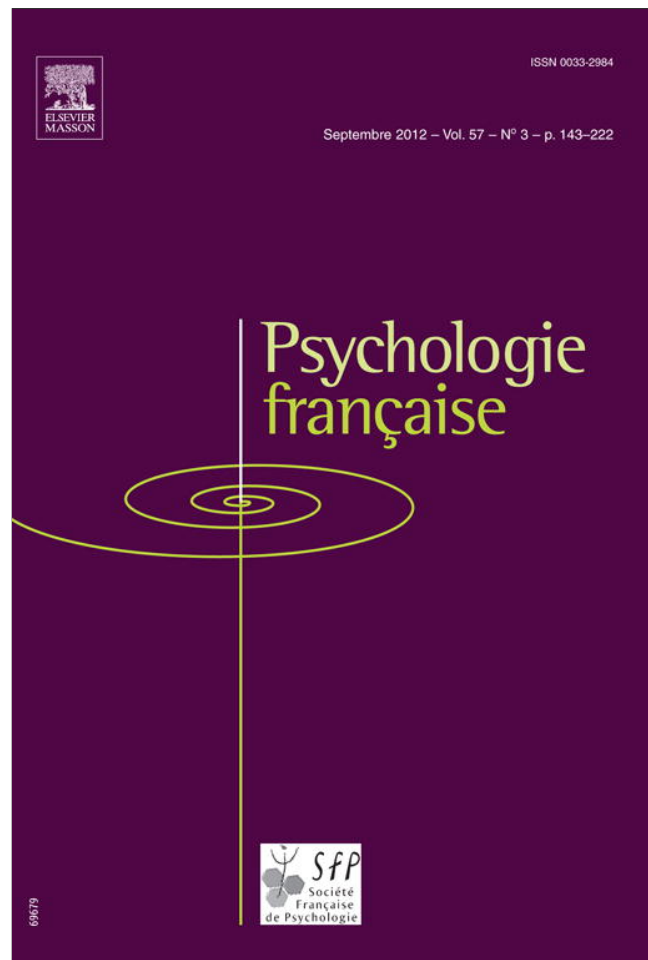


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



ELSEVIER
MASSON

Disponible en ligne sur www.sciencedirect.com

SciVerse ScienceDirect

et également disponible sur www.em-consulte.com



Article original

De l'intérêt de l'analyse statistique implicative (ASI) pour la recherche exploratoire en psychologie

In the interest of the statistical analysis implicative (ASI) for exploratory research in psychology

D. Pasquier^{a,*}, R. Gras^b

^a EA 2114, psychologie des âges de la vie, université François-Rabelais de Tours, 37000 Tours, France

^b Équipe connaissance et décision, laboratoire d'informatique de Nantes-Atlantique (LINA), UMR 6241, école polytechnique de l'université de Nantes, 44000 Nantes, France

INFORMATIONS

Historique de l'article :

Reçu le 28 août 2011

Accepté le 14 juin 2012

Mots clés :

Analyse statistique implicative

Recherche exploratoire

Approche corrélationnelle

RÉSUMÉ

Les psychologues utilisent fréquemment des approches corrélationnelles pour mener à bien l'analyse de leurs tableaux de données (corrélations, régressions, analyse en composantes principales...). De leur côté, les didacticiens pratiquent l'analyse statistique implicative (ASI). Dans ce papier, les auteurs présentent quelques arguments, illustrés d'exemples, en faveur d'une utilisation conjointe des deux approches méthodologiques. Si la corrélation permet d'identifier la polarité (le signe) et l'amplitude (la valeur absolue) des liens entre deux ou plusieurs variables, son aspect symétrique ne permet pas d'ordonner les variables. L'ASI complète l'approche corrélationnelle dans la mesure où la forme asymétrique de l'indice d'intensité de l'implication débouche sur une mise en ordre quasi-implicative des variables corrélées : $a \Rightarrow b$ « si la variable a est observée dans la population alors on note une propension à observer la variable b ». Reste à interpréter cet ordre implicatif, ce qui renvoie à une autre problématique.

© 2012 Société française de psychologie. Publié par Elsevier Masson SAS. Tous droits réservés.

* Auteur correspondant. 14, rue Aristide-Briand, 77370 Nangis, France.

Adresses e-mail dpasquier@avenireentreprise.fr (D. Pasquier), regisgra@club-internet.fr (R. Gras).

A B S T R A C T

Keywords:

Statistical implicative analysis
 Exploratory research
 Correlational approach

Psychologists frequently use correlational approaches to carry out the analysis of their data tables (correlations, regressions, principal component analysis...). In turn, didacticians practice implicative statistical analysis (ASI). In this paper, the authors present some arguments, illustrated with examples, for a joint use of two methodological approaches. If the correlation is used to identify the polarity (sign) and magnitude (absolute value) links between two or more variables, symmetrical appearance does not order the variables. The ASI full correlational approach since the asymmetrical shape of the index of intensity of involvement leads to an ordering of quasi-implicative correlated variables: $a \Rightarrow b$ "if the variable is observed in the population then there is a tendency to observe the variable b". It remains to interpret the order involved, which refers to another problem.

© 2012 Société française de psychologie. Published by Elsevier Masson SAS. All rights reserved.

1. Introduction

L'analyse statistique implicative (ASI) est une méthode d'analyse de données issue du croisement d'une population d'objets-sujets E et d'un ensemble de variables quelconques (booléennes, intervalles, floues, vectorielles. . .). À l'intersection d'une ligne quelconque et d'une colonne figure la valeur que prend un sujet-ligne selon une variable-colonne : par exemple, on écrit 1 si le sujet satisfait la variable booléenne et 0 dans le cas contraire. Cette méthode vise à attribuer une valeur d'intensité (de qualité) à des règles d'association implicative (non nécessairement strictes comme c'est en général le cas en sciences humaines et pour cela dites quasi-implicatives) du type : « Si la variable a est observée dans la population alors on note une propension à observer la variable b ». L'intensité, dite d'implication, que les variables a et b soient binaires ou numériques, varie de 0 à 1 et représente un critère de qualité de la règle. Plus elle est proche de 1, meilleure est la confiance en la qualité de la relation implicative. C'est en ce sens que nous dirons qu'une règle extraite par l'ASI est quasi-implicative par opposition à une règle stricte où une ou des prémisses entraînent, sans exception, une ou des conclusions ou conséquences.

Elaborée par Régis Gras (1979) puis développée par lui et ses collaborateurs (Gras et Larher, 1993 ; Gras 1996 ; Gras et al., 2001 ; Régnier et Gras, 2005 ; Gras et Kuntz, 2008 ; Gras et Régnier, 2009 ; Lahanier-Reuter, 2009 ; Verscheure et Chiocca, 2009), la méthode ASI fournit un cadre statistique permettant d'ordonner en des séquences quasi-implicatives des relations entre phénomènes. Sur le plan théorique, ce cadre s'appuie sur la probabilité de l'in vraisemblance de la liaison entre les variables dans une hypothèse d'indépendance a priori entre celles-ci. Qualité de la règle donnée par l'intensité et probabilité se conjuguent dans des domaines sémantico-conceptuels différents : le premier de l'ordre psychologique (confiance), le second de l'ordre mathématique (probabilités). On peut situer l'implication statistique par rapport à d'autres formes d'implication, logique ou mathématique. Contrairement aux tests d'hypothèse, bien que la valeur de l'intensité d'implication soit une probabilité – celle que le hasard conduise à plus de contre-exemples que la contingence – elle ne définit pas un seuil de réfutation d'une hypothèse. C'est une mesure de qualité de la règle comme il est habituellement pratiqué en analyse de données. Le chercheur l'apprécie de façon relative. Si la règle R1 admet pour intensité 0,92 sa qualité est meilleure que celle de la règle R2 d'intensité 0,89. Des simulations de l'intensité d'implication en fonction du cardinal n de la population montrent que des valeurs supérieures à 0,80 sont exceptionnelles pour n de l'ordre de 100.

À partir de quel seuil faut-il considérer qu'un indice d'intensité d'implication est statistiquement significatif ? La valeur 0,50 est évidemment celle à partir de laquelle la tendance implicative agit dans le sens de a vers b . Dans une optique simplement descriptive, tout indice supérieur à 0,50 peut donc être pris en compte et se révéler intéressant soit comme mise en évidence d'une niche implicative, d'une

« pépite de connaissance », soit comme source d'hypothèse : on dit alors qu'une connaissance extraite par l'ASI est vraie jusqu'à preuve du contraire. En deçà de 0,50, ce serait la tendance de la réciproque ($b \Rightarrow a$) qui pourrait éventuellement être retenue. Si $n(a)$ est inférieur ou égal à $n(b)$, le logiciel CHIC choisit de calculer l'intensité d'implication $b \Rightarrow a$, dont il est montré (Gras, 1996) qu'elle est plus forte que celle de $a \Rightarrow b$. On montre également que l'intensité de la contraposée non $b \Rightarrow$ non a est la même que celle de l'implication directe $a \Rightarrow b$. Il peut arriver que les indices d'implication entre deux variables soient tous deux supérieurs à 0,50 et très proches l'un de l'autre (par exemple 0,65 et 0,63), ce qui peut faire hésiter entre une conclusion d'équivalence des deux variables plutôt qu'une implication. Ce point reste à développer, par exemple sous la forme d'un test statistique de différence entre deux indices d'implication. Pour l'instant on ne pourra que dire qu'il existe une tendance positive à la validité de la règle établie à partir de l'indice le plus élevé. On peut aussi avoir recours à la méthode entropique pour lever le doute, méthode qui prend en compte la qualité de la contraposée de l'implication directe, mais méthode plus sévère en matière d'intensité d'implication (Gras et al., 1997).

2. Spécificité de l'implication statistique

L'implication logique entre deux propositions (vraies ou fausses) ou deux variables binaires a et b (égales à 0 ou 1) se définit à partir de la table de vérité des deux variables. Il y a implication « a entraîne b » dans tous les cas, sauf si a est vraie (ou égale à 1) alors que b est fausse (ou égale à 0).

En mathématique, la proposition a implique la proposition b (i.e. le théorème « a implique b » est vrai) si b est vérifiée dès que a est vérifiée. On n'évoque pas les deux autres cas : a non vérifiée et b vérifiée ou non, sauf si l'on doit fournir un contre-exemple au théorème en montrant que b est non vérifiée alors que a l'est.

Enfin, l'implication statistique, au sens de Gras, s'exprime en termes de règle : a implique statistiquement b si b a tendance à être vérifiée quand a l'est. La force ou la qualité de la règle s'apprécie à l'aide d'un indice (implication statistique) qui mesure l'in vraisemblance du nombre de contre-exemples à la règle sous l'hypothèse d'indépendance a priori des variables a et b et eu égard aux effectifs observés dans la population. Cette prise en compte des contre-exemples amène à privilégier le terme de quasi-implication. Comment interpréter une règle quasi-implicative ?

3. Interprétation de l'implication

On peut aborder l'implication stricte entre deux variables (des attributs par exemple) de manière intuitive à l'aide d'un diagramme de Venn (Fig. 1).

Dans notre modèle, A (respectivement B) représente les sujets qui possèdent le caractère a ou l'attribut a (respectivement b). Si l'ensemble A est inclus dans l'ensemble B , on pourra tirer les conclusions suivantes : la règle $a \Rightarrow b$ est strictement vraie ; et cela sans se soucier de la vérité de la réciproque $b \Rightarrow a$ qui pourrait être vraie ou fausse (fausse dans notre diagramme). En revanche, sur celui-ci, on

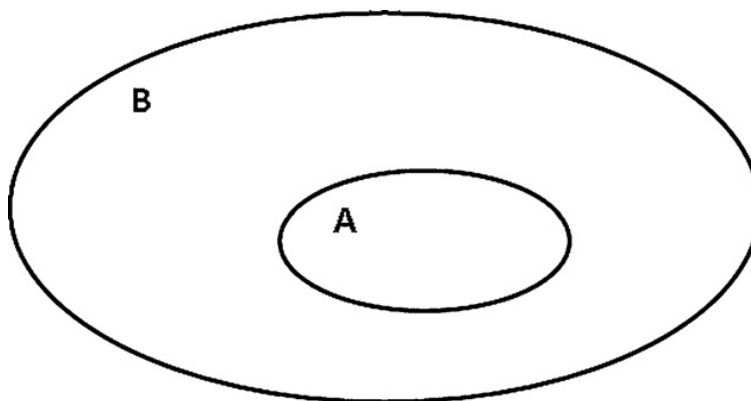


Fig. 1. L'implication représentée par un diagramme de Venn.

The implication represented by a Venn diagram.

constate que la contraposée non $b \Rightarrow$ non a est vraie ; on ne peut rien dire quant à la réciproque de la contraposée. Mais comment calcule-t-on un indice de quasi-implication statistique ?

4. Calcul de l'indice d'implication

Notons A et B les sous-ensembles respectifs de E d'individus qui vérifient respectivement les variables a et b (Fig. 2). Pour une règle quelconque $a \Rightarrow b$, observée dans E , l'ASI consiste à comparer le nombre de contre-exemples $n_{a\bar{b}}$ à cette règle observés dans $A \cap \bar{B}$, avec le nombre de contre-exemples qui apparaîtraient lors d'un choix aléatoire et indépendant de deux parties X et Y de E de mêmes cardinaux respectifs que A et B (Fig. 1) (Gras, 1979 ; Lebart et al., 2006). La variable aléatoire associée est notée $N_{a\bar{b}}$.

La qualité de la règle $a \Rightarrow b$ sera intuitivement d'autant meilleure que

$$\text{Prob} [N_{a\bar{b}} > n_{a\bar{b}}]$$

sera proche de 1 : autrement dit, en général, on observe ici plus de contre-exemples dans des circonstances aléatoires que l'on en a observés dans la contingence. Dans ce cas, le seul hasard conduit donc, en moyenne, à plus de contre-exemples que ce qui est observé.

La méthode de tirage au hasard de X et Y , dans une hypothèse a priori d'indépendance de a et b , conduit à différentes options pour la loi de la variable aléatoire $N_{a\bar{b}}$. Deux modélisations de cette variable sont généralement retenues en ASI conduisant à un modèle de Poisson et à un modèle binomial (Gras, 1979 ; Gras et Régnier, 2009). On centre et on réduit cette variable en la variable $Q(a, \bar{b})$; l'observation contingente (empirique), sa réalisation, est $Q(a, \bar{b})$. Par exemple, dans le cas du modèle de Poisson, on obtient l'indice de base pour des variables binaires :

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

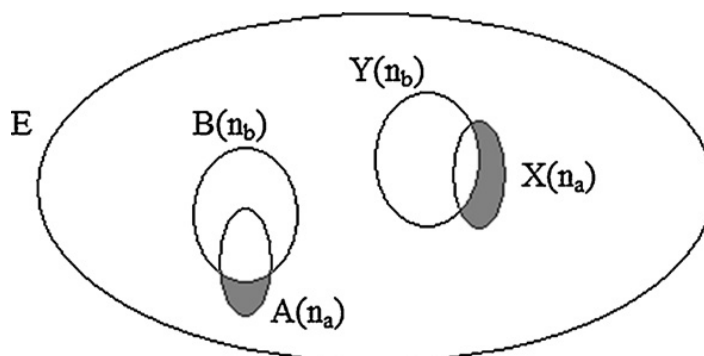
alors que

$$q(a, \bar{b}) = \frac{n_{a\bar{b}} - n_a n_{\bar{b}}}{\sqrt{\frac{n_a n_{\bar{b}}}{n}}}$$

est sa réalisation contingente.

En notant O l'effectif observé des contre-exemples et T l'effectif observable sous l'hypothèse d'indépendance des variables, cette réalisation s'écrit :

$$\frac{O - T}{\sqrt{T}}$$



Les parties grisées représentent les contre-exemples à l'implication $a \Rightarrow b$

Fig. 2. Représentation ensembliste de l'implication.

Together accounted for the involvement.

Tableau 1

Tableau de contingence de cancer et bronchite.
Contingency table of cancer and bronchitis.

	Bronchite : vrai	Bronchite : faux
Cancer : vrai	139	132
Cancer : faux	2912	3850

Le chiffre en gras exprime le nombre de contre-exemples (personnes atteintes d'un cancer et non d'une bronchite).

dont le carré est la contribution au Chi^2 d'indépendance.

L'intensité d'implication est alors définie par :

$$\varphi(a, b) = \text{Prob} [Q(a, \bar{b}) > q(a, \bar{b})]$$

dont la valeur gaussienne asymptotique, centrée et réduite est, en notant Z la variable aléatoire $Q(a, \bar{b})$ et z sa réalisation :

$$\varphi(a, b) = \frac{1}{\sqrt{2\pi}} \int_{q(a, \bar{b})}^{+\infty} e^{-\frac{t^2}{2}} dt = \text{Prob} [Z > z] \quad (1)$$

Plus $q(a, \bar{b})$ est négatif et grand en valeur absolue, meilleure sera la qualité de la règle car le niveau de confiance en la règle tend vers 1.

4.1. Exemple pour variables binaires

On reprendra un exemple concret pour variables binaires (Cadot, 2009, p. 226) tiré des données de l'enquête ASIA menée auprès de 7033 personnes. Dans cet exemple, on calcule l'intensité de la règle cancer (de la gorge) \Rightarrow bronchite et $n_{a\wedge\bar{b}}$ indique le nombre de contre-exemples, soit 132 personnes présentant un cancer sans pour autant être atteints de bronchite (Tableau 1).

Ainsi,

$$n_a = 271, n_{\bar{b}} = 3982, n_{a\wedge\bar{b}} = 132$$

Par suite,

$$q(a, \bar{b}) = \frac{n_{a\wedge\bar{b}} - \frac{n_a n_{\bar{b}}}{n}}{\sqrt{\frac{n_a n_{\bar{b}}}{n}}} = \frac{132 - \frac{271 \cdot 3982}{7033}}{\sqrt{\frac{271 \cdot 3982}{7033}}} = -1,73$$

L'indice d'intensité de l'implication correspondante tiré de la table de la loi normale centrée réduite prend la valeur de 0,96. On dira donc que le cancer de la gorge entraîne la bronchite avec l'intensité d'implication de 0,96.

4.2. Exemple pour variables numériques

Un second exemple concerne des données issues de variables numériques. Plus de 300 enseignants de mathématiques de lycée (nous n'en avons conservé que 20 pour la facilité des calculs) ont ordonné deux objectifs fondamentaux de leur enseignement parmi une douzaine d'autres selon leur préférence (Tableau 2) : « savoir-faire » (variable a) et « savoirs » (variable b). Cette préférence, ordonnée selon des valeurs croissantes, est marquée par des pondérations variant de 0 à 1 à l'issue d'une normalisation afin de ramener les rangs dans cet intervalle.

Tableau 2
Classement des objectifs.
Ranking of goals.

Professeurs	a savoir-faire	b savoirs
P001	0	0
P002	0,6	0,8
P003	0,6	0,8
P004	0	0,1
P005	0	1
P006	0,2	1
P007	1	0,4
P008	0,4	0
P009	0,6	0,8
P010	0	0,8
P011	0	0,6
P012	0	0,8
P013	0,8	1
P014	0	0
P015	0	0,1
P016	1	0
P017	0,6	1
P018	0,1	0
P019	0,4	0
P020	0	0

Par exemple, la variable b est classée plus importante¹ que la variable a par le professeur P017. Les calculs donnent les valeurs suivantes :

- $n = 20$;
- somme des rangs de a : $n_a = 6,3$;
- somme des rangs de b : $n_b = 9,2$;
- moyenne $m_a = 0,31$;
- moyenne $m_b = 0,46$;
- écart-type $s_a = 0,35$;
- écart-type $s_b = 0,42$;
- $v_a =$ variance de a ;
- $v_b =$ variance de b.

Notons a_i (resp. b_i) la valeur prise par la variable a (resp.b) par le sujet et $\bar{b}_i = 1 - b_{ii}$. La somme $\sum_{i \in E} a_i \bar{b}_i$ est obtenue comme somme des produits successifs :

$$0 \times 1 + 0,6 \times 0,2 + 0,6 \times 0,2 + 0 \times 0,9 + \dots + 0,4 \times 1 + 0 \times 1 = 2,86$$

Par suite appliquant la formule qui donne l'indice d'implication dans le cas des variables numériques (Lagrange, 1998 ; Gras et Régnier, 2009) :

$$\tilde{q}(a, \bar{b}) = \frac{\frac{1}{n} \sum_{i \in E} a_i \bar{b}_i - m_a m_{\bar{b}}}{\sqrt{\frac{v_a(v_b + m_{\bar{b}}^2 + m_a^2 v_b)}{n}}} + \frac{\frac{1}{20} \times 2,86 - 0,31 \times 0,54}{\sqrt{\frac{1}{20} [0,35^2(0,42^2 + 0,54^2) + 0,31^2 \cdot 0,42^2]}}$$

¹ On pourrait aussi considérer les variables modales comme des variables-rangs (Régnier et Gras, 2005). Dans ce cas, le vecteur-rang est constitué de composantes qui ne peuvent être posées comme indépendantes et par conséquent nécessite de respecter cette propriété pour le calcul de l'espérance et de la variance de la statistique «Coefficient de pension».

Le numérateur de l'indice est au signe près la covariance de a et b dont la symétrie est détruite pour la visée implicative par le dénominateur non symétrique en a et b et qui a la dimension d'une variance.

On obtient : $\tilde{q}(a,\bar{b}) = 0,44$ soit l'intensité d'implication : $\varphi(a,b) = 0,67$, valeur donnée par la table de la loi normale centrée réduite et validée par Classification hiérarchique implicative et cohésive (CHIC).

On dira donc, sur la base de cet échantillon de 20 professeurs, que si les enseignants de math placent les « savoir-faire » à un certain niveau d'importance parmi leurs objectifs didactiques, alors ils ont tendance à juger les « savoirs » à un niveau plus élevé.

Pratiquement, les calculs se font à l'aide d'un logiciel de traitement automatique d'extraction de règles, logiciel dénommé par l'acronyme CHIC (Couturier et Ag Almouloud, 2009).

5. Le logiciel CHIC

Ce logiciel a pour fonctions essentielles d'extraire automatiquement des données, un ensemble de règles d'association quasi implicatives entre les variables (on récupère la table complète des indices d'implication entre toutes les variables dans un fichier .csv), de fournir un indice de qualité de l'association, de représenter une structuration des variables obtenue au moyen de ces règles et de mettre en évidence la contribution (i.e. la « responsabilité ») des sujets ou des descripteurs de ces sujets par rapport aux règles ou aux familles de règles (chemin de graphe implicatif, classe de hiérarchie orientée).

Le graphe implicatif, de type « treillis » pondéré et sans cycle, représente la structure de tout ou partie des variables ; un arc orienté de a vers b représente la règle « a implique b » et sa pondération est la valeur de l'intensité d'implication de la règle.

Une originalité complémentaire de CHIC consiste à fournir, comme il est fait en analyse factorielle, la contribution ou la typicalité de tout sujet et de toute catégorie de sujets à la formation d'un arc implicatif entre deux variables a et b , c'est-à-dire la règle $a \Rightarrow b$.

Suite à cette présentation rapide de l'ASI et du logiciel CHIC, on abordera le thème central de l'article, à savoir la possibilité d'utilisation par psychologue en complément des méthodes corrélationnelles.

6. L'analyse statistique implicative et les approches corrélationnelles classiques

L'ASI complète les approches corrélationnelles classiques en orientant les liens entre les variables. En effet comme le souligne Brauer, 2000 :

« Alors que les recherches expérimentales possèdent généralement une validité interne très élevée, les études corrélationnelles souffrent chroniquement de l'impossibilité de pouvoir tirer des conclusions causales fermes ».

Cette impossibilité dérive de la symétrie du coefficient de corrélation : si a et b sont deux variables quelconques, les expressions aRb et bRa ne sont pas différenciées et l'ordre des deux variables reste indéterminé. En offrant la possibilité d'utiliser un indice d'intensité d'implication entre les variables corrélées l'ASI permet d'ordonner de manière implicative les variables corrélées, mais bien évidemment de façon non symétrique. Certes, la relation d'ordre quasi-implicatif n'exprime pas nécessairement une relation causale mais elle peut éventuellement constituer un premier pas dans le sens d'une hypothèse de causalité. Dans des travaux plus récents, l'ASI s'appuie non seulement sur la qualité de la règle directe, « si a alors b », mais aussi sur sa contraposée, « si non b alors non a ». L'implication peut aussi renvoyer à des relations d'appartenance dans le cas d'emboîtement de classes (homme \Rightarrow mortel). On peut aussi, par analogie avec l'analyse factorielle, interpréter une séquence implicative à la lumière de facteurs implicatifs latents.

Lorsque le chercheur considère a priori tous les modèles corrélationnels comme équivalents, l'ASI peut intervenir en tant qu'outil de fouille de données exploratoire. Dans une autre situation, le chercheur aura pu poser l'hypothèse d'un modèle quasi-implicatif particulier et chercher à le valider dans une perspective confirmatoire.

Dans ce papier, on examinera successivement un exemple à deux variables, puis un exemple à trois variables. On pourra constater l'intérêt de compléter l'approche corrélacionnelle par une ASI pour orienter les liens, ce qui permet également un gain de temps par l'économie de l'examen de toutes les possibilités. On se limitera à une approche en termes de variable à expliquer statistiquement sans entrer dans la problématique ni de la nature des relations ni de leur interprétation psychologique.

On utilisera des données relatives à 226 répondants salariés. On aura recours aux variables intention de poursuivre sa carrière dans l'entreprise (*carrière_intra*) et un indice de distance socio-normative calculé par rapport à un prototype personologique professionnel (*dsn_prof*) pour le premier exemple à deux variables. En variable supplémentaire, on prendra le niveau de diplôme des répondants. Pour l'exemple à trois variables, on adjoindra aux deux premières la variable pensée positive (*pensée_pos*). On complétera la présentation par l'introduction de variables psychologiques formant réseau nomologique : soutien du chef, qualité des relations avec les collègues, anxiété en entreprise et les quatre facettes de l'implication organisationnelle à savoir l'implication normative, l'implication affective, l'implication par intérêt et l'implication par manque d'alternatives.

7. Un exemple à deux variables : intention de poursuivre sa carrière dans l'entreprise et distance socio-normative

CHIC fournit la corrélation entre les deux variables qui est de 0,31. Si on peut affirmer le lien entre les deux processus, du fait de la symétrie du coefficient de corrélation, on ne peut pas dire si c'est l'intention de faire carrière au sein de l'entreprise qui se traduit par une distance socio-normative réduite (proposition P1) ou à l'inverse, si c'est une distance socio-normative réduite qui se traduit par l'intention de faire carrière au sein de l'entreprise (proposition P2). L'ASI aide à documenter l'incertitude quant aux deux propositions P1 et P2 par le biais de la table des indices d'implication entre les deux variables (Tableau 3).

On observe deux valeurs différentes, ce qui confirme l'asymétrie de l'indice d'implication. La règle quasi-implicative la plus marquée (0,72 vs 0,63) est en faveur de P2 : soit le chemin *dist_socio_nor* ⇒ *carrière_intra* qu'on retrouve à l'aide de CHIC sous la forme d'un graphe implicatif (Fig. 3).

On peut donc conclure que les répondants présentant une faible distance socio-normative auront tendance à déclarer vouloir poursuivre leur carrière au sein de l'entreprise ; et que, par la contraposée, les répondants qui déclarent ne pas vouloir poursuivre leur carrière dans l'entreprise auront tendance à présenter une distance socio-normative plus élevée. On ne peut rien dire sur l'inverse de ces deux conclusions.

L'ASI sous CHIC permet l'introduction de variables supplémentaires pour lesquelles on peut calculer la contribution à un chemin quasi-implicatif. Dans l'exemple présenté (Tableau 4), on a introduit les niveaux de diplôme et calculé la contribution de chaque niveau au chemin *dist_socio_nor* ⇒ *carrière_intra*. L'indice de contribution varie de 0 à 1. C'est une probabilité qui représente un seuil d'acceptabilité de la contribution du sujet ou de la catégorie de sujets à un chemin ou à un classe de variables. Plus il est proche de 0 plus grande est la contribution. On admet qu'un indice égal ou inférieur à 0,10 marque une contribution à prendre en compte dans l'interprétation des données. Ici, on s'aperçoit que la contribution des niveaux de diplômes supérieurs contribuent plus fortement à la qualité de la règle que les niveaux inférieurs, ce qui est particulièrement vérifié pour le niveau Bac. + 3. + 4.

Tableau 3

Table des indices d'implication pour les deux variables.

Table indices of involvement for both variables.

Variabes	Carrière_intra	Dist_socio_nor
Carrière_intra	0	0,63
Dist_socio_nor	0,72	0

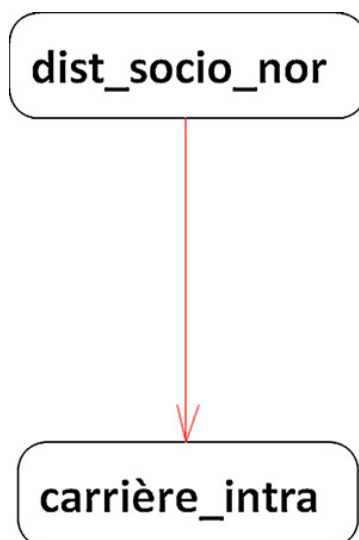


Fig. 3. Graphe implicatif $\text{dist_socio_nor} \Rightarrow \text{carrière_intra}$.

Graph implicative $\text{dist_socio_nor} \Rightarrow \text{carrière_intra}$.

Tableau 4

Contributions au chemin $\text{dist_socio_nor} \Rightarrow \text{carrière_intra}$.

Contributions to the path $\text{dist_socio_nor} \Rightarrow \text{carrière_intra}$.

Niveaux de diplôme	Contributions
Bac_+3_+4	0,09
Bac_+2	0,24
CAP_BEP	0,30
Bac	0,45
sans.diplôme	0,63

8. L'introduction d'une troisième variable dans le modèle : la tendance à la pensée positive

La nouvelle variable, la tendance à la pensée positive, est corrélée à hauteur de 0,19 avec les deux premières variables. On sait donc que les trois variables sont liées, mais les corrélations ne permettent pas de les ordonner.

Classiquement et par rapport à l'usage des régressions, on peut raisonnablement envisager trois hypothèses explicatives candidates. H1 : le score carrière_intra est statistiquement expliqué par les deux autres variables ; H2 : le score dist_socio_nor est statistiquement expliqué par les deux autres variables ; H3 : le score pensée_pos est statistiquement expliqué par les deux autres variables.

Les résultats des régressions (Tableau 5) confirment qu'on peut accepter les trois hypothèses, les deux premières étant équivalentes du point de vue du F de Snedecor, la troisième se révélant un peu moins puissante. Que peut apporter de plus le recours à l'ASI (Gras et al., 1997) ?

Le calcul des indices d'implication asymétriques (Tableau 6) introduit une méta règle : $(\text{dist_socio_nor} \Rightarrow \text{pensée_pos}) \Rightarrow \text{carrière_intra}$. Les répondants dont une distance socio-normative réduite prédispose à la tendance à la pensée positive seront ceux qui envisageront de poursuivre leur carrière dans leur entreprise. L'ensemble est visualisé par le graphe implicatif fourni par CHIC (Fig. 4).

La commande « mode cône » du logiciel permet de mettre en évidence l'existence d'une closure transitive de dist_socio_nor à carrière_intra , indiquée par la flèche pointillée (Fig. 5).

Au-delà d'une lecture du résultat terme à terme, de règle en métarègle, on peut, par analogie avec les approches factorielles, présupposer une variable latente orientée qui sous-tend le chemin implicatif, soit ici un facteur implicatif de conformisme socio-professionnel. CHIC permet de récupérer dans un fichier Excel les contributions des individus au chemin quasi-implicatif, ou au facteur implicatif si l'on accepte le postulat d'une variable latente susceptible de rendre compte des liens statistiquement quasi-implicatifs entre les trois variables.

Tableau 5
Régressions pour le modèle à trois variables.
Regressions for the three-variable model.

Équations	R multiple	R ² R ² ajusté	F	a1	a2
H1 : car- rière_intra = cte + a1* dist_socio_nor + a2* pensée_pos	0,34	0,11 0,10	14,71 <i>p</i> < 0,01	0,21 <i>t</i> = 4,47 <i>p</i> < 0,01	0,10 <i>t</i> = 2,15 <i>p</i> = 0,03
H2 : dist_socio_nor = cte + a1* carrière_intra + a2* pensée_pos	0,34	0,11 0,10	14,72 <i>p</i> < 0,01	0,39 <i>t</i> = 4,47 <i>p</i> < 0,01	0,14 <i>t</i> = 2,15 <i>p</i> = 0,03
H3 : pensée_pos = cte + a1* carrière_intra + a2* dist_socio_nor	0,23	0,05 0,04	6,73 <i>p</i> < 0,01	0,20 <i>t</i> = 2,15 <i>p</i> = 0,03	0,15 <i>t</i> = 2,15 <i>p</i> = 0,03

Tableau 6
Table des indices d'implication pour les trois variables.
Table of indices for the three variables involved.

Variabes	Carrière_intra	Dist_socio_nor	Pensée_pos
Carrière_intra	0	63	58
Dist_socio_nor	72	0	63
Pensée_pos	64	60	0

On utilisera le même sens de lecture que pour le [Tableau 3](#).

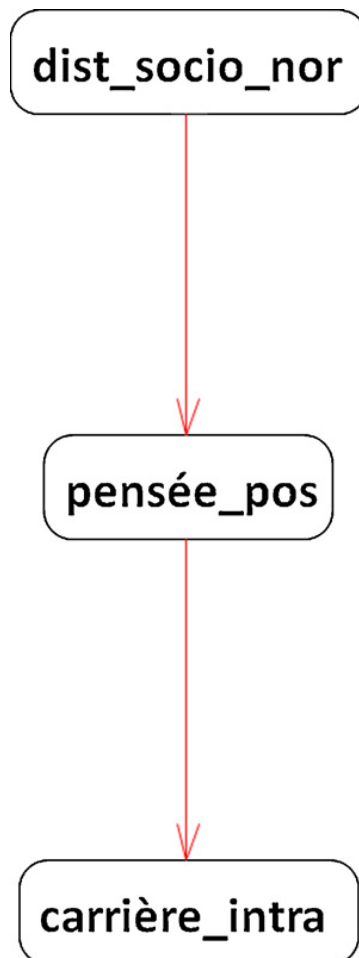


Fig. 4. Graphe implicatif dist_socio_nor ⇒ pensée positive ⇒ carrière_intra.
Graph implicative dist_socio_nor ⇒ Positive Thinking ⇒ carrière_intra.

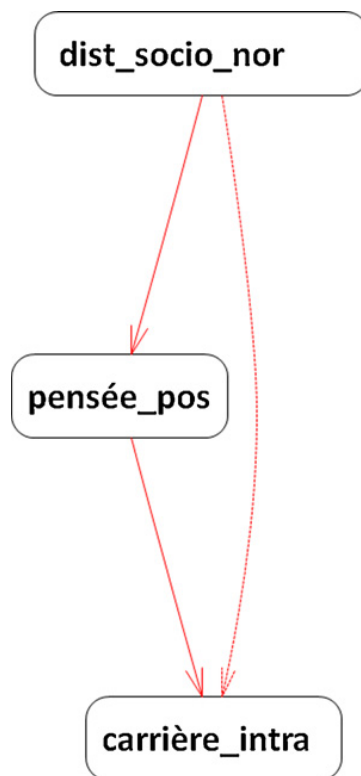


Fig. 5. Graphe implicatif en mode cône.

Graph mode implicative cone.

Il devient alors loisible d'inscrire ces contributions dans un réseau nomologique susceptible de donner du sens à ce facteur implicatif. Dans l'exemple présenté, on a calculé les corrélations entre contributions des répondants au facteur implicatif et un certain nombre de variables susceptible de l'éclairer (Tableau 7) : soutien du chef, qualité des relations avec les collègues, anxiété en entreprise et les quatre facettes de l'implication organisationnelle à savoir l'implication normative, l'implication affective, l'implication par intérêt et l'implication par manque d'alternatives.

On observe que ce facteur quasi-implicatif de conformisme socioprofessionnel est favorisé par de bonnes relations avec les collègues et par le sentiment d'être soutenu par le supérieur hiérarchique. A contrario, l'anxiété ressentie dans l'entreprise ainsi que l'implication organisationnelle sur le mode du manque d'alternatives interfèrent avec ce facteur.

Par comparaison, une analyse en composantes principales des trois variables débouche sur une même structure de résultats : un seul facteur est extrait de valeur propre 1,47 et expliquant 49% de la variance. Les corrélations des contributions des répondants à ce facteur et des variables du réseau nomologique sont du même ordre : 0,45 pour les relations avec les collègues, 0,33 pour le soutien du chef, -0,27 pour l'anxiété et -0,26 pour l'implication par manque d'alternatives. On constate donc une

Tableau 7

Corrélations entre les contributions au facteur implicatif et les variables du réseau nomologique.
Correlations between the contributions to the factors involved and the nomological network variables.

Variabiles	R contribution/variable
Relations collègues	0,39
Soutien du chef	0,37
Implication affective	0,12
Implication par intérêt	0,05
Implication normative	0,02
Anxiété entreprise	-0,23
Implication manque alternative	-0,26

certaine cohérence des résultats entre les deux approches. La différence est que l'ASI ordonne les variables du facteur de conformisme en une séquence quasi-implicative qui complète l'interprétation de la variable latente.

9. Conclusion

Notre propos n'est pas d'opposer l'ASI à la pratique traditionnelle du psychologue. Bien au contraire, on a vu que les exemples présentés illustrent une complémentarité entre approche corrélacionnelle et approche implicative. Cette complémentarité se situe dans la possibilité offerte par l'ASI d'ordonner de manière quasi-implicative les variables corrélées, ce qui enrichit la qualité descriptive des données, complète et approfondit l'explication statistique, ouvre la voie à une plus-value heuristique des perspectives interprétatives. Par exemple, la notion de facteur quasi-implicatif qui permet d'inscrire les contributions au facteur dans un réseau nomologique de variables susceptibles de l'éclairer, apporte une plus-value interprétative par rapport à la simple lecture des liaisons inter-variables.

On pourra mettre l'ASI en œuvre soit en amont de l'étude corrélacionnelle, soit en aval pour ordonner des corrélacions préalablement établies, tout aussi bien dans des approches exploratoires que confirmatoires.

Les exemples présentés relèvent d'une approche exploratoire sans modèle hypothétique prédéfini à vérifier. Pour le psychologue, s'ouvre également la possibilité de traiter des analyses confirmatoires sur un mode implicatif. Comme en analyse structurelle on établit le degré d'ajustement entre une structure à caractère théorique, ou au statut de modèle, posée a priori et une structure empirique établie à partir de données d'observation selon des indices globaux ou plus analytiques. Ces approches ont déjà été décrites (Spagnolo et al., 2009 ; Pasquier et Gras, 2011²).

Déclaration d'intérêts

Les auteurs déclarent ne pas avoir de conflits d'intérêts en relation avec cet article.

Remerciements

Nos remerciements à Évelyne Fouquereau pour sa relecture critique du texte.

Références

- Brauer, M., 2000. L'identification des processus médiateurs dans la recherche en psychologie. *Année Psychol.* 100, 661–681.
- Cadot, M., 2009. Graphe de règles d'implication statistique pour le raisonnement courant. Comparaison avec les réseaux bayésiens et les treillis de Galois. In: Gras, R., Régnier, J.C., Guillet, F. (Eds.), *Analyse statistique implicative, une méthode d'analyse de données pour la recherche de causalités*. Cépaduès, Toulouse, pp. 223–250.
- Couturier, R., Ag Almouloud, S., 2009. CHIC : utilisation et fonctionnalités. In: Gras, R., Régnier, J.C., Guillet, F. (Eds.), *Analyse statistique implicative, une méthode d'analyse de données pour la recherche de causalités*. Cépaduès, Toulouse, pp. 279–291.
- Gras, R., (1979). Contribution à l'étude expérimentale et à l'analyse de certaines acquisitions cognitives et de certains objectifs didactiques en mathématiques, Thèse d'Etat, Université de Rennes 1.
- Gras, R., Larher, A., 1993. L'implication statistique, une nouvelle méthode d'analyse de données. *Mathématiques, Informatique et Sciences Humaines* 120, 5–31.
- Gras, R., 1996. L'implication Statistique. La Pensée Sauvage, Grenoble.
- Gras, R., Briand, H., Peter, P., Philippe, J. (1997). *Implicative statistical analysis*. Proceedings of International Congress I.F.C.S., 96, Kobe, Springer-Verlag, Tokyo.
- 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* 154–155, 9–29.
- Gras, R., Kuntz, P., 2008. An overview of the Statistical Implicative. In: Gras, R., Suzuki, E., Guillet, F., Spagnolo, F. (Eds.), *Statistical Implicative Analysis*. Springer-Verlag, Berlin-Heidelberg, pp. 11–40.
- Gras, R., Régnier, J.C., 2009. Analyse implicative des variables binaires. Intensité implicative. Intensité entropique. In: Gras, R., Régnier et, J.C., Guillet, F. (Eds.), *Analyse statistique implicative, une méthode d'analyse de données pour la recherche de causalités*. Cépaduès, Toulouse, pp. 17–129.

² Pasquier, D., Gras, R. (2011). Analyse exploratoire a priori et analyse confirmatoire a posteriori : paradigme pour la comparaison de deux structures quasi-implicatives en analyse statistique implicative (ASI). Texte soumis.

- Lahanier-Reuter, D., 2009. Analyse implicative des données et didactiques disciplinaires. In: Gras, R., Régnier, J.C., Guillet, F. (Eds.), *Analyse statistique implicative, une méthode d'analyse de données pour la recherche de causalités*. Cépaduès, Toulouse, pp. 317–338.
- Lagrange, J.B., 1998. Analyse implicative d'un ensemble de variables numériques; application au traitement d'un questionnaire à réponses modales ordonnées. *Revue de statistique appliquée* 71–93.
- Lebart, L., Piron, M., Morineau, A., 2006. *Statistique exploratoire multidimensionnelle*, 4^e éd. Science sup, Dunod, Paris.
- Régnier, J.C., Gras, R., 2005. Statistique de rangs et analyse statistique implicative. *Rev. Stat. Appl.* LIII, 5–38.
- Spagnolo, F., Gras, R., Régnier, J.C., 2009. Mesurer l'écart entre une analyse a priori et la contingence en didactique. In: Gras, R., Régnier et, J.C., Guillet, F. (Eds.), *Analyse statistique implicative, une méthode d'analyse de données pour la recherche de causalités*. Cépaduès, Toulouse, pp. 165–174.

Quelques études utilisant l'ASI en psychologie

- Verscheure, I., Chiocca, C.M., 2009. Étude de représentations d'élèves en éducation physique et sportive. In: Gras, R., Régnier et, J.C., Guillet, F. (Eds.), *Analyse statistique implicative, une méthode d'analyse de données pour la recherche de causalités*. Cépaduès, Toulouse, pp. 369–386.

Pour en savoir plus

- Bellaj, T., Pasquier, D., 2009. Un exemple d'analyse implicative en neuro-psychologie: la comparaison de groupes contrastés. In: Gras, R., Régnier, J.C., Guillet, F. (Eds.), *Analyse statistique implicative, une méthode d'analyse de données pour la recherche de causalités*. Cépaduès, Toulouse, pp. 387–403.
- Mokoukolo, R., Pasquier, D., 2008. Stratégies d'acculturation: cause ou effet des caractéristiques psychosociales ? L'exemple de migrants d'origine algérienne. *Cah. Int. Psychol. Soc.* 79, 57–67.
- Pasquier, D., Temple, C., 2010. Pour une approche implicative confirmatoire; l'exemple des stratégies identitaires des migrants. In: Régnier, J.C., Spagnolo, F., Di Paola, B., Gras, R. (Eds.), *Analyse statistique implicative. Objet de recherche et de formation en analyse de données, outil pour la recherche multidisciplinaire*. GRIM, Université de Palerme, Palerme.
- Pasquier, D., Valéau, P., 2008. Des figures de la désirabilité sociale à l'équilibre psychologique. In: Loarer, E., Vrignaud, P., Mogenet, J.L., Cuisinier, F., Gottesdiener, H., Mallet, P. (Eds.), *Perspectives différentielles en psychologie*. Presses Universitaires de Rennes, Rennes, pp. 215–219.
- Pasquier, D., Valéau, P., 2011. Implication organisationnelle, anxiété et états affectifs au travail. *Cah. Int. Psychol. Soc.*, 5–36.
- Pellois, C., 2008. Performances cognitives et acquis scolaires: quelles interactions ? In: Loarer, E., Vrignaud, P., Mogenet, J.L., Cuisinier, F., Gottesdiener, H., Mallet, P. (Eds.), *Perspectives différentielles en psychologie*. Presses Universitaires de Rennes, Rennes, pp. 497–500.
- Site de Régis Gras : <http://math.unipa.it/~grim/homegras.03.htm>.
- Site du Groupe International d'Analyse Statistique Implicative : http://math.unipa.it/~grim/asi/asi_index.htm.
- Les textes relatifs au dernier congrès de l'ASI sont disponibles à l'adresse suivante : <http://sites.univ-lyon2.fr/asi5/>.
- Les textes de Daniel Pasquier et collaborateurs sont disponibles à l'adresse : <http://www.avenireentreprise.fr>.