



# REVUE DE MODIFICATION DU COMPORTEMENT

## SOMMAIRE

- Réduction du comportement "d'intervenir inutilement" en classe par l'utilisation d'un système d'auto-régulation de groupe 63  
*Dolorès L'HEUREUX-FOURNIER  
Frank VITARO  
Jacques CHAREST*
- Le comportement des élèves en classe: jugement des professeurs vs observation systématique 73  
*Michel LORANGER  
Richard CLOUTIER*
- Critique de livre 87  
*Claude BLAQUIERE*

REVUE DE MODIFICATION  
DU  
COMPORTEMENT

Revue trimestrielle publiée par l'A.S.M.C. Inc.

*Directeur en poste:*

André Soulières  
Service de Psychologie  
Hôpital Rivière-des-Prairies  
7070 boul. Perras  
Montréal, Qué.  
H1E 1A4

*Directeur nommé:*

Maurice Harvey  
A.S.M.C.  
6955 boul. Taschereau, suite 211  
Brossard, Qué.  
J4Z 1A7

*Responsable de la rubrique  
"Critique de livres":*

Pierre Baron  
Département de Psychologie  
Université de Moncton  
Moncton, Nouveau-Brunswick

*Conseil  
D'Administration  
De l'Association  
Des Spécialistes  
En Modification  
Du Comportement  
A.S.M.C. Inc.*

Président: Rodrigue Otis  
Président sortant: Léonard Goguen  
Secrétaire: JINETTE MARCIL-DENAUT  
Trésorier: Gilbert Leroux  
Représentant (Montréal): Henri Martin-Laval  
Représentant (Québec): Robert Ladouceur  
Représentant (Moncton): Pierre Baron  
Représentant (Sherbrooke): Rodrigue Otis  
Représentant (Nord-Ouest): Raymond Beausoleil

Port de retour garanti par A.S.M.C., 6955 boul. Taschereau, suite 211, Brossard,  
J4Z 1A7

Dépôt légal: Bibliothèque nationale du Québec et Bibliothèque nationale du Canada  
I S S N 0 3 8 3 - 0 5 6 X

Réduction du comportement "d'intervenir inutilement" en classe  
par l'utilisation d'un système d'auto-régulation de groupe

DOLORES L'HEUREUX-FOURNIER

*Commission scolaire La Vérendrye*

FRANK VITARO (1)

et

JACQUES CHAREST

*Centre d'Etudes universitaires dans l'Ouest québécois*

RESUME

*Neuf adolescents, âgés de quinze à dix-sept ans et fréquentant une classe d'initiation au travail, ont l'habitude "d'intervenir inutilement" durant le cours de français. Une procédure d'auto-régulation dans le contexte d'un programme de jetons déjà existant dans la classe, permet aux sujets de participer activement à la réduction du comportement indésirable chez tous les membres du groupe. Le co-aménagement des contingences, à l'intérieur d'un schème à changement de critères, de même que la fixation des seuils de performance par les sujets ont aidé ces derniers à atteindre une performance très acceptable. Des effets de généralisation sont notés dans d'autres situations.*

Plusieurs études ont démontré que le niveau de comportements de dissipation en classe peut être réduit par un aménagement systématique des contingences externes de renforcement (Allen, Hart, Buell, Harris et Wolf, 1964; Homme, DeBaca, Devine, Steinhorst et Rickert, 1963; O'Leary, Becker, Evans et Saundargas, 1969; Patterson, 1965; Schmidt et Ulrich, 1969; Thomas, Becker et Armstrong, 1968; Wasik, Senn, Welch et Cooper, 1969). Cependant, certains auteurs (Drabman, Spitalnik et O'Leary, 1973; Kazdin, 1975) constatent que les comportements académiques et sociaux soumis à l'évaluation d'intervenants adultes et acquis par aménagement externe des contingences de renforcement se limitent aux situations d'intervention et ne se maintiennent que peu de temps après le retrait du programme d'intervention.

1. Les demandes de tiré-à-part doivent être adressées à Frank Vitaro.

Devant ces résultats, certains auteurs ont exploré les systèmes d'auto-régulation pour modifier certains comportements indésirables en classe. Glynn (1970) indique que le renforcement déterminé par le sujet est aussi efficace que le renforcement déterminé par l'expérimentateur pour améliorer les performances académiques et diminuer les comportements d'indiscipline. Bandura et Perloff (1967) obtiennent des résultats similaires dans le cas de l'acquisition et du maintien de comportements moteurs. Les résultats de Lovitt et Curtiss (1969) vont également dans le même sens que ceux de Glynn (1970). Johnson (1970) et Johnson et Martin (1973) constatent que les comportements acquis par auto-régulation semblent plus résistants à l'extinction que ceux soumis à des contingences externes de renforcement. Bolstad et Johnson (1972) rapportent qu'une procédure d'auto-évaluation et d'auto-aménagement des contingences est aussi efficace qu'une procédure de renforcement externe pour réduire les comportements de dissipation en classe et plus efficace que celle-ci pour assurer le maintien de tels changements. Les résultats obtenus par Drabman, *et al.* (1973) laissent supposer qu'un système d'auto-évaluation et d'auto-renforcement favoriserait une diminution du nombre de comportements de dissipation tout en permettant une généralisation de ces changements à des périodes de la journée où aucune intervention n'est en application. Afin d'apprendre aux individus à contrôler leur environnement et devant les problèmes d'ordre éthique que peut susciter toute procédure visant à modifier le comportement d'une personne et devant la nécessité de prévoir non seulement l'installation d'un changement de comportement mais également son maintien et sa généralisation (Baer, Wolf et Risley, 1968; O'Leary et Drabman, 1971), les systèmes d'auto-régulation méritent d'être pris en considération et leurs effets, d'être systématiquement évalués.

Un autre aspect important d'un programme d'intervention en milieu scolaire est l'économie de temps et d'énergie pour celui qui administre le programme, en l'occurrence le professeur. Un programme d'intervention qui utilise un système d'auto-observation et d'auto-renforcement de groupe semble pouvoir assurer des résultats positifs sans mobiliser de façon indue le professeur. Plusieurs études ont démontré l'efficacité et l'économie engendrées par l'utilisation d'une contingence de groupe (le comportement de tout un groupe est soumis au renforcement) pour modifier des comportements d'indiscipline en classe (Ritchl, Mongrella et Presbie, 1972; Schmidt et Ulrich, 1969; Sulzbacher et Houser, 1968) et améliorer les performances d'ordre académique (Hamblin, Hathaway et Wodorski, 1971; Mc Carty, Griffin, Apolloni et Shores, 1977). Peu d'études ont cependant associé un système de contingence de groupe à une procédure d'auto-régulation.

Le but de la présente étude consiste à déterminer si la combinaison d'un système de co-aménagement en groupe des contingences de renforcement et d'une procédure d'auto-détermination des seuils collectifs de performance par les membres du groupe peut amener ces derniers à diminuer la fréquence de leurs "interventions inutiles" durant une activité de cours. Ce travail vise également à favoriser et à évaluer des possibilités de généralisation de ces changements à d'autres situations. Enfin, il y aura vérification du maintien de ces changements suite au retrait de l'intervention.

## Auto-régulation de groupe

### METHODOLOGIE

#### Sujets

Neuf adolescents dont les âges varient entre quinze et dix-sept ans fréquentent une classe d'initiation au travail (IAT). Ces garçons ont l'habitude "d'intervenir inutilement" durant le cours de français. "Intervenir inutilement" est défini opérationnellement de la manière suivante: interpeller un copain - répondre à une interpellation - faire des réflexions personnelles à haute voix - poser des questions inopportunes - riposter au professeur.

#### Procédure

Dans un premier temps, le professeur procède à une observation détaillée du comportement-cible pendant les périodes de français et de mathématiques (phase 1). Il note la fréquence d'apparition des "interventions inutiles" et indique les événements identifiables qui précèdent (antécédants) et accompagnent (conséquences) le comportement-cible. Les antécédants et conséquences peuvent être de nature physique ou sociale. Ils peuvent originer du professeur ou des pairs.

Dans un deuxième temps, il s'agit pour le professeur de mettre en place des stratégies qui favorisent les changements désirés (phase 2). L'objectif de cette phase consiste à réduire à moins de dix le nombre total d'interventions verbales inutiles et à moins de trois le nombre total de réponses d'attention aux interventions verbales inutiles. Le résultat des observations du niveau de base est présenté aux sujets. Tous constatent un minimum de 33 "interventions inutiles" par période de 60 minutes. Une moyenne de six antécédants et de huit conséquences provient des pairs. Pour favoriser la réduction de ces comportements, les sujets sont invités à ne pas porter attention (extinction) aux interventions inutiles de leurs compagnons plus de huit fois par période de 60 minutes (moyenne des conséquences provenant des pairs). Ils sont invités également à ne pas faire "d'intervention inutile" plus de 33 fois au cours d'une même période (performance minimale réalisée durant les sessions de niveau de base). Les sujets optent ensuite pour un système du type "démérite". Au début de chaque période d'intervention, les chiffres 8, 7, ., ., 0 sont inscrits au tableau et il est convenu qu'à chaque fois qu'un étudiant répond à une intervention d'un copain, un chiffre sera biffé. Les interventions inutiles sont notées par le professeur. Tout le groupe a accès au renforçateur lorsqu'il n'y a pas plus de huit réponses d'attention aux interventions inutiles. De plus, tout le groupe ne doit pas dépasser 33 "interventions inutiles" durant trois sessions consécutives. Si la série des trois sessions est rompue, dans un cas ou dans l'autre, une nouvelle série de trois sessions débute. Le renforçateur choisi par le groupe est le tirage d'un billet porteur d'un nombre d'"Edas" (système de jetons déjà existant dans la classe) dont la valeur peut varier entre dix et cinquante "Edas". La monnaie "édas", qui constitue un renforçateur conditionné généralisé, est toujours échangeable pour des activités renforçantes (périodes libres), de la nourriture (jus, chocolat) ou des objets (vêtements, articles de classe). Tout au long des sessions d'intervention, des exigences de plus en plus grandes sont fixées. En ce qui concerne les interventions inutiles, une première exigence consiste à ne pas dépasser le nombre de 33 pour une session de 60 minutes. Lorsque ce critère est respecté pendant trois sessions consécutives, le nombre maximal d'interventions

inutiles est fixé à 25 et ainsi de suite, jusqu'à ce que l'objectif fixé soit atteint. L'objectif terminal consiste en un nombre maximum de dix interventions par période de 60 minutes. La première exigence pour les réponses d'attention aux interventions inutiles est fixée à huit, ce qui correspond à la moyenne de niveau de base. Lorsque ce premier critère est atteint, de nouvelles exigences sont précisées jusqu'à ce que l'objectif terminal soit atteint. Celui-ci consiste en un maximum de trois réponses d'attention aux "interventions inutiles". Tous les critères intermédiaires de performance sont déterminés par les sujets eux-mêmes autant pour les "interventions inutiles" que pour les réponses d'attention à ces "interventions".

Dans le but de favoriser le maintien des changements de comportements obtenus, le professeur procède à une phase d'auto-enregistrement avec pairage (phase 3). Chaque sujet enregistre le nombre de fois que lui-même ou un compagnon intervient inutilement. Les résultats ne sont pas rendus publics comme à la seconde phase. Seul le professeur note la fréquence des réponses d'attention aux interventions inutiles. Le système de renforcement n'est plus en vigueur. Par contre, un entraînement à l'auto-enregistrement a lieu grâce à un système de pairage entre les observations des sujets et celles du professeur. Lorsqu'un sujet réussit à faire coïncider ses observations à un chiffre près avec celles du professeur, il reçoit trois "Edas". Six "Edas" sont remis lorsque les deux évaluations sont identiques. Le pairage est toujours fortement renforcé socialement. Cette phase d'auto-observation avec pairage s'étend sur une période de cinq sessions.

Au cours des cinq sessions suivantes, le comportement des sujets n'est plus soumis à des renforçateurs systématiques à l'exception des renforçateurs sociaux intermittents. Pendant cette quatrième phase, les sujets continuent cependant à observer leur comportement. Ces cinq dernières sessions constituent une phase d'auto-enregistrement. Le professeur continue ses observations systématiques pendant cette phase sans toutefois transmettre les résultats aux sujets.

Enfin, cinq périodes d'observation de la part du professeur constituent la phase de post-intervention (phase 5). Aucune intervention n'est effectuée durant cette phase. En principe, les conditions qui prévalent durant cette phase sont semblables à celles qui étaient en vigueur lors du niveau de base. Le professeur n'annonce cependant pas aux sujets que l'intervention est terminée.

Pendant les diverses phases d'intervention, les observations sont effectuées pendant le cours de français durant lequel l'intervention a lieu ainsi que pendant le cours de mathématiques afin de vérifier les effets de généralisation. Aucune intervention n'a lieu durant le cours de mathématiques.

## RESULTATS

Pendant les dix sessions de niveau de base, une moyenne de 54 interventions inutiles est atteinte au cours de français alors que cette moyenne n'est que de 17 interventions inutiles au cours de mathématiques. Une analyse des antécédants révèle qu'en moyenne 2.8 interventions inutiles semblent suscitées par le professeur alors

## Auto-régulation de groupe

que 6.3 semblent sous le contrôle des pairs. La moyenne des réponses d'attention aux interventions inutiles (conséquences) provenant du professeur se chiffre à 3 par période de soixante minutes alors que celles qui proviennent des pairs est de 8.4.

La figure 1 illustre le nombre d'interventions inutiles aux différentes phases d'intervention. Aux différentes étapes du schème à changement de critère utilisé dans la présente intervention (phase 2), nous observons une diminution du comportement-cible. Cette diminution se produit de façon parallèle à l'apparition de critères de performance de plus en plus exigeants. Les comportements-cibles dans les deux situations (cours de français et cours de mathématiques) demeurent sous le seuil de l'objectif terminal pendant les phases d'auto-observation avec pairage (phase 3) d'auto-observation seule (phase 4) et de post-intervention (phase 5).

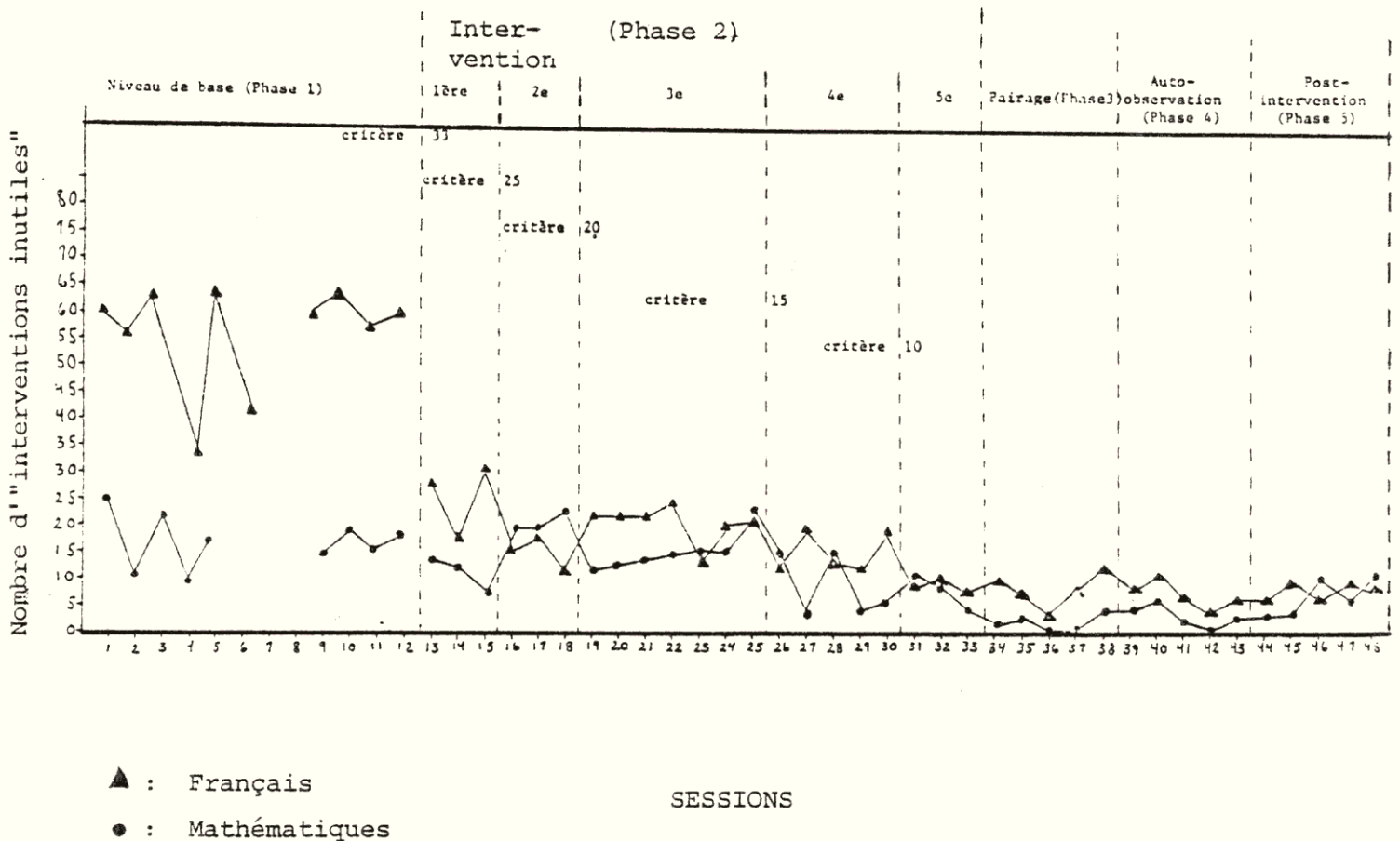


Fig. 1. Nombre d'interventions inutiles émises par les neuf élèves, par période de 60 minutes, durant les cours de français et de mathématiques.

Les pourcentages d'accord entre les observations effectuées par les sujets et celles effectuées par le professeur se situent à 75.2% en moyenne pour les cinq sessions d'auto-enregistrement avec pairage et de 82% en moyenne pour les cinq sessions d'auto-enregistrement seul.

Un test "t" pour échantillons reliés indique qu'il existe une différence statistiquement significative entre le nombre moyen d'interventions inutiles manifestées pendant les cinq premières sessions du niveau de base et celui des cinq dernières sessions d'auto-observation pendant le cours de français ( $t = 8.366$ ,  $p < .005$ ) et pendant le cours de mathématiques ( $t = 4.222$ ,  $p < .05$ ).

De plus, les résultats obtenus pendant le cours de français semblent se maintenir pendant une période de cinq sessions puisqu'il n'y a pas de différence significative entre la moyenne des résultats obtenus par les neuf sujets lors de la phase 5 de post-intervention ( $t = 0.823$ ,  $p > .05$ ) alors qu'il existe une différence significative entre les résultats du niveau de base et ceux de la phase post-intervention ( $t = 7.693$ ,  $p < .01$ ).

Il est important de souligner ici que la performance académique moyenne du groupe en français est de 60.5% pendant la phase de niveau de base, alors qu'elle atteint 83% pendant toute la durée de l'intervention pour retomber à 42% après l'intervention. Les résultats en mathématiques sont beaucoup moins variables: les performances moyennes sont de 75% avant le début de l'intervention, de 88% pendant l'intervention et de 89% durant la période post-intervention.

## DISCUSSION

Les résultats de la présente étude sont en accord avec ceux de Drabman et ses collaborateurs (1973), Bolstad et Johnson (1972) et Kazdin et Polster (1973). La contingence de groupe s'est avérée économique et efficace. La participation des sujets à l'analyse des résultats du niveau de base, à l'aménagement des contingences, à la fixation des seuils de performance et à l'enregistrement de leur propre comportement constituent un programme d'intervention efficace pour modifier les comportements sociaux indésirables en classe. Un tel programme, combiné à un système de renforcement social intermittent, favorise un maintien à court terme des changements de comportements obtenus et ces changements se généralisent à d'autres situations scolaires. En effet, les interventions verbales inutiles durant le cours de mathématiques comme durant le cours de français subissent des diminutions significatives entre les premières et les dernières sessions d'intervention.

Les observations systématiques de niveau de base ont permis de constater que les comportements indésirés étaient en partie maintenus par l'attention des pairs. Le professeur-expérimentateur a alors pu procéder à une extinction efficace des comportements verbaux indésirés en s'assurant la collaboration des pairs et en rendant les responsabilités collectives. Un système d'auto-contrôle et d'auto-aménagement, en plus de permettre d'atteindre et de maintenir les changements de comportements désirés, permet aux sujets dont le comportement est modifié d'apprendre à contrôler leur comportement et leur environnement tout en assurant eux-mêmes la responsabilité

## Auto-régulation de groupe

d'un tel contrôle. L'efficacité d'un tel programme d'intervention mérite d'être comparée à celle de d'autres types de programmes dans le cadre d'une méthodologie expérimentale plus rigoureuse afin de déterminer les variables nécessaires et suffisantes pour que l'efficacité de l'intervention soit optimale. Cependant, les bénéfices personnels que les sujets peuvent tirer d'un système d'auto-contrôle ne doivent pas être sacrifiés au profit d'une efficacité même augmentée.

Un schème à changement de critères permet d'harmoniser la procédure de façonnement et la nécessité de vérification expérimentale du lien de causalité possible qui peut exister entre l'intervention et les changements de comportements (Kazdin, 1973). Ce type de schème permet aussi d'éviter certains inconvénients d'ordre méthodologique que peuvent occasionner les schèmes plus classiques du type à renversement de procédures ou à phases initiales multiples. Il peut s'avérer en effet difficile d'utiliser un schème à renversement de procédures lorsque le but de l'intervention ne consiste pas seulement à installer les changements de comportements désirés mais aussi à favoriser et vérifier un maintien de tels changements. De même, lorsque le but de l'intervention consiste à faciliter et à vérifier une généralisation des changements de comportements à d'autres situations, d'autres sujets ou d'autres réponses, il devient alors embarrassant d'utiliser un schème à phases initiales multiples. Or, le but du présent programme d'intervention consiste en une installation, un maintien et une généralisation à d'autres situations des changements de comportements désirés.

Il est ainsi à noter qu'une réduction des comportements verbaux inappropriés en classe ne saurait se justifier qu'en fonction d'une volonté de créer des conditions favorables à l'apprentissage académique et social.

La présente expérience n'est pas sans présenter certaines lacunes méthodologiques. Ainsi, la présence d'un seul observateur (en occurrence l'expérimentateur) ne permet pas une vérification de la fiabilité des observations. L'ajout d'un observateur indépendant ignorant les hypothèses et les procédures aurait permis de solutionner un tel problème. Aussi, quelques périodes de relance auraient permis de vérifier à plus long terme, le maintien des changements de comportements réalisés.

### REFERENCES

- Allen, K.E., Hart, B., Buell, J.S., Harris, R.T. et Wolf, M.M. Effects of social reinforcement on isolate behavior of a nursery school child. *Child development*, 1964, 35, 511-518.
- Baer, D.M., Wolf, M.M., Risley, T. Some current dimensions of applied behavior analysis. *Journal of applied behavior analysis*, 1968, 1, 91-97.
- Bandura, A. et Perloff, B. Relative efficacy of self-monitored and externally imposed reinforcement systems. *Journal of personality and social psychology*, 1967, 7, 111-116.

- Bolstad, O.D. et Johnson, S.M. Self-regulation in the modification of disruptive classroom behavior. *Journal of applied behavior analysis*, 1972, 5, 443-454.
- Drabman, R.S., Spitalnik, R. et O'Leary, K.D. Teaching self-control to disruptive children. *Journal of abnormal psychology*, 1973, 82 (1), 10-16.
- Glynn, E.L. Classroom applications of self-determined reinforcement. *Journal of applied behavior analysis*, 1970, 3, 123-132.
- Hamblin, R.T., Hathaway, C. et Wodarski, J. Group contingencies, peer tutoring and accelerating academic achievement. In E. Ramp et B.L. Hopkins (Eds.) *A new direction for education: Behavior analysis 1971*, Lawrence: University of Kansas, 1971, p. 41-53.
- Homme, L.E., DeBaca, P.C., Devine, J.V., Steinhurst, R. et Rickert, E.J. Use of Premack principle in controlling the behavior of nurse school children. *Journal of the experimental analysis of behavior*, 1963, 6, 544.
- Johnson, S.M. Self-reinforcement vs. external reinforcement in behavior modification with children. *Developmental psychology*, 1970, 3, 148-149.
- Johnson, S.M. et Martin, S. Developing self-evaluation as a conditioned reinforcer. In B. Ashem et E.G. Poser (Eds.) *Behavior modification with children*. New-York: Pergamon, 1973, 69-78.
- Kazdin, A.E. Methodological and assessment considerations in evaluating reinforcement programs in applied settings. *Journal of applied behavior analysis*, 1973, 6, 1-23.
- Kazdin, A.E. *Behavior modification in applied settings*, Illinois: Dorsey Press, 1975.
- Kazdin, A.E. et Polster, R. Intermittent token reinforcement and response maintenance in extinction. *Behavior therapy*, 1973, 4, 386-391.
- Lovitt, T. et Curtiss, K.S. Academic response rate as a function of teacher and self-imposed contingencies. *Journal of applied behavior analysis*, 1969, 2, 49-54.
- McCarty, T., Griffin, S., Apolloni, T. et Shores, R.E. Increased peer-teaching with group-oriented contingencies for arithmetic performance in behavior-disordered adolescents. *Journal of applied behavior analysis*, 1977, 10, 313.
- O'Leary, K.D., Becker, W.C., Evans, M.B. et Saundargas, R.S. A token reinforcement program in a public school: A replication and systematic analysis. *Journal of applied behavior analysis*, 1969, 2, 3-13.
- O'Leary, K.D. et Drabman, R. Token reinforcement programs in the classroom: A review. *Psychological bulletin*, 1971, 75, 379-398.
- Patterson, G. An application of conditioning techniques to the control of a hyperactive child. In L.P. Ullman et L. Krasner (Eds.) *Case studies in behavior modification*. New-York: Holt, Rinehart and Winston, 1965, 370-375.

## Auto-régulation de groupe

- Ritchl, C., Mongrella, J. et Presbie, R.L. Group time-out from rock and roll music and out-of-seat behavior of handicapped children while riding a school bus. *Psychological reports*, 1972, 31, 967-973.
- Schmidt, G.W. et Ulrich, R.E. Effects of group-contingent events upon classroom noise. *Journal of applied behavior analysis*, 1969, 2, 171-179.
- Sulzbacher, S.I. et Houser, J.E. A tactic to eliminate disruptive behaviors in the classroom: group contingent consequences. *American journal of mental deficiency*, 1968, 73, 88-90.
- Thomas, D.R., Becker, W.C. et Armstrong, M. Production and elimination of disruptive classroom behavior by systematically varying teacher's behavior. *Journal of applied behavior analysis*, 1968, 1, 35-45.
- Wasik, B.H., Senn, K., Welch, R.H. et Cooper, B.R. Behavior modification with culturally deprived school children: Two case studies. *Journal of applied behavior analysis*, 1969, 2, 181-194.

Le comportement des élèves en classe:  
jugement des professeurs vs observation systématique

MICHEL LORANGER (1)

et

RICHARD CLOUTIER

Université Laval  
Québec

RESUME

*L'étude qui suit s'intéresse à l'analyse du comportement en milieu scolaire. Il s'agit d'une étude méthodologique qui confronte le portrait comportemental d'élèves en classe, obtenu par observation systématique avec celui que fournit le jugement du professeur. Les auteurs analysent la valeur globale et relative de ces deux sources d'information.*

L'évaluation est une étape cruciale dans toute démarche d'intervention. En milieu scolaire, il est maintenant établi que les conduites des élèves peuvent être reliées au rendement académique et que certaines d'entre elles, dont l'attention et la participation, facilitent le processus d'apprentissage (Hall, Lund et Jackson, 1968; Hops et Cobb, 1972) et que plusieurs autres, dont les conduites de jeu et de verbalisations non appropriées, entravent l'activité pédagogique (Becker, Madsen, Arnold et Thomas, 1967; O'Leary, Becker, Evans et Saudargas, 1969). Aussi cherche-t-on de plus en plus à contrôler et/ou à utiliser ces variables comportementales pour optimiser l'apprentissage des élèves.

Dans cette perspective, le chercheur et le praticien ont souvent recours à des outils d'évaluation dans le but de situer le niveau d'adaptation des conduites des élèves en classe. A cette fin, deux stratégies d'évaluation sont plus généralement utilisées: celle où l'information est recherchée auprès des professeurs à l'aide de questionnaires (Walker, 1971) et celle qui procède par l'observation systématique (Hamerlynck, Handy et Mash, 1973).

La première stratégie qui implique l'utilisation des jugements des professeurs est couramment utilisée, particulièrement parce qu'elle est économique. Le rationnel qui la sous-tend réside en ceci: comme les professeurs sont ceux-là même qui passent la majeure partie de leur temps avec les élèves, ils constituent des réservoirs appréciables d'information potentiellement utile pour l'intervention. Plusieurs questionnaires ont été développés pour le milieu scolaire afin d'exploiter

1. Les demandes de tiré-à-part doivent être adressées au premier auteur, à: Ecole de Psychologie, Cité universitaire, Québec G1K 7P4.

cette information (Walker, 1971). Cependant, la validité de l'information ainsi obtenue sur les élèves demeure contestée car plusieurs biais chez les évaluateurs peuvent l'affecter (Guilford, 1954; Rosenthal et Jacobson, 1968). Les effets de contraste, de halo, de générosité, d'attente, etc., en sont autant d'exemples. Bon nombre d'études se sont attardées à vérifier la fidélité des données obtenues par ces instruments. Or, comme le souligne Novick, Rosenfeld, Block et Dawson (1966), recourir à des indices de fidélité pour juger de la valeur de ces questionnaires peut être arbitraire. Par exemple, une consistance élevée dans le temps pourrait bien être le reflet d'un biais constant des évaluateurs. A la lumière de ces considérations, des études sur la validité de l'information obtenue sur les élèves par questionnaires adressés aux professeurs s'imposent. Ainsi, la littérature ne présente que très peu de données permettant de confronter le profil des conduites des élèves, obtenu par observation systématique avec le profil des conduites que trace le professeur à partir de son jugement (Walker, 1971).

La deuxième stratégie d'évaluation des élèves repose sur l'observation systématique des conduites en classe. Utilisée avec des précautions méthodologiques auxquelles font référence entre autres Johnson et Bolstad (1973), Jones (1973), Reid (1970), Romanczyk, Kent, Diament et O'Leary (1974), Soulières et McDonough (1978), cette technique semble fournir une image juste des fréquences d'émission de différentes conduites en classe. Elle permet d'établir un portrait de quantités d'émission de comportements; par exemple, le pourcentage de temps qu'un élève passe en classe à émettre différents types de comportements. Cependant, elle est onéreuse en temps et en ressources, de telle sorte qu'elle ne peut pratiquement pas se généraliser en milieu scolaire de façon courante.

Ainsi, d'une part on connaît mal la valeur réelle des jugements des professeurs sur leurs élèves, si ce n'est que plusieurs biais peuvent l'entacher, et d'autre part il apparaît difficile, en raison de son coût, de généraliser l'utilisation de l'observation systématique. Dans quelle mesure est-il possible de dépasser les biais et d'utiliser avec profit l'information dont dispose le professeur sur le comportement de ses élèves? Face à cette question, la présente recherche vise à étudier la valeur des jugements des professeurs, en les confrontant aux données issues de l'observation systématique. Dès lors, cette confrontation cherche à éprouver la capacité des professeurs à évaluer les taux réels d'émission de différentes conduites des élèves en classe. Une comparaison des résultats obtenus de ces deux sources d'évaluation permettra de discuter la valeur de chacune pour le processus d'intervention en classe.

#### METHODE (1)

L'expérience a été menée auprès de quatre professeurs de secondaire II régulier en français et en mathématiques. L'information recherchée est axée sur 17 conduites observables empruntées à la taxonomie et Cobb et Hops (1971). Cette dernière présente l'avantage d'avoir été éprouvée et de s'inscrire dans une tradition de recherche sur les conduites des élèves en classe, qui sont reliées à la tâche académique (Cartledge et Milburn, 1978).

1. *Le lecteur intéressé à une présentation plus exhaustive de la méthode utilisée peut se référer à Loranger (1977).*

## Jugement des professeurs vs observation systématique

1. Approbation (AP)
2. Obéissance (OB)
3. Conversation pertinente avec le professeur (CP)
4. Interaction pertinente avec les étudiants (IE)
5. Participation volontaire (PV)
6. Initiative d'un contact avec le professeur (IC)
7. Orientation visuelle appropriée (OA)
8. Destruction (DT)
9. Désapprobation (AP-)
10. Bruit (BR)
11. Désobéissance (OB-)
12. Jeu (JE)
13. Conversation non pertinente avec le professeur (CP-)
14. Interaction non pertinente avec les étudiants (IE-)
15. Lieu non approprié (LI)
16. Regarder autour (RA)
17. Regarder à proximité un objet non à propos (RP)
18. Ensemble des conduites appropriées (TCA) (de la première à la septième catégorie)
19. Ensemble des conduites non appropriées (TCNA) (de la septième à la dix-septième catégorie)

L'information plus précisément recherchée se situe à deux niveaux: 1. le pourcentage de temps que l'élève passe à émettre chacune de ces conduites (%: mesure absolue-ipsative); et 2. la position de l'élève dans la classe (par rapport à l'élève moyen) quant à la fréquence d'émission de ces mêmes conduites (NO: mesure relative-normative).

Les instruments développés pour parfaire cette comparaison sont au nombre de cinq. Mise à part l'observation systématique, les quatre autres instruments se présentent sous la forme de questionnaires d'évaluation. Le premier (échelle en pourcentage: %) est une évaluation détaillée des 17 premières conduites où le professeur répond à la question: "Quel pourcentage de temps l'étudiant passe-t-il depuis un mois durant vos cours, à émettre des comportements ...?" La même question est formulée dans un deuxième questionnaire d'évaluation globale qui ne compte que les deux dernières questions (nos 18 et 19 ci-haut) et qui sont relatives aux conduites "appropriées" et "non appropriées". On demande alors au professeur: "Quel est le pourcentage de temps ...?" La troisième évaluation est une évaluation détaillée des 17 catégories de conduites où le professeur doit juger l'élève en le

comparant à l'étudiant moyen dans une situation normale de classe, et ce sur une échelle en quatre points (très fréquemment - régulièrement - occasionnellement - rarement). On demande alors au professeur: "Comparé à l'étudiant moyen dans une situation normale de classe, à quelle fréquence, depuis un mois, durant vos cours, l'étudiant manifeste-t-il des comportements ...?" Enfin, ce dernier procédé (NO: échelle normative) est reproduit dans un quatrième questionnaire d'évaluation globale situant l'élève par rapport aux classes de conduites "appropriées" et "non appropriées".

### Observation systématique

L'observation des comportements en classe fut répartie sur un mois durant lequel 39 périodes de 55 minutes ont été couvertes pour les quatre groupes-classe. Le nombre moyen d'observations par sujet a été de 158.7 avec un écart-type de 20.7. Les étudiants ont été observés de l'avant de la classe par un miroir d'observation. Quatre psychologues et un conseiller en orientation ont procédé deux à deux à l'observation systématique en alternant dans leur rôle d'observateur ou de vérificateur. Branché sur le même appareil auditif qui émet un signal à toutes les six secondes, le vérificateur observe sur une cédule intermittente les conduites des élèves. Il est isolé de l'observateur par une paroi afin d'éviter la communication entre eux. Ce procédé a été utilisé afin d'atténuer les biais d'observation inhérents à la mise en présence des deux évaluateurs dans le même local (Johnson et Bolstad, 1973; Lipinski et Nelson, 1974; Romanczyk *et al.*, 1974). Les données des observateurs et des vérificateurs ont permis par la suite d'établir le pourcentage d'accord entre eux.

Préalablement à l'observation comme telle des conduites, les évaluateurs ont reçu un entraînement en trois phases, pouvant se résumer ainsi:

#### Phase 1

Entraînement à la différenciation des 17 catégories de comportements à l'aide d'un manuel spécifiquement élaboré à cette fin (Loranger, 1977). Les observateurs parviennent à atteindre 98% de concordance avec le critère, à un examen d'évaluation.

#### Phase 2

Entraînement par simulation sur bande vidéoscopique où l'on retrouve 18 séquences affichant chacune quatre comportements. Les 17 catégories y sont systématiquement représentées. Le matériel vidéoscopique a été construit en collaboration avec un professeur et quatre élèves qui participent à des jeux de rôle. L'entraînement à l'aide d'exemples visuels se poursuit jusqu'à ce que chaque observateur atteigne une performance de plus de 95% d'accord avec le critère (prototype de cotation).

#### Phase 3

Entraînement "in vivo" où les observateurs intègrent deux à deux le laboratoire et s'entraînent à l'observation des conduites en classe. Cet entraînement se poursuit jusqu'à ce que les deux observateurs atteignent un pourcentage d'accord total

## Jugement des professeurs vs observation systématique

de plus de 80%, celui-ci étant considéré dans ce type de recherche comme satisfaisant (Johnson et Bolstad, 1973).

### Procédé d'évaluation par les professeurs

Les professeurs ont été entraînés, tout comme les observateurs et au même moment, à la différenciation des 17 catégories de comportements (phase 1 de l'entraînement). Ils ont obtenu un niveau de 96% de concordance avec le critère. Après la période d'observation systématique, les quatre professeurs participants ont été impliqués au niveau de l'évaluation de leurs élèves (N = 76).

### Méthode d'analyse des données

Les méthodes de traitement utilisées sont aussi celles qui avaient été retenues par Cobb (1972) pour les mêmes besoins: 1. pourcentage d'accord entre les observateurs; 2. estimation de la stabilité des données par l'utilisation du coefficient de corrélation de Pearson, corrigé par la formule de Spearman et Brown, où les données des périodes impaires sont comparées aux données des périodes paires; 3. utilisation de programme BMDP2R (Dixon, 1975) pour les statistiques descriptives, le tableau des intercorrélations et les analyses de régression multiple.

## RESULTATS ET DISCUSSION

La présentation des résultats met en perspective les éléments suivants: 1. l'analyse des données de l'observation systématique; 2. la comparaison catégorie-catégorie des résultats de deux sources d'évaluation; 3. la comparaison ensemble-catégorie.

### Analyse des données de l'observation systématique

Les données de l'observation ont été confrontées aux jugements des professeurs. Aussi doivent-elles en premier lieu être soumises à une analyse de fidélité et de stabilité. C'est pourquoi la méthode du pourcentage d'accord est ici utilisée (Hamerlynck *et al.*, 1973) de même que la technique d'étude de stabilité utilisée pour les mêmes besoins par Cobb (1972).

Les résultats à ces deux niveaux d'analyse sont présentés au tableau 1. Il apparaît que les observateurs présentent au total un pourcentage d'accord de 83%. Ce niveau de consistance se compare à celui obtenu dans des travaux d'observation similaires (Cobb, 1970, 1972; Hamerlynck *et al.*, 1973). Il est aussi considéré entre autres par Johnson et Bolstad (1973) comme s'insérant dans une limite supérieure réaliste pouvant être atteinte avec ce genre de cotations. Quant à chacune des catégories, ce pourcentage d'accord fluctue de 0% à 94%. Quel est le pourcentage d'accord requis pour justifier la fidélité des données à chaque catégorie? A ce



## Jugement des professeurs vs observation systématique

niveau, la littérature est moins explicite. Néanmoins, certains auteurs (Johnson et Bolstad, 1973) suggèrent de retenir les seules catégories dont le pourcentage d'accord entre les observateurs est significativement plus élevé que celui qui serait obtenu par le simple hasard. Certes, il s'agit là d'un critère minimal. A partir de ce critère, les catégories AP, DT, AP-, BR, OB-, CP- doivent être rejetées au niveau de l'analyse du portrait des conduites individuelles des élèves.

L'étude de la stabilité des données, menée au niveau des catégories qui présentent un pourcentage d'accord satisfaisant, donne les résultats qui apparaissent au tableau 1. Les résultats aux "r" de Pearson, corrigés par la formule de Spearman et Brown, fluctuent de 0.02 à 0.93. L'utilisation des données requiert un seuil minimal de stabilité. Suite aux recherches de même type (Cobb, 1972; Johnson et Bolstad, 1973), il appert qu'un coefficient de corrélation significatif ( $p < .01$ ) peut constituer le critère minimal à retenir. C'est ainsi que la catégorie OB doit être écartée pour retenir les catégories CP, IE, PV, IC, OA, JE, IE-, LI, RA, RP, TCA, TCNA, au niveau de l'analyse des conduites individuelles des élèves.

Tout comme Cobb (1972), on observe une relation significative entre les probabilités d'apparition associées aux différentes catégories et leurs indices de pourcentage d'accord ( $r_s = 0.69$ ,  $p < .05$ ) et leurs indices de stabilité ( $r_s = 0.65$ ,  $p < .05$ ). Cobb (1972) suggérerait pour pallier à cette difficulté d'augmenter le nombre des observations par sujets. Dans la présente recherche, celui-ci est deux fois plus élevé que celui de Cobb (1972) et le même phénomène est remarqué, c'est-à-dire que les indices présentés (fidélité-stabilité) entre les observateurs ont tendance à croître parallèlement à l'augmentation de la fréquence d'apparition des comportements. Deux hypothèses explicatives de ce phénomène sont soulevées: 1. les définitions des catégories qui se présentent moins fréquemment sont moins opérationnelles lorsqu'utilisées au niveau de l'observation "in vivo". 2. ces catégories moins fréquentes peuvent apparaître dans la réalité avec un degré d'expression manifeste plus ou moins élevé, qui pourrait affecter l'indice de pourcentage d'accord. Il est à remarquer de plus que les catégories dont les indices sont insatisfaisants affichent des taux d'incidence très réduits (moins de .17%).

### Comparaison catégorie-catégorie

Le tableau 2 met en parallèle les fréquences moyennes de temps occupé par les différents comportements de 76 sujets, à partir des évaluations des professeurs et des observateurs, et les coefficients de corrélation obtenus entre ces différentes séries de mesures.

Au tableau 2, l'examen comparatif des données issues, de l'observation systématique d'une part, et des jugements des professeurs d'autre part, permet de constater qu'une vision différente des espaces - temps est allouée aux conduites appropriées (CP, IE, PV, IC, OA). Les professeurs accordent des fréquences beaucoup plus élevées à CP, IE, PV, IC, et moins élevées à OA, de telle sorte qu'il apparaît que les professeurs ne reproduisent pas les quantités physiques obtenues à l'observation systématique. A ce niveau donc, leurs jugements ne sont pas valides. Aux catégories non appropriées, JE et IE- ont des fréquences qui correspondent entre les évaluateurs; alors que LI, RA et RP présentent des résultats qui

TABLEAU 9

Pourcentages moyens de temps occupés par chaque catégorie de comportements à l'observation systématique, et aux jugements des professeurs aux échelles en pourcentage, et coefficients de corrélation ( $r$  de Pearson) mettant en relation terme à terme les items des questionnaires adressés aux professeurs (%), NO), et les catégories de l'observation systématique ( $N = 76$ ).

	CP	IE	PV	IC	OA	JE	IE-	LI	RA	RP	TCA	TCNA
$\bar{M}$ observation systématique OS	0.94	2.81	1.07	0.67	73.85	1.80	4.78	0.69	9.14	2.90	80.22	19.77
$\bar{M}$ jugements des professeurs échelle en %	7.15	8.68	9.95	7.28	20.75	1.72	4.46	0.09	5.04	5.43	72.45	27.55
$r$ Pearson entre OS et jugements des professeurs échelle en %	0.06	-0.32*	0.34*	-0.04	-0.13	0.18	0.39*	0.30*	0.27	0.18	0.44*	0.44*
$r$ Pearson entre OS et jugements des professeurs échelle NO	-0.02	-0.04	0.40*	0.24	0.25	0.10	0.08	0.52*	0.11	0.27	0.39*	0.33*

\*  $p < .01$

## Jugement des professeurs vs observation systématique

ne correspondent pas. Aux catégories plus globales (TCA et TCNA), les valeurs absolues correspondent. Le chi carré (voir Siegel, 1956) appliqué sur ces deux dernières variables ( $X_2 = 1.27$ ,  $p > .05$ ) indique qu'il n'y a pas de différence significative entre elles.

Il ressort ainsi que le professeur pourrait être habile à véhiculer un jugement ipsatif sur certaines conduites dont JE, IE-, TCA et TCNA, si les mesures sur les différents sujets étaient significativement reliées. Quant aux autres conduites, la valeur absolue de leurs jugements doit être écartée: en effet, il ne reproduisent pas les quantités physiques attendues pour celles-ci.

La confrontation des données des deux sources d'évaluation pour chacune des catégories peut être étudiée sous l'angle de la relation que ces séries de mesures entretiennent entre elles. Le tableau 2 fait état de ces relations en présentant les résultats au  $r$  de Pearson, entre les données issues de l'observation systématique (OS) et les jugements des professeurs aux différents questionnaires d'évaluation (échelle en %, et échelle normative, NO).

Il en ressort que quelques catégories présentent des relations significatives, plus particulièrement PV, IE-, LI, TCA, TCNA, et que l'utilisation des jugements des professeurs à celles-ci peut contribuer significativement à expliquer les comportements des élèves en classe. En pratique cependant, le pourcentage de variance du critère ainsi expliqué ne justifierait pas le recours unique à cette source d'information pour situer le niveau d'adaptation de ces conduites spécifiques des élèves. C'est comme un élément parmi d'autres que cette information s'avère utilisable. Dans la mesure où les résultats sont généralisables aux fiches descriptives du comportement de l'élève en classe, en usage dans plusieurs milieux québécois, l'information obtenue à chaque item de ces fiches possède un potentiel prédictif faible de la représentation quantifiée des fréquences des conduites des élèves.

### Comparaison ensemble-catégorie

La validité des jugements des professeurs peut être aussi étudiée sous l'angle de la relation entre l'ensemble des jugements des professeurs et les résultats à chacune des catégories de l'observation systématique. A cette fin, le programme d'ordinateur BMDP2R (Dixon, 1975) a été utilisé. Il permet, en spécifiant tour à tour chacune des catégories de l'observation systématique comme variable dépendante, d'utiliser les jugements des professeurs comme variables indépendantes dont on tire les meilleurs éléments pour expliquer le maximum de variance de la variable dépendante et ce, en développant une équation de régression multiple par étapes. Le tableau 3 fait l'état des résultats. Les coefficients de corrélation multiple, en fonction des différentes catégories, sont présentés à la colonne (R).

Les coefficients de détermination ( $R^2$ ) varient de 0.16 à 0.71. Les rapports F sont tous significatifs au seuil  $p < .01$ .

Ces résultats mettent en évidence la possibilité de recourir aux jugements des professeurs pour obtenir une quantité significative d'information sur les conduites spécifiques en classe, information qui ne semblait accessible auparavant que

TABLEAU 3

Indices de prédiction (R, R<sup>2</sup>, F) des données de l'observation systématique obtenus à partir des jugements des professeurs aux questionnaires en % et NO et les cinq premières variables des équations de prédiction

Variable prédite	1 NOM	2 NOM	3 NOM	4 NOM	5 NOM	Total R	R <sup>2</sup>	Rapport F	d1
CP	TCANO	TCNANO	ICNO	AP <sup>-</sup> NO	CP%	0.77	0.60	12.46*	8,67
IE	PV%	RA%	RPNO			0.60	0.36	13.37*	3,72
PV	CP <sup>-</sup> NO	AP <sup>-</sup> %	TCA%	CP%	CPNO	0.84	0.71	11.57*	13,62
IC	OB%	IC%	CP <sup>-</sup> %	IENO	AP <sup>-</sup> NO	0.69	0.48	10.62*	6,69
OA	CP <sup>-</sup> NO	OB <sup>-</sup> NO	TCAD%	RP%		0.60	0.35	9.77*	4,71
JE	TCNANO	DTNO	CP <sup>-</sup> NO	APNO	AP <sup>-</sup> NO	0.64	0.41	7.87*	6,69
IE <sup>-</sup>	IE <sup>-</sup> %	TCA%	TCNA%	AP <sup>-</sup> %	BRNO	0.75	0.56	7.49*	11,64
LI	LINO	AP <sup>-</sup> NO	RA%	RPNO	RANO	0.76	0.58	11.44*	8,67
RA	TCA%	OB <sup>-</sup> %				0.41	0.16	7.20*	2,73
RP	TCNANO	OB%	RA%	PV%	JE%	0.61	0.37	6.68*	6,69
TCA	CP <sup>-</sup> NO	TCA%	RA%	IE <sup>-</sup> %	PV%	0.77	0.59	5.41*	16,59

\* p &lt; .01

## Jugement des professeurs vs observation systématique

par le recours à des techniques d'observation systématique. Par exemple, les jugements des professeurs peuvent expliquer plus de 50% de la variance des résultats à différentes catégories telles CP (60%), PV (71%), IE- (56%), LI (58%), TCA (59%). Ainsi, le réservoir d'information que constitue le professeur, utilisé de cette façon, permet de spécifier le niveau d'adaptation des conduites des élèves en classe.

Le recours à une sélection des informations des professeurs permet une meilleure prédiction que l'utilisation de l'information catégorie-catégorie. Pratiquement toutefois, le recours à des équations de régression rend plus onéreuse et plus complexe l'utilisation de questionnaires demandés aux professeurs. De plus, on ne peut être assuré de la stabilité de ces équations et de leur valeur face à une reproduction éventuelle du même schéma d'expérience sur un autre échantillon.

Il ressort donc de ces résultats que les professeurs ont une expérience globale des conduites en classe, et qu'ils ne peuvent fournir un profil quantitatif exact de ces dernières comme le permet la technique de l'observation systématique. Cette caractéristique distinctive du jugement des professeurs par rapport à l'observation systématique peut s'expliquer par le fait que les jugements des professeurs ne résultent pas, comme pour l'observation systématique, d'un processus d'addition de petites unités d'observation; ils sont plutôt construits à posteriori à partir de leur expérience vécue en classe.

Cependant, le fait de pouvoir reproduire le pattern des résultats de l'observation systématique, à partir des informations des professeurs au moyen d'une analyse de régression statistique, implique que l'expérience globale des professeurs possède des fondements dans les comportements particuliers des élèves pendant les cours. Ces comportements sont ainsi présents au niveau de leur expérience. Cependant, ont-ils tous la même importance pour les professeurs? La technique de l'observation détermine quantitativement l'importance des conduites des élèves, mais elle ne permet pas de juger de l'importance relative de ces conduites quant à leur impact sur le système d'enseignement. Par exemple, 1% de conduites IE- peut avoir un impact sur le système d'enseignement qui est différent de celui introduit par l'émission de 1% de conduites OA. Ainsi, la technique de l'observation systématique est déficiente à fournir l'importance différentielle des différentes conduites, à fréquence égale, quant à leur impact sur la tâche des professeurs. A ce niveau, cette technique demeure incomplète pour traduire la réalité vécue en classe, et elle traite de façon indifférenciée l'impact des conduites au niveau de l'expérience psychologique des professeurs. La recherche dans ce domaine pourrait s'intéresser à pondérer les quantités physiques obtenues pour les resituer par une échelle comparative en tenant compte de l'expérience psychologique des professeurs et des élèves dans leur situation. Bref, si l'observation systématique quantifie mieux les comportements (fréquence, durée, etc.) que les professeurs, elle est insensible à l'impact différentiel des conduites sur la vie de la classe. Par exemple, si un élève en frappe un autre, la classe en sera beaucoup affectée, à fréquence égale de la conduite, que si le même élève affiche une orientation appropriée. A cet égard, les professeurs semblent plus sensibles à la signification ou la valeur environnementale des conduites que l'observation systématique.

Le schéma d'expérience utilisé ici, confrontant les jugements des professeurs aux données de l'observation systématique, permet, par l'étude de la variance commune de leurs résultats, d'aborder cette dernière question reliée à l'importance des différentes conduites des élèves au niveau de l'expérience vécue du professeur. Des sélections de l'information des professeurs (voir tableau 3) parviennent à prédire les différentes conduites (11 catégories) des élèves pendant les cours. Conséquemment, l'ensemble des conduites spécifiques des élèves se reflètent au niveau des jugements des professeurs, indiquant alors que ces comportements sont prégnants au niveau de leur expérience subjective. A titre d'exemple, 13 de leurs jugements sont reliés significativement ( $p < .01$ ) à PV; les comportements PV des élèves constituent donc une réalité bien représentée dans leurs jugements livrés à postériori.

Les valeurs prédictives associées aux différentes catégories de comportements, et présentées au tableau 3, sont inégales: parfois elles sont plus élevées, parfois elles sont plus faibles. Les catégories, dont la variance expliquée par les jugements des professeurs est plus élevée, sont susceptibles d'avoir un impact plus important au niveau de leur expérience vécue dans la classe. Ces catégories de l'observation qui ont le plus de variance expliquée par les jugements des professeurs sont CP, PV, IE-, LI, TCA (plus de 55% de variance commune).

Ces résultats ont des implications importantes quant à l'utilisation de la technique de l'observation systématique des comportements en classe. En effet, quoique cette technique s'avère donner la meilleure représentation des quantités (fréquence en pourcentage de temps-classe d'émission des comportements) de la réalité comportementale de l'élève, l'information qu'elle sous-tend ne véhicule pas les quantités "psychologiques" qui correspondent aux valeurs objectives particulières de chacune des conduites, ou du groupe de conduites. La présente recherche montre que les différentes conduites ont des valeurs différentes pour les professeurs. L'observation systématique ne fournit que les éléments chronologiques d'une réalité. Ces derniers constituent une base importante pour approfondir les mécanismes des conduites humaines. Trop souvent cependant, ces données sont substituées à la réalité globale de l'expérience psychologique vécue par les personnes évaluées. Nos résultats établissent une distinction entre la réalité des fréquences des conduites enracinées dans le temps, telle que reflétée par l'observation systématique, et la réalité psychologique vécue par le professeur et les élèves face aux conduites de ces derniers en classe.

Nos résultats mettent en évidence l'utilité de recourir à l'observation systématique des conduites de l'élève pour éclairer la signification des jugements des professeurs; et en même temps, la nécessité d'aller au-delà de l'observation systématique pour traduire le vécu de la classe. Cette recherche fait de plus ressortir un autre élément: peu importe les questions qui sont demandées aux professeurs, ces derniers semblent traduire leur expérience globale lorsqu'ils évaluent chacun des items. Les prédictions d'une catégorie à partir de différents jugements, s'avèrent plus élevés que les relations catégorie-catégorie entre les jugements des professeurs et l'observation systématique. Il apparaît ainsi que le nom et la définition des catégories utilisées n'ont que peu d'effet direct sur leur jugement. Qu'en est-il alors de leurs jugements sur la créativité, la coopération, le retrait social des élèves? Seraient-ils, à l'instar des résultats de la présente recherche, fortement reliés aux conduites PV, CP, IE et LI des élèves qui facilitent ou entravent la tâche même de l'enseignant? Nos données permettent de poser cette hypothèse

## Jugement des professeurs vs observation systématique

Dès lors, ne peut-on pas s'interroger sur l'utilité des différents questionnaires s'il ne peuvent spécifier à quelles conduites objectives des élèves en classe, ils réfèrent? Beaucoup de fiches descriptives sont touchées directement par cette question, ce qui peut expliquer le fait que fréquemment elles sont abandonnées au profit de nouvelles. Elles prolifèrent ainsi sans pour autant être améliorées.

En résumé, les résultats indiquent: 1- que si la connaissance des fréquences relatives de conduites spécifiques des élèves en classe est recherchée en s'adressant aux professeurs, il faut inférer cette réalité à partir de l'ensemble de leurs jugements (équations de prédiction); 2- et que la simple utilisation de fiches descriptives des comportements des élèves en classe, présente des indices faibles de validité, ce qui en nécessite une utilisation avec réserve; 3- que le recours unique à la technique de l'observation systématique néglige l'importance relative de différentes conduites à fréquence égale dans la situation de la classe.

Finalement, il appert que les portraits des deux sources d'évaluation présentent des complémentarités importantes pour la connaissance des phénomènes psychologiques impliqués dans la situation de la classe.

### REFERENCES

- Becker, W.C., Madsen, C., Arnold, C. et Thomas, D. The contingent use of teacher attention and praise in reducing classroom behavior problems. *Journal of special education*, 1967, 1, 287-307.
- Cartledge, G. et Milburn, J.F. The case for teaching social skills in the classroom: A review. *Review of educational research*, 1978, 2, 133-156.
- Cobb, J.A. *Survival skills and first-grade academic achievement*. Eugene: Center at Oregon for Research in Behavioral Education of the Handicapped, 1970.
- Cobb, J.A. Relationship of discrete classroom behavior to fourth-grade academic achievement. *Journal of educational psychology*, 1972, 63, 74-80.
- Cobb, J.A. et Hops, H. *Coding manual for subject/peer/teacher sequential interactions in academic survival skill settings*. Eugene: Center at Oregon for Research in Behavioral Education of the Handicapped, 1971.
- Dixon, W.J. (Ed.) *BMDP Biomedical computer programs*. Berkely: University of California Press, 1975.
- Guilford, J.P. *Psychometric methods*. New-York: McGraw-Hill, 1954.
- Hall, R., Lund, D. et Jackson, D. Effects of teacher attention on study behavior. *Journal of applied behavior analysis*, 1968, 1, 1-12.
- Hamerlynck, L.A., Handy, L.C. et Mash, E.J. (Eds.) *Behavior change*. Champaign, Ill.: Research Press, 1973.

- Hops, H. et Cobb, J.A. *Survival behaviors in the educational setting: their implications for research and intervention*. Eugene: Center at Oregon for Research in Behavioral Education of the Handicapped, 1972.
- Johnson, S.M. et Bolstad, O.D. Methodological issues in naturalistic observation: Some problems and solutions for field research, in L.A. Hamerlynck, L.C. Handy et E.J. Mash (Eds.): *Behavior change: Methodology, concepts and practice*. Champaign, Ill.: Research Press, 1973, pp. 7-67.
- Jones, R.R. Behavioral observation and frequency data: problems in scoring, analysis, and interpretation, in L.A. Hamerlynck, L.C. Handy et E.J. Mash (Eds.): *Behavior change*, Champaign, Ill.: Research Press, 1973, pp. 119-145.
- Lipinski, D. et Nelson, R. The reactivity and unreliability of self-recording. *Journal of consulting and clinical psychology*, 1974, 42, 118-123.
- Loranger, M. *L'évaluation des comportements en classe chez des élèves de secondaire II*. Thèse de doctorat. Université Laval, Québec, 1977.
- Novick, J., Rosenfeld, E., Block, D.A. et Dawson, D. Ascertaining deviant behavior in children. *Journal of consulting psychology*, 1966, 30, 220-238.
- O'Leary, K.D., Becker, W.C., Evans, M.B. et Saudargas, R.A. A token reinforcement program in the public school: a replication and systematic analysis. *Journal of applied behavior analysis*, 1969, 2, 3-13.
- Reid, J.B. Reliability assessment of observation data: a possible methodological problem. *Child development*, 1970, 41, 1143-1150.
- Romanczyk, R.G., Kent, R.N., Diament, C. et O'Leary, D.K. Measuring the reliability of observational data: a reactive process, in C.M. Franks et G.T. Wilson (Eds.): *Annual review of behavior therapy; Theory and practice*. New-York: Brunner/Mazel, 1974, pp. 403-416.
- Rosenthal, R.A. et Jacobson, L. *Pygmalion in the classroom*. New-York: Holt, Rinehart and Winston, 1968.
- Siegel, S. *Nonparametric statistics for the behavioral sciences*. New-York: Mc Graw-Hill, 1956.
- Soulières, A. et McDonough, C. Les méthodes d'observation dans l'analyse expérimentale du comportement. *Revue de modification du comportement*, 1978, 8, 15-26.
- Walker, H.M. *Early identification and assessment of behaviorally handicapped children in the primary grades*, 1971.

CRITIQUE DE LIVRE

Gordon L. Paul et Robert F. Lentz:  
*Psychosocial treatment of chronic  
mental patients: Milieu versus  
social-learning programs.*  
Cambridge, Mass.: Harvard  
University Press, 1977, 528pp.

Pour ceux et celles que le milieu psychiatrique intéresse, voici un excellent ouvrage de référence sur le traitement actif et continu des patients psychiatriques chroniques. Cet ouvrage, accompagné de graphiques et appendices (inventaires, grilles d'observation, etc.), est écrit dans un style simple, accessible à tous, malgré la présence de données statistiques complexes.

Après avoir soigneusement examiné pendant un an et demi l'état des approches traditionnelles de traitement (telles que la chimiothérapie et l'entretien de base des patients) en milieu psychiatrique, les auteurs en sont venus à la conclusion que ces approches sont insuffisantes et ne garantissent pas la réinsertion sociale des patients dans la communauté. Dans leur ouvrage, ils comparent les soins traditionnels donnés à une clientèle psychiatrique à deux nouvelles approches de traitement: la thérapie de milieu et la thérapie d'apprentissage social.

Le traitement psychosocial va plus loin que les simples soins d'hygiène de base. Au lieu d'entretenir les patients en faisant tout pour eux sous prétexte qu'ils ne comprennent pas, des procédures sont mises de l'avant pour les amener graduellement à prendre soin d'eux-mêmes sans l'aide extérieure continue des responsables de leur bien-être. Ces procédures, qui s'inspirent des principes du conditionnement opérant et de l'apprentissage social, visent à ce que les patients s'insèrent progressivement dans la communauté, par le biais de maisons de transition et de foyers, et à ce qu'ils vivent en milieu non-institutionnalisé. Cette désinstitutionnalisation, associée à l'acquisition d'habiletés de base essentielles à un comportement adapté dans la communauté, est l'objectif ultime de travail de Paul et Lentz.

La méthodologie utilisée est systématique et rigoureuse. Une clientèle de patients psychiatriques chroniques fut minutieusement

assignée à trois groupes distincts: le groupe A, traitement psychiatrique traditionnel; le groupe B, thérapie de milieu; et le groupe C, traitement d'apprentissage social. Les deux derniers groupes furent traités à l'extérieur des lieux de l'hôpital, dans des centres à caractère communautaire. Le programme global, depuis la conception du projet jusqu'à la rédaction finale de l'ouvrage, recouvre une période de 10 ans. Quant au programme de traitement comme tel, il s'échelonne sur une période de 6 ans, avec une relance de 1½ an pour les groupes B et C.

Pour résumer brièvement les résultats obtenus, l'approche d'apprentissage social s'avère de loin la plus efficace pour la réinsertion sociale des patients et le traitement des troubles de comportement. Vient ensuite la thérapie de milieu, qui donne également des résultats impressionnants. Quant à l'approche psychiatrique traditionnelle, elle apparaît en dernier lieu, tant du point de vue de la réinsertion sociale que du traitement des troubles de comportement. Les résultats de cette recherche exhaustive soulèvent des interrogations majeures concernant le bien-être des patients psychiatriques. Notamment, toute la philosophie thérapeutique des patients psychiatriques est remise en question, de même que les méthodes traditionnelles de traitement.

Une question qu'il convient de poser: est-ce que ce schème de traitement peut s'appliquer dans les milieux psychiatriques francophones canadiens? On peut certes apporter une réponse affirmative à une telle question, dans la mesure toutefois où il y a amélioration du côté des ressources financières, matérielles et humaines (subventions, maisons de transition, personnel qualifié, etc.).

En terminant, voici un aperçu rapide du contenu de l'ouvrage de Paul et Lentz. Une première partie relate la revue de littérature sur le milieu psychiatrique et ses méthodes de traitement. La deuxième partie contient le schème du traitement, c'est-à-dire population, échantillon, groupes, procédures de traitement. Les troisième, quatrième, cinquième et sixième parties présentent les résultats du programme à diverses étapes de son déroulement. La septième partie compare les traitements traditionnels aux deux types de traitement psychosocial. Les résultats d'une relance d'un an et demi avec les patients réinsérés dans la communauté sont donnés dans une huitième partie. Enfin, en neuvième partie, viennent les conclusions et recommandations.

L'ouvrage de Paul et Lentz est certes à recommander aux spécialistes en modification de comportement oeuvrant en milieu psychiatrique, à la fois à cause de son souci de précision, de simplicité et de rigueur scientifique.

Claude Blaquière  
Centre hospitalier de Restigouche  
Campbellton, N.-B.