaires sur le terrain!

Nous aurons donc à nous interroger sur ce qu'il en est des mécanismes identifiés, dans le cadre des arts plastiques. Ces indications sont autant d'incitations à interroger notre propre terrain pédagogique: La possibilité de comparer simultanément les oeuvres facilite-t-elle la tâche du P.A.P.? Celui-ci est-il tout autant tributaire des effets-source, ou bien la prégnance visuelle et sensible des objets à évaluer s'impose-t-elle prioritairement, laissant sans prise sur le jugement toute considération extérieure?

Quelques pistes d'expérimentation sont ici précisément suggérées.

(1) LANDSHEERE (G. de), Evaluation continue et examens, op. cit., page 40.

(2) MARC (P.), Autour de la notion pédagogique d'attente, op. cit., pages 107 à 139.

(3) BONBOIR (A.), La docimologie, Paris, P.U.F., 1972, page 18.

(4) LANDSHEERE (G. de), Ibid., page 56.

3.1.2.

LECTURE DE L'OEUVRE;
SEMIOTIQUE DES FAITS VISUELS.

D'une manière assez semblable à ce qui se produit lorsqu'un correcteur prend connaissance d'une copie écrite, appréhender un objet plastique est une activité de lecture.

Chacun connaît l'expression du peintre P. KLEE (1) comparant le contemplateur d'un tableau à un animal pâtureant une prairie.

L. MARIN, dans une importante étude (2), dira:

" Lire, c'est parcourir du regard (...) L'acte de lecture déroule
" ainsi un temps, une succession à l'intérieur de l'instant de vision".

Toutefois, ce circuit est " un circuit aléatoire: Il n'est jamais nécessaire".

P. FRANCASTEL précise de son côté (3):

" L'oeil balaie, d'abord au hasard, le champ figuratif; il est accroché,
" bientôt, par certains éléments qui le renvoient à des formes ou à
" des notions connues de son esprit ".

Le cheminement des yeux a été particulièrement bien mis en évidence par les expériences de l'américain YARBUS (4): A l'aide d'un faisceau lumineux qu'il put rendre solidaire du regard de sujets volontaires examinant une image, ce chercheur recueillit sur une plaque sensible située derrière l'image la trace du parcours des yeux de plusieurs lecteurs successifs.

(1) KLEE (P.), Credo du créateur, (traduction de Schöpferische Konfession, Berlin, 1920), in Théorie de l'art moderne, Paris, Gonthier, 1973, pages 34 à 42.

(2) MARIN (L.), Etudes sémiologiques, Paris, Klincksieck, 1971, pages 19 à 21.

(3) FRANCASTEL (P.), La figure et le lieu, Paris, Gallimard, 1967, page 39.

(4) YARBUS (A.L.), Eye movements and Vision, New York, Plenum Press, 1967, cité par MOLES (A.), L'image communication fonctionnelle, Tournai, Casterman, 1981, page 56.

L'étude comparative du mouvement des yeux a permis de montrer que plusieurs sujets, regardant le même tableau, en appréhendent les éléments de façon sensiblement voisine. Toutefois, elle met aussi en évidence les différences de parcours visuel d'un sujet à l'autre.

Ainsi, la lecture de l'oeuvre se présente d'emblée comme un mécanisme complexe. On prend pleine conscience de cette complexité lorsque l'on sait que, de surcroît, la trace la plus informelle est tout à fait capable de déclencher toute une série d'images dans cet oeil qui fonctionne inexorablement en continu, sauf à l'obstruer d'un mouvement de paupière!

La description de F. MOLNAR, autre spécialiste du mouvement des yeux, en donne une idée précise (1) :

- " Le tableau se présente comme un ensemble de stimulations excitant
- " plus ou moins - et en même temps - un très grand nombre d'éléments
- " rétinien qui envoient des signaux plus ou moins forts vers le cortex
- " visuel où ils sont traités "
- " Même les agglomérations pigmentaires non délimitées, c'est-à-dire
- " sans contour, qui se dégagent lentement en se confondant progressivement
- " avec l'entourage, sont perçues comme des formes "

Ces différentes analyses nous présentent la lecture de l'oeuvre visuelle comme un cheminement aléatoire non irréversible. S'il y a une certaine constance du parcours, ainsi que le montrent les expériences de YARBUS, chaque lecture est particulière, différente d'un sujet à l'autre, mais aussi pour un même sujet à quelques instants d'intervalle. Aucune lecture n'épuisera jamais l'image, on ne peut espérer que des vagues d'identification de plus en plus précises et l'on songe à cette expression de BENVENISTE, parlant de la phrase dans son "Traité de linguistique générale" (1974): Ne pourrait-on pas affirmer, en effet, que, comme la phrase, la lecture d'image est un "événement évanouissant"?

Errance du (ou des) sens, part importante et peu maîtrisable de l'action de ce regardeur, telles sont les conditions qui président à l'appréhension des objets visuels.

Un dernier élément vient conforter la thèse de l'infinitude des visions. Il porte sur le degré de conscience qui accompagne la perception d'un produit visuel:

De toute évidence, l'image ne communique qu'à partir du moment où le lecteur possède les outils conceptuels lui permettant

(1) MOLNAR (F.), "Quelques aspects psycho-biologiques de l'image", in Revue d'Esthétique, n°1, Paris, U.G.E., 10/18, 1976, pages 339 à 368.

d'"entrer" dans son discours, sinon l'image n'est qu'un amalgame de traces. La lecture implique donc nécessairement des codes de lecture et l'on peut penser qu'à cet égard le P.A.P. est de ceux qui possèdent la compétence nécessaire à son accomplissement. Toutefois, la réarticulation sémiotique qui en découle (c'est-à-dire, après la perception d'une texture plus ou moins précise et signifiante, la mise en ordre, la "traduction" en pensée consciente au travers d'un "texte qui la prend en charge et l'articule" (1)), l'acte technique qui pèse la relation entre le contenu signifié et le plan des expressions significatives (ce qui est véritablement l'objet de l'évaluation d'une production plastique), ne pourrait bien être, au mieux de sa mise en oeuvre, qu'à l'image de la part émergée de l'iceberg...

En effet, l'appréhension des objets visuels ne relève pas seulement de la simple lecture, c'est-à-dire d'un décodage que l'on pourrait qualifier de "mécanique" ou de convenu. Contrairement à la lecture d'un texte ou, plutôt, davantage que pour l'écriture, la perception de l'image implique le phénomène perceptif en sa totalité sensorielle. Alors, la perception de l'image ne relèverait pas immédiatement de l'ordre du discours, mais serait dotée, selon J. KRISTEVA (2),

" d'une modalité qu'on pourrait dire psychosomatique,
 " en prise directe sur le corps, (...), comme un moment
 " logiquement, génétiquement, productivement antérieur au symbolique,
 " mais qui fait dans celui-ci l'objet d'une relève par laquelle il s'intègre "

C'est bien ce qui ressort aussi des études scientifiques les plus récentes, ainsi que le confirme F. MOLNAR (3) :

" A la lumière de la neurophysiologie contemporaine, il devient évident
 " qu'une forme, avant d'être la représentation d'une chose réelle
 " ou imaginaire peut déclencher la réponse perceptive (...). L'excitation
 " affective n'a pas besoin d'être médiatisée par un contenu analysable
 " en termes signifiants (...). Il est probable qu'avant même que nous
 " arrivions à franchir le seuil de la perception purement formelle,
 " le processus qui déclenche la réponse esthétique - la réponse esthétique - est déjà entamé grâce à l'intervention de mécanismes psychophysologiques ou psycho-biologiques à partir, précisément, des
 " lignes et des formes "

(1) SCHEFER (J.L.), Scénographie d'un tableau, Paris, Seuil, 1969, texte II, pages 109 à 144: SCHEFER parle de "textualisation".

(2) KRISTEVA (J.), La révolution du langage poétique, Paris, 1974, citée par DAMISCH "Huit thèses pour (ou contre) une sémiologie de la peinture", Paris, MACULA n°2, 1977, pages 17 à 23.

(3) MOLNAR (F.), Quelques aspects psycho-biologiques de l'image, op. cit., pages 361 et 362.

Dans un autre article (1), F. MOLNAR rappelle la théorie qui fut élaborée par le physiologiste allemand HELMHOLTZ (1821-1894) :

" Le récepteur effectue quelque chose de semblable à une "fonction de calcul", à partir des stimulus. Mais ni les sensations, ni le calcul fait ne sont conscients. Seule, la conclusion perceptive est éprouvée ".

Moment antérieur à la position consciente du sujet, extérieur à toute signification ou opinion constituée, mais dont il y aurait à se préoccuper lorsque l'on connaît, par ailleurs, notre attirance spontanée vers la consonance cognitive qui nous poussera ensuite inconsciemment à sélectionner en priorité les indices congruents à notre première "impression".

Nous avons réuni là les éléments qui ont pu être à l'origine de ce que M. DUFRENNE, soulignant le rôle prépondérant du corps, nomme "l'intellection corporelle" (2) : Lire une image, c'est montrer le rapport qu'on entretient avec cette image. Au travers de la lecture, c'est aussi le lecteur qui se dévoile.

Imperceptiblement, à l'idée de l'image comme donnée objective susceptible d'être atteinte, vient se superposer l'idée d'une image n'existant que comme stimulus extérieur dont la prise en compte par le regard serait en grande partie l'affaire du pouvoir créateur du sujet lecteur, reformulation étrange du schéma de JAKOBSON où le récepteur serait avant tout... émetteur!

Ce que l'on sait de la lecture de l'image amène à s'interroger sur ce qu'il peut en être de l'évaluation des productions plastiques dès l'instant où l'on envisage que sa pratique est en premier lieu tributaire de l'activité du regard. On ne peut douter que les particularités signalées n'interviennent au moment de la prise de connaissance des objets à évaluer, ce qui amène à poser plusieurs jalons dans notre recherche des hypothèses de travail.

- En premier lieu, les différences dans les modalités de lecture de plusieurs sujets confirment notre crainte quant à la possibilité d'atteindre une meilleure homogénéité d'un correcteur à l'autre.

- En second lieu, le mécanisme même d'aperception des

(1) MOLNAR (F.), "Quoi de neuf dans la perception visuelle de la forme?", Paris, Année Psychologique, 1980, 80, pages 599 à 629.

(2) DUFRENNE (M.), Phénoménologie de l'Expérience esthétique, Paris, P.U.F., 1953, (2^e éd. 1967), 4^e partie, pages 657 à 677.

éléments de l'image reposant sur une chronologie d'émergence d'indices jamais fixée, y compris pour le même correcteur pour deux lectures successives, la hiérarchisation des indices, autrement dit ce qui appelle prioritairement l'attention du lecteur-évaluateur, peut être indéfiniment l'objet de modifications notoires. Cette observation invite à mieux connaître la psychologie de chaque lecteur, ce qu'il peut en être de ses attentes.

- Mais le point précédent rappelle aussi (fait connu, il est vrai, de tout plasticien) que "tout fait sens", y compris les taches les plus évanescentes, les accidents incontrôlés par le producteur de l'image, mais encore tout aussi bien (pourquoi pas?) la mauvaise qualité d'un papier, d'un support, son froissage, la fadeur de certaines gouaches aux pigments manquant de densité, c'est-à-dire autant d'éléments parasites, non comptabilisables formellement, mais porteurs d'influence pour qui ne saurait s'en prémunir.

- Enfin, de toutes les informations se rapportant au mécanisme de la lecture des images, la plus interpellante est certainement l'affirmation d'une antériorité de la sensation visuelle sur la perception consciente. Si l'on admet une action de l'image sur la personne, antérieure au nommé, si, comme l'expose F. MOLNAR, les grandes composantes de l'image en sont les actrices, dès lors se précise le processus qui peut conduire à une perception polarisée: L'entrée dans le document visuel apparaît alors comme amorcée, "guidée", s'élaborant dans la continuité de la phase préconsciente.

La lecture d'image se place là sur le terrain de la psychologie de la perception esthétique et nous invite à nous demander si les diverses dimensions plastiques de l'oeuvre n'exercent pas une action souterraine sur le jugement des professeurs.

3.1.3.

PSYCHOLOGIE DE L'ESTHÉTIQUE.

Pour FECHNER (1), fondateur de l'esthétique expérimentale, cette science se propose d'étudier

" tout ce qui pénètre en nous par les sens sous la forme d'un agrément
" ou d'un désagrément immédiats ".

Si l'esthétique peut se définir comme la science de la sensibilité, le mérite de FECHNER a été de montrer que le goût est directement tributaire d'une psychologie de la perception ou, plutôt, de la réaction, dont le schéma serait:

stimulus → personnalité → réponse.

Cette présentation met en avant l'importance du sujet-percevant, tributaire de quatre facteurs principaux:

- 1 - La tradition, c'est-à-dire son éducation culturelle.
- 2 - La réflexion personnelle.
- 3 - L'accoutumance à un mode d'expression plastique.
- 4 - L'exercice, entendu comme expérience acquise.

Aujourd'hui encore, la psychologie de l'esthétique, telle qu'elle a été définie, par exemple, par R. FRANCES, considère l'art au travers du filtre d'une personnalité, du prisme d'une conduite: Son objet est d' " étudier ces conduites de manière systématique en éprouvant sur
" elles l'effet de variables qui paraissent influentes, qu'elles
" appartiennent au stimulus ou au sujet " (2).

Un point fondamental est à énoncer d'emblée afin qu'il

(1) FECHNER (G.T.), Vorschule der Aesthetik, Leipzig, Breitkopf & Härtel, 1876.

(2) FRANCES (R.), Psychologie de l'esthétique, Paris, P.U.F., 1968, page 5.

ne soit pas la cause de confusions hâtives: L'intérêt perceptif a, très tôt, été considéré comme devant être séparé du plaisir esthétique.

Cette remarque est essentielle et travaille au plus haut point la question du jugement des productions plastiques scolaires. Le professeur, en effet, doit être en mesure de faire la différence entre des observations qui se rapportent aux critères qui étaient en jeu au travers de la tâche assignée (dont il juge le résultat) et l'aspect "séduisant" d'un travail (dont il y aura à préciser les composantes). Il y a tout lieu de penser qu'une partie des errances du jugement provient de phénomènes d'interférence d'un champ sur l'autre.

Un des maîtres de la psychologie, BERLYNE, distinguait très précisément (1) l'exploration cognitive (épistémique, analytique) de l'exploration récréative, liée au plaisir de la découverte ou de la reconnaissance. L'évaluation d'un travail scolaire ne ressortirait-il pas prioritairement au domaine de l'exploration analytique? Mais, d'un autre côté, que pourrait-on espérer d'un travail "juste" s'il n'était aussi une oeuvre se rattachant au sensible, tant parce qu'il est trace de l'investissement d'un auteur que parce qu'il est apte à interpeller le récepteur? D'ailleurs, BERLYNE présente ces deux activités comme successives, la recherche de l'agréable venant après l'activité de déchiffrement. (Toutefois, la théorie d'un impact de l'image antérieur au nommé amènerait aujourd'hui à nuancer cette analyse).

Dans un autre langage qui se réfère à la théorie de l'information, A. MOLES différencie le message sémantique, logique, structuré, énonçable, préparant des actions, du message esthétique, intraduisible, transposable qu'approximativement et préparant des états (2).

L'impact sur le sensible se ferait par le moyen du contenu figuré ainsi que par l'action des variables essentiellement plastiques que sont lignes, formes, composition, couleurs. De très nombreuses expériences ont eu pour objet de mesurer l'impact de ces différents facteurs, soit par l'étude du mouvement des yeux, soit par la mesure de la rapidité de choix entre deux "patterns", la mesure du temps d'exploration ou de la modification de l'état émotionnel au niveau

(1) BERLYNE (D.E.), "Les mesures de la préférence esthétique", Sciences de l'art, 1966, 3, pages 9 à 22.

(2) MOLES (A.), Théorie de l'information, op. cit., pages 196-197.

de la peau (réponse électrodermale). Nous avons donc affaire, dans la perception visuelle, à une activité qui serait à relier à une psychologie du comportement.

A. MOLES en énonce les fondements que l'on sentait sous-jacents dans la pensée de FECHNER:

- " L'individu est un système ouvert dont le comportement est déterminé
- " par la somme:
- " - d'un bagage héréditaire;
- " - des événements de son histoire particulière;
- " - de son environnement présent;
- " - d'une certaine quantité, limitée, d'aléatoire " (1).

Par ailleurs, sur le plan physique, l'activité sensorielle est dépendante de trois seuils: Un seuil de sensibilité minimale, un seuil de saturation maximale, un seuil différentiel permettant l'observation d'une variation du stimulus (2).

Toutefois, la compréhension du mécanisme de l'exploration visuelle est encore très partielle. Ce qui est certain, c'est la multiplicité des niveaux exploratoires et le fait que l'oeil infléchit spontanément son attention vers un autre niveau lorsque son investigation est jugée trop mince (3). Une impression défavorable pourrait découler d'un effort trop grand de l'observateur contraint de regarder une image lorsque celui-ci se trouverait réduit à chercher sur quel terrain il pourrait ressentir une certaine satiété visuelle.

Si l'on s'en réfère aux recherches récentes, F. MOLNAR, dans un article déjà cité (4), indique que l'on a repéré (CORCORAN, READ) l'existence de " toute une série de cellules répondant sélectivement à certaines formes d'excitations bien définies ". Cependant, si l'on connaît actuellement l'emplacement anatomique de quarante trois détecteurs spécifiques, on ne connaît pas les traits caractéristiques qui sont transmis au niveau supérieur du cerveau (dans une activité préconsciente) pour être confrontés et combinés à l'information déjà stockée.

C'est pour mieux connaître, à défaut du mécanisme physiologique, le comportement des individus face à une stimulation visuelle que se sont développées les expériences en matière de psychologie et de sociologie du goût.

(1) MOLES (A.), Théorie de l'information, op. cit., page 15.

(2) Ibid., page 26.

(3) Ibid., page 242.

(4) MOLNAR (F.), Quoi de neuf dans la perception visuelle de la forme?, op. cit., pages 599 à 629.

Toutes ces expériences portent sur l'interaction entre les aspects naturels et les aspects culturels et sociologiques. Ainsi que le rappelle R. FRANCES (1964), s'en tenir uniquement à la manipulation de variables du stimulus (exemple: assombrir un tableau)

- " ne permet pas d'aborder l'autre aspect du problème qui tient aux
- " variables du sujet: Que représentent, pour différents sujets, les
- " variations apportées au stimulus?" (1).

Avant d'aborder le détail des informations disponibles à l'heure actuelle, une dernière remarque porte sur l'esprit qui a présidé au recueil de ces informations. Depuis longtemps, comme le souligne Y. BERNARD (1973), deux thèses sont en présence (2):

- La thèse universaliste (EYSENCK, CHILD), qui affirme l'existence d'un facteur général "T" du goût, dépendant directement de la nature du cerveau humain et où, seuls, les facteurs secondaires s'appuyeraient sur des interprétations culturelles.
- La thèse réaliste ou relativiste (FRANCES), pour laquelle les différences de comportement peuvent être ramenées à des différences biologiques et sociologiques.

Quelles sont les principales informations dont nous disposons?

Nous tirons la plupart des données qui vont suivre de l'importante synthèse publiée en 1979 sous la direction de R. FRANCES (3).

1) Les critères du jugement esthétique.

Les motifs de choix les plus déterminants ont été mis à jour par R. FRANCES lors d'une étude portant sur l'appréciation de la peinture, tels qu'ils étaient exprimés dans des réponses orales. Sept motifs ont été dégagés, suivant deux polarités, selon qu'ils inspirent l'attraction du contemplateur ou son rejet:

- Réalisme
- Originalité
- Technique
- Nature du sujet
- Couleur et lumière
- Références purement subjectives
- Expression du sujet

(1) FRANCES (R.), "Variations génétiques et différentielles des critères du jugement pictural", 1^o colloque d'Esthétique expérimentale, Paris, Sciences de l'art, 1966, 3, pages 119 à 135.

(2) BERNARD (Y.), Psycho-sociologie du goût en matière de peinture, op. cit., pages 22 à 25.

(3) FRANCES (R.), Psychologie de l'art et de l'esthétique, op. cit., pages 118 à 132.

Cette étude ayant porté sur 1600 commentaires effectués par des sujets d'âges et d'origines très variés, il est fort probable que les raisons exprimées aient un caractère suffisamment général pour qu'elles puissent être utilisées comme objets d'expérimentation.

2) La question du facteur général de goût.

Selon certaines recherches, notamment celles de BURT (1924, 1940), l'on pourrait croire à l'existence d'un facteur général influençant tous les jugements artistiques. Il y aurait, par ailleurs, un facteur secondaire, bi-polaire, qui opposerait les classiques aux romantiques, les tenants de l'oeuvre dessinée, construite et raisonnée aux tenants de l'expression colorée plus spontanée. Selon les travaux d'EYSENCK, datant de 1940 et 1941, travaux qui confirment le facteur général (T = taste) et les facteurs individuels (K = kind), le facteur général aurait principalement à voir avec la composition de l'oeuvre, tandis que ce qui différencierait les spectateurs aurait trait au degré de réalisme, les sujets peu cultivés préférant les paysages figuratifs alors que les sujets experts seraient plutôt attirés par des genres et des factures moins conventionnels. Deux composantes de l'oeuvre se trouvent, ici, mises en question: La nature du sujet représenté et la technique, souvent confondue avec l'habileté de transcription figurative.

En outre, dans une publication datée de 1978 et intitulée "Rôle de la spécificité des éléments d'une oeuvre picturale dans l'identification de l'auteur", Y. BERNARD (1) note l'action importante de la couleur, notamment quant à sa faculté de caractériser le style d'un auteur. Les recherches sur le coloris sont multiples et attestent qu'il s'agit d'un facteur principal dont il y aura lieu de se préoccuper.

3) Intérêt perceptif et préférence esthétique.

Le mérite revient à BERLYNE (1960) d'avoir dégagé la notion de variables collatives (2). Il s'agit de variables qui dépendent de la collecte et de la comparaison d'informations provenant des objets à apprécier. Son travail a porté principalement sur l'étude de l'originalité et de la complexité.

Dans un ouvrage de synthèse publié un peu plus tard (3),

(1) Ces recherches sont citées par FRANCES (R.), *ibid.*, page 133.

(2) BERLYNE (D.E.), "Les mesures de la préférence esthétique", Paris, Sciences de l'art, 1966, 3, pages 9 à 22.

(3) BERLYNE (D.E.), Aesthetics and Psychobiology, New York, Appleton Century Crafts, 1971.

BERLYNE propose une explication du plaisir esthétique qu'il différencie de l'intérêt. Sa théorie est fondée sur la notion d'éveil ("arousal", autrement dit la capacité d'un objet à éveiller la vigilance d'un spectateur).

L'auteur observe ainsi que l'intérêt porté à l'objet est directement lié à la complexité de cet objet. Quant à l'agrément, il augmenterait, lui aussi, parallèlement au degré d'intéressement, mais seulement jusqu'à un certain seuil au-delà duquel, lorsque la sollicitation devient trop importante, il se mettrait à décroître.

A son tour, FRANCES a repris et approfondi ces expériences. L'intérêt suscité, la somme d'observations justifient, ici aussi, que cette question soit étudiée à nouveau, dans le contexte de l'évaluation scolaire.

La psychologie de l'esthétique a permis de faire émerger les caractéristiques plastiques qui interpellent prioritairement l'attention d'un sujet amené à porter un jugement esthétique. Il y a tout lieu de penser que ces catégories sont également opérantes dans le processus de l'évaluation, dès lors que la vigilance de l'évaluateur se trouve mise à défaut.

S'ajoutant aux effets déjà signalés par la docimologie générale, l'hypothèse de déterminants spécifiques à l'évaluation des objets plastiques pourrait, à la lumière de ce qui précède, se fonder sur l'interférence de réponses esthétiques spontanées portant sur des variables-stimuli, qu'elles appartiennent ou non à la grille de critères que l'évaluateur a construite a priori.

C'est donc bien dans ces données que doit s'originer une stratégie visant à l'élucidation des mécanismes en oeuvre, dès lors qu'on ne regarde plus exclusivement en direction de l'évaluateur mais aussi du côté des productions à évaluer.

3.2. METHODOLOGIE DE LA RECHERCHE.

3.2.1.

PROBLEMES RELATIFS AU RECUEIL DES DONNEES.

Notre matière d'enseignement s'est encore très peu prêtée à des approches théoriques. Tandis que de vastes étendues sont à défricher, notre angle d'attaque se révèle singulièrement mince.

Nous avons tenté, dans le chapitre précédent, de voir quels étaient les appuis théoriques sur lesquels il serait possible de fonder une docimologie des A.P.. Ces supports sont, à la fois, d'un inestimable secours, car susceptibles de vectoriser nos premières pistes de recherche, mais aussi à reforger partiellement.

Force est de constater, en effet, que les points d'ancrage dont nous disposons dans le cadre de la psychologie scolaire ressortissent tous au champ général (évaluation de devoirs écrits à dominante cognitive) et la première question est celle du transfert de leur application au domaine des arts plastiques.

Il faut convenir également que la plupart des observations dégagées par la psychologie de l'esthétique l'ont été avec une orientation sociologique, s'attachant à dresser des typologies comportementales du goût. Fonder quelques hypothèses sur certains de ces résultats est nécessairement faire le pari que ces données peuvent effectivement revêtir un caractère opératoire dans le processus de l'évaluation.

Il en est de même, enfin, pour ce qui concerne la sémiologie et la sémiotique des faits visuels: Leur apport permet de mieux connaître le mécanisme de la lecture des images, la nature des figures sémiotiques qui produisent le sens et conditionnent le parcours du regard. Cependant, l'évaluation d'une oeuvre scolaire ne présente probablement qu'une apparente analogie avec la contemplation des oeuvres artistiques.

Malgré l'importance de ce corpus théorique, de telles données ne sont donc pas suffisantes pour nourrir une réflexion sur l'évaluation en A.P.. Une étude du comportement réel de l'évaluateur est nécessaire. Malheureusement, l'entreprise n'est pas sans difficulté:

- Il ne peut être question de venir observer dans les classes comment procède le professeur, la présence d'un spectateur - fut-il toléré, voire invité - modifiant complètement les comportements des parties en présence. Il est, par ailleurs, totalement inimaginable de recourir à quelque caméra invisible!

- Par ailleurs, il serait illusoire de fonder trop d'espoir sur des entretiens oraux avec les P.A.P.: La question de l'évaluation atteint les replis les plus intimes de l'enseignant, provoque des blocages plus ou moins prononcés, suscite le plus souvent un réflexe défensif. Néanmoins, nous avons eu recours à un questionnaire largement ouvert (mais uniquement par la voie de l'écrit anonyme), pensant qu'il pourrait, malgré tout, apporter quelques indications supplémentaires, mais en nous gardant bien d'y accorder une valeur statistique. (Nous avons déjà fait état, dans la phase introductrice, de quelques observations qui en étaient issues). Pour compléter cette enquête au sujet de la manière dont l'évaluation est pratiquée et ressentie, nous avons aussi questionné les élèves.

Toutefois, seule la méthode expérimentale permet de dépasser le stade des constats épars. Pour recueillir les éléments dont nous avons besoin, nous avons donc procédé indirectement par le truchement d'évaluations en situation simulée.

Le recours à cette pratique n'est pas sans soulever un certain nombre de questions d'ordre méthodologique:

- En premier lieu, on ne peut manquer de s'interroger sur la légitimité de ces expériences, dites "de laboratoire", par rapport à la réalité de la classe sur le "terrain".

Peut-on admettre un morphisme entre la situation artificielle de professeurs réunis à l'occasion d'un stage durant lequel ils évaluent un lot de travaux sélectionnés par l'expérimentateur et l'activité qu'ils pratiquent réellement dans leur établissement dans le cadre d'un cursus éducatif qu'ils ont bâti eux-mêmes pour

des élèves dont ils connaissent le comportement et les difficultés?
Ce premier postulat est, à lui seul, très osé.

Les disparités de jugement qui pourraient être observées dans la solitude d'une expérimentation de laboratoire ne risquent-elles pas de revêtir un caractère caricatural par rapport à la situation réelle où le jugement de l'évaluateur se trouve tempéré par la réaction des élèves présents ou bien des collègues de jury?

- La seconde question se place au niveau des variables à étudier:

L'influence du milieu expérimental ne peut-il pas induire des comportements non spontanés?

Une première cause peut venir tout simplement d'une démobilisation progressive des sujets se prêtant à l'expérience, du fait de la fatigue ou de l'absence d'élèves.

Les attentes de l'expérimentateur, les premiers résultats recueillis que l'on souhaite voir se confirmer, un mot "lâché" incidemment par l'animateur, ne risquent-ils pas d'orienter imperceptiblement (et involontairement) la réaction des sujets-évaluateurs?

Le fait même de choisir certaines tâches comme support expérimental de l'évaluation, plutôt que d'autres, parce qu'elles paraissent mieux aptes à isoler la variable que l'on souhaite étudier, n'en arriverait-il pas à colorer l'expérience d'artificialité voire à provoquer des effets cumulatifs amenant à une hyper-sensibilisation à la variable en question?

Ensuite, peut-on seulement parler de variable "indépendante"? Peut-on isoler efficacement une variable lorsque celle-ci ressortit directement au domaine plastique, quand on sait qu'un objet plastique est avant tout perçu en sa globalité, que tout élément, fut-il isolable dans une suite diachronique par le processus du parcours visuel, n'existe véritablement dans l'oeuvre qu'amalgamé dans le dispositif de fusion qui en assure la cohérence et la multiplicité des lectures?

On ne peut rejeter totalement l'idée que la situation épurée du laboratoire puisse susciter des effets nouveaux qui en viendraient à se superposer à ceux ayant leur origine dans la variable dont on recherche l'influence.

- Un autre point est celui de l'éventail qualitatif des

travaux choisis. Peut-on parler d'un échantillonnage "normal" en se référant à la distribution de la courbe de LAPLACE et GAUSS concernant la répartition aléatoire des séries à grande numérosité? On sait que, dans le contexte éducatif, la référence à cette loi se trouve de plus en plus contestée comme étant une survivance d'idées supposant une distribution inégale des aptitudes. Elle reste pourtant un instrument utile pour doser l'homogénéité des lots de devoirs soumis à l'expérience. Nous en avons donc fait usage de manière empirique.

- Il reste, enfin, la question de la participation des professeurs. Dans une telle entreprise, on ne pouvait que faire appel à des enseignants volontaires (que nous remercions). Est-il utile de préciser que ceux-ci ont été rares?

C'est confrontés à l'ensemble de ces problèmes méthodologiques que nous avons tenté de mettre en place, de façon aussi rigoureuse qu'il nous semblait possible, les expériences d'évaluation dont nous allons maintenant esquisser la structure protocolaire.

3.2.2.

LA NOTATION EN LABORATOIRE :
PROTOCOLES EXPERIMENTAUX.

Des expériences de notation ont pu se pratiquer durant plusieurs journées de stage entre 1983 et 1985. A cet effet, il avait été proposé à trois Inspecteurs Pédagogiques Régionaux responsables de cinq académies (Lyon, Grenoble, Montpellier, Aix-Marseille, Nice) ainsi qu'aux Chefs de Mission à la Formation des Personnels une journée de stage durant laquelle, en contrepartie d'une information sur les problèmes de l'évaluation, les professeurs devraient être volontaires pour participer à une recherche et noter, suivant plusieurs protocoles expérimentaux, des séries de dessins.

Complétaient le dispositif plusieurs animations de stages de titularisation destinés aux Adjoints d'Enseignement ainsi qu'aux jeunes professeurs certifiés en année de formation pédagogique (C.P.R.).

Les sujets concernés étaient donc soit volontaires pour participer à cette recherche, soit fortement motivés dans le contexte d'une formation/titularisation.

Les travaux utilisés furent de deux ordres:

- Des travaux provenant de classes de 5° de collège, niveau intermédiaire du premier cycle de l'enseignement secondaire obligatoire et pouvant, sous cet angle, apparaître comme représentatifs de la production-type que l'on rencontre actuellement en nos classes d'A.P..

Les dessins étaient tous des travaux réalisés dans le contexte d'une pédagogie traditionnelle, c'est-à-dire celle où le professeur propose un thème inducteur de créativité à partir duquel doit être résolu un problème plastique: Objectifs et critères étaient donc particulièrement définissables. Partant volontairement d'un tel choix, il y avait lieu de penser que, si des effets pouvaient

être dégagés, ceux-ci ne pourraient que se faire sentir davantage dans le cas de travaux plus complexes.

- Suite à la vérification préliminaire des divergences, il a semblé nécessaire de fonder une partie de notre expérimentation sur des travaux de baccalauréat, au niveau le plus élevé, c'est-à-dire les sections artistiques A7 (devenues A3 depuis 1982), ceux-ci occasionnant probablement à l'évaluateur de plus amples difficultés, puisqu'il s'agit d'apprécier des "expressions libres sur thème". Trop complexes pour permettre l'isolement de facteurs de variation du jugement, ces travaux tirés des archives paraissaient, au contraire, particulièrement recommandés pour servir de support à une expérience portant sur la notation critériée.

Par ailleurs, afin de tester l'influence des variables-stimuli (variables de couleur, de complexité affectant les objets à évaluer), nous avons eu recours à des "dessins construits".

Certes, l'usage de "copies construites" n'est pas sans soulever un grave problème méthodologique. Il est facile, dans le cas de devoirs écrits, de multiplier les reproductions dactylographiées d'un même travail (variation du style, de l'orthographe, d'un type de démarche mathématique...). Il est, par contre, extrêmement difficile, voire impossible, de reproduire le dessin d'un élève, la photographie, même la plus fidèle, gommant le contact physique direct avec la matière picturale. Il ne reste donc que la procédure manuelle et l'on imagine facilement que peu de travaux pouvaient s'y prêter.

La simplicité de certains dessins de 5° a permis de procéder à deux tentatives en ce sens: On jugera de leur recevabilité.

Nous avons essayé de construire des plans d'expériences approchant le plus possible des conditions réelles de notation des dessins. Les lots étaient constitués d'une grosse partie de dessins "moyens" pour deux moindres parties de dessins plutôt "bons" et plutôt "faibles".

De manière à construire une situation réaliste et identique d'un stage à l'autre, les matériaux avaient été fixés sur de grandes planches de carton suivant une disposition bien définie visant à répartir de façon homogène les différences de niveau entre les dessins (les travaux sont souvent étalés côte à côte pour faciliter les comparaisons).

Afin de mettre à jour des facteurs systématiques déterminant l'évaluation, nous avons procédé à des expériences croisées où une modification était apportée à l'occasion du passage d'un groupe d'évaluateurs à l'autre.

Les protocoles expérimentaux comportaient donc généralement trois facteurs:

- Deux facteurs aléatoires:

- Les évaluateurs, partagés en deux groupes;
- deux lots de dessins.

- Un facteur systématique à deux modalités dont il y aurait à mesurer l'influence au travers de la variation des notes attribuées par les professeurs selon chacune des deux occurrences, autrement dit, en s'interrogeant sur la signification des différences relevées entre deux séries de moyennes.

3.2.3.

LA QUESTION DU TRAITEMENT
STATISTIQUE DES RESULTATS.

Après avoir soupesé quelques questions d'ordre logistique, on ne peut, en dernier lieu, passer sous silence les interrogations soulevées par l'interprétation des résultats.

Il ne s'agit pas tellement du dépouillement qualitatif des fiches-questionnaires dont nous avons volontairement limité la portée au rang de témoignages comportementaux destinés à broser un portrait plus vivant de la pratique de l'évaluation, mais cela concerne plutôt le résultat des expériences d'évaluation simulée.

L'hypothèse générale inspirant l'axiomatique de la partie expérimentale repose sur l'existence de disparités de jugement entre professeurs. Seulement, il convient, dès l'abord, de faire la distinction entre des causes d'ordre "mécanique" (sévérité, usage de la largeur de l'échelle de note) et celles dépendant plus directement de la psychologie de l'évaluant ou encore des caractéristiques propres aux objets à évaluer.

Nous n'avons pas cherché à travailler cette première série d'effets bien connue des chercheurs en docimologie car il ne nous est pas apparu de raison pouvant laisser supposer que ces comportements puissent être remis en cause par la spécificité plastique des travaux utilisés. C'est pourquoi, tout au contraire, nous avons tenu à en effacer les conséquences:

La procédure utilisée consiste à remplacer la note attribuée par un chiffre qui tient compte, à la fois, des différences de moyenne et d'écart-type entre les notations d'un même lot de travaux, unifiant

ces dernières par l'application d'un coefficient correcteur.

En l'occurrence, pour obtenir des notes standardisées, nous avons appliqué à chaque note recueillie la formule (1) :

$$N = \frac{x - m}{s} \times 10 + 50$$

où x = note réelle

m = moyenne des notes attribuées par le sujet

s = écart-type de l'éventail utilisé

N = note standardisée

Le choix du coefficient "...x 10) + 50" n'ayant pour but que de permettre le travail sur des entiers positifs.

Un préalable à l'examen des données chiffrées a donc été la réduction des notes. Cette pondération étant opérée, les variations qui continueraient à se manifester trouveraient leur origine ailleurs ailleurs que dans un usage non identique de l'échelle des notes.

Les autres variations enregistrées peuvent être fortuites, c'est-à-dire dépendre de causes variées, aléatoires, nées de circonstances superficielles ou changeantes. Elles peuvent, à l'inverse, dépendre de variables systématiques, c'est-à-dire de causes précises et dont on cherche à affirmer, par un raisonnement d'ordre scientifique, que les effets se reproduiront, avec un fort pourcentage de chances, dans une situation identique.

Aller plus avant amène inévitablement à rencontrer la statistique, cependant, il nous est apparu légitime d'exposer préalablement quelques scrupules:

Notre manque de familiarité avec ses procédures a contribué à la faire apparaître, à l'origine, comme une machine lourde, démesurée, intellectuellement trop onéreuse eu égard à la modestie des données que nous avons pu collationner. Mais, plus profondément, la pertinence de son utilisation nous semblait devoir être posée compte tenu du caractère fatalement artisanal de nos expérimentations.

Il ne devait pas être oublié, non plus, que les notes attribuées en situation simulée n'étaient que le résultat d'un dispositif artificiel et précaire.

Certes, le recours aux méthodes statistiques est nécessaire et il a donc bien été entrepris de tester la signification des écarts

(1) NOIZET (G.), CAVERNI (J.P.), Psychologie de l'évaluation scolaire, op. cit., page 204.

comparativement aux seuils établis par les statisticiens STUDENT et FISCHER (1).

Pourtant, exploiter sans retenue la positivité d'un test serait adopter un comportement que la docimologie elle-même s'est évertuée à fustiger. Ce serait appliquer à l'EVALUATION des données d'une recherche expérimentale une conduite maintes fois dénoncée, s'agissant de l'EVALUATION des productions scolaires: Le fait d'appliquer la rigueur et la crédibilité du calcul à des notes scolaires établies dans l'approximation, ce serait chercher, par la pureté d'un traitement scientifique valorisant, la caution nécessaire à des données dont on voudrait oublier l'origine incertaine...

Si la vérification statistique est un indicateur précieux, l'essentiel de la recherche présentée ici n'est donc pas composé desdits résultats statistiques par eux-mêmes. Les données ne sont utilisées qu'en tant qu'elles paraissent aptes, par la globalité de leurs tendances, à donner une image positive ou négative des procédures d'évaluation en oeuvre dans nos classes d'A.P..

Il reste la question du recours à des notes chiffrées. Celui-ci est légitime puisque telle est actuellement la disposition la plus généralement employée. Mais il ne faut pas perdre de vue que l'évaluation du P.A.P. passe souvent par le classement des travaux. C'est ce classement qui est, après coup, transformé en notation chiffrée. Nous y avons vu une invitation à examiner dans un second temps ce qu'il pouvait en être des résultats recueillis si l'on raisonnait en termes de rangs.

(1) REUHLIN (M.), Précis de statistique, Paris, P.U.F., 1976, pages 204 à 219.

EXPERIENCES (1):
RECONSIDERATION DES EFFETS
OBSERVES DANS LE CONTEXTE GENERAL.

4

4.1. DIVERGENCES ENTRE CORRECTEURS.

- 411 - Hypothèse de travail.
- 412 - Organisation de la collecte des données.
- 413 - Résultats.

4.2. INFLUENCE DES INFORMATIONS A PRIORI RELATIVES AUX PRODUCTEURS.

- 421 - Hypothèse de travail.
- 422 - Déroulement de l'expérience.
- 423 - Résultats.

4.3. EFFETS DE CONTRASTE.

- 431 - Hypothèse de travail.
- 432 - Déroulement de l'expérience.
- 433 - Résultats.

4.1. DIVERGENCES ENTRE CORRECTEURS.

4.1.1.

HYPOTHESE DE TRAVAIL.

Le travail entrepris ici repose en partie sur l'hypothèse de l'existence de divergences importantes entre correcteurs d'épreuves plastiques.

Nous avons vu précédemment combien ces divergences d'évaluation apparaissent un fait acquis et amplement généralisé. Avant d'aller plus au coeur de notre sujet, il nous faut vérifier qu'il y a bien aussi, en notre domaine, de larges désaccords entre correcteurs.

Cette confirmation est d'autant plus nécessaire que nous sommes en possession, au travers d'une recherche assez récente de M.H. CUENOT (1), de deux séries d'expérimentations tendant à prouver, tout au contraire, qu'il y aurait, dans la notation des travaux plastiques, une homogénéité de jugement fort acceptable.

Ces résultats sont d'autant plus inattendus qu'ils contredisent le vécu de tout jury d'examen artistique et qu'ils vont à l'encontre des observations de la docimologie générale qui a fourni depuis longtemps des chiffres sans appel à cet égard.

Ils sont également en contradiction avec le reste des approches dont il est rendu compte dans la même étude, dans la mesure où il est aussi mis en avant un certain nombre de causes possibles de la variation du jugement de P.A.P., au travers de corrélations observées entre la note et la sympathie accordée à l'élève (2) ou

(1) CUENOT (M.H.), Le P.A.P. et les problèmes de l'évaluation, op. cit..

(2) Ibid., pages 77 à 81: Plus un élève est apprécié, plus il a de chances d'avoir une bonne note en A.P..

du rôle de l'identité des sexes entre maître et élève (1), par exemple.

Toutefois, il est certain qu'on ne peut totalement écarter l'élément facilitateur que représente, en A.P., la faculté d'évaluer simultanément les travaux par juxtaposition: Le va-et-vient visuel permet une comparaison instantanée plus efficace que lors de la correction de copies où la même opération ne peut s'effectuer que par le canal de la mémoire.

Dans le constat des divergences que nous exposons en introduction, nous avançons l'idée qu'un certain accord pourrait provenir du fait qu'il s'agirait de travaux "libres" évalués selon des critères "libres", c'est-à-dire qui se réfèreraient finalement à des facteurs généraux du goût. C'est aussi l'opinion de M.H. CUENOT:

" L'évaluation des productions scolaires en A.P. se fait selon des
 " critères qui semblent communs. L'appréciation d'un dessin répond
 " donc à des normes esthétiques générales, quasiment identiques d'un
 " évaluateur à l'autre, quelque soit son âge, son origine et sa formation
 " en A.P. " (2).

Cependant, cette étude s'appuie sur l'observation des moyennes et des écarts-types relevés dans les notations de 4 correcteurs et des 14 professeurs de 14 classes différentes (soit 309 dessins). Si les moyennes par classe varient peu autour de 12/20 (0,4 points), si les écarts-types attestent d'une utilisation réduite de 5 degrés dans l'échelle, si les coefficients de corrélation varient, à quelques exceptions près, de 0.60 à 0.90 (3), le compte rendu ne précise pas si le calcul a été mené dessin par dessin ou simplement sur des moyennes.

Une deuxième série d'expériences (dont nous reprendrons le détail plus tard), portant sur la différence entre notation globale et notation critériée (4), fait apparaître également de forts coefficients de corrélation. Mais, ici encore, le support de ces expériences, (deux séries, respectivement sur le thème "plantes" et sur le thème "visages") ne livre aucune intention pédagogique spécifique... et l'on reste perplexe quant au choix des critères proposés:

- Qualité de l'expression
- Imagination
- Maîtrise du geste
- Harmonie des formes, qui, à l'exception du cri-

(1) Ibid., pages 82-83: Les P.A.P. ont une préférence pour les travaux provenant d'élèves du même sexe; les femmes sont, en général, moins sévères que les hommes.

(2) Ibid., page 97.

(3) Ibid., page 94.

(4) Ibid., pages 98 à 110.

tère n° 2 (ce que confirme le chercheur) sont loin d'être indépendants: On "note" donc une "appréciation" d'ensemble, largement tributaire des conventions du goût commun.

Notre hypothèse, quant à nous, est qu'il y a effectivement divergence de notation entre correcteurs, et ceci même dans des exercices scolaires (de maîtrise), tels qu'ils sont proposés dans les petites classes (5°, par exemple). Les écarts enregistrés sont probablement aggravés dans le cas d'épreuves plus complexes comme celles que l'on rencontre au niveau du baccalauréat.

C'est ce que nous avons cherché à montrer.

4.1.2.

ORGANISATION
DE LA COLLECTE DES DONNEES.

Nous n'avons pas, sur ce point préalable, construit d'expérience spécifique. Le premier stage d'évaluation organisé à Lyon en octobre 1984 était à tous égards l'occasion d'un premier test: Si les premiers résultats recueillis n'étaient pas encourageants par rapport aux diverses hypothèses que nous envisagions, l'ensemble de la formulation aurait été à reprendre.

Partant de ce constat, une mise en avant de divergences de notation pouvait être opérée par simple accumulation de notes sur quelques travaux.

1) Sur le terrain des exercices scolaires de 5°, le protocole d'une expérience nous a semblé pouvoir nous fournir des données assez bien encadrées:

Chacun sait que ce qui fait le plus problème au niveau de l'appréciation concerne principalement les travaux moyens, porteurs à la fois d'un certain nombre d'indices positifs et négatifs. Les travaux extrêmes, tout au contraire, facilitent un jugement plus tranché.

Précisément, l'expérience n° 2, traitant des effets de contraste, était fondée sur l'introduction d'ancres très polarisées, bien repérées par les professeurs: Ces derniers travaux paraissaient tout indiqués pour qu'y soit interrogée la question de l'homogénéité des jugements.

2) Pour ce qui concerne le niveau du baccalauréat, est-il nécessaire d'insister sur l'existence bien réelle de divergences?

Nous avons enquêté auprès du corps d'inspection afin de recueillir des informations administratives destinées à étayer nos suppositions.

4.1.3.

RESULTATS RECUEILLIS.

1) Exercice de maîtrise, niveau 5°. (1)

L'exercice de maîtrise que nous exploitons, ayant servi de support à l'étude des effets de contraste, consistait en un essai technique pictural (différenciation, par la couleur et le geste, du ciel et de l'eau de la mer). Les planches de 12 peintures similaires comportaient deux ancrés de qualité extrêmes. Celles-ci ont été globalement bien perçues par les professeurs puisque les moyennes varient en chiffres standardisés (et pour 76 notes) de : $M^+ = 63,17$
à : $M^- = 40,15$ tandis que
la moyenne générale (pour 360 notes) est de : $M = 51,12$.

Globalement, l'accord était réalisé quant à la perception du caractère positif ou négatif des ancrés, mais qu'en était-il dans le détail? (Et, répétons-le, après réduction a posteriori des écarts par la réduction des notes).

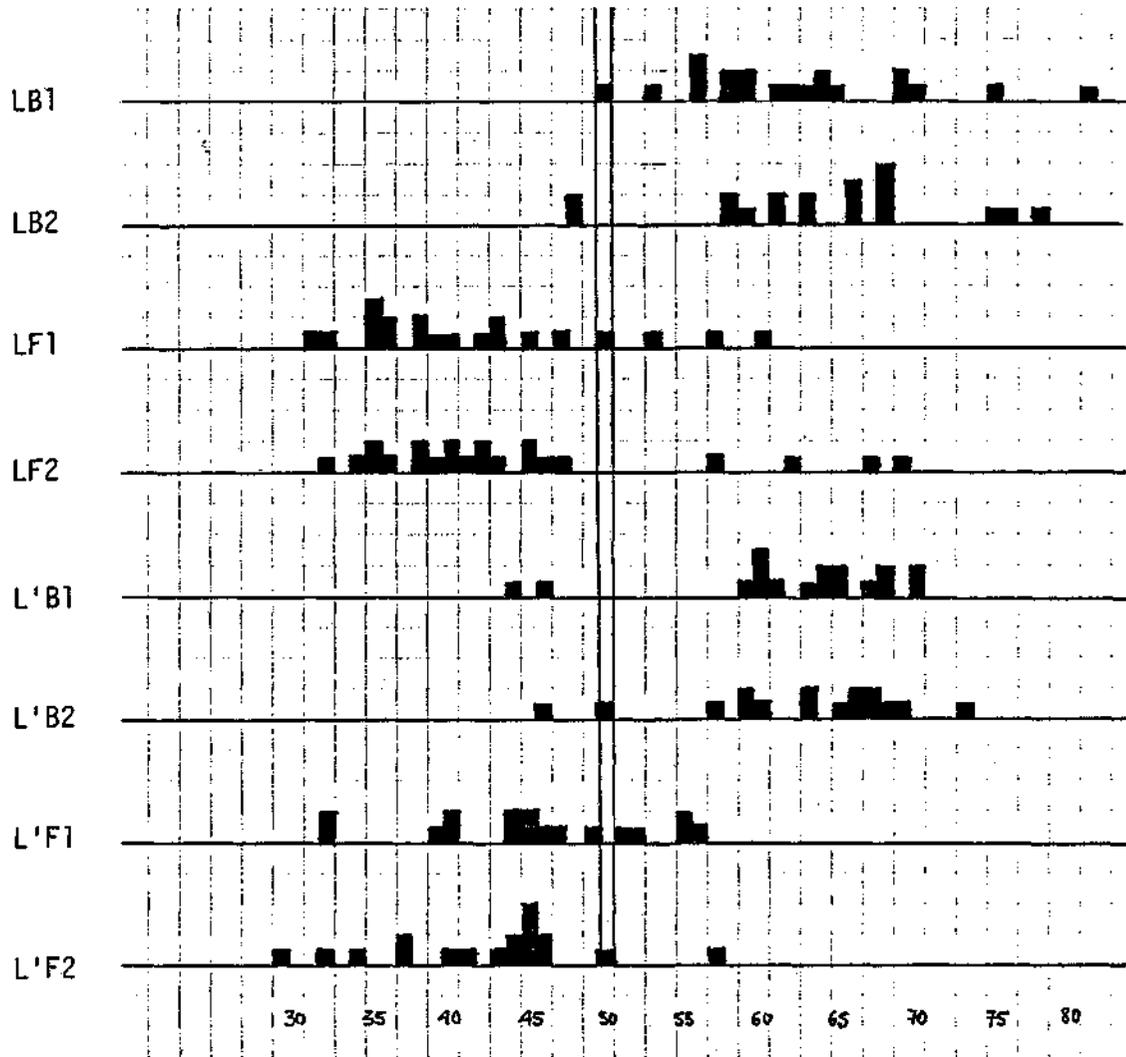
Nous présentons dans le tableau de la page suivante la répartition des notes correspondant à chacune des situations créées, à savoir deux ancrés fortes (B1 et B2) et deux ancrés faibles (F1 et F2) placées successivement dans deux lots (L et L') croisés avec deux groupes d'évaluateurs.

Cette expérience mettait à l'épreuve 38 professeurs répartis en deux groupes, ce qui permettait de recueillir 19 notes par travail.

Dans cette expérience, nous nous en tiendrons à l'aspect visuel de l'histogramme correspondant à la dispersion des notes:

(1) Cet exercice a également servi de support à l'expérience sur les effets de contraste. Voir § 4.3.

HISTOGRAMME DES NOTES RECUEILLIES PAR LES ANCRES, SUIVANT 8 OCCURENCES.



L'étalement en largeur de la notation apparaît immédiatement, même si, du fait probablement que les notes ont été préalablement réduites, peuvent s'observer çà et là quelques regroupements homogènes.

Si la situation sur l'horizontale laisse percevoir une inclinaison d'un côté ou de l'autre de la moyenne (50) qui correspond effectivement à la polarisation positive ou négative, on ne manquera pas de remarquer que, malgré la standardisation, d'importants chevauchements se produisent, certains correcteurs notant plus haut les ancrés négatives que d'autres les ancrés positives (Ceci est particulièrement net pour le couple LF2 / L'B1).

Cette simple approche laisse déjà apercevoir que l'éventail des notes attribuées à un même travail reste important après la standardisation de la notation, ce qui atteste de l'existence et de l'action d'autres déterminants.

2) Epreuves de baccalauréat.

Nous livrerons ici le résultat d'une observation conservée par L. FULCONIS, Inspecteur Pédagogique Régional en arts plastiques, à la suite d'un programme de sensibilisation à l'évaluation en direction des P.A.P. et datant de 1979 (1).

Les chiffres communiqués sont d'autant plus intéressants qu'ils sont exprimés en moyennes de jurys, ce qui signifie qu'ils ont déjà fait l'objet d'une pondération homogénéisant la notation de 7 à 12 professeurs différents.

Copie n°	NICE			CORSE		Ecart max. jurys	moyenne stages 48pers	Note correct. BAC	Ecart
	jury 1 11pers	jury 2 11pers	jury 3 12pers	jury 1 7 pers	jury 2 7 pers				
1	4	10	16	12	10	12	10,5	13	2,5
2	8	12	16	12	10	8	11,5	16	4,5
3	16	16	18	20	18	4	18	17	1
4	12	16	15	12	12	4	13,5	18	4,5
5	12	12	12	16	12	4	13	17	4
6	12	18	13	20	16	8	16,5	15	1,5
7	16	18	14	20	18	6	17,5	19	1,5
8	4	10	10	8	8	6	8	12	4
9	8	12	17	16	16	8	14	15	1
10	4	16	11	12	10	12	11	11	0

Ecart maximum moyen : 7,2

Recueillis dans un autre cadre que celui de la présente étude, ces chiffres donnent à penser: Peut-on imaginer que deux jurys composés de plus de dix personnes chacun peuvent diverger de 12 points sur 20 à propos du même travail ? (Copie n° 1). Certes, l'amplitude des désaccords est généralement moins forte, on peut observer, d'ailleurs, que les écarts varient considérablement d'un travail à l'autre, probablement selon la nature des problèmes posés et les types de modèles qu'ils monopolisent. Enfin, s'y observe la confirmation qu'un accord se réalise plus aisément dans le cas de productions extrêmes

(1) FULCONIS (L.), Evaluations sur 10 copies de BAC A7, sujets en main, Académies de Nice et Corse, 1979. (Tableau aimablement communiqué). Voir annexe n° 8.

(copie n° 3, excellente, ou copie n° 8, faible).

On peut, malgré cela, remarquer que le correcteur du baccalauréat s'en sort plutôt bien, comparativement à la note correspondant à la moyenne des 48 évaluateurs (écart maximum: 4,5 points).

Ceci est probablement dû à son attitude prudente, principalement quant à la responsabilité de sanctionner par une mauvaise note (il "donne" la moyenne à tout le monde) des élèves qu'il ne connaissait pas. Préoccupation qui ne concernait pas les stagiaires-évaluateurs qui descendirent jusqu'à 04/20.

Bien sûr, ces derniers résultats sont accentués par la non standardisation des notes. Mais, rappelons que la situation d'évaluation en vigueur à l'examen n'est pas davantage contrôlée.

Ceci paraît suffisant pour mettre en doute la question de l'homogénéité de la notation des P.A.P. et pour nous inciter à poursuivre plus avant le dévoilement des déterminants de l'évaluation des productions plastiques scolaires.

4.2. INFLUENCE DES INFORMATIONS A PRIORI RELATIVES AUX PRODUCTEURS.

4.2.1.

HYPOTHESE DE TRAVAIL.

Parmi les observations de la docimologie générale les plus largement démontrées, figure l'influence sur l'évaluateur des informations relatives à l'auteur du devoir: Le professeur prend en compte, dans son évaluation, la représentation qu'il se fait de son élève au travers de ce qu'il a appris à connaître peu à peu de celui-ci. Cette prise en compte porte tout aussi bien sur son identité que sur son apparence ou ce qui est perçu de son comportement.

M. GILLY (1980), analysant de nombreux travaux universitaires, notamment ceux d'EYSENCK et PICKUP (1968) (1) faisant ressortir que

" quel que soit le trait scolaire socialement désirable (les enseignants)
" jugent l'enfant en termes de sa conformité en situation scolaire ",

montre que l'enseignant juge, en fait, l'élève et non l'individu au travers de deux questionnements liés directement à la situation scolaire: La notion de persévérance est particulièrement valorisée par l'école ainsi que celle d'intégration sociale, de "participation":

" Dans sa représentation de l'élève, l'enseignant privilégie des valeurs
" cognitives et attitudes "morales" face au travail " (2).

Alors, qu'en est-il du P.A.P.? Est-il de ceux qui sont particulièrement sensibles au comportement studieux des élèves, et, en ce sens, réagit-il à l'identique de la majorité du corps enseignant, ou adopte-t-il plutôt une attitude opposée suivant laquelle les facultés créatrices sont prioritairement attendues aux dépens des descripteurs

(1) GILLY (M.), Maître-élève. Rôles institutionnels et représentations, op. cit., page 62.

(2) Ibid., page 97.

à connotation scolaire?

Il y a donc lieu de s'interroger sur ce qui valorise le plus l'image du "bon élève en A.P.". M. GILLY montre aussi que

- " les éducateurs extra-scolaires apprécient l'enfant différemment en
- " privilégiant, notamment, le facteur imagination tandis qu'inversement
- " ce facteur apparaît secondaire chez les enseignants, même dans le
- " cas où il s'agit de maîtresses de maternelle où pourtant les activités
- " créatrices occupent une place très importante " (1).

Autrement dit: Qui, de l'enseignant ou du créateur, a voix prépondérante chez le P.A.P.?

J.M. FABRE (2) regroupe toutes les informations, parfois connues expressément par les conditions d'enseignement (connaissances confidentielles recueillies sur les élèves et/ou leur famille, en classe, dans le dossier scolaire ou durant les conseils de classe, par exemple), ou encore ressenties beaucoup plus confusément par une série de traits comportementaux, sous la terminologie générale d'"effets-source".

Certes, prendre en compte dans l'évaluation des résultats les composantes individuelles de chaque élève n'est pas une mauvaise chose. Dans une optique formative, on peut même considérer que cette attitude est souhaitable: N'est-il pas légitime, en effet, d'apprécier les efforts d'un élève compte tenu des moyens qui lui sont propres, de ses conditions de travail, de l'équipement matériel que lui permettent ses ressources financières? Ainsi que le résume G. de LANDSHEERE,

- " ici on encourage en surestimant le travail; là on fait preuve d'une
- " sévérité exceptionnelle pour donner un choc que l'on espère salutaire.
- " Dans ce cas, le maître agit délibérément, en toute conscience " (3).

Mais cela devient regrettable si ces informations, en dehors de toute conscience des correcteurs, induisent des préjugés qui déstabilisent le jugement. Les observations des chercheurs ont fait apparaître depuis bien longtemps des faits non équivoques: Le voile formé autour d'un élève par la réputation qu'il s'est forgée auprès des professeurs peut masquer la réalité tandis que se perdurent les effets des premières impressions. La contagion des résultats va jusqu'à se faire sentir dans des épreuves pour lesquelles l'objectivité semblerait la plus assurée: Ainsi, a-t-on pu observer que,

(1) Ibid., page 120.

(2) FABRE (J.M.), Jugement et certitude, Berne, Peter Lang, 1980, page 40.

(3) LANDSHEERE (G. de), Evaluation continue et examens, op. cit., pages 40-41.

lors de la dictée orthographique, l'oubli de fautes par le correcteur était plus fréquent en faveur des bons élèves. C. CHASE (1968) a même constaté, dans les tests standardisés, l'influence de la qualité de l'écriture sur les scores obtenus par les candidats (1).

Précisons donc les informations dont nous disposons actuellement et qui peuvent nous permettre d'énoncer des hypothèses de travail quant à leur application au terrain des activités plastiques.

Plusieurs expérimentations émanant du Groupe de Recherche sur l'évaluation de l'Université de Provence ont permis d'identifier quelques types d'informations a priori:

1) Information sur le statut scolaire de l'élève.

BONNIOL, CAVERNI et NOIZET (2) ont montré que, en moyenne, les copies étaient mieux notées quand il était dit qu'elles provenaient d'élèves de bonnes classes que lorsqu'elles étaient attribuées à des élèves issus de classes faibles.

2) Information sur les performances antérieures.

CAVERNI, FABRE et NOIZET (3) ont pu mettre en évidence un effet de stéréotypie, sorte d'imitation par contagion des résultats: Les "bons" élèves recueillent des notes en moyenne de deux points supérieures à celles obtenues par les mêmes devoirs attribués à des élèves ayant généralement de faibles performances. On y montre une influence légère du livret scolaire interprétée comme étant la poursuite d'un objectif de consonance cognitive entre l'information sur le niveau de l'élève et les informations extraites de la copie. J.J. BONNIOL précisera en 1981 (4) que si les notes antérieures proviennent du même examinateur, le phénomène d'assimilation s'en trouve légèrement accentué.

3) Information sur le statut socio-économique.

NOIZET et CAVERNI (5) ont montré que, dans certains cas

(1) Ibid., page 42.

(2) BONNIOL (J.J.), CAVERNI (J.P.) et NOIZET (G.), "Le statut scolaire des élèves comme déterminant de l'évaluation des devoirs qu'ils produisent", Cahiers de Psychologie, 1972, 15, pages 93 à 104.

(3) CAVERNI (J.P.), FABRE (J.M.) et NOIZET (G.), "Dépendance des évaluations scolaires par rapport à des évaluations antérieures", Le Travail Humain, 1975, 38, pages 213 à 222.

(4) BONNIOL (J.J.), Déterminants et mécanismes des comportements d'évaluation d'épreuves scolaires, op. cit., pages 208 à 236.

(5) NOIZET (G.) et CAVERNI (J.P.), Influence de l'origine socio-économique des élèves sur l'évaluation de leur production scolaire, Université de Provence, 1975, 11 pages.

au moins, l'origine socio-économique de l'élève est susceptible d'influencer la note dans le même sens que celui de la catégorisation. Remarquons, à propos de ce facteur, que le niveau social n'est généralement pas sans interférence sur la relation de l'élève aux phénomènes culturels.

4) Information concernant l'origine ethnique.

Selon les expériences de AMIGUES, BONNIOL et CAVERNI (1), on observe une tendance à surévaluer les élèves dont le patronyme est, quant à son origine ethnique, différent du patronyme de l'évaluateur.

On voit combien ces indications peuvent stimuler notre curiosité quant à la nature de leur action dans le cas de l'évaluation de productions plastiques scolaires. Il y a peu d'arguments permettant de supposer une absence d'influence sur le jugement. Mais il reste à s'interroger sur le sens des effets occasionnés:

Constatera-t-on un effet d'assimilation à la polarité des représentations ou un effet prenant le contrepied de ce qui est généralement observé?

Si les expériences concluent majoritairement à un phénomène de "contagion", notons que cette tendance n'est pas sans appeler la réticence de certains chercheurs. J.M. FABRE (2) écrit à cet égard:

" Lorsque simultanément une bonne copie est associée à une information défavorable et une mauvaise copie à une information favorable, des comportements opposés pourraient aussi bien se manifester: Gratification de celui qui "s'est surpassé" et pénalisation de celui qui "s'est relâché". On peut noter que de telles conduites (même si elles ne se manifestent pas dans les expériences citées, au moins avec une fréquence suffisamment notable pour se traduire en effets moyens) font, elles aussi, l'objet de verbalisations de la part des enseignants".

Nous nous permettrons de reprendre à notre compte cette interrogation avec comme justificatif la référence à de nombreuses observations touchant au domaine de la "créativité" et nous irons jusqu'à poser:

- Le P.A.P. est-il un professeur comme les autres?

Chacun sait, en effet, qu'un individu créatif se reconnaît à sa capacité de produire une grande quantité d'idées. Dès l'aube

(1) AMIGUES (R.), BONNIOL (J.J.) et CAVERNI (J.P.), "Les comportements d'évaluation dans les systèmes éducatifs. Influence d'une catégorisation ethnique sur la notation de productions scolaires", Journal International de Psychologie, 1975, 10, pages 135 à 145.

(2) FABRE (J.M.), Jugement et certitude, op. cit., page 51.

de l'humanité (ARISTOTE, PLATON) la psychologie montre l'imagination comme dépendante de trois règles: Contiguïté, similitude et contraste.

L'individu créatif se trouverait donc plus apte à prendre le contrepied des idées généralement admises. A.F. OSBORN (1) et d'autres universitaires américains (CROPLEY, BARRON) le décrivent, eux aussi, comme tolérant, non conformiste, capable de prendre des positions peu partagées au risque de se tromper ou de paraître ridicule. Il est, par conséquent, tout à fait possible que le P.A.P. réagisse aux informations a priori d'une manière différente de la majorité du corps enseignant.

Il ne pouvait être question, dans une première étude, de tester une grande quantité de facteurs. Dans un souci d'efficacité, nous avons choisi, parmi les informations relatives au producteur, de porter notre intérêt plus particulièrement sur l'influence de la connaissance des résultats antérieurs de l'élève.

Toutefois, afin de voir ce qu'il en est de l'influence de données se rapportant au contexte directement scolaire, d'une part, et, d'autre part, de ce qui relève directement des capacités liées à l'expression plastique (les deux ayant été montrées par les chercheurs comme généralement séparées dans l'esprit des enseignants), nous avons entrepris de faire se croiser deux types d'information:

- Celle relative aux résultats de l'élève en "matières fondamentales";
- celle relative aux résultats de l'élève en arts plastiques.

L'expérience mise en place avait pour but de tenter de répondre à plusieurs interrogations:

- 1) En premier lieu, observe-t-on, dans l'évaluation des productions plastiques, un effet-source?
- 2) Ensuite, cet effet, à supposer qu'il existe, est-il dans le sens d'une assimilation aux performances antérieures et en cela semblables à ce qui est observé dans le contexte général, ou bien conduit-il à un comportement opposé?
- 3) La nature des performances (scolaires ou artistiques)

(1) OSBORN (A.F.), L'imagination constructive, Paris, Dunod, 1971, notamment "La technique du vice-verga", pages 250 à 252.

conduit-elle à des comportements différents?

Quel est, des deux domaines, celui qui retient davantage l'attention du P.A.P.? Se trouve-t-il prioritairement sensibilisé par les descripteurs qui lui sont spécifiques?

- 4) Enfin, les réactions recueillies sont-elles sujettes à variation selon les caractéristiques des professeurs concernés?

4.2.2.

DEROULEMENT DE L'EXPERIENCE.

Nous avons choisi pour servir de support à cette expérience un exercice ayant l'apparence, à la fois, d'un stimulus de créativité et d'un test d'apprentissage et nécessitant, par ailleurs, diverses qualités d'abstraction et d'intelligence dont il est convenu de créditer ceux qui, dans notre système scolaire traditionnel, sont considérés comme "bons élèves". Nous voulions, par ce choix, rendre vraisemblable la possibilité de distinguer et d'apprécier les deux modalités à partir desquelles était construite l'expérience et qui concernaient les performances en tant que plasticien, d'une part, les performances dans les matières fondamentales, d'autre part.

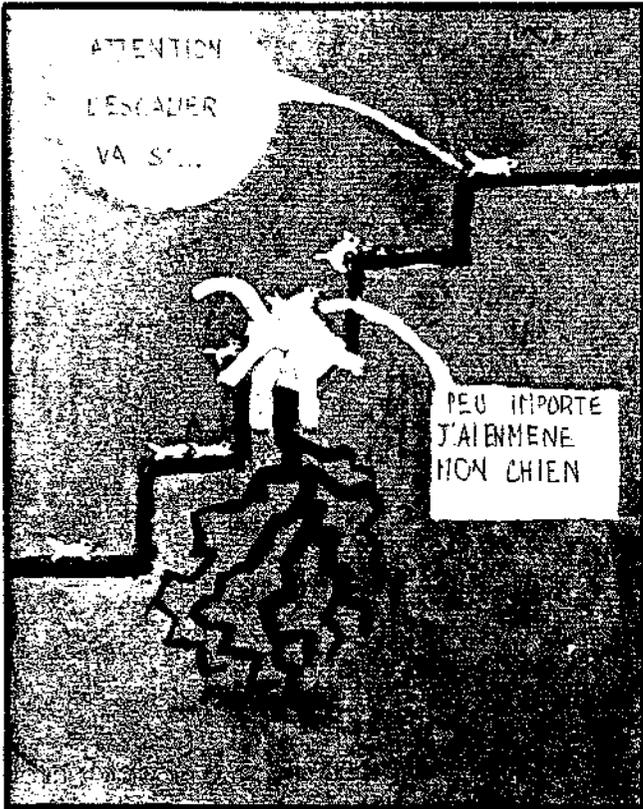
Ces travaux, de format 24 x 32, se lisaient tous verticalement sous l'allure générale d'une vignette unique de bande dessinée se présentant comme un récit achevé. Toutes renfermaient deux bulles situées identiquement dans le format, qui composaient le dialogue incongru suivant:

"- Attention! L'escalier va s'.....

"- Peu importe: J'ai emmené mon chien ".

Il était demandé d'inventer une scène imaginaire (toute situation fantastique, merveilleuse, anthropomorphique était, bien entendu, recevable) dans laquelle ce dialogue deviendrait, dès lors, empreint de logique, trouverait sa justification et sa cohérence. La recherche de l'idée nécessitait de bonnes capacités intellectuelles tandis que la présentation de la scène, en particulier son cadrage, mais aussi le choix des couleurs et l'écriture de l'ensemble se rapportaient à la compétence plastique de l'élève.

L'exercice imposé était de forme traditionnelle, ce qui



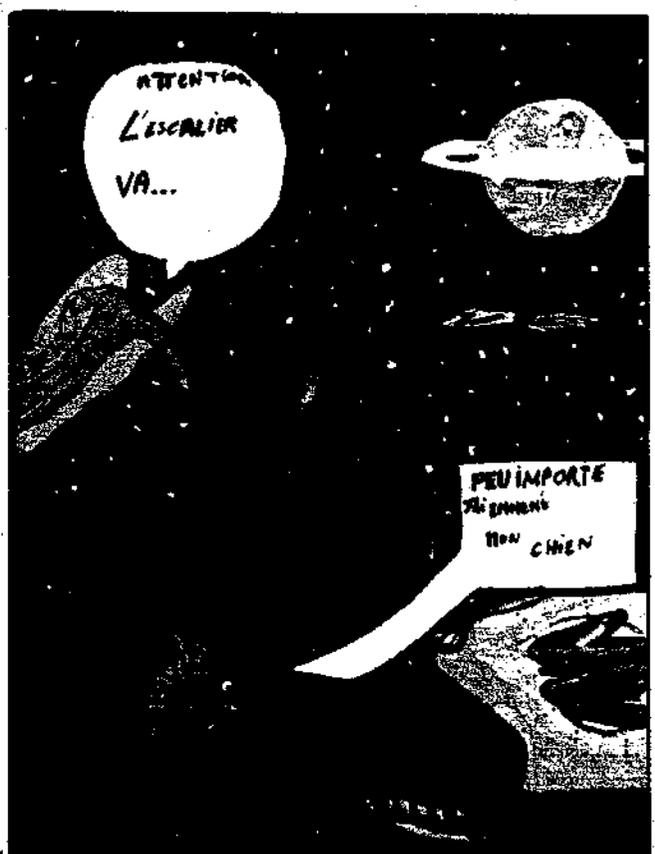
MAT. FONDAM. → *Bien*
 ARTS PLAST. → *Faible*



MAT. FONDAM. → *Bien*
 ARTS PLAST. → *Bien*



MAT. FONDAM. → *Faible*
 ARTS PLAST. → *Bien*



MAT. FONDAM. → *faible*
 ARTS PLAST. → *faible*

devait permettre aux évaluateurs de mieux cerner le produit attendu et les critères d'évaluation. Ces travaux de niveau 5° avaient été sélectionnés, d'après notre propre évaluation, dans un large éventail autour de la moyenne: Seuls avaient été écartés les travaux extrêmes pour lesquels il eut été difficile d'affirmer tour à tour qu'ils étaient l'oeuvre d'un bon élève puis d'un très faible.

Huit travaux (disposés comme pour les autres expériences sur un panneau de carton gris) étaient proposés à l'évaluation des professeurs. Le nombre choisi était peut-être restreint, mais il relevait du souci de ne pas multiplier les difficultés occasionnées par l'adjonction d'une double information supplémentaire portant sur les performances des auteurs. Trop d'éléments à prendre en compte auraient risqué, à nos yeux, d'entraîner des perturbations dans la prise en compte des données a priori et surtout provoquer le découragement des professeurs qui s'étaient portés volontaires pour cette tâche ingrate.

A	B	A'	B'
C	D	C'	D'

L'expérience a consisté à présenter ce même lot de huit travaux (A,B,C,D,A',B',C',D') à quatre groupes de correcteurs (G1, G2, G3, G4) en faisant varier chaque fois les informations concernant les deux niveaux de performance des auteurs (au moyen d'étiquettes adhésives déplaçables). Cette information indiquait par "+" ou "-" le niveau habituel en A.P. et en matières fondamentales, ce qui nous mettait en présence de quatre occurrences.

Le tableau ci-après rend compte du croisement des lots:

Groupes de correcteurs:	Niveau des performances:			
	++	+-	-+	--
G1	AA'	BB'	CC'	DD'
G2	BB'	AA'	BB'	CC'
G3	CC'	DD'	AA'	BB'
G4	DD'	CC'	DD'	AA'

Comme pour les autres expériences, l'éventail des professeurs a été, dans la mesure du possible, choisi au plus large. Il

s'est agi de 34 professeurs de différents statuts sollicités à l'occasion des cinq stages académiques.

Compte tenu de l'aspect fastidieux de l'opération, les professeurs n'ont pas fonctionné dans l'ensemble des quatre groupes ni face à la totalité des modalités d'information. C'est pourquoi l'expérience statistique repose finalement sur la moyenne de 58 notations pour les modalités extrêmes (M^{++} et M^{--}) et sur la moyenne de 78 notations pour les moyennes intermédiaires (M^{+-} et M^{-+}), davantage sujettes à caution.

Le protocole choisi comportait donc:

- Deux facteurs aléatoires, les 34 professeurs répartis en 4 groupes et les huit travaux.
- Le facteur systématique était constitué de l'information portant sur les résultats antérieurs des producteurs, présentée suivant 4 modalités.
- La variable dépendante était constituée des 4 moyennes des notes correspondant aux 4 occurrences de l'information.

4.2.3.

RESULTATS DE L'EXPERIENCE.

Les chiffres obtenus, correspondant à des notes réduites sont les suivants (1):

	MF +	AP +	MF -	AP +	MF +	AP -	MF -	AP -
	49,43		48,32		46,76		47,69	
Portant sur:	58 notes		78 notes		78 notes		58 notes	

1) Phénomène d'assimilation: Son existence.

L'assimilation de la notation aux performances antérieures est nettement affirmée par l'examen des deux résultats extrêmes.

Soit : $M^{++} = 49,43$
 $M^{--} = 47,69$
 $M^{++} - M^{--} = 1,73$

L'expérience se trouve vérifiée par le calcul du $|t|$ de STUDENT:

$$|t| = 3,68, \text{ significatif à } .01, \text{ pour } 57 \text{ degrés de liberté.}$$

Ainsi, contrairement à ce que l'on pouvait espérer de la prédominance de l'activité visuelle lors de l'évaluation des productions plastiques, contrairement à l'espoir que les P.A.P. seraient moins sensibles aux informations sur la réussite antérieure, la différence de performances est prise en compte: Comme nos collègues,

(1) Voir calculs en annexe n° 9.

nous notons plus favorablement les travaux provenant d'élèves réputés "bons" et plus sévèrement les productions des élèves "faibles".

2) Poids relatif des descripteurs.

Le second constat concerne la différence observable entre les deux moyennes intermédiaires. Cette différence signifie que l'inégalité dans les performances n'est pas perçue de manière équivalente selon qu'il s'agit d'être, pour un élève:

- Bon en A.P. et faible en matières fondamentales : $M^{-+} = 48,32$
- Bon en matières fondamentales et faible en A.P. : $M^{+-} = 46,76$

Cette différence est aussi statistiquement significative:

$$M^{-+} - M^{+-} = 1,55$$

(Différence statistiquement confirmée par: $|t| = 4,84$, significatif à .01 , pour 77 D.L.).

Cela signifie en premier lieu que le P.A.P. privilégie dans sa prise en compte ce qui concerne sa discipline. L'information sur le niveau antérieur en A.P. a davantage d'influence que celle concernant le niveau scolaire général, ce qui est plutôt rassurant. Le fait que les A.P. pèsent davantage est confirmé par la situation assymétrique des deux moyennes de part et d'autre de la moyenne générale:

$$\begin{aligned} M_{AP^+} &= 48,32 \\ M_g &= 48,02 \\ M_{AP^-} &= 46,76 \end{aligned}$$

Une troisième observation nous paraît digne d'intérêt. En effet, contrairement à la logique qui voudrait que les moyennes se présentent dans un ordre décroissant où $M^{++} > M^{-+} > M^{+-} > M^{--}$, l'on constate que la moyenne intermédiaire "MF⁺ AP⁻" est inférieure à la quatrième modalité où les performances sont annoncées faibles dans les deux matières:

$$\begin{aligned} M^{+-} &= 46,76 \\ M^{--} &= 47,69 \end{aligned} \qquad \underline{M^{+-} - M^{--} = - 0,93}$$

Ainsi, non seulement les A.P. seraient préoccupation majoritaire, mais ceux-ci seraient l'objet d'une relativisation exacerbée par rapport à l'ensemble des productions scolaires:

Si, d'un côté, de bonnes performances en A.P. n'effaceraient pas totalement l'impression négative produite par la faiblesse scolaire,

si l'élève bon en A.P. et faible ailleurs est légèrement racheté, de l'autre, on ne pardonnerait pas à un "bon" élève d'avoir de mauvais résultats en A.P.: Cette situation serait impitoyablement sanctionnée et considérée plus grave encore que si l'élève était faible partout.

Cette dernière considération chiffrée pourrait être interprétée de surcroît comme l'indice que les A.P. seraient pensés comme fonctionnant sur un registre parallèle à celui des autres matières scolaires: Quand on est bon élève, il n'y a pas de raison pour qu'on ne le soit pas aussi en A.P.. Pensée qui impliquerait l'idée d'une "fondamentalisation" possible de notre discipline dont il y aura lieu de soupeser plus avant les conséquences.

Toutefois, nous resterons au conditionnel sur ce dernier point car la pénalité supplémentaire comptabilisée sous le registre M^{+-} (-0,93) n'est pas tout à fait entérinée par la vérification statistique:

$|t| = 1,47$ pour 58 D.L., alors que pour un taux de $P = 0,10$, la valeur de $|t|$ devrait être supérieure à 1,67.

Néanmoins, la différence relative à la pondération des descripteurs invite à ne pas totalement ignorer cette interprétation puisqu'elle ressort également de l'examen de l'assymétrie des moyennes intermédiaires faisant ressortir que bon en "MF" et faible en "AP" est la situation la plus pénalisante.

3) Conclusion.

La répartition des évaluateurs dans les différents stages n'a pas permis de tester l'influence de la variable "statut-professeur".

Nous nous en tiendrons à des résultats globaux qui attestent une réelle prise en compte par les P.A.P. des informations relatives aux performances antérieures des producteurs, ainsi qu'une attention prioritaire, dans le sens d'une assimilation aux résultats antérieurs, accordée au domaine artistique.

4.3. EFFETS DE CONTRASTE.

4.3.1.

HYPOTHESE DE TRAVAIL.

Tout jugement est affaire de comparaison. Dans la situation traditionnelle où le P.A.P. a proposé à l'ensemble de la classe le même thème de travail, l'évaluation des résultats se fait généralement, comme dans la plupart des matières, comparativement aux autres productions du groupe.

Bien qu'il n'y ait pas de modèle de référence unique auquel confronter les travaux obtenus, il n'est pas exagéré de penser que les professeurs, pour peu que l'exercice proposé ne le soit pas pour la première fois, ont présent à l'esprit l'image plus ou moins précise d'un produit attendu, sinon fixé dans son contenu, du moins cerné par ce qui pourrait s'appeler une certaine "quantité de satisfaction minimale" qu'ils escomptent ressentir à la contemplation des résultats.

Ainsi, le professeur s'apprête-t-il à évaluer son lot de dessins avec une attente qui n'est pas sans référence à la courbe de GAUSS: Il s'attend à ce qu'une majorité d'élèves ait réussi à produire un travail qu'il convient de qualifier de "convenable", il espère rencontrer plusieurs "surprises agréables" (il les aura, d'ailleurs, repérées en classe durant la période d'exécution), il sait aussi que sa proposition a donné lieu à quelques "échecs".

Dès lors, la moyenne de la production va devoir être située à l'intérieur de ces extrêmes et l'on voit que se pose effectivement la question soulevée en introduction méthodologique: Le chiffre

attribué reflète certainement davantage une intention d'ordination; il ne peut réellement prétendre correspondre à tel ou tel degré d'une échelle d'intervalles. Ce n'est probablement qu'extensivement, sans que les notes puissent prétendre à valeur absolue, que ce qui marque plutôt la situation d'un rang est interprété comme mesure. Après tout, n'est-il pas courant de laisser un espace de plusieurs points (surtout, d'ailleurs, aux extrémités de la série) entre deux notes, marquant par là l'intention de signaler la grandeur de l'intervalle ressenti entre deux travaux jugés comme se suivant, mais d'assez loin, par la qualité? Il faudra y revenir.

La docimologie générale, prenant appui sur les observations de la psychophysique (notamment les tâches d'estimation absolue où l'on propose de situer des perceptions dans des catégories ordonnées comme par exemple de classer des objets du plus lourd au moins lourd) a montré que l'activité de comparaison s'effectue à partir d'éléments caractéristiques dénommés "ancres": Il s'agit, en psychologie de la perception, d'un objet facilement repérable. Ce stimulus peut parfois se signaler par sa répétition, mais le plus souvent il se dégage par l'exemplarité de sa situation extrême. Ainsi, dans le comportement d'évaluation, il y a tout lieu de penser que les travaux situés aux extrémités de la série aient une fonction d'ancrage. C'est ce qu'affirment G. NOIZET et J.P. CAVERNI (1978):

- " Une stratégie efficace, dans une tâche d'estimation absolue, consiste
- " à repérer un objet caractéristique, par exemple celui qui se trouve
- " à une des extrémités de la série " (1).

Ceci est en accord avec l'opinion de plusieurs chercheurs américains, VOLKMANN (1951), ERIKSEN et HAKE (1957), que mentionne J.M. FABRE (1980) (2).

C'est donc par rapport à des ancres que le comportement d'évaluation exerce son activité comparative.

Nous savons que de nombreuses recherches (HARVEY et CAMPBELL (1963) (3)) et, en France, celles de J.J. BONNIOL en 1965 (4), reprenant celles, plus anciennes, de H. PIERON, ont établi l'existence d'effets

(1) NOIZET (G.) et CAVERNI (J.P.), Psychologie de l'évaluation scolaire, op. cit., pages 109-110.

(2) FABRE (J.M.), Jugement et certitude, op. cit., page 31.

(3) Ibid., page 32.

(4) BONNIOL (J.J.), "Divergences de notation tenant aux effets d'ordre de correction", Cahiers de Psychologie, 1965, 8, pages 181 à 188.

de contraste tenant à l'ordre de correction ("effets séquentiels", dit J.M. FABRE) et montre tout particulièrement le score inférieur obtenu par une copie lorsqu'elle est examinée après une excellente vis-à-vis de laquelle elle apparaît quelconque. Ces expériences ont été reprises en 1971 par BONNIOL et PIOLAT (1) qui ont testé l'introduction d'ancres de forces variées à l'intérieur de lots de copies de mathématiques et d'anglais. Elles ont montré de manière définitive la variation d'appréciation consécutive à l'influence de ces ancres sur les autres devoirs.

Nous n'entrerons pas dans le détail des résultats diversifiés recueillis et largement commentés depuis lors. Ces quelques rappels étaient néanmoins nécessaires pour nous permettre d'asseoir notre hypothèse de recherche concernant l'existence d'effets de contraste dans l'examen des travaux plastiques. On ne peut, en effet, admettre leur extension au domaine des A.P. sans circonspection. Les recherches du cadre général ont permis de lier les effets de contraste à l'influence du travail immédiatement corrigé avant, c'est-à-dire par référence à une prise de connaissance linéaire et successive des devoirs.

Or, le principal avantage de la correction des objets plastiques est justement la simultanéité de la perception des travaux. Cette comparaison instantanée des productions sera-t-elle de nature à réduire les effets séquentiels, à permettre de mieux apprécier ce qui sépare ou rapproche certains travaux? Ou bien, des effets de contraste pourront-ils s'y observer tout autant?

On ne corrige pas des dessins un à un comme des copies, mais en les confrontant. Depuis toujours, le plasticien a su profiter de cette possibilité de comparaison simultanée des travaux: Dans l'atelier des artistes, les travaux étaient juxtaposés de manière à provoquer les comparaisons, soit suspendus à un fil, soit posés sur un large plan de travail, voire à même le sol pour permettre le maximum de recul et un jugement "d'ensemble".

Il est à remarquer que peu de P.A.P. disent employer cette méthode qui a pourtant fait ses preuves. Toute autre méthode ne peut pourtant qu'accentuer les effets que nous pourrions recenser ici. L'effet séquentiel dû à la succession des copies ne devrait pas avoir

(1) BONNIOL (J.J.) et PIOLAT (M.), "comparaison des effets d'ancrage obtenus dans une tâche d'évaluation", op. cit..

lieu d'être, pour ce qui nous concerne.

Par contre, le fait d'être placé à côté d'un "trop bon travail" n'est-il pas préjudiciable? Dès le collège (mais peut-être serait-il possible de recenser des exemples au niveau de l'enseignement primaire?), nos élèves savent bien qu'ils n'ont rien à gagner, lors d'un accrochage-bilan, à cultiver ce genre de proximité, nous avons pu maintes fois l'observer, ... sans toutefois que leur sens de la logique les conduise à rechercher le voisinage du plus faible: Seule la fascination de l'oeuvre exceptionnelle paraît frapper les esprits (ce qui tendrait à faire penser que l'ordination s'opère par rapport à une ancre positive).

Qu'en est-il réellement? Bien que ces réactions naturelles soient connues de tous, observe-t-on des effets de contraste simultanés dus à des phénomènes de contiguïté?

La réponse pourrait apparaître également négative: La possibilité d'un balayage global des travaux devrait permettre d'estomper les effets de voisinage. La vue d'ensemble devrait aussi permettre de repérer et de mesurer l'ampleur des écarts, d'en estimer le caractère "accidentel" et de tenir compte dans la note de la valeur exceptionnelle de quelques travaux sans que cela entraîne de réaction par contraste sur le reste du lot.

Mais on peut, ici aussi, craindre une réponse positive. Observer quelques travaux situés à l'écart, soit vers le haut, soit vers le bas de l'échelle mérite réflexion: Effectivement, s'il est possible de faire "aussi bien", c'est que le groupe-classe n'est pas représentatif de ce que l'on peut attendre "moyennement" et que sa production est particulièrement médiocre. Inversement, comme en d'autres matières, quelques travaux particulièrement faibles montrent que la tâche proposée n'était pas sans difficulté et rehaussent le mérite de ceux qui ont su produire quelque chose de "réussi".

Ces derniers propos n'ont rien de spécifique aux A.P., mais l'objet, pour nous, est de voir comment la spécificité du recueil des indices en situation de simultanéité (ainsi, peut-être que la personnalité du P.A.P.) peut modifier les observations réunies dans le cadre général.

Enfin, à supposer que des effets de contraste puissent se faire ressentir, il convient de s'interroger sur les modalités de leur action: S'agit-il d'un effet partiel, localisé à la zone

entourant l'ancre, ce qui justifierait le terme d'effet de proximité? Certaines recherches américaines, notamment celles de JOHNSON et KING (1964) citées par J.J. BONNIOL (1), montraient que l'effet diminue proportionnellement à l'éloignement de l'ancre. Il n'est pas incongru d'envisager qu'il puisse en être de même pour ce qui concerne la disposition spatiale des travaux.

Mais cet effet pourrait être global, le caractère singulier des ancres portant son action sur l'ensemble du lot soumis à la même présentation.

Résumons donc l'ensemble des hypothèses:

1. Existe-t-il, comme dans le cadre général, des effets de contraste affectant le jugement des productions plastiques? (A supposer qu'ils existent, peut-on apprécier la nature de leur amplitude, comparée à celle observée dans la correction des copies?).
2. Ces effets portent-ils leur action de manière globale ou leur importance est-elle fonction de la proximité des travaux ayant, dans le lot, un rôle d'ancrage?
3. Observe-t-on une variation des résultats en fonction du statut des évaluateurs?

(1) BONNIOL (J.J.), Déterminants et mécanismes des comportements d'évaluation d'épreuves scolaires, op. cit., pages 154 à 157.

4.3.2.

DEROULEMENT DE L'EXPERIENCE.

Nous avons choisi, dans un premier temps, de procéder à l'expérimentation suivant deux modalités.

- Le premier dispositif correspondait à la suspension des travaux sur un fil. Par son caractère linéaire, il comportait le maximum de points communs avec la prise de connaissance séquentielle pratiquée dans les matières fondamentales: Lorsque l'ensemble de la production d'une classe est alignée contre un mur, la longueur de l'affichage occasionne bien, effectivement, une prise de connaissance successive des travaux.

- La seconde modalité reposait sur un étalement au sol des travaux en un ensemble compact. La mise à plat des dessins est également largement utilisée, notamment lorsqu'il s'agit de comparer les oeuvres d'un petit groupe.

Le compte-rendu d'expérience qui suit ne concerne finalement que la seconde procédure.

En effet, les premiers résultats recueillis à Lyon en 1984 dans chacune des deux circonstances ont montré une similitude des effets produits.

Le dispositif de juxtaposition horizontale nous a paru pédagogiquement plus séduisant, comme permettant une cohabitation plus intimiste du maître et de ses élèves. Inversement, l'accrochage sur un fil est souvent le reflet d'une manière plus autoritaire de mener l'évaluation: Le fil, aménagé une fois pour toutes, fait partie de la structure de la classe. Il est le lieu du "jugement". Du fait du recul nécessaire, professeur et élèves sont en position frontale, en confrontation.

Mais surtout, le principal argument était d'ordre méthodologique: Tandis que le premier dispositif (linéaire) tendait à favoriser des phénomènes propres à l'appréhension séquentielle, le second dispositif, celui de la juxtaposition, contribuait, tout au contraire, à les atténuer. Observer néanmoins, dans ces conditions, des effets de contraste, devait mieux permettre, à notre sens, de penser que ces derniers ne pourraient être que plus importants en d'autres situations.

Deux lots de travaux ont donc été préparés et fixés sur deux planches de carton fort, à l'identique du reste des expériences.

Nous avons choisi comme support de cette expérience des exercices réalisés en classe de 5° de manière très conventionnelle: Il s'agissait d'un travail d'apprentissage technique portant sur le maniement de la peinture et dont le résultat se limitait à peindre "la mer surmontée d'un ciel". Selon nous, la très grande similitude des travaux devait permettre de mieux repérer tout travail nettement en dehors de l'éventail des résultats "courants" et qui pourraient faire office d'ancre (donc, de s'en prémunir), elle devait aussi favoriser la fixation d'un modèle de référence stable, conditions qui nous semblaient de nature à freiner l'apparition d'effets de contraste, notre souci étant de ne rien forger sur le plan protocolaire qui puisse en favoriser l'apparition.

Nous disposions donc de deux panneaux d'égale grandeur sur lesquels étaient disposées 12 peintures de format 24 x 32 représentant la mer. Les différences de tonalité avaient été réparties au mieux de manière à conférer aux deux ensembles la même homogénéité: Les tendances bleues et vertes alternaient avec les travaux menés dans des tonalités plus orangées ou plus mauves.

Nous avons indiqué que ce travail était celui d'élèves de 5° de collège n'ayant bénéficié d'aucun cours d'éducation plastique en 6°: Situé vers la fin du premier trimestre, l'objectif était, après l'apprentissage des mélanges de couleurs, de faire prendre conscience, dans l'acte de peindre, de l'importance de ces mélanges (nuances riches et personnelles), du rôle de la dilution selon les matières que l'on peint (ici, le vaporeux du ciel s'opposait au caractère plus ferme des vagues de la mer) ainsi que du rôle essentiel du geste de pinceau véhiculant ces couleurs. A cet effet, l'exercice



avait été mené en deux séances et préparé par le visionnement d'un nombre important de diapositives représentant des ciels et des étendues marines, de toutes tonalités, de toutes sortes, tout autant sous forme de photographies que de reproductions d'oeuvres peintes de diverses époques et civilisations. Les élèves devaient ensuite "inventer" le ciel de leur choix, puis la mer qui en portait le reflet, ajouter enfin un bateau ou quelque autre élément nageant ou vogant.

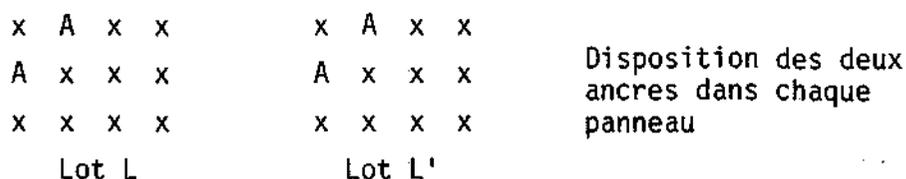
Les travaux présentés avaient été sélectionnés comme pouvant être situés entre le "médiocre" et le "convenable", c'est-à-dire amputés des productions les plus extrêmes. Les deux ensembles conservaient donc l'apparence d'une production habituelle de classe (les dessins provenaient, en fait, de plusieurs classes parallèles) mais laissant néanmoins une marge pour la mise en place des ancrés.

Reprenant en cela le dispositif utilisé en 1971 par BONNIOL et PIOLAT, nous avons, lors de la première phase, introduit dans ces planches deux couples d'ancres fortement typées positivement. Ensuite, après évaluation, les lots furent échangés d'un groupe d'évaluateurs à l'autre et, à cette occasion, les ancrés positives furent remplacées par des travaux particulièrement faibles.

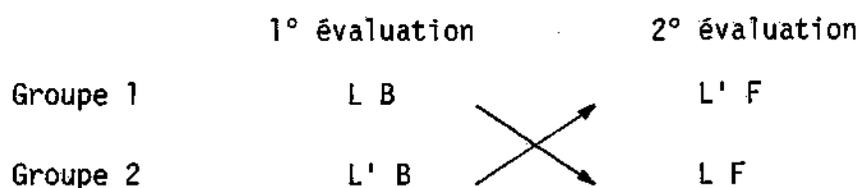
L'expérimentation a été menée dans deux académies. D'une part à Lyon où il s'est agi de 14 professeurs de tous statuts réunis volontairement dans le cadre d'un stage. D'autre part à Marseille où l'évaluation reposait sur 24 professeurs nouvellement certifiés présents dans le cadre du Centre Pédagogique Régional.

Nous avons divisé les groupes de professeurs en deux parties numériquement identiques mais sans intervenir dans la constitution des groupes de manière à ne pas heurter des affinités et à ménager une bonne ambiance de travail.

Conformément au schéma ci-après, l'ensemble des participants a évalué l'ensemble des oeuvres, dans une première étape accompagnées de deux ancrés positives, dans la seconde accompagnées de deux ancrés négatives.



Dispositif de croisement des lots:



Le croisement des lots permettait de gommer les différences d'appréciation d'un groupe à l'autre car n'était considéré que la moyenne des notes attribuées aux mêmes dessins dans chacune des deux situations.

En résumé, notre protocole comportait:

- Deux facteurs aléatoires: Les deux groupes hétérogènes de 19 professeurs et les 24 travaux d'élèves.
- Le facteur systématique était constitué par la variation de nature des ancrés.
- La variable dépendante était la moyenne des notes attribuées aux travaux contigus à ces ancrés selon que celles-ci étaient "bonnes" (B) ou "faibles" (F).

4.3.3.

RESULTATS DE L'EXPERIENCE.

1) Existence d'un effet de contraste global par juxtaposition.

La question préalable était celle de l'opérationnalité des ancrés utilisés. Effectivement, celles-ci ont bien été perçues, ainsi qu'en témoignent les moyennes des notes attribuées:

Moyenne générale portant sur 380 notes:	51,12
Moyenne des ancrés ⁺ (76 notes):	63,17
Moyenne des ancrés ⁻ (76 notes):	40,15

Tableau des résultats obtenus par chaque dessin en fonction des deux occurrences:

LB				LF				ΔL			
42	<u>60,7</u>	54	47,8	51,8	<u>39,3</u>	52,4	49,3	9,8	<u>21,4</u>	-1,2	1,5
<u>64,4</u>	56,4	47,1	52,3	<u>39,2</u>	52,9	55,6	54,7	<u>25,2</u>	-3,5	8,5	2,4
48,2	51	55,6	50,3	48,7	55	54,2	52,6	0,5	4	-1,4	2,3

Moyennes portant sur 190 notes:

MLB = 50,47	MLF = 52,72
-------------	-------------

L'B				L'F				$\Delta L'$			
46,1	<u>64,9</u>	49,5	46,4	53,6	<u>41,2</u>	51,9	51,1	7,5	<u>23,7</u>	2,4	4,7
<u>62,7</u>	52,3	48,9	53,7	<u>40,9</u>	48,9	54,5	59,2	<u>21,8</u>	-3,4	5,6	5,5
46,6	49,6	51,7	51,3	49,7	47,9	51	49	3,1	-1,7	-0,7	-2,3

ML'B = 49,61	ML'F = 51,68
--------------	--------------

Moyenne générale portant sur 380 notes:

<u>MB = 50,04</u>	<u>MF = 52,20</u>
-------------------	-------------------

Différence de moyenne entre les deux occurrences : $\Delta M = 2,16$

Cette différence est particulièrement significative. Le calcul du $|t|$ de Student donne (pour 379 D.L., il est vrai) le résultat suivant:

$$|t| = 10,96, \text{ significatif à } .001.$$

2) Existence d'un effet de proximité.

Nous avons pris soin de placer nos ancrés près d'un angle afin de pouvoir envisager la question de la proximité de l'ancré. L'effet de contraste tendant à s'atténuer dans le temps, on pouvait se demander si ce phénomène ne se vérifiait pas aussi dans l'espace.

Sur chacune des deux planches, l'emplacement privilégié correspondant à une proximité maximale des ancrés était celui de l'angle en haut à gauche (X).

(X)	A	x	x
	A	x	x
	x	x	x

L'examen des tableaux des différences enregistrées (ΔL et $\Delta L'$) permet de constater que ces amplitudes sont particulièrement fortes pour les travaux X et X': 9,8 et 7,5.

Moyenne de différence pour les travaux proches: 8,65 (38 n)

Moyenne pour les autres travaux (342 notes) : 1,46

Au premier abord, la différence entre la moyenne des amplitudes enregistrées sur les oeuvres proches et celle correspondant aux autres oeuvres semble particulièrement éloquente ($\Delta d = 7,19$).

Cependant, un examen plus attentif des tableaux des différences fait apparaître en d'autres lieux ne pouvant être considérés comme proches des amplitudes moyennes presque aussi importantes: 8,5 et 5,6.

Nous considérerons donc cette observation comme non vérifiée expérimentalement et nous nous bornerons à signaler comme "possible" une aggravation de l'effet de contiguïté en fonction de la proximité aux travaux extrêmes.

3) Influence du statut professionnel des évaluateurs.

Cette forte dépendance vis-à-vis des effets de contraste par contiguïté apparaît au premier abord comme assez surprenante. Certes, l'évaluateur, face à un travail excellent, est bien obligé de constater qu'un tel degré de réussite est possible, mais il devrait être en mesure, par son expérience, d'en peser le caractère "normal" ou "exceptionnel". Inversement, la faiblesse de certains travaux ne devrait pas suffire à assurer de la bonne qualité des autres productions.

Il nous est donc apparu que cette dépendance pouvait être en rapport avec l'expérience professionnelle des enseignants, un jeune professeur, faute de références antérieures suffisamment nombreuses, risquant de se laisser influencer à cet égard davantage que son aîné, mieux à même de relativiser ses observations.

Nous avons donc comparé les mesures effectuées à Lyon auprès de 14 professeurs en cours de carrière et celles de Marseille provenant de 24 professeurs stagiaires en formation pédagogique.

a) Moyenne des notes attribuées à chaque dessin par les professeurs stagiaires, reposant sur 240 notes:

$$\begin{array}{l} \text{MB} = 49,41 \\ \text{MF} = 52,26 \end{array} \qquad \underline{\Delta M = 2,85}$$

b) Moyenne des notes attribuées à chaque dessin par les professeurs en cours de carrière, reposant sur 140 notes:

$$\begin{array}{l} \text{MB} = 51,73 \\ \text{MF} = 52,56 \end{array} \qquad \underline{\Delta M = 0,83}$$

On observera que la différence de notation est beaucoup moins marquée dans le second groupe. La netteté de cet écart est confirmé par l'analyse de variance (1):

	S C	D L	C M	rapport F
contiguïté	51,51	1	51,51	32,8
ancienneté	28,14	1	28,14	17,9
C x An	10,17	1	10,17	6,4
résidu	82,15	52	1,57	
total	171,97	55		

} très significatifs

(1) Voir en annexe n°10 l'ensemble des calculs correspondant à cette expérimentation.

La prise en compte de l'influence de travaux contigus semble donc plus sensible chez les jeunes professeurs. La dépendance vis-à-vis des effets de contraste s'estomperait avec l'expérience professionnelle.

La comparaison aboutissant à la notation s'effectue bien par rapport à un produit attendu d'autant mieux solidifié que l'enseignant bénéficie d'une expérience antérieure. Dans ce cas, le niveau particulier du lot qu'il est en train de corriger ne modifie guère l'image qu'il s'est forgée de la fourchette de production: Tout travail "anormalement" faible ou excellent sera alors perçu comme exceptionnel et n'influencera pas le jugement porté sur les autres oeuvres. Le jeune professeur, quant à lui, construirait davantage son évaluation en fonction exclusive de l'éventail des niveaux constaté dans le lot qu'il corrige.

La conclusion de cette expérimentation débouche sur la mise en évidence d'un effet de contraste par contiguïté affectant la correction des travaux plastiques en situation de juxtaposition.

Cet effet diffère sensiblement de l'effet de contraste séquentiel repéré dans le contexte général, en raison, principalement, de la simultanéité de vision que permet la correction.

Un effet de proximité, s'il est possible, n'apparaît pas aussi nettement et laisse plutôt croire à un effet global sur l'ensemble du lot.

L'effet de contraste par voisinage semble pouvoir se maîtriser à mesure que l'expérience professionnelle s'accroît.

Pour en marquer la spécificité, nous proposerons son identification sous la dénomination: "Effet de contiguïté".

EXPERIENCES (2):
EFFETS LIES A LA PERCEPTION DES MESSAGES
VISUELS: EMERGENCE D'UNE SPECIFICITE.

5

5.1. INFLUENCE DE LA VARIABLE COLORIS.

- 511 - Hypothèse de travail.
- 512 - Déroulement de l'expérience.
- 513 - Résultats.

5.2. INFLUENCE DE LA VARIABLE COMPLEXITE.

- 521 - Hypothèse de travail.
- 522 - Déroulement de l'expérience.
- 523 - Résultats.

5.3. SYNTHÈSE DES RESULTATS: VERS LA PROPOSITION D'UN MODELE EXPLICATIF.

- 531 - Recueil des indices: Facteur évaluateur et représentations.
- 532 - Prégnance perceptive et préférence esthétique: Extension validante.
- 533 - L'évaluation en A.P.: Un mécanisme spécifique.

Dans une première série d'expériences, nous avons montré en quoi l'évaluation des productions plastiques se différencie du contexte général: Si des effets semblables peuvent être répertoriés, ils se présentent suivant des modalités particulières qui tiennent compte de données matérielles spécifiques.

Cependant, ces premières observations ne seraient pas suffisantes pour conclure à un fonctionnement différent. Poser l'hypothèse d'une spécificité de l'évaluation des productions plastiques implique que soient mis en avant des déterminants nouveaux liés à la dimension plastique des objets à évaluer.

Cette seconde phase d'expériences a pour objectif de faire apparaître l'influence de facteurs liés au contenu des travaux.

En effet, le recueil des indices nécessaires à l'évaluation est, dans le cas des arts plastiques, tributaire des difficultés rencontrées dans la lecture des faits visuels: L'activité perceptive du sujet évaluateur (variables de personnalité) en face d'un objet plastique (variables du stimulus) débouche sur une réaction comportementale où se mêlent probablement le cognitif et l'hédonique.

Il y a tout lieu de penser que le recueil des indices, bien qu'initialement guidé par une mise au net des critères d'évaluation, puisse se trouver influencé par la prégnance perceptive d'indices extérieurs au problème (quoique pertinents dans un autre cadre conceptuel) qui continueraient à s'imposer à l'évaluateur du fait d'une réceptivité particulière de celui-ci à leur égard.

Afin de mettre en évidence ces effets liés à des considérations d'ordre hédonique, nous avons pensé qu'il serait possible d'étudier quelles seraient les conséquences sur la notation de la variation systématique d'une "variable-stimulus" en principe inopérante par rapport aux consignes ayant guidé la réalisation d'un lot de travaux.

Le problème initial était celui du choix des variables pouvant être retenues comme support de ces expériences. Elles devaient répondre à plusieurs critères:

- Etre suffisamment représentatives des dimensions plastiques des oeuvres et, en cela, avoir déjà fait l'objet de nombreuses expériences convergentes en psychologie de l'esthétique mettant en avant leur importance toute particulière.

- Etre en oeuvre et facilement identifiables dans le cas de productions scolaires de l'enseignement secondaire.

- Permettre sans trop d'artificialité l'opération de fabrication de "dessins construits" nécessaire aux manipulations expérimentales.

Nous avons finalement retenu deux variables bi-polaires se rapportant au coloris et à la complexité des propositions, espérant que des résultats nets quant à la prise en compte de ces variables pourront être une invitation à avancer l'hypothèse d'une possible extension validante à un nombre plus important de variables.

5.1. INFLUENCE DE LA VARIABLE COLORIS.

5.1.1.

HYPOTHESE DE TRAVAIL.

La couleur est une variable dont l'importance n'échappe à personne: Elle est régulièrement citée pour argumenter des choix en matière de peinture. Nous savons aussi que tout objet visuel s'impose d'abord au regardeur au moyen d'une information colorée qui déclenche chez celui-ci les premières réactions perceptives. Les recherches contemporaines précisent même, nous l'avons vu, que l'excitation rétinienne se trouve amorcée dès l'apparition des moindres agglomérations pigmentaires et ceci avant toute prise de conscience.

Très tôt, la question du coloris a animé l'histoire de la peinture, opposant le plus souvent les partisans des tons rompus organisés à l'atelier aux utilisateurs des couleurs franches, mieux à l'image de la vie.

Très tôt également, les chercheurs en psychologie de l'esthétique se sont inquiétés de savoir ce qu'il pouvait en être du goût à cet égard.

Les premières recherches commencèrent par l'étude de la préférence individuelle en matière de couleur. L'imprécision des résultats tenait à un contrôle insuffisant des trois dimensions du stimulus. En effet, une couleur se définit au moyen de trois variables:

- La qualité (sa longueur d'onde dominante, appelée aussi tonalité).
- L'intensité (luminance, brillance, valeur).

- La vivacité (pureté, saturation) (1).

Une des premières études est celle de ST GEORGE (2), portant sur 500 élèves issus d'un collège américain. Cette enquête de 1938 révèle l'ordre de préférence suivant: Bleu (B), vert (V), rouge (R), jaune (J), orange (O), violet (W), blanc (Bc).

A cette même époque, de nombreuses expériences dues à R.B. et J.P. GUILFORD, étalées sur une période allant de 1933 à 1959, tendent à montrer sous une corrélation forte deux ordres de préférence:

- Les hommes préféreraient B, R, V, et bien plus loin O et J.
- Les femmes préféreraient R, V, B, puis J et O (il s'agit de couleurs pleinement saturées) (3).

Ces résultats vont dans le même sens que ceux de CHESKIN en 1948 (4) (les femmes préfèrent le rouge) et de CHOUNGOURIAN (5) en 1968 (les hommes préfèrent le bleu).

Nous citerons encore les recherches de H.J. EYSENCK (1941). Reprenant par l'analyse factorielle des travaux plus anciens (ALLESCH, 1925), l'auteur a pu retirer une observation particulièrement cohérente, quelque soient le sexe, le milieu social ou la race (6). L'ordre des préférences serait: B, R, V, puis W, O, J. EYSENCK en déduit l'existence d'un facteur général de goût, qu'il nomme facteur T, puis il met en évidence un deuxième facteur discriminant qui opposerait ensuite les partisans des couleurs vives à ceux des couleurs ternes.

Cet ordre est confirmé par GRANGER en 1965 (7) qui utilise des couleurs aux caractéristiques colorimétriques très précises (système MUNSELL).

L'ensemble de ces expériences américaines porte généralement sur des couleurs vives, en pleine saturation. Toutefois, une couleur se définit aussi par ses deux autres composantes, la vivacité et l'intensité. A cet égard, GUILFORD a pu montrer en 1934 que, pour

(1) FRANCES (R.), Psychologie de l'art et de l'esthétique, op. cit., page 69.

(2) ST GEORGE (M.W.), "Color preferences of college students with references to chromatic pull, learning & association", American Journal of Psychology, 1938, 5, pages 714 à 716.

(3) GUILFORD (J.P.) et SMITH (P.C.), "A system of color preferences", American Journal of Psychology, 1959, 72, pages 487 à 502.

(4) CHESKIN, cité par BERNARD (Y.), Psycho-sociologie du goût en matière de peinture, op. cit., page 24.

(5) CHOUNGOURIAN (A.), "Color preferences. A cross-cultural and cross-sectional study", Perceptual and Motor Skills, 1969, 28, pages 801-802.

(6) EYSENCK (H.J.), "A critical and experimental study of color preferences", American Journal of Psychology, 1941, 54, pages 385 à 394.

(7) GRANGER (G.W.), "The prediction of preference for color harmony", The Journal of General Psychology, 1956, 52, pages 213 à 222.

une même tonalité, la préférence augmentait en fonction de l'intensité et de la vivacité. Par ailleurs, les différences d'agrément dans le classement B V R W O J s'accuseraient dans les brillances faibles et s'atténueraient en brillancés fortes (1). GRANGER, en 1956 (2), confirme les travaux de GUILFORD. La pureté d'une couleur est appréciée jusqu'à un seuil élevé (chroma 8), mais l'intensité d'un coloris est préférée dans une brillance moyenne.

Bien entendu, on ne peut que se montrer circonspect vis-à-vis d'observations ne mettant en cause qu'une seule couleur, isolément. Une couleur est rarement vue seule (la plupart de ces expériences ont été réalisées sur un fond gris neutre de brillance moyenne). Sur la question des combinaisons de couleurs, il semblerait qu'il y ait un accord assez net pour les combinaisons ayant un intervalle de tonalité assez important.

De plus, ces observations ne sont pas indépendantes de considérations d'ordre génétique ou culturel. WOODS en 1956, JAMPOLSKY en 1954, FRANCES en 1966 (3), montrent que les enfants préfèrent les couleurs vives, chaudes, les assemblages bigarrés tandis que les adultes préfèrent les contrastes plus discrets. Y. BERNARD et J. BOYER, en 1976 (4), ont noté de leur côté que les sujets peu cultivés sont surtout sensibles à la couleur d'un tableau aux dépens de considérations plus techniques.

La préférence accordée à une couleur n'est donc qu'un artifice schématique permettant dans un premier temps de clarifier la très grande complexité des mécanismes d'attraction perceptive liés à la couleur.

Les observations recueillies par les chercheurs, tout en n'étant que très partielles, donnaient l'illusion de posséder une certaine unité. Mais cette cohérence n'est qu'apparente: Il nous faut maintenant faire état de travaux plus récents qui contredisent en partie les données antérieures et dont on peut penser que la fiabilité scientifique est tout autant recevable.

(1) GUILFORD (J.P.), "The affective value of color as a function of hue, tint and chroma", Journal of Experimental Psychology, 1934, 17, pages 342 à 370.

(2) GRANGER (G.W.), "An experimental study of color preference", The Journal of general Psychology, 1956, 52, pages 3 à 20.

(3) Cité par FRANCES (R.), Psychologie de l'art et de l'esthétique, op. cit., pages 69 à 72.

(4) Ibid., page 129.

Ainsi, NEMCSICS, en 1967 (1), portant son investigation sur 100.000 sujets échantillonnés avec rigueur, observe que, si les moins de vingt ans se comportent effectivement d'une manière assez proche de celles mises en évidence par les chercheurs précédemment cités, les plus de trente ans attribueraient une place plus importante à l'orangé, classant les couleurs ainsi: R O B V J W.

L'époque contemporaine a permis l'approche de la préférence esthétique au moyen d'autres outils d'investigation. Il s'agit notamment de la recherche des indices électrophysiologiques. Dans ce cadre-là non plus, les résultats obtenus ne viennent pas confirmer les observations recueillies antérieurement.

Signalons la recherche de SMETS, datant de 1973 (1), se fondant sur les rythmes électrocorticaux. L'ordre obtenu est radicalement différent: J O R B V W.

DARAKIS trouve en 1976 (1) des réponses électrodermales (rendant compte de l'impact affectif des couleurs) assez proches de celles de SMETS dans la mesure où elles observent également une attitude très différente à l'égard du jaune: J V W R. Cependant, ce revirement n'est pas du tout général et d'autres expériences portant sur la réponse électrodermale font apparaître la prédominance du vert et du rouge.

Au moment où nous cherchons à réunir les informations scientifiques qui vont nous permettre d'étayer nos hypothèses de travail, on ne peut que rester perplexe devant l'apparente incohérence des résultats en notre possession. Mais, à la lumière des travaux de F. MOLNAR, ces divergences pourraient n'être que superficielles.

F. MOLNAR lie la perception des objets visuels au parcours du regard à l'intérieur du stimulus (2). Deux ordres d'observations vont retenir notre attention. Le premier point concerne l'attirance spontanée du regard vers certaines tonalités de couleurs. MOLNAR montre que l'oeil est principalement attiré par les couleurs proches du jaune (O, V, J), ce qui pourrait expliquer l'importance observée pour cette couleur dans les expériences électrophysiologiques. Il se pourrait que cette attirance - de l'ordre du réflexe - soit ensuite,

(1) Recherches citées par FRANCES (R.), *ibid.*, pages 73 à 76.

(2) MOLNAR (F.), "Les mouvements exploratoires des yeux dans la composition picturale", *Sciences de l'art*, 1964, 1, pages 135 à 150.

à des degrés variables, surmontée voire contredite par la préférence consciente tributaire notamment de certains déterminants liés au sexe, à l'âge ou au niveau culturel du sujet. Pour corroborer cette hypothèse, nous nous appuyerons sur le fait qu'il existe de nombreuses recherches comme celles de OBST, TAYLOR, KNAPP que mentionne Y. BERNARD (1) qui font état, à la différence d'EYSENCK, de différences significatives liées à la catégorisation des sujets.

Compte tenu de ces observations aux directions parfois contradictoires, que peut-on retenir de suffisamment solide pour orienter nos investigations?

D'un côté, il apparaît certain que la dominante colorée des objets plastiques exerce une influence sur la résonance éveillée chez le sujet regardeur en fonction de ses affinités. Mais, d'autre part, le manque de certitude quant à la convergence vers un facteur général du goût identifiable ne permet pas de déduire des connaissances actuelles en psychologie de l'esthétique des mécanismes de fonctionnement propres à l'appréhension des travaux d'élèves.

De plus, le P.A.P. n'étant pas un adulte "comme les autres", mais se différenciant à plusieurs égards (formation universitaire, expérience professionnelle d'une cohabitation journalière avec les oeuvres plastiques) des sujets ayant pris part aux expériences mentionnées, il n'est pas du tout certain qu'il puisse y avoir identification de ses réactions à celles répertoriées par les chercheurs.

La première question est donc celle-ci: Y a-t-il réellement prise en compte de la dimension "coloris" lors de l'examen des travaux plastiques et est-ce que cette prise en compte continue à être opérante dans des situations où celle-ci n'est pas justifiée par les objectifs de l'évaluation?

Il reste ensuite à préciser sous quelles modalités la couleur se fait plus prégnante, quels en sont les aspects les plus discriminants, enfin, à s'interroger sur l'importance du biais en fonction des catégories d'évaluateurs.

Pour mettre le maximum de chances de notre côté sur le plan de l'authenticité des résultats, nous avons choisi de placer en opposition les indices cités majoritairement comme les plus positifs

(1) BERNARD (Y.), Psycho-sociologie du goût en matière de peinture, op. cit., page 25.

et les facteurs signalés comme les plus négatifs.

Pour ce qui concerne la qualité de la couleur, nous avons choisi le bleu, le vert et le rouge en opposition à des couleurs plus neutres comme gris, brun ou violet.

Pour ce qui tient à la vivacité, nous avons opposé les tons saturés aux couleurs neutres rabattues.

Nous n'avons pas fait intervenir l'intensité des coloris, les présentant d'une valeur moyenne correspondant à une même catégorie d'encre de crayons feutres.

Nous avons donc pensé établir une nette différence de séduction entre des catégories de couleurs franches et vives et des tons neutres très peu saturés. Cette opposition grossière ne pouvait manquer d'être ressentie par la majorité des sujets: La question était de savoir si cette probable perception affecterait le jugement de P.A.P. lorsque la couleur n'interviendrait pas en tant que critère pertinent d'évaluation.

Les différents niveaux de l'hypothèse de travail sous-tendant cette réflexion peuvent désormais être énoncés avec précision. Il y sera distingué trois étapes:

1) Le P.A.P. reste-t-il dépendant de la variable "coloris" indépendamment des objectifs pédagogiques définis? (Y a-t-il une corrélation significative entre la variation du coloris et les notes attribuées?).

2) Le P.A.P. se comporte-t-il dans son évaluation à l'identique de la majorité des sujets observés?

3) Les P.A.P. réagissent-ils de manière homogène ou leur sensibilité à l'effet de la couleur est-elle dépendante de leur statut?

5.1.2.

DEROULEMENT DE L' EXPERIENCE.

Pour tester l'influence du coloris sur l'appréciation des professeurs, il fallait que cette influence soit observée indépendamment du sujet représenté ou, plus généralement, des intentions expressives de son auteur.

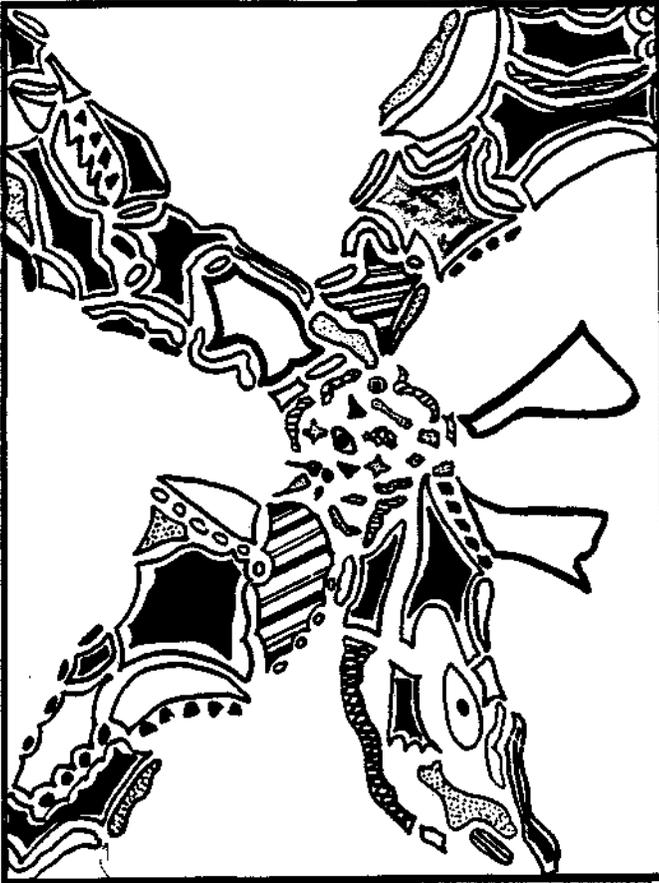
En effet, dans toute oeuvre plastique, la couleur n'est pas une variable qu'il y a lieu de juger "en soi". Il n'y a pas de couleur "belle" ou "laide". La couleur est toujours à considérer relativement à un ensemble décoratif ou par rapport à un contenu qu'elle accompagne pour servir l'expression. La prise en compte de la couleur dans un travail s'exprime donc plutôt en termes de cohérence-pertinence par rapport à des intentions.

L'objet de notre interrogation était tout autre, rappelons-le: L'appréciation des oeuvres scolaires n'est-elle pas perturbée par l'attirance des professeurs pour telle ou telle dominante colorée?

Pour mettre en avant l'interférence de considérations hédoniques, nous avons choisi comme support des compositions graphiques monochromatiques provenant de classes de 5° et exécutées aux crayons feutres.

L'exercice, présenté de manière traditionnelle comme tous ceux qui ont servi à mettre en valeur les effets liés à des biais d'ordre perceptif, consistait à créer un ensemble structuré de graphismes abstraits et à les disposer à l'intérieur du cadre de la feuille 24 x 32. L'objectif de ce travail était sous-tendu par plusieurs ambitions:

- 1) Entraînement à l'imagination par la création de formes.
- 2) Mise en oeuvre de la pensée logique par la structuration de l'ensemble: Les éléments devaient s'enchaîner pertinemment les uns aux autres suivant une certaine cohérence dont l'élève avait à définir



les règles, et non point être saupoudrés au hasard dans la surface.

3) Cet objectif comportait enfin des notions d'esthétique traditionnelle qui se fondaient sur l'unité/variété des formes et des masses ainsi que sur leur disposition qui devait être équilibrée à l'intérieur de la page.

Les critères d'évaluation étaient donc particulièrement bien circonscrits. En revanche, la nature de la couleur de réalisation (unique) était laissée à l'initiative de l'élève: Que l'exécution soit faite en rouge, vert, noir ou brun importait peu. La couleur ne devait donc pas entrer en ligne de compte.

Afin de ne pas mettre en éveil la vigilance des évaluateurs à propos du coloris et d'autre part parce que les réponses des P.A.P. auraient été à la fois trop complexes et trop difficilement vérifiables, nous n'avons pas cherché à faire apparaître directement une corrélation entre le goût exprimé par l'évaluateur en matière de couleurs et les notes attribuées aux graphismes. Plus simplement, nous avons choisi de mettre en opposition, afin d'en observer les conséquences, des coloris purs et saturés (B V R), jugés plus attirants, et des tonalités neutres.

Pour observer l'influence de la variation de la couleur, la méthode la plus rigoureuse consistait à proposer le même dessin, toutes choses égales par ailleurs, suivant deux colorations différentes. L'intention, totalement irréalisable dans la plupart des cas, était, par contre, aisée à mettre en place dans le cas de compositions graphiques: Il suffisait de décalquer les travaux d'un certain nombre d'élèves dans des encres de caractéristiques opposées.

Nous avons donc proposé à l'évaluation de deux groupes de professeurs deux panneaux de 12 dessins variés, plus précisément, composés de 3 "bons", 6 "moyens", 3 "faibles", et comportant des compositions aux coloris vifs et d'autres aux coloris ternes.

A l'occasion du croisement des groupes, nous avons procédé à la substitution des travaux tests (un par niveau qualitatif):

x B x x	x B' x x
x x x M	x x x M'
x F x x	x F' x x

La majeure partie des travaux (x) ne jouait donc simplement qu'un rôle de groupe témoin destiné à "noyer" les dessins-tests dans un ensemble à la présentation identique aux autres expériences. Seules furent prises en considération les moyennes des notes attribuées aux dessins-tests

- présentés suivant une tonalité terne;
- présentés suivant une tonalité vive.

Comme pour les expériences précédentes, nous avons recueilli nos informations à l'occasion de plusieurs stages académiques. Les notes collationnées proviennent de

- 14 professeurs titulaires de l'académie de Lyon;
- 10 professeurs titulaires de l'académie de Grenoble;
- 24 professeurs stagiaires de l'académie d'Aix-Marseille;
- 10 adjoints d'enseignement titularisables de la même académie.

Le protocole choisi comportait:

- Deux facteurs aléatoires: Les professeurs évaluateurs répartis en 2 groupes; les 2 panneaux de 12 travaux.
- Le facteur systématique était constitué par les 2 modalités de la variable coloris (vif - terne: V/T) affectant 2 couples de 3 dessins.
- La variable dépendante était la moyenne des notes sur 20 attribuées par chaque évaluateur à chaque dessin-test suivant chacune des deux occurrences.

5.1.3.

RESULTATS DE L' EXPERIENCE.

1) Influence globale de la variable coloris.

Pour mettre en évidence l'influence de la variable coloris, nous avons mis en parallèle les mesures obtenues par les 6 travaux en coloris vifs et leurs 6 homologues de tonalités ternes:

		terne	vif	différence
29) GR	B	60,79	61,12	0,33
	M	51,49	54	2,51
	F	33,82	39,89	6,03
(29) GR'	B	51,57	55,62	4,05
	M	46,41	49,59	3,18
	F	39,72	43,45	3,75

Moyennes portant sur 29 x 6 = 174 notes:

$$\underline{Mt = 47,30} \quad \underline{Mv = 50,61} \quad \underline{\Delta = 3,31}$$

Cette différence est statistiquement significative ainsi qu'en atteste le calcul du |t| de STUDENT (1):

$$|t| = 4,24, \text{ significatif à } .01.$$

Plusieurs observations sont déductibles de ce résultat:

- En premier lieu, il y a eu entre les correcteurs un accord moyen quant au niveau des dessins-tests (B, M, F), ce qui atteste d'une homogénéité globale de manipulation des critères et contribue à donner plus de poids aux variations dues au coloris.

- Les professeurs d'arts plastiques sont sensibles au

(1) Voir l'ensemble des calculs de cette série en annexe n° 11.

coloris général d'un travail. Ce qu'il faut bien appeler un "goût" pour telle ou telle tonalité continue à se manifester indépendamment des critères opérants dans l'exercice proposé aux élèves, ce qui confirme donc l'hypothèse que nous avançons.

Ainsi, cette première expérience relative aux variables-stimuli fait apparaître que même l'exercice technique le plus conventionnel n'est pas à l'abri des préférences de l'évaluateur.

Cette observation limite considérablement le rôle généralement accordé aux exercices de maîtrise et nous rend prudents à tous niveaux de jugement. Mais surtout, il rend perplexe quant à la part que doit prendre la subjectivité du goût pour ce qui concerne les oeuvres de libre expression où les critères de maîtrise sont plus difficiles à cerner et ne sont pas les seuls opérant.

Enfin, constatons que la préférence des P.A.P. se porte sur les tons vifs au détriment des coloris ternes, ce qui est une information importante compte tenu du fait que les tons rompus sont souvent le siège de raffinements de nuances plus subtils que les plasticiens apprécient généralement. Le comportement est donc identique à celui des sujets non spécialistes. Cela inciterait à penser que cette attitude est à lier à un comportement universel de l'individu correspondant au facteur général de goût. Ce comportement ne dépendrait pas de données acquises culturellement (et ici professionnellement) mais viendrait du plus profond de l'individu. Nous avons tout lieu de craindre alors que celui-ci risque d'être rebelle à toute tentative de correction. Toutefois, son caractère systématique laisse espérer que l'on puisse apprendre à s'en prémunir.

2) Influence du statut des évaluateurs.

a) Formation professionnelle.

Une autre interprétation pourrait être que l'attrance pour les coloris vifs, identique au comportement universel, serait due à une catégorie d'enseignants moins familiarisée avec les faits artistiques. D'une manière générale, cette incapacité à isoler une donnée visuelle aussi élémentaire que celle de la tonalité globale d'une oeuvre (il ne s'agit pas ici de rapports complexes de coloris en interaction avec l'expression d'un thème) a de quoi surprendre, même si nous en soupçonnions l'existence.

Nous avons donc recherché comment se présentaient les résultats précédents en regroupant les données suivant deux catégories:

- Les professeurs formés (C.P.R., certifiés, agrégés).
- Les professeurs peu formés (P.E.G.C., A.E.).

La comparaison des séries de moyennes donne le résultat suivant:

<u>Professeurs formés:</u>	<u>Professeurs peu formés:</u>
Moyenne "Terne" = 48,63	Moyenne "terne" = 45,50
Moyenne "vif" = 51,18	Moyenne "vif" = 49,77
Différence = 2,55	Différence = 4,27

On peut constater que la différence due à l'influence du coloris est nettement plus importante dans la catégorie des personnels peu formés.

Afin de vérifier cette observation, nous avons procédé à une analyse de variance portant sur le croisement du facteur C (coloris) et du facteur P (professeurs). Le tableau d'analyse de variance (1) se présente ainsi:

	D.L.	S.C.	C.M.	F	signification
Professeurs	1	309	309	4,85	significatif à .05
Coloris	1	698	698	10,94	significatif à .01
P x C	1	44	44	< 1	N.S.
Résidu	236	15042	63,7		
Total	239	16093			

Contrairement à notre attente, s'il y a vérification de l'influence du facteur coloris dans l'évaluation, si la différence de nature entre les deux groupes de professeurs produit bien une variation significative, les données recueillies n'ont pas permis de confirmer notre hypothèse portant sur le rôle de la formation professionnelle dans l'appréciation du coloris. Les chiffres vont néanmoins dans ce sens, ce qui inciterait à recommencer l'expérience, peut-être avec davantage de sujets, ou bien avec des sujets départagés avec davantage de rigueur.

(1) Voir annexe n° 11 b.

b) L'âge des évaluateurs.

L'attrance pour les coloris vifs pouvait amener à s'interroger également sur la question de l'âge de l'évaluateur, un sujet jeune appréciant généralement davantage les tons vifs saturés tandis que l'âge conduit souvent à préférer les teintes plus assourdis.

On peut considérer les recherches statistiques négatives à cet égard: L'écart des différences entre terne et vif est peu important d'un groupe à l'autre:

Différence T/V, moins de 35 ans: 2,99

Différence T/V, 35 ans et plus : 3,21

L'analyse de variance (1) faisant apparaître l'action du croisement coloris/âge n'est pas significative sur ce dernier point (il y a seulement confirmation de l'action de la variable coloris) :

F de Snédécour : Coloris : F = 9,63 significatif à .01

Age : F < 1 N.S.

C x A F < 1 N.S.

Conclusion.

Ainsi, le résultat principal de cette expérimentation porte sur l'influence effective du facteur coloris dans l'évaluation des travaux plastiques scolaires en dehors de toute référence aux critères pertinents.

Le calcul ne nous a pas donné confirmation d'une interaction due à l'âge ou à la formation professionnelle. L'hypothèse nulle n'est donc pas formellement rejetée, ce qui n'est pas une preuve mais seulement une présomption.

Le fait que l'attrance des P.A.P. (pour les coloris vifs) soit en conformité avec le goût général donne à penser que le comportement relèverait davantage de l'inné que de l'acquis, comme sujet à une poussée plus intime, ancrée dans l'intériorité des évaluateurs.

(1) Voir annexe n° 11 c.

5.2. INFLUENCE DE LA VARIABLE COMPLEXITE.

5.2.1.

HYPOTHESE DE TRAVAIL.

Nous avons retenu comme deuxième variable-stimulus la variable complexité, entendue sous l'aspect de la numérosité des éléments. Il s'agit là d'une variable intra-figurale telle que l'a définie R. FRANCES en 1970 (1), c'est-à-dire qui ne met pas en cause directement l'expérience personnelle de l'évaluateur relative à d'autres objets du même type. Cela signifie que si un effet se faisait sentir, celui-ci aurait de grandes chances d'apparaître plus important dans le cas où il se serait agi d'un indice de complexité portant sur une variable intra-sérielle.

Le choix de la variable complexité comme indicateur de l'interférence de variables-stimuli dans l'évaluation des objets plastiques est motivé par ailleurs par le fait que nous disposons, pour ce facteur, d'une base d'informations théoriques particulièrement riche. De plus, la totalité des observations recueillies par les chercheurs converge pour attester de son importance au niveau de l'acte perceptif. Il était donc difficile de ne pas entreprendre son étude dans le cadre de l'évaluation des productions plastiques scolaires.

Peut-être est-il utile de développer quelque peu les informations qui ont contribué à bâtir l'hypothèse de travail?

Evoquer la question de la complexité implique, ainsi

(1) FRANCES (R.), "Intérêt et préférence pour les stimuli de complexité variable", Journal de Psychologie, 1970, 2, pages 207 à 224.

que nous l'avons déjà signalé, de se référer à deux auteurs principaux, BERLYNE et FRANCES. La théorie de BERLYNE traite essentiellement de la variation de l'intérêt et de l'agrément en fonction de la complexité du stimulus:

- " L'intérêt commence à naître pour des stimuli ayant des valeurs d'éveil
- " plus élevées que ceux auxquels l'agrément est attaché et il continue
- " à croître pour des valeurs auxquelles l'agrément n'est plus obtenu " (1).

Autrement dit, la complexité d'un objet stimule notre curiosité, mais notre préférence serait toutefois liée à la quantité de complexité que notre entendement serait capable de tolérer. On obtiendrait ainsi, en fonction de la complexité de l'information, deux courbes au démarrage assez voisines:



Cependant, les expériences de BERLYNE étaient menées d'une manière globale sans que soient distinguées les catégories de regardeurs. FRANCES a repris et approfondi ces expériences, d'une part en introduisant une différenciation au niveau de l'âge et du degré d'instruction des sujets, d'autre part en affinant la notion de variable collative.

Pour ce qui a trait au premier point, FRANCES a pu montrer (2) que l'intérêt semble bien lié chez tous les sujets aux degrés de la complexité. Par contre, si l'attraction cognitive est uniforme, la préférence des ouvriers porte sur des stimuli plutôt simples tandis que celle des étudiants irait à des stimulations plus complexes.

FRANCES a entrepris l'approfondissement de la notion de variable collative. En effet, on pouvait reprocher aux recherches antérieures soit de simplifier exagérément les variables (recherches portant sur une seule dimension isolée artificiellement), soit de porter sur des stimuli trop complexes (préférence pour des tableaux). L'auteur montre (3) que les variables collatives sont définies par une mise en relation (collationnement)

- soit de différentes parties co-présentes dans le stimulus (régularité, équilibre, numérosité, homogénéité);
- soit de caractères du stimulus comparés à des caractères similaires

(1) BERLYNE (D.E.), *Aesthetics and Psychobiology*, op. cit., pages 213 à 216.

(2) FRANCES (R.), *Intérêt perceptif et préférence esthétique*, Paris, Editions du C.N.R.S. page 121.

(3) Ibid., pages 20 à 22.

rencontrés antérieurement pour d'autres objets (nouveaueté, étrangeté, incongruité...).

Il sépare ainsi information intra-figurale et information intra-sérielle. FRANCES fait ainsi remarquer que la complexité n'est qu'une des variables collatives (dites aussi structurales). Il conviendrait d'en citer d'autres, oscillant entre deux pôles:

familiarité / nouveauté
clair / ambigu
similarité / dissemblance, etc.

Les expériences de FRANCES manipulèrent six variables correspondant à ces différenciations. Elles ont porté à la fois sur l'intérêt et la préférence exprimés pour les mêmes stimuli. Elles ont pu montrer que la préférence est d'abord liée à l'âge pour ce qui est de la scolarité obligatoire, puis au degré d'aculturation, ce qui permet d'introduire la notion d'âge esthétique. Cela recoupe en partie les premières expérimentations menées avec H. VOILLAUME en 1964 (1) et celles d'Y. BERNARD (2) montrant que les personnes cultivées recherchent davantage les incongruités de formes, couleurs, proportions.

Ces recherches sont précieuses à plus d'un titre:

- En premier lieu, elles confirment que
 - " selon les normes culturelles des groupes, à un même degré d'éveil
 - " correspondent des valeurs hédoniques très différentes des mêmes stimuli"
- et démontrent " l'interaction du biologique et du social " (3).

Ce premier point inviterait à séparer dans les protocoles expérimentaux les réponses recueillies en fonction des types de formation reçues par les professeurs. Peut-on toutefois parler de différence de niveau universitaire? Il y aura lieu néanmoins de s'interroger sur l'effet de l'écart dû à la formation des P.A.P..

- Les recherches de FRANCES mettent aussi en lumière le double effet des variables collatives: cognitif et hédonique. Le jugement de l'enseignant se trouve-t-il principalement déterminé par rapport à l'intérêt cognitif au terme d'une lecture de l'objet exclusivement analytique et référée à une série de critères précisément répertoriés ou se trouve-t-il dépendant de considérations hédoniques

(1) FRANCES (R.) et VOILLAUME (H.), "Une composante du jugement pictural: La fidélité de représentation", Psychologie française, 1964, 9, pages 241 à 256.

(2) BERNARD (Y.), psychologie du goût en matière de peinture, op. cit., page 130.

(3) FRANCES (R.), Intérêt perceptif et préférence esthétique, op. cit., page 121.

dont on sait déjà qu'elles varient d'un individu à l'autre?

Cette interrogation invite à rechercher ce qui pourrait produire, au titre de la complexité, cet agrément. Les travaux d'A. MOLES sauraient peut-être apporter le complément que nous attendons:

Dans sa théorie de l'information (1), A. MOLES fait remarquer que la complexité n'est pas à relier uniquement avec la numérosité des stimuli ou, pour ce qui concerne les objets visuels, des éléments à appréhender:

" La valeur d'un message est d'autant plus grande que le message est susceptible d'apporter plus de modifications à ce comportement, c'est-à-dire, non qu'il est plus long, mais qu'il est plus nouveau, puisqu'il est déjà connu et intégré par le récepteur et appartient à son système intérieur ".

Dès l'instant où les conditions sont remplies au niveau du stimulus pour qu'il y ait perception (seuils de sensibilité, de saturation entre lesquels doit se situer le stimulus, quantification différentielle de l'importance de la sensation), l'information est liée à l'originalité. La quantité d'information se mesure à son imprévisibilité.

Pour ce qui concerne, ensuite, le mécanisme de la perception, A. MOLES fait observer que toute lecture d'un objet est sélective, ce qui limite aussi le caractère de la numérosité: Il est rare que l'oeil passe en revue toute l'information, sauf dans le cas où celle-ci est trop mince. En cas de surabondance, il y a prise en compte de certains indices seulement. MOLES affirme que

" Le but ontologique poursuivi par l'oeuvre d'art est précisément d'apporter "un peu trop" d'information, un peu trop d'originalité, c'est ce qu'on appelle la richesse perceptive de l'oeuvre d'art, mais ce trop doit être en quantité modérée " (2).

Inversement, dans les cas où il y a "insuffisance", l'oeil infléchit spontanément son attention vers un autre niveau où il trouvera meilleure satisfaction (3).

Toutefois, on peut penser que ce qui est valable dans le cas de l'examen d'une oeuvre isolée fonctionne, en situation d'évaluation comparative, au détriment du producteur d'une oeuvre trop pauvre: L'oeil aura tendance, en ce cas, à quitter le travail pour un autre. Il reste que les travaux de MOLES mettent, eux aussi, en lumière le rôle important joué par la complexité dans l'appréhension des oeuvres plastiques, tant au niveau de l'intérêt porté (il y a

(1) MOLES (A.) Théorie de l'information, op. cit., page 37.

(2) Ibid., page 240.

(3) Ibid., page 242.

recherche avide d'information de la part des terminaux rétiniens) que du plaisir esthétique qui serait de l'ordre de la découverte mais aussi de la surprise, de l'étonnement.

D'un autre côté, malgré nombre d'expériences qui font autorité en la matière, la complexité considérée sous l'angle de la numérosité des éléments et non point, selon l'analyse de MOLES, comme quantité d'information nouvelle n'est pas à l'origine de résultats totalement homogènes.

Si de nombreux chercheurs, ATTNEAVE (1957), VITZ (1966), DAY (1968), KIRLAND (1974), SMETS (1975), cités par FRANCES (1), livrent des résultats similaires à ceux de BERLYNE, d'autres chercheurs, tels MUNSINGER et KESSEN en 1964 (2), RAAB en 1975 (3) observent un intérêt égal pour les formes extrêmement simple ou extrêmement complexes. Il est probable que l'originalité des propositions doive être prise en compte, mais il peut s'agir aussi d'un phénomène de mode (cette hypothèse a été mise en avant par N. MEYER dès 1942): EYSENCK, en 1970 (4), après avoir montré l'attirance des jeunes artistes pour les figures simples, invoqua l'influence du "minimal art" en vigueur à cette époque.

Ces divergences ne sauraient mettre en doute le caractère essentiel de la complexité. Une raison supplémentaire de prendre en considération cette variable réside dans le fait qu'elle est aussi présentée indirectement par les spécialistes américains (GUILFORD, TORRANCE) comme indice majeur de créativité au travers des trois facteurs qui la définissent: La quantité (liée à la fluidité), la variété (flexibilité) et l'originalité (5).

La question qu'il y avait à se poser était de savoir si la complexité d'une proposition plastique en arrivait à interférer illégitimement durant la phase d'évaluation dans des situations où

(1) FRANCES (R.), Psychologie de l'art et de l'esthétique, op. cit., pages 47 à 97.

(2) Ibid., page 65.

(3) Ibid., page 78.

(4) EYSENCK (H.J.) & CASTLE (M.), "Training in art as a factor in the determinations of preference judgment for polygons", British Journal of Psychology, 1970, 61, pages 65 à 81.

(5) BEAUDOT (A.), Vers une pédagogie de la créativité, Paris, E.S.F., 1973, page 82.

celle-ci n'était pas désignée comme critère opérant relativement à la tâche considérée.

Nous avons donc recherché:

- Si une augmentation nette de la complexité entendue sous la forme de la numérosité des éléments (croyant à une influence plus marquée au cas où il s'agirait de différences intra-sérielles) entraînait une variation de la note attribuée aux oeuvres.

- Si une différenciation plus fine des réactions des professeurs évaluateurs pouvait s'opérer en fonction de leur caractéristiques.

5.2.2.

DEROULEMENT DE L' EXPERIENCE.

Comme pour l'expérience précédente, le problème pour l'expérimentateur consistait à trouver un type de production plastique pour lequel la complexité de formulation n'était pas un facteur à prendre directement en compte au niveau des critères d'évaluation.

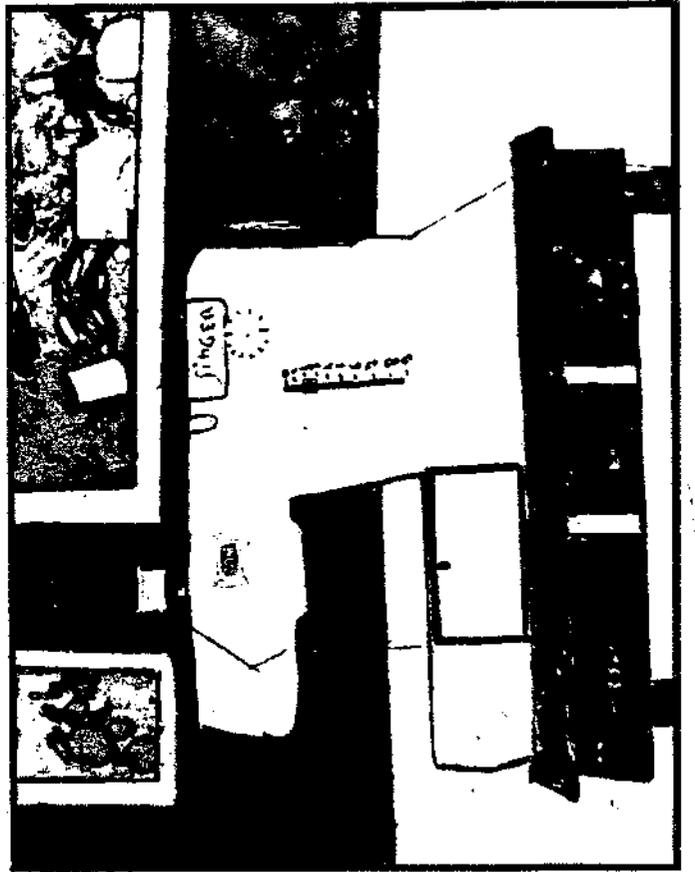
Nous avons opté pour un matériel constitué par des travaux de 5° intitulés "tracteur + machine à coudre".

Le travail était consécutif à une première réflexion concernant le rendu de l'espace au moyen de la perspective (celle-ci n'avait pas fait l'objet d'une approche mathématique mais n'avait été abordée qu'au moyen de constats simples). La tâche se présentait comme la réponse à une induction créatrice qui demandait d'imaginer un environnement respectant les lois élémentaires de la perspective qui permettrait de légitimer la grandeur relative de deux objets dessinés préalablement sur une feuille blanche sans indication de destination.

La numérosité des éléments convoqués intervenait peu dans un travail où étaient principalement à prendre en compte l'ingéniosité du dispositif permettant de réunir dans un même espace perspectif la représentation des deux objets et, d'autre part, le raisonnement d'ordre logique ayant présidé à la mise en place d'un espace visuellement cohérent.

Ici encore, nous étions en présence d'un travail très "scolaire", largement pratiqué sous des formes voisines par de nombreux P.A.P.. Les deux dimensions à apprécier étaient particulièrement bien définies et d'une identification courante.

Le dispositif mis en place était en tous points analogue à celui utilisé pour la variable coloris: Nous avons eu recours à des dessins construits. Il a été entrepris de reproduire plusieurs



dessins d'élèves de manière à disposer pour chaque travail d'un exemplaire dépouillé et d'un exemplaire complexe.

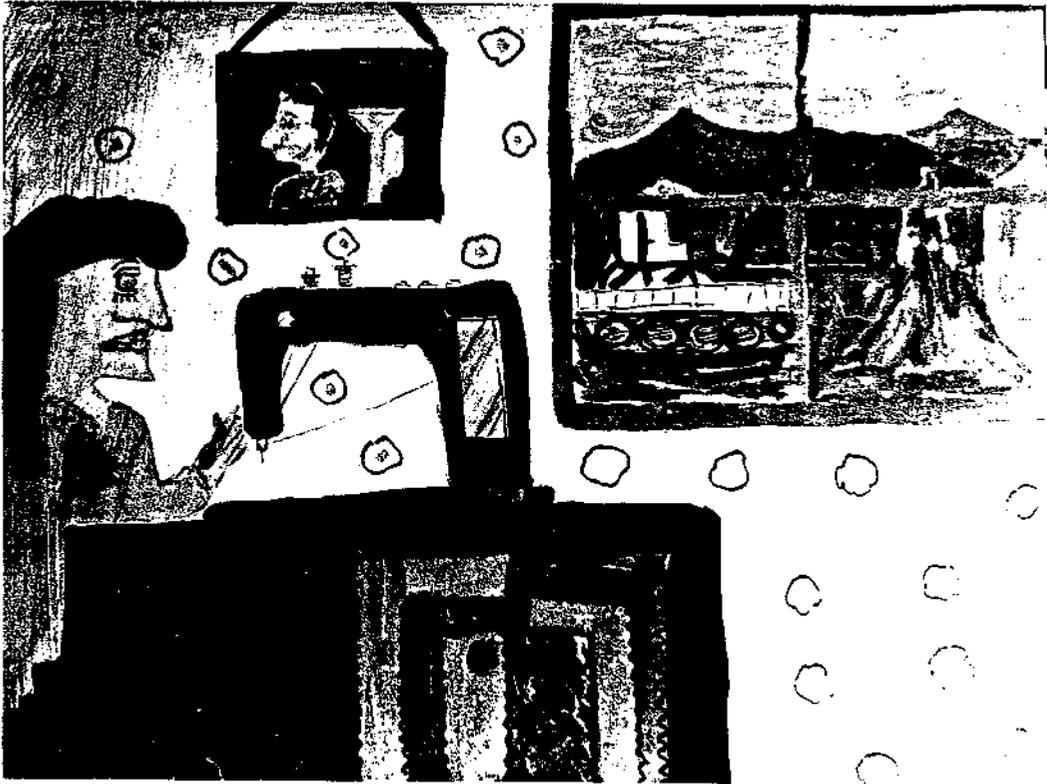
Cette procédure peut ici apparaître critiquable et nous avons touché, là, une des difficultés majeures qui grèvent la manipulation des objets plastiques: On ne peut, dans notre domaine, recourir à des copies reproduites de manière anonyme par le moyen d'une machine à écrire. Afin de minimiser les problèmes engendrés par cette reproduction, nous avons pris soin de choisir des travaux à l'écriture simple (sans effet de matières ni notes de styles trop particuliers). A cet égard, la peinture d'enfants de douze ans, essentiellement faite d'aplats colorés uniformément, a facilité notre tâche. Veillant à une copie fidèle, nous nous sommes, de fait, bornés à supprimer quelques figures répétitives dans les dessins les plus chargés que nous avons à notre disposition.

Comme précédemment, nous avons proposé successivement à l'évaluation de 2 groupes de professeurs 2 lots de 12 travaux. A l'occasion du croisement des lots, 3 dessins-tests ont été remplacés par leurs homologues plus complexes. Les notes collationnées proviennent de l'évaluation de 24 professeurs de différents statuts des académies de Lyon et Grenoble.

Le protocole choisi comportait:

- Deux facteurs aléatoires: Les professeurs évaluateurs répartis en 2 groupes, les 2 panneaux de 12 travaux.
- Le facteur systématique était constitué par les 2 modalités de la variable complexité (simple - complexe: S/C) affectant 2 couples de 3 dessins.
- La variable dépendante était la moyenne des notes sur 20 attribuées par chaque évaluateur à chaque dessin-test suivant chacune des deux occurrences.

Version "simple":



Version "complexe":



5.2.3.

RESULTATS DE L' EXPERIENCE.

1) Influence globale de la variable complexité.

Pour mettre en évidence l'influence de la variable complexité, nous avons mis en parallèle les mesures obtenues par les 6 travaux sous leur apparence simple et par leurs 6 homologues plus complexes:

	simple	complexe	différence
GR (12)	51,50	50,25	-1,25
	46,83	49,75	2,92
	52,25	55,83	3,58
GR' (12)	53,83	58,08	4,25
	45,75	47,75	2,00
	42,83	52,25	9,42

Moyennes calculées à partir de $12 \times 6 = 72$ notes:

$$\underline{M_s} = 48,83 \quad \underline{M_c} = 52,31 \quad \underline{\Delta} = 3,48$$

Cette différence est significative, ainsi qu'en atteste le calcul du $|t|$ de STUDENT (1):

$$\underline{|t|} = 9,66, \text{ très significatif à } .01 \text{ (pour 71 D.L.)}.$$

Comment interpréter ce résultat qui confirme nos hypothèse?

- Un constat global peut être formulé préalablement:
Le P.A.P. prend en compte dans son évaluation des critères ou, du moins, des informations, qui peuvent ne pas être explicitement requis par les données de la tâche à évaluer.

(1) Voir l'ensemble des calculs en annexe n° 12a.

Dans une seconde expérimentation portant sur l'action des variables-stimuli, nous observons que le P.A.P. reste sensible à d'autres paramètres, quel que soit le type de production à considérer et, parmi ceux-ci, à la complexité de formulation.

Comme pour la réaction au coloris, le type de variable (inta-figurale), assez peu lié à l'expérience professionnelle, le type de réaction, qui est conforme au plus grand nombre, tendraient à conforter notre opinion selon laquelle la dépendance vis-à-vis des registres constitutifs de l'oeuvre plastique serait davantage de l'ordre de l'inné que de l'acquis.

Cependant, peut-on, tout aussi aisément que pour la variable coloris, faire reproche aux professeurs de prendre en compte la dimension "complexité" dans leur évaluation d'une performance qui ne s'y réfère pas directement?

De par sa configuration, l'objet visuel est globalité et il est le résultat d'une opération de création. Chacun sait qu'une des dimensions de la créativité est justement la complexité de formulation.

On voit ainsi qu'il apparaît très délicat de vouloir porter un jugement de valeur sur la prise en considération d'un tel critère. Nous nous en garderons bien et nous en tiendrons au constat. Toutefois, si cette "explication" (ou justification) de la prise en compte était opérante, peut-être pourrait-on en avoir confirmation au travers de différences d'appréciation portées par les professeurs selon leur statut?

En ce cas, la sensibilité à la complexité pourrait être en rapport avec la formation professionnelle dans la mesure où une meilleure connaissance des faits artistiques habituerait à considérer la dimension "complexité" comme systématiquement opérante dans l'acte instaurateur. Une plus grande prise en compte de ce facteur pourrait aussi dépendre de l'âge, les professeurs les plus jeunes, encore imprégnés des critères d'évaluation utilisés dans l'enseignement artistique supérieur, pouvant continuer à s'y référer instinctivement.

2) Influence du statut des évaluateurs:

a) Formation professionnelle.

Nous avons repris les notes attribuées dans cette expérience en différenciant les professeurs certifiés et agrégés des personnels

peu formés. Les résultats sont les suivants (1):

<u>Professeurs formés:</u>	<u>Professeurs peu formés:</u>
Moyenne "simple" = 47,75	Moyenne "simple" = 50,79
Moyenne "complexe" = 52,30	Moyenne "complexe" = 51,99
Différence = 4,55	Différence = 1,20

On observe que la différence due à l'influence de la complexité est nettement plus importante chez les personnels formés. Cet écart porte sur 3,35 points standardisés.

Afin de vérifier cette information, nous avons procédé à une analyse de variance, interrogeant le croisement du facteur complexité (CX) et du facteur formation (F). Le tableau se présente ainsi:

	S.C.	D.L.	C.M.	F	
Formation	55,76	1	55,76	2,97	N.S.
Complexité	248,54	1	248,54	13,26	significatif à .01
F x CX	84,50	1	84,50	4,51	significatif à .05
Résidu	2174,21	116	18,74		
Total	2563,01	119			

Nous avons confirmation que le changement de complexité a suscité une différence significative. Si le partage opéré entre les deux catégories de professeurs n'est pas significatif par lui-même, l'interaction entre types de professeurs et degrés de complexité est significative au seuil de .05.

L'interprétation de cette information est, cependant, très délicate. Deux hypothèses semblent pouvoir être avancées, à l'opposé l'une de l'autre:

- D'un côté (interprétation rassurante pour un professeur en titre), on peut se dire que la réalité de l'appréhension des objets plastiques met à mal le principe de l'expérimentation dans la mesure où, comme nous le suggérons, toute création se réfère à la fluidité de l'imagination. Sous cet angle, la complexité de la formulation apparaît une dimension nécessaire, porteuse d'une plus-value qu'il

(1) Voir annexe n° 12b.

serait inconsideré de ne pas vouloir remarquer sous prétexte de s'en tenir à la vérification de l'acquisition d'un savoir ou d'un savoir-faire. Les professeurs formés se montreraient les mieux à même d'avoir le recul nécessaire et de ne pas pratiquer une évaluation "au ras des critères".

Mais inversement, on peut également tirer argument du fait que les professeurs peu formés plastiquement prendraient moins en compte l'indice complexité: Ayant une fréquentation moins directe avec le champ référentiel de l'art, ils ne songeraient pas à mesurer les travaux d'élèves à l'aune des chefs-d'oeuvre. Plus modestement et plus "pédagogiquement", ils s'en tiendraient à la prise en compte des critères correspondant aux objectifs assignés à la tâche.

b) L'âge des évaluateurs.

S'il ne s'agissait pas seulement d'une détermination de type global, la sensibilité à la complexité pourrait apparaître plus importante chez les jeunes professeurs dont la production universitaire s'est vue évaluée par le moyen, notamment, de ce critère. Nous avons donc cherché à interroger cette nouvelle différenciation (1):

<u>Jeunes:</u>	<u>Anciens:</u>
Ms = 46,77	Ms = 49,92
Mcx = 52,16	Mcx = 52,25
Δ = 5,39	Δ = 2,33

La différence semble effectivement plus marquée chez les jeunes professeurs (écart = 3,06).

Le tableau d'analyse de variance se présente ainsi:

	S.C.	D.L.	C.M.	F	
Complexité	447,38	1	447,38	23,78	significatif à .01
Age	78,90	1	78,90	4,19	significatif à .05
CX x A	69,91	1	69,91	3,71	significatif à .1
Résidu	2182,26	116	18,81		
Total	2778,45	119			

L'effet de la complexité est confirmé; la variation due à l'âge a

(1) Voir annexe n° 12c.

aussi suscité une différence de notation significative. L'interaction est, elle aussi, significative, mais à un faible taux (0,1): Toute interprétation trop franche à cet égard doit donc être écartée.

Comme pour la sous-question précédente, si l'on admet néanmoins une plus grande sensibilité des jeunes, nous sommes en présence de deux interprétations possibles: L'une, positive, qui conclurait à une meilleure prise en compte de la réalité d'une oeuvre plastique. L'autre, négative, expliquerait cette dépendance par un manque d'expérience professionnelle.

En conclusion, l'évaluation du P.A.P. est également dépendante d'une autre variable-stimulus, la complexité. Cette dépendance pourrait être en interaction avec l'âge et la formation des professeurs.

Cette observation viendrait confirmer l'expérience précédente portant sur le coloris et, plus globalement, irait dans le sens de l'hypothèse générale de cette deuxième série d'expérimentations, à savoir la persistance de l'influence des catégories critérielles au travers desquelles est perçu tout objet plastique lorsqu'il s'agit d'exercices scolaires visant principalement à la maîtrise d'objectifs préalablement déterminés.

Ceci nous permet-il d'avancer l'hypothèse d'une extension des observations recueillies? Il faudrait que d'autres séries d'investigations viennent consolider ces premières constatations. Il est certain que ces informations incitent à poursuivre dans la même direction.

D'ores et déjà, ces deux vagues d'expérimentations offrent la possibilité, par la synthèse des éléments collationnés, de tenter l'élaboration d'un schéma explicatif du processus de l'évaluation des travaux plastiques scolaires.

Ce sont ces réflexions que nous allons maintenant exposer.

5.3. SYNTHÈSE DES RESULTATS: VERS LA PROPOSITION D'UN MODELE EXPLICATIF.

5. 3. 1.

RECUEIL DES INDICES : FACTEUR EVALUATEUR ET REPRESENTATIONS

Les résultats recueillis permettent-ils d'expliquer un peu mieux le mécanisme de l'évaluation des épreuves plastiques? Peut-on dégager un enseignement propre à améliorer les procédures actuelles? C'est ce qu'il y a lieu de se demander maintenant.

La première série d'expérimentations concernait le questionnement des observations déjà mises à jour par la docimologie générale. Ces observations mettaient en cause de manière globale le comportement de l'évaluateur plasticien au niveau de la prise d'indices.

Le recueil des indices est une activité comparative. Il dépend naturellement de la configuration (du contenu) de l'objet visuel à appréhender, mais il est guidé par l'esprit dans lequel l'évaluateur se dispose à relever ses informations.

Il est probable que le phénomène d'attente soit teinté des informations qu'il a pu recueillir préalablement. J.J. BONNIOL (1) fait état à cet égard d'un phénomène de "facilitation perceptive" produit par tout ce qui peut aider à la recherche des indices, qu'il s'agisse d'informations préalables ou de points de repère immédiats offerts par des travaux-ancres. En A.P., davantage qu'en d'autres matières, l'évaluation repose sur un comportement perceptif, nous

(1) BONNIOL (J.J.), Déterminants et mécanismes des comportements d'évaluation d'épreuves scolaires, op. cit., pages 445-446.

Le savons: " L'image n'est rien d'autre que la lecture qui en est faite, l'image
 " n'est plus image d'un objet, mais image du travail de production
 " de l'image " (1).

Chaque lecture recrée à chaque instant l'image en un événement fugace, jamais identique, directement dépendant de la disponibilité du sujet regardeur et du système de représentations qu'il s'est forgé. Les expériences menées offrent une confirmation de ce que nous avançons dans nos pages introductives quant à la fragilité de l'évaluateur:

* Evaluateur qui reste victime d'effets perceptifs malgré les possibilités de comparaison offertes par la simultanéité de vision, si l'on en juge par les effets de contraste que nous avons mis en lumière. Effets qui pourraient peut-être se maîtriser puisque ceux-ci semblent liés en partie à l'expérience antérieure du sujet, les professeurs confirmés y étant moins sensibles.

* Evaluateur qui semble ne pouvoir faire abstraction de ses inclinations personnelles (qu'elles soient innées ou acquises), si l'on en juge par l'influence sur le jugement de variables telles que le coloris ou la complexité.

* Evaluateur qui se livre à une exploration polarisée idéologiquement, liée à l'idée qu'il se fait de sa mission d'enseignement (objectifs de maîtrise), de son rôle éducatif (encourager, sanctionner) eu égard à un individu dont il connaît déjà, par ailleurs, des informations extérieures portant sur son travail ou sa personne: Nous avons montré que le P.A.P. était sensible aux performances antérieures de l'élève en A.P. et que, s'il accorde dans le cadre de sa notation davantage de poids à ce qui concerne sa discipline, il l'articule très précisément aux performances scolaires générales en sanctionnant le "bon élève" qui ne le serait pas chez lui.

Le "facteur enseignant" est donc très important et il y a tout lieu de penser que le caractère concret, l'"objéité" des produits à évaluer, la possibilité de soupeser ceux-ci "hic et nunc" (dimensions que l'on opposerait aux caractéristiques "abstraites"

(1) SCHEFER (J.L.), "L'image: Le sens investi", in Communications, 15, 1970, pages 210 à 221.

des devoirs écrits) ne suffit pas à limiter l'exploration polarisée et la dépendance vis-à-vis des indices les plus prégnants: Effets d'ancrage et influence d'indices ayant une prise immédiate sur les représentations.

Le professeur est bien dépendant d'un " système organisateur normatif qui relève du caractère social de ses représentations " et où entrent proiritairement les facteurs aboutissant à une " forte désirabilité scolaire " (1) fondée sur l'attention, la persévérance, la participation, la confiance inspirée, l'intérêt affiché, la maturité...

Alors, la représentation que le professeur se fait du "bon comportement" d'un élève de collège en A.P. ou d'un élève de lycée en section artistique ne pourrait-elle pas, dans ces conditions, être déterminée par des considérations extérieures du même ordre que celles déjà relevées dans le cadre général comme la beauté ou la prestance physique (2) ? Et l'idée du "bon plasticien" ne pourrait-elle pas, de la même manière, être renforcée par des considérations proches de celles de C. CHASE dont nous avons déjà fait état et portant sur la qualité de l'écriture (ou la couleur de l'encre) ?

Prenons quelques exemples d'explorations polarisées (correspondant, elles-mêmes, à une polarisation des attentes) qui peuvent s'instaurer en A.P. aux dépens de la référence aux catégories critérielles directement opératoires:

- Le P.A.P. de collège n'aura-t-il pas tendance à sanctionner une expression plastique bâclée, approximative, dans une classe où les élèves sont agités et peu travailleurs alors que cette même observation dans le travail d'un élève sérieux et discipliné aurait pu être mise au compte d'une gestualité expressive?

- Au niveau du Baccalauréat, pouvons-nous affirmer que nous abordons avec le même esprit l'évaluation des travaux selon qu'ils proviennent des séries A, C ou G, par exemple?

- Ne sommes-nous pas sensibles à la générosité des moyens employés? Ne regrettons-nous pas le recours à des gouaches de médiocre

(1) GILLY (M.), Maître - élève. Rôles institutionnels et représentations, op. cit., pages 75 à 77.

(2) VENTURY (J.C.), Apparence physique / estimation de la réussite scolaire, Mémoire non publié, Université de Provence, U.E.R. de Psychologie, 1973.

qualité aux pigments colorés éteints? N'apprécions-nous pas l'emploi d'un papier de bonne tenue, de matériaux valorisants qui ne reflètent pas nécessairement l'ingéniosité manipulatoire de l'élève mais davantage son aisance financière?

En section artistique A3 ou F12, n'est-on pas sensible au fait qu'un élève ait déjà exposé ou "vendu", alors que le contact avec les institutions spécialisées dépend très largement de la nature du contexte familial?

Il y a donc lieu de nous interroger, en tant qu'évaluateurs, sur les facteurs qui peuvent conditionner notre regard d'enseignant-plasticien:

Notre enfance, l'éducation que nous avons reçue, notre lieu de résidence et le cadre dans lequel nous avons fait nos études, nos attentes quant aux traces de la réussite de notre enseignement, notre connaissance de l'actualité artistique et des polémiques qu'elle suscite, notre implication personnelle en tant que créateur, nos tendances profondes, enfin, en leur oscillation entre classique et baroque, ont peu à peu modelé notre manière de percevoir, d'accepter (ou de refuser) la "DIFFERENCE".

L'expression plastique est autre chose qu'un "jeu scolaire". Ses racines s'ancrent au plus profond de l'individu producteur de même que ses résonances se font entendre au coeur du regardeur. Mieux connaître notre comportement évaluateur c'est, d'une certaine façon, mieux mesurer ce qui nous sépare de nos élèves.

5. 3. 2.

PREGNANCE PERCEPTIVE ET PREFERENCE
ESTHETIQUE : EXTENSION VALIDANTE.

La deuxième série d'expériences posait la question de l'interférence de facteurs de goût dans la recherche des indices nécessaires à l'évaluation d'un travail plastique.

Nous avons apporté la preuve de l'influence de composantes visuelles importantes alors que la dimension considérée ne se trouve pas directement sollicitée par la vérification des objectifs assignés.

Les investigations reposaient sur deux variables: Le coloris et la numérosité. La positivité du test invite désormais à prendre avec davantage de sérieux l'hypothèse d'une extension validante de ces premières observations.

Le professeur, selon son degré de compétence plastique mais - semble-t-il - davantage suivant son inclination personnelle, resterait sous la dépendance de sa relation aux diverses dimensions plastiques de l'oeuvre.

Alors, ne pouvons-nous pas (et l'on mesure ici la modestie du travail accompli eu égard à l'ampleur de ce qu'il reste à envisager) nous interroger sur l'interférence de considérations portant sur les autres registres?

- Réalisme. Selon son ouverture d'esprit à l'art du XX^e siècle, comment le P.A.P. réagira-t-il au traitement de la figuration, à sa mise à mal, voire, en situation d'examen, à une réponse totalement abstraite au thème imposé?

FRANCES et VOILLAUME (1) ont montré l'effet déterminant de

ce critère: Le réalisme est apprécié par le grand public à tel point qu'il déteint sur les autres dimensions tandis que les personnes éduquées culturellement sont davantage sensibles à l'interprétation expressive.

On peut imaginer que dans les grandes classes le réalisme ne sera pas abordé pour lui-même mais corrélativement aux intentions de l'auteur. Mais en collège? Sa maîtrise ne risque-t-elle pas d'être prise en compte abusivement au titre d'indice de maturité?

Les effets de mode ne joueraient-ils pas, selon les tendances de l'actualité artistique? Les P.A.P. peuvent mesurer combien les épreuves de l'agrégation d'arts plastiques ont été marquées, dans les années 70, par l'influence de l'"abstraction analytique" et reconnaître que l'émergence récente de la "figuration libre" a permis que soit porté un autre regard sur les productions actuelles.

- Originalité. Comment sera-t-elle perçue? Pour qui connaît le milieu de l'art, cette interrogation pose en filigrane toute la question du concept d'avant-garde, tellement opérant en cette première moitié de siècle (concept de progrès historique, pénalisation du "déjà vu"), mais qui est actuellement l'objet d'un réexamen critique dont les pamphlets de J. ELLUL (1) et de J. CLAIR (2) ont probablement été les plus retentissants: "Aujourd'hui, être "moderne" ne serait-ce pas justement vouloir sortir de la modernité?" se demandait le premier dès 1980, tandis que le second, relisant les "Journaux intimes" de BAUDELAIRE, assimilait avant-garde à "conformité", se lamentait sur l'agonie de la peinture, la perte du métier et l'avènement du "n'importe quoi".

Sans vouloir nier qu'un certain degré d'originalité soit nécessaire, reste encore son appréciation. La singularité ne peut se mesurer qu'à l'aune d'un champ de référence variant d'un enseignant à l'autre.

- La technique d'expression. Il s'agit en premier lieu de la maîtrise technique et de la fascination plus ou moins grande qu'elle exerce sur l'enseignant. C'est un facteur de désirabilité

(1) ELLUL (J.), L'empire du non-sens, Paris, P.U.F., 1980, page 13.

(2) CLAIR (J.), Considération sur l'état des Beaux-Arts. Critique de la modernité, Paris, Gallimard, 1983, page 69.

notoire et le questionnaire introductif montrait qu'une majorité de professeurs privilégiait la technique par rapport aux intentions.

Mais, par ailleurs, comment l'enseignant gèrera-t-il son attirance ou son indifférence vis-à-vis du mode choisi: Mine de plomb, gouache, aquarelle, finesse miniaturiste ou, tout au contraire, travail de la pâte, incorporation de déchets, etc.?

- Le propos. Ne peut-on admettre, enfin, qu'il en soit de même pour ce qui se rapporte à la nature du sujet, au degré d'expressivité de l'agencement proposé, plus encore aux références subjectives que l'oeuvre peut provoquer chez le regardeur par l'évocation de thèmes, d'événements auxquels ce dernier est particulièrement sensible?

Alors, quelle interprétation peut-on avancer concernant les mécanismes en oeuvre? Prégnance perceptive et préférence esthétique nous paraissent étroitement liées tout en relevant d'origines différentes:

La prégnance perceptive de certaines parties de l'oeuvre peut dépendre uniquement du dispositif plastique mis en place: Le regard est sollicité par l'auteur en direction d'une zone de première importance, à des fins de communication. L'attention peut aussi être captée par l'utilisation exacerbée d'une dimension plastique qui provoque ainsi un effet de "style" (graphisme "électrique", sfumato...). Dans une telle hypothèse, on peut penser que l'évaluateur sera prioritairement impressionné par l'élément perçu de prime abord, d'autant plus profondément que l'on sait que la perception suscite chez le lecteur une réaction pré-consciente attestée par des mesures électrodermales. Cette première impression, de polarité positive ou négative, sera ensuite renforcée par la recherche d'une consonance cognitive au travers d'indices de confirmation: La préférence esthétique est donc amenée à se manifester relativement au caractère des indices les plus prégnants.

Mais l'hypothèse inverse peut aussi être avancée. L'exploration serait, dès l'abord, polarisée par les préférences esthétiques de l'évaluateur, ce qui se ramènerait au schéma préalablement proposé concernant les attentes et l'action des représentations. Dans l'oeuvre observée, le regardeur sélectionnerait inconsciemment, en plus des indices directement liés aux critères d'évaluation, d'autres perceptions

auxquelles il resterait, en tout état de cause, sensible, c'est-à-dire celles pour lesquelles s'élaboreraient le plus spontanément des phénomènes d'attraction ou de rejet.

Peut-être serait-il possible de nommer "effet d'appel rétinien" ce qui se rapporte aux dimensions visuelles du stimulus. Nous placerions sous ce terme tout effet d'ancrage perceptif d'origine plastique, que cet ancrage soit plutôt d'ordre "quantitatif" (contraste visuel volontairement déterminé par l'auteur de l'oeuvre) ou d'ordre "qualitatif" (indices dont la qualité suscite chez le regardeur une réponse privilégiée du fait de sa sensibilité particulière à leur égard).

Nous pensons que notre approche de l'effet "coloris" et de l'effet "complexité" rend licite la question d'une extension validante à d'autres dimensions de l'oeuvre qui viendraient agir en co-occurrence. Certes, nous devons rester pour l'instant au niveau des hypothèses et ces interrogations n'ont pour objet que de baliser la piste d'investigations ultérieures.

Néanmoins, l'importance des registres évoqués n'échappera pas. Leur relation directe au goût, à la préférence personnelle non plus, ce qui laisse mal augurer de la possibilité d'en maîtriser totalement les effets.

Ainsi, l'évaluation des productions plastiques peut être présentée comme doublément dépendante de mécanismes perceptifs: Si elle fonctionne globalement d'une manière assez proche du contexte général, elle s'avoue, de surcroît, influencée par des considérations hédoniques spontanées et diversifiées.

5. 3. 3.

L' EVALUATION EN ARTS PLASTIQUES :
UN MECANISME SPECIFIQUE.

Rassemblons nos informations: Qu'il s'agisse ou non d'une production artistique, l'opération d'évaluation consiste à extraire un certain nombre d'INDICES correspondant à des CRITERES pour aboutir à une DECISION.

C'est un comportement PERCEPTIF / COGNITIF, une activité de comparaison. Cependant, l'évaluation des objets plastiques s'avère plus difficile du fait de la présence de particularités aggravantes à chacun de ces deux niveaux:

- L'activité perceptive est tributaire d'erreurs dues à la sensibilité de l'évaluateur. Nous avons montré que le P.A.P. se trouve, en plus, sous la dépendance de phénomènes d'attraction spontanée en directions d'indices visuels correspondant à son goût naturel.

- L'activité cognitive est tributaire d'erreurs dues à l'appréciation comparative. Mais, dans le domaine des activités plastiques, les modèles de référence correspondant à un produit précisément attendu n'existent pas. Ou, inversement, nous pouvons dire que les "modèles" auxquels se réfèrent les professeurs sont infinis, multiformes, relevant d'idéologies plastiques parfois totalement contradictoires, à l'image de la diversité des démarches de l'art du XX^e siècle. L'activité cognitive s'exerce ici sur des bases instables.

La contribution de la phase expérimentale permet alors de proposer un tableau du mécanisme de l'évaluation des travaux plastiques mettant en avant les "moments" (ou stations) soumis à la subjectivité.

Le processus d'évaluation des productions plastiques

est, tout comme ailleurs, le lieu d'interactions sociales complexes. Il est, de plus, tributaire d'un système de représentations dépendant de modèles pédagogiques mouvants et surtout de normes esthétiques extrêmement diversifiées. Mais le recueil des indices n'est pas le seul lieu générateur de biais d'évaluation: Ce sont tous les niveaux du processus éducatif qui sont soumis à la subjectivité et dépendants de la relation maître-élève (1) :

1) Le niveau général des objectifs visés.

L'évaluation est utile relativement à certains besoins. Prolongeant les instructions officielles, le P.A.P. aura à préciser les besoins de l'élève en matière de compétence artistique et choisira en fonction de l'idée qu'il se fait de sa mission. Il choisira aussi en fonction de l'idée qu'il se fait du statut et du devenir social de ses élèves.

2) Le niveau pédagogique du travail proposé.

L'enseignant aura à construire une tâche appropriée pour mettre en évidence les indices utiles à son jugement, une tâche à la fois éducative (imaginer - interroger - manipuler) et révélatrice, selon lui, de comportements essentiels.

3) Le niveau du climat relationnel de la classe ou du lieu de production.

Le comportement du professeur sera-t-il apte à placer ses élèves dans des conditions matérielles et psychologiques telles que leurs productions soient réellement authentiques?

Que penser de la situation d'examen?

Peut-on parler d'égalité de traitement d'un élève à l'autre, d'une classe à l'autre, d'un établissement à l'autre?

Sur le versant producteurs, tous les élèves se trouvent-ils égaux face à la tâche? Selon les origines sociales ou géographiques, ne sont-ils pas différemment armés sur le plan des ressources documentaires ou culturelles, plus ou moins capables de répondre aux attentes du professeur, volubiles ou "bloqués"?

1) L'ensemble de ce chapitre a bénéficié pour son élaboration de deux apports théoriques importants:

- Le travail de P. PERRENOUD, "L'évaluation scolaire est-elle révélatrice ou génératrice de l'inégalité sociale devant l'école?", Genève, Service de la recherche sociologique, 1977, in L'évaluation formative dans un enseignement différencié, op. cit., pages 37 à 44.

- Le schéma de l'évaluation élaboré par J.J. BONNIOL (1981), op. cit., page 100.

4) Le niveau du recueil des indices.

Le professeur aura-t-il les moyens de dominer les phénomènes d'interférence perceptive? Outre les effets d'ordre général, face à un objet qui est globalité, ne va-t-il pas "tout" percevoir, en bloc, confusément? Ne glanera-t-il pas, au contraire, qu'une information filtrée par ses attentes ou ses prédilections? Saura-t-il juguler les effets d'appel rétinien, rejeter des indices qui pourraient être importants si le barème était différent mais qui sont, dans la situation donnée, hors des préoccupations? Comment ne pas être sensible au donné immédiat!

5) Le niveau de l'exigence dans la quête des indices.

De surcroît, saura-t-il doser ses attentes? Dans les petites sections, ne nous laissons-nous pas aller à imaginer un idéal de création nettement au-dessus des capacités correspondant à l'âge des enfants et à ériger en "norme" la production du plasticien confirmé? L'idée que l'on se fait du produit attendu est d'autant mieux circonscrite que l'expérience de l'enseignant lui a appris ce qu'il est possible d'obtenir dans telle ou telle circonstance.

Les P.A.P. déclarent souvent imaginer un "produit moyen", sorte de seuil minimal à atteindre, attestant que la proposition a été comprise et la consigne exécutée. Ils considèrent aussi que ce minimum est à la fois indispensable et totalement inintéressant: Il n'y a réellement oeuvre de création que dans la mesure où les consignes de base ont été sublimées, où l'élève a pu s'élever au-delà des contraintes dans une sorte d'élan spontané, a réussi à conférer à son travail une "plus-value" personnelle.

6) Le niveau de la transposition en note chiffrée.

Par delà la question de l'étendue et de la centration de l'échelle de mesure, on peut aussi s'interroger sur l'attribution des points.

La docimologie nous enseigne (1) que, dans le cas d'un modèle à atteindre ou à reproduire, l'évaluateur opère par soustraction au fil des erreurs rencontrées. Dans une tâche de construction, l'on procéderait plutôt par addition en fonction des difficultés surmontées.

Cette question revient à savoir si, en A.P., l'intensité

1) NOIZET (G.) et CAVERNI (J.P.), Psychologie de l'évaluation scolaire, op. cit., pages 145-146.

de la proposition visuelle (force, originalité...) l'emporte sur les maladroites du "rendu" ou si, inversement, une technique de haute virtuosité peut excuser l'indigence de l'idée. Si l'on s'en tient aux résultats du questionnaire dont nous avons fait état en introduction, les qualités techniques seraient relevées prioritairement à l'expression personnelle (36 contre 21) tandis que, parallèlement, l'évaluation relèverait d'un barème critérié plutôt que d'une estimation globale (48 contre 11). Ceci témoignerait du caractère "scolaire" de l'évaluation la plus généralement pratiquée.

S'il est possible de mettre en doute l'effective et généralisée référence à une grille de critères (explicable par le souci de paraître consciencieux), la priorité accordée à la technique semble cependant, elle aussi, contribuer à contredire toute relation affective à l'oeuvre (ce qui confirmerait aussi le caractère souterrain des effets d'appel rétinien) qui pourrait aboutir à mettre une note en fonction du degré d'émotion et de plaisir ressenti à son examen. Il est donc possible d'avancer l'hypothèse que le P.A.P. procéderait plutôt par retrait de points par rapport à l'idéal qu'il s'était fixé, ce qui rendrait la relation aux modèles particulièrement déterminante. Ceci aurait deux implications:

- La première serait que, dans un lot de travaux, serait placé en tête celui qui s'approcherait le plus de l'image mentale du "bon travail" intériorisé par l'enseignant.
- La seconde, dont il y aura lieu de peser toutes les conséquences, serait que la référence à ce butoir s'effectuerait par rétrogradation, ce qui signifierait que la notation en A.P. correspondrait, de fait, à l'ORDINATION des travaux à l'intérieur d'un groupe déterminé, ordination fondée ou non sur la pesée de l'oeuvre suivant différents critères mais dont la transposition en note chiffrée n'aurait aucune valeur d'absolu et serait susceptible d'une remise en question permanente dès la confrontation à une autre série de travaux, a fortiori, d'ancres.

7) Le niveau de l'interprétation des observations recueillies.

Enfin, reste la question de l'interprétation. Ici encore s'exerce la pression des modèles auxquels se réfère l'enseignant:

En fin de 3°, le "niveau" atteint permet-il d'encourager l'orientation vers une section artistique; en "C.A.P.E.S.", la production à laquelle le jury est confronté est-elle "digne" d'un futur enseignant plasticien?

La question est ici double: Le professeur saura-t-il

exploiter les indices recueillis, seront-ils utilisés efficacement tant pour conduire son enseignement que pour l'avenir des enseignés? Mais l'on doit aussi s'interroger sur la valeur prédictive des indices: Les constatations auront-elles une réelle utilité, le produit soumis à évaluation est-il, de par sa nature et les conditions de sa fabrication, porteur d'informations authentiques?

Probablement avons-nous oublié quelques lieux et causes d'exercice de la subjectivité. Nous en proposerons cependant le tableau récapitulatif ci-après:

NIVEAUX D'EXERCICE DE LA SUBJECTIVITE. (s).

- 1) Définition de la tâche. Constitution du produit-norme.
 - * Définition des objectifs:
 - s = L'idée que l'on se fait du rôle de l'éducation plastique.
 - * Mise en place de la tâche:
 - s = Représentation de la capacité de ses élèves.
Expérience professionnelle quant aux produits possibles.
 - * Choix des critères d'évaluation; produit-norme.
 - s = Ce que le professeur juge prioritaire (technique ou expression).
- 2) Le dosage des exigences: Le produit attendu.
 - * Conditions de la production et hauteur des objectifs assignés:
 - s = Perméabilité du jugement aux informations a priori / producteurs.
- 3) Codage du modèle de référence.
 - * Image d'un modèle de référence:
 - s = Errance du M.R. accentuée par le niveau variable des travaux reçus.
 - * Echelle de mesure:
 - s = Différence de sévérité.
Différence dans la hiérarchisation des critères opérationnels.
- 4) Lecture ou perception des indices.
 - * Les attentes:
 - s = Influence des informations a priori et effets de halo.
 - * L'activité perceptive:
 - s = Effets de contiguïté; effets d'appel rétinien.
 - * L'activité cognitive:
 - s = Niveau du champ plastique et nature des idéologies plastiques de référence. Préférences.
- 5) Prise de décision.
 - * Argumentation:
 - s = Avoir su repérer les indices pertinents.
 - * Interprétation, exploitation en vue d'une décision d'ordre éducatif:
 - s = La capacité à interpréter les indices est liée au degré prédictif de ces indices ainsi qu'à l'aptitude de la tâche à les mettre en avant.

Ainsi, le mécanisme de l'évaluation des productions plastiques, sans différer considérablement du schéma général, possède-t-il sa spécificité.

Cependant, cette recherche d'une meilleure intelligibilité des lieux d'exercice de la subjectivité n'est pas sans ouvrir la porte à de nouvelles interrogations:

- En premier lieu, dans ces conditions, peut-on espérer harmoniser les jugements en arts plastiques?

Une meilleure maîtrise des critères est-elle en mesure d'estomper la subjectivité?

La prise en compte du classement plutôt que de la note serait-elle la solution spécifique à une partie des écarts constatés?

- Mais, en second lieu, une harmonisation des jugements ne peut-elle apparaître illusoire, factice, voire contre nature? Par définition même de l'activité "expression plastique", le P.A.P. ne peut se limiter à considérer la réussite atteinte lorsqu'il y a maîtrise des apprentissages techniques ni même transfert (ou le ferait-il - le fait-il encore - qu'il se situerait délibérément dans une option pédagogique, voire politique, qui tend à s'amenuiser peu à peu) :

Le niveau attendu est celui de l'expression personnelle, c'est-à-dire justement la production d'un travail inattendu, manifestant une certaine autonomie (telle est, quant à nous, notre vision des choses), travail qui devrait dérouter, surprendre. Même si les codes employés sont connus du P.A.P., on peut prévoir que cette surprise sera perçue comme "plutôt agréable" ou "plutôt désagréable" compte tenu de la réceptivité de l'enseignant, de ses attitudes face à l'innovation, de ses prédilections.

Admettre ceci, c'est présenter l'évaluation comme lieu de friction entre deux registres de normes esthétiques: Celles des élèves, dépendantes des facteurs socio-culturels qui les entourent, celles du professeur, vectorisées par les idéologies auxquelles il se réfère. Le phénomène évaluation s'élabore à la rencontre de ces deux flux. On en vient donc tout naturellement à se demander s'il

n'est pas normal (sinon souhaitable) que les caractéristiques d'une expression donnée "parlent" mieux à certains évaluateurs qu'à d'autres.

Cette dernière observation ne peut manquer d'être présente au moment d'envisager les conditions d'une meilleure évaluation.