

# Psychologie du Travail



## et des Organisations

*Fondée en 1995 à l'initiative de Claude Lemoine et Michel Rousson, sous l'égide de l'Association Internationale de Psychologie du Travail de Langue Française, cette revue publie des articles originaux, des revues de questions, des comptes-rendus de recherches, y compris celles réalisées sur le terrain ou dans une perspective d'application. Elle présente également des comptes rendus d'ouvrages et des notes sur l'actualité du domaine.*

### **Audience**

**Psychologie du Travail et des Organisations** s'adresse aux enseignants, aux chercheurs et aux praticiens du domaine et à un plus vaste public : étudiants, responsables des ressources humaines, gestionnaires, ergonomes, médecins du travail...

### **Objectifs**

Les thèmes principaux concernent les aspects individuels, psychosociaux et structurels du travail et des organisations. À titre d'exemples non exclusifs, on peut citer les questions portant sur : la gestion et le développement des ressources humaines (formation, compétences, innovation), l'organisation et l'évaluation des systèmes (changement, communication, climat), les articulations hommes-organisations-techniques (représentations, aspects culturels, négociation, coopération, style de direction), la santé (bien-être, conditions de travail, stress, risques, sécurité), l'environnement, les aspects psychologiques liés à l'emploi et au non emploi (sélection, orientation, évaluation des personnes, insertion, identité, implication), le rôle du psychologue (expertise, conseil, mode d'intervention), l'épistémologie, la méthodologie et la déontologie.



**15<sup>ème</sup> congrès de l'Association Internationale de  
Psychologie du Travail de Langue Française**

***Entre tradition et innovation,  
comment transformons-nous l'univers du travail ?***

**Du 19 au 22 août 2008**  
Université Laval  
Ville de Québec  
Québec, Canada

*« Tout agent de changement, pour être efficace, doit être porteur autant de symboles de révision et de changement que de symboles de tradition et de stabilité ».*

Warren Bennis (à la suite de son mandat comme recteur de l'université de Cincinnati)

Ce congrès AIPTLF 2008 vise le partage et l'échange des savoirs et des praxis sur la transformation de l'univers du travail. Le psychologue du travail s'appuie sur les acquis de son domaine d'expertise pour faire face aux différentes réalités des organisations et, souvent, proposer des approches créatives et originales. Par ailleurs, les particularités des situations qu'il rencontre et l'évolution des environnements externes et internes l'amènent à remettre en cause, questionner, valider ces assises et, peut-être là aussi, à innover. Que ce soit suite à des initiatives individuelles ou à des pressions de l'environnement, le changement, ce délicat équilibre entre le passé et le futur, fait partie de notre quotidien.

Que deviennent les théories fondatrices de la profession lorsque confrontées à une redéfinition de la sécurité d'emploi, à l'avènement de programmes de participation des employés, à une génération de travailleurs ayant des valeurs individuelles, groupales et culturelles différentes, à des exigences accrues de productivité? Est-ce que de nouvelles théories émergent? Est-ce que les anciennes s'appliquent différemment? À la rigueur, les concepts de base sont-ils encore crédibles et applicables? Oscillant entre tradition et innovation, le psychologue est appelé à un questionnement très riche et sans fin.

Comment contribuons-nous à définir ou redéfinir l'univers du travail? Est-ce sous l'angle de la consolidation et du maintien des traditions ou bien est-ce sous l'angle du renouveau et de l'innovation? Notre contribution reflète-t-elle une adaptation réactive visant à fournir des outils pour composer avec les bouleversements ? Ou est-ce plutôt une contribution (pro)active qui vise à promouvoir des changements plutôt qu'à les subir ?

Ce congrès sera l'occasion de proposer des réflexions, des expérimentations et des solutions pour mieux faire face aux nouvelles réalités du travail en conjuguant les connaissances et leur application pratique. Comment, concrètement, cette connaissance nouvelle ou consolidée d'un aspect des comportements au travail peut-elle se traduire par des actions qui auront un impact sur la vie organisationnelle? Quelle est la contribution du psychologue du travail sur les pratiques qu'une organisation privilégie pour transiger avec ses ressources les plus précieuses, les êtres humains? Comment influencer les dirigeants qui décident de ces pratiques dans un contexte où les aspects économiques sont omniprésents ?

**Le psychologue du travail est-il cette valeur certaine pour la transformation de l'univers du travail autant en réponse à des tensions internes ou pressions externes qu'en tant qu'initiateur avant-gardiste de changement de l'univers du travail ?**

## Abonnement

Les demandes d'abonnement sont à envoyer à Claude Lemoine, AIPTLF, Université Lille 3, UFR Psychologie, BP 60149, F 59653 Villeneuve d'Ascq cedex. Tél.: 03 20 41 69 68

Tarif annuel 2008  
(4 numéros de 96 pages environ)

		France	UE	Autres Pays
Papier	Individuels	60 €	65 €	70 €
	Etudiants (- de 28 ans, justificatifs)	40 €	45 €	50 €
	Institutions	100 €	105 €	110 €
Internet	Individuels	45 €	50 €	55 €
	Etudiants (- de 28 ans, justificatifs)	20 €	25 €	30 €
	Institutions	85 €	90 €	95 €
Papier + Internet	Individuels	75 €	80 €	85 €
	Etudiants (- de 28 ans, justificatifs)	55 €	60 €	65 €
	Institutions	115 €	120 €	125 €

Les prix indiqués sont nets (ttc), les frais de virement sont à la charge du client. En cas de virement étranger, ajouter 10 Euros.

Règlement par chèque ou par virement bancaire à l'ordre de AIPTLF / IBAN : Crédit Mutuel, 969 av. République, F 59700 Marcq en Baroeul.  
N° FR761562 9027 2000 0448 8270 158

Adhésion AIPTLF : 20 € ([claudio.lemoine@univ-lille3.fr](mailto:claudio.lemoine@univ-lille3.fr) ou [christine.lagabrielle@u-bordeaux2.fr](mailto:christine.lagabrielle@u-bordeaux2.fr)). Pour les membres de l'AIPTLF, le justificatif d'abonnement permet d'obtenir une réduction sur la cotisation de l'année (10 euros au lieu de 20). Celle-ci est à régler directement à l'AIPTLF, Pr. C. Lemoine (cf. supra).

.....  
**Bon de commande**

Je m'abonne à **Psychologie du Travail et des Organisations** pour l'année 2008

Version papier

Version Internet

Version Papier et Internet

Individuels

Etudiants

Institutions

Nom de l'institution: .....

Nom: .....

Prénom: .....

Adresse:.....

Tel:..... E-Mail:.....

Je verse .....€       par chèque       par virement bancaire



# CONTENTS

## **Editorial**

**p 8**

## **Daniel Holender**

**p 9**

Use of the confidence intervals instead of the tests of significance: epistemological and practical aspects

## **C. Lagabrielle, A.M. Vonthron & D. Pouchard**

**p 43**

Contribution of The Theory of Planned Behavior to the prediction of taking up vocational training

## **Daniel Gilibert et Ingrid Banovic**

**p 55**

Nurses' representation of the determinants of the state of health and the points of view attributed to patients and caregivers: a socionormative approach.

## **Emmanuelle Camps**

**p 77**

Meaning of work and relations work life – out of work life among three categories teleworkers and traditional workers



# SOMMAIRE

## **Editorial**

**p 8**

## **Daniel Holender**

**p 9**

Utilisation des intervalles de confiance au lieu des tests de signification : aspects épistémologiques et pratiques

## **C. Lagabrielle, A.M. Vonthron & D. Pouchard**

**p 43**

Contribution du modèle du comportement planifié à la prédiction de l'entrée en formation professionnelle

## **Daniel Gilibert et Ingrid Banovic**

**p 55**

La représentation de l'évolution de santé par les infirmiers et les points de vue envisagés des patients et des aidants naturels: une approche socionormative.

## **Emmanuelle Camps**

**p 77**

Sens du travail et rapports vie de travail - vie hors travail chez trois catégories de télétravailleurs et chez les travailleurs classiques



## **Editorial**

Notre Revue « Psychologie du Travail et des Organisations » : PTO fait à nouveau peau neuve de sa couverture, année du Congrès de Québec et des fêtes qui l'accompagnent obligent. Comme vous le constaterez aussi PTO, possède désormais son logo, afin de se faire identifier, comme entité littéraire originale.

Désormais, la Revue est en ligne, après un petit temps d'absence.

Néanmoins, vous retrouverez gracieusement les quatre numéros de 2007 sur le site : [www.revue-pto.com](http://www.revue-pto.com)

Le présent numéro, un peu spécial, complète, par ses deux premiers articles les numéros 3 et 4 de l'année 2007.

Le premier article de Daniel Holender boucle «L' Hommage à Pierre Salengros », numéro spécial 4-2007

Le deuxième celui de C.Lagabrielle, A.M. Vonthron et D. Pouchard clôture le « Spécial Formation» numéro 3-2007.

Suivent deux autres articles D. Gilibert et Ingrid Banovic, et E. Camps dont les problématiques viendront éclairer les pratiques ou les recherches.

Le comité de lecture avec ce nouveau numéro a voulu donner à PTO l'essor que le marché économique impose à la publication.

Nous espérons continuer à vous satisfaire avec votre nouvel abonnement.

*Le Rédacteur en Chef*  
*Professeur G. Masclat*

# **Utilisation des intervalles de confiance au lieu des tests de signification : aspects épistémologiques et pratiques**

## **Use of confidence intervals instead of tests of significance: epistemological and practical aspects**

Daniel Holender  
Séminaire d'Analyse des Données,  
Université Libre de Bruxelles,  
Avenue F. D. Roosevelt, 50, CP 191,  
B-1050 Bruxelles  
Tel. : (322) 650 2031, Fax (322) 650 2209  
e-mail: holender@ulb.ac.be

### *Résumé*

L'utilisation des intervalles de confiance au lieu des tests de signification est fortement conseillée dans la cinquième édition du manuel de publication de l'*American Psychological Association* (2001). Cette possibilité ainsi que d'autres améliorations souhaitables dans les pratiques statistiques sont discutées à la lumière des grandes options théoriques sous-tendant l'inférence statistique et de la manière dont elles ont été traduites dans la pratique de la recherche en psychologie depuis une cinquantaine d'années. La suggestion d'une mise au ban complète de l'approche par les tests de signification est examinée puis rejetée. Ensuite, une démarche consistant à mettre en opposition des interprétations alternatives des données et à évaluer laquelle est la plus plausible par le calcul d'un rapport de vraisemblance est jugée intéressante et recommandable. Enfin, la richesse plus grande de l'approche par les intervalles de confiance par rapport à l'approche par les tests de signification est mise en évidence et illustrée par son application à une recherche visant à déceler l'absence plutôt que la présence d'un effet de la variable indépendante étudiée.

### *Abstract*

The use of confidence intervals instead of significance tests is strongly recommended by the fifth edition of the manual of the *American Psychological Association* (2001). This possibility as well as other improvements in statistical practice are discussed in the framework of the major theoretical options subtending statistical inference and the way they have been applied in psychology for about 50 years. First, the suggestion of a complete ban on statistical testing is examined and rejected. Next, a procedure consisting in measuring the fit of two competing models based on the likelihood ratio is judged interesting and commendable. Finally, the superiority of an approach based on confidence intervals instead of significance tests is assessed and illustrated by its application to an experimental study aiming to demonstrate the absence instead of the presence of an effect of the independent variable.

**Mots clés** : Induction ; Inférence statistique ; Intervalle de confiance ; Test de signification

**Key-words** : Confidence interval, Induction ; Significance test; Statistical inference

## 1. Introduction

Dans ses nouvelles recommandations en matière de traitement statistique des données, la cinquième édition du Publication Manual of the American Psychological Association (2001, ci-après manuel de l'APA) encourage l'utilisation des intervalles de confiance plutôt que des tests de signification. C'est à un double titre, en tant que chercheur en psychologie cognitive et en tant qu'enseignant des méthodes d'inférence statistique dans une faculté de psychologie, que je vais me livrer à un certain nombre de réflexions sur ces nouvelles directives de l'APA. L'objectif de l'article est de montrer comment ces nouvelles recommandations peuvent être utilisées comme un tremplin pour enfin aboutir à une formulation à la fois plus correcte et plus satisfaisante des conclusions que les méthodes usuelles d'inférence statistique autorisent.

Comme chercheur débutant, j'ai eu la chance d'être associé pour un temps au travail de Henri Rouanet et Dominique Lépine sur l'inférence fiducio-bayésienne (Rouanet et al. 1978) et, par là, d'être initié aux problèmes liés à l'utilisation des tests de signification. Les critiques majeures de cette approche de l'inférence statistique qui avaient déjà été formulées dans les années 1950 et 1960 ont été rassemblées dans un recueil édité par Morrison et Henkel (1970). Depuis lors, ces critiques ont été reprises et amplifiées dans de nombreuses publications visant à attirer l'attention des psychologues sur ce qui constitue actuellement un inventaire d'une douzaine d'interprétations plus ou moins erronées ou confuses des conclusions qui peuvent être tirées de l'obtention de résultats statistiquement significatifs (cf. Nickerson, 2000 pour une discussion nuancée).

Comme enseignant impliqué à des degrés divers dans la transmission d'un savoir appliqué lié à l'inférence statistique, j'ai toujours tenté de prévenir un certain nombre d'interprétations erronées en insistant au moins sur quatre propriétés importantes de la démarche fondée sur les tests de signification. Ces propriétés constituent les quatre premières parmi les sept propositions numérotées que contiendra cet article.

**Proposition 1.** Il faut distinguer entre « signification » au sens théorique ou pratique et « signification » au sens statistique.

**Proposition 2.** Des résultats statistiquement très significatifs n'impliquent pas nécessairement des différences de grandes amplitudes entre conditions.

**Proposition 3.** Un résultat non significatif n'implique pas nécessairement une absence d'effet (un effet nul).

**Proposition 4.** En étant libre d'augmenter indéfiniment le nombre d'observations, un effet d'amplitude aussi petite que l'on veut peut devenir aussi significatif que l'on veut avec une puissance aussi grande que l'on veut.

Maintenant qu'il existe une volonté plus manifeste de la part de l'APA de promouvoir des démarches alternatives de l'inférence statistique, il faut saisir l'occasion pour modifier l'enseignement de la statistique en conséquence. L'approche du traitement des données par les intervalles de confiance constitue un bon point de départ parce que, comme le soulignent Cumming et Finch (2001, 2005), leurs avantages principaux sur l'approche par les tests de signification sont les suivants :

- ils donnent une estimation ponctuelle de la valeur d'un paramètre et une estimation de la précision de l'estimation qui sont exprimées dans l'unité de mesure utilisée dans la recherche ;
- ils peuvent fournir la même information que les tests de signification usuels ;
- ils constituent les ingrédients de certaines méthodes de méta-analyse ;
- ils donnent une information sur la précision qui est équivalente, voire supérieure, à celle fournie par le calcul de la puissance du test de signification.

Mon adhésion à cette façon de voir les choses repose sur des préoccupations plus profondes liées à mon activité de recherche. Dans ce cadre, un problème fréquent est qu'il est tout simplement impossible de s'en tenir rigoureusement au canon de l'approche par les tests de signification selon lequel seul le rejet d'une hypothèse nulle est informatif. De fait, il existe beaucoup de circonstances dans lesquelles, implicitement au moins, cette obligation de suspendre son jugement face à des résultats statistiquement non significatifs est intenable. Par exemple, dans l'utilisation de la méthode des facteurs additifs de Sternberg (1969), c'est l'absence d'interaction ou, plus exactement, le non rejet de l'hypothèse nulle d'absence d'interaction, entre les effets de deux ou plusieurs variables sur le temps de réaction, qui autorise l'inférence (probable, mais non nécessaire) que chacune de ces variables influence sélectivement des stades successifs et disjoints (sans recouvrement temporel) dans la chaîne des opérations de traitement liant le stimulus à la réponse. Les données d'une de mes expériences montrant une interaction non significative entre les effets de deux variables sur le temps de réaction (Holender et Bertelson, 1975, expérience 4) ont servi d'illustration pour montrer comment l'inférence fiducio-bayésienne peut être utilisée pour valider le modèle des stades successifs du traitement (Rouanet et al. 1978). Un autre contexte fréquent dans lequel on peut être amené à « accepter » l'hypothèse nulle d'une absence d'effet est celui de la non réplication, motivée d'un point de vue théorique, d'un résultat publié, comme on en verra un exemple à la section 4.4.

La suite de l'article est divisée en trois parties principales. La section 2 est consacrée à une mise en perspective épistémologique des grandes options théoriques sous-jacentes à l'inférence statistique. La section 3 montre comment une approche hybride résultant d'une fusion entre des aspects partiellement incompatibles de ces approches théoriques sous-tend l'analyse statistique des résultats et la politique de publication en psychologie depuis une cinquantaine d'années. La section 4 est consacrée à la discussion de trois démarches visant au dépassement et à l'amélioration des pratiques courantes ; elle comprend : (1) l'analyse et le rejet d'une proposition de mise au ban complète de l'approche par les tests de signification (Hunter, 1997 ; Hunter et Schmidt, 1990 ; Schmidt, 1996), (2) l'analyse d'une démarche jugée prometteuse reposant sur le calcul des rapports de vraisemblance, une démarche qui autorise aussi bien de conclure en faveur qu'en défaveur de l'hypothèse nulle (Dixon, 1998, 2003 ; Dixon et O'Reilly, 1999 ; Glover et Dixon, 2004) et (3) l'illustration d'une démarche montrant qu'une utilisation judicieuse de la puissance et des intervalles de confiance permet, elle aussi, de symétriser le processus de décision.

## 2. Le problème de l'induction et les trois approches de l'inférence statistique

Les trois approches de l'inférence statistique présentées dans cette partie de l'article sont en rapport avec un problème épistémologique fondamental : comment peut-on justifier l'induction d'un point de vue logique ?

### 2.1. Problème de l'induction et solution par la logique déductive de Popper

Il existe des règles éprouvées qui garantissent la validité du raisonnement déductif. Ces règles se caractérisent par le fait qu'il est impossible de déduire une conclusion fautive de prémisses vraies. Le problème mis en évidence par Hume dans le livre I portant sur *l'entendement* du *Traité de la nature humaine* publié en 1739 est qu'il n'existe pas de règles comparables pour le raisonnement inductif. La critique de Hume concernant l'absence de preuve démonstrative de l'inférence inductive s'applique non seulement au cas des conjonctions certaines ou quasi certaines, illustré ci-dessous avec l'exemple des billes de billards (cf. section VI du livre I de Hume, 1995/1739 et section V de Hume, 1999/1748), mais aussi au cas des conjonctions probables (cf. section XII du livre I de Hume, 1995/1739 et section VI de Hume, 1999/1748). Au XVIII<sup>e</sup> siècle, à l'époque où Hume écrivait, il existait déjà une science formelle - la géométrie - qui avait été axiomatisée par Euclide au III<sup>e</sup> siècle avant notre ère et une science empirique qui venait d'être axiomatisée par Newton dans un ouvrage dont la première édition publiée en 1687 portait le titre de *Philosophiae Naturalis Principia Mathematica* (*Principes mathématiques de la philosophie naturelle*). L'objet de la philosophie naturelle dont il est question dans cet ouvrage est l'étude du rôle causal de différents types de forces (dont la gravitation) dans la détermination de différents types de mouvements, y compris les mouvements des corps célestes. Cette discipline empirique a été qualifiée de *mécanique rationnelle* par Newton et de *dynamique* par Leibniz.

La géométrie d'Euclide est fondée sur une série de définitions des notions qui seront utilisées, dont les notions de point, de ligne et de surface, et sur cinq axiomes. Les démonstrations des théorèmes reposent sur l'application des règles de la logique déductive. Ces règles garantissent notamment qu'il est impossible d'aboutir à des conclusions contradictoires. Si, après avoir démontré que la somme des angles d'un triangle est égale à deux angles droits, Euclide avait été amené à conclure, plus loin dans la dérivation des théorèmes, qu'elle est inférieure (ou supérieure) à deux angles droits, il aurait eu la certitude d'avoir commis au moins une erreur dans l'application des règles de la déduction et il serait revenu en arrière pour détecter à quelle étape cela s'était produit.

La théorie de Newton dont la forme finale se trouve dans la troisième édition des *Principia* (1999/1726) repose sur douze notions quantitatives (donc mathématisables), trois axiomes et six corollaires. Huit de ces notions reçoivent des définitions précises, dont celle de masse qui est une mesure de la quantité de matière qui dépend de la combinaison de la densité et du volume (définition 1) et celle de quantité de mouvement qui est une mesure du mouvement qui dépend de la combinaison de la vitesse et de la masse (définition 2). Les quatre notions ne recevant pas de définition précise - temps, espace, lieu (position dans l'espace) et mouvement - sont considérées comme suffisamment familières pour ne pas nécessiter de définition plus explicite, même si Newton attire l'attention sur trois distinctions

supplémentaires par rapport au sens commun qui interviendront dans leur utilisation (distinctions absolu/relatif, vrai/apparent et mathématique/commun).

Les trois axiomes et les six corollaires énoncent des *lois universelles* du mouvement dont Newton dit qu'elles sont acceptées par les mathématiciens et *confirmées* par l'expérience. La démonstration des autres lois se base sur les règles de la logique déductive, comme en géométrie. Les lois ainsi inférées de manière démonstrative sont aussi des lois empiriques susceptibles, en principe, d'être confirmées par l'expérience. La restriction imposée par le « en principe » est nécessaire parce qu'il se pourrait que certaines des lois déduites ne soient pas testables empiriquement pour des raisons pratiques.

L'obstacle principal dans l'élaboration d'un savoir empirique certain est que, contrairement à ce qui se passe en géométrie, la cohérence interne du système déductif ne garantit que partiellement la validité d'une théorie comme celle de Newton. Il faudrait en plus que les lois universelles, du moins celles qui sont susceptibles de confirmation empirique, correspondent *toujours* à des faits *vérifiables* par l'observation et l'expérimentation. Or, les vérifications (confirmations) de ces faits s'expriment sous forme d'énoncés non universels alors que les lois induites sont exprimées par des énoncés universels supposés valables en tout temps et en tout lieu. Comment justifier que le futur se conformera au passé sinon en recourant au principe de l'induction lui-même, ce qui nous enferme dans un cercle vicieux.

Malgré cette critique logique émise par Hume à l'encontre de l'induction de lois générales reposant sur la simple énumération de cas nombreux et jamais contradictoires, tout le monde croit que le soleil se lèvera demain et qu'une bille de billard A en mouvement rapide va, en la heurtant, causer la mise en mouvement d'une bille B qui était immobile avant le choc. Dans une situation de pari où 1 euro serait gagné si la bille B se mettait en mouvement dans les conditions énoncées et 10.000 euros dans le cas contraire, tout le monde choisirait de gagner 1 euro. Si, après avoir été heurtée par la bille A, la bille B restait immobile, tout le monde penserait qu'une cause cachée s'est opposée au cours normal des choses (peut-être que la bille B contient un noyau ferreux et qu'un puissant électro-aimant est dissimulé sous la table de billard). D'un point de vue psychologique subjectif, la force de conviction du modèle déterministe des sciences empiriques, ce que Hempel (1966) qualifie de modèle d'explication *déductive nomologique*, n'est donc pas réellement mise à mal par la critique de Hume, même si le problème épistémologique qu'il a soulevé demeure entier.

Ce que Popper (1991/1971, chapitre 1) appelle sa « solution au problème de l'induction » est une solution par élimination élaborée dans les années 30 dans un ouvrage célèbre dont le titre français est *La logique de la découverte scientifique* (1973/1959). Popper conteste le statut d'inférence inductive qui est attribué aux théories scientifiques et aux lois universelles. Selon lui, ces théories et ces lois ont le statut de conjectures, c'est-à-dire d'hypothèses qui peuvent être soumises à des tests. Ce changement de perspective permet à Popper d'adopter une position purement déductiviste qui repose sur le fait qu'un énoncé non universel peut entrer en contradiction avec des énoncés universels. L'énoncé universel « tous les cygnes sont blancs » ne peut pas être vérifié empiriquement mais il peut être réfuté par un énoncé comme « quelques cygnes ne sont pas blancs ». Or, un tel énoncé non universel est à la fois vérifiable et réfutable empiriquement. Une bonne théorie est une théorie qui spécifie a priori quelles seraient ses conditions de réfutation. Plus il y en a, meilleure est la théorie puisqu'elle offre plus de possibilités d'être soumise à des tests cruciaux ; plus souvent une

théorie aura effectivement été soumise à des tests cruciaux tout en y résistant, mieux elle sera corroborée.

Bien que la « solution » reposant sur la logique déductive apportée par Popper au problème de l'induction bénéficie d'une diffusion très large, elle n'est que peu admise et elle n'a pas empêché le développement d'une logique inductive reposant sur différents aspects de la notion de probabilité (Hacking, 2001, voyez Hacking et Dufour, 2004 pour l'adaptation française à laquelle je me réfère par la suite). C'est dans ce contexte que s'inscrit la discussion sur les différentes approches de l'inférence statistique qui suit.

## 2.2. Inférence statistique selon Fisher

La démarche du test de signification proposée par Fisher repose sur des travaux commencés au début des années 1920. Voici un résumé de cette démarche basé sur les deux premiers chapitres de *The design of experiments* (1935) :

formulation d'une hypothèse nulle ponctuelle concernant l'effet d'une certaine manipulation expérimentale sur un phénomène ;

calcul de la probabilité que le résultat observé s'écarte du résultat attendu sous l'hypothèse nulle d'une quantité aussi grande ou plus grande que celle qui est observée ;

rejet de l'hypothèse nulle si cette probabilité est faible – disons  $p \leq 0,05$  ;

non conclusion en faveur de l'hypothèse nulle lorsque ce critère n'est pas atteint ;

prise en compte des résultats non significatifs.

Cette démarche a été mise au point pour traiter des données expérimentales provenant de petits échantillons, des échantillons dont les effectifs sont trop faibles pour fournir des estimations précises des paramètres des populations dont ils sont extraits. Dans ces conditions, tout ce que l'on peut espérer, c'est de parvenir à une décision positive ou négative concernant la plausibilité d'une hypothèse. Le rejet d'une hypothèse nulle prend la forme d'une disjonction entre deux propositions : ou bien l'hypothèse nulle est vraie et quelque chose de très improbable s'est produit, ou bien l'hypothèse nulle est fausse.

Cependant, Fisher insiste sur le fait qu'aucune expérience isolée, aussi significative soit-elle, ne suffit à la démonstration d'un phénomène naturel. Un phénomène n'est expérimentalement démontrable que si l'on dispose d'une procédure expérimentale qui n'échoue que rarement à fournir des résultats statistiquement significatifs. On peut augmenter la sensibilité d'une expérience en augmentant la taille de l'échantillon. Une manière indirecte d'arriver à cet objectif consisterait à répéter l'expérience plusieurs fois puis à globaliser l'analyse en prenant en compte aussi bien les expériences qui ont conduit au rejet qu'au non rejet de l'hypothèse nulle.

Pour Fisher, l'approche par les tests de signification constitue une solution au problème de l'induction dans le contexte restreint des disciplines non encore constituées théoriquement. Il faut aussi souligner une différence fondamentale entre l'approche de Fisher et l'approche de Popper. Ce que Popper propose de soumettre à des tests expérimentaux sont des théories et des lois bien établies dont tout le monde espère qu'elles résisteront aux tests cruciaux, de manière à s'en trouver encore mieux corroborées. Ce que Fisher propose de soumettre à des tests statistiques sont des hypothèses nulles jugées très peu plausibles a priori dont tout le monde espère le rejet au profit d'hypothèses alternatives jugées plus plausibles.

### 2.3. Inférence statistique selon Neyman-Pearson

L'approche de Neyman-Pearson repose sur des prémisses différentes de celles de Fisher. Elle permet de se *comporter* de manière inductive sans toutefois contribuer à résoudre le problème logique du raisonnement inductif, même dans les limites étroites fixées par Fisher. Neyman et Pearson ont développé leur théorie des tests de signification dans une série de publications conjointes parues dans les années 1930 ; Neyman seul a développé la théorie des intervalles de confiance vers la fin de cette même décennie (voyez Neyman, 1977 pour une présentation non technique des deux théories). Plus tard, dans les années 1950, la théorie de Neyman-Pearson sera étendue en une théorie générale de la décision qui s'applique au diagnostic dans l'incertitude dans n'importe quelle situation conduisant à une décision binaire de type positif ou négatif (Swets et al., 2000). Ces deux aspects de la démarche initiée par Neyman et Pearson sont envisagés successivement.

#### 2.3.1. Cas de l'inférence statistique

Neyman et Pearson ne se satisfont pas d'une démarche dans laquelle seul un type d'erreur – rejeter l'hypothèse nulle ponctuelle dans le cas où elle est vraie – est pris en compte. Ils veulent aussi que les méthodes d'inférence statistique permettent de choisir entre plusieurs hypothèses concurrentes. Le cas le plus simple concerne le choix entre deux hypothèses généralement notées  $H_0$  et  $H_1$ .  $H_0$  est l'hypothèse nulle ponctuelle de Fisher,  $H_1$  est une hypothèse ponctuelle alternative. Une des hypothèses est considérée comme correcte ; la démarche du test de signification sert à déterminer laquelle.

Techniquement, le processus de décision revient à centrer une distribution d'échantillonnage de la statistique étudiée sur la valeur attendue sous  $H_0$  et une autre sur la valeur attendue sous  $H_1$ . C'est le recouvrement partiel entre les deux distributions qui rend le diagnostic incertain. Le critère de décision consiste en la valeur de la statistique de test en deçà de laquelle on accepte l'hypothèse nulle et au-delà de laquelle on la rejette. La représentation graphique de ce schéma de décision est suffisamment familière pour qu'on puisse se dispenser de la reproduire encore ici. Plaçons-nous dans le cas d'un test unilatéral destiné à rejeter l'hypothèse nulle en faveur d'une hypothèse alternative indiquant une valeur ponctuelle sous  $H_1$  supérieure à celle attendue sous  $H_0$ . Le critère de décision partage chaque distribution en deux régions dont les surfaces respectives représentent des probabilités complémentaires.

- Sous la courbe centrée sur  $H_0$  :
  - la surface située à gauche du critère représente la probabilité de ne pas rejeter  $H_0$  en faveur de  $H_1$  *conditionnellement* au fait que  $H_0$  soit vraie. Il s'agit de la probabilité d'une décision correcte notée  $1 - \alpha$  qu'on appelle le niveau de confiance du test ;
  - la surface située à droite du critère représente la probabilité de rejeter  $H_0$  en faveur de  $H_1$  *conditionnellement* au fait que  $H_0$  soit vraie. Il s'agit de la probabilité d'une décision erronée notée  $\alpha$  correspondant à l'erreur de première espèce (ou erreur de type I) qui est appelée niveau de signification du test.

- Sous la courbe centrée sur  $H_1$  :
  - la surface située à gauche du critère représente la probabilité de ne pas rejeter  $H_0$  en faveur de  $H_1$  *conditionnellement* au fait que  $H_0$  soit fausse. Il s'agit de la probabilité d'une décision erronée notée  $\beta$  appelée erreur de deuxième espèce (ou erreur de type II) ;
  - la surface située à droite du critère représente la probabilité de rejeter  $H_0$  en faveur de  $H_1$  *conditionnellement* au fait que  $H_0$  soit fausse. Il s'agit de la probabilité d'une décision correcte notée  $1 - \beta$  appelée puissance du test.

A cause du recouvrement des courbes centrées sur  $H_0$  et  $H_1$ , les probabilités des deux types de décisions erronées ( $\alpha$  et  $\beta$ ), bien que non complémentaires, varient en sens opposé et, pareillement, pour les probabilités des deux types de décisions correctes ( $1 - \alpha$  et  $1 - \beta$ ). La démarche de Neyman-Pearson consiste à minimiser les deux types d'erreurs. En général, ils souhaitent maintenir la probabilité de l'erreur de première espèce à une valeur très faible ( $\alpha = 0,05$  comme chez Fisher, voire souvent  $0,01$ ) tout en minimisant la probabilité de l'erreur de deuxième espèce ( $\beta$ ). Une formulation équivalente consiste à dire qu'ils veulent maximiser la puissance du test (puisque  $\beta$  et  $1 - \beta$  sont des probabilités complémentaires), c'est-à-dire maximiser la probabilité de rejeter l'hypothèse nulle si elle est fausse. Une fois  $\alpha$  choisi, le moyen le plus simple pour atteindre la puissance désirée consiste à augmenter le nombre d'observations.

Nous verrons à la section 3.1 que la démarche de Neyman-Pearson est souvent présentée comme s'il ne s'agissait que d'un simple perfectionnement de celle de Fisher. En réalité, ces deux démarches diffèrent pour au moins les trois raisons suivantes :

- La démarche de Neyman-Pearson est symétrique ; elle autorise aussi bien une décision en faveur de  $H_0$  au détriment de  $H_1$  (donc, une acceptation de  $H_0$ ) qu'un rejet de  $H_0$  en faveur de  $H_1$  (donc, une acceptation de  $H_1$ ). En revanche, l'approche de Fisher est asymétrique ; elle autorise seulement le rejet de  $H_0$  au profit d'une gamme d'hypothèses alternatives dont les valeurs ne sont pas spécifiées.
- Chez Neyman-Pearson, le niveau de signification n'est pas fixé une fois pour toute. C'est par une évaluation des coûts relatifs associés aux décisions erronées et des bénéfices relatifs associés aux décisions correctes qu'un niveau de signification et une puissance sont choisis. Ces préoccupations sont complètement étrangères à Fisher qui préconise de s'en tenir à un niveau de signification unique (souvent  $0,05$ , parfois  $0,01$  selon ses écrits) et pour qui les notions d'erreur de deuxième espèce et de puissance sont vides de sens puisqu'aucune hypothèse alternative ponctuelle n'est formulée.
- Neyman et Pearson donnent une interprétation qu'ils qualifient de *comportementale* au rejet ou au non rejet de  $H_0$ . En rejetant  $H_0$  au profit de  $H_1$ , on se comporte comme si  $H_0$  était fausse et  $H_1$  vraie même si on ne croit pas nécessairement que c'est le cas en réalité et sans disposer d'aucune indication sur les probabilités que  $H_0$  soit fausse et  $H_1$  vraie. Ils adoptent en cela une approche strictement *fréquentiste* (voyez section 2.4) qui exclut d'attribuer des probabilités aux valeurs possibles du paramètre étudié, cette interdiction s'appliquant donc aussi aux valeurs spécifiques du paramètre attendues sous  $H_0$  et sous  $H_1$ . En revanche, pour Fisher, le rejet de l'hypothèse nulle est interprété de manière *épistémique*, ce rejet rendant l'hypothèse nulle non seulement

moins probable (au sens *fréquentiste* du terme), mais aussi moins plausible en tant qu'explication causale du phénomène étudié.

C'est surtout cette troisième caractéristique de l'approche de Neyman-Pearson qui est à l'origine de la polémique très violente qui va les opposer à Fisher pendant plus de trois décennies. Les positions sont en effet inconciliables. Pour Fisher, l'inférence statistique doit contribuer à l'induction des théories scientifiques, du moins dans le cadre des disciplines non encore théorisées où les connaissances préalables sont peu nombreuses et incertaines. Les préoccupations de type utilitariste en rapport avec les coûts et bénéfices associés aux décisions correctes et erronées n'ont pas à entrer en ligne de compte dans le développement des théories scientifiques, même si ces préoccupations sont légitimes dans des domaines appliqués comme, par exemple, dans le contexte du contrôle de qualité des produits industriels.

### 2.3.2. Cas de la théorie de la décision

Durant les années 50, la théorie de Neyman-Pearson a été étendue de manière à fournir une solution générale au problème de la prise de décision dans l'incertitude. Le développement le plus connu des psychologues s'est d'abord incarné dans la *théorie de la détection* du signal en psychophysique (Tanner et Swets, 1954 ; Swets et al., 1961). Cette approche a ensuite été généralisée à la situation de diagnostic dans l'incertitude se traduisant par une décision positive ou négative selon qu'une certaine condition soit présente ou absente (Swets, 1996 ; Swets et al., 2000). Le tableau 1 indique les changements terminologiques par rapport au cas de l'inférence statistique vu précédemment. La courbe qui représente la distribution d'échantillonnage centrée sur  $H_0$  dans le cas du test de signification peut aussi représenter :

- la distribution des états liés à l'absence de stimulus (cas appelé *bruit*), dans une expérience de détermination du seuil absolu ;
- la distribution des valeurs d'une variable correspondant à des personnes ne souffrant pas d'une pathologie (donc, des personnes qualifiées de saines) en médecine et en psychologie ;
- la distribution des scores à un test de sélection correspondant à des personnes jugées *inaptés* à exercer une activité professionnelle spécifique.

La courbe qui représente la distribution d'échantillonnage centrée sur  $H_1$  dans le cas du test de signification représente alors :

- la distribution des états liés à la présence d'un stimulus (cas appelé *signal + bruit*) dans une expérience de détermination du seuil absolu ;
- la distribution des valeurs d'une variable correspondant à des personnes malades (cas appelés *pathologiques*) en médecine et en psychologie ;
- la distribution des scores à un test de sélection correspondant à des personnes jugées *aptés* à exercer une activité professionnelle spécifique.

Pour fixer les idées, détaillons le cas de la sélection du personnel en utilisant l'exemple fictif inventé par Anastasi (1994, pp. 166-168) pour illustrer à la fois le processus de validation d'un test de sélection du personnel et le processus de prise de décision dans l'incertitude. Anastasi imagine que les 100 candidats à un emploi aient passé un test d'aptitude. Les 100 candidats sont tous engagés pendant une période d'essai, quel que soit le résultat au

test. A l'issue de la période d'essai, le rendement au travail est mesuré selon les critères habituels de l'entreprise. Cette procédure permettrait d'évaluer la validité prédictive du test d'aptitude en calculant la corrélation entre les scores au test et la mesure de la performance au travail. Dans l'exemple développé par Anastasi, cette corrélation est légèrement inférieure à 0,70 et les données sont présentées sous forme d'un diagramme de dispersion avec les scores au test en abscisse et la mesure de la performance au travail en ordonnée.

Une ligne horizontale traversant le nuage de points indique le seuil minimum de performance au travail qui permet de distinguer les 60 candidats considérés comme suffisamment aptes des 40 considérés comme inaptes. Ceci illustre un taux de succès de 60% pour un taux de sélection de 100% (donc en engageant pour la période d'essai tous les candidats qui se sont présentés à l'embauche). Une ligne verticale traversant le nuage de points indique le seuil minimum d'aptitude basé sur les scores au test au-delà duquel 45 candidats auraient été sélectionnés et 55 rejetés. En prenant les deux critères en considération, les 100 candidats se répartissent comme suit dans les quatre catégories définies au tableau 1 : 33 rejets valides, 22 rejets erronés, 38 acceptations valides et 7 acceptations erronées. En n'engageant pour la période d'essai que les 45 candidats jugés aptes en fonction de leurs scores au test, l'entreprise améliorerait son taux de succès qui passerait de 60% en l'absence de test à 84% (38 acceptations valides/45 candidats présélectionnés). Du point de vue des candidats, l'utilisation de cette procédure de sélection apparaît moins favorable puisque parmi les 55 d'entre eux qui seraient refusés à l'embauche sur base du test, 22 (soit 40%) auraient été jugés aptes à l'emploi après la période d'essai.

		Etat du monde	
		Négatif	Positif
		$H_0$ vraie ( $H_1$ fausse) Bruit Sains Inaptes	$H_1$ vraie ( $H_0$ fausse) Signal + bruit Malades Aptes
Décision	Négative	Non rejet valide de $H_0$ Abstentions correctes Vrais négatifs Rejets valides	Erreur de type II Omissions Faux négatifs Rejets erronés
	Positive	Erreur de type I Fausse alerte Faux positifs Acceptations erronées	Rejet valide $H_0$ Détection correcte Vrais positifs Acceptations valides

Tableau 1

Terminologies de la théorie de la décision appliquées au test de signification, à la détection d'un signal en psychophysique, au diagnostic médical et psychologique et à la sélection du personnel.

## 2.4. Approche bayésienne de l'inférence statistique

Prenons comme point de départ l'énoncé de deux croyances erronées souvent associées à l'utilisation des tests de signification. Ces croyances constituent les propositions 5 et 6 numérotées en continuité avec celles déjà énumérées dans l'introduction. Supposons qu'un résultat soit significatif à un niveau exact de  $p = 0,05$ .

**Proposition 5.** Il est erroné de croire que  $p = 0,05$  est la probabilité que l'hypothèse nulle soit vraie et que son complément  $1 - p = 0,95$  est la probabilité que l'hypothèse alternative soit vraie ;

**Proposition 6.** Il est erroné de croire que la probabilité  $1 - p = 0,95$  est la probabilité de réplication du résultat significatif.

Ces deux interprétations erronées proviennent d'une triple confusion. La première confusion porte sur la distinction entre probabilité au sens absolu, non conditionnel, et probabilité au sens conditionnel. Le niveau de signification exact de  $p = 0,05$  représente la probabilité que la valeur de l'estimation du paramètre étudié que l'on peut symboliser par  $D$  (parce qu'il s'agit d'une valeur calculée sur les *données* empiriques) s'écarte de la valeur ponctuelle avancée sous l'hypothèse nulle si  $H_0$  est vraie, donc conditionnellement au fait que  $H_0$  soit vraie, ce que l'on symbolise comme suit :  $P(D|H_0)$ . Cependant, cette probabilité conditionnelle ne nous indique rien sur la probabilité non conditionnelle que  $H_0$  soit vraie. Donc, en cas de rejet de l'hypothèse nulle pour le seuil de signification exact de 0,05, nous ne savons pas plus quelle est la probabilité associée à l'hypothèse nulle a posteriori que nous ne le savions a priori. C'est pareil en ce qui concerne la probabilité associée à n'importe quelle hypothèse alternative ponctuelle qui, elle aussi, demeure inconnue tant en cas de rejet que de non rejet de l'hypothèse nulle. Or, sans connaître la probabilité associée à la valeur du paramètre qui a conduit au rejet, on ne peut pas connaître la probabilité de réplication du résultat significatif.

La deuxième confusion porte sur la distinction entre ce premier type de probabilité conditionnelle  $P(D|H_0)$  et la probabilité conditionnelle inverse  $P(H_0|D)$ , la probabilité que l'hypothèse nulle soit vraie conditionnellement à l'obtention d'une estimation de la valeur  $D$  du paramètre juste suffisante pour conduire au rejet de cette hypothèse nulle. Or, ces deux probabilités conditionnelles sont généralement différentes et l'une ne peut pas être dérivée de l'autre, sauf via le théorème de Bayes (Cohen, 1994 ; Dixon et O'Reilly, 1999 ; Nickerson, 2000). Voici comment ce théorème se traduit dans le cas de la détermination de  $P(H_0|D)$  :

$$P(H_0|D) = P(H_0) \times P(D|H_0) / P(D)$$

Le problème, c'est que les valeurs des deux probabilités non conditionnelles  $P(H_0)$  et  $P(D)$  sont inconnues et qu'il n'est généralement pas possible de les estimer de manière objective, ce qui nous conduit à la troisième confusion.

La troisième confusion porte sur la distinction entre la notion de probabilité au sens fréquentiste et la notion de probabilité au sens épistémique. Il serait plus correct de parler d'ignorance plutôt que de confusion parce que cette distinction est à peine effleurée, voire complètement passée sous silence, dans la plupart des cours standards de statistique dispensés aux psychologues. Or, la dualité de la probabilité apparaît dès l'émergence de cette notion au milieu du XVIIe siècle (Hacking, 1975/2002). La distinction entre ces deux types de

probabilité et les conséquences de cette distinction sont remarquablement bien expliquées au chapitre 11 de Hacking et Dufour (2004) ; en voici un bref résumé.

Au sens épistémique, la probabilité indique le degré de *croyance*, de *confiance* ou de *crédibilité* que l'on peut associer à des propositions du type « le réchauffement climatique actuel est en grande partie causé par l'activité humaine ». Au sens fréquentiste, la probabilité désigne la fréquence avec laquelle un dispositif aléatoire comme le jet répété d'une pièce de monnaie tend à fournir, par exemple, le résultat *face*. La probabilité est ici conçue comme une idéalisation de la notion de fréquence relative. Si la pièce est parfaitement équilibrée, la fréquence relative de l'événement *face* va progressivement se stabiliser autour de 0,50 au fur et à mesure de l'augmentation du nombre de jets. La probabilité au sens fréquentiste correspond à une caractéristique physique de la pièce qui est vraie ou fausse indépendamment de nos croyances. Cette propriété physique un peu particulière peut être qualifiée de *disposition* ou de *propension* parce que contrairement à d'autres propriétés physiques comme la masse ou la forme, elle ne se manifeste pas en permanence mais seulement quand la pièce est placée dans les circonstances appropriées, c'est-à-dire soumise à des jets successifs dont les résultats sont aléatoires.

Le point important souligné par Hacking et Dufour (2004) est que nous effectuons souvent un va-et-vient entre l'acception fréquentiste et l'acception épistémique de la probabilité en utilisant une sorte de règle qu'ils qualifient de *principe fréquentiste*. Pour comprendre ce principe, il faut contraster trois manières de justifier le crédit que l'on accorde à une proposition. Dans une situation fréquentiste pure, une proposition du type « cette pièce particulière semble bien équilibrée » peut être vérifiée empiriquement. En effet, on aurait des raisons de croire que cette proposition est vraie si, à l'issue d'une série de quelques dizaines ou de quelques centaines de lancers successifs, les fréquences relatives des deux événements possibles étaient proches de 0,50. A cause de cette possibilité de vérification empirique, la probabilité au sens fréquentiste a souvent été qualifiée d'*objective*.

Dans une situation épistémique pure, une proposition du type « il y a au moins huit chances sur dix pour que le réchauffement climatique actuel soit causé par l'activité humaine » exprime le degré de confiance élevé qu'une personne accorde aux différents arguments qui appuient cette thèse. Les notions de fréquence, de tendance et de propension ne s'appliquent pas puisque le réchauffement climatique actuel est un événement unique. Comme le degré de confiance accordé à une proposition varie d'une personne à l'autre, la probabilité au sens épistémique a souvent été qualifiée de *subjective*.

La situation de va-et-vient entre les deux acceptions de la notion de probabilité se produit, par exemple, dans la situation où une pièce serait lancée une seule fois et où un observateur prédirait que l'événement *face* a une chance sur deux de se produire pendant que la pièce est encore en l'air ou alors qu'elle a déjà atterri, mais à un endroit dissimulé à son regard. Le jugement de l'observateur sur le résultat de ce lancer unique de la pièce est épistémique au même titre que le jugement sur le réchauffement climatique actuel mais la justification de ce jugement repose sur la connaissance de la probabilité de l'événement *face* au sens fréquentiste. L'observateur attribue donc la valeur de la probabilité fréquentiste de l'événement *face* à la probabilité épistémique de l'événement *face* dans le cas du lancer unique d'une pièce particulière. C'est l'utilisation de ce lien entre probabilité fréquentiste et probabilité épistémique pour justifier une croyance que Hacking et Dufour (2004) qualifient de principe fréquentiste.

Le principe fréquentiste a une portée générale ; il intervient systématiquement dans toutes les situations de décision dans l'incertitude évoquées à la section 2.3.2. En outre, dans ces situations appliquées, la décision s'accompagne généralement d'une action dans un sens ou dans l'autre. Reprenons l'exemple de la sélection du personnel. La probabilité de .84 d'adaptation à la situation professionnelle des candidats ayant dépassé un certain score au test a été établie de manière empirique. Confronté à un candidat particulier qui dépasse le seuil au test, l'employeur utilise le principe fréquentiste pour lui attribuer une probabilité de 0,84 de réussite future dans l'activité professionnelle et il agit en conséquence en l'engageant. C'est en ce sens que Neyman et Pearson diraient que l'employeur se comporte de manière inductive.

Or, comme le souligne Birnbaum (1977), il n'y a pratiquement pas d'exemple dans les publications conjointes de Neyman et Pearson dans les années 1930 où une conclusion de type scientifique soit l'objectif principal. En effet, les situations étudiées par Neyman et Pearson sont pratiquement toujours des situations où les probabilités sont estimées empiriquement. Ceci n'empêche pas Neyman (1957) d'encore soutenir que la notion de comportement inductif s'applique aussi à la situation de la recherche scientifique et qu'elle est beaucoup plus appropriée que la notion de raisonnement inductif de Fisher. Il justifie sa position en montrant que Fisher lui-même, quand il se livre à des analyses statistiques plutôt qu'à des réflexions philosophiques, utilise les tests de signification pour se comporter d'une manière conforme aux résultats de l'expérience dans la poursuite de la recherche. Neyman oublie cependant que la démarche de Fisher avait pour objectif principal de tirer des conclusions d'expériences uniques ou peu nombreuses dans des domaines non constitués du savoir.

La même remarque vaut pour l'interprétation correcte de la signification d'un intervalle de confiance spécifique. L'erreur à éviter à tout prix est de dire qu'il y a une probabilité de 0,95 que la vraie valeur du paramètre étudié soit comprise dans les limites de l'intervalle de confiance observé. En effet, cette valeur du paramètre étant fixe, bien qu'inconnue, elle ne peut pas être probabilisée. Une autre formulation erronée consiste à dire qu'il y a une probabilité de 0,95 pour que l'intervalle observé recouvre la vraie valeur du paramètre. Or, les limites de cet intervalle de confiance sont aussi des valeurs fixes obtenues dans une expérience particulière, donc pas des valeurs probabilisables. Cependant, cette seconde erreur est moins grave que la première si on envisage la connexion entre cette interprétation et l'interprétation correcte via le principe fréquentiste. L'interprétation correcte est qu'en répétant un grand nombre de fois une expérience basée sur un même nombre d'observations, 95% des intervalles de confiance recouvriraient la vraie moyenne. On peut voir ce processus à l'œuvre dans des simulations sur ordinateur qui constituent de merveilleux outils didactiques. Mais, ces simulations sont artificielles parce que dans le contexte d'une recherche réelle, la valeur du paramètre est inconnue et on prélève un seul échantillon que l'on soumet à une manipulation expérimentale, pas une grande quantité d'échantillons.

Les deux approches de l'inférence statistique reposant sur les tests de signification, celles de Fisher et celle de Neyman-Pearson, ainsi que l'approche par les intervalles de confiance de Neyman, s'inscrivent dans un courant résolument fréquentiste. Elles excluent la perspective bayésienne parce que le recours à des probabilités qualifiées de personnelles ou de subjectives semble incompatible avec l'objectivité qui doit être la marque de l'approche scientifique. Avec le recul actuel, les raisons de l'incompréhension entre Fisher, d'une part, et

Neyman et Pearson, d'autre part, sont devenues très claires. Fisher s'intéresse à des expériences uniques dans lesquelles le principe fréquentiste ne peut pas s'appliquer au sens littéral du terme alors que la plupart des situations envisagées par Neyman et Pearson, en conduisant à des estimations des probabilités au sens fréquentiste, autorisent le recours au principe fréquentiste au sens littéral. Cependant, appliqué aux expériences uniques, le principe fréquentiste apparaît tout aussi abstrait et métaphorique chez Neyman-Pearson que chez Fisher.

### 3. Pratique de l'inférence statistique en psychologie

Selon Cohen (1994), ce qui est erroné dans l'approche par les tests de signification est « qu'elle ne nous indique pas ce que l'on veut savoir, et nous voulons tellement savoir ce que l'on veut savoir que, pour éviter le désespoir, nous croyons néanmoins qu'elle nous l'indique » (p. 997, ma traduction). Ce que l'on veut savoir, c'est, au vu de nouvelles données, quelle est la probabilité que l'hypothèse nulle soit vraie, donc  $P(H_0/D)$ . Or, le test de signification nous indique la probabilité des données si l'hypothèse nulle est vraie,  $P(D/H_0)$ . Lecoutre (2005 ; Lecoutre et Poitevineau, 2000) exprime la même chose en disant qu'en interprétant erronément les résultats des tests comme nous indiquant « ce que l'on veut savoir », nous nous comportons comme des « bayésiens qui s'ignorent ». C'est pourquoi Lecoutre pense, comme Rouanet (1996 ; Rouanet et al., 1978, 2000), qu'on en viendra inévitablement à adopter des méthodes d'inférence statistique d'inspiration bayésienne, tout en admettant que ce changement de perspective prendra du temps.

En attendant, ce qui peut être qualifié de quatrième approche de l'inférence statistique est une approche hybride qui a été développée dans les manuels de statistique rédigés pour des praticiens ayant des connaissances très élémentaires en mathématiques, ce qui ne devrait pas constituer un handicap, et aucune ou très peu de formation en logique, en philosophie des sciences et en épistémologie, ce qui est nettement plus regrettable. Les directives en matière d'inférence statistique que l'on trouve dans les éditions successives du manuel de publication de l'APA et dans les éditoriaux des journaux scientifiques même les plus exigeants en cette matière reflètent aussi cette approche hybride.

#### 3.1. Approche hybride des manuels de statistique

L'institutionnalisation en psychologie d'un type particulier de logique hybride de l'inférence statistique a été remarquablement bien analysée par Gigerenzer et Murray (1987, chapitre 1 ; voir aussi Gigerenzer et al., 1989, chapitre 3) et reprise de manière humoristique par Gigerenzer (1993) dans un chapitre intitulé « le *surmoi*, le *moi* et le *ça* dans le raisonnement statistique ». Cette logique hybride se traduit par une utilisation quasi aveugle et mécanique, encore renforcée par la grande disponibilité de logiciels puissants de traitement statistique des données, d'une approche de l'inférence statistique qui ignore la complexité de ses origines et qui oscille entre une interprétation fréquentiste et une interprétation épistémique de la probabilité. Brièvement, on peut dire que l'approche hybride se fonde principalement sur l'approche de Fisher décrite à la section 2.2 (le *moi*), mais en ignorant la subtilité. Elle y associe des éléments et une terminologie qui émanent de l'approche de Neyman-Pearson (le *surmoi*), comme s'il s'agissait d'un perfectionnement de celle de Fisher, donc, en passant sous

silence les profondes différences théoriques évoquées à la section 2.3.1 qui sous-tendent ces deux approches. Cette fusion boiteuse entre deux conceptions partiellement incompatibles de l'inférence statistique, mâtinées d'interprétations bayésiennes implicites (le *ça*) est à l'origine de nombreuses confusions dont celles qui ont été signalées dans ce qui précède sous la forme des propositions 1 à 6.

### 3.2. Evolution des directives de l'APA en matière d'inférence statistique

La première édition du manuel de l'APA datant de 1952 indiquait simplement qu'une attention spéciale doit être accordée aux tests de signification ainsi qu'à la logique de l'inférence et de la généralisation à partir de données empiriques. Un bon point de départ pour tracer l'évolution de ces directives depuis lors est constitué par l'éditorial publié par Melton (1962) à l'issue de 12 années passées dans la fonction d'éditeur du *Journal of Experimental Psychology*.

#### 3.2.1. Editorial de Melton (1962)

Melton (1962) résumait les critères qu'il avait utilisés pour accepter ou rejeter les articles expérimentaux qui lui étaient soumis. Ces critères portaient sur (1) l'adéquation du schéma expérimental pour la collecte des données sur le problème énoncé, (2) l'adéquation des mesures et des tests statistiques employés, (3) la validité de l'argumentation logique liant les données aux conclusions et (4) la confiance qui peut être accordée dans la possibilité de reproduire les résultats de l'expérience dans des conditions décrites. Sur ce dernier point, Melton faisait état de ses fortes réticences par rapport à l'acceptation à la publication de résultats significatifs seulement au niveau 0,05, mais il affirmait aussi que c'est une erreur de croire qu'il a accordé un rôle décisif uniquement à l'atteinte d'un niveau de signification de 0,01. D'autres critères ont été pris en compte, comme le degré de nécessité logique qui lie le résultat à la théorie et à d'autres résultats de la littérature ainsi que la cohérence des résultats à travers une série d'expériences. Ces critères ont joué un rôle important dans le rejet de *résultats négatifs*, mais l'argument principal pour un tel rejet a souvent été le manque de sensibilité de la manipulation expérimentale (évaluée comment ? Melton ne nous le dit pas).

Deux conclusions peuvent être dégagées de ce bref résumé de la position de Melton (1962). La plus négative est que l'éditeur d'une des revues les plus réputées pour sa rigueur méthodologique a contribué à véhiculer l'erreur grossière qui consiste à croire que le niveau de signification est un indicateur de la probabilité de réplification des résultats (cf. proposition 6 à la section 2.4). L'autre conclusion est que, bien que sa position soit surtout inspirée par Fisher, elle est beaucoup moins subtile et elle s'en écarte fortement en excluant pratiquement la possibilité de publier des résultats non significatifs (qualifiés de négatifs), alors que Fisher (1935) plaidait pour leur publication et leur évaluation combinée avec celle des résultats significatifs (cf. section 2.2).

#### 3.2.2. Deuxième édition du manuel de l'APA (1974)

Dans la seconde édition du manuel de l'APA datant de 1974, il est demandé de rapporter les résultats sous forme de tableaux ou de figures, en incluant ceux qui vont à

l'encontre de l'hypothèse. Les résultats des tests de signification doivent comporter des informations concernant le nom du test utilisé, la valeur de la statistique de test, le nombre de degrés de liberté et le niveau de signification. Suivent des exemples de formes de rédaction appropriées dans le texte. Par exemple, le groupe 1 est significativement différent du groupe 2,  $t(22) = 2.62, p < 0,01$  ; l'analyse de la variance montre un effet significatif du facteur étudié,  $F(1, 34) = 123,07, p < 0,001$ . Ces exemples sont révélateurs du fait que le rapport du niveau de signification exact n'est pas requis mais que, implicitement, un crédit plus grand est sans doute accordé à un résultat significatif à 0,001 plutôt qu'à 0,01. Enfin, il est aussi recommandé de ne pas inférer une tendance quand un résultat atteint presque un des niveaux de signification usuels. De tels résultats doivent être considérés comme reflétant uniquement des fluctuations aléatoires, ce qui est une position qui favorise l'assimilation d'un résultat non significatif à l'absence totale d'effet.

Pour ce qui est de l'acceptation des « résultats négatifs », trois cas sont distingués. Le premier est celui où une théorie bien établie prédit clairement qu'un résultat significatif aurait dû être trouvé. Le poids de la preuve incombe au chercheur ; il est d'autant plus grand que le corpus de données soutenant la théorie est plus grand. Le second cas est celui où le résultat non significatif provient de la correction d'une erreur méthodologique. Le troisième cas concerne l'obtention d'un résultat non significatif à partir d'une réplique exacte, à l'échantillon près, d'une expérience. Ce cas est jugé peu intéressant, du moins tant que le résultat non significatif n'est pas confirmé sur plusieurs échantillons différents.

### 3.2.3. Quatrième édition du manuel de l'APA (1994)

Les recommandations de la quatrième édition du manuel de l'APA sont en continuité avec celles des éditions précédentes, tout en devenant plus exigeantes sur plusieurs points. Il faut que les statistiques soient suffisantes. Dans le cas de tests concernant une ou plusieurs moyennes ceci implique que, d'un point de vue descriptif, les moyennes doivent être associées à des mesures de variabilité comme l'écart type ou la variance. Alternativement, on peut fournir les moyennes accompagnées du carré moyen de l'erreur et des degrés de liberté associés à l'effet testé. Il est demandé aussi de se préoccuper de la puissance du test d'hypothèse, en particulier quand on souhaite ne pas rejeter l'hypothèse nulle, et de se préoccuper de la taille des effets.

### 3.2.4. Cinquième édition du manuel de l'APA (2001)

Les années 1990 ont été fertiles en discussions et critiques de l'utilisation des tests de signification. L'essentiel des questions soulevées avait déjà été traité dans trois articles publiés dans le *Psychological Bulletin* (Bakan, 1966 ; Lykken, 1968 ; Rozeboom, 1960), une des revues scientifiques majeures en psychologie, et dans un article moins accessible de Meehl (1967). Ces critiques ont été largement ignorées ainsi que leurs reprises par d'autres dans les années 1970 et 1980. En revanche, leurs prolongements dans les années 1990 se sont avérés un peu moins faciles à contourner et ont eu des conséquences sur les recommandations actuelles de l'APA. Dans la deuxième édition de leur manuel de statistique, Loftus et Loftus (1988) avaient été parmi les premiers à prôner une approche alternative par les intervalles de confiance.

Cette entreprise s'est poursuivie dans une série de publications importantes (Loftus, 1993a, 1993b, 1996, 2002 ; Loftus et Masson, 1994 ; Masson et Loftus, 2003).

En réponse à la reprise de ces critiques par Cohen (1990, 1994), l'APA a réuni un comité d'experts dont la mission était de fournir de nouvelles propositions en vue de la révision de la quatrième édition du manuel de l'APA. Avant d'entamer cette révision, le comité a jugé utile de diffuser un rapport préliminaire afin de susciter la discussion. Ce rapport signé Wilkinson *and the Task Force on Statistical Inference* a été publié en août 1999 dans l'*American Psychologist* ; la cinquième édition du manuel de l'APA a suivi en 2001.

En matière de présentation statistique, le manuel se place immédiatement en retrait par rapport aux positions nettement plus engagées du rapport du comité d'experts. Comme remarque préliminaire, il est dit que le champ de la psychologie n'a pas un avis unique sur les différentes controverses entourant l'utilisation des tests de signification et que ce n'est pas le rôle du manuel de publication de résoudre ces controverses, ni de prendre position en la matière. La plupart des recommandations sont les mêmes que dans la quatrième édition sauf sur le point important suivant :

le fait de rapporter des intervalles de confiance (pour des estimations des paramètres, pour des fonctions des paramètres comme des différences entre moyennes et pour des tailles d'effets) peut constituer un moyen très efficace de présentation des résultats. Comme les intervalles de confiance combinent de l'information sur la localisation et la précision et qu'ils peuvent souvent être utilisés directement pour inférer les niveaux de signification, ils constituent, en général, la meilleure stratégie de rapport [des résultats]. L'utilisation des intervalles de confiance est donc fortement recommandée (Manuel de l'APA, 5<sup>ème</sup> édition, 2001, p. 22, ma traduction).

#### **4. Au-delà des tests de signification**

##### 4.1. Puissance des tests de signification

Bien que la notion de puissance fasse partie de la logique hybride véhiculée par les manuels de statistique depuis le début des années 60, elle a été largement ignorée dans la pratique. Dans l'immense majorité des publications dans les différents domaines de la psychologie, les auteurs se contentent de contrôler le risque de première espèce et de ne publier ou, du moins, de n'interpréter que les effets significatifs. Or il est connu, depuis l'étude empirique effectuée par Cohen en 1962, que la puissance des tests de signification est généralement trop faible pour autoriser des conclusions fermes à partir des études individuelles. Cependant, il est nécessaire de bien comprendre sur quelles bases repose ce diagnostic afin de prévenir tout nouveau malentendu.

Deux conditions doivent être remplies pour pouvoir se livrer à une comparaison des puissances atteintes dans des études de nature différente (Cohen, 1962, 1988). D'abord, il faut, pour chaque type de test de signification, disposer d'un indice de la taille des effets qui s'exprime par des nombres « purs », c'est-à-dire des nombres qui ne sont pas associés à des unités de mesure. Ensuite, il faut pouvoir établir des correspondances entre les valeurs des différents indices associés à différents types de tests de signification. Illustrons cette démarche dans le cas de la comparaison de deux moyennes et dans celui de l'étude du degré de liaison linéaire entre deux variables.

Dans le cas d'une différence entre deux moyennes, la taille brute de l'effet est mesurée dans la même unité que la variable dépendante (par exemple en points de QI ou en temps de réaction en msec). Il suffit de standardiser cette différence en la divisant par l'estimation de l'écart type commun aux deux conditions pour obtenir un nombre pur qui n'est plus associé à aucune unité de mesure. Ce nombre est un indice de la taille relative de l'effet appelé  $d$  par Cohen. Cohen (1962, 1988) propose comme convention d'appeler « petits », « moyens » et « grands » des effets dont les ordres de grandeur de  $d$  sont respectivement de 0,20, 0,50, et 0,80. Ces valeurs de  $d$  seraient obtenues, par exemple, avec un test de QI dont l'écart type est de 15 pour des différences moyennes de 3, 7,5 et 12 points entre deux groupes. Avec des temps de réaction moyens et un écart type estimé de 80 msec, des différences de 16, 40 et 64 msec entre les moyennes de deux groupes indépendants donneraient aussi des valeurs de  $d$  de 0,20, 0,50 et 0,80.

Dans le cas de la mesure du degré de liaison linéaire entre deux variables variant de manière continue, le coefficient de corrélation  $r$  de Pearson constitue un indice adéquat de la taille de l'effet puisque  $r$  est une mesure standardisée, donc un nombre pur. Cohen montre que des  $d$  de 0,20, 0,50 et 0,80 correspondent approximativement à des valeurs de  $r$  de 0,10, 0,30 et 0,50 (cf. Cohen, 1988, p. 82). Cette procédure de mise en correspondance des différents indices de la taille des effets s'est poursuivie de manière à conférer une signification comparable à des effets dont les tailles sont conventionnellement qualifiées de petites, moyennes et grandes quel que soit le type de test de signification utilisé.

Dans son analyse de 70 études publiées en 1960 dans le *Journal of Abnormal and Social Psychology*, Cohen (1962) a montré que les puissances moyennes pour détecter des effets petits, moyens et grands étaient respectivement de 0,18, 0,48 et 0,83 (pour un niveau de signification de 0,05 au test bilatéral). Un quart de siècle plus tard, Sedlmeier et Gigerenzer (1989) trouvaient des résultats similaires dans leur analyse de 54 articles publiés en 1984 dans la partie de la même revue scientifique maintenant appelée *Journal of Abnormal Psychology*, les puissances médianes étant de 0,14, 0,44 et 0,90. Ce diagnostic se généralise assez bien à l'ensemble des domaines de la psychologie (voyez Maxwell, 2004 pour d'autres références). On tend à résumer cette information en disant que la puissance pour le rejet d'une hypothèse nulle en faveur d'une hypothèse alternative portant sur un effet de taille moyenne est de l'ordre de 0,40 pour un test bilatéral.

Cette conclusion devrait être préoccupante parce qu'elle est révélatrice des tensions sous-jacentes à l'interprétation des données inférentielles. La démarche des psychologues consiste à vouloir rejeter des hypothèses nulles auxquelles ils ne croient pas en étant très prudents à propos du risque de les rejeter à mauvais escient (quand elles sont vraies), donc en maintenant le niveau de signification à des valeurs inférieures à 5%. Ce faisant, ils acceptent, implicitement du moins, un risque de l'ordre de 60% de ne pas rejeter ces hypothèses nulles quand elles sont fausses en faveur d'hypothèses alternatives vraies représentant des effets dont les tailles sont *conventionnellement* qualifiées de moyennes par Cohen (1988).

C'est par l'omission de l'adverbe « conventionnellement » mis en exergue dans la conclusion du paragraphe précédent qu'un nouveau malentendu, ou même un double malentendu peut surgir ; un double malentendu qui tire son origine de la propriété 1 des tests de signification (cf. introduction) qui met en garde contre la confusion entre la notion de « signification » au sens statistique et la notion de « signification » au sens théorique ou

pratique. Tout chercheur devrait être capable d'avancer une valeur de la taille absolue (la taille non standardisée) de l'effet qu'il juge digne d'intérêt pour servir d'hypothèse alternative. Il peut pour cela recourir à des arguments théoriques ou se fonder sur une évaluation des tailles des effets habituellement considérées comme critiques dans son domaine d'expertise. C'est seulement en l'absence de tels critères que Cohen suggère de recourir à son opérationnalisation de la taille des effets, tout en mettant en garde contre les risques associés à cette pratique ; voici ce qu'il dit :

pour les indices de la taille des effets associés à chaque test statistique, l'auteur propose, *comme une convention*, des valeurs de la taille des effets qui servent de définitions opérationnelles pour les adjectifs qualificatifs « petit », « moyen » et « grand ». Ceci est une opération dangereuse à maints égards : les définitions sont arbitraires, des concepts qualitatifs tels que 'grand' sont tantôt compris comme absolus, tantôt comme relatifs ; et donc, ils courent le risque d'être mal compris (Cohen, 1988, p. 12, ma traduction, les italiques sont de l'auteur).

Cette précaution permet d'éviter un premier malentendu dans l'interprétation de la notion de puissance dans le contexte d'une recherche particulière associée à l'utilisation d'un test statistique spécifique. Le second malentendu possible provient de l'équivalence établie entre les tailles conventionnelles des effets obtenus dans des études utilisant des tests statistiques différents. Or, cette équivalence fondée sur les relations d'égalité au sens mathématique qui peuvent être établies entre différentes statistiques de test sont surtout intéressantes dans le contexte des études empiriques portant sur la puissance générale des recherches dans un domaine du savoir, comme dans celles réalisées par Cohen (1962) et Sedlmeier et Gigerenzer (1989). En revanche, cette équivalence est trompeuse au niveau des recherches individuelles.

Prenons un exemple. Dans des expériences standards dont les latences moyennes de réponses varient souvent dans une gamme de 500 à 700 msec (cf. section 4.4 pour un exemple), j'aurais tendance à considérer comme théoriquement négligeables des effets inférieurs à 10 msec, même s'ils sont statistiquement significatifs. Ce jugement est indépendant du fait que cette différence de 10 msec serait conventionnellement qualifiée de « petite » dans le contexte d'une étude inter-groupes d'écart type commun de 80 msec ou de « moyenne » dans le contexte d'une étude intra-groupe dont l'écart type de l'effet serait de 20 msec. Par contre, dans les domaines appliqués de la psychologie, il est probable (ou du moins souhaitable) que beaucoup de praticiens se refuseraient à utiliser des tests psychologiques affichant une validité prédictive mesurée par un coefficient de corrélation de 0,30, même si cet effet est conventionnellement qualifié de « moyen » dans l'opérationnalisation de Cohen et même si cette corrélation est statistiquement significative. En effet, une telle corrélation implique que seulement 9% de la variance du critère sont expliqués par le prédicteur. L'examen du diagramme des résidus montrerait donc une variabilité résiduelle du critère qui serait quasi aussi grande que sa variabilité totale.

Enfin, il faut se rappeler que toute cette discussion a lieu dans un contexte strictement fréquentiste, un contexte dans lequel les valeurs du paramètre étudié ne sont pas probabilisées. Ici aussi, il faut se garder de confondre probabilité au sens absolu et probabilité au sens conditionnel. La seule interprétation correcte des études empiriques portant sur la puissance

atteinte par les recherches en psychologie constitue la septième et dernière proposition numérotée de cet article.

**Proposition 7.** La puissance pour le rejet d'une hypothèse nulle *conditionnellement* à l'intérêt pour des effets dont les tailles sont *conventionnellement* qualifiées de moyennes est de l'ordre de 0,50, voire plus faible.

#### 4.2. Pour une mise au ban des tests de signification

A la question de savoir s'il faut bannir les tests de signification posée par Shrout (1997) dans l'introduction d'une section spéciale consacrée à ce problème dans *Psychological Science*, Hunter (1997) répondait catégoriquement oui. Il estime en effet que l'utilisation des tests de signification a eu des conséquences désastreuses pour les sciences sociales en empêchant l'élaboration d'un savoir cumulatif. Cet état des choses serait dû essentiellement au manque de puissance statistique des études individuelles ainsi qu'à des croyances erronées concernant l'interprétation des résultats des tests de signification. Le chercheur qui croit naïvement que le niveau d'erreur est de 5% se trompe complètement puisque ce risque d'erreur est conditionnel au fait que l'hypothèse nulle soit vraie. Or, une hypothèse nulle ponctuelle est pratiquement toujours fautive (voyez Meehl, 1967). Donc, en réalité, le risque de première espèce a une probabilité nulle. En revanche, comme la puissance des tests de signification est généralement inférieure à 0,50 et que les résultats s'expriment par une décision binaire de rejet ou de non rejet, autant dire que le niveau de précision atteint par le processus de décision est, en moyenne, inférieur à celui qui serait atteint en jouant à pile ou face.

Cette argumentation ainsi que le remède proposé ont été développés dans de nombreuses publications (Schmidt, 1996 ; Schmidt et Hunter, 1997) ainsi que dans un livre entier consacré aux méthodes de méta-analyses (Hunter et Schmidt, 1990). Prenons la simulation effectuée par Schmidt (1996) pour illustrer leur démarche. Elle est basée sur une étude de la validité prédictive d'un test de sélection portant sur 1455 employés du tri postal (Schmidt et al., 1985). La corrélation bisérielle de point entre le score au test et un critère dichotomique de réussite dans l'activité de tri postal n'est que de 0,22, mais cette corrélation est significativement différente de 0. Sachant que l'effectif médian des études de validation en sélection du personnel est de 68, Schmidt et al. ont prélevé aléatoirement (sans remise) 21 échantillons de 68 données parmi les 1455 disponibles. Ceci leur a permis de se livrer à une simulation de ce que seraient les résultats de 21 études différentes de validation portant sur une même aptitude professionnelle évaluée avec un même test. Seules 8 corrélations (38%) étaient significativement différentes de 0. En concluant à l'absence de validité prédictive dans les cas où les corrélations sont non significatives, on se tromperait dans 62% des cas puisque l'hypothèse nulle d'une corrélation de 0 est fautive alors que celle d'une corrélation de 0,22 est vraie.

En revanche, si pour chaque échantillon, on rapportait la valeur du coefficient de corrélation (qui est une mesure de la taille de l'effet) associée à son intervalle de confiance (plutôt que le résultat du test de signification), on s'apercevrait qu'il existe un recouvrement important entre les intervalles dans les différentes études. Ensuite, en se livrant à une méta-analyse des 21 résultats, on conclurait que le coefficient de validité moyen est égal ou proche de sa valeur réelle de 0,22 et que toute la variabilité entre les résultats (conduisant tantôt à des

corrélations significatives, tantôt à des corrélations non significatives) proviendrait des erreurs d'échantillonnage. Cette conclusion est à la fois correcte et inévitable quand on se livre à une méta-analyse des résultats obtenus par le processus artificiel de simulation décrit ci-dessus. Cependant, depuis une trentaine d'années, Hunter, Schmidt et leurs associés ont défendu l'idée que cette conclusion peut être étendue aux revues de la littérature portant sur la validité prédictive de tests dans les situations naturelles de sélection du personnel.

La thèse de Hunter et Schmidt est la suivante. Quand on se livre à une revue de la littérature de type narratif, on s'aperçoit que des tests apparemment comparables censés prédire la performance dans des situations professionnelles apparemment comparables fournissent néanmoins des coefficients de validité très variables, dont certains sont statistiquement différents de 0 et d'autres pas. L'interprétation traditionnelle de tels résultats est de considérer que la validité prédictive des tests est spécifique des situations qui ont servi de support au processus de validation, ce qui a pour inconvénient majeur que ce processus de validation devrait être sans arrêt recommencé pour s'adapter à la spécificité des situations. Or, selon Hunter et Schmidt, cette conclusion est erronée. En fait, la validité des tests est généralisable parce que la variabilité inter-études des coefficients de validité est largement due aux erreurs d'échantillonnage et à d'autres sources adventices de variations qui peuvent être contrôlées et corrigées grâce au processus de méta-analyse (Hunter et Schmidt, 1990).

En définitive, la puissance trop faible des tests de signification dans les études individuelles et l'inventaire des rejets et non rejets de l'hypothèse nulle dans les synthèses des résultats de plusieurs études contribueraient à donner l'image d'un savoir non cumulatif. Le remède proposé par Hunter (1997) et Schmidt (1996, Schmidt et Hunter, 1997) consiste, pour les études individuelles, à rapporter les estimations ponctuelles de l'amplitude (la taille) des effets avec leurs intervalles de confiance au lieu des tests de signification. Dans l'intégration des résultats de la littérature, ces données doivent servir d'ingrédients pour des techniques de méta-analyses, comme celles qu'ils ont initialement développées pour s'opposer à l'hypothèse de la spécificité des situations dans l'établissement de la validité des tests de sélection du personnel (Hunter et Schmidt, 1990). Selon eux, ces méthodes sont de portée tout à fait générale, elles peuvent aussi s'appliquer aux études reposant sur une méthodologie expérimentale plutôt que corrélationnelle.

Sur ce dernier point, comme praticien des méthodes expérimentales et des revues de la littérature de type narratif, je suis résolument convaincu du contraire, mais ce point ne sera pas développé ici. Notez simplement que l'élaboration de la méthode de méta-analyse décrite ci-dessus s'est effectuée dans le contexte étroit de la défense des procédures de généralisation de la validité prédictive (*validity generalization procedures*) des tests psychologiques. Or, certains aspects spécieux de cette argumentation ont été dénoncés par James et al. (1986) et la plupart des arguments développés par Hunter (1997), Schmidt (1996) et Schmidt et Hunter (1997) en faveur de l'utilisation de leur méthode de méta-analyse ont été jugés fallacieux par Mulaik et al. (1997). L'espace manque pour entrer dans les détails mais disons qu'une des erreurs importantes commises par ces auteurs est de ne pas avoir apprécié correctement le caractère conditionnel des conclusions qui peuvent être tirées des études empiriques de la puissance atteinte dans les recherches en psychologie, comme exprimé dans la proposition 7 à la fin de la section précédente.

#### 4.3. Utilisation des rapports de vraisemblance pour évaluer des explications alternatives

La perspective développée par Dixon (1998, 2003 ; Dixon et O'Reilly, 1999 ; Glover et Dixon, 2004) n'a pas suscité un grand intérêt jusqu'à ce jour. Pourtant, elle mérite qu'on s'y arrête parce qu'elle répond à beaucoup de critiques adressées à la pratique de l'inférence statistique selon l'approche hybride véhiculée par les manuels de statistique. Dixon et ses collaborateurs rappellent une vérité première : ce qui intéresse le chercheur n'est pas tellement la décision de rejeter ou de ne pas rejeter l'hypothèse nulle mais la possibilité de choisir laquelle parmi plusieurs explications théoriques différentes rend le mieux compte des données. Or, la description abstraite des données empiriques sous forme d'une collection d'effets principaux et d'interactions significatifs et non significatifs n'offre qu'un lien très indirect avec les explications théoriques dont le chercheur souhaite évaluer les plausibilités relatives. Pour rendre ce lien plus direct, Dixon propose de recourir au *rapport de vraisemblance maximum* plutôt qu'au test de signification.

Envisageons le cas le plus élémentaire où seulement deux explications alternatives d'un même phénomène sont comparées et appelons ces explications modèle 1 ( $M_1$ ) et modèle 2 ( $M_2$ ). Admettons aussi que les chercheurs d'une communauté scientifique se partagent en deux groupes : ceux qui pensent que le modèle 1 est plus plausible que le modèle 2 et ceux qui pensent, au contraire, que le modèle 2 est plus plausible que le modèle 1. Ces croyances peuvent être traduites par le rapport  $P(M_1)/P(M_2)$  entre les probabilités épistémiques associées a priori à chaque modèle, ce rapport ayant une valeur supérieure à 1 pour les chercheurs du premier groupe et une valeur inférieure à 1 pour ceux du second groupe.

En se fondant sur une approche bayésienne, on évaluerait comment ce rapport est modifié suite à une nouvelle prise de données en utilisant le théorème de Bayes comme suit :

$$\frac{P(M_1|D)}{P(M_2|D)} = \frac{\frac{P(M_1) \times P(D|M_1)}{P(D)}}{\frac{P(M_2) \times P(D|M_2)}{P(D)}}$$

Après simplification dans le membre de droite, on obtient :

$$\frac{P(M_1|D)}{P(M_2|D)} = \frac{P(M_1)}{P(M_2)} \times \frac{P(D|M_1)}{P(D|M_2)}$$

Le deuxième terme  $P(D|M_1)/P(D|M_2)$  du membre de droite est le rapport de vraisemblance. C'est par l'intermédiaire de ce rapport de vraisemblance que le théorème de Bayes permet d'établir une relation entre le rapport des probabilités a priori  $P(M_1)/P(M_2)$  et le rapport des probabilités a posteriori  $P(M_1|D)/P(M_2|D)$  des deux modèles. Pour éviter de prêter le flanc aux critiques habituelles associées à l'utilisation du théorème de Bayes, donc de devoir se prononcer sur les probabilités a priori des deux modèles, Dixon et O'Reilly (1999) proposent de se limiter à utiliser le rapport de vraisemblance maximum dans une optique purement descriptive.

Le rapport de vraisemblance maximum entretient aussi une relation avec la démarche classique du test de signification. Par exemple, le test d'une hypothèse nulle au sens de Fisher correspond au cas où le modèle 1 ne prédit pas de différence entre les conditions et le modèle 2 prédit une différence d'une certaine amplitude (mais dont la valeur est laissée indéterminée). Le rapport de vraisemblance n'est rien d'autre que le rapport entre deux densités de probabilités. Admettons que la distribution d'échantillonnage soit normale. Si l'hypothèse nulle est vraie, la densité de probabilité pour la valeur attendue sous cette hypothèse nulle (modèle 1) est de 0,3989. Si la différence observée est nulle, la vraisemblance maximum de cette valeur du paramètre s'obtient en centrant aussi la distribution d'échantillonnage sur cette différence nulle, ce qui donne un rapport de vraisemblance  $M_1/M_2 = 0,3989/0,3989 = 1$ . Si la différence observée est faible, le rapport de vraisemblance reste proche de 1. Si le test est juste significatif au niveau 0,05 bilatéral, la densité de probabilité de cette valeur du paramètre sur la courbe centrée sur l'hypothèse nulle (modèle 1) est de 0,0584 (pour une valeur de  $z$  de +1,96 ou de -1,96), ce qui donne un rapport de vraisemblance  $M_1/M_2 = 0,0584/0,3989 = 0,146$  en défaveur de l'hypothèse nulle (modèle 1), soit en inversant, un rapport  $M_2/M_1 = 0,3989/0,0584 = 6,830$  en faveur de l'existence d'un effet de l'amplitude observée dans l'expérience (modèle 2).

Dixon (1998) a montré que le rapport de vraisemblance est en relation quasi linéaire avec le niveau de signification, du moins dans la gamme des niveaux de signification usuels et pour les distributions d'échantillonnage courantes. Il suggère d'adopter le rapport de  $M_2/M_1 = 10$  comme critère, ce qui correspond en gros à un niveau de signification inférieur à 0,05. Cependant, ce rapport doit être utilisé seulement pour résumer l'information apportée par une expérience, pas comme un critère de décision. Notez que ce rapport peut aussi bien être calculé dans un sens que dans l'autre ( $M_1/M_2 = 0,10$ ), donc aussi bien en faveur qu'en défaveur de l'hypothèse nulle. En outre, on peut décider de suspendre son jugement quand le rapport  $M_2/M_1 < 10$  ou le rapport  $M_1/M_2 > 0,10$ .

Enfin, rappelons que, dans sa formulation originale, l'approche de Neyman-Pearson permet aussi de symétriser le processus de décision. En se livrant à quelques réflexions sur le développement de leur théorie, Pearson (1962) a rappelé que Neyman et lui avaient d'abord exploré le degré avec lequel l'inférence et le processus de décision pouvaient être basés sur le rapport de vraisemblance. Cependant, ils ont finalement choisi d'exprimer les résultats en termes de régions de rejet parce que cette manière de faire s'accordait mieux avec leur idéal d'objectivité, donc en évitant de faire entrer des jugements personnels dans le processus de décision (et aussi pour se placer dans le prolongement de l'approche de Fisher).

#### 4.4. Exemple d'utilisation des intervalles de confiance au lieu des tests de signification

L'avantage de l'utilisation des intervalles de confiance sur les tests de signification sera illustré par une recherche effectuée en collaboration avec Katia Duscherer (Duscherer et Holender, 2002). Notre point de départ était un résultat de Fox (1996) montrant un effet d'amorçage sémantique négatif dans une tâche de décision lexicale avec amorces inconscientes. Nous doutions de la validité de ce résultat et nous pensions que l'effet d'amorçage devrait disparaître une fois un certain nombre de corrections apportées à la procédure. C'est ce qui s'est produit et c'est ici que les difficultés ont commencé. Comment faire accepter à la

publication une recherche qui repose sur l'hypothèse qu'un effet ne doit pas se produire, donc sur l'acceptation plutôt que le rejet d'une hypothèse nulle, et sur l'obtention d'un résultat non significatif ?

Par rapport aux attitudes très négatives en ces circonstances décrites aux sections 3.2.1 et 3.2.2, une certaine évolution s'est quand même produite. Par exemple, dans son éditorial du *Journal of Experimental Psychology: General*, Hunt (1990) a énoncé trois conditions pour l'admission à la publication de résultats négatifs : (1) il faut de sérieuses raisons théoriques pour justifier que l'absence d'effet d'une manipulation soit informative ; (2) il faut que le test statistique soit suffisamment puissant pour garantir la fiabilité de l'effet nul ou quasi nul ; (3) il faut montrer que la procédure expérimentale et le matériel sont suffisamment sensibles pour fournir des résultats positifs dans des conditions différentes mais pertinentes par rapport à la question étudiée.

Avant de poursuivre, il est indispensable de décrire plus en détails la situation expérimentale dont il est question ici. La tâche de décision lexicale consiste à indiquer le plus rapidement possible si un stimulus cible constitue un mot ou un pseudomot (une séquence de lettres orthographiquement légale mais ne constituant pas un mot dans la langue du test). Lors de chaque essai d'une expérience standard d'amorçage sémantique, le stimulus cible est précédé d'un stimulus amorce constitué par un mot, les deux stimuli étant présentés en vision centrale de manière à être parfaitement lisibles, donc disponibles à la conscience. Ce stimulus amorce n'appelle généralement aucune réponse. On observe alors un effet d'amorçage sémantique de signe positif : le temps de réaction moyen est plus court pour les mots cibles qui sont précédés par des mots amorces sémantiquement associés (ex : amorce *pain*, cible *beurre*) que pour ceux précédés par des mots amorces non sémantiquement associés (ex : amorce *pain*, cible *docteur*). Pour autant que les associations sémantiques soient fortes, l'amplitude des effets d'amorçage rapportés dans la littérature varie dans une gamme de 40 à 60 msec.

Fox (1996, expérience 1) a apporté deux modifications principales à la procédure standard. Le stimulus amorce était constitué d'un chiffre présenté en vision centrale et d'un mot amorce présenté en deux exemplaires en vision périphérique, une fois au-dessus et une fois en dessous du chiffre. Pour se faire une idée concrète des distances impliquées, imaginez que si le chiffre occupait la ligne que vous lisez pour le moment, le mot amorce serait présenté trois lignes plus haut et trois lignes plus bas, les lignes intermédiaires étant laissées blanches. Lors de chaque essai, deux réponses manuelles consécutives étaient requises, la première pour indiquer si le chiffre amorce était inférieur ou supérieur à cinq et la seconde pour indiquer si le stimulus cible était un mot ou un pseudomot. La nécessité de répondre au chiffre obligeait les participants à maintenir la fixation. Dans ces conditions, un seul participant sur les 19 testés a dû être écarté pour s'être aperçu de la présence de mots dans le stimulus amorce.

Fox (1996) prédisait qu'avec des mots amorces indisponibles à la conscience, l'effet d'amorçage s'inverserait, ce qui fut le cas dans son expérience 1 : le temps de réaction moyen était plus court pour les mots cibles précédés par des mots amorces non associés (ex : amorce *pain*, cible *docteur*) que pour ceux précédés par des mots amorces sémantiquement associés (ex : amorce *pain*, cible *beurre*). Cet effet que l'on peut qualifier d'effet d'amorçage sémantique négatif, puisque son signe est inversé par rapport à celui obtenu dans la situation standard, avait une amplitude de 35 msec.

Nous avons confirmé que la procédure de Fox garantit l'indisponibilité à la conscience des mots amorces. Cette affirmation devrait être accompagnée d'une série de réserves parce que le scepticisme quant à l'existence d'effets inconscients repose généralement sur l'insuffisance des preuves de l'indisponibilité à la conscience des stimuli critiques (Holender, 1986a, 1986b ; Holender et Duscherer, 2004). Ici, nous admettons que ces preuves sont suffisantes. En revanche, nous pensions satisfaire pleinement à la condition 1 de Hunt en disposant de raisons théoriques suffisantes pour affirmer qu'aucune des conditions requises pour engendrer un effet d'amorçage négatif ne sont réunies dans l'expérience 1 de Fox et cela, que les mots amorces soient ou ne soient pas disponibles à la conscience. Nous avons longuement étayé cette position dans l'introduction de notre article (Duscherer et Holender, 2002) qui contenait aussi une revue exhaustive de la littérature sur les effets d'amorçage sémantique observés dans des conditions comparables à celles étudiées par Fox.

Voyons ensuite le degré avec lequel nous avons pu satisfaire à la condition 3 de Hunt. Nous avons d'abord effectué une expérience préliminaire pour nous assurer qu'un effet d'amorçage de grande amplitude pouvait être obtenu avec notre matériel. Nous avons obtenu un effet d'amorçage positif de 47 msec, ce qui est conforme aux effets obtenus dans la littérature avec la procédure standard. Idéalement, il aurait fallu pouvoir vérifier directement nos hypothèses en obtenant un effet d'amorçage sémantique négatif substantiel dans une réplique exacte de l'expérience de Fox. La comparaison avec l'effet nul obtenu dans notre expérience 1 (Duscherer et Holender, 2002) effectuée avec notre procédure corrigée aurait permis de clore la question. Cette voie nous était malheureusement fermée parce que nous ne disposions pas de toute l'information nécessaire pour reproduire exactement l'expérience de Fox.

Finalement, c'est le degré avec lequel nous avons pu satisfaire à la condition 2 de Hunt qui s'est avéré décisif. Il nous fallait montrer que notre résultat non significatif ne témoignait pas simplement d'un manque de puissance du test statistique. Notre objectif a été d'obtenir une puissance de 0,95 pour un effet de 20 msec pour un test bilatéral au seuil de signification de 0,05.

Le choix d'un test bilatéral correspond aux recommandations les plus courantes en psychologie (Cohen, 1965 ; Maxwell, 2004) ; il se justifie aussi par le fait que Fox prédisait un effet d'amorçage négatif, nous prédisions un effet nul ou quasi nul et d'autres chercheurs prédiraient un effet d'amorçage positif. Le choix de la taille de l'effet est plus contestable. Il s'agit d'un choix partiellement opportuniste lié au fait que Fox avait trouvé un effet substantiel de 35 msec. Le critère plus exigeant de 20 msec nous permettait de parer aux critiques portant sur le manque éventuel de puissance de notre étude comparée à la sienne. Toutefois, ce critère de 20 msec est trop peu exigeant par rapport à notre objectif qui était de pouvoir conclure en faveur d'un effet nul. En revanche, le choix d'une puissance de 0,95 au lieu de la puissance conventionnelle de 0,80 est cohérent avec notre objectif. En voulant démontrer l'absence d'effet, il est logique d'exiger que le risque de deuxième espèce ne soit pas plus grand que le risque de première espèce, du moins si on adopte l'approche de Neyman-Pearson dans sa forme originale (cf. section 2.3.1) plutôt que dans sa forme dévoyée par l'approche hybride de l'inférence statistique (cf. section 3.1), donc, en s'autorisant aussi bien de conclure en faveur de l'hypothèse nulle qu'en faveur de l'hypothèse alternative. Cet objectif est encore mieux satisfait

en recourant au calcul du rapport de vraisemblance selon la méthode préconisée par Dixon et O'Reilly (1999) décrite à la section précédente.

Expérience	<i>n</i>	Effet d'amorçage	Limites de confiance pour 1- $\alpha = 0,95$	Puissance pour 20 msec	Taille de l'effet pour une puissance de 0,95
1	32	-1 (24)	-8,9 ; +8,7	0,99	16
1a	16	0 (25)	-13,5 ; +13,5	0,84	25
1b	16	-1 (24)	-13,3 ; +12,3	0,87	23
2	32	0 (19)	-6,4 ; +7,4	0,99	13
2a	16	-6 (19)	-16,3 ; + 4,3	0,99	19
2b	16	+7 (17)	-2,1 ; +16,1	.0,99	16
PA	16	+5 (25)	-8,4 + 18,2	.0,80	24

Tableau 2

Effets d'amorçage (en msec) avec les écarts types entre parenthèses, limites des intervalles de confiance à 95%, puissances pour déceler un effet de 20 msec et tailles requises des effet pour une puissance de 0,95 pour le test bilatéral au seuil de 0,05 dans les expériences 1 et 2 et l'expérience de pseudo-amorçage (PA) de Duscherer et Holender (2002).

Nous n'avons pas pu atteindre une puissance de 0,95 en utilisant l'unité minimale de 16 participants requise par notre plan expérimental. En définitive, l'expérience 1 (Duscherer et Holender, 2002) comprenait deux répliquations (appelées expériences 1a et 1b) basées sur notre procédure corrigée, chacune comprenant 16 participants, de l'expérience 1 de Fox (1996) qui comprenait 18 participants. Le tableau 2 montre que l'effet d'amorçage moyen était virtuellement nul, tant dans l'expérience 1 globale que dans les deux sous-expériences composantes. Cependant, les effets individuels étaient assez variables (écart type de l'effet de 24 msec), certains participants montrant des effets positifs et d'autres des effets négatifs de tailles non négligeables. Pour vérifier que ces fluctuations de l'effet étaient aléatoires, nous avons effectué une expérience de pseudo-amorçage basée sur 16 participants dans laquelle les amorces étaient supprimées mais où l'effet d'amorçage était calculé comme si elles avaient été présentées. L'effet de pseudo-amorçage de 5 msec était non significatif et l'écart-type de cet effet était le même que dans les sous-expériences 1a et 1b. Enfin, dans l'expérience 2, le chiffre était supprimé de l'amorce, ce qui entraînait que les participants fournissaient une seule réponse lors de chaque essai au lieu de deux dans l'expérience 1. Ici aussi, l'expérience globale donnait un effet d'amorçage nul, même si les sous-expériences fournissaient des effets non significatifs de -6 et +7 msec.

Le tableau 2 montre aussi que les tailles des effets requises pour obtenir une puissance de 0,95 sont de 16 msec dans l'expérience 1 et 13 msec dans l'expérience 2. En outre, pour chaque sous-expérience, nous avons calculé le rapport de vraisemblance selon la méthode de Dixon et O'Reilly (1999) présentée à la section 4.3 en prenant comme valeurs associées aux modèles 1 et 2 des effets d'amorçage de 0 et 20 msec. Dans tous les cas, le modèle 1 était plus de 10 fois plus plausible que le modèle 2. Notez qu'à la fois l'information

sur les puissances calculées a posteriori et l'information fournie par les rapports de vraisemblance doivent être envisagées d'un point de vue strictement descriptif.

Le plaidoyer pour l'utilisation des intervalles de confiance à la place des tests de signification s'est accompagné de diverses recommandations en matière de représentations graphiques ainsi que d'instructions concernant l'établissement des intervalles tant pour les plans expérimentaux à mesures répétées que non répétées (Estes, 1997 ; Loftus, 2002 ; Loftus et Masson, 1994 ; Masson et Loftus, 2003 ; Saville, 2003). Dans la figure 1, les résultats des expériences 1 et 2 de Duscherer et Holender (2002) sont repris en suivant la procédure suggérée par Cumming et Finch (2005, voyez leur figure 3)

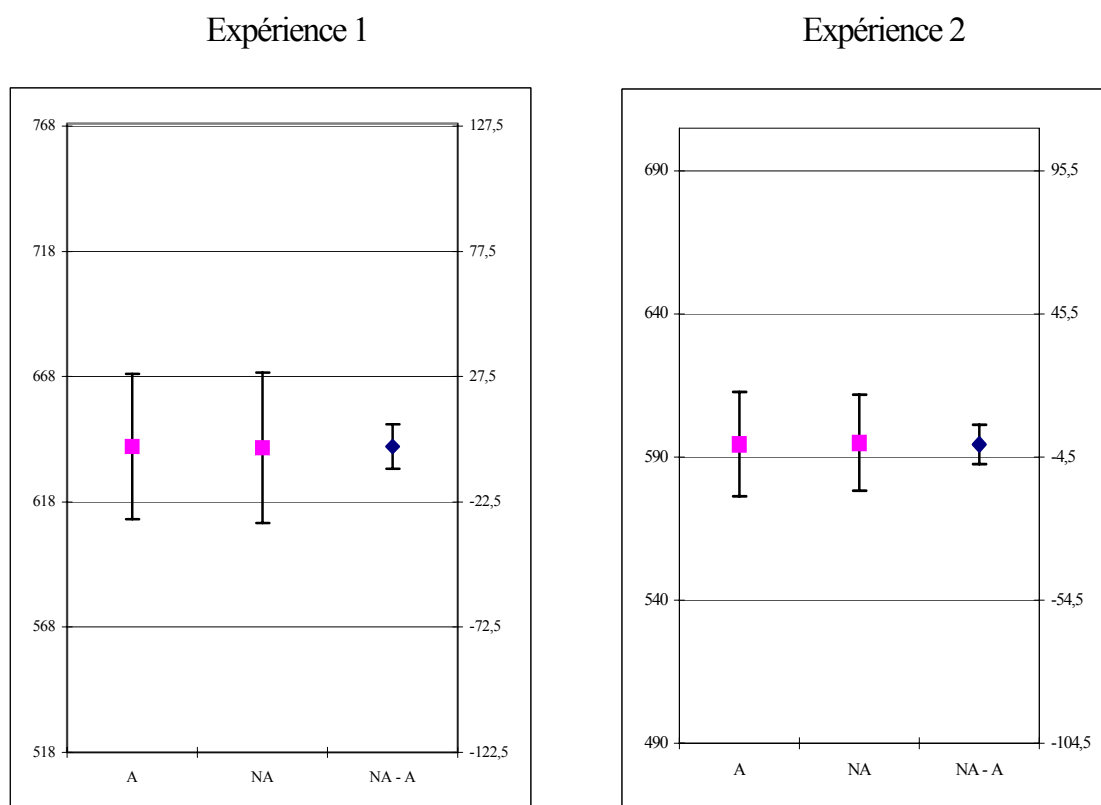


Fig. 1. Temps de réaction moyens (en msec) et intervalles de confiance pour les essais associés (A) et non associés (NA) ainsi que pour l'effet d'amorçage (NA - A) dans les expériences 1 et 2 de Duscherer et Holender (2002). L'origine de l'axe de droite est alignée sur la moyenne des essais non associés <sup>(1)</sup>.

Pour chaque expérience, l'ordonnée gauche de la figure 1 permet de représenter les temps de réactions moyens pour les essais associés ( $M_A$ ) et non associés ( $M_{NA}$ ). Les barres d'erreurs représentent les intervalles de confiance à 95% autour de chacune de ces moyennes. L'ordonnée droite sert à représenter l'effet d'amorçage moyen calculé par la différence entre les moyennes des essais non associés et des essais associés ( $M_{NA} - M_A$ ).

<sup>1</sup> Je remercie Bernard Fekete de vari pour son aide dans la construction de la figure 1. Idéalement, il faudrait pouvoir rendre l'échelle de l'ordonnée gauche des figures indépendante du niveau des moyennes et de la largeur des intervalles de confiance autour de ces moyennes. Il faudrait aussi pouvoir indiquer le repère de l'origine sur l'ordonnée droite. Nous n'y sommes pas encore parvenus ; Geoff Cumming non plus (communication personnelle du 17-09-07).

L'origine de l'ordonnée droite (correspondant à un effet d'amorçage nul) est alignée avec la moyenne des essais non associés. Les barres d'erreurs entourant l'effet d'amorçage représentent l'intervalle de confiance à .95 de la différence entre les moyennes qui, dans le cas qui nous occupe, proviennent de mesures répétées sur les mêmes participants. Voyons comment cette figure peut être utilisée pour dégager les avantages de l'approche par les intervalles de confiance soulignés par Cumming et Finch (2001, 2005), dont il a été fait état dans l'introduction.

D'un point de vue descriptif, la figure 1 permet de s'apercevoir en un coup d'œil que la performance globale et l'effet d'amorçage sont moins variables dans l'expérience 2 que dans l'expérience 1. On peut aussi voir que les effets d'amorçage sont non significatifs puisque les limites de confiance recouvrent l'origine de l'ordonnée droite (et aussi parce que la moyenne  $M_A$  est à l'intérieur de ces limites de confiance alors que l'origine de l'ordonnée droite se trouve au niveau de la moyenne  $M_{NA}$ ). Enfin, on dispose aussi, toujours en un seul coup d'œil, d'une information sur la précision des estimations qui est au moins aussi intéressante que l'information sur la puissance, voire plus intéressante. Ce dernier point mérite d'être développé.

Voyons d'abord la valeur de 29 msec de la demi-largeur de l'intervalle de confiance pour la moyenne  $M_A = 640$  msec de l'expérience 1. Cette valeur nous indique qu'en testant la conformité de cette moyenne avec une moyenne attendue sous l'hypothèse nulle de  $640 + 29 = 669$  msec, le test de signification basé sur la valeur de la distribution  $t$  à 31 degrés de liberté serait tout juste significatif pour le niveau de 0,05 en bilatéral. Ce serait pareil pour le test de conformité par rapport à une moyenne attendue de  $640 - 29 = 611$  msec. L'intervalle de confiance peut donc nous fournir exactement la même information que le test de signification. Il est aussi possible de traduire cette information en une estimation de la puissance du test. Nous ne pouvons pas connaître la valeur de la moyenne de la population mais la valeur la plus vraisemblable de cette moyenne est la moyenne trouvée dans l'échantillon, donc 640 msec. Ceci implique que par rapport à l'hypothèse nulle de 669 msec, l'hypothèse alternative de 640 msec, soit une valeur située à une distance d'une demi-largeur de l'intervalle de confiance, peut être décelée avec une puissance approximative de 0,50. La puissance approximative pour déceler une différence de 58 msec, soit une valeur située à une distance d'un intervalle de confiance entier de 640 msec, serait de 0,975.

Le même raisonnement vaut pour la puissance pour déceler une différence d'une certaine amplitude entre deux moyennes, mais ici, deux cas doivent être envisagés. Dans le cas d'observations indépendantes et en admettant que les valeurs des intervalles de confiance soient les mêmes autour des deux moyennes, ce serait une erreur de croire que la seconde moyenne serait juste significativement différente de la première pour une différence de 29 msec. En effet, l'intervalle de confiance autour de la différence entre les deux moyennes est  $\sqrt{2}$  fois plus large que la moyenne des intervalles de confiance individuels. Il faudrait donc une différence de  $29 \times 1,41 = 40,9$  msec pour que les deux moyennes diffèrent. Dans le cas des observations répétées, comme à la figure 1 gauche, on bénéficie de la très forte corrélation positive entre les moyennes des essais associés et non associés, si bien qu'il suffit d'une différence de 8,9 msec pour que la différence entre les deux moyennes soit significative.

Supposons qu'on ait trouvé un effet d'amorçage négatif de 8,9 msec tout juste significatif à .05. Nous aurions été forcés de conclure dans le même sens que Fox (1996) en dépit de la modicité de cet effet comparé au sien qui valait 35 msec. L'intérêt de la présentation

des résultats sous forme de graphique incluant les intervalles de confiance autour des effets est qu'on peut aussi dériver de l'information sur la probabilité de reproduire un résultat significatif (Cumming et Maillardet, 2006 ; Killeen, 2005). Par exemple, admettons que l'effet observé soit de 8,9 msec dans la population, quelle serait la probabilité de reproduire cet effet juste significatif dans une deuxième expérience. La réponse est environ 0,50 puisque la demi-largeur de l'intervalle de confiance correspond à une puissance d'environ 0,50, donc aussi à une erreur de deuxième espèce d'environ 0,50.

## 5. Conclusions

L'approche hybride de l'inférence statistique telle qu'elle s'est développée en psychologie depuis une cinquantaine d'années contient un certain nombre de vices rédhibitoires dont la nature et les conséquences ont souvent été dénoncées. Certains vont jusqu'à penser que le carcan imposé par la pratique des tests de signification est largement responsable du caractère non cumulatif de la recherche (par exemple, Hunter, 1997 ; Schmidt, 1996 ; Schmidt et Hunter, 1997). Les critiques de ces pratiques formulées dans les années 1960 (Bakan, 1966 ; Lykken, 1968 ; Meehl, 1967 ; Rozeboom, 1960) sont passées inaperçues en dehors d'un tout petit cercle de psychologues intéressés par les problèmes méthodologiques. Depuis lors, la reprise et l'amplification de ces critiques, ainsi que le développement de solutions alternatives, sont devenus un genre en soi donnant lieu à de nombreuses publications dans des revues largement diffusées, ce qui les rendent plus difficiles à ignorer involontairement. A cela, il faut ajouter la prise de position officielle de l'APA contre l'utilisation abusive des tests de signification (*Wilkinson and the Task Force on Statistical Inference*, 1999).

Malheureusement, l'espoir suscité par cette initiative de l'APA (Lecoutre et Poitevineau, 2000) a été suivi d'une grande déception lors de la parution de la cinquième édition du manuel de publication. Dans des communications personnelles recueillies par Fidler (2002), de nombreux protagonistes du débat pensent que l'ambiguïté de la position de l'APA risque d'être utilisée comme un encouragement à ne rien changer aux pratiques anciennes (voyez aussi Finch et al., 2002). Par exemple, à la page 22 du manuel, l'utilisation des intervalles de confiance est « fortement recommandée » parce qu'il s'agit de « la meilleure stratégie de report » (cf. section 3.2.4). Cependant, deux paragraphes plus haut, il est écrit que ce n'est pas le rôle du manuel de publication de prendre position dans les débats entourant l'utilisation des tests de signification. En outre, la tentative de promotion de l'utilisation des intervalles de confiance par Loftus (1993b) durant ses quatre années de mandat comme éditeur de *Memory and Cognition* n'a pas vraiment été un succès, pas plus que d'autres initiatives du même genre en médecine. On apprend, par exemple, que beaucoup de chercheurs tendent à confondre les barres d'erreurs qui représentent des intervalles de confiance avec celles qui représentent des erreurs standards (Belia et al., 2005). Même quand les intervalles de confiance sont inclus de manière systématique dans les rapports de recherche (surtout dans les tableaux, beaucoup plus souvent que dans des figures), ils ne sont pas réellement utilisés pour justifier l'interprétation des données, ce qui fait dire à Fidler et al. (2004) qu'en amenant les chercheurs à utiliser les intervalles de confiance, les éditeurs ne parviennent pas pour autant à les faire penser.

Certes, mais c'est d'abord à l'enseignement universitaire qu'il incombe d'apprendre à penser, pas aux éditeurs de journaux scientifiques. Une formation en statistique ne peut pas se limiter à faire apprendre l'utilisation quasi mécanique des procédures statistiques. La mission première d'une telle formation est de prévenir et d'empêcher les interprétations erronées et les conclusions abusives que l'utilisation non raisonnée de ces méthodes peut engendrer. Mais ceci est insuffisant, cet enseignement se doit aussi d'aborder des questions plus substantielles, à caractère épistémologique, comme celles évoquées dans les sections 2 et 3 du présent article. Le problème de l'induction ne se pose plus dans les termes formulés par Hume à l'époque du triomphe du modèle déterministe des sciences empiriques, alors que le concept de probabilité venait juste d'émerger. Maintenant que *l'empire de la probabilité* ('*the empire of chance*' Gigerenzer et al., 1989) s'étend sur pratiquement tous les domaines du savoir, la nature du problème de l'induction a elle aussi changé, le raisonnement inductif apparaissant comme probabiliste par essence. En quatrième de couverture de *An introduction to probability and inductive logic* de Hacking (2001), l'ouvrage est qualifié d'accessible à des étudiants variés, pas seulement aux étudiants en philosophie, même s'il n'ont pas reçu de formation en logique symbolique. L'accessibilité de l'ouvrage devrait être encore plus grande pour les étudiants en psychologie qui reçoivent une formation substantielle en statistique ; il me paraîtrait donc normal que les contenus de cet ouvrage soient progressivement distillés au gré de cette formation.

## Références

- American Psychological Association, Council of Editors. (1952). Publication manual of the American Psychological Association. *Psychological Bulletin*, 49 (Suppl., Pt. 2), 389-449.
- American Psychological Association. (1974). *Publication manual of the American Psychological Association* (2<sup>ème</sup> éd.). Washington, DC.
- American Psychological Association. (1994). *Publication manual of the American Psychological Association* (4<sup>ème</sup> éd.). Washington, DC.
- American Psychological Association. (2001). *Publication manual of the American Psychological Association* (5<sup>ème</sup> éd.). Washington, DC.
- Anastasi, A. (1994/1988). *Introduction à la psychométrie*. Guérin Universitaire, Montréal [traduction des parties I et II de *Psychological testing* (6<sup>ème</sup> éd.), 1988. Macmillan, New York].
- Bakan, D. (1966). The test of significance in psychology research. *Psychological Bulletin*, 66, 423-437.
- Belia, S., Fidler, F., Williams, J. & Cumming, G. (2005). Researchers misunderstand confidence intervals and standard error bars. *Psychological Methods*, 10, 389-396.
- Birnbaum, A. (1977). The Neyman-Pearson theory as decision theory, and as inference theory; with a criticism of the Lindley-Savage argument for Bayesian theory. *Synthese*, 26, 19-49.
- Cohen, J. (1962). The statistical power of abnormal-social psychological research : A review. *Journal of Abnormal and Social Psychology*, 69, 145-153.
- Cohen, J. (1965). Some statistical issues in psychological research. In : Wolman, B.B. (Ed.), *Handbook of Clinical Psychology*. McGraw-Hill, New York, pp. 95-121.
- Cohen, J. (1988). *Statistical analysis for the behavioral sciences* (2<sup>ème</sup> éd.). Lawrence Erlbaum Associates, Hillsdale, NJ.

- Cohen, J. (1990). Things I have learned (so far). *American Psychologist*, 45, 1304-1312.
- Cohen, J. (1994). The Earth is round ( $p < .05$ ). *American Psychologist*, 49, 997-1003.
- Cumming, G. & Finch, S. (2001). A primer on the understanding, use, and calculation of confidence intervals that are based on central and noncentral distributions. *Educational and Psychological Measurement*, 61, 532-574.
- Cumming, G. & Finch, S. (2005). Inference by eye : Confidence intervals and how to read pictures of data. *American Psychologist*, 60, 170-180.
- Cumming, G. & Maillardet, R. (2006). Confidence intervals and replication : Where will the next mean fall ? *Psychological Methods*, 11, 217-227.
- Dixon, P. (1998). Why scientists value  $p$  values. *Psychonomic Bulletin & Review*, 5, 390-396.
- Dixon, P. (2003). The  $p$ -value fallacy and how to avoid it. *Canadian Journal of Experimental Psychology*, 57, 189-202.
- Dixon, P. & O'Reilly, T. (1999). Scientific versus statistical inference. *Canadian Journal of Experimental Psychology*, 53, 133-149.
- Duscherer, K. & Holender, D. (2002). No negative semantic priming from unconscious flanker words in sight. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 839-853.
- Estes, W.K. (1997). On the communication of information by displays of standard errors and confidence intervals. *Psychonomic Bulletin & Review*, 4, 330-341.
- Fidler, F. (2002). The fifth edition of the APA publication manual : Why its statistics recommendations are so controversial. *Educational and Psychological Measurement*, 62, 749-770.
- Fidler, F., Thomason, N., Cumming, G., Finch, S. & Leeman, J. (2004). Editors can lead researchers to confidence intervals, but can't make them think : Statistical reform lessons from medicine. *Psychological Science*, 15, 119-126.
- Finch, S., Thomason, N. & Cumming, G. (2002). Past and future American Psychological Association guidelines for statistical practice. *Theory and Psychology*, 12, 825-853.
- Fisher, R.A. (1935). *The design of experiments*. Hafner Publishing Company, New York.
- Fox, E. (1996). Cross-language priming from ignored words : Evidence for a common representational system in bilinguals. *Journal of Memory and Language* 35, 353-370.
- Gigerenzer, G. (1993). The superego, the ego, and the id in statistical reasoning. In : Keren, G., Lewis, C. (Eds.), *A handbook for data analysis in the Behavioral sciences : Methodological issues*. Lawrence Erlbaum Associates, Londres, pp. 311-339.
- Gigerenzer, G. & Murray, D.J. (1987). *Cognition as intuitive statistics*. Lawrence Erlbaum Associates, Londres.
- Gigerenzer, G., Swijtink, Z., Porter, T., Daston, L., Beatty, J. & Krüger, L. (1989). *The empire of chance*. Cambridge University Press, Cambridge.
- Glover, S. & Dixon, P. (2004). Likelihood ratios : A simple and flexible statistic for empirical psychologists. *Psychonomic Bulletin & Review*, 11, 791-804.
- Hacking, I. (2002/1975). *L'émergence de la probabilité*. Seuil, Paris [traduction de *The emergence of probability*. Cambridge University Press, 1975].
- Hacking, I. (2001). *An introduction to probability and inductive logic*. Cambridge University Press, Cambridge.
- Hacking, I. & Dufour, M. (2004). *L'ouverture au probable : éléments de logique inductive*. Armand Colin, Paris.
- Hempel, C.G. (1966). *Philosophy of natural science*. Prentice-Hall, Englewood Cliffs, NJ.

- Holender, D. (1986a). Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking : A survey and appraisal. *Behavioral and Brain Sciences*, 9, 1-23.
- Holender, D. (1986b). Conceptual, experimental, and theoretical indeterminacies in research on semantic activation without conscious identification. *Behavioral and Brain Sciences*, 9, 50-66.
- Holender, D. & Bertelson, P. (1975). Selective preparation and time uncertainty. *Acta Psychologica*, 39, 193-203.
- Holender, D. & Duscherer, K. (2004). Unconscious Perception : The need for a paradigm shift. *Perception and Psychophysics*, 66, 872-881
- Hume, D. (1995/1739). *L'entendement. Traité de la nature humaine*, livre I et appendice. GF Flammarion, Paris [traduction de *Treatise on human nature*, livre I, 1739].
- Hume, D. (1999/1748). *Enquête sur l'entendement humain*. Le Livre de Poche, Paris [traduction de *Philosophical essays concerning human understanding*, 1748].
- Hunt, E. (1990). The traditional editorial statement. *Journal of Experimental Psychology : General*, 119, 3-4.
- Hunter, J.E. (1997). Needed : A ban on the significance test. *Psychological Science*, 8, 3-7.
- Hunter, J.E. & Schmidt, F.L. (1990). *Methods of meta-analysis : Correcting errors and bias in research findings*. SAGE publications, London.
- James, L.R., Demaree, R.G. & Mulaik, S.A. (1986). A note on validity generalization procedures. *Journal of Applied Psychology*, 71, 440-450.
- Killeen, P.R. (2005). An alternative to null-hypothesis significance tests. *Psychological Science*, 16, 345-353.
- Lecoutre, B. (2005). Et si vous étiez un bayésien qui s'ignore ? *Revue MODULAD*, 32, 92-105.
- Lecoutre, B. & Poitevineau, J. (2000). Aller au-delà des tests de signification traditionnels : vers de nouvelles normes de publication. *L'Année Psychologique*, 100, 683-713.
- Loftus, G.R. (1993a). A picture is worth a thousand *p* values : On the irrelevance of hypothesis testing in the microcomputer age. *Behavioral Research Methods, Instruments, & Computers*, 25, 250-256.
- Loftus, G.R. (1993b). Editorial comment. *Memory and Cognition*, 21, 1-3.
- Loftus, G.R. (1996). Psychology will be a much better science when we change the way we analyze data. *Current Directions in Psychological Science*, 5, 161-171.
- Loftus, G.R. (2002). Analysis, interpretation, and visual presentation of experimental data. In : Pashler, H. (Ed.), *Stevens' handbook of experimental psychology* (Vol. 4). John Wiley and Sons, New York, pp. 339-390.
- Loftus, G.R. & Loftus, E.F. (1988) (2<sup>ème</sup> éd.). *Essence of statistics*. Alfred A. Knopf, New York.
- Loftus, G.R. & Masson, M.E.J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, 1, 476-490.
- Lykken, D.T. (1968). Statistical significance in psychology research. *Psychological Bulletin*, 70, 151-159.
- Masson, M.E.J. & Loftus, G.R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology*, 57, 203-220.
- Maxwell, S.E. (2004). The persistence of underpowered studies in psychological research : Causes, consequences, and remedies. *Psychological Methods*, 2, 147-163.
- Meehl, P.E. (1967). Theory testing in psychology and physics : A methodological paradox. *Philosophy of Science*, 34, 1103-1115.

- Melton, A.W. (1962). Editorial. *Journal of Experimental Psychology*, 64, 553-557.
- Morrison, D.E. & Henkel, R.E. (1970). (Eds.). *The significance test controversy*. Butterworths, Londres.
- Mulaik, S.A., Raju, N.S. & Harshman, R.A. (1997). There is a time and a place for significance testing. In : Harlow, L.L., Mulaik, S.A., Steiger, J.H. (Eds.), *What if there were no significant tests ?* Lawrence Erlbaum Associates, Londres, pp. 65-115.
- Newton, I. (1999/1726). *The Principia : Mathematical principles of natural philosophy* [une nouvelle traduction par Cohen, I.B. et Whitman, A. assistés par Budens, J. de la 3<sup>ème</sup> édition de *Philosophiae Naturalis Principia Mathematica*, Londres, 1726].
- Neyman, J. (1957). « Inductive behavior » as a basic concept of philosophy of science. *Revue de l'Institut International de Statistique*, 25, 7-22.
- Neyman, J. (1977). Frequentist probability and frequentist statistics. *Synthèse*, 36, 97-131.
- Nickerson, R.S. (2000). Null hypothesis significance testing : A review of an old and continuing controversy. *Psychological Methods*, 5, 241-301.
- Pearson, E.S. (1962). Some thoughts on statistical inference. *Annals of Mathematical Statistics*, 33, 394-403.
- Popper, K.R. (1973/1959). *La logique de la découverte scientifique*. Payot, Paris [traduction de *The logic of scientific discovery*. Hutchinson, Londres, 1959].
- Popper, K.R. (1991/1979). *La connaissance objective*. Flammarion, Paris [traduction de *Objective knowledge*. Oxford University Press, Oxford, 1979].
- Rouanet, H. (1996). Bayesian methods for assessing importance of effects. *Psychological Bulletin*, 119, 149-158.
- Rouanet, H., Bernard, J.-M., Bert, M.-C., Lecoutre, B., Lecoutre M.-P. & Le Roux, B. (2000). (2<sup>ème</sup> éd.). *New ways in statistical methodology : From significance tests to Bayesian inference*. Peter Lang, Berne.
- Rouanet, H., Lépine, D. & Holender, D. (1978). Model acceptability and the use of Bayes-fiducial methods for validating models. In : Requin, J. (Ed.), *Attention and Performance VII*. Lawrence Erlbaum Associates, New York, pp. 687-701.
- Rozeboom, W.W. (1960). The fallacy of the null hypothesis significance test. *Psychological Bulletin*, 57, 416-428.
- Saville, D.J. (2003). Basic statistics and the inconsistency of multiple comparison procedures. *Canadian Journal of Psychology*, 57, 167-175.
- Schmidt, F.L. (1996). Statistical significance testing and cumulative knowledge in psychology : Implications for training of researchers. *Psychological Methods*, 1, 115-129.
- Schmidt, F.L. & Hunter, J.E. (1997). Eight common but false objections to the discontinuation of significance testing in the analysis of research data. In : Harlow, L.L., Mulaik, S.A., Steiger, J.H. (Eds.), *What if there were no significant tests ?* Lawrence Erlbaum Associates, Londres, pp. 37-64.
- Schmidt, F.L., Ocasio, B.P., Hillery, J.M. & Hunter, J.E. (1985). Further within-setting empirical tests of the situational specificity hypothesis in personnel selection. *Personnel Psychology*, 38, 509-524.
- Sedlmeier, P. & Gigerenzer, G.T. (1989). Do studies of statistical power have an effect on the power of studies. *Psychological Bulletin*, 105, 309-316.
- Shrout, P.E. (1997). Should significance tests be banned ? Introduction to a special section exploring the pros and the cons. *Psychological Science*, 8, 1-2.

- Sternberg, S. (1969). The discovery of processing stages : Extension of Donders' method. In : Koster, W.G. (Ed.), *Attention and Performance II, Acta Psychologica*, 30, 276-315.
- Swets, J.A. (1996). *Signal detection theory and ROC analysis in psychology and diagnostics : Collected papers*. Erlbaum, Mahwah, NJ.
- Swets, J.A., Dawes, M.R. & Monahan, J. (2000). Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest*, 1, 1-26.
- Swets, A., Tanner, W.P. & Birdsall, T.G. (1961). Decision processes in perception. *Psychological Review*, 68, 301-340.
- Tanner, W.P. & Swets, J.A. (1954). A decision-making theory of visual detection. *Psychological Review*, 61, 401-409.
- Wilkinson, L. and the Task Force on Statistical Inference, APA Board of Scientific Affairs, 1999. Statistical methods in psychology journals : Guidelines and explanations. *American Psychologists*, 54, 594-604.

# **Contribution du modèle du comportement planifié à la prédiction de l'entrée en formation professionnelle**

## **Contribution of The Theory of Planned Behavior to the prediction of taking up vocational training**

C. Lagabrielle\*, A.M. Vonthron\* & D. Pouchard\*\*

\*Laboratoire de psychologie EA 4139, équipe « Psychologie Sociale des Insertions »,  
Université Victor Segalen-Bordeaux2, France

\*\* *Observatoire régional des études, AFPA-CROP Aquitaine*  
Mail : [christine.lagabrielle@wanadoo.fr](mailto:christine.lagabrielle@wanadoo.fr)

### *Résumé*

La démarche de recherche présentée ici vise à approcher divers facteurs explicatifs des conduites d'engagement en formation grâce à un modèle inspiré de la théorie du comportement planifié (Ajzen, 1985). Un échantillon de 127 sujets adultes devant entamer par la suite un cycle formatif qualifiant a répondu à une série de questionnaires relatifs au modèle théorique choisi auxquels s'ajoutaient des variables sociobiographiques, des variables attitudinales et des variables contextuelles en lien avec la formation. En outre, la conduite effective (entrée/non-entrée en formation) a été prise en compte. Sur l'ensemble de l'échantillon, 60,6% a effectivement intégré le stage prévu. Les résultats majeurs montrent que le modèle du comportement planifié est pertinent dans la mesure où l'intensité de l'intention est prédictive de la conduite d'intégration. Les différentes dimensions du modèle agissent de façon différenciée selon qu'il s'agit d'expliquer l'intention d'entrée en formation ou le comportement effectif.

### *Abstract*

The research analysis various factors of the commitment in professional training by a model inspired of the theory of planned behavior (Ajzen, 1985). A sample of 127 adults having to start a professional training cycle has answered questionnaires relating to the selected model, sociobiographic, attitudinal and contextual variables. Moreover, effective behavior (entry/non-entry in the session) was taken into account. On the whole of the sample, 60,6% actually integrated the forecasted training course . The major results show that the model of the planned behavior is relevant insofar as the intensity of the intention predicts the behavior of integration. Various dimensions of the model act in differentiated ways: some of them explain the intention of entry in training course and others explain the effective behavior.

**Mots-clés** : Formation professionnelle, comportement planifié, attitude, norme subjective, perception de contrôle

**Keywords**: Professional training, planned behavior, attitude, subjective norm, perceived behavior control

## **1. Problématique**

Actuellement, la formation professionnelle est rendue indispensable par le contexte très évolutif du monde du travail (nécessité de maintien ou de déploiement des compétences, redéfinition des postes et des activités afférentes) et du marché de l'emploi (mobilité accrue dans l'emploi, réorganisation des qualifications et des métiers). L'ampleur du chômage a, de plus, opéré « une torsion des finalités de la formation » (Carré & Caspar, 1999) conduisant à des exigences accrues en termes d'efficacité des systèmes de formation. Il en résulte une injonction politico-économique de plus en plus soutenue « d'une formation tout au long de la vie » qui se confronte aux souhaits variables et singuliers de développement professionnel des personnes. La maîtrise des changements professionnels et l'adaptation récurrente des travailleurs à de tels changements se heurtent à une double contrainte : d'une part, répondre rapidement et activement, parfois de manière mécaniste, aux intérêts immédiats du marché et d'autre part, tenter de redéfinir la place du travailleur dans un environnement professionnel et organisationnel en évolution. Finalités purement économiques et finalités plus humanistes se côtoient donc pour faire face à la nécessité urgente d'articuler la formation aux changements professionnels que rencontrent les organisations et les individus (Lagabriele, 1999).

Parallèlement au développement des besoins et de l'offre de formation professionnelle continue, il est constaté un taux conséquent de non-aboutissement des parcours formatifs que ce soit durant la phase d'intégration ou au cours de leur déroulement. Ce constat amène de nombreuses interrogations tant pour les décideurs, prescripteurs de formation que pour les organismes prestataires de programmes. Les questions soulevées concernent notamment les attentes des différents acteurs à l'égard de la formation professionnelle, l'intentionnalité d'inscription dans un parcours formatif, les défections ou l'engagement effectif dans un stage ainsi que les réalités de l'insertion sociale et professionnelle produite. Plus généralement, force est de souligner la complexité des problématiques en jeu relevant de la prédiction des conduites formatives ou de la compréhension des processus d'engagement mis en œuvre.

Pour cela, plusieurs cadres conceptuels peuvent être considérés. L'objectif est ici d'envisager, dans un cadre socio-cognitif, l'apport de la théorie du comportement planifié (Ajzen, 1985, 1991 ; Ajzen & Madden, 1986) pour expliquer un moment-clé de l'engagement, sujet à de nombreux désistements, à savoir l'entrée effective en formation.

## **2. Cadre théorique**

### **2.1. Les différents modèles explicatifs de l'engagement en formation**

A la suite de Piolat (1980), la démarche d'engagement peut se définir comme un ensemble plus ou moins coordonné de conduites qu'un individu développe pour

rechercher les conditions favorables à l'exercice de certaines activités qu'il valorise. Selon Dupont et al (1980), les intérêts désignent les tendances orientées vers différents domaines d'activités ou d'expériences alors que l'engagement est l'action même qui concrétise ces tendances. L'engagement en formation serait donc de l'ordre de l'action dès lors qu'il se traduit par des actes concrets, observables, parmi lesquels l'acte de candidature à une formation particulière et la participation effective et contractualisée à un parcours formatif débouchant sur des issues professionnelles.

L'analyse des phénomènes d'engagement en formation a donné lieu à de multiples travaux s'inscrivant dans différents champs théoriques. Certaines recherches s'orientent vers l'identification de facteurs antécédents (caractéristiques individuelles, professionnelles, institutionnelles, sociales) facilitant ou inhibant l'engagement effectué (Demol, 1995). D'autres adoptent un point de vue dynamique prenant en compte l'existence de besoins, de tensions que l'engagement en formation aurait pour fonction de réduire ou de motivations auxquelles l'engagement en formation répondrait (Nuttin, 1980 ; Boutinet, 1998 ; Battistelli et al, 1993, 1994, 1998, 2000). D'autres encore étudient les spécificités relatives aux conduites d'engagement en tant qu'elles participent à la socialisation (Dubar, 1995) ou à la personnalisation (Malrieu, 1989) ou selon qu'elles relèvent des processus d'anticipation et de projets (Michel, 1989). Enfin, plusieurs perspectives considèrent le rôle actif de construction de la réalité vécue par l'individu dans ce qu'il entreprend (Baubion-Broye et al, 1998 ; Dupuy, 1993, 1998) et la part d'intentionnalité sollicitée dans ses conduites d'engagement en formation (Perron, Lecomte & Desharnais, 1990). Il semble toutefois que cette question de l'intentionnalité à l'égard de la formation et des conduites qui en découlent soit moins explorée que les autres orientations citées. Dans cette perspective, la théorie du comportement planifié offre des travaux conséquents permettant d'appréhender les mécanismes socio-cognitifs expliquant de nombreuses conduites volitives mises en œuvre par les individus.

## 2.2. La théorie du comportement planifié et l'engagement en formation

La théorie de l'action raisonnée (Fishbein & Ajzen, 1975) est basée sur l'idée que l'individu se conduit habituellement de façon stratégique, qu'il tient compte des informations auxquelles il a accès et qu'il évalue l'implication de ses actions (bénéfices escomptés rapportés aux inconvénients perçus). Les auteurs insistent sur le fait que les individus agissent généralement en concordance avec leur intention. Dans ce cadre conceptuel, l'intention est définie comme la fermeté avec laquelle la personne souhaite s'engager dans un comportement visé. Elle s'appuierait d'une part sur l'attitude du sujet à l'égard du comportement à adopter et d'autre part sur la norme subjective. La dimension attitudinale renvoie aux avantages et inconvénients que perçoit la personne dans la mise en œuvre d'une conduite, ainsi plus les avantages prévalent (notamment les conséquences estimées positives à agir) plus l'attitude est favorable. La dimension de norme subjective, quant à elle, a trait à la façon dont le sujet considère que l'adoption du comportement est valorisée par son entourage social et à l'importance qu'il accorde à cette approbation.

Ajzen en 1985 enrichit ce premier modèle et propose la théorie du comportement planifié qui ajoute la perception de contrôle comme troisième facteur prédictif de l'intention comportementale. Cette dernière relève des croyances de l'individu que l'adoption du comportement envisagé est plus ou moins sous son contrôle. Ainsi, plus le comportement dépend strictement de la personne, plus celle-ci estime disposer des moyens nécessaires pour adopter le comportement voulu. Ajzen & Madden (1986) insistent sur le fait que la perception de contrôle peut prédire, non seulement l'intentionnalité mais aussi, directement, le comportement adopté. Pour que la relation entre la perception de contrôle et le comportement soit directe et non médiatisée par l'intention, il faut que, d'une part, la perception de contrôle et le contrôle comportemental réel soient très proches et que, d'autre part, le comportement étudié ne soit pas exclusivement sous le contrôle de l'individu. Si ces conditions ne sont pas réunies, la théorie du comportement planifié est alors réduite à celle de l'action raisonnée. Il semble que la situation de formation professionnelle puisse être analysée de façon plus pertinente grâce au modèle du comportement planifié dans le sens où, effectivement, l'individu, bien que décidant de ses conduites d'engagement en formation dans un cadre volitif, ne maîtrise pas pour autant l'ensemble des conditions (favorables ou défavorables) de la situation. Jaccard (1975), dans un autre domaine, a souligné les différents facteurs favorisant l'intensité de la relation « intention-comportement » : la rapidité de l'enchaînement entre les deux, la spécificité de l'intentionnalité, la faible complexité, la familiarité, ou encore le degré d'auto-détermination de la conduite envisagée. De la même manière, vis-à-vis des situations formatives, on peut supposer que la relation entre une intention d'entrée en formation et le comportement effectif d'intégration soit soumise à des éléments contextuels tels que le délai séparant un accord donné pour un stage et son démarrage ou encore le fait d'avoir ou non des expériences antérieures de formation continue.

L'étude présentée s'attache donc à mettre à l'épreuve le modèle théorique du comportement planifié afin d'identifier ses apports dans la compréhension de l'entrée effective ou non en formation.

### **3. Méthodologie**

#### **3.1. Population et procédure**

Les sujets pris en compte lors de l'étude de terrain sont issus de plusieurs régions françaises et doivent entamer dans l'année un cycle formatif qualifiant<sup>1</sup>. Les secteurs d'activités, métiers et niveaux pris en compte sont divers. Les personnes ont été interrogées par questionnaires auto-administrés, un mois après leur participation à un dispositif d'orientation formative se concluant par leur admission à un stage précis. 127 sujets ont répondu à une série de questionnaires. Nous avons également pu répertorier les conduites d'intégration ou de non intégration dans les formations ciblées que les

---

<sup>1</sup> L'étude a été réalisée en partenariat avec l'AFPA (Association pour la Formation Professionnelle des Adultes) et plus particulièrement avec la cellule Observatoire et Etudes du CROP Aquitaine.

sujets ont mises en œuvre par la suite. Pour cet échantillon, les formations ont démarré dans un délai de moins de 6 mois pour 41,7%, de 6 à 12 mois pour 48,8% et de plus d'une année pour 9,4%.

L'échantillon est constitué de 80,2% d'hommes, l'âge moyen est de 31,6 ans et les niveaux de formation initiale sont variés (du niveau Vbis correspondant à un niveau de fin d'études au collège jusqu'au niveau III et plus c'est-à-dire au-delà de deux années d'études supérieures). 40,9% est en emploi et parmi les personnes au chômage, on compte 34,7% de chômeurs de moins de 6 mois, 40% de 6 à 12 mois et 25,3% de plus d'un an. La durée moyenne de l'activité professionnelle est de 10,3 années et 45,6% des sujets est sans expérience de la formation professionnelle continue. Pour ceux qui ont expérimenté la formation professionnelle, le dernier stage effectué a été mené à son terme par 83,6% des personnes. Parmi les personnes interrogées avant le démarrage d'une formation, 60,6% a effectivement intégré par la suite le stage prévu.

### 3.2. Mesures des variables

Outre les données socio-biographiques et contextuelles liées à la formation qui ont été recueillies, les sujets ont été interrogés à l'aide d'échelles élaborées spécifiquement pour cette étude et portant sur :

- *leur attitude générale à l'égard de la formation professionnelle* grâce à 9 propositions constituées d'antonymes (intéressant/inintéressant ; valorisant/dévalorisant ; utile/inutile etc.), mesurées par une échelle bipolaire en 7 points. L'alpha de Cronbach est de .99 ;
- *la valeur accordée à la formation retenue* par une échelle de valorisation en 10 points ;
- *la probabilité d'obtention d'un emploi grâce à la formation prévue* (en pourcentages sur 10 degrés).

De même, les différentes échelles relatives au modèle du comportement planifié proprement dit ont également été élaborées spécifiquement pour cette étude selon la procédure de mesure préconisée par les auteurs :

- *une échelle d'attitude à l'égard du comportement d'entrée en formation* composée de 9 items (en 7 points) concernant les conséquences perçues de l'entrée en formation (exemple d'items : restreindre vos activités, enrichir vos relations, augmenter votre niveau de stress et de fatigue etc.) et 9 items associés (en 7 points également) concernant la valeur accordée à ces conséquences. La première sous-échelle obtient un alpha de Cronbach de .97 et la seconde de .97 également ;
- *une échelle de mesure de la norme subjective* comprenant 4 propositions (en 7 points) relatives à la croyance d'approbation par l'entourage personnel et professionnel ainsi que 4 autres items (en 7 points également) portant sur la motivation à agir dans le sens exprimé par cet entourage. Pour la première échelle, le sujet doit évaluer la probabilité qu'ont différents membres de son entourage (conjoint, membres de la famille etc.) d'approuver son entrée dans la formation

- prévue et indiquer dans la seconde, son degré d'accord pour agir dans le sens de son entourage (conjoint, membres de la famille etc.). La première sous-échelle obtient un alpha de Cronbach de .97 et la seconde de .98 ;
- une échelle de mesure de perception de contrôle à l'égard du comportement d'entrée en formation constituée de deux sous-échelles : d'une part, 15 propositions (en 7 points) pouvant être perçues comme défavorables à l'intégration en formation (exemple d'items : horaires de la formation, obligations familiales, revenus pendant le stage, trouver un emploi entre-temps) et d'autre part, 15 items (en 7 points également) correspondant à l'importance accordée à chacune. La consigne de la première échelle est la suivante : « jusqu'à quel point croyez-vous probable ou improbable que chacun des facteurs suivants puissent nuire à votre entrée dans la formation prévue ? ». Pour la seconde, la consigne est la suivante : « par rapport à votre entrée dans la formation prévue, quelle importance accordez-vous aux dimensions suivantes ? ». L'alpha de Cronbach est de .94 pour la première sous-échelle et de .83 pour la deuxième ;
  - une estimation de la fermeté d'intention d'intégrer la formation en 7 degrés<sup>2</sup>.

On constate que la plupart des échelles construites ont des alphas de Cronbach particulièrement élevés ce qui signifie qu'il existe une redondance trop forte entre les items de chaque questionnaire. A des fins d'utilisation ultérieure de cet outillage, il conviendrait donc de les réduire.

#### 4. Résultats

##### 4. 1. Résultats descriptifs

Variables	Moy	E.T	2	3	4	5	6	7
1. Attitude générale vis-à-vis de la formation	5,85/7	0,67	.46 **	.36 **	.32 **	.23 *	-.38 **	.25 **
2. Valeur accordée à la formation retenue	8,54/10	1,51		.40 **	.28 **	.25 **	-.23 *	.23 *
3. Proba obtention emploi par formation	7,72/10	2,06			.28 **	.19 *	-.36 **	.32 **
4. Attitude compt d'entrée en formation	24,86/49	5,88				.29 **	-.25 *	.24 **
5. Norme subjective	34,03/49	10,71					-.13	.25 **
6. Perception de contrôle	12,84/49	7,10						-.14
7. Intention d'entrer en formation	6,61/7	0,86						

\* < .05 ; \*\* < .01

Tableau 1

Moyennes, écarts-types et corrélations pour les variables étudiées

<sup>2</sup> Les questionnaires complets sont disponibles sur demande adressée au premier auteur.

Les résultats présentés dans le tableau 1 indiquent une attitude globalement favorable envers la formation professionnelle et très favorable à l'égard du stage programmé qui doit être intégré par la suite. De plus, la formation prévue est considérée majoritairement comme un bon moyen d'accéder ensuite à un emploi. Ceci étant, il faut remarquer que, malgré cette perception positive de la formation et de ses liens avec l'emploi, l'attitude spécifique à l'égard du comportement d'entrée reste plus mitigée ( $m = 24,86/49$ ). Les personnes envisagent un certain nombre de conséquences néfastes associées au fait d'intégrer une formation alors même qu'elles considèrent pouvoir contrôler elles-mêmes fortement leur entrée en formation (dans la mesure où le résultat moyen de  $12,84/49$  correspond à une perception de peu d'obstacles ou d'obstacles de peu d'importance pouvant entraver l'intégration). La norme perçue provenant de l'entourage est plutôt en faveur d'une approbation de l'entrée en formation. Enfin, au moment où on les interroge, elles déclarent une ferme intention d'intégrer le stage prévu.

Concernant les liens entre les différentes variables prises en compte, il convient de souligner que les dimensions attitudinales étudiées ne sont pas indépendantes les unes des autres mais ne sont pas dans tous les cas fortement corrélées. Parmi les liens les plus élevés, nous notons que les sujets sont d'autant plus favorables à la formation professionnelle en général qu'ils accordent de la valeur à celle qu'ils doivent intégrer ( $r = .46$ ) et qu'ils estiment que la formation est un bon moyen d'obtenir un emploi ( $r = .36$ ). Dans la même logique, ils associent également la valorisation de la formation qui a été retenue à la facilitation de l'accès à l'emploi ( $r = .40$ ). On peut supposer que ces liens attitudinaux sont l'expression d'une certaine consonance cognitive. A l'examen des liens négatifs les plus élevés, il apparaît que moins la personne perçoit d'obstacles à son entrée en formation et plus elle valorise la formation en général ( $r = -.38$ ) et l'obtention d'un emploi par ce moyen ( $r = -.36$ ).

#### 4.2. Résultats concernant la prédiction de l'intention et du comportement d'entrée en formation

Dimensions	Intention d'entrée en formation	$\beta$
1. Attitude générale formation		.04
2. Valeur accordée à la formation retenue		.02
3. Proba obtention emploi par formation		.22* ( B = .085 ; ES = .039)
4. Attitude compt d'entrée en formation		.08
5. Norme subjective		.28** ( B=.022 ; ES = .008)
6. Perception de contrôle		-.03
<b><i>R<sup>2</sup> ajusté : 0,14</i></b>		

p<.05 ; \*\*p<.01

Tableau 2

Analyses de régression multiple prédisant l'intention d'entrée en formation

Note : les coefficients  $\beta$  associés aux dimensions dans le tableau sont des bêtas standardisés

Les résultats présentés dans le tableau 2 indiquent que seules la norme subjective et dans une moindre mesure la probabilité d'obtention d'un emploi grâce à la formation participent à la prédiction de l'intention d'entrée en formation. Pour autant cette prédiction reste très modérée. Si on se réfère au modèle du comportement planifié, l'intention d'entrée en formation devrait être expliquée à la fois par la norme subjective (ce qui est le cas) mais aussi par l'attitude à l'égard du comportement d'entrée et par la perception de contrôle (ce qui n'est pas vérifié ici).

Enfin, l'impact potentiel de caractéristiques personnelles (sexe, âge, niveau de qualification), professionnelles (type d'itinéraire professionnel, durée de l'expérience, durée d'inactivité pour ceux qui sont au chômage, expérience antérieure de la formation, interruption d'un stage précédent) ou formatives (niveau de qualification du stage, délai avant l'entrée en formation, domaine professionnel du métier visé) a été examiné. Le seul effet significatif qui apparaît sur l'intention d'entrée concerne la durée d'inactivité des chômeurs : lorsque la durée de chômage augmente, la fermeté de l'intention d'entrée diminue ( $F = 5,499$  ;  $p = .001$ ).

	<i>DDL</i>	<i>Somme des carrés</i>	<i>Carré moyen</i>	<i>F</i>	<i>p</i>
1. Attitude générale formation	1	1,96	1,96	4,57	<b>.03</b>
2. Valeur accordée à la formation retenue	1	8,35	8,35	3,57	<b>.06</b>
3. Proba obtention emploi par formation	1	25,33	25,33	5,97	<b>.02</b>
4. Attitude compt d'entrée en formation	1	14,38	14,38	.397	.53
5. Norme subjective	1	656,06	656,06	5,80	<b>.02</b>
6. Perception de contrôle	1	458,10	458,10	9,77	<b>.002</b>
7. Intention d'entrer en formation	1	8,38	8,38	12,88	<b>.000</b>

Tableau 3

ANOVA relatives aux dimensions influençant la conduite d'intégration de la formation (entrée effective vs non entrée dans le stage prévu)

Les résultats du tableau 3 révèlent que seule l'attitude à l'égard de l'entrée en formation ne prédit pas significativement la conduite ultérieure d'intégration (vs non intégration) dans la formation prévue. Il convient de souligner que la fermeté de l'intention (bien qu'elle soit ici mesurée largement à distance du comportement constaté) s'avère déterminante dans l'action d'entrée dans le stage programmé, conformément à la modélisation théorique de l'action planifiée. Comme suggéré dans la littérature, la perception de contrôle a une influence directe très significative sur le

comportement d'entrée en formation.

Parmi les caractéristiques personnelles prises en compte et leur impact sur la conduite d'entrée effective en formation, seul le type d'itinéraire professionnel influence le comportement ultérieur (Khi deux de Pearson = 15,266 ;  $p = .009$ ) : les sujets qui ont alterné emploi et chômage sont ceux qui intègrent le plus souvent le stage prévu (à 81,1%) suivis des sujets qui ont toujours travaillé (à 53,8%). Les personnes déclarant vivre de « petits boulots » sont celles qui entrent le moins fréquemment dans la formation programmée (à 33,3%).

## **5. Discussion**

L'étude menée s'attachait à introduire une perspective socio-cognitive susceptible de prédire l'entrée effective en formation professionnelle de sujets retenus pour intégrer un stage programmé ultérieurement. Elle a permis d'aborder la question des désistements à l'entrée des stages puisque, dans l'échantillon pris en compte, il y a une personne sur quatre environ qui n'a pas intégré la formation dans laquelle elle était inscrite. Si le modèle d'Ajzen (1985) a déjà fait ses preuves pour comprendre les décrochages et le maintien dans différents domaines (santé, éducation, insertion professionnelle), aucune étude à notre connaissance ne l'a utilisé pour comprendre l'intégration ou non dans un parcours formatif. Les résultats obtenus montrent l'intérêt de l'approche théorique du comportement planifié pour la compréhension des mécanismes à l'œuvre lors de ce moment-clé de l'engagement en formation. Comme attendu dans ce cadre conceptuel, l'intention s'avère prédictive des comportements futurs observés bien que l'enchaînement entre intention et comportement soit de l'ordre de plusieurs mois. Ceci conforte l'idée selon laquelle l'inscription des adultes en cours de vie active dans des dispositifs de formation qualifiante relève de conduites finalisées.

Cependant, le modèle proposé par Ajzen (1985) ne fonctionne pas dans sa globalité. Il est nécessaire de rappeler que la dimension « perception de contrôle » a été ajoutée à un premier modèle, celui de l'action raisonnée, dans la mesure où la plupart des actions entreprises ne sont pas strictement volitives. Dans le cadre de la formation continue, cette dimension a d'autant plus de sens que l'intégration dans un stage comporte de multiples paramètres (économiques, liés aux dispositifs formatifs etc.) qui ne dépendent pas de l'individu. A ce propos, nos résultats indiquent que la perception de contrôle à l'égard du comportement d'entrée agit sur la conduite elle-même (comme suggéré par Ajzen & Madden, 1986) mais pas sur l'intention. Ainsi, la force de l'intention est indépendante du contrôle perçu ( $\beta = -.03$ ) comme des caractéristiques formatives considérées (niveau de qualification du stage, délai d'attente avant le début de stage, domaine professionnel du métier visé par la formation). Madden, Ellen & Ajzen (1992) avaient également montré que la perception de contrôle prédit le comportement de façon significative lorsqu'il s'agit d'une conduite faiblement maîtrisée par l'individu. Ici aussi, la perception de contrôle a un impact direct sur le comportement d'entrée dans une formation sans médiatisation par l'intention.

Par ailleurs, nous observons que la norme subjective influence l'intention d'entrée en formation mais prédit également la conduite ultérieure. Dans le cas présent, cette dimension paraît à même d'enrichir la compréhension des mécanismes à l'œuvre dans l'action finale. Ceci est d'autant plus intéressant que la question de l'approbation sociale est peu prise en compte dans les études portant sur l'engagement et le maintien en formation (Vonthron et al, à paraître). Or, les résultats présentés ici attestent d'un besoin chez l'individu d'avoir des projets et des conduites qui soient perçues favorablement par l'entourage pour pouvoir se réaliser. Finalement, la décision d'entrée en formation dépasse une approche individuelle. Il convient d'insister sur le fait que « se former » en cours de vie active, sollicite, au-delà de l'individu lui-même, le positionnement de son réseau social.

Enfin, concernant les attitudes mesurées, il faut rappeler que notre étude prend en compte d'une part l'attitude à l'égard d'un objet (attitude à l'égard de la formation) et d'autre part l'attitude à l'égard d'un comportement spécifique (ici le comportement d'entrée en stage). On constate que l'attitude à l'égard du comportement d'entrée en formation n'explique ni l'intention ni la conduite effective d'intégration alors que cette dimension joue un rôle majeur dans la prédiction d'autres types de conduites liées notamment au domaine de la santé (Schifter & Ajzen, 1985 ; Gatch & Kendzierski, 1990). Ce résultat questionne d'autant plus que si l'attitude spécifique à l'égard de l'entrée en formation n'a pas de poids sur le comportement, une attitude beaucoup plus générale vis-à-vis de la formation a, elle, une incidence sur l'action entreprise plusieurs mois après. Ce qui *in fine* va à l'encontre des travaux qui, depuis les années 1960, font remarquer la faible relation observée entre l'attitude à l'égard d'un objet et les comportements plus spécifiques adoptés vis-à-vis de celui-ci (Wicker, 1969). Dans ce contexte formatif particulier, il semble bien que ce soit pourtant l'attribution de valeur à « l'objet formation » et non les conséquences perçues de l'entrée en formation plus ou moins favorables (mesurées par l'attitude à l'égard du comportement d'entrée) qui détermine la conduite effective. Finalement, la formation est sans doute suffisamment valorisée socialement pour que cette valorisation porte l'individu au-delà des difficultés potentielles perçues.

Concernant les variables autres que celles du modèle étudié, nous pouvons retenir que la probabilité d'obtention d'un emploi grâce à la formation intervient aussi bien dans l'explication de l'intention que de la conduite ultérieure. Ce résultat va dans le sens des travaux qui montrent combien le lien formation-travail-trajectoire professionnelle est prioritaire pour les personnes, dans la mesure où elles sont de plus en plus soumises à une injonction de formation pour maintenir ou développer leur employabilité (Barbier, 1980 ; Carré & Caspar, 1999 ; Fond-Harmand, 1995). A ce propos, notre échantillon est d'ailleurs composé pour moitié de personnes au chômage. Il semble souhaitable de porter une attention toute particulière au fait que, dans ces populations au chômage, ce sont les publics les plus en précarité d'emploi (c'est-à-dire ceux qui ont vécu jusque-là de « petits boulots »), qui, finalement, moins que les autres, intègrent les parcours formatifs.

En conclusion, ces premiers résultats incitent à explorer davantage les apports de modèles socio-cognitifs quant au domaine de l'engagement en formation professionnelle. Plusieurs questions mériteraient d'être approfondies par la suite et parmi elles : le poids de la valorisation de la formation par rapport à des attitudes plus spécifiques à son égard ou encore le poids de la norme sociale sur l'engagement formatif. La façon dont se structure (ou se déstructure ?) le rapport à la formation des populations particulièrement en difficulté d'insertion socioprofessionnelle nécessiterait également d'être davantage étudiée.

### **Références**

- Ajzen, I. (1985). From intentions to actions : a theory of planned behavior. In J. Kuhl et J. Beckmann (Eds.), *Action-control : from cognition to behavior*, (pp.11-39). Heidelberg.
- Ajzen, I. & Madden, T.J. (1986). Prediction of goal-directed behavior: Attitudes, intentions, and perceived behavioral control, *Journal of Experimental Social Psychology*, 22, 453-474.
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision processes*, 50, 179-211.
- Barbier, J.M., (1980). Situation de travail, poursuite d'un projet de formation et procès de transformation personnelle, *Revue Française de sociologie*, XXI, 409-438.
- Battistelli, A., Majer, V. & Odoardi, C. (1993). La motivazione nella formazione degli adulti, *Risorsa Uomo*, 1, 233-261.
- Battistelli, A., Majer, V. & Odoardi, C. (1994). Il significato della formazione : sua influenza sulle motivazioni degli adulti verso le attività formative, *Risorsa Uomo*, 2, 43-58.
- Battistelli, A., Majer, V. & Odoardi, C. (1998). Une échelle pour évaluer la motivation dans la formation des adultes, *Psychologie du travail et des organisations*, 4, 226-234.
- Battistelli, A. & Odoardi, C. (2000). *Les composantes de la motivation à apprendre chez les participants à la formation pour l'emploi*. In B. Gangloff (ed.). *Satisfactions et souffrances au travail*, Paris: L'Harmattan.
- Baubion-Broye A. (1998). *Événements de vie, transitions et construction de la personne*. Saint Agne, Eres.
- Boutinet, J.P. (1998). L'engagement des adultes en formation et ses formes de légitimation, *Education Permanente*, 136, 91-100.
- Carré, P. & Caspar, P. (1999). *Traité des sciences et techniques de la formation*. Paris : Dunod.
- Demol, J.N. (1995). L'entrée en formation. *Education permanente*, 125 (4), 27-37.
- Dubar, C. (1995). *La socialisation. Construction des identités sociales et professionnelles*. Paris : Armand Colin.
- Dupont, J.B., Gendre, F., Berthoud, S. & Descombes, J.P. (1980). *La psychologie des intérêts*. Paris : PUF Le psychologue.
- Dupuy, R., (1993). Valorisations, attentes et investissements en formation. *L'orientation scolaire et professionnelle*, 22, n° 4, 371-392.

- Dupuy, R., (1998). Transitions et transformation des identités professionnelles, le cas des adultes en situation de formation continue, in A. Baubion-Broye, *Événements de vie, transitions et construction de la personne*, Eres.
- Fishbein, M. & Ajzen, I. (1975). *Belief, attitude, intention and behavior : an introduction to theory and research*. Don Mills: Addison-Wesley.
- Fond-Harmand, L. (1995). Approche biographique et retour aux études. *Education Permanente*, 125, p 7-25.
- Gatch, C.L. & Kendzierski, D. (1990). Predicting exercise intentions : the theory of planned behavior, *Research Quarterly for exercise and sport*, 61, 100-102.
- Jaccard, J. (1975). A theoretical analysis of selected factors important to health education strategies, *Health Education Monographs*, 3, 152-167.
- Lagabriele, C. (1999). *Transitions identitaires chez des salariés en formation professionnelle qualifiante*. Thèse de Doctorat en Psychologie, Université Victor Ségalen-Bordeaux2.
- Madden, T.J., Ellen, P.S. & Ajzen, I. (1992). A comparison of the theory of planned behavior and the theory of reasoned action, *Personality and Social Psychology Bulletin*, 18, 3-9.
- Malrieu, P. (1989). *Dynamiques sociales et changements personnels*. Paris : Editions du CNRS.
- Michel, S. (1989). *Peut-on gérer les motivations ?* Paris: Presses Universitaires de France.
- Nuttin, J. (1980). *Théorie de la motivation humaine*. Paris : Presses Universitaires de France.
- Perron, J., Lecomte, C. & Desharnais, R. (1990), *Transitions de carrière à l'âge adulte, Etude d'un modèle conceptuel*, projet soumis et accepté au C.R.R.H, non publié.
- Piolat, M., (1980). *L'engagement en formation, conception de soi et modèles de rôle chez l'adulte face au changement*, Thèse de Doctorat de 3ème cycle, Université de Provence.
- Schifter D.B. & Ajzen, I. (1985). Intention, perceived control and weight loss : an application of the theory of planned behavior. *Journal of Personality and social psychology*, 48, 843-851.
- Vonthron, A.M., Lagabriele, C. & Pouchard, D. (2006). Le maintien en formation professionnelle qualifiante : effets de déterminants motivationnels, cognitifs et sociaux, *Orientation Scolaire et Professionnelle* (à paraître).
- Wicker, A.W. (1969). Attitudes versus actions : The relationship of verbal and overt behavioral responses to attitude objects. *Journal of Social Issues*, 25, 41-48.

# **La représentation des causes de l'état de santé par les infirmiers et les points de vue envisagés des patients et des aidants naturels : une approche socionormative.**

## **Nurses' representation of the determinants of the state of health and the points of view attributed to patients and caregivers : a socionormative approach.**

Daniel Gilibert et Ingrid Banovic

[Daniel.Gilibert@u-bourgogne.fr](mailto:Daniel.Gilibert@u-bourgogne.fr)

Daniel GILIBERT, Université de Bourgogne

Pôle AAFE, SPMS, B.P. 26513, 21065 Dijon Cédex France

### *Résumé*

Cette étude porte sur la représentation des déterminants de l'état de santé lors de prises en charges thérapeutiques. Nous décrivons ici la représentation d'infirmiers hospitaliers expérimentés (n=44) et les différences de points de vue qu'ils envisagent avec leurs usagers. En se basant sur la finalité du soin et l'asymétrie existante dans la relation soignant-soigné, nous faisons d'une part l'hypothèse que ces infirmiers exhiberont une représentation de l'état de santé centrée sur le soin à apporter : les déterminants de l'état de santé devant être perçus comme ayant un impact général sur la vie des patients tout en étant remédiables, c'est à dire susceptibles d'évoluer et contrôlables par l'aide apportée. D'autre part, on peut s'attendre à ce qu'ils attribuent au patient une position et une attitude passive : les déterminants de son état de santé devant être perçus comme incontrôlables par lui. Les résultats et l'analyse des représentations qu'ils rapportent comme étant celles du patient et de sa famille semblent justifier la perception et le rôle soignant : les représentations supposées de ces derniers seraient convergentes mais moins centrées sur le problème de santé et minimiseraient l'implication du patient, notamment en cas d'évolution défavorable. Ils sont discutés au regard du Health Belief Model et de la norme sociale d'internalité, ainsi que des pratiques professionnelles en milieu hospitalier.

### *Abstract*

This study examines the representation of the determining factors of patients' states of health in therapeutic care situations. We describe the average representation of experienced hospital nurses (n=44) and the differences which they consider to exist between their own and their patients' points of view. Taking as our starting point the purpose of the care provided and the asymmetry of the carer/patient relationship, we hypothesize, on the one hand, that nurses will have a representation of their patients' state of health which is based on the care to be provided, with the factors determining the state of health being necessarily perceived as having a general impact on patients' lives while still being remediable, i.e. capable of changing and being controlled by the provided care. On the other hand, we expect them to attribute a passive position and attitude to their patients, with the factors determining their state of health being necessarily perceived as being out of their own control. The results and the analysis of the representations reported as being those of the patients and their families seem to justify this perception and role of carer: the supposed representations of the latter would appear to be convergent but less focused on health problems and would tend to minimize the patient's involvement, in particular when the state of health declines. The results are discussed with reference to the Health Belief Model and the social norm of internality as well as the professional practices observed in hospital environments.

**Mots-clefs** : Dimensions Causales – Norme Sociale – Maladie – Soignants – Health Belief Model

**Keywords** : Causal dimensions – Social norms - Illness – Caregivers -Health Belief Model

## **1. Introduction**

Les recherches en psychologie de la santé abordent la question de la prise en charge thérapeutique comme conditionnée par un système où les points de vue de différents acteurs interviennent et interagissent. Elles traitent parfois des interactions entre soignants et patients et, depuis peu, de l'implication des aidants naturels. Cette perspective dépasse la seule prise en compte de la représentation médicale de la maladie et de son traitement. Elle vise à prendre en compte les représentations plus «naïves» du patient et de ses aidants naturels, intervenant également dans le déroulement de la prise en charge thérapeutique. Nécessairement en interaction dans la mise en place des comportements de soins, les représentations des différents acteurs semblent co-exister, s'influencer mutuellement et peuvent parfois interférer les unes avec les autres. On peut alors s'interroger sur la façon dont ces acteurs s'accordent en général sur une représentation commune de la maladie, de son origine et des actions de soins à mettre en place. Si traiter une telle question demeure complexe et ambitieux, on peut tout au moins s'intéresser, dans un premier temps, à la façon dont les soignants se représentent les déterminants de l'évolution thérapeutique en relation avec les points de vue des autres acteurs concernés (patients, familles). La prise en charge de la maladie étant largement conditionnée par le savoir médical des soignants et leur présence dans le choix des soins, on peut alors se demander: 1/ quels sont, du point de vue des soignants, les déterminants de l'état de santé du patient, 2/ s'ils s'attendent à ce que leur représentation soit partagée par le patient et les aidants naturels (la famille), 2/ si leur analyse causale de la maladie et les points de vue attendus de ces usagers seront les mêmes selon que l'état de santé s'améliore ou non. Cette contribution concerne donc la représentation que des soignants ont de la maladie et de la prise en charge thérapeutique en fonction des trois points de vue susceptibles d'intervenir dans le système de soins: leur point de vue de soignant, celui du soigné et celui de son aidant naturel.

En partant du constat de l'asymétrie existante dans la relation du soignant avec le patient, dont la santé est supposée dépendante du soin prodigué, des hypothèses peuvent être posées : 1/ en ce qui concerne la perception que les soignants ont des déterminants de l'état de santé : ces déterminants devraient être perçus comme décisifs, susceptibles d'évoluer, dépendants de l'aide reçue plutôt que du contrôle du patient, 2/ en ce qui concerne le point de vue du malade (et de sa famille) qui serait supposé moins centré sur le problème : il sous-estimerait le fait que ces déterminants sont décisifs et remédiables, et plus passif : il surestimerait l'intervention d'éléments extérieurs, incontrôlables par lui-même (en accord avec les travaux sur les points de vue attendus dans les relations asymétriques) 3/ enfin, lorsque l'état de santé s'est détérioré, les soignants devraient avoir une perception moins optimiste des déterminants de l'état de santé et évoquer des points de vue plus fortement divergents. Afin de servir cet objectif de recherche, nous proposons dans un premier temps de passer rapidement en revue les travaux soulignant l'importance d'une approche des relations soignant-patient-aidants naturels. Dans un second temps, nous rappellerons quelques travaux montrant que les

explications, ainsi que les points de vue envisagés pour autrui, sont dépendants des rôles professionnels et de l'asymétrie existant dans la relation.

### ***L'intervention des aidants naturels, des patients et des soignants dans la prise en charge thérapeutique***

En posant la question de la prise en charge thérapeutique comme résultant d'un système composé des interactions entre patient, soignant et aidant naturel, les travaux en psychologie de la santé tentent d'éclairer le rôle spécifique de ces acteurs dans l'observance et l'éducation thérapeutique(1). Ils amènent à s'intéresser aux interactions possibles entre soignant et patient (Eider, Ayala, Harris, 1999; Fischer, Tarquinio, 2006; Goldring, Taylor, Kemeny, Anton, 2002; Hessen-Klemens & Lapinska, 1984). Ainsi, concernant l'observance thérapeutique, trois facteurs semblent importants: (a) le soutien que le patient peut trouver dans son entourage familial et personnel, (b) les attentes que le patient peut avoir quant à ses possibilités de guérison, l'incitant à mettre en œuvre des comportements adéquats, et enfin (c) la qualité de la relation thérapeutique avec le soignant.

Concernant tout d'abord *la participation des aidants naturels*, un élément illustratif de l'importance d'une approche «interactive», est que le soutien, ressenti par le malade en provenance de son entourage, semble intervenir sur les conséquences de la maladie et du traitement (Dew, Goyocolea, Harris, Lee, Zomak, Dunbar-Jacob & Rotondi, 2004; Pibernik-Okanovic, Prasek, Poljicanin-Filipovic, Pavlic-Renar, Metelko, 2004; van Wijngaarden, Schene, & Koeter, 2004).

*Le système de croyances et de représentations des patients* serait déterminant dans la compréhension et la mise en œuvre de comportements d'observance thérapeutique. A ce sujet, le *Health Relief Model* (Rosenstock, 1974; Rosenstock, Strecher & Becker, 1988) souligne l'importance des évaluations faites par les patients concernant l'utilité des prescriptions et du respect du traitement médical. Ce modèle insiste plus précisément sur l'importance du sentiment d'efficacité personnelle du patient, c'est-à-dire son sentiment de pouvoir contrôler, par des efforts personnels, l'évolution de sa maladie (Goldring, Taylor, Kemeny & Anton 2002; Gao, Nau, Rosenbluth, Scott & Woodward, 2000; Bradley, 1995; Home & Weinman, 1999). La façon dont les patients conçoivent l'origine de ce qui leur arrive serait une variable importante. Ainsi le «*Locus of control*» différencierait les individus en deux catégories extrêmes: ceux qui envisagent les événements comme essentiellement dépendant de facteurs externes (chance, destin, etc.) et ceux qui les envisagent comme essentiellement dépendant de facteurs internes (liés à des caractéristiques et comportements personnels). La vulnérabilité face à des agents stressants serait d'ailleurs plus importante pour les individus fatalistes, percevant ce qui leur arrive comme provenant de facteurs externes. Pour Rosenstock et al. (1988), les échecs dans l'observance des prescriptions ou dans l'adoption de comportements de soins seraient également liés au fait que les individus évaluent mal l'impact réel de la maladie sur leur vie. Ces échecs pourraient enfin être liés au fait que les efforts à faire par le patient pour éviter de telles conséquences sont supérieurs au bénéfice direct du soin pour le patient (Home, James, Petrie, Weinman &

Vincent, 2000; Home, & Weinman 1999; Bradley, 1994; Norman, 1995; Shiloh, Rashuk-Rosenthal & Benyamini, 2002).

Enfin, *la qualité de la relation thérapeutique avec le soignant* est le facteur prépondérant dans l'observance. Le patient peut rechercher ou percevoir des éléments dans la relation avec le médecin qui vont le rassurer ou le soutenir psychologiquement. La qualité de cette relation clinique serait un élément fondamental dans la mise en place de comportements de soins. Hessen-Klemens et Lapinska (1984) ainsi que Squier et Jackson (1990) ont montré que plus le médecin encourage et soutient son patient plus celui-ci est observant dans la mise en œuvre des comportements de soins. La chaleur humaine et l'empathie seraient donc essentielles pour l'observance. De même pour Rosenstock et al. (1988), la qualité de la relation renforcerait la perception qu'a le patient d'un contrôle personnel sur l'évolution de la maladie et ses attentes positives quant aux bénéfices d'un comportement observant. L'interaction ayant lieu lors de la relation thérapeutique pourrait contribuer à un soutien psychoaffectif pour le patient, soutien apporté par le médecin lorsque celui-ci retire également une satisfaction personnelle de sa relation avec le patient (Squier, Jackson, 1990; Hessen-Klemens, Lapinska, 1984; Gao, Nau, Rosenbluth, Scott, Woodward, 2000).

### ***Asymétrie de la relation de soin, rationalisation des rôles et points de vue attendus***

A notre connaissance, aucun travail empirique ne s'est jusqu'alors intéressé à la façon dont les soignants eux-mêmes percevaient les interactions au sein de la triade soignant- patient-aidant naturel, notamment en ce qui concerne la prise en charge et l'évolution de la maladie.

Des travaux tentent pourtant d'analyser de façon générale la relation soigné-soignant et de proposer des modèles interprétatifs de cette relation (Le Sommer-Péré, M., 2000). Un premier modèle est un modèle de relation pouvant être qualifié de «paternaliste» où le soignant, représenté dans une posture bienveillante, connaît et sait ce qui est bien pour le patient, du fait de son savoir et de son expertise médicale. Le patient y est représenté dans une posture passive de dépendance et d'attente vis-à-vis du soignant et du soin. Un second modèle plus récent se veut contractuel et soutient l'idée du libre choix du patient, de son adhésion préalable au projet de soin et de son droit à l'information. Un dernier modèle, de relation pédagogique, rappelle que le malade est en situation de crise et que son état de souffrance altère temporairement la rationalité de sa décision et ses capacités de coopération. Le soignant aura la responsabilité de la conception du programme de soin et des conditions de sa mise en œuvre, mais devra les expliciter au patient tout au long du soin. En fait, quelque soit le modèle envisagé, l'objectif du soin est toujours la guérison et celle-ci passe nécessairement par l'aide du soignant et sa maîtrise des soins médicaux, représentant une asymétrie difficilement contournable dans la relation entre soignant et patient. Il est vraisemblable que les soignants élaborent leur représentation de l'évolution thérapeutique et de ses déterminants sur la base de cette représentation de la relation soigné-soignant dont l'objectif est la guérison par le biais de leur maîtrise des actions de soins à mettre en place.

D'une façon générale, il est fréquent que les pratiques professionnelles déterminent largement les représentations que les sujets ont d'un objet (Abric, Flament, Guimelli, Mardellat, Morin & Singery, 2001). Dans de nombreux secteurs d'activités, il est observé que les travailleurs sociaux rendent cohérente avec leur pratique l'analyse causale des situations auxquelles ils sont confrontés. Ainsi, dans le domaine de l'accidentologie routière, les différents protagonistes tendent à minimiser leur propre rôle dans la survenue des accidents et se centrent sur des déterminants justifiant leur position ou leur rôle professionnel (Kouabenan, 2002). Dans les activités de maintien de l'ordre, les policiers tendent à attribuer les comportements déviants à des déterminants internes et stables aux personnes qu'ils doivent interpeller (par exemple leur personnalité ou prédispositions déviantes). Les surveillants de prison, quant à eux, attribuent ces mêmes conduites déviantes à des facteurs essentiellement stables dans le temps, justifiant ainsi l'incarcération (facteurs qui sont alors aussi bien d'ordre interne qu'externe aux personnes les produisant ; Abdellaoui, Kouabenan & Gilibert, 1998).

Dans le domaine de l'éducation enfin, les enseignants tendent à considérer que les performances scolaires sont majoritairement liées aux efforts des élèves (caractéristiques internes et contrôlables par ces derniers) et ce d'autant plus qu'ils les ont notés de façon polarisée ou qu'ils leur est demandé de rendre compte du niveau scolaire de l'élève (Matteucci & Gosling, 2004 ; Gosling, 1994 ; 1995). Les explications des enseignants tendent aussi à être différentes en fonction de la valeur scolaire de l'élève : alors que les résultats d'un élève ayant un bon niveau sont attribués aux efforts qu'il aurait produits, ainsi qu'à la qualité de l'encadrement pédagogique, les résultats d'un mauvais élève sont essentiellement attribués à un manque d'efforts au niveau scolaire (à des éléments plus fréquemment internes, instables et spécifiques au domaine scolaire). Les explications évoquées permettent ainsi de justifier les pratiques évaluatives et la qualité de l'encadrement pédagogique.

Outre le fait que les causes envisagées varient en fonction du niveau scolaire de l'élève, il semble enfin que les enseignants s'attendent à des explications bien différentes de la part des bons et des mauvais élèves ; explications pouvant justifier la moindre performance de ces derniers. D'après les enseignants, le point de vue des mauvais élèves à la différence de celui des bons élèves, serait que les performances (bonnes ou mauvaises) sont imputables à des facteurs extérieurs ou incontrôlables par eux. Dans ces travaux basés sur un « paradigme d'identification », les sujets, les parents ou les enseignants, doivent envisager et rapporter le point de vue d'un bon et d'un mauvais élève (Dubois & Beauvois, 1996; Beauvois, Gilibert, Pansu & Abdellaoui, 1998; Dubois, 2003). Dans ce domaine, comme dans d'autres domaines de la perception sociale, il apparaît que les personnes les moins valorisées socialement, ici les mauvais élèves, se voient attribuer les croyances externes qui sont peu valorisées socialement (que la valeur soit définie sur la base du statut professionnel, académique, scolaire ou de façon plus affective, Gilibert & Cambon, 2003 ; Pansu, Tarquinio & Gilibert, 2005). Rappelons qu'en matière d'attributions causales, les croyances les moins valorisées sont relatives à une détermination externe plutôt qu'interne de nos résultats, ou encore à des facteurs incontrôlables plutôt que contrôlables par soi-même.

Etant donnée le caractère asymétrique de la relation de soin, on peut alors s'attendre à ce que les déterminants de l'état de santé soient décrits par les soignants de façon cohérente avec leur objectif de soin et leur maîtrise du soin et de la prise en charge du malade. On peut ainsi s'attendre (hypothèses unilatérales) à ce que :

1/ les soignants se représentent les déterminants de l'état de santé comme : a) ayant un impact général sur la vie du patient, mais b) demeurant toutefois remédiables (le soin ayant pour finalité d'améliorer l'état de santé du patient), c'est-à-dire instables et susceptibles d'évoluer et c) contrôlables par l'aide d'autrui, mais a) n'étant pas contrôlables par le patient ni particulièrement lié à lui-même (l'état de santé étant supposé dépendre du soin prodigué).

2/ les soignants se représentent le point de vue du malade comme moins centré sur le problème (sous-estimant le fait que ces déterminants sont décisifs et remédiables) et plus passif (surestimant l'intervention d'éléments extérieurs et incontrôlables par lui-même) ; justifiant ainsi leur présence sur les soins. Non seulement le patient surestimerait des croyances qui apparaissent comme néfastes dans le cadre du *Health-Belief Model* (extermalité, incontrôlabilité) mais il sous-estimerait aussi des croyances utiles (généralité de l'impact des déterminants de l'état de santé, instabilité, contrôlabilité par autrui).

3/ lorsque l'état de santé s'est détérioré, les soignants devraient avoir une perception moins optimiste des déterminants de l'état de santé et surtout rapporter des points de vue plus fortement divergents de la part du malade et de sa famille : ceux-ci seraient perçus comme ayant des croyances externes et d'absence de contrôle, particulièrement prononcées dans les cas où la prise en charge s'est avérée difficile et l'évolution de la maladie la moins favorable (en accord avec les travaux sur la norme sociale d'internalité).

## 2. Méthode

### Population:

La population était constituée de quarante-quatre soignants, dix hommes et trente-quatre femmes, infirmier(e)s, diplômé(e)s d'état depuis plusieurs années (infirmiers anesthésistes, puéricultrices, infirmiers, infirmiers psychiatriques, infirmiers de bloc opératoire, manipulateurs-radio), en formation pour devenir soit cadres de santé, puéricultrices ou anesthésistes. Leur âge moyen est de 31,5 ans, dont plusieurs années de service (en moyenne 8 ans).

### Procédure:

La passation du questionnaire, lors de sessions de formation, était présentée comme relevant d'une recherche dont l'objectif était de mieux comprendre les déterminants des prises en charge thérapeutique et des comportements de soin.

### ***Phase 1 : évocation de deux cas cliniques contrastés en termes d'évolution***

Après quelques informations signalétiques (spécialité, âge, expérience), il leur était demandé de se remémorer deux cas cliniques de leur domaine de spécialité qu'ils avaient suivis et qu'ils avaient bien connus, atteints de la même pathologie avec le même niveau initial de gravité. Pour l'un des patients, la prise en charge thérapeutique devait s'être plutôt bien passée avec une évolution favorable de la maladie et, pour l'autre, la prise en charge devait avoir été plus difficile avec une évolution défavorable de sa maladie. Ils devaient tout d'abord rapporter le nom de la pathologie et le degré de gravité initial des symptômes. Les pathologies rapportées couvrent un vaste ensemble de pathologies somatiques extrêmement variées (peu de pathologies psychiatriques ont été évoquées). Ils devaient également, et pour chacun des deux cas séparément, mentionner la durée de suivi. Le suivi de ces cas cliniques a duré en moyenne 202 jours, qu'il se soit bien ou mal passé, avec une très forte variabilité dans chaque cas du fait de la variété des pathologies (bien passé:  $m=202.7$  e-t.=539.6 vs. Mal passé:  $m=201.8$  e-t.=413.9).

### ***Phase 2 : point de vue des soignants sur les déterminants de l'état de santé et points de vue rapportés pour les patients et leurs familles***

Ils étaient avertis, qu'afin d'être exhaustifs dans la recherche des déterminants possibles de l'état de santé, qu'ils devraient non seulement donner leur propre point de vue sur les déterminants de l'état de santé des patients, mais qu'ils devraient également rapporter ce qu'était ou leur a semblé être le point de vue de ces patients et de leurs familles. Après avoir évoqué de façon qualitative les déterminants possibles de l'évolution de l'un de ces deux cas, il leur était demandé à partir d'échelles dimensionnelles de caractériser les éléments qu'ils estimaient être déterminants de cette évolution de la maladie du patient (évolution soit favorable ou soit défavorable, selon l'ordre de réponse des questionnaires aléatoirement distribués aux infirmiers). Ces échelles dimensionnelles sont basées sur les distinctions habituellement faites entre facteurs de causalité (Weiner, 1985; Fontayne, Martin-Krumm, Buton & Heuzé, 2003) : 1/ le lieu d'origine de la cause interne ou externe au patient, 2/ la stabilité temporelle des causes (permanentes ou aléatoires), 3/ la généralité de leur impact dans la vie du patient, et 4/ la contrôlabilité de causes: contrôlables ou non par le patient. Une cinquième dimension relative à la contrôlabilité de ces causes par l'aide d'autrui a paru pertinente dans la mesure où l'étude porte sur une relation d'aide et de soin apporté par autrui.

Après avoir donné leur opinion sur les déterminants de l'évolution de la maladie, les soignants étaient invités à indiquer, sur un format identique, ce qui leur avait semblé être l'opinion du patient sur les déterminants de sa maladie, puis ce qui leur avait semblé être l'opinion de sa famille(2). La procédure de réponse était répétée alors pour le second cas clinique contrasté en terme d'évolution de la maladie (soit favorable soit défavorable).

### ***Phase 3 : questions de vérification concernant les cas cliniques et l'évolution thérapeutique***

Dans un dernier temps les sujets complétaient un questionnaire de description clinique pour chacun de ces deux cas. Il permettait d'évaluer la gravité de la maladie, son évolution au cours de la prise en charge et l'impact du traitement thérapeutique (*Clinical Global Impression Scale*, voir partie « matériel ») ainsi que l'impact de la maladie sur le fonctionnement psychosocial du patient (*Global Assessment of Functioning Scale*).

### **Matériel**

#### ***Représentation des déterminants de l'état de santé et de leurs caractéristiques causales.***

Dans un premier temps, il était demandé aux soignants, en fonction de leur pratique clinique et leur expérience de la maladie, de mentionner par écrit des éléments qui leur avait parus déterminants de l'apparition des symptômes, de la consultation et de l'évolution favorable (vs. défavorable) de sa maladie pour chaque cas clinique(3).

Concernant plus particulièrement l'évolution (favorable vs. défavorable) du patient, il leur était alors demandé si ce qui avait déterminé cette évolution a) *Était liée à des éléments extérieurs à lui-même.* b) *Était liée à lui, à sa personnalité, à sa façon d'être,* c) *Était stable et peu susceptible d'évoluer dans le temps* d) *Était aléatoire et susceptible d'évoluer dans le temps,* e) *Était contrôlable, maîtrisable par lui,* f) *N'était pas contrôlable, n'est pas maîtrisable par lui,* g) *N'était susceptible d'affecter son comportement que dans peu de situations,* h) *Était susceptible d'affecter son comportement dans d'autres situations de la vie,* i) *Était contrôlable, maîtrisable grâce à l'aide apportée par autrui,* j) *N'était pas contrôlable, maîtrisable grâce à l'aide apportée par autrui.* En face de chacune de ces affirmations, les sujets indiquaient leur degré d'accord (sur une échelle d'opinion de Likert en 7 points allant de « pas du tout » à « tout à fait » par la suite codée de -3 à +3). Ces réponses permettaient d'évaluer leur perception de l'origine de la maladie sur les cinq dimensions attributives retenues pour cette étude: le lieu d'origine de l'évolution thérapeutique, interne/externe au patient ; la contrôlabilité/incontrôlabilité des éléments déterminants par le patient ; la stabilité/l'instabilité dans le temps de ces déterminants ; l'impact général/spécifique de ces déterminants sur le comportement du patient et enfin leur contrôlabilité/incontrôlabilité par l'aide d'autrui. Cette méthode de mesure est similaire à la version révisée de mesure des attributions causales permettant de d'investiguer les dimensions causales les plus pertinentes dans les théories de l'attribution (*Causal Dimension Scale*, adaptée en français par Fontayne, Martin-Krumm, Buton & Heuzé, 2003). Conformément à celle-ci, nous traiterons de façon statistique la différence entre les deux items antagonistes pour chacune des cinq dimensions (e.g. : l'attribution au patient moins l'attribution aux éléments externes).

### **Questions de vérification concernant les cas cliniques et l'évolution thérapeutique**

Afin de vérifier que chaque cas correspondait bien à une évolution favorable ou défavorable de la maladie, il leur était demandé de produire une évaluation clinique de chaque patient après sa prise en charge par le biais de la *Clinical Global Impression Scale* en plusieurs échelles fréquemment utilisées dans le domaine médical (Guy, 1976)(4). Les valeurs de la première échelle (item 1) ont été comparées à une variante présentée en tout début de questionnaire pour situer la gravité des symptômes (la pathologie et les symptômes devaient être identiques entre les deux cas ; le niveau de gravité initiale des symptômes ( $m=5.43$ ) au moment de la prise en charge correspondant en moyenne à des cas se situant entre «manifestement» et «gravement malade»). Il apparaît que les symptômes sont jugés tout aussi sévères en fin de traitement dans le cas d'évolution défavorable ( $m=5.16$ ,  $F(1,43)=1.2$  ns) mais qu'il s'atténuent significativement en cas d'évolution favorable ( $m=4.59$ ,  $F(1,43)=8.69$ ;  $p<.01$ ). La seconde échelle, évaluant l'amélioration de l'état de santé, atteste d'une évolution positive dans le cas favorable ( $m=2.04$  ; «Fortement amélioré») et d'une stagnation dans le cas défavorable ( $m=3.82$ ; «pas de changement», ces évolutions étant bien différentes;  $F(1,43)=32.0$ ;  $p<.0001$ ). La différence entre les deux patients est confirmée par la troisième et dernière mesure de l'effet du traitement (CGI item3 : 3.98 vs. 8.24,  $F(1,41)=27.1$ ;  $p<.0001$ ) avec une «disparition importante des symptômes» rapportée pour le cas favorable alors qu'un effet seulement «minime» est rapporté pour le cas défavorable (1.50 vs. 2.52,  $F(1,41)=29.6$   $p<.0001$ ), sans qu'il n'y ait plus des effets secondaires rapportés pour le cas le plus défavorable (1.98 vs. 2.14,  $F(1,41)=0.8$  ns).

En toute fin de questionnaire, les soignants devaient évaluer l'impact de la maladie sur le fonctionnement du patient dans sa vie sociale, à l'issue de la prise en charge (*Global Assessment of Functioning Scale* (5)). Cette mesure atteste également d'une différence d'évaluation en terme de fonctionnement psychosocial, entre les deux cas ( $F(1,42)=60.65$ ;  $p<.0001$ ) avec une évaluation plutôt positive du cas favorable  $m=71.14\%$ : «Quelques symptômes légers ou une certaine difficulté dans le fonctionnement social, professionnel ou scolaire mais fonctionne assez bien de façon générale et entretient plusieurs relations interpersonnelles positives» et une évaluation clairement négative pour le cas défavorable  $m=40.84\%$  : «Existence d'une certaine altération de la communication - discours par moment inadapté - ou déficience majeure dans plusieurs domaines (le travail, l'école, les relations familiales, le jugement)».

### **3. Résultats**

Pour chacune des 5 dimensions attributives, la différence entre les deux items antagonistes a été traitée par analyse de variance selon un plan entièrement intrasujet, en fonction de «l'évolution» de l'état de santé (cas Favorable vs. cas Défavorable) et en fonction «des points de vue envisagés» sur l'origine de l'état de santé (le point de vue du Soignant lui-même vs. le point de vue qu'il rapporte comme étant celui du Patient concerné vs. le point de vue qu'il rapporte comme étant celui de sa Famille). Les moyennes sont présentées dans le tableau 1. Une première lecture descriptive des moyennes (une différence de 1 point avec zéro,

est significative à un seuil  $p < .05$ ) indique que les éléments contribuant à l'état de santé sont globalement perçus comme ayant un impact général (affectant le comportement du patient dans d'autres situations de la vie), instables (susceptibles d'évoluer dans le temps) et clairement contrôlables par l'aide apportée par autrui, mais plutôt incontrôlables par le patient (pour cette dernière dimension  $p < .10$ ).

**Tableau 1 :** Moyennes et écart-types des valeurs des échelles dimensionnelles pour caractériser l'évolution thérapeutique (favorable ou non) en fonction des points de vue envisagés par les soignants (eux-mêmes, rapporté pour le patient, rapporté pour la famille).

	Point de vue du soignant lui-même		Point de vue rapporté pour le patient		Point de vue rapporté pour la famille	
	en cas d'évolution Favorable.	en cas d'évolution Défav.	en cas d'évolution Favorable .	en cas d'évolution Défav .	en cas d'évolution Favorable.	en cas d'évolution Défav .
<i>Ce qui a déterminé l'évolution de la maladie était...</i>						
-Général « <i>Etait susceptible d'affecter son comportement dans d'autres situations de la vie</i> » vs « <i>que dans peu de situations</i> »	2,25 <sup>ab</sup> (2,90)	2,77 <sup>a</sup> (2,64)	<b>1,11<sup>c</sup></b> (2,82)	<b>1,77<sup>bc</sup></b> (2,50)	1,86 <sup>bc</sup> (2,59)	<b>1,75<sup>bc</sup></b> (2,73)
-Stable et peu susceptible d'évoluer « <i>Etait stable et peu susceptible d'évoluer dans le temps</i> » vs. « <i>aléatoire et susceptible d'évoluer dans le temps</i> »	-1,91 <sup>bc</sup> (2,78)	-2,16 <sup>c</sup> (2,81)	<b>-0,84<sup>a</sup></b> (3,28)	<b>-1,29<sup>a</sup></b> (3,12)	-1,72 <sup>bc</sup> (2,69)	-1,59 <sup>abc</sup> (2,97)
-Contrôlable par l'aide d'autrui « <i>Etait contrôlable, maîtrisable grâce à l'aide apportée par autrui</i> » vs. « <i>n'était pas contrôlable...</i> »	2,64 <sup>ab</sup> (2,82)	1,91 <sup>b</sup> (2,60)	2,35 <sup>ab</sup> (2,74)	<b>1,15<sup>c</sup></b> (3,11)	2,68 <sup>a</sup> (2,64)	2,05 <sup>ab</sup> (2,95)
-Interne au patient « <i>Etait lié à lui, à sa personnalité, à sa façon d'être</i> » vs. « <i>lié à des éléments extérieurs au patient</i> »	-0,43 <sup>a</sup> (2,93)	-0,16 <sup>a</sup> (2,92)	-0,57 <sup>a</sup> (2,61)	<b>-1,45<sup>b</sup></b> (3,10)	-0,39 <sup>a</sup> (2,58)	<b>-1,34<sup>b</sup></b> (2,66)
-Contrôlable par le patient « <i>Etait contrôlable, maîtrisable par lui</i> » vs. « <i>n'était pas contrôlable, n'était pas maîtrisable par lui</i> »	-0,80 <sup>bc</sup> (3,78)	-0,66 <sup>abc</sup> (3,43)	<b>0,25<sup>a</sup></b> (2,85)	-1,57 <sup>c</sup> (3,22)	-0,03 <sup>ab</sup> (3,35)	-1,32 <sup>c</sup> (3,34)

**Note :** Dans ce tableau, les moyennes ne portant en indice aucune lettre commune, sont statistiquement différentes à un seuil  $p < .05$  (avec un test  $F$  de Fischer, ou à un seuil  $p$ -bilatéral  $< .10$ , avec un test post-hoc LSD). Sont rapportées en gras les moyennes divergeant du point de vue des soignants pour le cas d'évolution concerné.

Les résultats seront présentés en trois temps : 1/ les effets de « l'évolution » de l'état de santé sur la perception de ses déterminants, 2/ les effets des « points de vue envisagés » révélant des différences entre le point de vue du soignant et le point de vue rapporté par ce soignant pour le patient et sa famille et 3/ l'interaction de ces deux variables permettant de voir si les différences de points de vue sont dépendantes de l'évolution de l'état de santé.

### ***Perception des déterminants de l'état de santé en fonction de son évolution (favorable vs. défavorable)***

Le fait que l'état de santé ait évolué favorablement ou non amène des différences de description globale des déterminants de cette évolution. Les déterminants tendent à être décrits comme étant moins contrôlables par l'aide d'autrui lorsque l'évolution a été défavorable plutôt que favorable (1,70 vs. 2.56 vs. 0.03,  $F(1,43)=4.8$   $p<.02$ ), à être plus incontrôlables par le patient (-1.18 vs. -0.19,  $F(1,43)=8.1$   $p<.01$ ) et d'autant moins liés au patient (-0.98 vs. -0.46,  $F(1,43)=2.5$   $p<.06$ ). Néanmoins si l'on considère isolément les réponses des soignants, indépendamment des points de vue qu'ils rapportent pour le patient et sa famille, seule la différence de contrôlabilité par l'aide d'autrui tend en partie à être significative dans les réponses qu'ils endossent (2.64 vs. 1.91,  $F(1,43)=2.7$   $p<.06$ ).

### ***Différences entre les différents points de vue envisagés par le soignant (son propre point de vue, le point de vue du patient et celui de sa famille)***

Si on s'attache à étudier les différences globales de point de vue indépendamment de la qualité de l'évolution thérapeutique (favorable vs défavorable), les soignants rapportent des points de vue différents du leur pour le patient et sa famille en ce qui concerne quatre dimensions causales sur cinq ( $F(2,86)=6.2$   $p<.01$  pour la généralité de l'impact des éléments déterminants l'évolution de la maladie,  $F(2,86)=3.83$   $p<.02$  pour leur stabilité,  $F(2,86)=2.45$   $p<.05$  pour la contrôlabilité par l'aide d'autrui et  $F(2,86)=2.24$   $p<.06$  pour le lieu d'origine de ces éléments). La lecture du tableau 1 indique que les soignants s'attendent à des points de vue du patient et de sa famille qui sont moins polarisés que le leur mais allant néanmoins dans le même sens.

Si l'on fait abstraction de la qualité de l'évolution, peu de différences sont envisagées avec le point de vue de la famille qui occuperait une position intermédiaire entre le point de vue du patient et celui du soignant. Selon les soignants, la famille sous-estimerait pourtant l'impact général des déterminants de santé comparativement à eux-mêmes (1.81 vs. 2.51,  $LSD$   $p$ -bilatéral $<.05$ ).

En comparaison avec leur point de vue, les soignants présente le patient comme moins centré sur la résolution de son problème de santé : il sous-estimerait l'impact des déterminants de santé (1.44 vs. 2.51,  $LSD$   $p$ -bilatéral $<.01$ ), le fait qu'ils soient remédiables, c'est à dire susceptibles d'évoluer (-1.07 vs. -2.04,  $LSD$   $p$ -bilatéral $<.01$ ) et contrôlables par l'aide d'autrui (1.75 vs. 2.27,  $LSD$   $p$ -bilatéral $<.10$ ).

Son point de vue serait également plus passif, dans la mesure où il sous-estimerait le fait que ces déterminants sont, au moins en partie, liés à lui-même (-1.01 vs. -0,30, *LSD p-bilatéral* <.05).

Aucune interaction significative n'apparaît pourtant pour la contrôlabilité de cette évolution par le patient lui-même. Cette variable est pourtant d'intérêt, tant au regard des travaux sur le *Health Relief Model* que du point de vue de la valeur socialement attribuée au contrôle personnel dans le cadre de la norme sociale d'internalité. On peut néanmoins s'attendre à une différence de point de vue dépendante de l'évolution (favorable ou non) de la maladie.

***Interaction des deux variables : différences entre les points de vue envisagés en fonction de l'évolution (favorable vs. défavorable)***

Des divergences apparaissent entre les trois points de vue envisagés en fonction de l'évolution favorable ou défavorable de la maladie, pour ce qui est de la contrôlabilité perçue de la maladie par le patient (interaction  $F(2,86)=2.9$   $p<.05$ ) ainsi que du lieu d'origine de la maladie (interaction  $F(2,86)=2.5$   $p<.05$ ). Les moyennes sont représentées dans les figures 1 et 2. Aucune autre divergence significative entre les points de vue, liée à l'évolution de la maladie, n'est observée pour les autres dimensions.

Pour la contrôlabilité de la maladie, le point de vue du soignant lui-même ne semble pas changer que l'évolution ait été ou non favorable : dans les deux cas, le soignant est indécis et tendrait plutôt à penser que ce qui a déterminé l'évolution était plutôt incontrôlable par ces deux patients (6) (voir figure 1). Le soignant présente cette incontrôlabilité comme marquée dans l'esprit du patient et de la famille lorsque l'évolution a été défavorable plutôt que favorable (pour les patients : -1.57 vs. 0.25, *LSD p-bilatéral* <.01 et pour les familles : -1.32 vs. -0.02, *LSD p-bilatéral* <.05). En cas d'évolution favorable, le patient est même présenté comme percevant sa maladie comme plus contrôlable que ne le faisait le soignant (0,25 vs. -0.80, *LSD p-bilatéral* <.08).

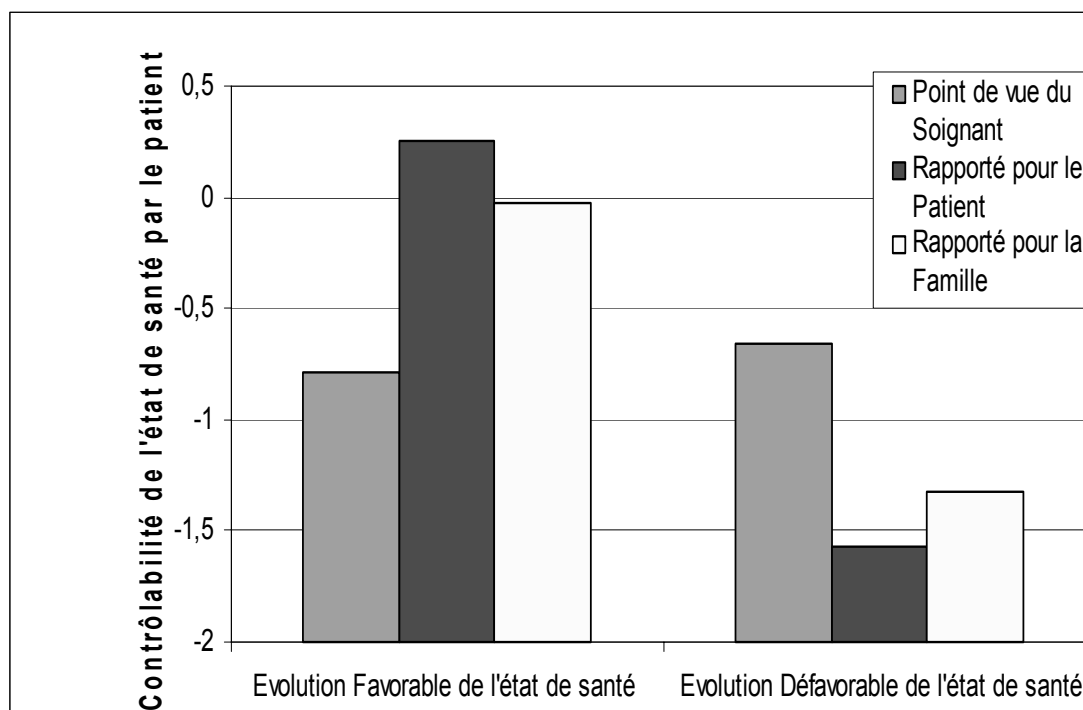


Figure 1

Moyennes pour la contrôlabilité des déterminants de l'état de santé en fonction des points de vue envisagés par les soignants (eux-mêmes, selon la famille, selon le Patient) en fonction de l'évolution (favorable vs. défavorable).

En ce qui concerne l'origine de la maladie (figure 2), le soignant estime qu'elle est tout autant liée au patient qu'à la situation extérieure, et ce quelque soit l'évolution. En cas d'évolution favorable, aucune différence n'apparaît entre les trois points de vue envisagés. En revanche, en cas d'évolution défavorable, des points de vue différents du leur sont rapportés pour le patient et sa famille. Dans ce cas, les soignants présentent le point de vue du patient comme contestant le fait que cette évolution défavorable soit liée à la lui (-1.45 vs.-0.16, *LSD p-bilatéral*<.01). Selon eux, la famille, insisterait sur l'origine situationnelle de cette évolution défavorable, de façon plus prononcée qu'ils ne le feraient eux-mêmes (-1.34 vs.-0.16, *LSD p-bilatéral*<.01).

Patient et famille sont ainsi présentés comme plus externes dans leurs croyances lorsque l'évolution est défavorable que lorsqu'elle est favorable (pour les patients : -1.45 vs.-0.57, *LSD p-bilatéral*<.05 et pour les familles : -1.34 vs. -0.39, *LSD p-bilatéral*<.05).

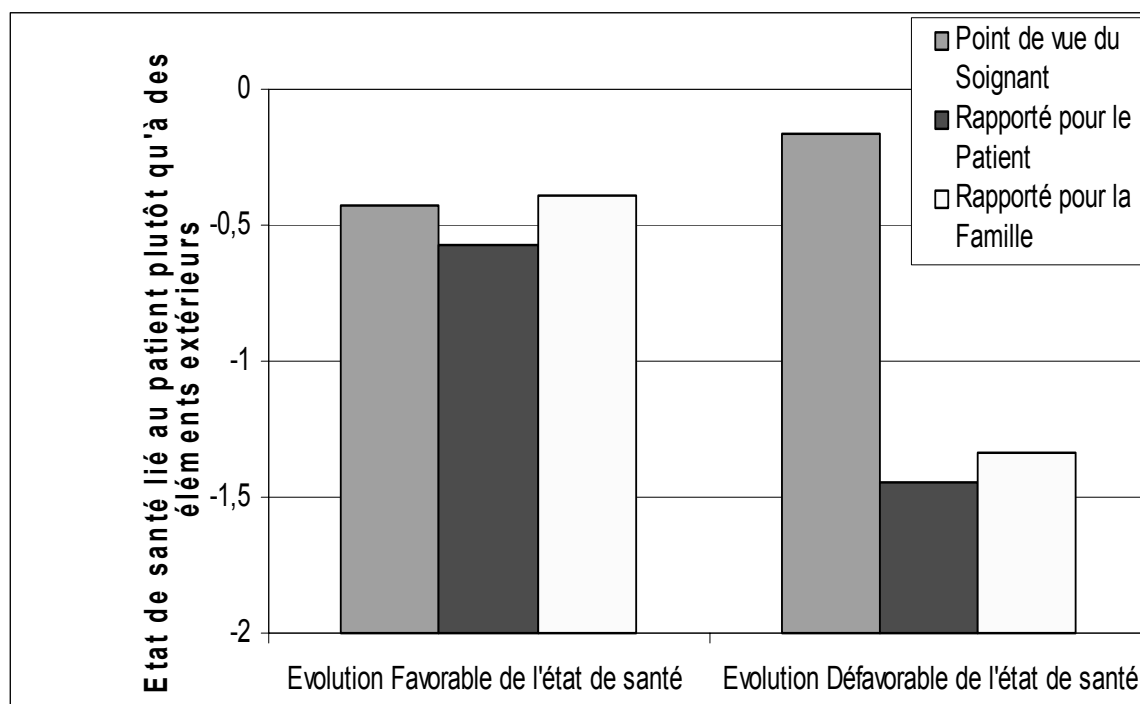


Figure 2

Moyennes pour l'origine des déterminants de l'état de santé en fonction des points de vue envisagés par les soignants (eux-mêmes, rapporté pour le patient, rapporté pour la famille) en fonction de l'évolution (favorable vs. défavorable).

#### 4. Conclusion

Les résultats observés sont globalement conformes aux attentes que l'on pouvait avoir en se basant sur la nature asymétrique de la relation de soins. Trois éléments essentiels se dégagent et seront discutés successivement. Le premier élément concerne la perception que les soignants ont des déterminants de l'état de santé lors de la prise en charge. Le second concerne les points de vue qu'ils rapportent comme étant ceux des patients indépendamment de l'évolution de leur état de santé. Le dernier a trait aux points de vue rapportés comme étant ceux des patients et de ses aidants naturels lorsque la maladie suit une évolution défavorable.

#### *Une représentation de l'évolution de l'état de santé ancrée dans la nécessité du soin et l'état d'incapacité du patient*

Concernant tout d'abord la représentation des déterminants de l'état de santé, elle semble pour le moins cohérente avec la pratique de soin hospitalier. Rappelons que la relation soigné-soignant reste asymétrique du fait des connaissances médicales, de la maîtrise du protocole de soin par les soignants et des places occupées par les différents acteurs (Le Sommer-Péré, 2000). Une représentation paternaliste et savante du soin continuerait d'influencer les attitudes des soignants et des malades en dépit de la nécessité de faire participer le patient à la décision de soin. Rappelons qu'elle dresse le

portrait d'un malade passif, nécessitant un soin, en attente face à un soignant responsable et bienveillant, avec le risque d'une dépendance au discours et aux décisions médicales.

Conformément à cette représentation paternaliste, il a été observé que, du point de vue des soignants, le patient n'est pas en mesure de contrôler seul ce qui a déterminé sa maladie et son évolution, alors que ces déterminants auraient pourtant un impact général sur sa vie. Cette représentation est cohérente avec une vision du patient diminué, en attente de soins et n'étant pas en mesure de remédier par lui-même à son état de santé. Cela suggère aussi que, pour les soignants, le patient reste au second plan et ne ferait essentiellement que subir le soin et l'évolution de sa maladie. Par ailleurs, les soignants envisagent les éléments faisant évoluer l'état de santé comme remédiables : instables dans le temps et contrôlables par l'aide d'autrui. De façon cohérente avec leur objectif de soin, les soignants envisagent donc l'état de santé comme maîtrisable. Etant les initiateurs des soins, l'évolution de la maladie est de leur point de vue contrôlable par l'aide apportée et non par le patient, dont le rôle est relégué au second plan (le patient étant « pris en charge » par eux).

Cette représentation du suivi thérapeutique, par les soignants, justifie leur fonction de soin (la maladie ayant un impact général, étant susceptible d'évoluer mais contrôlable par l'aide apportée) ainsi que la position de chacun dans une relation définie comme asymétrique (la maladie étant contrôlable par l'aide apportée et non par le patient lui-même). Cette représentation de l'évolution de santé est néanmoins à relativiser au regard du contexte hospitalier. En effet, rappelons que les cas cliniques évoqués sont des cas le plus souvent graves et ayant nécessité une ou plusieurs hospitalisations. Ceci pourrait expliquer le peu de contrôle supposé du patient sur l'évolution de sa maladie et son impact, supposé important. Ces paramètres (pratique soignante avec ou sans prises en charges successives, représentation la contribution possible du patient au soin et de la relation avec lui, prise en charge ambulatoire) mériteraient d'être explorés dans une recherche ultérieure.

Enfin, de façon consistante avec les travaux sur les attributions causales de la maladie (Weiner, 1985), les soignants attribuent d'autant moins au patient un contrôle sur l'évolution la maladie, favorable ou non, que celle-ci relève d'une pathologie grave ou lorsqu'ils la considèrent comme invalidante pour le fonctionnement psychosocial du patient. Le fait, sans doute le plus troublant dans la représentation des déterminants de la prise en charge thérapeutique, est qu'elle ne semble quasiment pas varier en fonction de l'évolution de cet état de santé (amélioration effective ou stagnation). Les déterminants de l'état de santé tendent tout au plus à être perçus comme légèrement moins contrôlables par l'aide apportée en cas d'évolution défavorable. Aucune différence n'est observée concernant leur lieu d'origine, leur contrôlabilité par le patient, leur stabilité ou leur impact sur la vie du patient. Les recherches en matière d'analyse causale montrent pourtant de façon récurrente des effets liés à la désirabilité du fait expliqué, les explications divergeant facilement selon que les sujets aient à analyser une réussite ou un échec. On aurait ainsi pu s'attendre à ce que les déterminants soient perçus par les

soignants eux-mêmes comme peu susceptibles d'évoluer, plus lourds de conséquences et moins contrôlables par le patient en cas d'évolution défavorable.

Il est vraisemblable que les infirmiers, de par leur fonction dans le cadre hospitalier, n'ont pas d'analyse causale à produire suite à l'évolution. Leur analyse causale de la maladie semble définie a priori sur la base du savoir médical, en fonction de la pathologie et de la gravité des symptômes (7). Il est aussi possible qu'invoquer des connaissances médicales et la technicité du rôle professionnel puisse permettre aux soignants de faire face aux affects pouvant être liés à l'issue défavorable de la maladie. Les éléments contribuant à l'évolution favorable ou non de la maladie, seraient alors à rechercher auprès d'autres soignants effectuant une action de prévention, un bilan du traitement ou qui suivraient régulièrement le patient : les médecins ou les infirmiers ayant des activités de secteur, de réseaux ou en libéral, voir les aidants naturels et les patients eux-mêmes.

### ***Des points de vues supposés convergents de la part des patients mais moins centrés sur la résolution du problème de santé***

Le second élément ressortant de ces observations concerne les points de vue qu'ils attendent globalement des patients. Certes les patients, en accord avec les soignants, reconnaîtraient l'impact des déterminants de la santé dans leur vie ainsi que leur caractère remédiable, c'est-à-dire aléatoire dans le temps et contrôlable par l'aide d'autrui (« *maladie-du-malade* », Canguilhem, 1972 ; Pedinelli, 1987). Néanmoins, ils sous-estimeraient globalement l'importance de ces aspects. Au regard des travaux sur le Health Belief Model, on peut considérer que les soignants confèrent aux patients des croyances moins adaptées que les leurs. Les patients, moins que les soignants percevraient l'impact des éléments pouvant contribuer à leur état de santé ; ils sous-estimeraient en partie l'aide qu'ils pourraient trouver dans leur environnement, notamment soignant. Cette représentation, en quelque sorte moins experte ou moins centrée sur le problème de santé, pourrait justifier le rôle des soignants ou leur présence dans le soin à apporter : tout en reconnaissant la nécessité, le patient en serait moins convaincu, quelque soit d'ailleurs l'évolution de son état de santé.

Tout se passe donc comme si les soignants se représentaient le point de vue du patient comme proche du leur, le confortant en partie mais justifiant néanmoins leur fonction. On peut pourtant s'interroger sur la convergence effective de ces points de vue si les patients et les aidants naturels étaient réellement interrogés. En effet dans d'autres secteurs, comme celui de l'éducation, on sait que les points de vue peuvent diverger entre enseignants, élèves et parents (Gosling, 1992). Il est vraisemblable que les infirmiers passent peu de temps avec leur patient pour le consulter quant à la façon dont il perçoit l'origine de sa maladie : l'activité diagnostique revient au médecin ; lorsqu'ils disposent d'un peu de temps avec le patient c'est, vraisemblablement pour lui expliquer le point de vue médical sur l'origine de sa maladie et de son traitement (en accord avec un modèle de relation pédagogique plutôt que contractuel).

### ***Un patient qui manifesterait son absence de contrôle en cas d'évolution défavorable de son état de santé***

L'élément le plus frappant concerne les représentations que les soignants rapportent comme étant celles des patients, concernant l'origine et la contrôlabilité des éléments faisant évoluer la maladie (mesures qui sont particulièrement pertinentes tant au regard du *Health Belief Model*, qu'au regard des travaux la valeur sociale des attributions causales). Si les soignants ne prennent pas position quant à savoir si ces éléments sont liés au patient ou à son environnement, ils s'attendent clairement à ce que le patient sous-estime ou minimise le fait que ces éléments puissent être liés à lui-même. Concernant les aidants naturels, ils s'attendent à ce que ceux-ci insistent sur les déterminants liés au contexte. De plus, cette différence avec leur propre point de vue n'est observée que lorsque l'évolution s'est avérée défavorable.

En fait, d'après les réponses données par les soignants, seuls les points de vue des patients et de leur famille seraient sensibles à l'évolution favorable ou non de la maladie. Dans les cas d'évolution défavorable, ces derniers manifesteraient une réaction « affective » : le patient préférerait voir les raisons de l'échec thérapeutique comme n'étant pas liées à lui-même et sa famille tendrait à le dédouaner de la responsabilité de cet échec thérapeutique, en évoquant la présence de facteurs contextuels. De la même façon pour la contrôlabilité des déterminants de l'état de santé par le patient, les soignants s'attendent à ce que les patients et la famille présentent l'évolution comme étant particulièrement incontrôlable par le patient lui-même lorsqu'elle est défavorable. Pour ces deux mesures (internalité et contrôlabilité), les soignants semblent donc attribuer une attitude passive et fataliste au patient qui ne guérit pas. Dans les cas d'évolution défavorable, on suppose donc que le patient voit son évolution comme essentiellement déterminée par des éléments situationnels et incontrôlables par lui, sans que l'on puisse déterminer si, du point de vue des soignants, une telle perception par les patients serait la cause ou la conséquence de l'évolution de leur état de santé (ou à long terme les deux).

Du point de vue du système de relations entre soignant et patient, on peut s'interroger sur la façon dont les soignants réagissent face à des cas de patients prétendant n'avoir aucun contrôle sur l'évolution de leur maladie, notamment lorsqu'elle n'est pas particulièrement invalidante. Dans bien d'autres secteurs du travail, on sait par exemple que les employeurs, les enseignants ainsi que les travailleurs sociaux émettent des pronostics, des jugements ou prennent des décisions plus fréquemment défavorables vis à vis des personnes faisant état de croyances externes ou d'un faible sentiment de contrôle sur ce qui leur arrive (Desrumaux-Zagrodnicki & Masclat 2001, Louche, Pansu & Papet 2001 ; Dubois, 2003 ; Pansu, 2006) (8). On peut ainsi se demander quel peut être l'impact de cette attitude sur la pratique soignante et la prise en charge (en termes de soutien émotionnel apporté ou non, de qualité de suivi, de choix du traitement médical ou de participation des patients à la décision thérapeutique).

Quoi qu'il en soit, la représentation de l'évolution de la prise en charge thérapeutique que les soignants ont donnée ici à voir, semble extrêmement normée par le savoir et leur rôle médical. Cette relation de soin asymétrique est sous tendue par une bienveillance à l'égard du patient et un objectif de guérison. Ainsi, les soignants demeurent extrêmement bienveillants en ne rejetant jamais la responsabilité de l'évolution de la maladie sur le patient (celle-ci n'étant ni spécifiquement liée au patient, ni contrôlable par lui et ce d'autant plus que la maladie est grave ; leur interprétation médicale ne variant d'ailleurs quasiment pas en fonction de l'évolution thérapeutique). Ils considèrent par ailleurs que ce qui détermine l'évolution de santé du patient, a un impact très général sur la vie du patient. En accord avec leur rôle de soignant, ils considèrent les déterminants de cette évolution comme remédiables et contrôlables par l'aide apportée au patient. De plus, selon eux, les patients en conviendraient sans en être aussi conscients. L'évolution défavorable de la maladie affecterait essentiellement les points de vue supposés du patient et de sa famille : le patient envisagerait ne pas pouvoir contrôler sa maladie ; sa famille partagerait son point de vue et mettrait en avant sur les circonstances de la maladie et de sa prise en charge.

Au niveau organisationnel, les évolutions qu'a connues la fonction infirmière semblent effectivement restreindre l'activité et l'analyse des infirmiers à l'application des soins. Les progrès techniques de la médecine ont réduit la durée des prises en charge ; leur nombre a augmenté (les pathologies devenant plus souvent chroniques que directement létales) alors que l'on fait face à difficulté à pourvoir les postes vacants (nombreux départs à la retraite accompagnés d'une diminution des vocations durables). La priorité alors donnée aux actes de soin (désormais quantifiés), plutôt qu'à l'échange avec les patients, peut avoir conditionner directement la perception que les infirmiers ont de l'état de santé de leurs patients (rôles nécessaires et prépondérants des soins et du soignant) ainsi que leur perception des différents points de vue.

### **Note de fin**

(1) S'il est évident que les points de vue des patients et de ses aidants naturels sont pris en compte dans la pratique médicale afin d'établir les symptômes dont souffre le patient, la place qui leur est attribuée dans le soin est le plus souvent périphérique: la littérature existante sur le sujet mentionne majoritairement la prise en compte éventuelle de leur point de vue pour estimer l'impact du traitement sur leur qualité de vie (Arnold, Ranchor, Sanderman, Kempen, Ormel & Suurmeijer, 2004; Conroy, Guillemin & Kaminsky, 2002; Parker, Baile, de Moor & Cohen, 2003), l'impact de la prise en charge au domicile sur le fardeau, la charge ou l'épuisement (burden) des aidants naturels ou encore les informations fournies par la famille sur l'évolution de l'état de santé de patients dans l'impossibilité permanente ou temporaire de le faire (Andrieu, Balardy, Gillette-Guyonnet, Bocquet, Cantet, Albarède, Vellas, Grand, 2003 ; Gwyther, 1998; van Wijngaarden, Schene & Koeter, 2004; Reine, Lançon, Avorio, Duplan, Siméoni, Aghababian, & Auquier, 2004 ; Magliano, Fiorillo, De Rosa, Malangone, & Maj, 2005).

(2) Seize questionnaires n'ont pas été pris en compte pour les analyses ici présentés, dans la mesure où certains infirmiers ont estimé ne pas être en mesure de rapporter le point de vue du patient et celui de sa famille (le plus souvent) pour l'un des deux cas requis. Pour les parties complétées leurs réponses ne diffèrent statistiquement pas des autres.

(3) D'un point de vue qualitatif, les réponses fournies par les soignants sont difficilement traitables. Concernant leur propre point de vue, elles sont essentiellement fondées

sur l'état somatique (malformation congénitales, sur poids, tabac...) ou contextuelle (stress, travail difficile physiquement...). Le tableau clinique décrit est celui de la pathologie considérée et du traitement, assez indépendamment de son effet (le traitement lui-même et éventuellement l'état clinique antérieur plus complexe que prévu). Leur point de vue reste essentiellement centré sur le domaine médical même si l'influence de l'environnement familial est parfois évoquée (notamment dans les prises en charge pédiatriques). Le point de vue qu'ils supposent être celui des malades et des aidants naturels témoigne d'assez peu de différences avec leur propre point de vue (parfois avec une difficulté à énumérer autant d'éléments qualitatifs et à se décentrer de leur propre point de vue). L'absence a priori de divergence au niveau qualitatif se manifesterait notamment lorsque la prise en charge s'est bien passée: les malades et leur famille manifesteraient leur conscience des symptômes et viendrait dans les services de soin se faire aider (incapacité à respirer seul, le désir de guérir, le risque de perdre son autonomie...). En revanche, en ce qui concerne l'évolution défavorable, certains éléments seraient mis en avant et pas uniquement dans le registre d'un état clinique déjà difficile. Dans ce dernier cas, des facteurs familiaux sont rapportés pour les malades comme le divorce, le temps mis avant de se faire soigner ou la négligence de leur état somatique. La famille quant à elle, remettrait parfois en cause de la qualité et de l'efficacité des soins, invoquerait la fatalité liée souvent à l'état clinique déjà mauvais, remettrait en cause de la combativité du malade ou évoquerait sa propre difficulté à apporter une aide au malade (détresse, épuisement; absence de visite...).

(4) La Clinical Global Impression est constituée de 3 items indépendants utilisés dans presque toutes les études où on évalue une prise en charge médicale ou l'efficacité d'un médicament. Le principal avantage de ces trois items est leur simplicité d'emploi et le fait qu'ils soient adaptés à toutes les pathologies. Le premier item (mesure de la gravité du trouble actuel du patient) et le second (mesure de l'amélioration globale) sont quottés sur sept paliers de réponse. Le troisième item (mesure combinée de l'effet clinique et des effets secondaires) dispose de 16 réponses possibles. Cet item mesure à la fois le bénéfice thérapeutique (sur quatre paliers) et l'absence d'effets secondaires néfastes induits par le traitement (sur quatre paliers aussi). Ces trois items donnent lieu à trois mesures indépendantes et sont utilisés indépendamment les uns des autres. L'intérêt d'utiliser un tel outil est double. En début de questionnaire l'item 1 CGI, permettait aux soignants de situer le degré de gravité; le niveau de gravité devant être explicitement identique au départ dans le choix de leurs deux patients, seule l'évolution devant différer (évolution également vérifiable par les items 2 et 3).

(5) La Global Assessment of Functioning Scale (GAF) est une échelle d'origine psychiatrique (American Psychiatric Association) qui évalue le fonctionnement psychologique, social et professionnel de la personne (atteinte d'une pathologie psychiatrique dans sa version initiale). Il est demandé au praticien effectuant le suivi du patient d'évaluer le fonctionnement psychologique, social et professionnel global de celui-ci sur une échelle de 1 à 100%. Des indicateurs clairs et concis sont offerts pour chaque tranche de 10 points de l'échelle. La GAF constitue le résultat opérationnel de l'axe V de l'évaluation multiaxiale du DSM-IV. Il s'agit d'une forme légèrement modifiée du Global Assessment Scale et peut être considérée comme fiable. Afin de la rendre plus facilement utilisable dans des situations cliniques non psychiatriques, toutes les références à un trouble mental ont été retirées ou remplacées par une référence simple :« la maladie ». La forme initiale, pour le reste, a été conservée. La G.A.F. est pertinente ici car elle donne une vision complémentaire à la CGI de l'évolution des deux patients en élargissant l'évaluation des seules conséquences somatiques de la maladie au fonctionnement social du sujet.

(6) Indépendamment de l'évolution de la maladie (favorable ou non), la gravité initiale des symptômes semble être un facteur prédictif de l'incontrôlabilité des déterminants de santé par le patient. En effet les soignants considèrent que la maladie était d'autant moins contrôlable par le patient que ses symptômes étaient graves (pour le cas d'évolution favorable :  $R = -.39$  avec la

contrôlabilité et  $R=+.37$  avec l'incontrôlabilité ; de même globalement pour l'évolution défavorable :  $R=-.31$  et  $R=0.13ns$ ). De la même façon, ils considèrent que la maladie était d'autant moins contrôlable par le patient que son fonctionnement psychosocial est altéré (mesuré par la G.A.F. ; pour le cas d'évolution favorable  $R=+.42$  et  $R=-.33$  ; pour le cas défavorable  $R=+.39$  et  $R=-0.18ns$ ).

(7) Les infirmiers doivent avant tout suivre la bonne application du traitement prescrit par le médecin. Le diagnostic d'une telle évolution de santé et du bénéfice des traitements choisis, ne relève a priori pas de leur rôle, mais de celui d'un médecin référent. On peut pourtant s'inquiéter du fait que les infirmiers hospitaliers ne puissent différencier les déterminants d'une évolution effectivement défavorable et de ceux d'une évolution favorable, pour anticiper une telle évolution dans leur soin.

(8) Au-delà des choix alors à faire dans le traitement (que les soignants peuvent aisément contrôler par l'application strict d'un protocole) on peut se demander si, face à l'attitude de tels patients, leur satisfaction et la qualité de la relation n'est pas altérée de façon suffisamment perceptible par le patient pour amener ces derniers, au travers d'un cercle vicieux, à une moindre observance et une moindre croyance en l'effet du traitement (Hessen-Klemens & Lapinska, 1984 ; Squier & Jackson, 1990).

## Références

- Beauvois, J.-L., Gilibert, D., Pansu, P., Abdelhaoui, S. (1998). Internality attribution and intergroup relations. *European Journal of Social Psychology*, 28, 123-140.
- Andrieu, S., Balardy, L., Gillette-Guyonnet, S., Bocquet, H., Cantet, C., Albarède, J.-L., Vellas, B., Grand, A., le groupe REAL.FR. (2003). Charge ressentie par les aidants informels de patients atteints de la maladie d'Alzheimer au sein de l'étude REAL.FR: méthode de mesure et facteurs associés. *La revue de Médecine interne*, 24, 351s-359s.
- Arnold, R., Ranchor, AV., Sanderman, R., Kempen, GI., Ormel, J., Suurmeijer, TP. (2004). The relative contribution of domains of quality of life to overall quality of life for different chronic diseases. *Qual Life Res.*, 13, 883-96.
- Benyamini, Y., Leventhal, H., Leventhal, EA. (2004). Self-rated oral health as an independent predictor of self-rated general health, self esteem and life satisfaction. In *Social Science & Medicine*, 59, 1109-1116.
- Bradley, C. (1994). Measures of perceived control in diabetes. In *Handbook of Psychology of Diabetes*. London: Harwood Academic Publishers.
- Bradley, C. (1995). Health beliefs and knowledge of patients and doctors in clinical practice and research. *Patient Education and Counseling*, 26, 99-106.
- Canguilhem, G. (1972). *Le normal et le pathologique*. Paris : PUF.
- Conroy, T., Guillemin, F., Kaminsky, MC. (2002). Mesure de la qualité de vie des patients atteints de cancer colorectal métastatique : techniques et principaux résultats. *La revue de médecine interne*, 23, 703-716.
- Desrumaux-Zagrodnicki, P., Masclet, G. (2001). Stratégies d'explications internes et externes en fonction du contexte de chômage ou de recrutement, du niveau professionnel et du statut social. *Bulletin de Psychologie*, 54, 295-305.
- Dew, MA., Goyocoola, J., Harris, RC., Lee, A., Zomak, R., Dunbar-Jacob, J., Rotondi, A. (2004). An internet-based intervention to improve psychosocial outcomes in heart transplant recipients and family caregivers: development and evaluation. *J Heart Lung Transplant*, 23, 745-58.
- Dubois, N. (2003). *A sociocognitive approach to social norms*. London: Routledge

- Dubois, N., Beauvois, J.-L. (1996). Internality, academic status and intergroup attributions. *European Journal of Psychology of Education*, 11, 329-341.
- Elbogen, EB., Swanson, JW., Swartz, MS. (2003). Effects of legal mechanisms on perceived coercion and treatment adherence among persons with severe mental illness. *J Nerv Ment Dis.*, 191, 629-37.
- Elder, JP., Ayala, G., Harris, S. (1999). Theories and intervention approaches to health-behavior change in primary care. *Am J Prev Med*, 17, 275-284.
- Fischer, G-N., Tarquinio, C. (2006). *Les concepts fondamentaux de la psychologie de la santé*. Paris: Dunod.
- Fontayne, P., Martin-Krumm, C., Buton, F., Heuzé, J.-F. (2003). Validation française de la version révisée de l'échelle de mesure des attributions causales (CDSII). *Cahiers Internationaux de Psychologie Sociale*, 58, 59-72.
- Gao, X, Nau, DP., Rosenbluth, SA., Scott, V., Woodward, C. (2000). The relationship of disease severity, health beliefs and medication adherence among HIV patients. *AIDS Care*, 12, 387-98.
- Gilibert, D., Cambon, L. (2003). Paradigms of the sociocognitive approach. In N., Dubois (Ed.), *A sociocognitive approach to social norms*, (pp 38-69), London: Routledge.
- Goldring, A., Taylor, S., Kemeny, M., Anton, P. (2002). Impact of health beliefs, quality of life, and the physician-patient relationship on the treatment intentions of inflammatory bowel disease patients. *Health Psychology*, 21, 219-228.
- Gosling, P. (1994). The attribution of success and failure: the subject/object contrast. *European Journal of Psychology of Education*, 9, 69-83.
- Gosling, P. (1995). Description vs. Explanation and academic evaluation. *European Journal of Psychology of Education*, 10, 41-59.
- Graham, D., Kunik, M., Doody, R., Snow, A. (2005). Self-reported awareness of performance in dementia. in *Cognitive Brain Research*, 2005, in press. Available online at [www.sciencedirect.com](http://www.sciencedirect.com).
- Guy W. (1976) *Early Clinical Drug Evaluation (ECDEU) assesment manual for psychopharmacology*. National Institute of Mental Health, Rockville.
- Gwyther, LP. (1998). Social issues of the Alzheimer's patients and family. in *The American Journal of Medecine*, 104, 17S-21S.
- Louche, C., Pansu, P., Papet, J. (2001). Normes de jugement et appréciation du personnel. *Bulletin de Psychologie*, 54, 369-374.
- Hessen-Klemens, I., Lapinska, E. (1984). Doctor-patient interaction, patients' health behavior and effects of treatment. *Social Science and Medicine*, 19, 9-18.
- Horne, R., James, D., Petrie, K., Weinman, J., Vincent, R. (2000). Patient's interpretation of symptoms as a cause of delay in reaching hospital during acute myocardial infarction. *Heart*, 83, 388-93.
- Horne, R., Weinman, J. (1999). Patient's beliefs about prescribed medicines and their role in adherence to treatment in chronic physical illness. *Journal of Psychosomatic Research*, 47, 555-567.
- Jakson, J., Kincey, J., Fiddler, M., Creed, F., Tomeson, B. (2004). Differences between out-patients with physical disease and those with medically unexplained symptoms with respect to patient satisfaction, emotional distress and illness perception. *Br J Health Psychol.*, 9, 433-46.
- Kouabenan, D.R. (2002). Occupation, driving experience and risk and accident perception. *Journal of Risk Research*, 51, 49-68.

- Magliano, L., Fiorillo, A., De Rosa, C., Malangone, C., Maj, M., the National Mental Health Project Working Group. (2005). Family burden in long-term diseases: a comparative study in schizophrenia vs. Physical disorders. *Social Science & Medicine*, 61, 313-322.
- Le Sommer-Péré, M. (2000). Relation soigné-soignant: quels modèles de référence? *Perspective soignante*, 9, 67-78.
- Matteucci, M-C., Gosling, P. (2004). Italian and French teachers faced with pupil's Academic Failure: The "Norm of Effort". *European Journal of Psychology of Education*, 19, 147-166.
- Norman, P. (1995). Health locus of control and health behavior : an investigation into the role of health value and behaviour-specific efficacy beliefs. *Personn. Individ. Diff.*, 18, 213-218.
- Parker, PA., Baile, WF., de Moor, C., Cohen, L. (2003). Psychosocial and demographic predictors of quality of life in large sample of cancer patients. *Psychooncology*, 12, 183-93.
- Pedinelli, J-L. (1987). Le « travail de la maladie ». *Psychologie Médicale*, 19, 1047-1052.
- Pibernik-Okanovic, M., Prasek, M., Poljicanin-Filipovic, T., Pavlic-Renar, I., Metelko, Z. (2004). Effects of an empowerment-based psychosocial intervention of quality of life and metabolic control in type 2 diabetic patients. *Patient Educ Couns.*, 52, 193-9.
- Politis, C. (1998). The psychosocial impact of chronic illness. In *Ann N Y Acad Sci*, 30, 349-54.
- Pucheu, S., Consoli, S., D'Auzac, C., Français, P., Issad, B. (2004). Do health causal attributions and coping strategies act as moderators of quality of life in peritoneal dialysis patients? In *Journal of Psychosomatic Research*, 56, 317-322.
- Reine, G., Lançon, C., Avorio, A., Duplan, S., Siméoni, MC., Aghababian, V., Auquier, P. (2004). Comment mesurer la charge des aidants naturels de personnes souffrant de schizophrénie ? *Ann Méd Psychol*, 162, 453-61.
- Rosenstock, I.M. (1974). Historical origins of the health belief model. *Health Education Monographs*, 2, 354-386.
- Rosenstock, I.M., Strecher, V.J., Becker, M.H. (1988). Social learning theory and the health belief model. *Health Education Quarterly*, 15, 175-183.
- Shiloh, S., Rashuk-Rosenthal, D., Benyamini, Y. (2002). Illness causal attributions : an exploratory study of their structure and association with other illness cognitions and perceptions of control. *Journal of Behavioral Medicine*, 25, 373-394.
- Squier, D.J., Jackson, T.C. (1990). A model of empathic understanding and adherence to treatment regimens in practionner-patient relationships. *Social Science and Medicine*, 30, 325-339.
- Van Wijngaarden, B., Schene, A., Koeter, M. (2004). Family caregiving in depression: impact on caregivers' daily life, distress and help seeking. *Journal of Affective Disorders*, 81, 211-22
- Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, 92, 548-573

# **Sens du travail et rapports vie de travail - vie hors travail chez trois catégories de télétravailleurs et chez les travailleurs classiques**

## **Meaning of work and relations work life – out of work life among three categories teleworkers and traditional workers**

Emmanuelle Camps

[camps.emmanuelle@wanadoo.fr](mailto:camps.emmanuelle@wanadoo.fr)

Développement, Cognition, Acquisition, Université Paul Valéry, rte de Mende,

B.P. 5043

34199 Montpellier cedex 5.

### *Résumé*

Le télétravail, appelé aussi travail à distance, se décline selon trois formes de base : le télétravail à domicile, le travail en mode alterné et le télétravail nomade. Ces formes de télétravail sont comparées entre-elles et à la forme d'emploi classique pour identifier leurs effets respectifs sur la centralité relative du travail, la représentation sociale du travail et les rapports vie de travail - vie hors travail. Le télétravail nomade s'accompagne d'une plus forte centralité du travail alors que le travail en mode alterné conduit à une augmentation de la centralité de la famille. Sur le plan de la représentation, le travail ne peut être envisagé sans la notion de contraintes chez les télétravailleurs nomades et les travailleurs classiques. Au niveau des échanges entre la vie de travail et la vie hors travail, les travailleurs classiques séparent leurs domaines de vie alors que les travailleurs en mode alterné et les télétravailleurs à domicile les concilient. Les télétravailleurs nomades quant à eux favorisent les échanges entre les domaines de vie. Cette étude corrobore l'idée selon laquelle il est préférable, non pas de raisonner globalement pour étudier le travail à distance, mais de tenir compte des spécificités de chacune des formes de télétravail implantées ou à instaurer au sein d'une organisation.

### *Abstract*

Telework, namely so remote work, is composed of three basic forms: home based-telework, alternated telework and mobile telework. These three basic forms of telework and the traditional form of work are compared to identify their respective impact on work centrality, work social representation, and on relations between work life and out of work life. The mobile teleworkers present a high work centrality and the alternated teleworkers, a high family centrality. Concerning the representation, only mobile teleworkers and home based teleworkers cannot think the work without the basic knowledge of constraint. For the exchanges between work of life and out of work life, traditional workers separate their domains of life even though alternated teleworkers and home based-teleworkers combine it. Mobile teleworkers develop exchanges between some domains of life.

This study proves that it is not necessary to reason globally but to consider the specificity of each telework forms introduced or to implant in a organization.

**Mots clés** : forme de télétravail, travail classique, centralité du travail, représentation du travail, échanges entre les domaines de vie.

**Key-Words**: telework form, traditional work, work centrality, work representation, exchanges between domains of life.

Le télétravail, appelé travail à distance, émerge progressivement en France dans les entreprises à partir des années 1980. Plusieurs facteurs sont selon Kouloumdjian (2000, pp.351-355), à l'origine de son développement :

- la montée en puissance des technologies de l'information et de la communication,
- l'informatisation, le déploiement des réseaux (internet, intranet) et de la téléphonie mobile dans les entreprises,
- la mondialisation,
- la tertiarisation du travail.

L'implantation du télétravail au sein des entreprises provoque un changement organisationnel qui pose un certain nombre d'interrogations tant sur un plan juridique, technique, que psychologique. En effet, le télétravail constitue une situation particulière d'activité, qui conduit à analyser son impact sur la relation au travail et sur les rapports vie de travail - vie hors travail. En voici les principales caractéristiques.

### **1. Les caractéristiques de la situation de télétravail**

Le télétravailleur est un salarié qui travaille à distance de son entreprise de rattachement. Il est éloigné géographiquement de son responsable hiérarchique et de ses collègues de travail. Il maintient un contact régulier virtuel grâce aux nouvelles technologies de l'information et de la communication pour recevoir et envoyer le travail demandé (Breton, 1994 ; accord cadre sur le télétravail, 2002). Le télétravail n'est donc pas une nouvelle activité en soi mais une nouvelle forme de travail qui rompt avec les repères traditionnels du salariat : le travailleur n'est plus obligé d'exécuter ses tâches à l'intérieur de l'entreprise et il échappe au contrôle direct de son supérieur hiérarchique.

La typologie des formes de base du télétravail repose sur la différenciation des lieux d'exercice de l'activité professionnelle. Trois grandes catégories sont distinguées : le télétravail à domicile, le télétravail nomade et le travail en mode alterné ou télépendulaire.

- Le télétravailleur à domicile exerce son activité professionnelle chez lui, avec du matériel appartenant en général à son employeur.
- Le télétravailleur nomade effectue ses tâches à distance à différents endroits, sans revenir à son bureau (chez un client, à la maison, dans une filiale de l'entreprise, dans un train, un avion...)
- Le télépendulaire occupe son poste alternativement, selon un rythme prédéfini, sur son lieu de travail et à domicile ou dans un télécentre, ou dans un bureau satellite, etc.

Chaque forme de télétravail est singulière. Par comparaison aux travailleurs en mode alterné et aux télétravailleurs nomades, les télétravailleurs à domicile sont plus souvent isolés d'un collectif de travail et les barrières entre leur vie professionnelle et leur vie privée sont plus étanches. Ces contextes plus ou moins spécifiques, des situations de télétravail, peuvent donc induire des types de relation au travail, eux -

mêmes plus ou moins spécifiques (sens du travail, représentation sociale du travail), ainsi que de nouveaux rapports entre la vie de travail et la vie hors travail. C'est l'objet de l'étude qui suit.

## **2. Le télétravail et la relation au travail**

Joulain (2006) relève un certain nombre de facteurs dont l'âge qui agissent sur la relation au travail. La nouvelle recherche qui est ici proposée, s'inscrit dans la thématique des rapports vie de travail-vie hors travail et analyse la relation au travail selon la forme de télétravail pratiquée. Précisément, il s'agit de s'intéresser au sens du travail et à la représentation sociale du travail dans ces situations particulières d'activité.

### **Le sens du travail : définitions**

Le modèle «the Meaning of Work» (1987) qui a été validé dans un certain nombre de recherches et la conception proposée par Morin (1997) définissent le sens du travail dans la littérature.

D'après l'équipe du Mow (1987), le sens du travail comporte cinq dimensions :

- la centralité du travail : elle renvoie à la place que l'individu fait au travail dans l'absolu. Elle est également constituée, sur un plan relatif, par l'importance accordée au travail par rapport à celle attribuée à d'autres domaines de vie (famille, loisirs, activités sociales).

- Les motifs pour lesquels l'individu travaille (statut, prestige, contacts, occupation...)

- L'importance attachée aux aspects instrumentaux (salaire, possibilités de promotion...) et aux aspects expressifs (autonomie, variété, intérêt...) du travail.

- L'orientation normative correspondant à l'intériorisation par l'individu des normes sociales concernant le travail.

- L'identification à des rôles que l'individu accomplit en travaillant.

Selon Morin (1997), le sens du travail se décrit grâce à trois composantes :

- la signification qui se rapporte aux représentations du travail, lesquelles impliquent la centralité du travail (absolue et relative),

- l'orientation qui fait référence aux motivations qui incitent la personne à travailler,

- la cohérence qu'une personne trouve dans son travail qui correspond au degré d'équilibre qu'elle atteint en travaillant.

Le sens du travail pour tel ou tel groupe s'étudie à partir des significations qui lui sont collectivement attribuées (Harpaz, 1985 ; Mow International Research Team, 1987). Le cadre théorique des représentations sociales permet d'approcher les significations assignées par un groupe à un objet social. L'étude des représentations du

travail permettra donc de saisir les significations que les groupes attribuent au travail, donc d'appréhender la notion de sens du travail.

### **La représentation sociale du travail**

Jodelet (1989, p.36) définit la représentation sociale «comme une forme de connaissance, socialement élaborée, ayant une visée pratique et concourant à la construction d'une réalité commune à un ensemble social». Pour compléter cette définition, Abric (1994, p.13) ajoute : «elle est à la fois le produit et le processus d'une activité mentale par laquelle un individu ou un groupe reconstitue le réel auquel il est confronté et lui attribue une signification spécifique».

La théorie du noyau central développée par Abric (1987) affirme que les représentations sociales fonctionnent comme une entité mais avec deux composantes : le système central et le système périphérique. Ils se caractérisent par un rôle spécifique mais complémentaire l'un par rapport à l'autre :

- le noyau central constitue une structure particulièrement stable qui résiste au changement, il fonde ainsi la permanence de la représentation (Abric, 1994). Il assure également la base commune, c'est-à-dire collectivement partagée, des représentations sociales et par conséquent sa fonction consensuelle.

- Les éléments périphériques sont sous la dépendance du noyau central. Ils sont caractérisés par une plus grande souplesse et forment l'interface entre le noyau central et la réalité concrète propre à l'objet (Abric, 1994). Ils permettent ainsi d'intégrer dans la représentation, des variations individuelles déterminées par des expériences spécifiques.

Deux représentations sont considérées comme différentes si et seulement si leurs éléments centraux sont différents. L'étude de la représentation sociale du travail va donc consister à identifier, dans l'ensemble des opinions et croyances relatives à l'objet travail, les éléments appartenant au noyau.

Les caractéristiques de la situation de télétravail permettent non seulement d'analyser la relation au travail mais aussi de s'interroger sur les rapports entre la vie de travail et la vie hors travail, à partir du modèle théorique du système des activités proposé par Curie et Hajjar (1987). Il est à constater en effet que le télétravailleur ne réalise plus seulement son travail dans les murs clos de l'entreprise. Dans le cas du télétravail à domicile, le salarié travaille chez lui, donc la séparation entre sa vie de travail et sa vie hors travail est susceptible d'être abolie. Le travailleur en mode alterné quant à lui, exerce son activité professionnelle selon un rythme prédéfini en partie dans son entreprise et en partie à son domicile ou dans un télécentre. Enfin, le télétravailleur nomade se déplace régulièrement, il travaille dans des endroits différents y compris dans son logement. Ces nouvelles configurations d'organisation du travail atténuent chacune spécifiquement la séparation physique et psychologique entre la vie de travail et la vie hors travail, donc rendent par conséquent pertinentes l'étude des relations entre la sphère de vie professionnelle et extra-professionnelle en fonction de la forme de télétravail pratiquée.

### **3. Le télétravail et les rapports vie de travail – vie hors travail**

Curie et Hajjar (1987) créent le modèle du système des activités pour rendre compte dans leur ensemble des interactions entre la vie de travail et la vie hors travail. Ils évitent ainsi d'analyser les relations en terme d'interdépendance unilatérale d'une sphère de vie vis-à-vis de l'autre. L'intérêt de ce modèle pour la recherche est d'étudier les relations réciproques entre les domaines de vie professionnel, familial, personnel et social. Ces échanges de ressources ou de contraintes vont dépendre du ou des lieux dans lesquels s'exerce l'activité professionnelle, donc de la forme de télétravail.

#### ***Le système des activités***

D'après le modèle de Curie et Hajjar (1987) (cf. également Baubion-Broye, 1998 ; Curie, 2000), les activités des individus forment un système, dans la mesure où chacune d'elles peut constituer une ressource ou une contrainte pour l'ensemble des autres. Dès lors, on doit considérer qu'elles sont interdépendantes : afin de réaliser une activité dans un secteur de vie (professionnel par exemple), la personne peut s'appuyer sur d'autres activités relatives à des domaines différents, mais elle doit aussi tenir compte des obstacles éventuels que constituent certaines des activités « extérieures » à ce secteur. Un ensemble d'activités relevant de divers domaines de vie (professionnel, familial, personnel et social, etc.) sont ainsi intégrées dans une perspective systémique. Les activités sont constituées en sous-systèmes (ou domaines) spécifiques et relativement autonomes entre lesquels se produisent des échanges d'ordre matériel, temporel, énergétique, affectif et informationnel. Chaque sous-système est relativement autonome : il a des objectifs propres à la réalisation desquels peuvent concourir des moyens spécifiques. Les sous systèmes sont dépendants les uns des autres : des échanges de ressources et de contraintes s'effectuent entre-eux de sorte que les activités d'un sous-système peuvent servir (ou gêner) le déroulement des activités des autres sous-systèmes.

L'ensemble du système obéit à des régulations qui peuvent se produire aussi bien à un niveau local (au sein d'un domaine) selon le modèle d'action, qu'à un niveau central (entre domaines) en fonction du modèle de vie des sujets.

### **4. Problématique**

Louche & Moliner (2001) ont examiné l'influence du télétravail sur le sens du travail et la représentation sociale du travail. Les résultats montrent que le télétravail à domicile s'accompagne d'une forte centralité du travail par rapport au travail classique. Au niveau des attentes, les télétravailleurs négligent certains aspects expressifs du travail et sur le plan de la représentation, ils sont plus sensibles au caractère valorisant du travail, alors que les travailleurs classiques sont davantage attachés au caractère ludique et contraignant du travail.

Cette recherche permet d'apprécier combien le télétravail agit sur la relation de travail mais elle ne concerne qu'une seule forme de télétravail et omet d'étudier les relations entre les sphères d'activité. Or Lancry et al. (2000) observent que le télétravail

à domicile peut permettre une meilleure adéquation entre vie professionnelle et vie privée. Par conséquent dans la recherche présentée, d'autres formes de télétravail sont étudiées et comparées entre-elles par rapport au fonctionnement du système des activités.

L'objectif est d'identifier d'une part le sens du travail et les représentations du travail selon la forme de travail pratiquée, d'autre part d'analyser le fonctionnement du système des activités en fonction également de la forme de travail adoptée (télétravail à domicile, travail en mode alterné, télétravail nomade, travail classique).

## **5. Hypothèses**

### **5.1 Hypothèse sur la centralité relative du travail**

D'après les entretiens réalisés au cours de la pré enquête, les télétravailleurs à domicile ainsi que les travailleurs en mode alterné choisissent de pratiquer cette forme d'organisation du travail pour pouvoir concilier leur vie professionnelle et leur vie familiale, donc pour atteindre les objectifs qu'ils se sont fixés dans ces deux domaines de vie. Par conséquent, les degrés de centralité à l'égard du travail et de la famille seront très élevés chez ces deux groupes de salariés. Les résultats obtenus par Louche et Moliner (2001), selon lesquels le travail est très important pour ceux qui télétravaillent à domicile, servent aussi de référence.

Par contre, les télétravailleurs nomades, salariés en déplacement, argumentent leur intérêt à continuer à télétravailler pour atteindre en priorité leurs objectifs professionnels. Grâce à cette forme d'organisation du travail, ils ne sont plus en effet obligés de retourner à leur bureau avant ou après un rendez-vous chez le client, donc ils peuvent terminer un travail dans les délais quels que soient le lieu et l'heure. Dans ce contexte, ils privilégieront la sphère professionnelle par rapport à la sphère familiale.

Quant aux travailleurs classiques, ils présentent une centralité du travail et de la famille très élevée (Camps, 2006).

Un certain nombre de recherches effectuées sur ce thème révèlent d'autre part que les loisirs et les activités sociales arrivent pour les salariés en dernière position dans leur hiérarchie des valeurs après la famille et le travail (Louche et Moliner, 2001 ; Camps, 2006).

A partir de l'ensemble de ces données, l'hypothèse suivante peut être énoncée : la centralité du travail ne variera pas selon la forme d'emploi pratiquée, elle sera élevée chez l'ensemble des salariés. La centralité des loisirs et celle des activités sociales ne différeront pas non plus selon la forme d'emploi considérée. Elles seront faibles chez l'ensemble des personnes interrogées. La centralité de la famille sera élevée auprès des quatre groupes de salariés mais elle sera moins forte chez les télétravailleurs nomades.

## 5.2. Hypothèse sur la représentation sociale du travail

Dans plusieurs études, l'élément rémunération se révèle constitutif du noyau central de la représentation sociale du travail de plusieurs populations : les cadres (Flament, 1994), les télétravailleurs à domicile et les travailleurs classiques (Louche & Moliner, 2001), les travailleurs en mode alterné et les travailleurs classiques (Camps, 2006).

Le système périphérique est quant à lui sensible au contexte immédiat, il est mouvant et évolutif, donc permet l'intégration des expériences individuelles et, par conséquent, il variera selon la forme de travail pratiquée. Camps (2006) montre en effet que le besoin est un élément périphérique plus saillant de la représentation du travail chez les travailleurs classiques que chez les télétravailleurs. Dans cette recherche, les travailleurs classiques seront donc plus sensibles à certains aspects contraignants du travail. Les télétravailleurs à domicile intégreront également des aspects plutôt négatifs (nécessité, contraintes, besoin, manque de temps libre) dans leur système périphérique en raison de la non légitimité de leur statut pour certains managers. Par contre, les travailleurs en mode alterné et les télétravailleurs nomades seront plus sensibles aux aspects positifs du travail (épanouissement, plaisir, utilité sociale, confiance en soi) puisque la forme de télétravail qu'ils pratiquent, semble faciliter leur intégration dans l'équipe.

D'après les éléments qui viennent d'être évoqués, la seconde hypothèse qui en découle est la suivante : l'élément rémunération sera constitutif du noyau central de la représentation sociale du travail chez l'ensemble des salariés. En revanche, le système périphérique variera selon la forme de travail pratiquée. Les travailleurs classiques et les télétravailleurs à domicile seront plus sensibles aux aspects contraignants du travail (nécessité, contraintes, besoin, manque de temps libre) tandis que les travailleurs en mode alterné et les télétravailleurs nomades seront plus sensibles aux aspects positifs du travail (épanouissement, plaisir, utilité sociale, confiance en soi).

## 5.3. Hypothèse sur les rapports vie de travail - vie hors travail

Le travail en mode alterné permet une présence au domicile, même s'il se traduit par une implication limitée et mesurée (travailler chez soi le mercredi pour être présent auprès de ses enfants par exemple). Le télétravail à domicile favorise, dans la durée, l'ajustement au mieux des exigences professionnelles et des impératifs familiaux (s'occuper d'un enfant handicapé par exemple) en jouant sur la souplesse des horaires que permet l'autonomie. Ainsi, qu'il s'agisse d'assurer un soutien scolaire, d'être présent pendant une période raisonnable (mettons, à partir de 18h30) ou de garder les enfants quand ils sont malades ; hommes ou femmes trouvent dans le travail en mode alterné et le télétravail à domicile une solution leur permettant de résoudre le conflit entre deux types d'exigences : achever son travail dans des délais convenables et profiter de sa famille.

Le télétravail nomade est pratiqué par le salarié en raison de la nature de son activité professionnelle qui l'oblige à travailler le plus souvent à l'extérieur. Il évolue donc dans un espace où les domaines de vie sont séparés physiquement.

Camps (2006) constate par ailleurs que les travailleurs classiques établissent des barrières entre leur vie professionnelle, familiale, personnelle et sociale.

Par conséquent, d'après les caractéristiques rattachées aux situations de travail, l'hypothèse suivante est posée : les télétravailleurs à domicile et les travailleurs en mode alterné tenteront de concilier leur vie de travail et leur vie hors travail alors que les télétravailleurs nomades et les travailleurs classiques établiront une séparation entre ces deux sphères de vie.

## **6. Contexte**

Cette recherche s'est déroulée dans une grande entreprise privée de 145264 salariés, pionnière pour la mise en place du télétravail. Le nombre de télétravailleurs nomades (4000 salariés) est plus important que le nombre de travailleurs en mode alterné (400 salariés) et surtout de télétravailleurs à domicile (30 salariés). En effet, la forme nomade du télétravail correspond à une évolution logique des postes de travail associés à des métiers itinérants, en effectif important dans le groupe concerné. Celle-ci n'induit pas une totale refonte de l'organisation du travail et un changement profond des mentalités à l'inverse d'autres formes de télétravail, comme par exemple le télétravail à domicile. Ces différences constatées nécessitent de comparer ces trois formes de télétravail.

### **6.1. Population**

La population est constituée de trois échantillons de télétravailleurs (58 télétravailleurs nomades, 45 travailleurs en mode alterné, 30 télétravailleurs à domicile) et d'un échantillon de 60 travailleurs classiques.

A l'exception des télétravailleurs nomades qui sont majoritairement des hommes (81,03%), la mixité caractérise les autres catégories de travailleurs. La plupart des salariés sont âgés de 41 à 50 ans, les travailleurs en mode alterné sont légèrement en plus grand nombre dans la tranche d'âge 31-40 ans. Tous ont un niveau de formation supérieur au bac, principalement de 3ème cycle chez les travailleurs en mode alterné, les télétravailleurs nomades et les travailleurs classiques. Majoritairement, ils ont donc le statut de cadre, voire de cadre supérieur.

La filière commerciale est représentative de cette population, en effet le service commercial/marketing est celui qui comprend le plus gros des effectifs. Seuls les télétravailleurs à domicile sont en plus grand nombre dans la filière administrative et sont rattachés au service informatique.

Les salariés interrogés ne managent pas d'équipe, ne travaillent pas à temps partiel, ont un emploi stable et plus de dix ans d'ancienneté dans l'entreprise. Ils vivent en couple, ont des enfants et leur conjoint travaille.

## 6.2 Procédure empirique

L'intervention sur le terrain se déroule en deux phases :

- la première phase est la réalisation d'une pré enquête : conduite d'entretiens auprès d'un échantillon de 10 télétravailleurs et de 10 travailleurs classiques. Chaque interviewé dispose d'un guide d'entretien qu'il complète par téléphone avec l'interviewer. Les thèmes abordés se rapportent pour chaque domaine de vie aux changements occasionnés par le télétravail, et en conséquence aux avantages et aux inconvénients qu'il génère. Les entretiens visent aussi à identifier les motivations du salarié à télétravailler, et à recueillir les conseils qui seraient à promulguer auprès de télétravailleurs potentiels.

- La deuxième étape correspond au recueil des données sur le terrain : les questionnaires élaborés, sont envoyés par e-mail auprès des trois groupes de télétravailleurs et auprès des travailleurs classiques.

## 6.3. Instrument utilisé

Le questionnaire sert à mesurer les composantes du sens du travail et la structure des échanges entre les domaines de vie.

La première partie concerne la centralité relative du travail. Les employés doivent répartir 100 points entre quatre domaines de vie (travail, loisirs, activités sociales et familiales). Selon cette procédure utilisée dans des recherches antérieures (Mow International Research Team, 1987), le nombre de points accordé à un domaine traduit l'importance qu'il a pour l'individu.

La deuxième partie traite des échanges entre les domaines de vie. Les questions posées correspondent à une transposition de l'exercice 3 «structure des échanges entre les domaines de vie» de la technique ISA (Inventaire du Système des Activités)<sup>3</sup> sous forme papier crayon. Inspirées de celles proposée par Megemont (2000) dans sa thèse, des modifications sont apportées afin de faciliter la passation de l'outil. Pour chaque objectif prioritaire sélectionné dans chaque sous-système, le sujet doit indiquer s'il constitue une aide, un obstacle pour atteindre l'objectif des deux autres sous-systèmes, ou bien si ne coexiste aucun lien entre eux, sur une échelle en 7 points : de 1 «obstacle important» à 4 «sans rapport et de 4 «sans rapport» à 7 «aide importante.»

La troisième partie se rapporte à la représentation sociale du travail. Une liste de 16 items est utilisée, largement inspirés de Salmaso, Pombéni (1986) et de Flament (1994). Le questionnaire met en œuvre une technique de mise en cause (Moliner, 1994). Celle-ci permet de repérer les éléments centraux d'une représentation sociale, c'est à dire, ceux qui sont indispensables à la reconnaissance de l'objet, selon la théorie des représentations sociales. Cette technique de mise en cause consiste à décrire un objet ou une situation présentant une contradiction avec un élément de la représentation (par

---

<sup>3</sup> Pour une description précise de la technique, voir Curie, Hajjar, Marquie, & Roques, 1990.

exemple, une activité qui ne donne lieu à aucune rémunération). Les sujets doivent alors décider si l'objet qu'on leur a décrit rentre bien dans le champ de la représentation (cette activité est-elle un travail ?) En théorie, si la contradiction concerne un élément central de la représentation, les sujets doivent répondre par la négative. Selon cette procédure, est considéré comme central un élément dont la mise en cause provoque un pourcentage de réponses négatives voisin de 100%. Dans le questionnaire, les sujets passent en revue 16 cas correspondant à 16 mises en cause de la représentation (par exemple «Simon exerce une activité professionnelle mais cela ne lui permet pas d'avoir des relations sociales.») Pour chaque cas, les sujets répondent sur un échelle en 4 points, depuis 1 (c'est une situation très vraisemblable) jusqu'à 4 (c'est une situation totalement invraisemblable). De cette manière, un score élevé indique la forte centralité d'un item.

## 7. Les résultats

### 7.1. La centralité du travail

#### 7.1.1. Télétravail et centralité relative

La centralité relative du travail est mesurée, afin d'identifier l'importance accordée au travail par rapport aux autres sphères d'activité, selon la forme d'emploi pratiquée. Les sujets doivent donc répartir 100 points entre les 4 domaines de vie.

CENTRALITE RELATIVE	Moyennes et (écarts-types)			
	TWD n = 26	TA n = 43	TWN n = 58	TC n = 60
Famille	<b>37,77</b>	<b>44,56</b>	<b>39,14</b>	<b>40,88</b>
Travail	<b>31,58</b>	<b>30,16</b>	<b>37,21</b>	<b>30,73</b>
Loisirs	21,96	18,21	16,41	19,00
Activités sociales	8,31	7,19	7,24	9,38
Total	100	100	100	100

Tableau 1

Centralité relative du travail chez les télétravailleurs à domicile (TWD), les travailleurs en mode alterné (TA), les télétravailleurs nomades (TWN) et les travailleurs classiques (TC)

On calcule les moyennes par groupe et on applique l'épreuve du U de Mann et Whitney. Il en résulte quatre relations significatives selon la forme d'emploi pratiquée :

- la centralité relative du travail entre les TWN et les TC (**Z = 3,22 ; p = 0,00**),
- la centralité relative des loisirs entre les TWD et les TWN (**Z = 2,08 ; p = 0,04**) ainsi qu'entre les TWN et les TC (**Z = 1,69 ; p = 0,09**),
- la centralité relative des activités sociales entre les TWN et les TC (**Z = 1,89 ; p = 0,06**).

Conformément à la première hypothèse, les télétravailleurs et les travailleurs classiques accordent une grande importance au travail. Le degré de centralité à l'égard de ce domaine d'activité est significativement plus élevé chez les télétravailleurs nomades que chez les travailleurs classiques.

Contrairement à la première hypothèse, les télétravailleurs nomades accordent de l'importance à la famille. Elle est valorisée par l'ensemble des salariés et bien plus fortement par les travailleurs en mode alterné. Elle est placée largement en tête par rapport au travail chez ce groupe de télétravailleurs ainsi que chez les travailleurs classiques et les télétravailleurs à domicile.

Les loisirs sont moins valorisés par les salariés. Les télétravailleurs nomades y accordent significativement moins d'importance que les télétravailleurs à domicile et les travailleurs classiques. Ainsi qu'il a été supposé, les activités sociales présentent un degré faible de centralité chez l'ensemble des salariés. Les télétravailleurs nomades y accordent significativement moins d'importance que les travailleurs classiques.

### 7.1.2. Télétravail et représentation sociale du travail

Eléments centraux de la représentation sociale du travail	Fréquence de réfutation (FR.)			
	TWN *n = 58	TC *n = 60	TA **n = 45	TWD ***n = 26
Rémunération	<b>0,91</b>	<b>0,87</b>	<b>0,87</b>	<b>0,88</b>
Contraintes	<b>0,83</b>	0,75	0,78	<b>0,85</b>

\*n = 58 & n = 60, dmax = 0,18 ; \*\*n = 45, dmax = 0,20, \*\*\*n = 26, dmax = 0,27

Tableau 2

Représentation sociale (RS) du travail chez les télétravailleurs nomades (TWN), les travailleurs classiques (TC), les travailleurs en mode alterné (TA) et les télétravailleurs à domicile (TWD)

La rémunération constitue un item central chez l'ensemble des salariés. Le test de Kolmogorov-Smirnov confirme la centralité de cet élément. La fréquence des réponses négatives est en effet proche de 100% dans les trois échantillons (rémunération : TWN=0,91, TC=0,87, dmax = 0,18 ; TA=0,87, dmax = 0,20 ; TWD=0,88, dmax = 0,27).

Les contraintes font partie du système central de la représentation sociale du travail des télétravailleurs nomades et des télétravailleurs à domicile. Le test de Kolmogorov-Smirnov confirme la centralité de cet élément. La fréquence des réponses négatives est en effet proche de 100% dans ces deux échantillons (contraintes: TWN = 0,83, dmax = 0,18 ; TWD = 0,85, dmax = 0,27).

La seconde hypothèse est donc en partie vérifiée étant donné que tous les salariés confèrent le statut central à l'élément rémunération.

Eléments périphériques de la RS du travail	Moyennes et (Ecart-types)			
	TWN n = 58	TC n = 60	TA n = 45	TWD n = 26
Manque de temps libre	3,02 (0,87)	2,90 (0,93)	2,93 (1,01)	2,65 (1,06)
Nécessité	2,41 (0,92)	2,38 (0,88)	2,22 (0,88)	2,23 (1,07)
Besoin	2,26 (0,95)	2,49 (0,88)	2,20 (0,89)	2,35 (0,98)
Utilité sociale	2,16 (0,89)	2,00 (0,69)	1,82 (0,78)	2,04 (0,99)
Confiance en soi	1,86 (0,60)	1,95 (0,75)	1,82 (0,65)	1,73 (0,67)
Source d'épanouissement	1,59 (0,59)	1,65 (0,73)	1,53 (0,62)	1,65 (0,84)
Plaisir	1,45 (0,45)	1,42 (0,62)	1,44 (0,62)	1,27 (0,63)

Tableau 3

Eléments périphériques de la représentation sociale (RS) du travail chez les télétravailleurs nomades (TWN), les travailleurs classiques (TC), les travailleurs en mode alterné (TA) et les télétravailleurs à domicile (TWD)

Contrairement à la deuxième hypothèse sur la représentation sociale du travail, il n'y a aucun des éléments périphériques qui soit significativement plus saillant dans un groupe de travailleurs que dans un autre. Les travailleurs classiques et les télétravailleurs à domicile ne sont pas plus sensibles aux aspects contraignants du travail (nécessité, besoin, manque de temps libre). Les travailleurs en mode alterné et les télétravailleurs nomades ne sont pas davantage attachés aux aspects positifs du travail (épanouissement, plaisir, utilité sociale, confiance en soi).

### 7.1.3. Télétravail et rapports vie de travail - vie hors travail

Les travailleurs classiques se distinguent des deux autres catégories de télétravailleurs pour gérer leur système des activités, ils segmentent les domaines de vie alors que les travailleurs en mode alterné et les télétravailleurs à domicile assurent une activation des échanges entre les différents domaines. En effet, d'après les rapports qui s'établissent entre le domaine de vie familial d'une part, et le domaine de vie professionnel, personnel et social d'autre part, les travailleurs classiques séparent leur vie familiale des autres sphères de vie (**F/PS**<sup>4</sup> : **TC** : Fr. Ctrtes<sup>5</sup> = 0,24 ; **Fr. SR**<sup>6</sup> = **0,51**, Fr. Rscs<sup>7</sup> = 0,25) alors que les travailleurs en mode alterné estiment que le domaine de vie familial apporte également des ressources aux autres domaines (**F/PS** : **TA**: Fr. Ctrtes = 0,18 ; Fr. SR = 0,39 ; **Fr. Rscs** = **0,43**) ; (F/PS : TC – TA ; ddl = 2 ;  $\chi^2 = 7,85$  ; **p < 0,02**).

On obtient des résultats similaires d'après les rapports qui s'instaurent entre le domaine de vie professionnel d'une part, et les domaines de vie familial, personnel et social d'autre part. Les travailleurs classiques séparent aussi ces domaines (**P/FS**<sup>8</sup> : **TC** : Fr. Ctrtes = 0,26 ; **Fr. SR** = **0,37** ; **Fr. Rscs** = **0,37**), alors que les télétravailleurs à domicile établissent essentiellement des relations d'aide (**P/FS** : **TWD** : Fr. Ctrtes = 0,10 ; Fr. SR = 0,33 ; **Fr. Rscs** = **0,58**) ; (P/FS = TC - TWD : ddl = 2 ;  $\chi^2 = 7,60$  ; **p < 0,05**).

Concernant les télétravailleurs nomades, les résultats sont plutôt mitigés. Ce groupe de salariés segmente tout autant qu'il concilie ses activités. Sa stratégie dépend en effet des domaines de vie considérés et du sens des échanges (**F/PS** : Fr. Ctrtes = 0,23 ; **Fr. SR** = **0,42** ; Fr. Rscs = 0,34 et **PS/F**<sup>9</sup> : Fr. Ctrtes = 0,17 ; Fr. SR = 0,35 ; **Fr. Rscs** = **0,47**).

La dernière hypothèse est vérifiée d'après les stratégies mises en œuvre par les travailleurs classiques, les travailleurs en mode alterné et les télétravailleurs à domicile, pour gérer leur système des activités. En revanche, un résultat plus nuancé est à constater chez les télétravailleurs nomades : ils sont amenés à séparer tout comme à concilier leurs sphères de vie.

---

<sup>4</sup> Domaine de vie familial (F) vis-à-vis des domaines de vie professionnel (P), personnel et social (S)

<sup>5</sup> Contraintes

<sup>6</sup> Sans rapport

<sup>7</sup> Ressources

<sup>8</sup> Domaine de vie professionnel (P) vis-à-vis des domaines de vie familial (F), personnel et social (S)

<sup>9</sup> Domaines de vie professionnel (P), personnel et social (S) vis-à-vis du domaines de vie familial (F)

## **8. Discussion et conclusion**

### 8.1. Télétravail et centralité relative du travail

Le travail est central chez l'ensemble des salariés mais cette centralité est significativement plus forte chez les télétravailleurs nomades par comparaison aux travailleurs classiques. En effet, ce groupe de télétravailleurs considère le nomadisme comme une évolution positive de leur fonction, donc ils seront d'autant plus attachés à leur travail.

L'importance accordée à la famille est élevée chez l'ensemble des salariés. Elle est plus forte chez les travailleurs en mode alterné et elle arrive nettement en tête par rapport au travail chez les travailleurs classiques et les télétravailleurs à domicile.

Les loisirs arrivent en troisième position chez l'ensemble des salariés. Ils accordent une moindre valeur à ce domaine d'activité ainsi qu'aux activités sociales qu'ils négligent très largement. Une centralité élevée dans la famille et dans le travail induiraient-elles une faible centralité dans les loisirs et dans les activités sociales et inversement ? (Image des vases communicants).

La classification des domaines de vie dans chaque forme d'emploi autorise les remarques suivantes : les travailleurs qui exercent leur activité professionnelle, toute ou en partie à domicile, valorisent davantage le domaine pour lequel ils ont choisi de pratiquer le télétravail. Les télétravailleurs nomades qui ont consenti, et non choisi de télétravailler, rééquilibrent le niveau de centralité entre le domaine de vie professionnel et le domaine de vie familial. Enfin les travailleurs classiques attribuent de l'importance au domaine de vie le plus perturbé ; d'après les entretiens, il s'agit du domaine de vie familial en raison du temps qui ne peut lui être octroyé.

Famille et travail constituent les deux sphères de vie principales, loisirs et activités sociales sont loin de les supplanter. La valeur du travail ne serait donc pas en voie de disparition. Travailler autrement ne semble pas traduire l'engouement pour une révolution culturelle du temps libre (Méda, 1995).

### 8.2. Télétravail et représentation sociale du travail

L'ensemble des personnes interrogées considère que le travail n'est pas concevable sans la notion de rémunération, résultat classique qui corrobore ceux obtenus dans un certain nombre d'études sur la représentation sociale du travail (Flament, 1994).

Louche et Moliner (2001) constatent que les travailleurs classiques sont plus sensibles aux aspects ludiques et contraignants du travail alors que les télétravailleurs sont plus centrés sur les aspects valorisants du travail. On ne relève pas ces différences concernant les éléments périphériques de la représentation sociale du travail dans cette étude. Toutefois, il est à souligner que l'élément «contraintes» appartient à des

composantes différentes de la représentation sociale du travail en fonction de la forme de travail considérée. Ainsi il fait partie du système central de la représentation sociale du travail des télétravailleurs nomades et des télétravailleurs à domicile, contrairement aux travailleurs en mode alterné et aux travailleurs classiques qui intègrent cet élément dans leur système périphérique. La forme d'emploi selon le type d'organisation où elle est exercée, n'induit pas la même représentation sociale du travail. Les télétravailleurs nomades et les télétravailleurs à domicile sont plus sensibles aux aspects contraignants du travail, les travailleurs en mode alterné et les travailleurs classiques sont plus centrés sur les aspects valorisants, liés au mode d'exécution de leur travail. Par conséquent, la forme de travail influence la représentation sociale du travail puisque d'après la théorie du noyau central (Abric, 1987), deux représentations sont différentes si leurs éléments centraux sont différents.

### 8.3. Télétravail et échanges

Les travailleurs classiques se distinguent des deux autres types de télétravailleurs pour gérer leur système des activités, ils segmentent les domaines de vie alors que les travailleurs en mode alterné et les télétravailleurs à domicile assurent une activation des échanges entre les domaines. La segmentation des sous-systèmes apparaît dans la forme classique de travail qui établit une séparation entre les sphères de vie, alors que la conciliation des sous-systèmes se produit dans des formes de télétravail qui favorisent l'imbrication des sphères de vie.

Les télétravailleurs nomades se distinguent des autres travailleurs. Ils sont autant amenés à développer des échanges d'aide entre les domaines qu'à établir des séparations. Probablement, le mode particulier du nomadisme qui permet de travailler n'importe où n'importe quand, modifie différemment les pratiques entre la vie de travail et la vie hors travail. En effet, il ne conduit pas toujours à un rapprochement des espaces de vie comme le télétravail à domicile et le travail en alternance domicile-travail, c'est à dire à un nécessaire échange entre les domaines.

Les résultats de cette recherche incitent à la répliquer mais selon une nouvelle perspective d'approche. Il serait en effet intéressant d'effectuer un travail longitudinal, de façon à recueillir les données en aval et en amont de la mise en place du télétravail, pour répondre à la question qui suit : le fonctionnement du système des activités des salariés induit-il la pratique du télétravail ? Ou bien la pratique du télétravail détermine-t-elle le mode de fonctionnement du système des activités des salariés ?

### **Références**

- Abric, J.-C. (1987). *Coopération, compétition et représentations sociales*. Cousset : Delval.  
Abric, J.-C. (1994). Les représentations sociales : aspects théoriques.  
In J.-C.Abric (Ed.), *Pratiques sociales et représentations* (chap.1, pp.11-36). Paris : Presses Universitaires de France.

- Association Française du Télétravail et des Téléactivités. (s.d.). *Accord cadre sur le télétravail*. Pages consultées, le 16 août, 2002  
[http://europa.eu.int/comm/employment\\_social/news/2002/jul/145\\_fr.html](http://europa.eu.int/comm/employment_social/news/2002/jul/145_fr.html)
- Baubion-Broye, A. (1998). *Événements de vie, transitions et construction de la personne*. Saint Agne : Erès.
- Breton, T. (1994). *Le télétravail en France : situation actuelle, perspectives de développement et aspects juridiques*. Paris : La Documentation Française.
- Camps, E. (2006). *Télétravail : sens du travail, rapports vie de travail - vie hors travail*. Thèse de doctorat en psychologie, Université Paul Valéry, Montpellier III.
- Curie, J. (2000). *Travail, personnalisation, changements sociaux*. Toulouse : Octares.
- Curie, J., Hajjar, V. (1987). Vie de travail, vie hors travail : la vie en temps partagé. In C. Lévy-Leboyer & J.-C. Sperandio (Eds.), *Traité de psychologie du travail* (chap. 3, pp.37-55). Paris : Presses Universitaires de France.
- Curie, J., Hajjar, V., Marquie, H., & Roques, M. (1990). Proposition méthodologique pour la description du système des activités. *Le travail humain*, 53(2), 103-118.
- Flament, C. (1994). Le plaisir et la rémunération dans la représentation du travail. *Cahiers Internationaux de Psychologie Sociale*, 23, 61-69.
- Harpaz, I. (1985). Meaning of Working Profiles of Various Occupational Groups. *Journal of Vocational Behavior*, 26, 25-40.
- Jodelet, D. (1989). Représentations sociales : un domaine en expansion. In D. Jodelet (Ed.), *Les représentations sociales* (pp.31-61). Paris : Presses Universitaires de France.
- Joulain, M. (2006). *Approche psychosociale du rapport au travail*. Habilitation à diriger des recherches, Université de Tours.
- Kouloumdjian, M.-F. (2000). Travail à distance. In J.-C. Bernaud & C. Lemoine (Eds.), *Traité de Psychologie du Travail & des Organisations* (chap. 10, pp.345-386). Paris : Dunod.
- Lancry, A., Baugnet, L., Come, K., & Cottinet, A. (2000). Télétravail salarié à domicile et identité professionnelle. *Psychologie du Travail et des Organisations*, 6(1-2), 163-183.
- Louche, Cl., & Moliner, P. (2001). Sens et représentations du travail chez des télétravailleurs et des travailleurs classiques. *Cahiers Internationaux de Psychologie sociale*, 50, 70-79.
- Méda, D. (1995). *Le travail. Une valeur en voie de disparition*. Paris : Aubier.
- Megemont, J.-L. (2000). *Processus de socialisation et de construction de l'identité dans une situation de mobilité professionnelle : le cas des élèves ingénieurs du Centre d'Etudes Supérieures Industrielles*. Thèse de doctorat en psychologie, Université de Toulouse, le Mirail.
- Moliner, P. (1994). Les méthodes de repérage et d'identification du noyau des représentations. In C. Guimelli (Ed.), *Structures et transformations des représentations sociales*.(ch.7, 199-232). Lausanne, Delachaux et Niestlé.
- Morin, E. M. (1997). Le sens du travail pour des gestionnaires francophones. *Psychologie du Travail et des Organisations*, 3(3-4), 26-45.
- MOW International Research Team (1987). *The meaning of working*. New York: Academic Press.
- Salmaso, P., & Pombéni, L. (1986). Le concept de travail. In W. Doise & A. Palmonari (Eds.), *L'étude des représentations sociales* (pp.196-208). Neuchâtel : Delachaux et Niestlé.

## **Note importante aux auteurs**

Les articles proposés qui présentent une recherche empirique, un travail de synthèse, une revue critique de la littérature... doivent suivre les normes de publication de la revue. Les travaux doivent satisfaire aux critères de vérification scientifique et être présentés dans un langage clair.

Deux évaluateurs en assurent l'expertise, à l'issue de laquelle ils sont acceptés, avec ou sans modifications, ou refusés.

**La présentation de l'article doit se conformer en tous points aux recommandations suivantes. En cas de non conformité de la présentation le document sera retourné aux auteurs.**

**Le texte est adressé par mail, en fichier Word (selon les consignes de mise en page indiquées) à la rédaction :**

**Rédacteur en Chef** : Georges MASCLET  
572 rue Longue  
59450 Sin le Noble

Tél. Domicile : 03-27-98-78-26 et 06-60-58-78-26  
Email : g.masklet@revue-pto.com *ou* GMasc@aol.com

**Assistante de rédaction** : Lysiane MASCLET  
Email : l.masklet@revue-pto.com *ou* lisedesin@aol.com

N'hésitez pas à consulter également le site de la revue sur:

[www.revue-pto.com](http://www.revue-pto.com)

**Dans le souci d'unifier la présentation  
des articles soumis pour publication et de répondre  
aux règles de bibliographie, nous portons à votre attention  
les points suivants :**

Votre article doit être envoyé au format de la revue :

- fichier
- mise en page
- papier
- taille personnalisée (largeur : 16.5 ; hauteur 24)

**1<sup>ère</sup> page**

**Titre (14 Times new roman Gras ; en français)**

**Titre (14 Times new roman Gras ; en anglais)**

Prénom Nom\* & Prénom Nom \*\* (11 Times new roman)

\*Université de ..., Laboratoire..., Adresse, mail

\*\*Université de ..., Laboratoire..., Adresse, mail

(10 Times new roman)

*Résumé*

Le résumé s'écrit en 10 Times new roman, interligne simple. Il ne dépasse pas 10 lignes ou 120 mots. Le titre «Résumé» est en Italique.

Le résumé indique la question traitée, son but, la référence théorique, donne le genre de méthodologie utilisée, indique le contexte et éventuellement la population, le résultat principal attendu ou obtenu.

*Abstract*

Le résumé doit être traduit en anglais

Mots-clés : Pas plus de 5, sur une ligne.

Key-words : traduction des mots clefs

## **A partir de la 2ème page**

### **1. Consigne concernant le texte**

Le texte comprend une introduction, des informations sur les hypothèses, la méthode, les résultats, la discussion, une conclusion.

La présentation fait apparaître clairement le plan (parties, sous parties).

Vérifiez l'orthographe et la grammaire.

Evitez les paragraphes trop longs.

### **2. Consigne concernant la mise en page**

**Cliquer sur «mise en page» dans «fichier». Pour obtenir un texte sur 20 cm de haut maximum et 11,8 cm de large, il faut mettre : Haut : 3 cm ; Bas : 5,8 cm ; Gauche : 6,2 cm ; Droite : 3 cm**

#### 2.1 Les titres

Voir l'exemple suivant:

### **1. Titre niveau 1 (en 11 Times new roman gras)**

#### 1.1 Titre niveau 2 (en 11 Times new roman ; retrait 1,25cm)

##### 1.1.1 Titre niveau 3 (en 11 Times new roman ; retrait 2,5cm)

#### 2.2 Le texte

Le texte est écrit en 11 Times new roman, interligne simple. Tout le texte doit être traité en « justifié », comme ici.

La première ligne d'un paragraphe commence avec un retrait de 1.25 cm.

Tout paragraphe est suivi d'une ligne en blanc (et une seule). Il y a toujours une ligne en blanc avant un titre (et une seule). Il n'y a pas de point à la fin d'un titre.

#### 2.3 Les tableaux et graphiques

Un tableau doit être lisible et entrer dans la largeur de la page. Il est placé dans la page sans déborder sur la suivante. Il ne doit pas être placé à la fin du document écrit. Il se place dans le sens de la lecture

Il comprend une légende placée en dessous et un n° : tableau 1, par exemple. Les chiffres et les indications sont en 10 Times new roman. On peut contraster gras et non gras (par exemple texte en gras, chiffres non gras).

Si les tableaux sont en couleurs ils seront à la charge de l'auteur.

### **3. Consigne concernant la bibliographie**

La bibliographie est rédigée en 10 Times new roman

Les auteurs cités en bibliographie doivent impérativement être des auteurs cités dans le texte.

Les normes sont celles de l'APA : le titre d'un ouvrage est en italique, de même que le nom d'une revue.

Suivre les exemples suivants :

#### ***Références***

Di Fabio, A., Majer, V. & Taralla, B. (2006). Corrélatifs de la Teacher Self-Efficacy : caractéristiques personnelles et attitude envers le travail. *Psychologie du travail et des organisations*, 12, n°4, 263-277

Pithon, G. & Gangloff, B. (2005). *Evaluer pour former, orienter et apprécier le personnel*. Paris, L'Harmattan

Savoie, A. & Brunet, L. (2000). Les équipes de travail. In : Bernaud, J.L. & Lemoine, C. (Eds). *Traité de psychologie du travail et des organisations*. Paris, Dunod, 171-202

**L'ensemble ne doit pas dépasser 30 000 signes environ.**

