

Évaluation de *Psychological Abstracts* par l'entremise de trois systèmes en mode direct : comparaison des résultats d'une recherche bibliographique

The evaluation of *Psychological Abstracts* using three on line systems: Comparing the results of a literature search

Evaluación de *Psychological Abstracts* utilizando tres sistemas directos: comparación de los resultados de una investigación bibliográfica

Alain Pagé

Volume 26, numéro 2, juin 1980

URI : <https://id.erudit.org/iderudit/1054248ar>

DOI : <https://doi.org/10.7202/1054248ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Association pour l'avancement des sciences et des techniques de la documentation (ASTED)

ISSN

0315-2340 (imprimé)

2291-8949 (numérique)

[Découvrir la revue](#)

Citer cet article

Pagé, A. (1980). Évaluation de *Psychological Abstracts* par l'entremise de trois systèmes en mode direct : comparaison des résultats d'une recherche bibliographique. *Documentation et bibliothèques*, 26(2), 83-95.
<https://doi.org/10.7202/1054248ar>

Résumé de l'article

La plupart des principales banques d'information sont offertes, de nos jours, par plus d'un distributeur commercial. La sélection d'un système donné repose sur certains critères de comparaison parmi lesquels il faut favoriser la cohérence d'une banque (file consistency) et les stratégies d'interrogation des systèmes qui affectent directement les résultats d'une recherche bibliographique et, conséquemment, la qualité du service à offrir à un chercheur. Afin de vérifier ces deux aspects, une recherche bibliographique est menée à l'intérieur de la banque *Psychological Abstracts* par l'entremise de trois systèmes de repérage de l'information en mode direct : DIALOG/LIS, STAIRS/BRS et ORBIT/SDC. Cette recherche s'établit selon deux approches particulières : une première investigation à l'intérieur de la catégorie « descripteur » uniquement, puis une seconde à l'intérieur des différentes catégories de type « sujet » de l'index principal de la banque *Psychological Abstracts*.

Évaluation de *Psychological Abstracts* par l'entremise de trois systèmes en mode direct: comparaison des résultats d'une recherche bibliographique*

Alain Pagé

Bibliothèque

Université du Québec à Trois-Rivières

La plupart des principales banques d'information sont offertes, de nos jours, par plus d'un distributeur commercial. La sélection d'un système donné repose sur certains critères de comparaison parmi lesquels il faut favoriser la cohérence d'une banque (file consistency) et les stratégies d'interrogation des systèmes qui affectent directement les résultats d'une recherche bibliographique et, conséquemment, la qualité du service à offrir à un chercheur. Afin de vérifier ces deux aspects, une recherche bibliographique est menée à l'intérieur de la banque Psychological Abstracts par l'entremise de trois systèmes de repérage de l'information en mode direct: DIALOG/LIS, STAIRS/BRS et ORBIT/SDC. Cette recherche s'établit selon deux approches particulières: une première investigation à l'intérieur de la catégorie "descripteur" uniquement, puis une seconde à l'intérieur des différentes catégories de type "sujet" de l'index principal de la banque Psychological Abstracts.

The evaluation of Psychological Abstracts using three on line systems: Comparing the results of a literature search

Nowadays, most of the principal information data bases are offered by more than one commercial distributor. Selecting a system depends upon certain criteria, primarily file consistency and inquiry techniques. These criteria directly affect the results of a literature search and consequently the quality of service available to researchers. To verify these two aspects, a literature search was carried out on the same data base, Psychological Abstracts, using three on line systems for information retrieval: DIALOG/LIS, STAIRS/BRS and ORBIT/SDC. This research used two approaches. The first involved an inquiry into the descriptor category. The second consisted of an inquiry into the different categories of subjects included in the main index of Psychological Abstracts.

Evaluación de Psychological Abstracts utilizando tres sistemas directos: comparación de los resultados de una investigación bibliográfica

La mayoría de los principales bancos de datos son ofrecidos por más de un distribuidor comercial. La selección de un sistema se basa en algunos criterios de comparación entre los cuales se debe favorecer la coherencia del banco y las estrategias de interrogación del sistema; estos factores afectan directamente los resultados de una investigación bibliográfica y, por consiguiente, la cualidad del servicio ofrecido al investigador. Para verificar esta teoría, una investigación bibliográfica se conduce en el banco Psychological Abstracts con tres sistemas de recuperación de la información: DIALOG/LIS, STAIRS/BRS y ORBIT/SDC. Esta investigación se establece según dos métodos particulares: una primera investigación en la categoría "Descriptor" únicamente, y una segunda dentro de las distintas categorías temáticas del índice principal del banco Psychological Abstracts.

*N.D.L.R. À cause des changements constants dans ce domaine, il est important de noter que l'auteur a déposé son article en août 1979.

Les services de repérage automatisé de l'information ont pris beaucoup d'ampleur au cours des cinq dernières années. Ce grand essor de la recherche bibliographique en mode direct a été rendu possible grâce d'une part à la disponibilité de banques considérables répertoriant la littérature de différents champs de spécialisation surtout depuis les cinq à dix dernières années, et ce à un coût abordable, résultat de la forte concurrence que s'imposent les grands distributeurs de banques et grâce d'autre part à l'amélioration importante des réseaux de télétransmission de données tant aux États-Unis (Tymshare et Telenet) qu'au Canada (Datapac).

Les bibliothécaires et documentalistes ont emboîté le pas dans cette voie du modernisme et on peut considérer comme achevée la phase d'implantation des principaux systèmes dans la plupart de nos institutions universitaires et de recherche, phase qui s'est échelonnée sur une période de trois ans (1975-1977).

À cette époque, la raison principale qui poussait les institutions à s'abonner à tel et tel système était surtout d'ordre pratique: des banques importantes qui pouvaient répondre en tout ou en partie aux besoins documentaires des chercheurs n'étaient fournies très souvent que par un seul distributeur. Cependant, cette situation est de moins en moins vraie; on constate en effet qu'une majorité de banques d'information considérées comme majeures sont maintenant offertes par deux, trois et même quatre distributeurs (Tableau 1).

L'utilisateur des systèmes de repérage automatisé de l'information se trouve maintenant exposé à un problème de taille: quel système choisir dans un tel cas?

Au cours des deux dernières années, les auteurs ont passablement débattu cette question, sans toutefois en arriver à une conclusion définitive¹. Tellement de facteurs², variables dans le temps, sont en

Tableau I: Liste de quelques banques disponibles dans différents systèmes

Banques \ Systèmes	CAN/OLE	DIALOG	ORBIT	STAIRS
Biosis	X	X	X	X
Chemical abstracts	X	X	X	X
Comprehensive dissertation index		X	X	X
Compendex	X	X	X	
Eric		X	X	X
Inform		X	X	X
Inspecc	X	X	X	X
NTIS	X	X	X	X
Psychological abstracts		X	X	X
Social science citation index		X	X	X

cause et affectent ce que Pemberton appelle "la valeur d'une recherche"! "Value is what we seek in everything we purchase, and often it is a combination of many factors, some of them subtle, some of them almost unnoticed."³ L'utilisateur a sans doute la même réaction que l'acheteur lorsque vient le temps de choisir un système pour répondre à une demande particulière de recherche bibliographique. Son choix est sûrement guidé par des considérations qui ont un rapport direct avec le coût d'un système: taux horaire d'une banque, temps moyen de connexion par recherche, coût de l'impression en différé, taux d'abonnement au système. De plus, l'utilisateur tient probablement compte d'un certain nombre de raisons immédiates, instinctives, que lui dictent son flair, son expérience et sa connaissance des divers systèmes au moment de l'entrevue ainsi que sa familiarité à manipuler les différentes commandes d'un système lors de la recherche au terminal. Le fait de se sentir à l'aise avec un système a, d'ailleurs, une incidence directe sur le coût d'utilisation de ce système. Plus un usager manipule les diverses commandes d'un système avec

1. Le lecteur n'a qu'à consulter les articles publiés dans des revues telles *Online Review*, *Online*, *Journal of the American Society for Information Science* pour constater cet état de fait.

2. Pemberton en définit quatorze. Voir à ce sujet Jeffery K. Pemberton, "The inverted file - A column by the publisher", *Online*, vol. 1, no. 4 (October 1977), 6-8.

3. *Ibid.*, 6.

dextérité, moins le temps de connexion requis pour l'exécution de la recherche au terminal est long.

Bien que toutes ces considérations aient leur importance, il n'en demeure pas moins que le choix d'un système doit avant tout être relié aux possibilités qu'offre chacun d'entre eux de répondre le plus adéquatement possible aux attentes documentaires d'un chercheur. Dans la mesure où la même banque est disponible dans différents systèmes, que la période couverte est semblable, ne doit-on pas choisir le système qui permettra d'obtenir les meilleurs résultats même si, pour ce faire, il en coûte un peu plus cher?

Les deux critères pour mesurer la qualité d'une recherche dans une banque sont l'exhaustivité et la pertinence⁴. La recherche de l'exhaustivité soulève deux questions précises auxquelles nous allons tenter de répondre tout au cours de cet article:

1. Une banque fournie par un distributeur correspond-elle intégralement à la version originale de l'éditeur, ce que les auteurs appellent la cohérence d'une banque (file

consistency)?⁵ Une étude de Rouse et Lannom⁶ nous renseigne davantage sur ce point. Les deux auteurs ont comparé les résultats d'une recherche bibliographique menée à l'intérieur de la banque *Eric* par l'intermédiaire de trois systèmes de repérage de l'information en mode direct: DIALOG/LIS (Lockheed Information Systems), ORBIT/SDC (System Development Corporation) et STAIRS/BRS (Bibliographic Retrieval Services). Sur les seize descripteurs interrogés, ORBIT a fourni treize réponses⁷ différentes par rapport aux deux autres systèmes, tandis que DIALOG et STAIRS ne montraient entre eux qu'une seule réponse divergente. Malgré une équivalence au niveau des caractéristiques d'interrogation des trois systèmes, il est surprenant de constater une incohérence du nombre de références rattachées aux mêmes descripteurs d'un système à l'autre.

2. Par l'utilisation de certaines ressources de chacun des systèmes, quels sont les facteurs qui affectent, qui modifient les résultats d'une recherche d'un système à l'autre? En d'autres termes, en utilisant les possibilités d'interrogation de chacun des systèmes, le résultat bibliographique est-il différent? Si oui, sous quel rapport?

Ces deux aspects sont importants; ils remettent en cause la qualité du service à offrir aux chercheurs. Si, pour une recherche donnée, on obtient un résultat quantitatif moindre à cause de l'utilisation d'un système plutôt qu'un autre, ou parce que certaines possibilités du système ont été mal exploitées, comment justifier alors auprès du chercheur une qualité de service à laquelle il est en droit de s'attendre?

4. Le taux de rappel (exhaustivité) et le taux de précision (pertinence) sont deux mesures qui permettent d'évaluer l'efficacité du repérage de l'information. Le calcul de ces taux s'obtient de la façon suivante:

Taux de rappel: $\frac{\text{Nombre de références pertinentes obtenues}}{\text{Nombre de références pertinentes de la banque}} \times 100$

Taux de précision: $\frac{\text{Nombre de références pertinentes}}{\text{Nombre de références repérées}} \times 100$

La mesure du taux de rappel est difficile à obtenir. Madeleine Balcer, dans un article publié dans la revue *Documentaliste*, vol. 16, no 2 (mars-avril 1979), 59, sous le titre "Réactions de l'utilisateur face à l'utilisation du système de repérage en mode dialogué, BADADUQ", résume bien les raisons qui causent l'échec d'un taux de rappel absolu, i.e. 100%. Ces raisons bien souvent échappent, du moins en partie, au contrôle de l'utilisateur d'un système documentaire. Ainsi, lorsqu'il sera fait mention d'exhaustivité, il doit être bien entendu que ce ne sera pas en terme de valeur absolue. Il faudra lui prêter le sens de "tendance à l'exhaustivité".

5. Une des façons de vérifier la cohérence d'une banque est de comparer les résultats d'une recherche d'un système à l'autre.

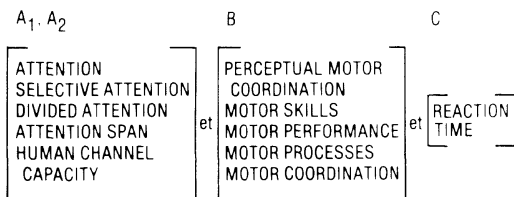
6. Sandra H. Rouse and Laurence W. Lannom, "Some differences between three on-line systems: impact on search results", *Online Review*, vol. 1, no. 2 (June 1977), 117-132.

7. L'interrogation d'un descripteur dans une banque d'information bibliographique amène l'ordinateur à fournir une réponse qui correspond à l'impression au terminal d'un résultat quantitatif représentant le nombre de notices à l'intérieur desquelles l'on retrouve ce descripteur. En plus du terme réponse, on emploie souvent l'expression résultat de recherche.

Approche

Pour tenter d'apporter un éclairage sur ces questions, nous allons comparer les résultats d'une recherche bibliographique menée à l'intérieur de la banque *Psychological Abstracts* par le truchement de trois systèmes de repérage de l'information en mode direct: DIALOG/LIS, STAIRS/BRS et ORBIT/SDC. L'expérience porte sur une période de recensement bibliographique de *Psychological Abstracts* qui s'échelonne de janvier 1967 à avril 1979. Le sujet de l'interrogation concerne les exigences requises en attention par un individu qui exécute deux mouvements simultanément et le temps de réaction demandé pour l'accomplissement de chacune de ces tâches⁸. Le plan de concepts⁹ s'élabore comme suit:

Tableau II: Plan de concepts



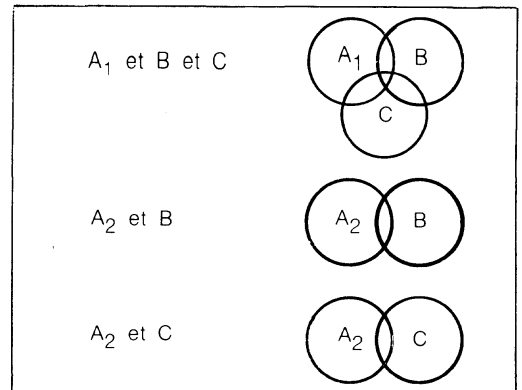
A₁: sont inclus dans cet ensemble les cinq descripteurs ci-haut mentionnés;

A₂: sont inclus dans cet ensemble les quatre descripteurs mentionnés ci-haut, sauf le descripteur ATTENTION.

Ce plan de concepts dont les descripteurs sont tirés du thesaurus propre à cette banque¹⁰ montre trois combinaisons à exécuter dans chacun des systèmes.

Ces combinaisons sont:

Tableau III: Combinaisons logiques



Il convient tout d'abord de mettre en oeuvre une combinaison à trois ensembles qui a pour effet de faire intervenir chacun des trois aspects, soit l'attention sélective, l'habileté motrice et le temps de réaction. Le descripteur ATTENTION a été inclus dans la combinaison de ce premier ensemble, même si le sujet étudié concerne l'attention sélective à proprement parler. Cependant, étant donné la restriction qu'impose une combinaison à trois ensembles, il est opportun d'étendre cet ensemble au terme ATTENTION sans risquer d'obtenir un résultat bibliographique trop volumineux. Cela permettra possiblement de repérer quelques références pertinentes de plus qui n'auraient pu être dépistées autrement. Les deuxième et troisième combinaisons, des intersections à deux ensembles, excluent le descripteur ATTENTION dans le premier bloc de concepts; sinon, le nombre de références repérées grâce à ces deux combinaisons dépasserait largement la limite acceptable sans pour autant que les documents soient relatifs à l'attention sélective comme telle.

La recherche s'établit selon deux approches particulières: tout d'abord, une première investigation à l'intérieur de la catégorie "descripteur" uniquement; puis, une seconde à l'intérieur des différentes catégories de type "sujet" de l'index principal, incluant la catégorie "descripteur". Le vocabulaire est contrôlé pour ces deux approches, les descripteurs provenant pour la plupart du thesaurus de la banque *Psychological Abstracts*. Cependant, la seconde approche comporte deux autres descripteurs qui s'ajoutent à l'ensemble C du plan de concepts initial, soit RESPONSE

8. Nous désirons remercier Pierre Black, étudiant en sciences de l'activité physique à l'Université du Québec à Trois-Rivières, qui nous a aimablement permis d'utiliser son sujet de recherche comme base de données à cette étude et qui, de plus, a généreusement fourni de son temps pour la détermination de la pertinence des références repérées.

9. Le plan de concepts est un petit diagramme qui montre les agencements logiques des termes les uns par rapport aux autres.

10. *Thesaurus of Psychological Index Terms*, 2nd ed., Washington, American Psychological Association, (c1977), 300 p.

SPEED et RESPONSE TIME. Ces deux descripteurs libres, synonymes de REACTION TIME, permettront peut-être de repérer quelques références pertinentes de plus.

L'analyse des résultats, selon la première approche, servira à vérifier la cohérence de la banque *Psychological Abstracts* à l'intérieur de chacun des trois systèmes. Une comparaison des résultats obtenus par l'une et l'autre approche révélera ensuite si nous avons, par l'utilisation de la seconde approche, augmenté le nombre de références repérées et surtout le nombre de références pertinentes. Cette approche est plus longue à exécuter au terminal et produit un plus grand nombre de références au total; de ce fait, elle est plus coûteuse qu'une recherche dans la seule catégorie "descripteur". Il convient donc de s'assurer que les résultats justifient les moyens. Enfin, pour la seconde approche, si les résultats diffèrent d'un système à l'autre, il est nécessaire de déterminer les causes de ces différences. Elles trouvent sans doute leur origine au niveau des diverses stratégies d'interrogation des systèmes, ce qui n'exclut pas la possibilité, encore là, d'un manque de cohérence de la banque.

Caractéristiques des systèmes

Une comparaison de systèmes est possible dans la mesure où tous les éléments du problème, c'est-à-dire le plan de concepts, la banque consultée, la période couverte et les caractéristiques de recherche propres à chacun des systèmes, sont similaires. Les trois premiers éléments du problème ne causent aucune difficulté. Quant au quatrième élément, les caractéristiques de systèmes¹¹, il convient de s'y attarder afin d'établir, dans la mesure du possible, les équivalences de recherche, ou sinon, d'en noter les différences. Avant toute chose, cependant, il importe d'étudier la structure de la banque à l'intérieur des systèmes.

Structure de la banque *Psychological Abstracts*

Psychological Abstracts est une des nombreuses banques offertes par chacun des trois systèmes. Cette banque est composée, d'après ORBIT, de 282,675 notices bibliographiques (janvier 1967 - avril 1979). Une notice bibliographique pour sa part est constituée d'un certain nombre de catégories (Tableau IV) dont les éléments servent alternativement à deux fonctions spécifiques: une première fonction permet, grâce à l'utilisation des éléments de certaines catégories, de repêcher les informations désirées lors de la recherche au terminal; une deuxième fonction permet, suite à une commande d'impression ou d'affichage, de visualiser les éléments d'un ensemble de notices bibliographiques afin de juger de leur pertinence.

Tableau IV: Échantillon d'une notice bibliographique (*Psychological Abstracts/STAIRS*)

AN	09349 59-5
AU	Stratton, Richard K.
IN	Florida State U. Col of Education.
TI	Selective attention in a multiple-response reaction-time task.
SO	Perceptual and motor skills. jun 77 vol 44 (3, pt 1) 889-890.
LG	en.
YR	77.
CC	2380.
PT	10.
DE	Selective-attention. Reaction-time. Motor-performance.
AB	Consistently higher correlations (.74-.89) were obtained between hand and foot reaction times in a button-pressing task for 7 college volunteers assigned to a contralateral group than for 7 other ss assigned to an ipsilateral group (.51-.80).
ID	Ipsilateral vs contralateral hand & foot, selective attention in multiresponse rt task, college students.

Index principal

La recherche d'un mot, à défaut de mentionner l'abréviation d'une catégorie en particulier à la suite de ce mot, se fait dans l'index principal. Le Tableau V montre les différentes catégories de *Psychological Abstracts* qui constituent l'index principal de chacun des systèmes.

11 *A Brief Guide to Dialog Searching*, Palo Alto (Calif.), Lockheed Information System, 1976. *Guide to Dialog Databases*, Palo Alto (Calif.), Lockheed Information System, 1977, chap. 11, *BRS System Reference Manual*, Schenectady (N.Y.), Bibliographic Retrieval Services, c1978. *Orbit IV is Coming!* Toronto, Infomart, 1978, 14 p. *Orbit Quick Reference Guide*, Santa Monica (Calif.), SDC Search Service, 1978.

Tableau V:
Index principal des abréviations de la banque *Psychological Abstracts* à l'intérieur des systèmes DIALOG, STAIRS et ORBIT

	DIALOG	STAIRS	ORBIT
. Numéro de résumé	- -	. AN .	-
. Auteur/s		. AU .	
. Affiliation	/ CS	. IN .	-
. Titre	/ TI	. TI .	/ TI
. Source	-	. SO	
. Code de classification de P.A.		. CC .	-
. Descripteur	/ DE	. DE MJ .	/ IT
. Résumé	/ AB	. AB .	(/ BI)
. Mot-clé	/ ID	. ID .	/ IW
. ISSN / ISBN	-	. IS .	
. Coden	-	. CD .	

Il va de soi que seulement quatre catégories de l'index principal sont, dans la plupart des cas, à l'origine des références repérées lors d'une recherche de mots/sujets. Ces catégories, titre, descripteur, mot-clé et résumé sont toutes disponibles dans l'un ou l'autre des trois systèmes.

Recherche dans la catégorie "Descripteur"

Une recherche dans la catégorie "descripteur" s'effectue grâce à l'utilisation des termes du thesaurus de *Psychological Abstracts*. Ces descripteurs sont de deux types: le descripteur uniterme (ATTENTION) et le descripteur multiterme (SELECTIVE ATTENTION).

Descripteur uniterme

Si on désire que la recherche soit exécutée à l'intérieur de la catégorie "descripteur" uniquement, il est nécessaire de faire suivre le descripteur uniterme de l'abréviation appropriée, soit /DE (DIALOG), /IT (ORBIT) ou .MJ. (STAIRS). L'équivalence n'est cependant pas possible pour les trois

systèmes. DIALOG se distingue des deux autres systèmes du fait qu'il ne peut chercher le descripteur uniterme si celui-ci fait aussi partie d'un descripteur multiterme, par exemple ATTENTION et SELECTIVE ATTENTION. Dans un tel cas, les résultats de la recherche correspondent aux références repérées grâce au mot ATTENTION comme descripteur uniterme, d'une part, et comme élément de descripteurs multitermes, d'autre part.

Descripteur multiterme

La recherche d'un descripteur multiterme dans la catégorie "descripteur" ne pose aucun problème d'équivalence entre les trois systèmes. Cependant, on doit noter une particularité propre au système STAIRS: un trait d'union doit apparaître entre chacun des mots d'un descripteur multiterme. Il n'est pas nécessaire de faire suivre le descripteur multiterme par l'abréviation: /DE (DIALOG), /IT (ORBIT) et .DE. (STAIRS). La catégorie "descripteur" est la seule à l'intérieur de laquelle il est permis de rechercher des expressions de deux mots et plus sans que ces mots ne soient réunis par un opérateur logique ou par une stratégie quelconque.

Recherche dans les catégories de l'index principal¹²

Les descripteurs unitermes et multitermes peuvent aussi se retrouver ailleurs qu'à l'intérieur de la seule catégorie "descripteur". L'omission de l'abréviation d'une ou de quelques catégories à la suite d'un descripteur uniterme entraîne l'ordinateur à chercher ce mot dans toutes les catégories disponibles de l'index principal. Même si l'analyse documentaire n'a pas fait ressortir ces termes, il convient sûrement, dans une recherche à caractère exhaustif, d'examiner ces références et de juger de leur pertinence.

Recherche d'expressions (Full-text searching)

Les mots d'un descripteur multiterme doivent, pour leur part, être réunis de façon à ce qu'ils demeurent près l'un de l'autre.

12. L'ensemble de ce paragraphe fait référence au Tableau VI.

Ainsi, les termes REACTION et TIME, s'ils ne sont pas adjacents, perdent leur valeur sémantique et ne méritent plus d'être considérés pour les fins de la recherche. Pour cette raison, les systèmes ont développé une stratégie particulière qui permet justement d'unir deux termes voisins.

Ainsi, un (W) entre deux termes, dans le système DIALOG, unit ces deux termes selon l'ordre prescrit. À défaut de spécifier une catégorie quelconque à la suite de cette expression, la recherche est faite dans l'une ou l'autre des catégories de l'index principal, incluant la catégorie "descripteur".

Le système STAIRS, pour sa part, permet une recherche de mots adjacents au moyen de l'abréviation ADJ entre ces mots, comme par exemple SELECTIVE ADJ ATTENTION. Comme pour DIALOG, si aucune abréviation de catégorie n'est mentionnée à la suite de l'expression, l'ordinateur exécute

la recherche dans chacune des catégories de l'index principal. Dans certains cas, l'expression recherchée correspond à une partie de titre de périodique. Ainsi, la recherche de l'expression MOTOR ADJ SKILLS, dans le système STAIRS, permet le repérage de toutes les notices rattachées au périodique *Perceptual and Motor Skills*. Dans ce cas-ci, il convient d'éliminer la recherche de cette expression à l'intérieur de la catégorie "source" en utilisant une stratégie particulière d'exclusion de catégorie: MOTOR ADJ SKILL\$1..SO.

Quant au système ORBIT, l'opération qui assemble deux termes à proximité l'un de l'autre s'appelle STRINGSEARCH et, même si elle peut se comparer à celles des deux autres systèmes sur le plan des résultats, on ne peut que constater son manque de souplesse lors de l'interrogation au terminal. En effet, alors que l'opération de recherche se limite à une seule étape

Tableau VI: Recherche à l'intérieur des catégories de l'index principal de la banque *Psychological Abstracts*

DIALOG/LIS	STAIRS/BRIS	ORBIT/SDC
1: # ATTENTION	1: ATTENTION	1: ATTENTION
2: # SELECTIVE (W) ATTENTION	2: SELECTIVE ADJ ATTENTION	2: SELECTIVE ATTENTION
3: # DIVIDED (W) ATTENTION	3: DIVIDED ADJ ATTENTION	3: DIVIDED ATTENTION
4: # ATTENTION (W) SPAN	4: ATTENTION ADJ SPAN	4: ATTENTION SPAN
5: # HUMAN (W) CHANNEL (W) CAPACITY	5: HUMAN ADJ CHANNEL AHJ CAPACITY	5: HUMAN CHANNEL CAPACITY
6: \$2-5/OR	6: 2 OR 3 OR 4 OR 5	6: 2 OR 3 OR 4 OR 5
7: \$1 + 6	7: 1 OR 6	7: 1 OR 6
8: # MOTOR (W) SKILL	8: MOTOR ADJ SKILL\$1..SO.	8: PERCEPTUAL MOTOR COORDINATION
9: # MOTOR (W) SKILLS	9: MOTOR ADJ PERFORMANCE	9: MOTOR SKILLS
10: \$8 + 9	10: MOTOR ADJ PROCESSES	10: MOTOR PERFORMANCE
11: # MOTOR (W) PERFORMANCE	11: MOTOR ADJ COORDINATION	11: MOTOR PROCESSES
12: # MOTOR (W) PROCESSES	12: 8 OR 9 OR 10 OR 11	12: MOTOR COORDINATION
13: # MOTOR (W) COORDINATION	13: REACTION ADJ TIME\$1	13: 8 OR 9 OR 10 OR 11 OR 12
14: \$10-13/OR	14: RESPONSE ADJ SPEED\$1	14: REACTION TIME
15: # REACTION (W) TIME	15: RESPONSE ADJ TIME\$1	15: RESPONSE AND SPEED #
16: # REACTION (W) TIMES	16: 13 OR 14 OR 15	16: STRS (TI, IT, ST, AB) :RESPONSE # SPEED:
17: \$15 + 16	17: 7 AND 12 AND 16	17: RESPONSE/TI AND TIME # /TI
18: # RESPONSE (W) SPEED	18: 6 AND 12	18: RESPONSE/IW AND TIME # /IW
19: # RESPONSE (W) SPEEDS	19: 6 AND 16	19: 17 OR 18
20: \$18 + 19		20: STRS (TI, IT, ST, AB) :RESPONSE # TIME:
21: # RESPONSE (W) TIME		21: 14 OR 16 OR 20
22: # RESPONSE (W) TIMES		22: 7 AND 13 AND 21
23: \$21 + 22		
24: \$17 + 20 + 23		
25: \$7*14*24		23: 6 AND 13
26: \$6*14		24: 6 AND 21
27: \$6*24		

dans le système STAIRS, à une ou deux étapes dans le système DIALOG, cette opération prend souvent une dimension hors de portée dans le système ORBIT: il est fréquent que l'utilisateur de cette stratégie doive donner une quinzaine d'instructions à l'ordinateur avant d'arriver à un résultat équivalent. Étant donné que le coût d'une recherche dépend en grande partie du facteur "temps d'interrogation au terminal", la stratégie STRINGSEARCH devient quasi inabordable. Pour les besoins de cette étude, nous ne faisons intervenir la stratégie STRINGSEARCH qu'au niveau des deux expressions RESPONSE SPEED et RESPONSE TIME.

Dans STAIRS et DIALOG, la recherche de l'expression MOTOR COORDINATION au niveau de l'index principal englobe déjà toutes les références qui pourraient être repérées grâce à l'expression PERCEPTUAL MOTOR COORDINATION. C'est la raison pour laquelle, sur le plan stratégique, cette dernière expression est totalement ignorée.

Troncature

Lors d'une recherche de descripteurs unitermes et multitermes dans les catégories de l'index principal, il convient de prévoir toutes les formes possibles d'entrée pour ces termes: l'adjectif, le nom, le singulier, le pluriel. La troncature permet, en une seule opération, d'inclure dans la recherche tous les mots dont les premières lettres sont identiques au terme en question.

La troncature s'exprime, dans DIALOG, par le point d'interrogation "?" à l'intérieur ou à la fin d'un terme. Cependant, il n'est pas permis de tronquer deux termes unis par un "W". L'expression doit alors apparaître deux fois, avec et sans "s", pour les formes d'entrée au singulier et au pluriel. Dans STAIRS, la troncature se traduit par le signe de dollar "\$" et un chiffre à la suite du mot; le chiffre indique le nombre de lettres que l'ordinateur peut ajouter au terme tronqué. La troncature, dans ORBIT, se définit par le dièse "#" et par deux points ":"; le dièse à la fin d'un mot permet l'ajout d'une lettre à ce mot tandis que les deux points permettent l'ajout d'un nombre illimité de lettres.

Stratégie de recherche

Une des stratégies proposées pour cette expérience est l'interrogation d'un seul descripteur à la fois par ligne au terminal, suivant ainsi le modèle du plan de concepts. Cette stratégie n'est pas nécessairement la plus rapide; en effet, les trois systèmes acceptent la recherche de plusieurs descripteurs séparés par des opérateurs logiques (ou, et, sauf) à l'intérieur d'une même opération d'interrogation. L'ordinateur exécute alors les commandes une à une et sa réponse inclut l'ensemble des termes interrogés. Cette stratégie ne nous permet pas cependant de connaître le nombre de références attribuées à chacun des descripteurs, élément essentiel pour les fins de cette étude. La démarche utilisée selon l'approche de recherche à l'intérieur de la catégorie "descripteur" se fait en dix-sept étapes réparties comme suit:

- .1-5: recherche des descripteurs de l'ensemble A;
- .6-7: combinaison des descripteurs de l'ensemble A;
- .8-12: recherche des descripteurs de l'ensemble B;
- .13: combinaison des descripteurs l'ensemble B;
- .14: recherche du descripteur de la colonne C;
- .15-17: intersection des ensembles.

Pour sa part, la recherche à l'intérieur des catégories de l'index principal, telle que proposée au Tableau VI, laisse voir un plus grand nombre d'étapes de réalisation au terminal ainsi qu'une utilisation accrue des possibilités offertes par chacun des trois systèmes.

Il faut noter que la stratégie développée s'est strictement limitée à une question de vocabulaire et que tout autre intervenant en a été exclu, simplement parce qu'un nombre suffisant d'éléments sont déjà en cause et permettent amplement de vérifier les données à l'origine de cette expérience. Qu'il suffise de mentionner que les catégories d'une notice bibliographique peuvent pour la plupart intervenir lors de la

préparation d'une stratégie de recherche; ainsi, par exemple, une recherche bibliographique peut être limitée à une langue en particulier, à un certain nombre d'années et à un type de publication. D'autres éléments à l'intérieur des catégories d'une notice peuvent en permettre le repérage, comme les noms des auteurs et de leur organisme, le titre de la publication, le code de classification de *Psychological Abstracts*, le coden et l'ISSN/ISBN.

Les conditions de recherche au terminal prévalant pour cette expérience ont varié quelque peu par rapport aux recherches habituelles, en ce sens qu'il n'y a pas eu de véritable dialogue entraînant parfois des modifications substantielles au plan de concepts et à la stratégie initiale. Étant donné le caractère expérimental de cette recherche, tout l'aspect interactif a été mis de côté, ce qui a eu pour effet de diminuer passablement le temps de connexion¹³. En contrepartie, la stratégie adoptée lors de cette expérience, c'est-à-dire un seul descripteur à la fois par ligne au terminal, est sûrement plus longue d'exécution. Bien que ces conditions soient artificielles, il n'en demeure pas moins qu'elles sont semblables pour chacun des trois systèmes; le temps de connexion correspond alors à ce qu'on pourrait appeler "la performance brute d'un système", c'est-à-dire le temps que met l'ordinateur pour répondre à chacune des interrogations de l'utilisateur. C'est pourquoi nous allons fournir le temps de connexion et le coût rattachés aux deux approches de recherche menées à l'intérieur des trois systèmes. Même si on ne peut utiliser ces chiffres comme éléments de base pour le choix d'un système, à cause du caractère expérimental de la recherche et aussi parce qu'il faut tester un plus grand nombre de fois dans des conditions de recherche habituelles pour ainsi établir une moyenne de temps de recherche, ces données numériques peuvent toujours servir comme éléments d'information s'ajou-

tant aux indications fournies par d'autres auteurs¹⁴.

Analyse des résultats

Cohérence

L'examen des résultats de la recherche (ou des réponses) menée à l'intérieur de la catégorie "descripteur" de la banque *Psychological Abstracts* (voir Tableau VII) ne permet pas de déceler une incohérence quelconque d'un système à l'autre. Les chiffres obtenus, sauf pour le descripteur ATTENTION, sont en tous points semblables. Afin de vérifier si ce fait n'est rattaché qu'à la banque *Psychological Abstracts*, les descripteurs utilisés par Rouse et Lannom lors de leur expérience de 1977¹⁵ sont reconduits à nouveau à l'intérieur de la banque *Eric* de chacun des trois systèmes. Là encore, on constate une similitude au niveau des résultats, alors qu'en 1977, quatorze descripteurs sur seize montraient une différence d'un système à l'autre.

Particularités de système

Les résultats de la recherche menée à l'intérieur des différentes catégories de type sujet de l'index principal (voir Tableau VII) diffèrent pour quelques expressions d'un système à l'autre¹⁶. Après vérification des références de quelques-unes de ces expressions (DIVIDED ATTENTION, ATTENTION SPAN et RESPONSE SPEED/S), on peut dégager trois points particuliers:

A) Deux incohérences dans le système DIALOG: on ne retrouve l'expression DIVIDED ATTENTION dans aucune catégorie

13. En effet, les statistiques de repérage automatisé de l'information des trois dernières années (juin 1976 - mai 1979) à la bibliothèque de l'Université du Québec à Trois-Rivières permettent d'établir la moyenne de temps de connexion à une banque à 12.9 minutes. Durant cette période, 780 banques des systèmes ORBIT, DIALOG, STAIRS, CAN/OLE et ELHILL ont été consultées.

14. Lyse Pérusse, *Étude comparative des systèmes de repérage de l'information bibliographique en mode dialogué*, Québec, Université du Québec, 1979, p. 27-31. Ryan E. Hoover, "A comparison of three commercial online vendors", *Online*, vol. 3, no. 1 (January 1979), 12-21. Johanna C. Ross, "Searching the chemical literature via three online vendors: A comparison", *Journal of the American Society for Information Science*, vol. 30, no. 2 (March 1979), 103-106.

15. Sandra H. Rouse and Laurence W. Lannom, "Some differences . . .", 124.

16. Il a déjà été mentionné que seule la stratégie STRINGSEARCH permet la recherche d'expressions dans ORBIT. Sauf pour deux expressions, RESPONSE SPEED et RESPONSE TIME, la recherche de tous les termes est exécutée, dans ce système, à l'intérieur de la catégorie "descripteur".

Tableau VII: Résultats - Psychological Abstracts, janvier 1967 - avril 1979

Termes	Recherche dans la catégorie "descripteur"			Recherche dans les catégories de l'index principal		
	DIALOG	STAIRS	ORBIT	DIALOG	STAIRS	ORBIT
1. Attention	3 137	2 783	2 783	7 320	7 316	2 783
2. Selective attention	289	289	289	538	538	289
3. Divided attention	76	76	76	109	108	76
4. Attention span	27	27	27	124	125	27
5. Human channel capacity	152	152	152	154	154	152
6. 2-5/ou	497	497	497	874	874	497
7. 1 ou 6	3 239	3 239	3 239	7 420	7 416	3 239
8. Perceptual motor coordination	154	154	154	—	—	154
9. Motor skill	—	—	—	261	—	—
10. Motor skills	166	166	166	466	—	166
11. 9 ou 10	—	—	—	663	663	—
12. Motor performance	2 203	2 203	2 203	2 344	2 340	2 203
13. Motor processes	2 592	2 592	2 592	3 834	3 834	2 592
14. Motor coordination	109	109	109	390	391	109
15. 8-14/ou	5 076	5 076	5 076	6 652	6 651	5 076
16. Reaction time	2 257	2 257	2 257	2 702	—	2 257
17. Reaction times	—	—	—	433	—	—
18. 16 ou 17	—	—	—	2 826	2 831	—
19. Response speed	—	—	—	97	—	—
20. Response speeds	—	—	—	37	—	—
21. 19 ou 20	—	—	—	117	119	114
22. Response time	—	—	—	257	—	—
23. Response times	—	—	—	177	—	—
24. 22 ou 23	—	—	—	404	403	117
25. 16-24/ou	—	—	—	3 158	3 168	2 386
26. A ₁ et B et C	16	16	16	30	30	16
27. A ₂ et B	13	13	13	39	39	13
28. A ₂ et C	38	38	38	80	79	38

Tableau VIII: Pertinence des résultats selon l'approche utilisée

	Documents repérés		Documents pertinents		Taux de précision	
	Cat. desc.	Index princ.	Cat. desc.	Index princ.	Cat. desc.	Index princ.
A ₁ et B et C	16	30	11	14	69%	46%
A ₂ et B	13	39	9	15	69%	38%
A ₂ et C	38	80	10	15	26%	19%
Total	65*	137*	28*	38*	43%	27,7%

*Les documents en double ne sont comptés qu'une fois.

de la notice 42-08749 et pourtant, elle a été repérée grâce à cette expression. Quant à la notice 52-12784, l'expression ATTENTION SPAN s'y retrouve au niveau du résumé et malgré cela, ces deux termes n'ont pu en permettre le repérage.

B) La liste des mots sans signification est différente dans DIALOG et STAIRS, ce qui explique l'absence de deux références dans le système DIALOG par rapport au système STAIRS suite à la recherche de l'expression RESPONSE SPEED. Les suites de mots "response or speed" et "response has upon the speed" apparaissent dans les résumés des notices 60-10802 et 43-12195. Ces deux références ont été repérées dans STAIRS et non dans DIALOG et ORBIT. STAIRS ne considère pas, à tort d'ailleurs, les mots séparant RESPONSE et SPEED comme étant significatifs.

C) La stratégie STRINGSEARCH du système ORBIT permet la recherche de deux termes avec ou sans un certain nombre d'espaces, de lettres ou de chiffres entre ceux-ci. Ainsi, un dièse entre RESPONSE et SPEED ne laisse place qu'à un espace blanc entre ces deux termes. Pour leur part, les systèmes DIALOG et STAIRS acceptent la présence de signes (comme la virgule et la parenthèse) et de mots sans signification entre deux termes. C'est la raison pour laquelle l'expression RESPONSE SPEED draine cinq références de plus dans le système STAIRS que dans ORBIT.

Sans avoir vérifié les références des autres expressions qui engendrent une différence quantitative, on peut supposer les mêmes raisons pour expliquer ces quelques références en plus ou en moins à l'intérieur des systèmes.

Recherche dans la catégorie "descripteur" et recherche dans toutes les catégories de l'index principal

Le Tableau VIII montre que sur 65 références repérées grâce à la recherche dans la seule catégorie "descripteur", 28 ont été déclarées pertinentes, soit un taux de précision de 43%. Ce taux de précision relativement intéressant a été rendu possible grâce au bon travail des documentalistes qui ont indexé ces 65 documents.

La recherche dans toutes les catégories de l'index principal à l'intérieur des systèmes DIALOG et STAIRS¹⁷ a permis de repérer 72 nouvelles références par rapport aux résultats obtenus lors de la première approche. De ce nombre, il n'y a que dix notices qui ont été déclarées pertinentes pour un taux de précision de 13,9%. Ce faible taux s'explique par le fait que les 72 références ont été repérées grâce aux termes contenus dans toutes les catégories de l'index principal, à l'exception de la catégorie "descripteur". Ce sont des expressions qui apparaissent dans le titre, le résumé ou les mots-clés et qui n'ont pas

17. La recherche d'expressions dans toutes les catégories de l'index principal, sauf pour les expressions RESPONSE SPEED et RESPONSE TIME, n'a pas été réalisée à l'intérieur du système ORBIT. Pour obtenir une certaine équivalence avec les deux autres systèmes, il aurait fallu utiliser la stratégie STRINGSEARCH pour chacune des expressions, ce qui aurait eu pour effet d'augmenter considérablement le temps de connexion. Ainsi, les résultats qui apparaissent dans la dernière colonne du Tableau VII correspondent aux résultats obtenus suite à la recherche dans la catégorie "descripteur".

Tableau IX: Temps et coûts de recherche

	Recherche dans la catégorie "descripteur"			Recherche dans les catégories de l'index principal		
	DIALOG	STAIRS	ORBIT	DIALOG	STAIRS	ORBIT
Date	7/6/79	7/6/79	12/6/79	7/6/79	7/6/79	12/6/79
Heure de départ	9h22	8h46	8h46	9h29	9h50	8h57
Temps en minutes	4min17s	4min44s	9min	7min12s	7min06s	17min
Coûts estimés en \$ U.S.	11,67\$	10,92\$	18,14\$	23,30\$	19,85\$	27,41\$

été retenues par le documentaliste lors de l'analyse du document. Il demeure cependant que dix références pertinentes sur un total de 38 n'auraient pu être repérées autrement que par cette approche de recherche.

Quelle importance doit-on accorder à ces dix références par rapport à l'ensemble des références déclarées pertinentes? Un examen des 28 références permet de les classer ainsi: 8 références sont très pertinentes, 16 sont moyennement pertinentes et 4 sont peu pertinentes. Sur les dix références repérées en plus, suite à la recherche des descripteurs dans les catégories de l'index principal, trois sont considérées moyennement pertinentes et sept peu pertinentes. Pour obtenir un tel résultat, il en a coûté 23,30\$ U.S. dans DIALOG et 19,85\$ U.S. dans STAIRS, soit approximativement deux fois le montant d'une recherche menée dans la catégorie "descripteur" seulement. Parmi les trois références considérées comme moyennement pertinentes, deux auraient pu être repérées lors de la première approche si le descripteur PERCEPTUAL MOTOR PROCESSES avait été inclus dans l'ensemble B du plan de concepts. Quant à la troisième référence, une notice qui traite davantage de l'attention partagée, l'analyste lui a attribué le descripteur ATTENTION plutôt que DIVIDED ATTENTION.

Temps et coûts de recherche

Les recherches ont été exécutées tôt en matinée et sensiblement aux mêmes heures. Les chiffres montrent (voir Tableau IX) qu'il y a peu de différences au niveau des temps d'exécution de la recherche au terminal à l'intérieur des systèmes DIALOG et STAIRS. Quant aux coûts, ils s'équivalent

à peu près même si les taux d'escompte consentis par les distributeurs aux usagers selon leur volume d'utilisation n'ont pas été calculés.

Là où la différence de temps est importante, c'est entre ORBIT et les deux autres systèmes: plus de quatre minutes selon la première approche et environ dix minutes pour la seconde approche. Afin de vérifier la pertinence de ces temps, la même recherche a été reconduite dans la catégorie "descripteur" des systèmes ORBIT (26 juillet, 8h51), DIALOG (27 juillet, 10h) et STAIRS (27 juillet, 10h08). Les résultats de ce second test montrent des temps d'exécution de 5 min 55s (DIALOG), de 4 min 14s (STAIRS) et de 9 min (ORBIT). DIALOG a été plus lent de 1 min 38s par rapport au premier test. Ce ralentissement s'explique par le fait que l'ordinateur a pris respectivement 54 et 45 secondes pour répondre à deux interrogations particulières, alors qu'habituellement le temps de réponse est assez rapide. L'ordinateur du système STAIRS, pour sa part, a été plus efficace de 30 secondes; le temps d'exécution de la recherche équivaut au temps établi par l'ordinateur du système DIALOG lors du premier test, soit 4 min 14s. Le temps d'attente moyen suite à une interrogation au terminal est à peine de 1,5s. Quant au système ORBIT, il ne montre aucune différence de temps entre le premier et le second test. Trois causes principales ont été notées pour expliquer cette performance relativement lente du système (si on la compare à celles des deux autres systèmes): 1. Le temps d'attente moyen suite à l'interrogation d'un descripteur ou d'une combinaison logique se situe à neuf secondes; 2. Le temps global pour l'exécution de la commande STORAD (commande de mise en mé-

moire d'une adresse donnée) est de 1 min 17s; 3. Enfin, la commande d'impression de la bibliographie en mode différé prend en moyenne 45 secondes à réaliser. Il faut remarquer à propos de ce dernier point que l'impression des trois bibliographies en mode différé, dans le système ORBIT, a été exécutée en une seule commande plutôt que trois, comme pour les deux autres systèmes.

Conclusion

Cette étude met en évidence les résultats de deux recherches bibliographiques; la première a été réalisée selon deux approches distinctes, l'une dans la catégorie "descripteur" uniquement, l'autre dans toutes les catégories de l'index principal, à l'intérieur de la banque *Psychological Abstracts* des systèmes DIALOG/LIS, STAIRS/BRS et ORBIT/SDC. La seconde a été effectuée dans la catégorie "descripteur" à l'intérieur de la banque *Eric* des trois mêmes systèmes. L'analyse des résultats soulève trois points particuliers: A) Il faut remarquer, tout d'abord, la cohérence des banques *Psychological Abstracts* et *Eric* lors d'une recherche de termes contrôlés dans la catégorie "descripteur" de chacun des trois systèmes. La similitude des réponses contraste avec les résultats qu'avaient obtenus Rouse et Lannom en 1977. Cependant, les deux incohérences décelées lors de la recherche des expressions DIVIDED ATTENTION et ATTENTION SPAN dans toutes les catégories de l'index principal du système DIALOG laissent planer un doute sur la cohérence de la banque *Psychological Abstracts* à l'intérieur de ce système.

B) Cette recherche nous a aussi permis de constater une certaine qualité de l'indexation des documents contenus dans la banque *Psychological Abstracts*. 28 références pertinentes sur un total de 38 ont été repêchées grâce au travail des indexeurs. De ce nombre, 24 (85,7%) ont été déclarées très utiles pour les fins de la recherche. Parmi les dix références pertinentes obtenues suite au repérage mené dans les catégories de type sujet de l'index principal (titre, mots-clés et résumé), et en supposant que le descripteur PERCEPTUAL MOTOR PROCESSES ait été inclus dans le plan de concepts original, une seule référence,

considérée utile, aurait alors échappé à l'utilisateur. Ainsi, sur un total de 27 références jugées moyennement et très pertinentes, 26 (96,3%) auraient pu être repêchées selon l'approche de recherche dans la catégorie "descripteur"¹⁸.

C) Par voie de conséquence, une recherche à caractère exhaustif peut tout aussi bien être exécutée selon l'une ou l'autre approche. Cependant, la recherche de termes dans la catégorie "descripteur" doit être favorisée au détriment de la recherche d'expressions dans les catégories de l'index principal, à cause surtout d'un coût de recherche moins élevé, mais aussi parce que l'utilisateur se retrouve avec une bibliographie répondant davantage à ses attentes (taux de précision plus élevé).

N'ayant pu déceler une différence marquée entre chacun des trois systèmes au niveau des résultats de recherche, le choix d'un système pour l'utilisation de la banque *Psychological Abstracts* s'appuie alors davantage sur des considérations d'ordre monétaire. En exploitant une stratégie de recherche appropriée aux possibilités d'interrogation propres à chacun, le système retenu est celui qui entraîne les frais les moins élevés.

18. Bien sûr, un plus grand nombre de recherches exécutées dans cette banque selon chacune des deux approches aurait permis de mieux évaluer la qualité du travail de l'indexeur. Cependant, le fait d'avoir obtenu pour cette recherche un taux de précision aussi élevé que 96,3% nous laisse supposer un bon rendement général.