

## VARIABLE NODALE ET CONE IMPLICATIF

Dominique LAHANIER-REUTER<sup>1</sup>, Régis GRAS<sup>2</sup>, Marc BAILLEUL<sup>3</sup>

### RÉSUMÉ.

Dans le cadre de l'Analyse Statistique Implicative, il est procédé à l'extraction de règles implicatives pondérées par une mesure appelée intensité d'implication. Un graphe dit implicatif permet de représenter l'ensemble des relations non symétriques qui associent les variables. Au sein du graphe, on observe quelquefois, à partir d'une variable centrale, des relations amont et aval telles qu'il soit possible d'isoler ces deux types de part et d'autre de ce nœud, ce confluent, sous la métaphore de cône implicatif à deux grappes. Donner une signification à ce cône, sous la condition d'une certaine homogénéité et de la connexité de l'ensemble, permet de l'isoler conceptuellement du tout en des termes de causalité-conséquences. Dans cette communication, nous définissons selon deux approches un critère d'homogénéité d'un cône et des conditions d'existence de relations causales. Nous illustrons le propos théorique par l'examen de plusieurs exemples significatifs.

*Mots-clés.* Analyse statistique implicative, règle, graphe, nœud, cône, sommet du cône, arête, cause et conséquence,

### ABSTRACT

Inside the Statistical Implicative Analysis frame, implicative rules may be extracted while pondered by a measure called Intensity of Implication. The set of the non-symmetrical relationships that associate the different variables may be represented by a graph so called implicative. Inside the graph, relationships upstream and downstream from a central variable parted on either sides of this node can be sometimes noticed. This situation of a confluent is metaphorically said a two clusters implicative cone. On condition that the set is rather homogeneous and linked, giving conceptual meaning to this cone allows some kind of causalities/consequences isolation of it from the whole set. In this communication, we define criteria for the homogeneity of a cone and conditions for existence of causal relationships. Some significant examples illustrate our theoretical discourse.

*Keywords* Statistical Implicative Analysis, Rule, Graphic, Node, Cone, Edge, Top of the cone, Causes and consequences

## 1 Introduction

Disposant d'un corpus de données croisant sujets et variables, l'Analyse Statistique Implicative offre des réponses à certains objectifs du chercheur :

- Extraire de cet ensemble de données où les sujets sont éléments d'un ensemble discret E et les variables sont de natures diverses, des règles

---

<sup>1</sup> Équipe Théodile CIREL, Université Lille 3, Villeneuve d'Ascq, France, E-mail: [dominique.lahanier@univ-lille3.fr](mailto:dominique.lahanier@univ-lille3.fr)

<sup>2</sup> École Polytechnique de l'Université de Nantes, Équipe DUKE Data User Knowledge, Laboratoire d'Informatique de Nantes-Atlantique (LINA), UMR 6241, : [regisgra@club-internet.fr](mailto:regisgra@club-internet.fr)

<sup>3</sup> CERSE (Centre d'Études et de Recherches en Sciences de l'Éducation), Université de Caen Basse-Normandie, courriel : [marc.baillleul@unicaen.fr](mailto:marc.baillleul@unicaen.fr)

- implicatives par exemple du type : « si la variable a est instanciée alors généralement la variable b l'est aussi », dans le cas de variables booléennes,
- Affecter à chaque règle une valeur numérique de [0,1] croissante avec la qualité prédictive de la règle,
  - Représenter par un graphe dit implicatif, non symétrique, pondéré et sans cycle, l'ensemble des règles de qualité au moins égale à un certain seuil d'acceptabilité,
  - Interpréter en donnant du sens aux arêtes du graphe, mais aussi aux structures connexes qui le constituent.

C'est ainsi qu'il existe, pour un seuil d'intensité d'implication donné, certaines sous-structures connexes où des nœuds du graphe admettent isolément des antécédents (des « pères ») et des successeurs (des « fils »). Ces sous-structures sont extraites « à la main » par l'utilisateur lui-même au vu du graphe implicatif donné par CHIC. Antécédents et successeurs peuvent éventuellement être interprétés respectivement comme causes et conséquences du sommet du graphe dans une perspective causale.

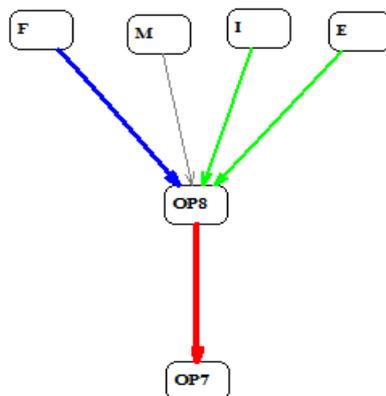


Figure 1

Par exemple, sur la figure 1, on a extrait, d'un graphe implicatif d'une quarantaine de variables et au seuil 0.69, la sous-figure ci-dessus que nous appelons « *cône implicatif* », avatar géométrique car il admet un sommet et deux nappes dont l'amont et l'aval représentent respectivement les « pères » et les « fils ». Le sommet de ce cône, OP8, admet pour « pères » (pluriel, comme c'est bizarre !) ou « causes » les variables F, M, I et E. Il n'a qu'un « fils » ou « conséquence » : OP 7. Ce sommet joue un rôle que l'on peut désigner comme « confluent » ou « variable nodale » de variables. Le mot « cône » a été retenu car il symbolise aussi, comme un entonnoir, un objet de type déversoir » où confluent divers affluents et d'où repartent de nouveaux réseaux absorbant les flux de l'amont. Jean-Marc Lévy-Leblond, dans (J.-M. Lévy-Leblond, 1996, p. 320) parlerait même d'un **champ causal** ou un écheveau enchevêtré qui « ne laisse plus individualiser les fils continus au long desquels pourrait se concevoir la propagation d'influences causales ». Au sujet des mots constitutifs de la phrase, François Gaudin dans (Sève, L., 2005) dirait que l'on voit « s'y dessiner une topologie du sens ».

Ce cône pouvant se prolonger le long de ses deux nappes, on constate que la notion de « cause » (ou de « pères ») et de « conséquence » (ou « d'effet » ou de « fils ») est toute relative. Un père n'est père qu'en tant qu'il a un fils. Nous rencontrons ici aussi le

raisonnement logique de type dialectique où l'analyse a pour fonction la résolution de la contradiction. La cause détermine l'effet tout autant que l'effet détermine la cause (cf L.Sève, 2005, p.71). Son sens ne se construit que dans son rapport avec l'effet.

Notre intention est d'utiliser ce cône pour mettre en évidence un ensemble de causes non liées entre elles afin de pouvoir identifier leurs rôles, leurs responsabilités dans l'apparition de l'attracteur. Autrement dit, nous nous intéresserons de façon privilégiée aux cônes dont les « pères » du sommet sont indépendants ou très faiblement dépendants. Ce sommet pourrait définir une sorte de « bassin d'attraction » au sens de la théorie des catastrophes. Sous cette restriction *d'indépendance*, nous nous limiterons donc ici, un graphe étant donné comme image d'un ensemble de règles, à ce type de structure conique et à son interprétation. Mais celle-ci exige que ce cône ait des chances d'offrir une certaine consistance sémantique et, pour cela, que sa constitution présente des caractères analytiques révélateurs d'une certaine homogénéité implicative. *N'y observerait-on pas un phénomène comparable à celui que L. Sève désigne par « propriété émergente » ?* Cette problématique constituera notre premier paragraphe.

## 2 Indice d'homogénéité d'un cône implicatif

### 2.1 Rappels

Au cours de la formalisation de l'implication statistique, nous avons introduit un indicateur de qualité de la mesure d'une règle de type  $a \Rightarrow b$ . Dans le cas des variables binaires, nous avons comparé les occurrences des contre-exemples à la règle à celles qui serait observables au hasard si les variables étaient indépendantes. Pour quantifier cette proposition, aux sous-ensembles A et B de E des observations de sujets effectives et respectives de a et b, nous avons associé des parties X et Y de mêmes cardinaux que A et B. En modélisant un tirage au hasard dans E des parties X et Y, nous avons comparé le cardinal aléatoire  $N_{a\bar{b}}$  des contre-exemples à l'implication (soit  $\text{card}(X \cap \bar{Y})$ ), au cardinal  $n_{a\bar{b}}$  (soit  $\text{card}(A \cap \bar{B})$ ) des contre-exemples observés. L'intensité d'implication qui mesure la qualité de l'implication est alors la probabilité que le hasard conduise à plus de contre-exemples que la contingence soit :

$$\phi(a,b) = \Pr[\text{Card}(X \cap \bar{Y}) > n_{a\bar{b}}] = 1 - \Pr[\text{Card}(X \cap \bar{Y}) \leq n_{a\bar{b}}] \quad (1)$$

Si l'on centre et l'on réduit la variable aléatoire  $\text{Card}(X \cap \bar{Y})$  sous la forme :

$$Q(a, \bar{b}) = \frac{\text{Card}(X \cap \bar{Y}) - \frac{n_a n_{\bar{b}}}{n}}{\sqrt{\frac{n_a n_{\bar{b}}}{n}}} \quad (2)$$

Celle-ci dans des conditions à grands effectifs est une variable gaussienne  $N(0,1)$  en tant que valeur approchée d'une variable aléatoire de Poisson centrée et réduite (cf. Partie 1, Chapitre 1 de Gras et Régnier, 2013). Sa réalisation contingente est  $q(a, \bar{b})$ . Par suite, l'intensité d'implication devient dans ce cas gaussien :

$$\varphi(a, b) = 1 - \Pr[Q(a, \bar{b}) \leq q(a, \bar{b})] = \frac{1}{\sqrt{2\pi}} \int_{q(a, \bar{b})}^{\infty} e^{-\frac{t^2}{2}} dt \quad (3).$$

## 2.2 Rôle fonctionnel des intensités d'implication des arêtes

Une première approche pourrait consister à déterminer la *variance implicative* intra-classe du cône implicatif (cf. Partie 2, Chapitre 2 de Gras et Régnier, 2013). Cette variance du cône, basée sur une interprétation géométrique de l'implication statistique, offre une mesure qui permet d'apprécier la consistance d'une grappe connexe du graphe comme l'est un cône implicatif. L'absence de programmation des calculs associés dans CHIC nous incite, puisque nous visons des calculs conduits à la main, à retenir une autre méthode plus intuitive et aisée sur le plan calculatoire.

Une deuxième approche, plus simple et plus intuitive, consiste à examiner la consistance d'un tout à partir de celles de l'ensemble de ses arêtes, même si un tout est plus informatif que l'ensemble de ses parties. Ainsi, un cône implicatif étant donné, une mesure naturelle pour en qualifier la qualité d'homogénéité se nourrit de l'ensemble des qualités implicatives des arêtes. Il est alors possible de mesurer cette qualité à travers, par exemple, la moyenne arithmétique ou la moyenne géométrique (afin de valoriser une intensité nulle pour viser un rejet) de toutes les intensités d'implication des arêtes. Cependant, cette méthode simple a le défaut de ne pas accorder à la mesure un caractère relatif comme nous l'obtenons pour l'intensité  $\varphi(a, b)$  qui est une probabilité.

Aussi, nous faisons appel à une troisième approche qui a la vertu épistémologique de s'harmoniser avec la philosophie de la modélisation employée pour évaluer la qualité d'une quasi-implication. Comme l'intensité d'implication, elle présente aussi l'avantage de fournir une échelle de mesure (une probabilité, nombre de  $[0, 1]$ ), c'est-à-dire une valeur ordinale où le « zéro » sera signe d'une hétérogénéité maximale du cône et le « un » sera signe d'une excellente homogénéité dissymétrique ou d'un flux implicatif régulier et riche dans ses liaisons.

Associions à chacune des  $p$  arêtes une variable aléatoire relative à l'implication  $a \Rightarrow b$  ; on la note  $Q(a, \bar{b})$  (formule (2) du §2.1). Elle conduit, nous l'avons vu, à une mesure de la qualité de l'implication entre les deux extrémités de l'arête  $a \rightarrow b$ . Notons de façon générique,  $Q_i$ ,  $i = 1$  à  $p$ , l'indice d'implication aléatoire dont la réalisation est  $q_i$ . Nous savons que les  $p$  variables  $Q_i$ , supposées indépendantes, suivent approximativement la loi normale  $N(0, 1)$ . Or, ces  $p$  variables indépendantes, suivant la même loi satisfont, en conséquence, les hypothèses du théorème central limite.

Posons  $S_p = \sum_1^p Q_i$  et sa forme centrée-réduite  $T_p = \frac{S_p - E(S_p)}{\sigma(S_p)}$  où  $E(S_p)$  est l'espérance de  $S_p$ , somme des espérances des variables  $Q_i$  qui sont toutes nulles et  $\sigma(S_p)$  est l'écart-type de  $S_p$ , soit  $\sqrt{p}$ .

Par suite,  $T_p = \frac{S_p}{\sqrt{p}}$  converge en loi vers la variable aléatoire de loi  $N(0,1)$ . Pour une valeur de  $p$  convenable, par exemple  $p \geq 4$ <sup>4</sup>,  $T_p$  est donc approximable par la variable gaussienne centrée réduite  $N(0,1)$ .

Comparant cette valeur attendue à la valeur observée  $t_p = \frac{\sum_1^p q_i}{\sqrt{p}}$ , nous obtenons alors un indice probabiliste pour mesurer, sur son versant d'entrée, la qualité du cône à savoir :  $\Pr[T_p \geq t_p] = \frac{1}{\sqrt{2\pi}} \int_{t_p}^{\infty} e^{-\frac{u^2}{2}} du$

### 3 Recherche des conditions de causalités sans liaison entre elles

Si l'on veut identifier les « pères » du sommet du cône comme étant des causes ne dérivant pas de phénomènes causaux communs ou proches, comme le seraient par exemple « il pleut » et « le sol est mouillé », il nous faut exiger que ces variables amont ne soient pas liées ou ne le soient que faiblement. Si tel est le cas, on pourra affirmer que telle variable (par ex. le sommet du cône en l'occurrence) dépend essentiellement de telles causes séparées, tout au moins dans les conditions expérimentales sources des données traitées. Ces données qu'elles soient quantitatives ou qualitatives, pour un traitement avec CHIC, sont numériques ou binaires ou décimales. Envisager une étude la corrélation linéaire de Bravais-Pearson est donc justifié puisque le rho de Spearman n'aurait été indiqué que si nous avions conservés les rangs des valeurs des variables. Or, dans l'exemple initial, la relation d'ordre des variables ordinales –des préférences– a été commuée en relation d'ordre sur  $R$ . Aussi, nous écartons le rho de Spearman.

Un test de significativité du coefficient de corrélation linéaire peut alors être joint à l'analyse interprétative pour réfuter la liaison entre deux variables amont. Par exemple, utilisant ce test, si le nombre de sujets est 150, on dira que la liaison est significative au risque de 5% si on observe une corrélation  $r$  supérieure à 0.16. Mais si  $r = 0.10$ , on admettra l'absence de liaison entre les deux variables considérées. Il est possible aussi de pratiquer un test de corrélation partielle<sup>5</sup> entre  $d$  variables « pères ». Pour  $d=5$ , le risque de 5% est atteint pour une corrélation partielle de 0.164.

Il en est de même pour les « fils » en aval du sommet qui seront désignées comme conséquences non liées entre elles de ce sommet.

## 4 Retour sur l'exemple initial

### 4.1 Les données

---

<sup>4</sup> Pour un nombre d'arêtes plus petit, l'indice envisagé risque de n'avoir qu'un rôle indicateur de tendance.

<sup>5</sup> Rappelons que la corrélation partielle se calcule de façon récurrente.

Par exemple, si  $d=3$  :  $r_{1,2,3} = \frac{r_{1,2} - r_{1,3} r_{2,3}}{\sqrt{(1-r_{1,3}^2)(1-r_{2,3}^2)}}$ . Pour obtenir  $r_{1,2,3,4}$ , on remplace dans la formule précédente les corrélations simples par les corrélations partielles.

Reprenons l'exemple du § 1 Fig. 1. Ce cône est obtenu par extraction au sein d'un graphe implicatif de données issues d'un questionnaire présenté à des professeurs de mathématiques de classes terminales de lycée (17-18 ans). Nous avons recueilli et analysé (Bodin et Gras., 1999) les réponses de 311 professeurs, à des classements attendus (de 1 à 6) portant sur quinze objectifs que ces professeurs assignent à leur enseignement (A, B, C, ...O)<sup>6</sup> et sur leurs opinions relatives à dix phrases susceptibles d'être communément énoncées (OP1, OP2,..OPX)<sup>7</sup>. Les 26 variables correspondantes ne sont pas binaires, mais ordinales (valeurs (1, 0.8, 0.6, 0.4, 0.2, 0.1, 0) pour les objectifs et (1, 0.5, 0) pour les opinions). Ainsi l'analyse intègre l'intensité des attitudes, d'un choix prioritaire d'un objectif à un choix plus secondaire, voire non retenu.

La consigne était la suivante :

« A votre avis, quels sont les objectifs essentiels de la mission d'un professeur de mathématiques dans la série pour laquelle vous répondez. Pour répondre à cette question, classez par ordre préférentiel décroissant de 1 à 6 (1 : le plus important,...) six des objectifs majeurs de cette formation en les choisissant parmi les objectifs proposés ci-dessous :

**A-** acquisition de connaissances

.....

**E-** développement de l'imagination et la créativité

**F-** développement de la capacité à prouver et valider sa preuve.....

**I-** développement de l'esprit critique.....

**M-** développement de la capacité à mathématiser et à formaliser.....

Puis, Voici quelques opinions entendues dans la salle des profs. Vous pouvez être d'accord, ou un peu d'accord ou pas d'accord avec l'une ou l'autre. Entourez votre choix :

**1-(OP1)** C'est vrai que les math constituent un instrument de sélection excessif.

D'ACCORD            UN PEU D'ACCORD            PAS D'ACCORD

.....

A la sortie de la terminale de la série sur laquelle vous répondez, un élève devrait....

**7-(OP7)** ....pouvoir reconnaître si un nombre entier écrit dans la base 10 est divisible par 4.

D'ACCORD            UN PEU D'ACCORD            PAS D'ACCORD

**8-(OP8)** ....pouvoir donner un exemple ou un contre-exemple personnels à l'affirmation :

"si deux applications  $f$  et  $g$  sont strictement croissantes sur un intervalle, l'application produit  $fxg$  y est également croissante".

D'ACCORD            UN PEU D'ACCORD            PAS D'ACCORD

---

<sup>6</sup> Par exemple, E symbolise l'objectif : « développement de l'imagination et de la créativité »

<sup>7</sup> Par exemple, OP4 symbolise : « Pour corriger, j'aime bien un barème très détaillé sur les résultats à obtenir »

## 4.2 L'homogénéité du cône

Le logiciel CHIC nous fournit les valeurs des intensités d'implication des variables en jeu dans le graphe de la Fig. 1. De ces valeurs de l'intensité d'implication qui sont des probabilités donc des nombres de  $[0,1]$  affichées par CHIC, nous pouvons remonter à la valeur  $q_i$  prise par  $Q_i$  où  $i$  est l'un des couples qui nous intéressent du cône implicatif. Il suffit de lire la table de la loi normale  $N(0,1)$  de façon inverse.

Règle $i$	Intensité d'implication $\varphi_i$	Indice d'implication observé $q_i$
F -> OP8	0.96	- 1.75
M -> OP8	0.86	- 1.08
I -> OP8	0.93	- 1.48
E -> OP8	0.91	- 1.34
F -> OP7	0.81	- 0.88
M -> OP7	0.61	- 0.28
I -> OP7	0.75	- 0.68
E -> OP7	0.81	- 0.88
OP8 -> OP7	1	- 4.00

Appliquons les relations établies dans le 2.2 pour  $p = 9$ ,  $t_9 = \frac{\sum_1^9 q_i}{\sqrt{9}} = - 4.13$

$$D'où \Pr[T_9 \geq t_9] = \frac{1}{\sqrt{2\pi}} \int_{-4.13}^{\infty} e^{-\frac{u^2}{2}} du \cong 1.$$

Ceci signifie que la qualité d'homogénéité sémantique et dissymétrique du cône, eu égard à notre modèle aléatoire, est excellente puisqu'une distribution au hasard des indices d'implication conduirait à des sommes de  $Q_i$  bien supérieures à celle qui est observée à savoir  $- 4.13$ . Autrement dit, le « flux implicatif » qui traverse le cône respecte bien la dissymétrie attendue de l'implication et cela dans le même sens puisque les valeurs des  $q_i$  sont toutes négatives.

Remarquons, en passant, que la liaison  $OP8 \Rightarrow OP7$  donne du sens à  $OP7$  comme engageant la réflexion personnelle et critique de l'élève et non pas l'application d'un critère appris en cours puisque c'est précisément ce qu'exprime l'objectif  $OP8$ . La compétence de l'élève qui lui est associée apparaît comme « émergente » de la conjugaison des objectifs qui seraient réalisés en amont. Cette interprétation est renforcée par la nature de  $I$ , une des « causes » amont où l'esprit critique est explicitement visé au lycée. Cette remarque illustre le propos tenu dans l'avant-dernier paragraphe du § 1.

## 4.3 Les liaisons entre les « causes » et entre les « conséquences »

Dans cet exemple, seules les éventuelles « causes » sont plurielles : F, M, I et E. Mais nous n'avons qu'une conséquence : OP7. Les coefficients de corrélation linéaire - à la limite d'applicabilité en raison des faibles effectifs en jeu- entre les variables « pères » sont les suivantes (toujours données par CHIC) :

$$r(F ; M) = 0.13 ; r(F ; I) = - 0.04 ; r(F ; E) = 0.1 ;$$

$$r(M ; I) = - 0.1 ; r(M ; E) = - 0.07 ; r(I ; E) = 0.15$$

Faisant référence à la table des valeurs critiques du coefficient de corrélation d'un échantillon issu d'une population normale, nous constatons qu'un coefficient au hasard R, pour tous les couples en jeu, dans le cas d'une liaison entre ses termes, est supérieur à 0.13 avec un risque de 5%. Ceci est le cas pour (F et M) et (I et E), mais non pour les autres couples dont l'affirmation d'une liaison conduirait à un risque supérieur à 10%. On peut donc faire l'hypothèse d'ordre cognitif que deux causes fondamentales se conjuguent pour impliquer OP8 puis OP7 : l'un de type non normatif dans le cadre scolaire (imagination E, esprit critique I), l'autre de type cartésien (preuve F, formalisation M), à l'intérieur desquels des modalités seraient liées. L'examen des tâches nécessaires à la résolution de OP8 (exemple, contre-exemple, concepts mathématiques) montre que celle-ci nécessite la mise en commun des deux types causaux identifiés en tant que « pères ».

Cependant, l'examen des corrélations partielles montre que admettre l'absence de liaison globale reste légitime et donc que les « causes » présentent des caractères spécifiques non inclusives. En effet, dans les limites de seuil de 5% (0.164) et même à 10% (0.137), on obtient des corrélations partielles en permettant la validation :

$$r(F, M, I) = 0.127 ; r(M, I, E) = -0.092 ; r(F, I, E) = - 0.056 ; r(F, M, E) = 0.143 ;$$

$$\text{et finalement } r(F, M, I, E) = 0.131.$$

## **5 Un deuxième exemple**

### **Relations et Co variations**

D'un point de vue statistique la problématique principale que nous abordons est celle de la définition des liens statistiques entre variables lorsqu'ils sont déduits des mises en relation de leurs variations. Il s'agit de donner un sens statistique aux questions suivantes : si les valeurs d'une des variables différencient deux groupes de sujets (si ces groupes de sujets sont affectés par un changement de valeurs de cette variable ou par une modification importante de ces valeurs) en est-il de même pour l'autre variable? Et ces changements sont-ils comparables en importance? Sont-ils significatifs ou négligeables? etc.

Nous proposons de différencier les réponses à ces questions selon un critère particulier, celui de la symétrie exigée ou non des mises en relation des variations des variables. Ainsi nous distinguons les liens statistiques dont la définition inclut une exigence de réciprocité entre les variations des variables intéressées, et ceux dont la définition n'accorde pas le même statut aux variables considérées : les variations de la seconde (disons « impliquée » ou « seconde » ou « fils ») ne sont étudiées qu'au regard des variations de la première (disons le « père » ou l'« initiale »).

Ainsi nous comparons des traitements « classiques » qui permettent d'établir des relations symétriques entre variables tels ceux que le coefficient de corrélation linéaire  $r$  permet (ou le  $\chi^2$ ) aux traitements issus de l'ASI, qui permettent au contraire d'établir des dissymétries.

Ces problématiques et les décisions que nous venons d'évoquer sont inscrites dans le champ de la statistique. Notre projet est de montrer ici l'intérêt qu'elles peuvent présenter dans le champ des didactiques disciplinaires ou plus largement, dans celui des sciences de l'éducation.

Pour ce faire, il est nécessaire de considérer l'interprétation ou les points de la modélisation qui posent problème d'un point de vue didactique.

Il a été dit plus haut que la problématique des liens entre variables reposait sur l'identification de modifications, de changements de valeurs d'une des variables sur un *groupe* de sujets. Par conséquent, les changements de valeurs à étudier sont plutôt à considérer du point de vue d'un groupe de sujets et non pas d'un individu isolé.

Ceci nous amène à interroger les groupes pertinents pour les études didactiques. Nous envisageons ici quelques-uns des groupes pertinents du point de vue didactique. Pour que cette pertinence soit reconnue, nous ne relevons que des groupes relativement stables afin que leurs frontières soient identifiables ainsi que leurs variations. Il nous semble que parmi ces groupes font consensus les groupes classes, les groupes de niveaux scolaires, les groupes de classes disciplinaires (nous pensons par exemple aux groupes d'élèves pour les enseignements de langues). Il nous semble également que sont moins consensuels, au regard toujours de leur pertinence pour les didactiques, les groupes identifiés par des choix pédagogiques revendiqués des enseignants ou encore les regroupements d'élèves selon leur « catégorie sociale »<sup>8</sup>.

C'est pourquoi nous avons choisi dans ce cadre, de présenter une étude dans laquelle ces derniers regroupements seront l'objet d'une attention particulière.

## **5.1 La recherche « conscience disciplinaire »**

Nous disposons d'un vaste corpus, construit à partir de la passation de trois questionnaires. Ces questionnaires ont été conçus et analysés dans le cadre d'une recherche en didactiques, dirigée par Cora Cohen-Azria, qui a donné depuis lieu à un ouvrage collectif. Cette recherche a pour but de questionner ce que nous avons appelé la « conscience disciplinaire » d'élèves de CM1 et CM2. Pour le dire très rapidement, nous appelons « conscience disciplinaire » les façons dont chaque élève identifie (ou ignore) les disciplines scolaires, les façons dont il se munit pour identifier ces différents espaces disciplinaires, les façons dont il se reconnaît sujet de ces espaces et y agit. Ainsi par exemple, les façons de reconnaître dans quelle discipline on travaille – ce qui n'est pas toujours aussi évident que l'on peut le croire – ce qui y est valorisé et ce qui y est dénigré, ce qui y est possible, ce qui ne l'est pas, sont autant de dimensions de cette conscience disciplinaire : déclarer qu'en Français « ce que le prof attend c'est surtout du bla bla bla » tandis qu'en Mathématiques « c'est surtout un raisonnement correct »,

---

<sup>8</sup> Guillemets, car ce sont plutôt celles de leur famille, avec tous les risques que cela comporte quant à l'interprétation du terme famille.

est une trace de la conscience disciplinaire du Français et des Mathématiques (en tant que disciplines scolaires) du lycéen interrogé. Durant cette recherche, nous avons tenté d'identifier certaines dimensions de ces consciences disciplinaires, dans quatre disciplines : le Français et les Mathématiques (disciplines scolaires reconnues comme « très importantes » et « très présentes »), les Sciences et l'Histoire-Géographie (qui le sont sans doute un peu moins). Nous avons choisi de les étudier auprès d'élèves de l'école primaire, pour plusieurs raisons : celle tout d'abord de l'organisation institutionnelle, en France, des disciplines scolaires, qui varie fortement entre l'école maternelle, l'école primaire, le collège et les divers types de lycées (nous mettons en note quelques différences). Les autres raisons tiennent à l'ancrage du travail didactique de notre équipe de recherche (Théodile CIREL) dans l'exploration des effets des modes de pédagogies alternatives (références) ainsi que dans l'intérêt que nous avons vu à explorer ces degrés de conscience disciplinaire à un moment du curriculum « charnière ».

C'est pourquoi ces questionnaires ont été soumis à quelques trois cents élèves de CM1 et CM2. Ces trois questionnaires les interrogent sur les façons dont ils dénomment, identifient, re-construisent... les disciplines scolaires choisies: le Français, les Mathématiques, les Sciences et l'Histoire-Géographie (en note, nous expliquons que quelques questions plus globales sont importantes aussi). Ils sont pratiquement uniquement constitués de questions ouvertes.

Nous nous focalisons ici sur les variables issues de l'étude de quatre questions :

- (1) Qu'est-ce que tu as appris d'important en Mathématiques ? (Questionnaire 1)
- (2) Qu'est-ce que tu as appris d'important en Français ? (Q1)
- (3) Qu'est-ce qui est important pour réussir en Mathématiques ? (Q3)
- (4) Qu'est-ce qui est important pour réussir en Français ? (Q3)

Nous avons relevé les expressions/ locutions (terme à fixer) des réponses des élèves (présents à Q2 et à Q3). Et nous avons défini ainsi un ensemble de 51 variables binaires : elles valent 0 ou 1 selon que cette expression/locution a été employée par l'élève dans la 1<sup>e</sup>, 2<sup>e</sup>, 3<sup>e</sup> ou 4<sup>e</sup> question. Par exemple la variable FIOortho correspond à la question (2) (F pour Français, I pour Important Appris) et à l'usage du terme « orthographe » dans la réponse de l'élève.

## **5.2 L'importance du flux implicatif et de la dissymétrie**

Nous avons fait passer ces questionnaires dans des établissements implantés dans des quartiers « socialement défavorisés », et dans d'autres qui le sont au contraire dans des zones « socialement favorisées ».

Et de fait, savoir s'il est possible, à partir des origines ou des milieux sociaux des élèves d'induire des conséquences quant à leurs représentations de l'école et des disciplines scolaires est une question très étudiée et particulièrement polémique. Les discours d'opinion abondent sur ce problème récurrent qu'ils tranchent dans un sens ou dans un autre. Du point de vue scientifique, en sociologie de l'éducation, dans les différentes didactiques (et peut-être plus particulièrement en didactique du français), les références sont nombreuses également (voir les travaux du groupe ESCOL, ceux de Lahire...). Formulons autrement notre question : est-ce que les expériences sociales

possibles et effectives des *enfants* sont déterminantes pour leurs compréhensions, lectures de l'organisation de l'école où ils sont *élèves* ? Cohen-Azria, Lahanier-Reuter et Reuter, 2013).

Reprenons les termes que nous venons d'utiliser : « induire des origines sociales des conséquences sur les représentations ». Ceci recouvre quelques questions suivantes :

- est-ce que le type de quartier est considéré dans cette question comme un facteur/déclencheur de certaines dimensions des représentations? (corrélée ou « cause », « père » dans le schéma implicatif à des locutions/termes qui seraient distincts : par exemple, « être dans un milieu défavorisé » (est corrélé à ) implique que/ a pour conséquence « dire que pour réussir en maths, il faut savoir compter » et « pour réussir en Français il faut bien écouter », tandis que « être dans un milieu favorisé » (est corrélé à ) implique que/ a pour conséquence « dire que pour réussir en maths, il faut faire des exercices, s'entraîner » et « pour réussir en Français il faut bien écouter »). Le type d'établissement est alors « antécédent », « père » de certaines dimensions.
- est-ce qu'à un type d'établissement correspond un certain type de représentation ? (presque une fonction, donc des relations corrélées/implicatives comme ci-dessus, mais l'exigence de ne relever aucun terme/locution partagée par des élèves de milieux différents./pas de consensus possible, pas d'attitude commune entre ceux qui appartiennent à des types d'établissements différents).
- est-ce qu'un type d'établissement est particularisé/singularisé par plusieurs dimensions des représentations ? Ce qui est assez différent de ce qui précède. Dans ce cas, ce sont certaines dimensions qui seront propres à des élèves de ce type d'établissements : elles ne seront pas communes à tous les élèves de ces derniers, mais ne se relèveront que dans ceux-ci (tout cela doit être tempéré par « statistiquement »). Par exemple, le fait de reconnaître la lecture comme moyen de réussite en français est un discours que seuls tiennent des élèves des établissements d'un type particulier. Pour l'analyse que nous menons, ces dimensions apparaîtront alors comme des « pères », des antécédents des types d'établissement. La situation est l'inverse de la première.

Nous avons donc interrogé les discours des élèves qui ont été recueillis sous cet angle : les caractéristiques de ces discours, autour des contenus importants enseignés en Mathématiques et en Français et autour des contenus importants pour réussir dans ces deux disciplines, sont confrontées aux inscriptions des établissements scolaires qu'ils fréquentent, dans des quartiers « socialement favorisés » ou au contraire « socialement défavorisés ».

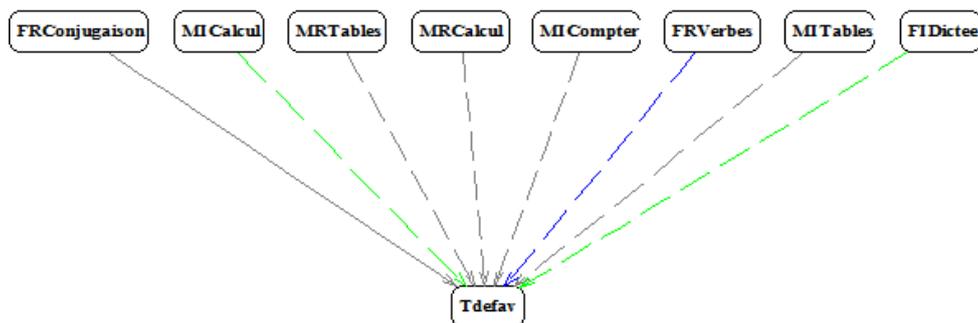
A nos variables précédentes s'en ajoutent donc deux : « Favorisé » et « Très Défavorisé », qui sont encore des variables binaires. Les cônes implicatifs sont centrés tour à tour sur ces deux variables.

### **5.2.1 Les établissements situés dans des quartiers « très défavorisés »**

#### **Les facteurs « pères »**

Commençons par le cône implicatif qui apparaît sous CHIC au seuil .60 centré sur la variable «Très Défavorisé ».

Figure 2 : Cône implicatif au seuil .60 centré sur la variable « Très défavorisé »



Nous constatons que le fait d’être inscrit dans un établissement (public) situé dans un quartier «Très défavorisé » n’a **aucune conséquence implicative**, mais apparaît comme une « conséquence » implicative.

Il semble donc que dans les établissements étudiés, situés dans des quartiers très défavorisés, *on puisse identifier des discours d’élèves qui sont propres à ces établissements.*

- Les contenus importants en Mathématiques sont compter/calculer/les tables de multiplication (MICompter, MICalcul, MITables);
- Il est important pour réussir en Mathématiques de [savoir] les tables de multiplication (MRTables) ;
- Les contenus importants en Français sont [faire] des dictées (FIDictées);
- Il est important pour réussir en Français de [savoir] les verbes et/ou les conjugaisons (FRVerbes, FRConjugaison).

Ces discours, spécifiques (au sens statistique du terme) des établissements étudiés situés dans des quartiers très défavorisés, sont cohérents. Ils mettent en évidence des constructions des deux disciplines essentiellement centrées sur des contenus :

- emblématiques de l’école primaire (compter, calculer, les tables de multiplication ou les dictées, les verbes et les conjugaisons) ;
- ancrés dans les représentations communes ;
- utiles ;
- gérables/contrôlables par les parents ou l’entourage ;

Pour le dire très brièvement, ces traits sont ceux d’une conscience disciplinaire qui serait celle du « certificat d’études », c’est-à-dire d’un enseignement non basé sur des sous-disciplines, destinés à des enfants dans lesquels on lit l’adulte, les futurs travailleurs et non pas de futurs collégiens par exemple.

### Les corrélations partielles

Si nous considérons à présent les corrélations partielles entre ces différents facteurs (voir Annexe), seules les variables (MRCalcul, MRTables) et (FRConjugaison,

FRVerbes) peuvent être considérées comme couples de variables dépendantes au seuil de risque de 5%. Remarquons tout d'abord que ces couples de variables concernent des réponses distinctes aux deux questions :

qu'est-ce qui est nécessaire selon toi pour réussir en Mathématiques ?

qu'est-ce qui est nécessaire selon toi pour réussir en Français ?

Il n'y a pas trace de corrélation entre des réponses concernant des disciplines scolaires différentes, ce qui conforte notre hypothèse de conscience disciplinaire, c'est-à-dire du poids des disciplines scolaires dans les « attitudes » des élèves.

Plus précisément, les deux conditions à la réussite en Mathématiques telles qu'elles sont principalement énoncées par des élèves des établissements très défavorisés (la maîtrise du calcul *vs* la connaissance des tables de [multiplication]) seraient des conditions exclusives<sup>9</sup> ( $r = -0,23$ ). Cette tendance à l'exclusion peut être interprétée comme l'indice de deux modes d'appréhension de la réussite, de ce qui est attendu par l'enseignant, de ce qui est évalué en mathématiques par ces élèves. Y aurait-il une différence entre concevoir les Mathématiques comme l'art de calculer (au sens de « jongler avec les chiffres) et les résumer par un « slogan » (il faut savoir [ses] tables) ? Ou n'est-ce là qu'une caractéristique du discours des élèves, qui en quelque sorte « choisissent » l'un des termes (calculer ou les tables) de façon tendanciellement exclusive tant ils leur paraissent synonymes ou redondants ?

Il est plus simple nous semble-t-il d'interpréter la corrélation positive entre les deux facteurs de réussite en Français, [maîtriser] les conjugaisons et [connaître/savoir] les verbes, qui sont en effet parfois rapprochés dans le discours des élèves. Certains d'entre eux choisissent donc de considérer ces deux termes comme non redondants, non synonymes, mais plutôt sans doute comme complémentaires.

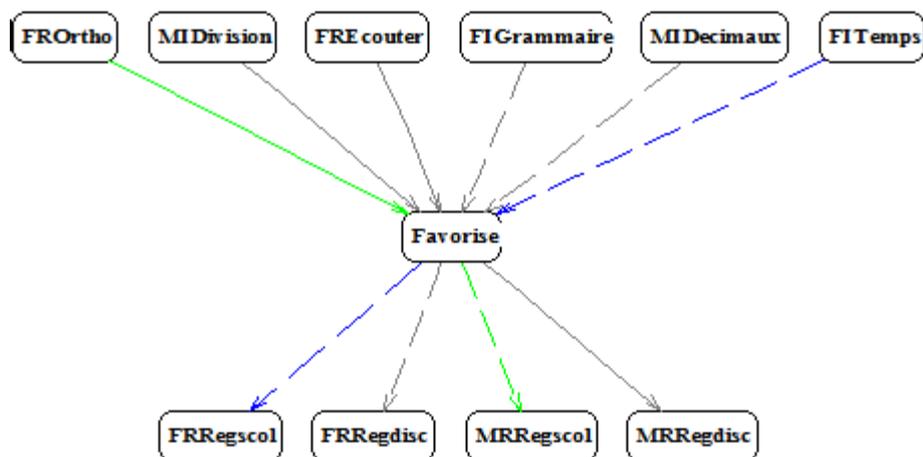
### **5.2.2 Les établissements situés dans des quartiers favorisés**

Considérons à présent les établissements situés dans des quartiers considérés comme plutôt favorisés.

Figure 2 : Cône implicatif à .60 centré sur la variable « Favorisé »

---

<sup>9</sup> Ou des façons de dire exclusives plutôt.



Cette fois, les établissements « favorisés », centre d'un bassin d'attraction, sont à la fois des « pères » pour certaines dimensions et des « fils » pour d'autres (on retrouve la relation dialectique où la cause détermine l'effet et réciproquement, voir page 2).

Commençons par les discours qui sont ceux d'élèves de ces établissements, et d'eux (statistiquement) uniquement :

Les contenus importants en Mathématiques sont la division et les nombres décimaux (MIDivision, MIDécimaux) ;

- Pour réussir en mathématiques... (il n'y en a pas de cités) ;
- Les contenus importants en Français, c'est la grammaire et les temps (FIGrammaire, FITemps) ;
- Pour réussir en Français, [il faut bien] écouter (FREcouter).

Par conséquent, les contenus qui sont désignés comme importants en Mathématiques sont des contenus qui sont *nouveaux* (au CM1 et au CM2), la division et les nombres décimaux. La perception de cette nouveauté présuppose son ancrage dans un apprentissage continu, ce qui particularise sans doute la conscience disciplinaire de ces élèves en ce qui concerne les Mathématiques. Ceux qui le sont en Français sont pour l'un, une sous-discipline (la grammaire) pour l'autre une désignation (scolaire?), les temps.

Les discours de ces élèves d'établissements (publics) « favorisés » privilégient parfois des contenus disciplinaires qui portent la marque de l'école: apprendre des temps (peu employés parfois hors de l'école, comme le passé antérieur), la grammaire, la division (qui est vraiment un savoir nouveau et complexe), les nombres décimaux (qui ne sont pas toujours les plus usités à l'extérieur de cette école, pensons par exemple à l'écriture 0,9764). Nous avançons, avec beaucoup de prudence, comme précédemment, l'idée que ces discours sont les traces de consciences disciplinaires qui seraient de l'ordre de l'apprentissage de contenus académiques, dont la maîtrise inscrit son possesseur dans une trajectoire/curriculum.

Cette fois, deux variables seulement sont corrélées positivement, au seuil de risque de 5% : Le fait de citer les décimaux comme contenus importants vus en Mathématiques et celui de citer les Temps comme contenus importants vus en Français : une vision globale des savoirs disciplinaires conforme à celle des programmes ?

### **5.2.3 Comparaison**

Si nous comparons les « antécédents » mis au jour des deux types d'établissements (favorisés et très défavorisés), il nous semble qu'ils montrent des différences sensibles. Identifier des sous-disciplines et non pas des objets, identifier les savoirs nouveaux comme importants, ceux qui sont bien évalués (la maîtrise de la division) et non pas ceux qui sont intéressants à maîtriser (bien compter, savoir ses tables de multiplication) et que l'entourage ou l'opinion commune désigne comme tels seraient des positions (?) qui permettraient de parier à presque coup sûr sur l'appartenance à un établissement inscrit dans un quartier « favorisé ».

En ce qui concerne les « conséquences » de l'inscription dans un établissement « favorisé », elles présentent des aspects similaires, qui concernent les conditions de réussites :

- Que ce soit en Français ou en Mathématiques, les élèves de ces établissements s'entendent (statistiquement toujours) sur le fait que l'on réussit en obéissant à des règles scolaires : [bien] écouter/ ne pas faire de bêtises/faire ce que le maître/la maîtresse dit de faire etc.
- Que ce soit en Français ou en Mathématiques, les élèves de ces établissements s'entendent (statistiquement toujours) sur le fait que l'on réussit en obéissant à des règles cette fois disciplinaires : en Français en consultant son Bled, en Mathématiques en tenant bien droit sa règle pour tracer, etc.

Sur le corpus dont nous disposons, cette analyse confirme des résultats déjà établis : il y a des modes d'appréhension variables de l'école, ici de la structuration des enseignements scolaires en enseignements disciplinaires. Et ces variations, de ce qui est pour nous des traces des consciences disciplinaires, sont réglées (soumises ?)/ expliquées par les variations des milieux sociaux dominants des quartiers dans lesquels les établissements scolaires sont implantés. C'est le respect formel des règles qui semble dériver de l'appartenance au milieu favorisé : une institutionnalisation acceptée ?

Mais il nous est aussi possible, par le recours à ce type d'analyse, de détruire (ou de reconsidérer ?) la symétrie des relations (statistiques) entre dimensions de la conscience disciplinaire et type d'environnement social dans lequel les élèves évoluent. Ces relations (à une variation conséquente du milieu correspond une variation conséquente des dimensions reconstruites) comme nous venons de le voir, ne sont pas, loin de là, des relations simples de cause à effet. En montrant que certaines de ces dimensions sont des caractéristiques « partagées » tandis que d'autres semblent plutôt être le fait de « groupes d'élèves », l'ASI démontre aussi son intérêt.

Cependant, notre analyse n'est pas terminée. En effet, notre projet global est l'évaluation (dans un sens très large) les effets des modes de travail pédagogiques dans des quartiers populaires (voir Reuter etc.).

### 5.2.4 Les pédagogies dont se réclament les enseignants

Par conséquent, les variations les plus intéressantes pour nous sont celles des modes de travail pédagogiques : dans le corpus dont nous disposons, deux établissements (dans des quartiers « très défavorisés ») revendiquent plus ou moins l'inscription dans une démarche pédagogique particulière. Dans le troisième et dernier des établissements inscrits dans ce type de quartier, les enseignants en revanche, ne revendiquent pas de démarche pédagogique spécifique.

Voici un tableau descriptif des inscriptions pédagogiques déclarées par les enseignants et les codages :

Mode de travail pédagogique	Établissements <sup>10</sup>	Codages des classes correspondantes
Mouvement Freinet	Hélène Boucher (HB)	HB1
		HB2
Fonctionnement collectif : échanges de pratiques, de groupes de classe etc.	Pascal (Pa)	Pa1
		Pa2
Pas d'inscription particulière	Veermer (Ve)	Ve

Nous nous posons donc la question de savoir si aux différentes pratiques pédagogiques des maîtres correspondent des différences entre les discours des élèves et des [reconstructions de leurs] conscience disciplinaire.

L'analyse a été menée sur le même corpus, en centrant cette fois les cônes implicatifs sur les 5 classes de ces trois établissements.

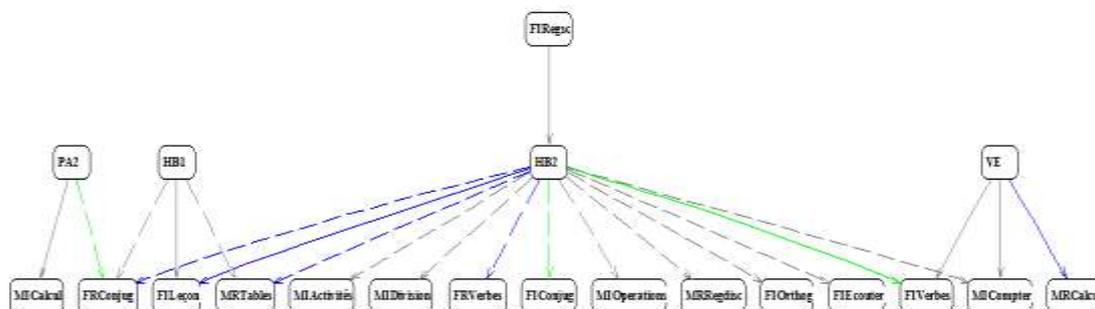


Figure 3

<sup>10</sup> Les noms ont été modifiés, sauf celui d'Hélène Boucher.

La particularité de la classe HB2 « saute aux yeux ». C'est tout d'abord celle où les discours sont les plus variés. C'est aussi celle où les discours des élèves sont originaux : ainsi la dimension « les contenus importants en Mathématiques sont les activités spécifiques telles les recherches mathématiques (MIActivités). Nous notons enfin que quelques-unes des « conséquences » c'est-à-dire quelques-unes des dimensions manifestées par l'appartenance à cette classe sont celles qui particulariseraient des discours des élèves des établissements situés dans un autre type de quartier : MRReglesdisciplinaires, MIDivision, FIOrthographe, Flécouter.

Nous ne creuserons pas davantage ici cette singularité, pour laisser la place à une question qui illustre la relation dialectique cause-effet. Cette classe HB2 nous est familière : elle est celle d'un enseignant qui est un des piliers de ce mouvement de l'ICEM, qui a été un des premiers enseignants de cette équipe<sup>11</sup>. Faut-il en conclure que ce dernier peut être qualifié d'expert ? (et, parallèlement, qualifier le second enseignant de l'établissement HB, l'enseignant de la classe HB2 de « novice » ?). Nous hésitons à attribuer ces étiquettes à des sujets : ne faut-il pas plutôt parler dans la classe HB2 de pratiques pédagogiques réellement alternatives ?

## 6 Conclusion

A la faveur de l'analyse non symétrique d'un corpus de données et à travers la représentation en graphe implicatif de l'ensemble des règles d'association obtenues, nous avons isolé un sous-graphe constitué d'une sorte de cône à deux nappes. Son intérêt réside dans la mise en évidence du rôle de bassin d'attraction que remplit le sommet du cône. Dans le cas où la nappe supérieure supposée causale est constituée de sommets indépendants, on dispose d'une confluence causale des pères du sommet comme effet résumé de plusieurs causes et, en un même temps dialectique, explicatif de celles-ci. De la même façon, l'examen des « fils » de ce sommet permet de donner un sens de conséquences éclatées de cet attracteur que joue le sommet. Plusieurs situations concrètes, dont l'une plus fouillée portant sur la notion de « conscience disciplinaire », ont permis d'illustrer ces démarches.

Il serait intéressant de se priver de l'indépendance des « causes » pour, au contraire, examiner en quoi le sommet pourrait relever de l'agglutination de ces causes qui ne seraient alors que des modalités faisant en quelque sorte « cause commune ». Cette autre problématique rejoindrait celle de cet article dans notre ambitieuse entreprise de l'élucidation des relations causales entre phénomènes car comme l'écrit Francis Bacon : « *Vraiment connaître, c'est connaître par les causes* » et par ailleurs : « *Savoir pour prévoir, prévoir pour pouvoir* ».

---

<sup>11</sup> Nous avons suivi cette expérience (une école entière fonctionnant en pédagogie Freinet depuis 2001. L'enseignant responsable de la classe HB1 est un jeune enseignant, qui a rejoint cette école depuis un an, mais qui s'inscrit lui aussi dans ce mouvement pédagogique.

## Références

- [1] Bailleul, M. & Godard, S. (2009), Derrière les réseaux de variables, il y a des individus... à écouter!, In Gras R. (dir.), Régnier J-C., Marinica C., Guillet F., *L'analyse statistique implicative, Méthode exploratoire et confirmatoire à la recherche de causalités*, Cépaduès Ed., Toulouse, 437-453.
- [2] Bodin, A. & Gras, R. (1999), Analyse du préquestionnaire enseignants avant EVAPM-Terminales, *Bulletin de l'Association des Professeurs de Mathématiques*, n° 425, 772-786.
- [3] Briand, H., Sebag, M., Gras, R. et Guillet F. (eds), (2004), *Mesures de Qualité pour la Fouille de Données*, RNTI-E-1, Cépaduès, Toulouse.
- [4] Cohen-Azria, C., Lahanier-Reuter, D., Reuter, Y. (2013), *La conscience disciplinaire*, P.U.R., Rennes.
- [5] Gras R. (dir.) Régnier J.-C., Marinica C., Guillet F (eds.), (2013), L'analyse statistique implicative, Méthode exploratoire et confirmatoire à la recherche de causalités, Cépaduès Editions, Toulouse.
- [6] Gras, R. (dir.), Almouloud, S. Ag, Bailleul, M., Larher, A., Polo, M., Ratsimba-Rajohn, H., Totohasina, A. (coll.), (1996), *L'implication statistique. Nouvelle méthode exploratoire de donnée*, La Pensée Sauvage, Grenoble.
- [7] Gras, R., Kuntz, P. et Briand, H. (2001), Les fondements de l'analyse statistique implicative et quelques prolongements pour la fouille de données, *Mathématiques et Sciences Humaines*, n° 154-155, 9-29.
- [8] Gras, R., Suzuki, E., Guillet, F., Spagnolo, F. (eds), (2008), *Statistical Implicative Analysis, Theory and Applications*, Springer.
- [9] Guillet, F. & Hamilton, H. (eds), (2007), *Quality Measures in Data Mining*, Springer.
- [10] Lahire, B. (1993), *Culture écrite et inégalités scolaires*, P.U.L., Lyon.
- [11] Lévy-Leblond, J.-M. (1996), *Aux contraires*, Gallimard, Paris.
- [12] Orus, P., Zemora, L., Gregori, P., (2009), *Universitat Teoria y Aplicaciones del Analisis Estadístico Implicativo*, Eds : Jaume-1, Castellon (Espagne).
- [13] Pasquier, D. & Gras, R. (2012), De l'intérêt de l'Analyse Statistique Implicative (A.S.I.) pour la recherche exploratoire en psychologie, *Psychologie Française*, Elsevier-Masson, 161-17.
- [14] Régnier, J-C., Bailleul, M., Gras, R. (eds.), (2012), *L'Analyse Statistique Implicative : de l'exploratoire au confirmatoire*, Université de Caen, Caen.
- [15] Sève, L. (2005), *Émergence, complexité et dialectique*, Odile Jacob, Paris.

## Annexes

Coefficients de corrélation (Figure 2)

VIII Colloque International –VIII International Conference  
A.S.I. Analyse Statistique Implicative —Statistical Implicative Analysis  
Radès (Tunisie) - Novembre 2015  
<http://sites.univ-lyon2.fr/ASIS/>

	MR Tables	FR Conjugaison	FR Verbes	MI Tables	MI Calcul	MI Compter	FI Dictées
MR Calcul	-0,23	0,07	0,08	-0,02	0,20	0,04	0,05
MRTables		0,14	0,19	0,06	0,01	0,02	0,11
FRConjugaison			0,28	0,05	0,09	0,02	0,08
FRVerbes				0,21	0,04	0,03	0,14
MITables					-0,06	-0,03	0,07
MICalcul						0,06	-0,05
MICompter							-0,04

Corrélations partielles

$$r(\text{MRC}, \text{MRT}, \text{FRC}) = \mathbf{-0,24} \quad r(\text{MRC}, \text{MRT}, \text{FRV}) = \mathbf{-0,25}$$

$$r(\text{MRC}, \text{FRC}, \text{FRV}) = 0,05 \quad r(\text{MRT}, \text{FRC}, \text{FRV}) = 0,09$$

Pour la Figure 3

	FREcouter	FROrtho	MIDivision	MIDécimaux	FITemps	FIGrammaire
FREcouter		-0,06	-0,02	0,02	0,12	0,11
FROrtho			0,05	0,2	0,09	0,18
MIDivision				0,09	0,22	0
MIDécimaux					0,32	0
FITemps						-0,08
FIGrammaire						

Cette fois, deux variables seulement sont corrélées positivement, au seuil de risque de 5% : Le fait de citer les décimaux comme contenus importants vus en Mathématiques et celui de citer les Temps comme contenus importants vus en Français : une vision globale des savoirs disciplinaires conforme à celle des programmes ?