UNIVERSITÉ FRANÇOIS - RABELAIS

DE TOURS

ÉCOLE DOCTORALE Santé, Sciences et Technologies

EA 2114 Psychologie des Ages de la Vie

THÈSE présentée par :

Olivier SOREL

soutenue le : **04 décembre 2009**

pour obtenir le grade de : **Docteur de l'université François - Rabelais**

Discipline/ Spécialité : PSYCHOLOGIE

APPROCHE DEVELOPPEMENTALE DU RAISONNEMENT BAYESIEN

-Analyse quantitative et qualitative selon le format de présentation et le niveau scolaire-

THÈSE dirigée par :

Mme PENNEQUIN Valérie	Professeure, Université F. - Rabelais Tours
M. FONTAINE Roger	Professeur, Université F. - Rabelais Tours

RAPPORTEURS :

M. MOUTIER Sylvain	Professeur, Université Paris Descartes
M. VALLEE-TOURANGEAU Frédéric	External examiner, Université de Kingston

JURY :

Mme BLAYE Agnès	Professeure, Université de Provence Aix en Provence
M. FONTAINE Roger	Professeur, Université F. - Rabelais Tours
M. MOUTIER Sylvain	Professeur, Université Paris Descartes
Mme PENNEQUIN Valérie	Professeure, Université F. - Rabelais Tours
M. VALLEE-TOURANGEAU Frédéric	External examiner, Université de Kingston

Je dédie ce travail de recherche à ma mère qui nous a quittés pendant sa réalisation.

Table des Matières

Remerciements

J’adresse en premier lieu mes remerciements à Valérie Pennequin pour son encadrement depuis mon mémoire de Maîtrise, sa disponibilité, sa rigueur et son implication dans ce travail de recherche. Je la remercie de m’avoir soutenu et encouragé dans les périodes de doutes.

Je remercie également Roger Fontaine pour la qualité scientifique de sa direction et pour les longues conversations qui ont bien souvent fait naître des questions que je n’avais pas à l’idée.

Je remercie les membres du laboratoire -permanents et doctorants- pour leur sympathie et leurs encouragements.

Je remercie tous les participants ayant donné de leur temps et permis la réalisation de ces travaux. Je suis reconnaissant aux parents d’élèves d’avoir donné leur accord concernant la participation de leurs enfants. Je salue leur accueil chaleureux et leur bonne humeur.

Une pensée particulière pour toutes les personnes proches de moi qui m’ont soutenu, encouragé, et conseillé. Merci d’avoir été selon les moments source de motivation, d’échanges ou partenaires sportifs ou échiquéens pour me défouler.

Un grand merci à ma famille qui a toujours été là en cas de besoin, notamment dans des moments difficiles que nous avons traversés ensemble. Un merci particulier à Audrey.

Un merci complice à Lucas qui m’a laissé dormir (un peu) les nuits qui précédèrent la soutenance…

Je remercie enfin les membres du jury pour l’intérêt qu’ils ont porté à mon travail.

Résumé

Emettre des inférences bayésiennes, c'est-à-dire réviser son jugement quant à l'apparition d'un événement, et passer d'une probabilité *a priori* à une probabilité *a posteriori* est une activité quotidienne qui, en théorie requiert la formule de Bayes. Toutefois, en pratique, l'être humain est sensible à différents biais et n'utilise pas toujours à bon escient les informations dont il dispose. Zhu & Gigerenzer (2006) ont montré qu'un contexte en fréquences naturelles favorisait davantage une révision de jugement qu'un contexte présenté en probabilités conditionnelles, les fréquences rendant explicite le taux de base (Hoffrage, Gigerenzer, Krauss & Martignon, 2002).

Le but de ce travail est de préciser dans quelle mesure le format de présentation et les exigences quant au jugement attendu influencent les performances bayésiennes de collégiens et lycéens français. Des cohortes de 20 participants de classe de sixième à première ont été soumises à des problèmes bayésiens. Les trois premières expériences font de façon progressive abstraction du nombre. Les résultats confirment l'effet facilitateur des fréquences sur les probabilités. L'analyse des stratégies utilisées par les participants suggère qu'il ne faut pas se contenter d'une cotation dichotomique : réponse bayésienne versus non bayésienne. En effet, nos résultats précisent qu'avec l'avancée scolaire les participants commettent des erreurs quantitatives de moins en moins éloignées de la réponse théorique, et estiment qualitativement de plus en plus finement l'occurrence du dit événement.

La quatrième expérience tente de faire le lien entre les fonctions exécutives de bas niveau, la vitesse de traitement et le niveau scolaire en mathématiques, d'une part, et les performances bayésiennes, d'autre part. Les résultats montrent que la vitesse de traitement et l'inhibition prédisent modérément les performances bayésiennes des collégiens, mais pas celles des lycéens.

La dernière expérience est une analyse longitudinale des performances des participants testés à dix-neuf mois d'intervalle. Les résultats étayent ceux de l'analyse transversale, puisque la majorité des participants présente des performances accrues.

Mots clés : inférences bayésiennes, fréquences naturelles, probabilités conditionnelles, raisonnement probabiliste.

Résumé en anglais

Emit bayesian inferences, in other words revise his judgment about the appearance of an event, and change from prior probability to posterior probability, is a daily activity which, in theory, requires the Bayes' rule. However, in practice, human is sensible to different bias and he doesn't use systematically wisely all information at his disposal. Zhu & Gigerenzer (2006) showed that when data are presented in natural frequencies, bayesian performances increase. The reason is that base rate information is contained in natural frequencies (Hoffrage, Gigerenzer, Krauss, & Martignon, 2002).

The aim of this work is to specify in what measure the format of presentation and the experimenter's request about the judgment to product influence the bayesian performances of French schoolchild from the beginning to the end of secondary school. Groups of 20 participants from sixth grader to eleventh grade were tested on bayesian problems. The first three experiments progressively disregard number. Results confirm the easier effect of frequencies on the probabilities. The analyse of strategies used by participants indicates that we don't have to be content with dichotomous quotation: bayesian versus no bayesian answer. In fact, our results specify that with the school advance participants make quantitative mistakes which are less and less distant of the theoretic answer, and estimate qualitatively more and more precisely the event occurrence.

The fourth experience tries to make the link between the executive functions of low level, the processing speed and the school level in mathematical, on the one hand, and the bayesian performances, on the other hand. Results show that processing speed and inhibition predict moderately bayesian performances during the first middle of secondary school, but don't during the second.

The last experience is a longitudinal analyse of participants' performances tested nineteen months later. Results support the transversal analyse ones, since the majority of participants produce better performances.

Key words: bayesian inferences, natural frequencies, conditional probabilities, probabilistic reasoning.

Introduction

Si l'on peut admettre aisément que certaines espèces animales font montre de raisonnement, il semble tout naturel de l'associer à l'homme. En effet, le raisonnement est une activité quotidienne qui permet à l'être humain d'aller au-delà de l'information immédiatement disponible. Bruner (1974) popularise l'idée selon laquelle l'homme, à partir de données -perceptives ou verbales- est capable de faire des inférences, en d'autres termes d'aboutir à une information qui n'était pas immédiatement mobilisable, ni explicitement présente, voire même, dans certains cas, totalement absente. Le raisonnement prend alors la forme d'un processus cognitif permettant de produire une conclusion, qui se veut nouvelle, d'après des propositions initiales, des prémisses.

Cette idée n'est pourtant pas si nouvelle, puisque avant le psychologue, le philosophe s'est déjà intéressé à cette question, au travers de réflexions sur la logique et sur la rhétorique. Dans un premier chapitre, nous présenterons donc des concepts proches de celui de raisonnement, telle la logique formelle. Ceci nous permettra de comparer un raisonnement logique, valide, à un raisonnement psycho-logique qui peut être soumis à différents biais. Nous présenterons également une taxonomie afin de lister et de définir les différents types de raisonnement que nous pouvons mettre en œuvre. Ainsi, nous aborderons les différences entre le raisonnement hypothético-déductif, plus connu sous le raccourci de raisonnement déductif et le raisonnement inductif, en nous attardant sur le raisonnement probabiliste.

Dans un second chapitre, nous détaillerons davantage le raisonnement probabiliste qui nous intéresse tout particulièrement ici. Nous présenterons les travaux de Thomas Bayes et son apport considérable au vaste domaine des probabilités, tant en économie, en mathématiques, en informatique, en médecine et en psychologie cognitive. En effet, une inférence bayésienne permet de réviser son jugement et ainsi de passer d'une probabilité *a priori* à une probabilité *a posteriori*, grâce à l'intégration d'un nouvel élément ou d'une nouvelle information dans notre jugement, notre raisonnement. Ces notions explicitées, nous ferons le point sur les modèles existants et aborderons l'analyse fréquentiste et la rationalité écologique. Nous verrons l'effet facilitateur des fréquences naturelles par rapport aux probabilités conditionnelles (Hoffrage, Gigerenzer, Krauss, & Martignon, 2002). Ensuite, ce chapitre fera le bilan des connaissances que nous avons quant au développement des habiletés cognitives précoces des enfants et leurs capacités à inférer et réviser leur jugement (Zhu & Gigerenzer, 2006). Enfin, nous aborderons le raisonnement selon deux modes -logique et analytique versus heuristique et intuitif-. Le développement du raisonnement probabiliste sera donc appréhendé d'une manière structuraliste et d'une manière fonctionnelle.

Les apports de ces travaux nous permettront de présenter la problématique de notre recherche dans le chapitre 3. Nous proposerons une analyse détaillée des performances bayésiennes de participants français scolarisés de la classe de sixième à la classe de première. Plus précisément, nous testerons les réponses bayésiennes quantitatives, qualitatives et heuristiques des participants selon le format et leur classe.

Les chapitres 4 à 8 traiteront respectivement des cinq expérimentations menées durant ce travail.

Au cours de la première expérimentation (chapitre 4) les problèmes présentés aux participants sont quantitatifs et sous deux formats : en fréquences naturelles et en probabilités conditionnelles. Les réponses attendues de la part des participants sont également quantitatives. Les résultats confirment l'effet facilitateur du format fréquentiste sur le format probabiliste et permettent une analyse des stratégies utilisées par les participants. Nous discuterons la notion de progrès et de développement des performances d'un point de vue qualitatif selon les stratégies utilisées.

Pour la seconde expérimentation (chapitre 5), des participants équivalents doivent répondre aux problèmes précédemment cités mais leur réponse est qualitative, sous la forme d'une croix sur un continuum libre. Là encore, les fréquences amènent de meilleures performances. De même, ces performances intuitives évoluent au cours des années de collège et de lycée.

Allant plus loin dans l'abstraction du nombre, la troisième expérience (chapitre 6) propose aux participants des problèmes avec des données qualitatives (peu versus beaucoup pour chaque donnée pertinente). Les participants doivent à nouveau estimer leur réponse sur un continuum libre. Un modèle factoriel permet de rendre compte de l'utilisation faite des différentes données présentées dans le contexte. Nos résultats sont discutés autour de la notion d'algèbre cognitive.

Au regard des précédents résultats qui montrent la difficulté intrinsèque de la formule mathématique de Bayes, la quatrième expérimentation (chapitre 7) vise à savoir si le niveau en mathématiques des participants et les fonctions exécutives de bas niveau -flexibilité, inhibition et remise à jour en mémoire de travail- (Miyake, Friedman, Emerson, Witzki, & Howerter, 2000) et la vitesse de traitement sont des prédicteurs fiables de leurs performances bayésiennes. La discussion de ces résultats s'articule autour de la théorie du double processus et des approches structuraliste et fonctionnelle du développement cognitif de l'enfant et de l'adolescent.

La cinquième expérimentation (chapitre 8) est une étude longitudinale. Elle se déroule avec les mêmes participants, dix-neuf mois après la première expérience. Elle a pour but d'affiner les résultats obtenus lors des deux premières expériences transversales.

Le chapitre 9 conclura notre travail en synthétisant les résultats obtenus au cours des cinq expérimentations. Ce sera l'occasion de discuter ces résultats et leurs éventuelles applications à la lumière des travaux antérieurs ayant conduit à l'élaboration de cette thèse. La seconde partie de ce chapitre nous permettra d'évoquer de possibles perspectives de recherches sur le raisonnement bayésien.

Première partie

Partie théorique

Chapitre 1 - De la logique formelle au(x) raisonnement(s)

La logique, inventée par Aristote il y a plus de deux mille ans, a beaucoup évolué. D’abord liée au langage, elle est devenue, au fil des siècles, sœur siamoise des mathématiques modernes.

I. La dialectique et la rhétorique

La forme première de logique est la dialectique grecque. Deux interlocuteurs dialoguent sous la forme d’un échange au cours duquel se succèdent questions et réponses. Prenons pour illustrer ces échanges *le Cornu* du philosophe Diodore Cronos :

> « Tu as ce que tu n’as pas perdu ?
> - Oui.
> - Mais as-tu perdu des cornes ?
> - Non !
> - Donc tu as des cornes ? »

Répondre respectivement oui et non à la première et la deuxième question amène à répondre oui à la troisième question, d’un point de vue logique.

Si la dialectique est une forme de discours et de raisonnement, il en est de même pour la rhétorique, son double. On parle de rhétorique lorsque le discours est continu et sous forme de monologue, comme c’est le cas avec le sophisme du *menteur* qui suit :

> « Si tu dis que tu mens et que ce que tu dis est vrai, alors tu mens et tu dis vrai. »

Les termes de sophisme et de paralogisme sont utilisés pour rendre compte d’un raisonnement fallacieux, qui, en apparence, semble être un raisonnement valide. Lorsque le raisonnement n’est pas valide sciemment, avec la volonté de tromper, on préfère le terme de sophisme. Celui de paralogisme est utilisé quand la personne commettant un raisonnement invalide ne le fait pas volontairement.

Le mot « logique » fut inventé plus tard par un autre philosophe, Xénocrate. Il provient de l’adjectif grec logikos, logikê au féminin, dérivé de logos, qui signifie à la fois raison, langage et raisonnement. D’après Xénocrate, est donc logique ce qui est rationnel, ce qui relève du langage ou ce qui est raisonné.

Figure 1.1. Typus Logice de Gregor Reisch

Dans le *Typus Logice* de Gregor Reisch (figure 1.1), issu de son ouvrage Margarita Phylosophica, que nous pourrions qualifier d’encyclopédie, la *logique* est personnifiée, armée de son épée *syllogimus*, et chassant le lièvre *problema* avec ses deux chiens *veritas* et *falsitas*. Cette esquisse allégorique met en avant un outil au service de la logique : le syllogisme.

II. Les syllogismes

Quand Aristote développe la notion de syllogisme, il la définit comme « un raisonnement dans lequel, certaines choses étant posées ; quelque chose d’autre que ce qui a été avancé résulte nécessairement de ce qui a été avancé » (Aristote, 1983). En ce sens, Aristote semble vouloir proposer un mode de déduction, un système logique, qui s'appuie sur une structure bien particulière. Blanché (1973) et Lear (1980) précisent qu'Aristote a défini le syllogisme avant ses *Analytiques*, mais que dans ces ouvrages, il a défini à nouveau ce concept, le considérant comme l'essence même de la démarche déductive. Dans un syllogisme, on passe des deux premières propositions à la troisième. En d'autres termes, la conclusion se déduit des deux prémisses. Notre propos n'est pas ici de détailler ce que le psychologue et le logicien appellent aujourd'hui le raisonnement catégorique, à savoir l'étude des différents types de syllogismes. Toutefois, nous allons en présenter quatre afin de définir des concepts qui nous seront utiles pour la lecture du chapitre 2 et de la suite de ces travaux.

Tous les hommes sont mortels.
Socrate est un homme.
Donc Socrate est mortel.

Ce premier syllogisme, qu'il n'est probablement plus nécessaire de présenter, tant il est connu, est attribué à Guillaume d'Occam. Il est depuis passé dans les mains des logiciens et psychologues. Binet (1902) l'a utilisé comme outil de sélection dans le domaine scolaire. Pour lui, ce syllogisme est un modèle tant descriptif et explicatif, que normatif. Pour ce chercheur, la possibilité de passer des deux prémisses à la conclusion certaine que Socrate est mortel permet d'appréhender l'essence même de tout raisonnement. Plus tard, cet exemple est repris dans de nombreuses études, à la fois de logique formelle et d'analyse philosophique des travaux d'Aristote (Lukasiewick, 1972), et de recherches en psychologie cognitive portant sur le raisonnement (Oléron, 1989).

Si tu n'es pas avec moi, alors tu es contre moi.
Tu n'es pas avec moi.
Donc tu es contre moi.

Avec ce deuxième syllogisme, nous retrouvons la notion de *sophisme*. En effet, si la structure en elle-même du syllogisme respecte les règles de logique, de déduction au sens aristotélicien du terme, la première prémisse est piège et laisse apparaître l'envie de manipuler l'interlocuteur. Présenté de la sorte, cette phrase laisse penser qu'il n'y a pas d'alternative à la dichotomie avec/contre moi. Or, il y a évidemment la possibilité d'être neutre, c'est-à-dire ni avec ni contre.

Tous les chats sont mortels.
Socrate est mortel.
Donc Socrate est un chat.

Ce troisième syllogisme illustre la notion évoquée précédemment de paralogisme. On peut penser, en effet, que la personne qui donne la conclusion *Socrate est un chat* a commis un biais de raisonnement puisqu'il n'est pas admis de remonter du particulier au général. L'erreur commise ici est de considérer comme équivalent *Tous les chats sont mortels* et *Tous les mortels sont des chats* (Chapman & Chapman, 1959 ; Revlis, 1975).

III. Validité d'un raisonnement et véracité d'une proposition

Tout ce qui est rare est cher.
Un cheval bon marché est rare.
Donc un cheval bon marché est cher.

Ce quatrième, et dernier, syllogisme présenté ici nous amène à définir deux notions : la validité et la véracité. Ce syllogisme est valide mais présente toutefois une conclusion fausse. Nous pourrions même préciser qu'elle est illogique, dans le sens où *bon marché* et *cher* sont tous les deux antinomiques. Pour Politzer & Braine (1991) un argument valide peut avoir des prémisses vraies ou fausses. Ce qui compte c'est le passage des deux prémisses à la conclusion. En d'autres termes, la validité d'un raisonnement, d'un syllogisme par exemple, dépend du respect des règles de la logique que nous appellerons formelle. Gochet & Gribomont (1990) précisent qu'un raisonnement vrai, valide, peut avoir une conclusion fausse. La sémantique utilisée est donc très importante dans l'étude du raisonnement (Begg & Harris, 1982 ; Newstead, 1989).

IV. La rationalité humaine

La présentation de ces deux notions -validité et véracité- nous invite à évoquer la notion de rationalité humaine. Quand le logicien considère l'erreur comme anecdotique , accidentelle et ne remettant pas en cause le concept de *logique humaine*, le psychologue cognitiviste préfère, quant à lui, étudier l'erreur, le biais de raisonnement (Sells, 1936 ; Wilkins, 1928 ; Woodworth & Sells, 1935). Toutefois, ce courant de recherche a surtout connu un réel développement dans les années 70 et 80, sous l'impulsion des recherches menées dans le cadre du *programme des heuristiques et des biais* (Tversky & Kahneman, 1974 ; Kahneman, Slovic, & Tversky, 1982), que nous détaillerons dans le chapitre 2. Pour Oléron (1989, p.5) « L'étude de la validité du raisonnement est l'affaire de la logique. […] La psychologie, au contraire, s'intéresse aux raisonnements tels qu'ils se déroulent, sans se préoccuper de telles règles, les raisonnements faux ayant pour elle autant d'intérêt que les autres. ». Si l'intérêt pour l'erreur est partagé par les chercheurs en psychologie, son incidence quant à la rationalité et la logique humaine est non consensuelle. Ainsi, certains auteurs considèrent l'être humain comme non logique, puisque son comportement est déterminé par des biais (Sells, 1936). Johnson-Laird & Byrne, (1991) défendent l'idée selon laquelle malgré l'utilisation de raisonnements n'obéissant pas aux règles de la logique classique, les individus sont rationnels.

Evans (1989) adopte une position plus nuancée pensant que le comportement humain est complexe, et régi, à la fois par des processus logiques et non logiques. Evans & Over (1997, p.34) plaident en la faveur d'une rationalité humaine à deux niveaux : « Il est généralement admis que les individus sont dans une large mesure rationnels quand ils doivent accomplir un but personnel, mais qu'ils ont une rationalité limitée à raisonner ou à agir suivant des raisons sous-tendues par un modèle normatif. ». De ce point de vue, l'homme aurait plusieurs types de raisonnement.

V. Vers différentes catégories de raisonnement

Les syllogismes ont été les premiers outils expérimentaux pour appréhender le raisonnement (Störring, 1908 ; Thorndike, 1922 ; Wilkins, 1928), dans le sens où « le syllogisme est un schème de raisonnement […] » (Piaget, 1972, p.359). Mais de quel raisonnement parle-t-on ici ? Le raisonnement qui est l'objet d'étude du psychologue, n'est pas tant l'enchaînement logique des prémisses à une conclusion (Dowek, 1995) -comme décrit précédemment-, que l'ensemble des processus cognitifs mis en œuvre pendant ces activités de logique (Oléron, 1989).

Plutôt que de traiter le raisonnement de façon unique, George (1997) suggère d'adopter une vision plus large et plurielle, en intégrant divers modes inférentiels. Johnson-Laird (1983) propose de catégoriser les différentes formes de raisonnement. Pour lui, il convient de distinguer trois catégories de raisonnement : la déduction, l'induction et les autres formes d'inférences. Aristote, quant à lui, avait déjà distingué les arguments déductifs, inductifs et abductifs (Peirce, 1974).

A. L'abduction

Josephson & Josephson (1994) définissent l'abduction comme le fait d'émettre une hypothèse explicative. L'abduction est un cas particulier de l'induction, soit une inférence vers l'explication la plus plausible. On peut en d'autres termes la considérer comme un processus d'élaboration de théories. Elle ne sert pas à rechercher des lois déjà connues, mais à créer, à produire de nouvelles idées (Peirce, 1974).

B. L'induction

Comme l'abduction, l'induction est en quelque sorte une généralisation, un passage du particulier au général. Peirce (1974) distingue deux formes d'induction : l'induction classique et l'induction statistique. La première forme, l'induction classique, aussi appelée induction par énumération, est l'idée selon laquelle ce qui est vrai pour un cas particulier est généralisable à l'ensemble de situations présentant les mêmes caractéristiques. La seconde forme, l'induction statistique, est légèrement différente de la première. Elle consiste en la généralisation des proportions déjà observées aux mêmes proportions pour les cas à venir. Michalski (1991) ajoute à cette taxonomie la spécialisation inductive qui fait appel à des connaissances qui permettent d'affiner les données et restreindre le champ des possibles. Ainsi, l'induction est généralisante ou spécialisante, c'est-à-dire qu'elle procède du particulier au général, ou du particulier au particulier. D'un point de vue plus large, le raisonnement inductif consiste à prendre comme point de départ des faits particuliers associés entre eux et à tirer de ces associations une proposition générale énonçant la probabilité que de telles associations se manifestent en d'autres occasions (Fortin, 1996).

C. La déduction

Comme l'a défini Aristote, la déduction est étroitement liée à celle de validité. Le raisonnement déductif est une opération mentale qui consiste à prendre comme point de départ une proposition de portée générale et à tirer une hypothèse portant sur des cas particuliers (Fortin, 1996).

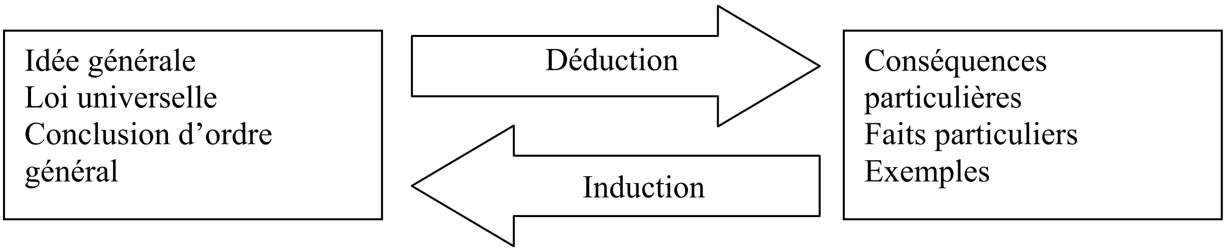


Figure 1.2. Déduction et induction, deux processus inverses

La conclusion d'un raisonnement déductif est, la plupart du temps, certaine, valide. Toutefois, une déduction peut être émise dans un contexte incertain, un espace ouvert. On parle de raisonnement plausible (Pólya, 1954). Dès lors, les conclusions peuvent être obtenues par l'utilisation des probabilités, et plus spécifiquement par l'utilisation d'inférences bayésiennes, propos du chapitre 2.

Précisons que pour Chater & Oaksford (1999), tout raisonnement est probabiliste, dans le sens où, même pour des syllogismes qui réfèrent à un raisonnement déductif, les individus utilisent des stratégies quotidiennes de raisonnement probabiliste.

Chapitre 2 - L'apport bayésien au raisonnement probabiliste

De façon quotidienne, nous effectuons des jugements probabilistes. Ces jugements sont souvent immédiats et spontanés (Baratgin, 2002) et portent sur des événements divers. Prenons quelques exemples afin d'illustrer nos propos. Penser que l'équipe sportive que l'on soutient va gagner le match de ce soir et avoir le sentiment qu'il va faire beau ce même soir, sont deux jugements incertains, intuitifs et ne s'appuient pas tant sur une approche cognitive que conative. Le sentiment de certitude est pourtant présent sans qu'une analyse objective des données (feuille de match et prévisions météorologiques, respectivement, par exemple) ait vraiment lieu pour le sujet. Quand ce même sujet, un homme naïf -nous reviendrons sur cette notion plus tard dans ce chapitre-, prend connaissance d'informations nouvelles concernant les deux événements pré cités, il devrait en tout état de cause réviser son jugement, c'est-à-dire revoir son jugement et changer la probabilité qu'il a attribué à l'événement en question. Ainsi, lorsque notre sujet apprend que le butteur le plus performant de l'équipe sera absent pour le match, pour cause de blessure au genou, et qu'il voit un voisin avec un parapluie humide, il devrait revoir à la baisse l'éventualité de fêter la victoire de son équipe sous le soleil. La formule de Bayes est l'outil, qu'en théorie, il devrait utiliser pour ce faire.

Ce chapitre va nous permettre de développer les notions relatives au raisonnement probabiliste et de décrire l'évolution théorique des paradigmes rendant compte des performances humaines en ces situations de raisonnement incertain. Dans une première partie, nous ferons une rapide description de ces paradigmes, allant des travaux pionniers dans le domaine, à l'utilisation et la reconnaissance de la formule de Bayes comme modèle normatif, qui fait suite à une période de recherche de modèle fort, à proprement parler. Dans une seconde partie, nous aborderons deux approches différentes, conséquences de l'utilisation normative de la formule de Bayes. La première approche est une vision pessimiste de l'homme et de sa non rationalité : l'école *heuristiques et biais* (Tversky & Kahneman, 1973, 1974). La seconde prône une certaine résurgence d'optimisme défendant l'idée selon laquelle le décalage entre les performances bayésiennes théoriques et celles de l'homme naïf, tout comme celles de sujets avertis d'ailleurs, n'est pas tant dû à son irrationalité qu'à des effets de contexte et de présentation : l'approche de la rationalité écologique (Gigerenzer, 1991a, 1991b).

I. Les premiers modèles descriptifs du raisonnement probabiliste

Les premiers travaux portant sur le raisonnement probabiliste remontent à la fin du XIX [e] siècle et portent davantage sur la représentation que les individus se font du hasard et la représentation qu'ils ont des probabilités que sur le raisonnement probabiliste (Baratgin, 2002). Bresson (1965) explique que les individus ont des difficultés à se représenter certaines probabilités. Ainsi, il décrit une certaine propension à surestimer les probabilités faibles et à sous estimer les probabilités élevées. Il serait, en quelque sorte, plus aisé de traiter des nombres intermédiaires que des nombres très petits ou très grands. Les premiers travaux portant sur le raisonnement probabiliste sont attribués à Piaget et Inhelder (1951, 1975) et à Cohen (1963).

A. Le modèle piagétien du raisonnement probabiliste

Dans la conception piagétienne, le nombre est considéré comme un instrument logique de l'intelligence humaine (Bideaud, 1997). « Le développement intellectuel est celui de conduites « adaptatives » endogènes dont l'organisation structurale se rapproche progressivement (par isomorphisme) des formes élaborées étudiées par la logique et les mathématiques » (Bideaud, Meljac & Fisher, 1991). En d'autres termes, l'objectif de Piaget et de ses collaborateurs est avant tout de décrire les constituants logiques du nombre. Piaget & Szeminska (1941) décrivent deux systèmes logiques qui se construisent en même temps : le système cardinal et le système ordinal. Le premier cité -le système cardinal- résulte du groupement logique des classes. Un nombre cardinal est un entier N rendant compte de la quantité d'objets dans un ensemble. Ce système repose sur la notion d'inclusion de classes. Le second système cité -le système ordinal- est la résultante des groupements logiques pouvant être effectués entre divers objets. C'est un système de relations entre les nombres, de relations entre les quantités, rendant possible, par exemple, la sériation. Le parachèvement de ces deux systèmes a pour aboutissement le nombre achevé. En parallèle de ces travaux sur le nombre, Piaget & Inhelder (1951, 1975) ont travaillé sur le jugement probabiliste des enfants. Ils ont décrit trois stades de développement du jugement des probabilités. Avant 7/8 ans, les enfants présentent un raisonnement pré logique et ne font pas la différence entre le déductible et le non déductible. Vers 7/8 ans, l'enfant commence à différencier le hasard et les opérations. Il entrevoit la déductibilité. Passé 8 ans, l'enfant tend vers une synthèse entre les opérations et le hasard, ce qui lui permet vers 12 ans de comprendre la notion de probabilité, les lois combinatoires et les rudiments des grands nombres. Il est « capable d'opérations dont deux au moins sont propres au présent niveau : une combinatoire permettant de tenir compte de toutes les associations possibles entre les éléments en jeu, et un calcul de proportions, si élémentaire soit-il, permettant (ce qui échappe aux sujets des niveaux précédents) que des probabilités telles que 3/9 ou 2/6, etc., sont égales entre elles. » (Piaget & Inhelder, 1966). Notons que dans les travaux de Piaget, nulle allusion à la règle de Bayes n'est faite.

B. Les paradigmes expérimentaux de Cohen sur le raisonnement probabiliste

Cohen (1963) a mené de nombreux travaux expérimentaux auprès d'enfants, d'adolescents et d'adultes, et portant sur le jugement probabiliste. Il défini la « psychologie humaniste » et « les probabilités psychologiques ». Ces deux concepts lui permettent d'introduire une distinction conceptuelle et un nouveau cadre : il convient pour lui de dissocier les probabilités théoriques des probabilités de l'homme de la rue, sujet lambda n'ayant pas les connaissances pour traiter correctement les probabilités. Il décrit différents résultats caractéristiques de cet homme de la rue, ce sujet naïf que nous évoquions au début de ce chapitre. Il montre que les enfants âgés de 4 à 9 ans ne respectent pas la contrainte de complémentarité et sont parfois, au contraire, suradditifs ou sous-additifs. En d'autres termes, ils ont tendance à donner des probabilités, qui une fois sommées, dépassent ou n'atteignent pas le total théorique de 1. Ces résultats semblent cohérents avec ceux trouvés par Piaget & Inhelder (1966) sur leurs deux premiers stades. Cohen (1963) décrit également l'illusion du joueur à propos des jeux de hasard. Pour le classique *Pile ou Face*, il montre que les participants s'attendent davantage à obtenir un lancer *Pile* après une série de *Face*, et inversement. Il interprète cela comme une attente à un équilibre, une équipartition, un 50/50 : autant de *Pile* que de *Face*. Or, sur une petite série de lancers indépendants, cette loi des grands nombres ne s'applique pas. Il décrit également une négligence des anciennes informations au profit des plus récentes dans les tâches d'échantillonnage. Par exemple, à la suite de différents tirages de boules avec remise dans une urne, les individus ont une propension à négliger les premiers tirages et imaginer une urne composée de boules de couleur correspondante aux derniers tirages. Enfin, Cohen (1963) marque le début des travaux portant sur la perception de la prise de risque. Précisons, là encore, que dans ces travaux, nous ne retrouvons pas de référence à la formule de Bayes.

II. L'émergence des modèles mathématiques

Dans les années 1950, les psychologues vont emprunter aux mathématiques des modèles qu'ils vont chercher à appliquer à la théorie de la prise de décision (Bresson, 1965 ; Hammond, McClelland & Mumpower, 1980 ; Von Winterfeldt & Edwards, 1986).
Plus précisément, quelques modèles peuvent être, en quelque sorte, considérés comme marquants pour les études ultérieures du raisonnement probabiliste.

A. La théorie de la prise de décision de Edwards

Edwards (1954), et plus tard Edwards, Lindman, & Savage (1963), vont élaborer une théorie de la prise de décision censée transcender les disciplines et être valide en économie, en théorie des jeux, en statistiques, notamment en probabilités. Avec ces auteurs, nous retrouvons l'idée d'un écart entre une théorie comportementale, celle de notre homme naïf, et une théorie normative, correspondant aux lois mathématiques, c'est-à-dire à l'utilisation des probabilités dans les règles.

B. L'étude de Meehl sur la fiabilité du jugement clinique

Meehl (1954) pose le fait que l'homme naïf n'est pas le seul à s'écarter de la rigoureuse utilisation des probabilités, et que le sujet expert, ici le clinicien, commet aussi des erreurs dans ses jugements probabilistes. Il commet en effet des erreurs dans ses prédictions en ne tenant pas compte de toutes les informations statistiques à sa disposition.

C. L'approche fonctionnaliste de Hammond

En reprenant le modèle fonctionnaliste de Brunswik, Hammond (1955) et Hammond, McClelland, & Mumpower, (1980) ont développé un modèle de *jugement quasi rationnel* qui est en quelque sorte un compromis entre rationalité et intuition. L'homme adopte un mode de raisonnement qu'il juge adapté à son environnement.

D. La rationalité limitée de Simon

Simon (1956) développe l'idée selon laquelle la prise de décision de l'homme est la résultante, à la fois de ce qu'il observe et de son expérience. Ensuite, avec ses capacités cognitives limitées, au niveau du traitement de l'information, l'homme prend des décisions plus ou moins rationnelles.

E. Le sens commun du jugement selon Heider

Heider (1958) s'intéresse à l'attribution causale. Il développe une théorie d'attribution dans laquelle, l'homme naïf, selon ses propres connaissances et croyances, interprète les événements qui surviennent autour de lui, et leur attribue une cause probable, telle l'environnement, l'individu lui-même, ou le contexte.

Ces différents modèles, qu'il serait abusif de qualifier de modèles du raisonnement probabiliste, ont néanmoins contribué à en faire une première description normative. Dans cette mouvance, le théorème de Bayes s'est imposé comme modèle normatif dès le début des années 1960.

III. L'avènement du raisonnement bayésien en tant que modèle de référence

Avant de présenter les recherches utilisant le théorème de bayes, il nous semble judicieux de le présenter ici, à la suite d'un rapide exposé sur les probabilités.

A. Théorie des probabilités

Afin d'appréhender plus précisément l'apport de la formule de Bayes, attardons nous quelque peu sur la théorie des probabilités. Pour Daston (1988) les probabilités et le raisonnement humain sont les deux côtés de la même pièce. Laplace (1814, 1951), quant à lui, considère que la théorie des probabilités se réduit dans le langage commun à du calcul. Toutefois, pour maîtriser le calcul de probabilités, il convient d'en connaître les règles. Notre présentation ne porte pas sur la notion de probabilité antérieure à l'axiomatisation proposée par Kolmogorov (1950). Une probabilité est appelée *p*. Son ensemble de définition va de 0 à

1 [0 ; 1]. 0 signifie que l'événement n'arrive jamais et 1 qu'il arrive toujours. Entre ces deux bornes, nous trouvons une infinité d'intermédiaires. Ces deux particularités sont appelées règles de convexité. La somme des probabilités de toutes les possibilités est égale à 1, c'est la règle d'addition. Par exemple, si je sais qu'un colis doit m'arriver de façon certaine entre lundi et mercredi, la somme des probabilités d'une livraison lundi, mardi et mercredi est de 1. Il y a 100% de chances que mon colis m'arrive un de ces trois jours. La probabilité de voir se réaliser un événement parmi deux événements indépendants -on parle d'événements indépendants dès que l'apparition du premier n'a aucune incidence sur la survenue du second- l'un de l'autre est tout simplement la somme des deux probabilités. On parle alors d'union. Par contre, la probabilité de voir se réaliser ces deux événements indépendants est le produit des deux probabilités. On parle alors d'intersection.

B. Règle de Bayes

Le théorème de bayes s'appuie sur ces axiomes mais concerne des événements non indépendants. Nous devons ce théorème, comme son nom l'indique, au révérend Thomas Bayes (1763), même s'il arrive que son origine soit controversée et des fois attribuée à Laplace (1774).

Figure 2.1. Révérend Bayes

La règle de Bayes permet la révision probabiliste, et est utilisée dans le cadre de probabilités conditionnelles binaires (malade / non malade sachant que je présente les symptômes). En d'autres termes, elle nous permet de réviser notre jugement dès lors que nous obtenons une nouvelle information. Ceci équivaut à passer d'une probabilité *a priori* -précédent la prise en compte du nouvel élément- à une probabilité *a posteriori* -faisant suite à la prise de connaissance de la nouvelle information-. Cette interprétation épistémique de la règle de

Bayes est souvent décrite comme le passage de notre degré de croyance initiale à notre degré de croyance révisée.

$$P(A/B) = \frac{P(B/A)\,P(A)}{P(B/A)\,P(A) + P(B/A^C)\,P(A^C)}$$

<u>Figure 2.2.</u> La règle de Bayes

P(A/B) est la probabilité conditionnelle de A sachant B. Ce qui revient à dire que c'est la probabilité *a posteriori* de A après avoir pris connaissance de B.

P(A) est la probabilité *a priori*, la probabilité marginale de A qui précède toute information sur B.

P(B/A) est la fonction de vraisemblance de A, c'est-à-dire la probabilité d'avoir B sachant que A s'est produit.

P(B) est la probabilité *a priori* de B.

A^C est le complémentaire de A c'est-à-dire « non-A ».

La règle de Bayes peut aussi s'écrire sous une forme réduite :

$$P(A/B) = \frac{P(A)\,P(B/A)}{P(B)}$$

<u>Figure 2.3.</u> Forme réduite de la règle de Bayes

C. L'impact de la règle de Bayes dans la psychologie du raisonnement

Cette règle de Bayes marque un tournant considérable dans l'utilisation des probabilités et rempli un manque certain. Il faudra toutefois attendre de nombreuses années pour qu'elle soit utilisée dans les recherches de psychologie. Elle va néanmoins devenir incontournable dès les années 1960 (Peterson & Beach, 1967) et être à l'origine de nouveaux paradigmes expérimentaux. Cette simple formule -pas tant dans le sens d'aisée à appliquer que dans le sens petite, nous y reviendrons- est considérée comme la norme, la référence. Un sujet est cohérent (au sens bayésien) s'il respecte les axiomes énoncés par Kolmogorov (1950) lors de l'utilisation des probabilités, et, si dans le même temps, il révise conditionnellement au nouveau message appris les probabilités, conformément à la règle de Bayes (De Finetti, 1937).

Ce rapprochement n'est pas le simple fait des psychologues. Il découle également de l'intérêt que les probabilistes commencent, à cette époque, à porter à la psychologie. La première étude expérimentale sur le raisonnement que nous pouvons maintenant appeler bayésien, est menée par un probabiliste (Rouanet, 1961). Les associations entre psychologues et probabilistes accordent à la règle de Bayes le statut de modèle d'intégration de l'information (Edwards, Lindman, & Savage, 1963). Ceci permet d'appréhender le rapport à l'environnement de l'individu.

D. Deux hypothèses concernant la non utilisation de bayes par les individus

En tant que nouveau modèle fort et normatif, la règle de Bayes devient un outil qui permet de revisiter les expériences sur le raisonnement probabiliste. Les psychologues, notamment les théoriciens de la prise de décision vont l'utiliser afin de rendre compte du rapport de l'être humain aux probabilités (Fischhoff & Beyth-Marom, 1983 ; Slovic & Lichtenstein, 1971 ; Von Winterfeld & Edwards, 1986). Deux approches différentes, deux visions de la rationalité humaine vont se succéder. La première, plutôt pessimiste, voit le jour dans les années 1970 et 1980. C'est l'école *des heuristiques et des biais* (Tversky, & Kahneman, 1974 ; Kahneman, Slovic, & Tversky, 1982). La seconde vision est, quant à elle, plus optimiste et envisage l'homme comme un être rationnel, écologique et surtout fréquentiste. C'est l'approche fréquentiste et la rationalité écologique (Gigerenzer, Todd, & the ABC Research Group, 1999 ; Todd, & Gigerenzer, 2000).

1. L'école *heuristiques et biais* de Tversky et Kahneman

Dans les années 1970 et 1980, Tversky et Kahneman -prix Nobel d'économie en 2002- développent l'idée selon laquelle les paradigmes utilisés jusqu'alors pour appréhender le raisonnement probabiliste humain ne reflètent pas la réalité, la vie quotidienne des participants. Les protocoles sont jugés trop artificiels et abstraits. Ainsi, ils délaissent des expériences comme le tirage des urnes (Peterson & Beach, 1967) et proposent d'étudier le raisonnement probabiliste à l'aide de situations expérimentales proches de la vie courante. La règle de Bayes est, dans ce cadre, utilisée non plus comme modèle descriptif mais comme mesure de l'écart entre l'utilisation de la règle de Bayes et la rationalité humaine limitée. Pour les auteurs de cette école des *heuristiques et biais* (Tversky, & Kahneman, 1974), aussi connue sous le nom de théorie psychologique de la décision (Hammond, McClelland, & Mumpower, 1980), l'être humain utilise des stratégies de simplification, dans son rapport à l'environnement. Le but de ces chercheurs est de définir et décrire ces stratégies intuitives, « rapides et frugales » (Chater & Oaksford, 1999), appelées heuristiques. Les principales heuristiques décrites sont l'heuristique de réprésentativité, celle de disponibilité, et l'ancrage et ajustement.

a. L'heuristique de représentativité

Cette heuristique consiste en la recherche d'une image, d'une représentation interne que les individus ont de l'objet, de l'échantillon à évaluer. La probabilité qu'ils attribuent à cet objet dépend donc de la similarité entre l'échantillon et la population parente à laquelle il appartient et de la saillance de ses caractéristiques (Kahneman, & Tversky, 1972). Cette heuristique est donc utilisée dès lors que les individus doivent émettre un jugement d'appartenance, d'inclusion. Tversky & Kahneman (1982) distinguent dans ce concept le jugement de représentativité et le jugement par représentativité.

➢ Le jugement de représentativité

Pour rendre compte d'une probabilité et émettre une prévision, l'individu extrapole les données d'un petit échantillon qu'il a à sa connaissance à la population parente. Selon l'aspect plus ou moins représentatif de l'échantillon, l'individu est soumis à deux biais : le sophisme des petits échantillons, et l'illusion du joueur.

◆ Le sophisme des petits échantillons

Un échantillon est d'autant plus représentatif de sa population parente, qu'il présente en quantité de nombreuses caractéristiques avec elle, et qu'il est large. Lorsque l'échantillon est petit, la loi des grands nombres ne peut s'appliquer. Prenons pour illustrer ce sophisme, ce paralogisme, en accord avec la distinction faite entre ces deux concepts dans le chapitre 1, deux paradigmes expérimentaux énoncés par Kahneman & Tversky (1972) :

> « Dans une certaine région, il y a à peu près 1000 bébés qui naissent chaque jour et en moyenne 50 % des nouveau-nés sont des garçons. Le pourcentage exact varie néanmoins d'un jour sur l'autre. Selon vous, quel est le pourcentage de jours où le nombre de bébés garçons sera :
>
> - Inférieur à 50 garçons,
> - Entre 50 et 100 garçons,
> - […]
> - Supérieur à 950 garçons. »

Le même problème est présenté pour une population de 100 et 10 bébés, avec des catégories équivalentes (entre 0 et 5 garçons, ou 0 garçon respectivement). Les participants répondent de manière identique aux trois versions du problème.

> « Il y a deux maternités dans une certaine ville. Dans la première, environ 45 bébés naissent chaque jour, dans la seconde, 15 bébés. Comme vous le savez, à peu près 50 % des nouveau-nés sont des garçons. Le pourcentage exact varie néanmoins d'un jour sur l'autre. Il peut être tantôt supérieur, tantôt inférieur. Au cours d'une année donnée, chaque hôpital a compté le nombre de journées où ce pourcentage était supérieur à 60 %. Selon vous, quelle est la maternité qui a compté le plus grand nombre de telles journées ? ».

La grande majorité des participants estime qu'il n'y a pas de différences entre les deux maternités et que les deux établissements peuvent avoir des journées de naissance non équilibrées en termes de genre. Pour les raisons évoquées précédemment, ce mode de raisonnement est un paralogisme, les chercheurs de ce courant parlent de biais.

Les auteurs montrent que les participants négligent la taille de l'échantillon. Or cette dernière est essentielle. En effet, plus l'échantillon est grand, plus il est représentatif de sa population mère. Dans le même temps, il est plus aisé de dévier statistiquement de la norme quand l'échantillon est petit. Par exemple, pour un échantillon de deux enfants (un garçon et une fille) ajouter une troisième personne fait passer le sexe ratio à 33 %. Pour obtenir le même phénomène avec un échantillon de vingt enfants (dix garçons et dix filles), il faut ajouter dix garçons ou dix filles. En résumé, le cas particulier a beaucoup plus de poids lorsque l'échantillon est petit. Notons que cet aspect est connu des chercheurs en psychologie, puisque bon nombre des tests statistiques qu'ils utilisent sont sensibles à la taille des échantillons.

- L'illusion du joueur

Ce biais a déjà été évoqué par Cohen (1963). Il est caractérisé par une représentation du hasard relativement ordonnée. Ainsi, lors d'une série de lancers d'une pièce, la série PPFPFF est plus probable que la série PPPPPP ou encore que la série PFPFPF. Pour les individus, les deux dernières séquences ne reflètent pas assez l'aléatoire. Or les tirages étant indépendants les uns des autres, chacune des séquences présentées a autant de chances de sortir, soit une chance sur 64 (2^6). Kahneman & Tversky (1972) ont proposé un problème équivalent qu'ils jugeaient plus familier :

> « Toutes les familles de 6 enfants d'une ville sont étudiées. Dans 72 familles l'ordre exact des naissances filles/garçons est FGFGGF. Quelle est votre estimation du nombre de famille dans lesquelles l'ordre exact des naissances est GFGGGG ? »

Là encore les individus estiment comme peu probable le fait d'avoir une seule fille dans une fratrie de six enfants.

- Le jugement par représentativité

Le jugement par représentativité est la cause de deux biais : l'erreur de conjonction et la négligence du taux de base.

- L'erreur de conjonction

Conformément à la théorie des probabilités, la règle de conjonction nous dit que si deux événements sont indépendants, la probabilité de rencontrer les deux en même temps est plus faible que celle d'en rencontrer un des deux. Ainsi, la probabilité que la prochaine personne que je rencontre dans la rue soit une femme (événement A) avec un chapeau (événement B) est plus faible que celle de rencontrer une femme, la contrainte du chapeau restreignant les possibilités. Tversky & Kahneman (1982) ont mis en exergue cette erreur typique grâce au problème Linda :

> « Linda a 31 ans. Elle est célibataire, c'est une fille très brillante qui n'a pas la langue dans sa poche. Elle est diplômée en philosophie. Quand elle était étudiante, elle se sentait profondément concernée par les problèmes de discrimination raciale et de justice sociale ; elle a également participé à des manifestations antinucléaires.
>
> Quelle est la probabilité que Linda soit :
>
> - Employée de banque ?
> - Militante dans un mouvement féministe ?
> - Employée de banque et militante dans un mouvement féministe ? »

Presque 90 % des sujets accordent une probabilité plus élevée à la troisième phrase, qui est la conjonction des deux premières et par conséquent moins probable que les deux autres individuellement. Certains auteurs font une distinction entre la présence explicite de la conjonction *et* dans l'énoncé (Daniel a été admis à faire médecine ET a terminé ses études) et une présence implicite de cette conjonction (Daniel a terminé ses études de médecine -il a donc été admis-) (Politzer & Noveck, 1991). Ceci souligne l'importance du langage dans les problèmes probabilistes (Dulany & Hilton, 1991), idée sur laquelle nous reviendrons quand nous aborderons l'approche fréquentiste. Schwartz, Strak, Hilton & Naderer (1991) ont réussi à obtenir de meilleures performances en mettant en exergue dans les énoncés les informations statistiques. Par exemple, ils ont dit aux participants que des statisticiens devaient réaliser la même tâche qu'eux. En quelque sorte ils les ont orientés vers une approche statistique. D'autres stratégies facilitatrices sont également utilisées (Gigerenzer, Hell & Blank, 1988 ; Baratgin & Noveck, 2000), comme, par exemple, le fait de présenter les informations psychologiques -les aspects du portrait- après les informations statistiques (Kroznick ; Li & Lehman, 1990).

- La négligence du taux de base

Le fait de négliger le taux de base est la seconde caractéristique du jugement par représentativité. Le taux de base peut être défini comme étant un nombre (pourcentage, mais aussi fréquence, nous y reviendrons) représentant une caractéristique présente dans une population. Les chercheurs nomment ces informations les informations *a priori*. De nombreuses études montrent que les participants ne tiennent que très peu compte de ce taux de base (Bar-Hillel, 1990 ; Lynch & Ofir, 1989 ; Manis, Dovalina, Avis, & Cardoze, 1980 ; Wells & Harvey, 1978), voire l'ignorent totalement (Pollard & Evans, 1983 ; Evans & Bradshaw, 1986). Lorsqu'ils ignorent l'information relative à une caractéristique d'un échantillon (Nisbett & Borgida, 1975), les individus fondent, en contrepartie, leur jugement davantage sur la similarité entre la personnalité des individus de l'échantillon et l'image prototypique qu'ils ont en tête de la catégorie pré citée (Ginossar & Trope, 1987).

Le problème des avocats et des ingénieurs (Kahneman & Tversky, 1973) est couramment utilisé pour rendre compte de cette négligence à l'égard du taux de base (Fagley, 1988 ; Nisbett & Borgida, 1975).

> « Un panel de psychologues a interrogé et fait passer des tests de personnalité à 30 ingénieurs et 70 avocats qui ont tous très bien réussi dans leur spécialité. Sur la base de ces informations, des descriptions des 30 ingénieurs et des 70 avocats ont été écrites. Vous trouverez sur vos questionnaires cinq descriptions disponibles. Pour chaque description, veuillez indiquer votre probabilité que la personne décrite soit un ingénieur (ou un avocat), sur une échelle de 1 à 100.
> La même tâche a été effectuée par un panel d'experts qui ont été très précis dans leurs estimations de probabilités des différentes descriptions. Vous recevrez un bonus proportionnel à la proximité de vos estimations à celles des experts. »

Un second groupe voit le même problème lui être présenté, cette fois avec des taux inversés quant aux proportions ingénieurs / avocats. Les expérimentateurs proposent deux types de descriptions aux participants : les descriptions informatives (Jack, par exemple) et les non informatives (Dick, par exemple). Les premières renvoient aux représentations que l'on a des ingénieurs et des avocats de façon stéréotypée, et les secondes sont plutôt neutres et semblent correspondre aussi bien aux deux professions.

> « Jacques est un homme de 45 ans. Il est marié et a quatre enfants. Il est en général conservateur, prudent, et ambitieux. Il n'a pas d'intérêt pour les problèmes politiques et sociaux et passe la plupart de ses moments de liberté à ses nombreux passe-temps qui incluent la menuiserie, la voile et les puzzles. »

> « Dick est un homme de 30 ans ; il est marié, sans enfant. Homme de grande aptitude et hautement motivé, il promet tout à fait de réussir dans son domaine. Il est bien apprécié de ses collègues. »

Avec la première description, Jacques correspond à la représentation que les individus ont d'un ingénieur, et non celle d'un avocat. La seconde description ne nous permet pas de choisir aisément entre les deux professions. Rappelons que l'énoncé nous donne 30 % d'ingénieurs et 70 % d'avocats. Ainsi, tenir compte de cette information revient à répondre en utilisant ces valeurs aux questions posées, que la probabilité que Jack, ou Dick, soit un ingénieur est de 30 %, et un avocat de 70 %. Les participants ne tiennent pourtant pas compte de ces taux de base et se laissent influencer par la ressemblance entre la description donnée et leur représentation de la profession. Cela va les conduire à donner une probabilité élevée pour Jacques, quant à la possibilité qu'il soit ingénieur, alors qu'il n'y a que 30 % d'ingénieurs dans l'échantillon évoqué. Et dans le cas de Dick, dont la description est relativement neutre, les participants vont estimer la probabilité qu'il soit ingénieur, ou avocat, à 50 %, c'est-à-dire une chance sur les deux professions possibles (Kahneman & Tversky, 1973). Dans les deux cas, les participants ne tiennent pas compte du taux de base : dans le premier cas, ils négligent les informations statistiques (taux de base) et utilisent davantage des informations psychologiques (leurs représentations), et dans le second cas, l'absence d'informations psychologiques ne les amène pas à utiliser davantage les informations statistiques, puisqu'ils donnent un pourcentage de 50 %, et non de 30 % ou 70 %, comme cela doit être, en référence à l'énoncé. Notons que ces résultats ne sont pas systématiquement répliqués (Koehler, 1996).

Un autre problème, non moins connu, est également proposé par Kahneman & Tversky (1973) et par Tversky & Kahneman (1980, 1982). Il s'agit du problème des taxis (Bar-Hillel, 1980 ; Lyon & Slovic, 1976 ; Fischhnoff, Slovic, & Lichtenstein, 1979).

> « On compte dans une certaine ville, 85 % de taxis verts et 15 % de taxis bleus. Un accident a lieu dans lequel un taxi est impliqué. Un témoin affirme que le taxi était bleu. Une série de tests montre que le témoin est fiable à 80 % : il identifie correctement la couleur d'un taxi, qu'il soit vert ou bleu, dans 80 % des cas, et se trompe dans les 20 % restants. Quelle est la probabilité que le taxi impliqué soit effectivement bleu ? ».

La plupart des sujets interrogés répondent à la question posée 80 % de chances que le taxi soit bleu, alors que la réponse attendue, la difficile réponse normative (Birnbaum, 1983) obtenue avec la règle de Bayes est de 41 %, si l'on tient compte de la proportion de taxis bleus dans la ville, c'est-à-dire du taux de base. Une cause avancée pour expliquer ces résultats est le fait que les individus utilisent une probabilité inverse (Villejoubert & Mandel, 2002). En effet, bon nombre de participants, quand on leur demande d'estimer la probabilité conditionnelle de A sachant B -p(A/B)-, répondent l'inverse, c'est-à-dire donnent la fonction de vraisemblance de A : la probabilité d'avoir B sachant A -p(B/A)-. Si ce constat est unanime, il est important de préciser que la cause n'est pas consensuelle. La question posée est la suivante : parmi la négligence du taux de base et l'inversion de la probabilité, quelle est la cause et quelle est la conséquence ? Pour de nombreux auteurs, le biais d'inversion de la probabilité est la conséquence directe de la négligence du taux de base (Bar-Hillel, 1980 ; Dawes, Mirels, Gold, & Donahue, 1993 ; Kahneman & Tversky, 1973 ; Pollard & Evans, 1983). Pour d'autres, le rapport de causalité est inversé et la négligence du taux de base, voire son ignorance, vient de la non maîtrise des probabilités et de la confusion entre les deux probabilités inverses présentées précédemment (Hamm, 1993 ; Koehler, 1996 b ; Wolfe, 1995). Ces derniers auteurs ont comme argument les résultats suivants : lorsque des individus sont entraînés à distinguer les deux probabilités inverses, ils tiennent davantage compte du taux de base (Wolfe, 1995). Ceci va dans le sens d'une confusion de la part des participants entre les probabilités conditionnelles. Pour Hogarth & Einhorn (1992) les participants utilisent un modèle algébrique et ainsi n'utilisent pas de façon adéquate le taux de base lors de l'intégration de la nouvelle information.

L'heuristique de représentativité est donc une hypothèse sérieuse, à l'origine de nombreux travaux, pour rendre compte des écarts entre l'utilisation normative de la théorie des probabilités, plus particulièrement de la règle de Bayes, et l'utilisation plus intuitive des mêmes probabilités, par les sujets non avertis.

L'école des *heuristiques et des biais* (Tversky, & Kahneman, 1974 ; Kahneman, Slovic, & Tversky, 1982) a également proposé l'heuristique de disponibilité et l'ancrage et ajustement, qui sont nettement moins liées au raisonnement bayésien. Pour cette raison, la description que nous allons en faire sera relativement sommaire.

b. L'heuristique de disponibilité

Cette heuristique est davantage liée aux processus mnésiques qu'au raisonnement. En effet, lorsque des individus tentent d'évaluer la probabilité d'un événement futur, ils recherchent en mémoire la survenue d'événements analogues. Selon la facilité avec laquelle ils parviennent à se remémorer des événements similaires, ils vont estimer comme plus ou moins probable l'apparition de l'événement en question. En d'autres termes, selon la disponibilité mnésique de faits analogues, ils vont estimer la fréquence d'apparition de celui à évaluer. Ceci est un biais de raisonnement. Il est admis que les événements tragiques comme les accidents d'avion marquent les esprits. Aussi, bon nombre de personnes pensent qu'il est plus risqué de prendre l'avion que de prendre la voiture, car ils se remémorent facilement le dernier crash très médiatisé. Dans le respect de la règle de Bayes, il nous faudrait présenter les taux de base -que nous n'avons pas-, à savoir le nombre de personnes prenant l'avion et la voiture comme moyen de transport, afin d'en tenir compte dans notre calcul de la probabilité d'avoir un accident en utilisant un de ces deux moyens de déplacement. Tversky & Kahneman (1973) ont proposé le paradigme suivant pour appréhender cette heuristique de disponibilité :

> « La fréquence d'apparition des lettres en anglais est étudiée. Un texte typique est sélectionné, et la fréquence relative d'apparition des différentes lettres de l'alphabet en première et en troisième positions est calculée. Les mots de moins de trois lettres sont exclus du compte. On va vous présenter différentes lettres et vous devrez juger si cette

> lettre apparaît plus fréquemment en première ou en troisième position et estimer le rapport des deux fréquences.
> Considérez la lettre R (respectivement K, L, N, V). Est-ce que R apparaît plus fréquemment
> En première position
> En troisième position
> Mon estimation du rapport des deux fréquences est ______ ».

La majorité des individus pense qu'il est plus probable de rencontrer des mots avec un R en première position qu'en troisième. Les résultats sont les mêmes avec les autres lettres présentées. La réalité est tout autre. Tversky & Kahneman (1973) pensent que ces estimations de fréquence sont dues à la facilité des participants à évoquer un mot ayant la lettre comme initiale. Présenter une lettre peut être ici considéré comme un amorçage, qui va faciliter l'accès en mémoire. Les mots sont alors davantage disponibles.

c. L'heuristique d'ancrage et ajustement

Pour Kahneman & Tversky (1974), l'heuristique d'ancrage et ajustement fonctionne comme une « ancre » du jugement. Les individus fondent leur jugement sur une première information, et ne font que l'ajuster par la suite. Ainsi, cette heuristique est en quelque sorte une première impression qu'il n'est pas aisé -ou souhaité- de se défaire et qui oriente par rapport à la tâche en cours. Pour mettre en exergue cette heuristique, Kahneman & Tversky (1973) ont utilisé deux paradigmes. Le premier consiste à demander aux participants d'estimer le nombre de pays africains présents aux Nations Unies. Avant de leur poser cette question, les chercheurs leur demandent de tourner une roue portant les nombres de 1 à 100 et de juger si le nombre de pays est supérieur ou inférieur à celui donné aléatoirement par la roue. Ils ont mis en évidence une corrélation entre le nombre aléatoire donné par la roue et le nombre proposé par les participants. Pour les auteurs, les participants accordent de l'importance au nombre de la roue -qui est pourtant le fruit du hasard-, et ne font qu'ajuster leur estimation par rapport à ce nombre qui persiste.

Ils ont utilisé un second problème pour illustrer cette heuristique ancrage et ajustement. C'est le problème du calcul mental. Les chercheurs ont demandé à deux groupes de participants d'estimer le résultat d'un produit. Le premier groupe doit évaluer la version a du problème en cinq secondes, et le second groupe doit évaluer la version b avec le même temps imparti.

> Version a : « Vous allez disposer d'exactement 5 secondes pour répondre à la question suivante : évaluer le produit 1 x 2 x 3 x 4 x 5 x 6 x 7 x 8 ».
> Version b : « Vous allez disposer d'exactement 5 secondes pour répondre à la question suivante : évaluer le produit 8 x 7 x 6 x 5 x 4 x 3 x 2 x 1 ».

En seulement cinq secondes, les participants ne peuvent réaliser le calcul dans son intégralité, mais ne peuvent qu'en faire une estimation. Le résultat du produit est de 40320. Les sujets soumis à la version a estiment le produit à environ 512, et ceux de la version b à environ 2250. Kahneman & Tversky (1974) avancent deux arguments pour montrer que les participants ont utilisés l'heuristique d'ancrage et ajustement pour réaliser leur calcul. Le premier est le fait que les estimations des deux groupes sont de loin inférieures au résultat théorique. Rappelons que la notion de commutativité de la multiplication nous donne le même résultat pour les deux versions. Leur second argument est le fait que les individus du premier groupe estiment plus faiblement le produit que le second groupe. Pour eux, les participants ont basé leur estimation sur le premier produit, à savoir respectivement 1 x 2 et 8 x 7 pour les deux versions, et ont seulement ajusté leur estimation sur ce premier sous produit. Ainsi, les participants ayant un premier sous produit petit estiment faiblement le produit total, et inversement pour les participants de la version b.

Pour Tversky & Kahneman (1974), l'heuristique d'ancrage et ajustement est à l'origine de deux biais relatifs à l'utilisation des probabilités : la surestimation des probabilités conjonctives, et la sous-estimation des probabilités disjonctives.

- Surestimation des probabilités conjonctives

Les protocoles visant à mettre en évidence cette surestimation des probabilités conjonctives nous ramènent aux classiques en la matière, à savoir les tirages de boules dans les urnes. Les participants doivent choisir laquelle des situations est la plus probable parmi les deux suivantes :

> Situation A : Sortir sept fois de suite dans un tirage aléatoire (avec remise) une boule noire d'une urne contenant neuf boules noires et une boule blanche.
> Situation B : Sortir une boule blanche dans un seul tirage aléatoire d'une urne contenant cinq boules noires et cinq boules blanches.

La probabilité de voir survenir la situation A est $P(A) = 0,9^7 = 0,4783$. En effet, il y a neuf chances sur 10 de tirer une boule noire, soit 0,9 en fréquence, à chaque tirage. Ayant neuf tirages avec remise, il faut appliquer les règles de probabilités et ainsi multiplier cette probabilité par le nombre de tirages indépendants. La probabilité de voir survenir la situation B est $P(B) = 0,5$. En effet, il y a une chance sur deux de tirer une boule noire sur un seul tirage puisqu'il y a dans l'urne autant de boules noires que de blanches. La situation B est donc plus probable. Toutefois, la majorité des participants estime la situation A comme plus probable (Bar-Hillel, 1973). Pour Tversky & Kahneman (1974) les sujets sont biaisés par l'heuristique d'ancrage et ajustement. Ils s'ancrent sur la probabilité du premier tirage et ne font que comparer 0,9 à 0,5. Même s'ils ajustent ensuite cette estimation, ils ne revoient pas suffisamment à la baisse la probabilité de la situation A.

- Sous-estimation des probabilités disjonctives

De la même manière qu'ils surestiment les probabilités conjonctives, les individus sous-estiment les probabilités disjonctives. A nouveau les paradigmes de tirages de boules dans une urne sont utilisés. Les participants doivent dire quelle situation est la plus probable parmi les deux suivantes :

> Situation A' : Sortir au moins une fois dans sept tirages aléatoires (avec remise) une boule blanche d'une urne contenant neuf boules noires et une boule blanche.
> Situation B' : Sortir une boule blanche dans un seul tirage aléatoire d'une urne contenant cinq boules noires et cinq boules blanches.

Notons que ces situations sont les complémentaires des deux situations précédentes.

La probabilité de voir la situation A' est P(A') = 1 - P(A), soit 0,5217. La probabilité de rencontrer la situation B' est, selon le même principe de complémentarité, de 0,5. Cette fois, la première situation est donc la plus probable. Pourtant, les participants jugent la situation B' plus probable que la situation A' (Bar-Hillel, 1973). L'explication de l'école heuristiques et biais est la même : les individus fondent leur jugement uniquement sur le premier tirage (Tversky & Kahneman, 1974).

➢ L'heuristique d'ancrage et ajustement et le conservatisme

L'heuristique d'ancrage et ajustement est également, pour Tversky & Kahneman (1973), une forte hypothèse explicative du phénomène de conservatisme, bien connu par les probabilistes et les chercheurs en théorie comportementale de la décision. Ce biais qu'est le conservatisme est en quelque sorte l'opposé de la négligence du taux de base. Avec la négligence du taux de base, nous avons vu que les participants ne tiennent pas, ou peu, compte des informations de base, alors qu'a contrario, avec le conservatisme, ils ont une forte tendance à se focaliser sur l'information de départ et à ne pas suffisamment réviser leur jugement, par rapport aux préconisations de la règle de Bayes. La probabilité révisée qu'ils proposent est plus proche de la probabilité *a priori* que ne l'est la probabilité *a posteriori* calculée avec la formule de Bayes.

Le paradigme des urnes et des boules (Peterson & Beach, 1967) illustre le biais de conservatisme :

> « Deux urnes sont remplies de nombreux jetons de poker. La première urne contient 70 % de jetons rouges et 30 % de jetons bleus. La seconde urne contient 30 % de jetons rouges et 70 % de jetons bleus. L'expérimentateur jette une pièce non biaisée en l'air pour sélectionner une des deux urnes de sorte que la probabilité *a priori* de la sélection de chaque urne est de 0,5. Il tire ensuite une succession de jetons de l'urne sélectionnée. Supposons que l'échantillon obtenu contient 8 jetons rouges et 4 jetons bleus (situation A). Quelle est votre probabilité révisée que l'urne sélectionnée soit l'urne à dominante rouge ? ».

Les probabilités *a priori* sont de 0,5 puisque le tirage de la pièce non truquée donne équivalente les chances de sélectionner chacune des deux urnes. L'application stricto sensu de la règle de Bayes donne une probabilité révisée de 0,967. Or, la grande majorité des participants ne propose comme probabilité *a posteriori* qu'une valeur de 0,75. L'école heuristiques et biais pense que, là encore, les individus se focalisent sur le premier tirage, celui de la pièce, et ne font qu'ajuster leur probabilité pour réviser leur jugement.

2. L'après l'école *heuristiques et biais*

L'école *heuristique et biais* a eu un impact considérable sur les recherches menées, tant en psychologie que dans d'autres disciplines, telles la philosophie et le champ d'études des probabilités.

Toutefois, elle a également connu bon nombre de critiques (Anderson, 1991 ; Cohen, 1981). Les plus argumentées dans la littérature s'appuient sur divers arguments. Le premier argument est l'utilisation du modèle bayésien comme référent. Est-ce vraiment un modèle applicable et transposable à la psychologie cognitive et à l'étude du raisonnement et de la prise de décision ? Cette question est encore débattue.

Un autre argument avancé est la valeur limitée des heuristiques. Pour de nombreux auteurs (Anderson, 1991 ; Shafir, Smith & Osherson, 1994) ces heuristiques sont mal définies et sont donc difficilement modélisables. Elles sont la description du comportement humain, disons une description des stratégies utilisées par les participants pour raisonner de façon probabiliste, mais n'en sont pas une réelle explication (Anderson, 1991).

Un autre argument est souvent évoqué. Il ne concerne, cette fois, pas tant les heuristiques que les biais qui leur sont associés. Les biais ne sont pas explicitement définis et sont souvent confondus avec les notions de sophisme et d'erreurs. Certains évoquent même la notion de stratégies. De la même manière que pour certains c'est la loi qui provoque le délit, certains auteurs avancent l'idée selon laquelle le modèle normatif utilisé est la cause des biais, et si on ne prend plus comme référence une norme théorique, alors les biais disparaissent (Hammond, 1996 ; Oaksford & Chater, 2001). Les modèles théoriques, notamment les systèmes normatifs, sont inadaptés pour rendre compte des inférences probabilistes humaines car ils renvoient à des compétences computationnelles que les individus, avec leur rationalité limitée, semblent ne pas avoir (Politzer, 2002).

Les biais seraient également dépendants les uns des autres. Plus précisément, ils seraient corrélés entre eux. Un biais peut donc en cacher un autre (Beyth-Marom &

Fischhnoff, 1983 ; Baratgin & Noveck, 2000), un autre parfois plus ancien et décrit bien avant l'émergence de l'école *heuristiques et biais* (Poulton, 1994).

Enfin, ces heuristiques et biais semblent sensibles à différents aspects liés aux paradigmes qui les mettent en exergue. Ils ont sensibles à la présentation, comme nous l'avons vu précédemment : la façon séquentielle dont sont présentées les informations aux participants a une importance réelle car elle influence nettement les performances (Kroznick ; Li & Lehman, 1990). Les biais sont aussi dépendants du contexte. Ainsi, Poulton (1994) trouve que les participants ont de meilleures performances si le problème des maternités leur est présenté sous un isomorphisme renvoyant au sport, un contexte plus familier. Les biais sont également dépendants des connaissances du sujet dans le domaine étudié. Par exemple, s'ils doivent estimer la présence d'une caractéristique dans un échantillon, la connaissance qu'ils ont quant à la probabilité de cette caractéristique dans la population parente oriente leurs performances. Ils ont comme des connaissances intuitives quant à l'homogénéité des traits de la population, et donc de l'échantillon (Nisbett, 1993). La robustesse des biais dépend également du format de présentation, puisque l'utilisation de fréquences naturelles améliore les performances probabilistes.

3. L'approche fréquentiste et l'idée de rationalité écologique

A la suite des travaux de l'école *heuristiques et biais*, un programme d'aussi grande envergure a vu le jour. Le porte-étendard est Gerd Gigerenzer qui a une approche plus optimiste de l'être humain et de sa rationalité. Inspiré d'une idée développée par les associationistes tels que Hume (1739 ; 1951) il défend l'idée selon laquelle l'homme, comme les autres animaux d'ailleurs, se sont adaptés, au cours de leur histoire, grâce à des représentations de fréquences d'événements et non pas des représentations d'événements isolés. Hasher & Zacks (1984) ont avancé, en adéquation avec cette idée, l'hypothèse du processus automatique fréquentiste. Pour ces chercheurs, l'homme enregistre naturellement les fréquences des événements passés, et ceci plus facilement que des pourcentages (Gigerenzer, 1991b, 1993). Gould (1992, p. 469) considère que nos esprits ne sont pas construits pour travailler avec les règles des probabilités. Gigerenzer pense donc que l'être humain est rationnel car adapté à son environnement. En effet, les animaux évoluent dans un environnement auquel ils s'adaptent selon des événements contingents survenus en fréquences (Brunswik, 1939 ; Gallistel, 1990 ; Shanks, 1991). Certaines espèces d'animaux, comme les insectes, les rats ou encore les canards, semblent être très sensibles aux fréquences des événements survenus (Gallistel, 1990 ; Real, 1991 ; Real & Caraco, 1986), comme s'ils étaient de « bons statisticiens intuitifs » (Gigerenzer & Hoffrage, 1995, p. 5). Dans le débat qui oppose son point de vue à l'école *heuristiques et biais* (Gigerenzer, 1996 ; Kahneman & Tversky, 1996), Gigerenzer considère la rationalité comme un verre d'eau, et Kahneman et Tversky voient le verre à moitié vide, alors que lui le voit à moitié plein.

Sa vision va plus loin que la remise en cause de la règle de Bayes comme modèle normatif. Pour lui, les heuristiques ne sont ni des biais ni des erreurs mais des outils appartenant à la boîte adaptative de l'homo sapiens. Les heuristiques sont en accord avec l'environnement humain. Pour Brunswik (1939) et Anderson (1991) la correspondance entre le problème posé et l'environnement humain est une condition nécessaire, certes non suffisante, pour avoir une analyse rationnelle. Un isomorphisme structurel est possible entre l'environnement et le raisonnement bayésien si le paradigme est présenté sous forme de fréquences naturelles et non plus de probabilités conditionnelles (Gigerenzer & Murray, 1997). Illustrons ce format avec le paradigme de la mammographie (Eddy, 1982) présenté dans un premier temps en probabilités conditionnelles -avec des pourcentages- et par la suite en fréquences naturelles :

> « 1 % des femmes de 40 ans qui participent à un dépistage de routine ont un cancer du sein. 8 femmes sur les 10 avec un cancer du sein auront des mammographies positives. 9,6 % des femmes sans cancer du sein auront aussi des mammographies positives. Une femme appartenant à cette tranche d'âge a eu lors d'un dépistage de routine une mammographie positive. Quelle est la probabilité qu'elle ait réellement un cancer du sein ? _____ % ».

> « 10 femmes de 40 ans sur 1000 qui participent à un dépistage de routine ont un cancer du sein. 8 femmes sur les 10 avec un cancer du sein auront des mammographies positives. 95 femmes sur les 990 sans cancer du sein auront aussi des mammographies positives. Une femme appartenant à cette tranche d'âge a eu lors d'un dépistage de routine une mammographie positive. Quelle est la probabilité qu'elle ait réellement un cancer du sein ? ______% ».

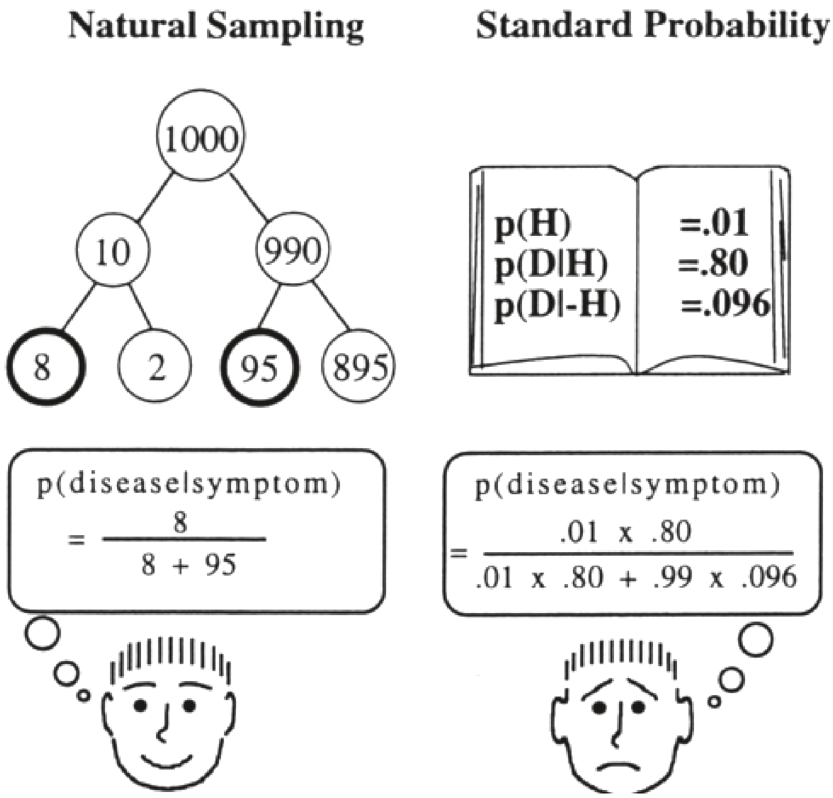


Figure 2.3. Représentation de l'information selon le format (extrait de Gigerenzer & Hoffrage, 1995)

Gigerenzer (1991b) trouve qu'une présentation en fréquences naturelles améliore les performances des individus et les libèrent des tunnels mentaux et de leur inconscient cognitif (Piattelli-Palmarini, 1995). La règle de Bayes et l'utilisation des fréquences nous donne comme rapport 0 ,078. Toutefois, avec la forme fréquentiste (cf. figure 2.3), seuls les nombres 8 et 95 sont utilisés :

$$P(\text{maladie/symptôme}) = 8 / (8+95) = 0{,}078$$

La probabilité *a posteriori* ne dépend donc pas des taux de base. Kleiter, Krebs, Doherty, Garavan & Chadwick (1997) estiment qu'il est alors rationnel d'ignorer l'information relative au taux de base. Cosmides & Tooby (1996) ont montré que la négligence du taux de base disparaissait en format fréquentiste, 76 % des participants réussissent le problème en fréquences, contre seulement 12 % en probabilités. Les formats peuvent être équivalents d'un point de vue mathématique -cette question est encore débattue et source de polémique, nous y reviendrons ultérieurement- mais pas d'un point de vue psychologique (Feynman, 1967). Martignon, Vitouch, Takezawa, & Forster (2003) ont montré, que bien qu'équivalents, les chiffres romains et les chiffres arabes ne se manipulent pas aussi aisément les uns par rapport aux autres. Pour Feynman (1967), différentes représentations de la même formule mathématique peut évoquer diverses images mentales, amener de nouvelles représentations mentales, et mener à de nouvelles solutions.

Quel est l'enjeu de cette distinction entre les fréquences naturelles et les probabilités conditionnelles ? Il n'est pas seulement théorique, ni même méthodologique, mais a une réelle importance quant au débat sur la rationalité / irrationalité humaine. En effet, si une des conditions amène les individus à réussir les problèmes proposés, alors il n'est plus autant légitime de parler d'irrationalité humaine. L'alternative est donc une rationalité limitée soumise à un effet de contexte, et de présentation. Ce dernier s'applique, semble-t-il, tant aux sujets naïfs -l'homme de la rue- qu'aux sujets experts, censés être passés maître en la matière de probabilité.

a. L'impact des fréquences naturelles sur les performances de sujets naïfs

La règle de Bayes est utilisée dans de nombreux domaines. Elle permet, par exemple, en informatique de rendre autonomes certains logiciels, tels les antivirus. En effet, la détection de virus se fait selon les bases de données statistiques mises à disposition du logiciel, qui calcule, d'après la règle de Bayes, la probabilité que le fichier entrant est un virus. Un logiciel censé jouer aux échecs se contente d'évaluer la valeur d'un coup en cherchant dans sa base de données le nombre de fois où un coup équivalent a fait gagner ou perdre la partie au joueur le réalisant. Plus l'on enrichit la base de données -on donne de nouvelles informations- plus le logiciel révise son jugement.

La météorologie est également une science statistique qui s'appuie sur les données passées. C'est en effet en s'appuyant sur les pressions atmosphériques et la force des vents, par exemple, que les météorologues estiment la probabilité d'avoir un week-end ensoleillé. L'indice de fiabilité dépend donc directement des proportions de cas semblables aux conditions météorologiques du jour. Si à chaque fois que ces conditions ont été réunies il a plus, alors il pleuvra sans doute. Si ce n'est pas le cas, les météorologues vont intégrer à leurs données cette possibilité. La prochaine fois que des conditions équivalentes se présenteront, leur estimation sera plus fiable -car la taille de l'échantillon est plus grande- et la probabilité qu'il pleuve plus faible -car pas sûre à 100 %-.

La question qui nous intéresse ici est la suivante : que comprennent les individus qui s'informent sur la météorologie et comment interprètent-ils les probabilités qui leur sont communiquées ?

Murphy, Lichtenstein, Fischhoff & Winkler (1980) ont montré que les interprétations sont multiples quant à une phrase comme « 30 % de chances de pleuvoir demain ». Beaucoup de météorologistes communiquent leurs estimations sous la forme de probabilités car ils sont convaincus que la probabilité est le format adéquat pour juger des risques et qu'elle fournit davantage d'informations (Monahan & Steadman, 1996 ; Murphy & Winkler, 1971). Pour ces mêmes professionnels, une information quantitative est à la fois plus dense et moins ambiguë qu'une information qualitative du genre « il va probablement pleuvoir demain ». Si l'information est plus riche et davantage en adéquation avec la notion d'incertitude, le fait de rencontrer différentes interprétations nous laisse penser que le choix de présentation n'est pas si judicieux qu'il y parait. En effet, va-t-il pleuvoir demain dans 30 % des endroits concernés par la page météo, va-t-il pleuvoir 30 % de la durée de la journée ? (Gigerenzer, Hertwig, van den Broek, Fasolo & Katsikopoulos, 2005). Les individus ne sont pas familiarisés avec la notion de probabilité et ne savent pas comment l'interpréter. Toutefois, cette hypothèse n'est pas la seule : les individus éprouvent des difficultés à interpréter les probabilités quand celles-ci ne renvoient pas explicitement à (aux) événement(s) auquel(s) elles se réfèrent (Murphy, Lichtenstein, Fischhoff & Winkler, 1980). Gigerenzer, Hertwig, van den Broek, Fasolo & Katsikopoulos, (2005) préconisent une information diffusée en fréquences telle « 3 fois sur 10 que les météorologistes font cette prédiction, il pleut ». Cet exemple sur la météorologie a son équivalent en médecine. Face à un médecin, comme face à un météorologue, les individus naïfs sont en difficulté et ne savent comment interpréter l'information qui leur est présentée. Gigerenzer (2002) a mis en avant des données similaires à celles sur la météorologie. Un psychiatre, qui prescrit un antidépresseur à des patients dépressifs, les avertit qu'ils peuvent développer des troubles sexuels dans 30 à 50 % des cas. En plus du stress provoqué par la situation, ces individus sont assez perplexes quant à la façon d'interpréter cette information. Cette proportion renvoie-t-elle à la proportion de troubles sexuels chez les patients soignés par ce médicament, réfère-t-elle à la proportion de troubles sexuels chez un patient par rapport au nombre total de rapports ? Si le pourcentage se réfère à l'ensemble des patients, est-ce que ce sont tous les patients au niveau national ou seulement les patients de ce psychiatre ? Dans ce cas de figure aussi, Gigerenzer (2002) préconise d'utiliser une communication plus claire avec des fréquences qui vont préciser l'échantillon de référence. Il est ainsi plus limpide pour le patient de comprendre que 3 à 5 des patients sur 10 qui ont pris ce traitement présentent des troubles secondaires.

Pour vérifier l'effet facilitateur des fréquences sur les performances bayésiennes, Zhu & Gigerenzer (2006) ont fait passer des problèmes bayésiens sous les deux formats à des adultes. Leurs résultats sont probants : dans la condition probabiliste, les adultes répondent correctement à 47 % des problèmes, alors que dans la condition fréquentiste, ils réussissent environ 76 % des problèmes. Cette différence significative montre, pour Zhu & Gigerenzer (2006), que si l'on souhaite appréhender la rationalité humaine et maximiser les performances des individus, il faut présenter les données en fréquences naturelles et non en probabilités.
Ces différents travaux confirment la thèse selon laquelle les individus ne seraient pas faits pour traiter les probabilités (Gould, 1992) mais auraient une disposition à assimiler les fréquences. Notons toutefois, que ces travaux concernent des individus naïfs, qui n'ont pas l'habitude de rencontrer des probabilités. Qu'en est-il des sujets que l'on peut considérer comme experts en probabilités ? Nous venons déjà de voir ici que les sujets experts s'ils manipulent les probabilités, ne savent pas forcément les utiliser pour communiquer de façon limpide. Les travaux suivants vont nous montrer que les experts peuvent même se méprendre dans l'utilisation des probabilités.

b. L'impact des fréquences naturelles sur les performances de sujets experts

De nombreux travaux montrent que les sujets experts qui utilisent les probabilités au quotidien ne le font pas de façon adéquate (Budescu & Wallsten, 1995 ; Gigerenzer, Hoffrage & Kleinbölting, 1991). Par exemple, Gigerenzer, Hoffrage & Ebert (1998) ont montré que des médecins rencontrant une population à risques quant au virus du SIDA n'estiment pas de façon correcte le risque de contamination et commettent des biais. Pour ces chercheurs, certains médecins communiquent à leurs patients une probabilité erronée. Ils montrent que les conseillers habitués à présenter les risques de contamination, ne présentent pas les risques du test de dépistage en lui-même, souvent car ils ne les pèsent pas eux-mêmes. Plus précisément, les conseillers n'entrevoient pas la possibilité que le test soit un faux positif, et ne tiennent pas compte de la prévalence du virus selon la population, en d'autres termes, du taux de base. Gigerenzer, Hoffrage & Ebert (1998) ont montré que pour la plupart des conseillers, un test positif signifie avec certitude que le patient est infecté, et cela indépendamment du nombre de personnes infectées dans la population à laquelle appartient le patient. En d'autres termes, les médecins ne tiennent pas compte de la prévalence du virus. Dans leur expérimentation, les chercheurs ont soumis les médecins à deux cas de figure.

Dans le premier, un patient appartenant à une population qui ne présente pas de risques particuliers est soumis à un test. Dans cette première population, 1 individu sur 10 000 a le virus du SIDA. Cet individu aura un test positif. Toutefois, un autre individu non infecté aura également un test positif -c'est le faux positif-.

Dans le second cas de figure, une autre population, plus exposée au virus présente la prévalence suivante : 150 individus sont porteurs et leur test sera, à tous, positif. Un autre individu aura également un test positif, sans avoir contracté la maladie.

Dans le premier cas, deux individus auront un test positif et un seul sera atteint réellement. Ceci équivaut à une chance -au sens statistique du terme, évidement- sur deux, soit 50 % de risques, d'être infecté en ayant un test positif.

Dans le second cas, 151 individus présenteront un test positif, et pour 150 d'entre eux ce test aura raison. La probabilité d'être infecté en ayant un test positif est cette fois bien plus élevée : 150/151 = 0,9934.

Gigerenzer, Hoffrage & Ebert (1998) ont montré que les conseillers médicaux ne tentaient pas compte de cette différence (un seul parmi 20 prend cette information en considération). Or l'enjeu pour un patient pour lequel le test est positif est essentiel. Les conséquences peuvent être désastreuses (Gigerenzer, 1998). Les auteurs préconisent une fois de plus une communication des données statistiques en termes de fréquences, tant pour la facilité de manipulation pour les professionnels que pour la facilité de compréhension pour les patients. Ceci réduit les erreurs de physiciens et ainsi le taux de réussite passe de 10 % à 46 % si on change les probabilités en fréquences (Hoffrage & Gigerenzer, 1998).

D'autres professionnels sont autant concernés par les probabilités et non moins à l'abri d'une utilisation non erronée. Gigerenzer (1998) cite un cas d'enquête policière. Dans cet exemple, deux experts, l'un spécialisé dans les groupes sanguins et l'autre dans les fibres textiles, ont estimé réciproquement à 97,3 % et 99,94 % les chances de culpabilité de l'individu. Les données étaient présentées en probabilités et en pourcentages et ont été mal utilisées. Les auteurs pensent que l'utilisation de fréquences aurait permis d'estimer la culpabilité entre 33 % et 50 %. Pour la petite histoire, l'accusé fut innocenté par son alibi. A l'opposé, le raisonnement probabiliste peut aussi faire avancer la loi. Dans un article récent, Domenici, Toni, Spinetti, Rocchi & Presciuttini (2007) ont évoqué le cas d'un double infanticide italien. Les enquêteurs ont utilisé la règle de Bayes pour définir les possibilités de liens entre les parents et les enfants retrouvés morts et pour définir les probabilités *a priori* et *a posteriori*. Cette méthodologie, couplée à un test ADN a permis de conclure à la culpabilité des parents. Les auteurs espèrent voir l'inférence bayésienne de plus en plus utilisée dans les cours pénales d'Italie.

L'inférence bayésienne est un outil de nombreux professionnels et est utilisée dans divers domaines d'expertise. Nous avons cité la météorologie, la médecine diagnostique, la police judiciaire, l'informatique. Ces domaines sont très variés et l'avènement de l'ère informatique ne va qu'augmenter l'utilisation de cette inférence. Si les experts humains semblent avoir besoin d'une formulation fréquentiste pour exceller dans l'utilisation de données statistiques, les ordinateurs, eux semblent très efficients avec les probabilités. Campos & Romero (2008) ont par exemple mis au point un logiciel qui s'appuie sur la règle de Bayes pour créer un Thésaurus autonome.

c. L'effet d'une formation aux fréquences pour une maîtrise du raisonnement probabiliste

Pour que l'être humain soit davantage efficace dans l'utilisation des probabilités, Gigerenzer & Hoffrage (1995) ont proposé d'entraîner les individus à l'utilisation des fréquences naturelles dans le but de maîtriser la règle de Bayes. C'est dans cette perspective que Sedlmeier & Gigerenzer (2001) ont mis en place un programme d'entraînement de représentation des fréquences. Enseigner la façon de se représenter les choses est une alternative à l'enseignement des règles en elles-mêmes (Arkes, 1981). Pour Falk & Konold (1992), l'enseignement du raisonnement bayésien doit en premier lieu porter sur la façon d'insérer les probabilités des événements dans la règle de Bayes. Cet entraînement viserait donc une meilleure compréhension de la règle à proprement parlé. Dans cette perspective, la littérature est riche en programmes d'entraînement, notamment dans le domaine des mathématiques. Par exemple, Fong & Nisbett (1991) ont proposé un entraînement à la règle concernant la loi des grands nombres en probabilités. Ils ont montré un effet limité de l'entraînement. De plus, cet effet modéré, s'atténue considérablement dès deux semaines (Ploger & Wilson, 1991 ; Reeves & Weisberg, 1994). Pour cette raison, Sedlmeier & Gigerenzer (2001) ont proposé un entraînement sur les représentations des problèmes bayésiens plutôt que sur la règle de Bayes. Les auteurs ont proposé deux types de représentations informatisées : une grille de fréquences et un arbre de fréquences comme présenté précédemment. Les auteurs ont comparé quatre groupes expérimentaux : le premier est entraîné à se représenter les données sous la forme d'une grille de fréquences, le second est entraîné à se représenter les données sous forme d'un arbre de fréquences, le troisième reçoit un entraînement à l'utilisation de la règle de Bayes et le quatrième est un groupe contrôle qui ne reçoit pas d'entraînement. Une grille de fréquences est une matrice carrée de dix cases de côté, soit une matrice de cent cases. Dans cette matrice les individus doivent schématiser les deux critères (nature du test et fait d'être porteur ou non de la maladie) en utilisant des symboles et couleurs différents. Leurs résultats vont dans le sens de ceux de Sedlmeier (1999), à savoir un meilleur apprentissage en fréquences qu'en probabilités. Comme l'indique le titre de leur recherche -Teaching Bayesian reasoning in less than two hours-, en moins de deux heures, les chercheurs arrivent à former les participants, des étudiants universitaires, au raisonnement bayésien, plus précisément à l'utilisation de la règle de Bayes. Les trois groupes expérimentaux présentent un apprentissage positif et une amélioration de leurs performances. De plus, les bénéfices sont transférables, dans les trois conditions, à de nouveaux problèmes. Les effets de cet entraînement sont stables et se maintiennent dans le temps (Kurzenhäuser & Hoffrage, 2002). Les auteurs remarquent néanmoins que l'entraînement à la représentation du problème sous forme de grille ou d'arbre de fréquences améliorent davantage les performances, immédiatement après la séance de training, que l'entraînement à l'utilisation de la règle de Bayes. De plus, les deux conditions d'entraînement à la représentation donnent des effets plus durables dans le temps. Ces résultats laissent entrevoir l'intérêt pédagogique et éducatif d'utiliser de tels outils informatisés pour enseigner les mathématiques, notamment les statistiques, aux élèves (Paulos, 1988 ; Sedlmeier, 1999). Cette approche est également envisageable avec un autre public : les professionnels utilisant les probabilités de façon quotidienne. Cette formation serait bénéfique à deux niveaux : d'une part elle offrirait une meilleure compréhension au niveau individuel pour le praticien, ce qui lui permettrait de ne plus se fourvoyer, et d'autre part, lui fournirait un outil pour transmettre l'information au patient, dans le cas d'un test positif, par exemple.

d. Distinction entre fréquences normalisées et fréquences naturelles

Il n'y a pas de consensus concernant l'effet facilitateur des fréquences par rapport aux probabilités dans le raisonnement bayésien. Par exemple, pour Evans, Handley, Perham, Over, & Thompson (2000) les fréquences naturelles ne sont en aucun cas facilitatrices. Au contraire, pour ces chercheurs, les participants négligent davantage le taux de base, quand le problème leur est présenté en fréquences, par rapport à un contexte en probabilités. Notons que ces auteurs utilisent, en plus du format probabiliste, deux conditions en fréquences : un format fréquentiste qu'ils nomment *facile* et un format qu'ils qualifient de *difficile.* Ils utilisent dans leur recherche le classique problème du test médical, d'après Casscells, Schoenberger, & Graboys (1978), repris par Cosmides & Tooby (1996) :

> Version probabiliste originale :
> « Si l'on teste une maladie dont la prévalence est de 1/1000, avec un test qui a un taux de faux positif de 5 %, quelle est la chance qu'une personne présentant un test positif ait réellement la maladie, d'après ce que vous savez à propos des signes et symptômes de la maladie ? »

Pour Evans, Handley, Perham, Over, & Thompson (2000), les termes de prévalence et de taux de vrais/faux positifs sont délicats à aborder. Ne pas les comprendre peut amener les participants à se tromper dans l'élaboration de leur réponse. Mais la difficulté serait alors sémantique et non liée à l'utilisation des probabilités. Ils ont donc présenté trois versions équivalentes de ce problème, en prenant soin d'épurer le style d'expression, et donc de faciliter la compréhension et la représentation du problème :

> Version fréquentiste facile :
> « Un américain sur 1000 a la maladie X. Un test a été développé pour savoir si les gens ont cette maladie. Chaque fois que le test est donné à une personne qui a la maladie, le test est positif. Mais quelquefois, le test est positif alors que la personne n'est pas malade. Plus précisément, parmi 1000 personnes n'ayant pas la maladie, 50 vont malgré tout avoir un test positif pour la maladie X. Un échantillon de 1000 participants à la loterie nationale est sélectionné et on leur fait passer le test médical. Combien de personnes ayant un test positif auront réellement la maladie X ? ______ of ______. »

> Version fréquentiste difficile :
> « Un américain sur 1000 a la maladie X. Un test a été développé pour savoir si les gens ont cette maladie. Chaque fois que le test est donné à une personne qui a la maladie, le test est positif. Mais quelquefois, le

> test est positif alors que la personne n'est pas malade. Plus précisément, parmi 20 personnes n'ayant pas la maladie, 1 va malgré tout avoir un test positif pour la maladie X. Un échantillon de 1000 participants à la loterie nationale est sélectionné et on leur fait passer le test médical. Combien de personnes ayant un test positif auront réellement la maladie X ? ______ of ______. »

Notons que cette version difficile diffère de la version facile par le taux de base. Dans cette version difficile, il est également donné sous forme de fréquences, mais sous la forme d'une fraction irréductible. Ceci nécessite donc une étape de plus par rapport à la version simple, version pour laquelle le taux de base est directement proposé sur le nombre de personnes au total, soit 1000.

> Version probabiliste :
> « Un américain sur 1000 a la maladie X. Un test a été développé pour savoir si les gens ont cette maladie. Chaque fois que le test est donné à une personne qui a la maladie, le test est positif. Mais quelquefois, le test est positif alors que la personne n'est pas malade. Plus précisément, 5 % de personnes n'ayant pas la maladie vont malgré tout avoir un test positif pour la maladie X. Un échantillon de 1000 participants à la loterie nationale est sélectionné et on leur fait passer le test médical. Combien de personnes ayant un test positif auront réellement la maladie X ? ______ of ______. »

Les auteurs ont proposé ces trois versions avec une double question : en fréquences, comme présenté ci-dessus, et en probabilités de la façon suivante :

> « Combien de personnes ayant un test positif auront réellement la maladie ? ______% »

Ils ont ensuite comparé les bonnes réponses (environ 2 % ou 1/51, respectivement selon que la question est posée en probabilités ou en fréquences). Lorsque la question est posée en termes fréquentistes (______sur _____), les auteurs trouvent une différence significative entre les trois conditions : Il y a davantage de réponses correctes quand l'énoncé est proposé en fréquences version difficile qu'en fréquences version facile, et il y a davantage de réponses correctes quand l'énoncé est proposé en fréquences version simple qu'il n'y en a en version probabiliste. Lorsque la question est posée en termes probabilistes (______%), les auteurs trouvent à nouveau une différence entre les nombres de réponses correctes des trois groupes, mais selon un ordre différent de difficulté : la version la mieux réussie est la version en fréquences simples, puis la version en probabilités et enfin la version en fréquences difficiles. Evans, Handley, Perham, Over, & Thompson (2000) confirment leur hypothèse selon laquelle une présentation en fréquences n'est pas toujours synonyme de performances accrues. Pour ces chercheurs, la relation entre le format de présentation, le format de la question et la valeur du taux de base sont des éléments à prendre en considération. Ils précisent que ces différences de réponses correctes sont en partie dues au fait que selon le format les participants seraient plus ou moins sensibles au biais de la négligence du taux de base (Cosmides & Tooby, 1996).

Dans une seconde partie de leur recherche, Evans, Handley, Perham, Over, & Thompson (2000) ont testé l'impact d'une nouvelle formulation de la question posée, sur les mêmes problèmes. Ils ont ainsi proposé la question sous forme de probabilité individuelle et sous forme de proportion :

> Probabilité individuelle :
> « Quelle est la chance qu'une personne avec un test positif soit réellement atteinte de la maladie ? ______% ».
>
> Proportion :
> « Quelle proportion de gens qui auront un test positif pour cette maladie, en seront réellement atteints ? _____% ».

Les auteurs ont montré que le fait de poser une question sous le format de proportion, à taux de base équivalent, donnaient des performances intermédiaires : le nombre de bonnes réponses est en proportion plus élevé qu'en probabilités, mais moins qu'en fréquences. Cet ordre est le même quant au biais de négligence du taux de base.

Ces travaux s'inscrivent dans un large débat sur l'impact des fréquences. Sont-elles facilitatrices pour traiter des problèmes de raisonnement probabiliste et sont-elles équivalentes à des probabilités ? Macchi & Mosconi (1998) pensent que les fréquences ne sont pas facilitatrices. Plus précisément, ces auteurs suggèrent que si les fréquences peuvent amener de meilleures performances, ceci n'est pas lié directement aux fréquences en tant que format sémantique mais à la simplification donnée par un tel format : le taux de base est explicite. En utilisant des fréquences qui ne donnent pas de façon transparente le taux de base, Macchi & Mosconi (1998) et Lewis & Keren (1999) ne reproduisent pas les résultats de Gigerenzer & Hoffrage (1995) et concluent à un effet non facilitateur des fréquences sur les probabilités. Johnson-Laird, Legrenzi, Girotto, Legrenzi, & Caverni (1999) tentent de clore le débat en concluant leurs travaux ainsi : les fréquences ne sont pas le garant d'un raisonnement bayésien correct. Girotto & Gonzalez (2001) vont également dans ce sens. Ils émettent une autre hypothèse pour rendre compte de meilleures performances en fréquences dans certaines études (Gigerenzer & Hoffrage, 1995). Pour Girotto & Gonzalez (2001), ceci s'explique par le fait d'avoir en fréquences des sous catégories, incluses dans les catégories de départ. Ceci est à rapprocher du taux de base, mais le lien n'est pas suggéré par ces auteurs. Pour Brase (2002), cette simplification entre les catégories et les sous catégories ne devrait pas permettre d'attribuer un effet facilitateur aux fréquences, en tant que telles. Cet effet, serait en quelque sorte, un bénéfice secondaire. Pour Macchi (2000), le fait de rencontrer des différences en termes de performances, selon que le poids d'une sous catégorie est plus ou moins important dans deux problèmes présentés en fréquences, montre que le format n'est pas un élément crucial. Fiedler, Brinkmann, Betsch, & Wild (2000), Macchi (2000) et Evans, Handley, Perham, Over, & Thompson (2000) ont utilisé des problèmes probabilistes avec des nombres purs absolus (Hoffrage, Gigerenzer, Krauss, & Martignon, 2002). Ceci signifie que le taux de base n'est pas donné dans la version probabiliste, puisque la probabilité de rencontrer le second élément est donnée dans l'absolu et n'est pas donnée en tant qu'inclusion du premier élément. Par exemple, Fiedler, Brinkmann, Betsch, & Wild (2000) utilisent la version suivante du problème de la mammographie en probabilités :

> « Cette étude contient les données de 1000 femmes. 99 % des femmes n'ont pas de cancer du sein, et 1 % en a un. Parmi les femmes sans cancer du sein, 10 % ont une mammographie positive et 90 % ont une mammographie négative. Et parmi les femmes ayant un cancer du sein, 80 % ont une mammographie positive et 20 % une mammographie négative. Quelle est la probabilité d'avoir un cancer du sein quand la mammographie est positive ? ».

Pour Hoffrage, Gigerenzer, Krauss, & Martignon (2002), cette version est la version probabiliste à utiliser. Elle ne donne pas de façon explicite le taux de base. Pour ces auteurs, il ne faut donc pas présenter le problème de la façon suivante :

> « Cette étude contient les données de 1000 femmes. 99 % des femmes n'ont pas de cancer du sein, et 1 % en a un. Parmi les 10 femmes sans cancer du sein, 10 % ont une mammographie positive et 90 % ont une mammographie négative. Et parmi les 990 femmes ayant un cancer du sein, 80 % ont une mammographie positive et 20 % une mammographie négative. Quelle est la probabilité d'avoir un cancer du sein quand la mammographie est positive ? ».

En effet, cette version donne deux informations supplémentaires : le nombre de femmes ayant et n'ayant pas de cancer du sein. Cet apport est, pour ces auteurs, réservé au format en fréquences.

Girotto & Gonzalez (2001) distinguent les probabilités des fréquences selon quatre critères :

- Le type d'information : le format fréquentiste renvoie à des fréquences observées et le format probabiliste à des probabilités.
- Le domaine de la tâche : le format fréquentiste demande aux participants de prédire une fréquence et le format probabiliste leur demande d'évaluer une probabilité conditionnelle.
- La forme de la question : le format fréquentiste invite les participants à réfléchir sur un quotient de fréquences, une proportion, tandis que le format probabiliste n'y invite pas.
- La structure de l'information : le format fréquentiste fournit aux participants les taux de base, ce qui n'est pas le cas avec le format probabiliste.

Cette distinction faite, Girotto & Gonzalez (2001) émettent l'hypothèse selon laquelle les performances en fréquences et en probabilités sont équivalentes. Ils ont donc proposé un moyen de faire varier les critères précédents indépendamment les uns des autres. Ainsi, le format n'est plus tout fréquentiste ou tout probabiliste, mais renvoyant plutôt, par exemple, à une structure d'information précise, ou un domaine de tâche particulier. Pour ce faire, ils ont utilisé une nouvelle terminologie : « le nombre de chances sur ». Ceci leur permet de confirmer le rôle des représentations mentales (Baratgin & Noveck, 2000 ; Macchi & Girotto, 1994 ; Nickerson, 1996) et le type de question (Girotto & Gonzalez, 2001 ; Hilton, 1995 ; Jones, Taylor, & Frisch, 1995 ; Macchi, 1995) dans l'inférence probabiliste. De plus, pour ces auteurs, les individus sont donc capables de raisonner avec des probabilités si le problème les oriente vers une réflexion en quotient et non en nombre isolé, et si le problème leur fournit de façon explicite les taux de base. Mais on peut se poser la question, si répondre à ces deux critères équivaut à donner de réelles probabilités. Pour Hoffrage, Gigerenzer, Krauss, & Martignon (2002), la réponse est non. Ils précisent que fournir le taux de base revient à confondre les deux formats.

Pour la suite de ces travaux nous nous référerons à ce dernier point de vue lorsque nous évoquerons des problèmes en probabilités.

Cette distinction entre probabilités et fréquences étant faite, il convient de revenir au débat concernant l'effet facilitateur des fréquences par rapport aux probabilités. Le fait que les fréquences permettent de meilleures performances reste discuté, malgré un regard harmonisé sur les probabilités. Hoffrage, Gigerenzer, Krauss, & Martignon (2002) proposent un argument pour clore le débat. Les différentes recherches ne se réfèrent pas à une seule et même définition de fréquences. Pour Hoffrage, Gigerenzer, Krauss, & Martignon (2002), les recherches dont les résultats ne soulignent pas un effet facilitateur des fréquences (Evans, Handley, Perham, Over, & Thompson, 2000 ; Girotto & Gonzalez, 2001 ; Johnson-Laird, Legrenzi, Girotto, Legrenzi & Caverni, 1999 ; Macchi, 2000) traitent de fréquences normalisées, tandis que l'approche fréquentiste stipule que l'effet facilitateur des fréquences concerne des fréquences naturelles. Quelle est la différence entre ces deux types de fréquences et comment cette distinction intervient sur les performances des participants en résolution de problème bayésien ? Conformément à Kleiter (1994), un échantillon naturel est un groupe représenté par la fréquence d'apparition d'une caractéristique dans l'environnement. L'apparition et la non apparition de deux caractéristiques peuvent se représenter sous forme d'un arbre de fréquences (cf. Figure 2.4.). Ces fréquences sont donc pour Hoffrage, Gigerenzer, Krauss, & Martignon (2002) des fréquences naturelles. Ces auteurs nous font remarquer qu'il n'est pas possible de représenter les fréquences de survenue des deux caractéristiques en arbre lorsque les fréquences sont des fréquences normalisées. Ces chercheurs illustrent ces différences avec une variante du classique problème de test médical, présenté sous les deux conditions de fréquences :

> Fréquences naturelles
> « Parmi 1000 patients, 40 sont infectés. Parmi ces 40 patients infectés, 30 ont un test positif. Parmi les 960 patients non infectés, 120 présentent également un test positif ».

> Fréquences normalisées
> « Parmi 1000 patients, 40 sont infectés. Parmi 1000 patients infectés, 750 ont un test positif. Parmi 1000 patients non infectés, 125 ont également un test positif ».

La représentation en arbres est aisée en fréquences naturelles car les données de la seconde caractéristique, ici le résultat du test, sont incluses dans les résultats des répartitions de la première caractéristique, ici le fait d'être malade ou non. Les taux de base sont donc explicites. *A contrario*, avec des fréquences normalisées, il n'est pas possible de faire un arbre de fréquences, comme le montre la figure suivante :

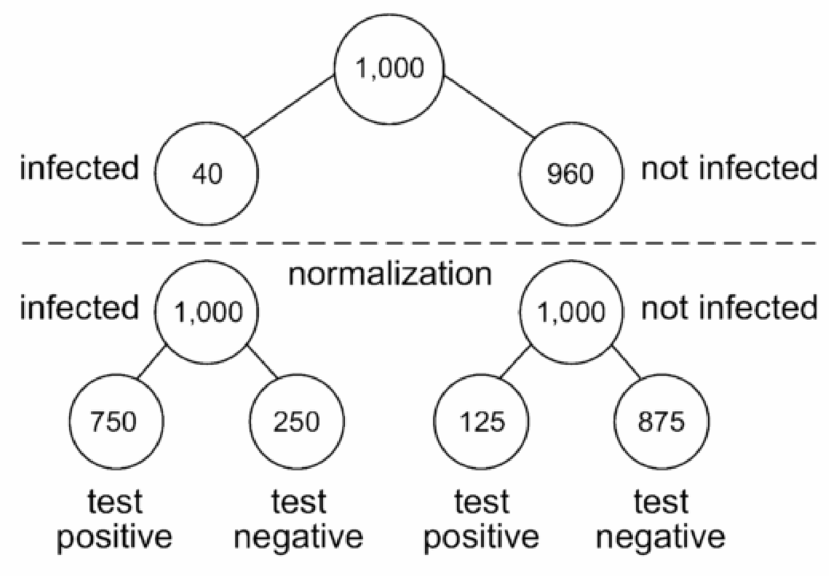


Figure 2.4. Transcription de fréquences normalisées en arbre (extraite de Hoffrage et al. 2002)

Cette distinction entre les deux types de fréquences permet d'éclairer les recherches précédemment citées. Les fréquences naturelles ont donc un effet facilitateur, et les études ne trouvant pas cette simplification computationnelle utilisent des fréquences normalisées et non des fréquences naturelles (Hoffrage, Gigerenzer, Krauss, & Martignon, 2002). Dans la suite de ces travaux, nous ferons référence aux fréquences naturelles et non aux fréquences normalisées, conformément aux distinctions qui viennent d'être apportées, chaque fois que évoquerons le concept de fréquences.

e. Une analyse qualitative des différentes stratégies grâce aux fréquences

Un autre intérêt de l'approche fréquentiste préconisée par Gigerenzer et son équipe est de rendre compte des réponses non bayésiennes. Ces auteurs ne se contentent pas d'une dichotomie réponses bayésiennes versus réponses non bayésiennes. Leur analyse n'est pas seulement quantitative, en nombre de bonnes réponses. Selon, les données utilisées par le participant, les auteurs décrivent les différentes réponses non bayésiennes, en fonction des données qui sont utilisées. Regardons de plus près un exemple de problèmes sous les deux formats qui ont été utilisés par Zhu & Gigerenzer (2006).

> Problème des maux de tête en probabilités
>
> « Dans un hôpital, 60 % des patients ont un rhume. Parmi ces 60 % de patients qui ont un rhume, 70 % ont des maux de tête. Parmi les 40 % de patients sans rhume, 30 % ont aussi des maux de tête. Tu rencontres les patients avec maux de tête, quelle est la probabilité qu'une personne ait un rhume ?
> _____________%. »
>
> Problème des maux de tête en fréquences
>
> « Dans un hôpital, 60 des 100 patients ont un rhume. Parmi ces 60 patients qui ont un rhume, 42 ont des maux de tête. Parmi les 40 patients qui n'ont pas de rhume, 12 ont aussi des maux de tête. Tu rencontres les patients avec maux de tête, combien d'entre eux ont un rhume ?
> _______ sur _______. »

Le second format -en fréquences- est facilitateur. Quand on utilise la règle de Bayes comme présentée précédemment avec le paradigme de la mammographie, on obtient pour ce problème des maux de tête un taux de 78 % (42 sur 54 en fréquences) de personnes présentant un rhume parmi les patients avec des maux de tête.

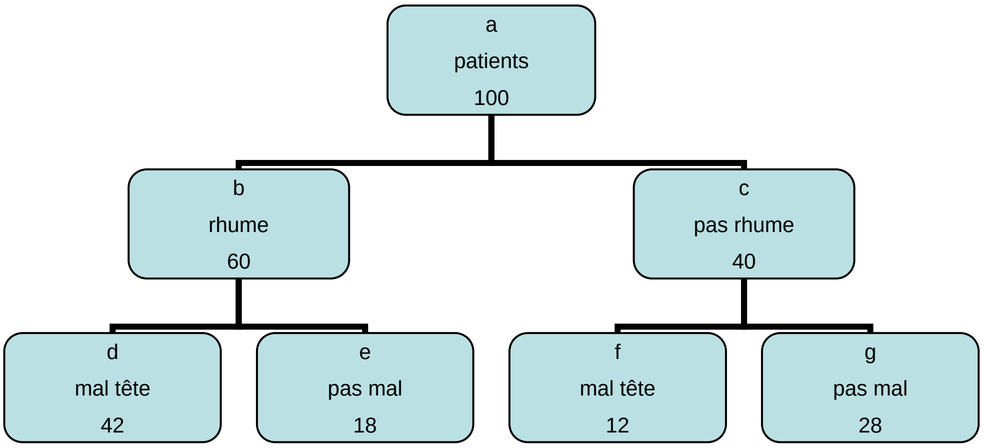


Figure 2.5. Représentation en arbre du problème des maux de tête en fréquences

La représentation en arbre (Figure 2.5.) nous offre une représentation simplifiée des données du problème. Le nombre de patients (a) est divisée en deux catégories selon le caractère rhume : ceux qui ont un rhume (b) et ceux qui n'en ont pas (c). Le nombre b est ensuite également divisé en deux sous catégories selon le critère maux de tête : ceux qui ont mal à la tête (d) et ceux qui ne l'ont pas (e). De la même manière, le nombre c est décliné en deux sous catégories : ceux qui ont des céphalées sans avoir de rhume (f) et les autres qui n'ont ni rhume ni mal de tête (g).

Selon les données utilisées par les participants dans le format fréquentiste, Gigerenzer & Hoffrage (1995) et Zhu & Gigerenzer (2006) ont proposé une taxonomie des différentes stratégies rencontrées. Ces stratégies sont en quelque sorte des biais, ou des sophismes car des utilisations erronées de la règle de bayes, sous sa forme fréquentiste, qui est pourtant facilitatrice (Gigerenzer & Hoffrage, 1995 ; Kleiter, 1994). Ces dernières sont résumées dans le tableau suivant :

Stratégies	Données utilisées	Définitions
Bayésienne	d / (d + f)	Jugement révisé conformément à la règle de Bayes
Pré-Bayésienne	b / (d + f)	Surestimation de l'intersection des deux événements positifs
Fonction de vraisemblance	d / b	Inversion des deux événements
Apparition conjointe	d / a	Ne considère que les événements positifs et ne tient pas compte de l'information non p (faux positifs)
Evidence	(d + f) / a	Ne tient plus compte de la probabilité *a priori* mais uniquement de la nouvelle information
Conservatisme	b / a	Ne révise pas son jugement

Tableau 2.1. Différentes stratégies selon les données utilisées

Ces stratégies sont ordonnées dans le tableau, de celle qui réfère à une utilisation correcte de la règle de Bayes à celle qui s'en éloigne le plus.

- Stratégie bayésienne

En fréquences, et avec l'arbre, il semble assez aisé de répondre à la question posée, à savoir : « Imagine que tu rencontres les patients avec des maux de tête, combien d'entre eux ont un rhume ? ». Les personnes ayant mal à la tête sont les personnes du groupe d et du groupe f. Ils sont donc 42 + 12 = 54. Parmi ces 54 personnes, seules celles du groupe d ont un rhume. Ainsi répondre correctement à la question posée, c'est-à-dire raisonner de façon bayésienne équivaut à répondre 42 sur 54. Toute fraction équivalente, telle 21 sur 27, sera bien évidement acceptée comme réponse correcte. Des fractions approchées qui ne seraient pas exactement égale à celles-ci ne sont pas considérées comme bonnes réponses, même si leur valeur approchée est quasi identique.

➢ Stratégie Pré-bayésienne

La stratégie pré-bayésienne est celle qui se rapproche le plus de la réponse attendue selon Zhu & Gigerenzer (2006). Elle a le bon dénominateur mais un numérateur approché. Les individus utilisant cette stratégie se focalisent sur les deux critères présents dans la question : le nombre de personnes ayant un rhume et le nombre de personnes ayant mal à la tête. Ceci les fait surestimer l'intersection des deux événements positifs. Ainsi, ils répondent le nombre de personnes ayant un rhume, soit le groupe b de 60 personnes, sur le nombre de personnes ayant des maux de tête, c'est-à-dire la somme des groupes d + f = 54. Leur réponse est donc 60 sur 54. Nous voulons apporter une précision ici quant au bien fondé de cette stratégie. Dans le type de répartition présent ici, à savoir beaucoup de personnes présentant le premier critère -le rhume- et beaucoup présentant le second critère dans les deux sous groupes -le mal de tête-, alors répondre de façon pré-bayésienne amène le participant à donner une fraction avec un numérateur plus élevé que le dénominateur. Ceci revient à dire qu'il y a plus de personnes avec telle caractéristique que de personnes présentent. A notre avis, dans ce cas de figure, une telle réponse n'est pas moins mauvaise qu'une des autres stratégies. Le taux de base a donc une incidence sur le bien fondé de cette stratégie (Evans, Handley, Perham, Over, & Thompson, 2000).

➢ La stratégie fonction de vraisemblance

Utiliser cette stratégie revient à inverser les deux caractéristiques évoquées dans la question. Dans le problème des maux de tête, la question porte sur le nombre de personnes ayant un rhume parmi les personnes ayant mal à la tête. Or cette fonction de vraisemblance revient à répondre d / b, soit le nombre de personnes ayant mal à la tête chez celles ayant un rhume. Comme vu précédemment, les individus inversent les probabilités conditionnelles (Villejoubert, Mandel, 2002) et confondent la probabilité conditionnelle de A sachant B -p(A/B)- et la probabilité d'avoir B sachant A -p(B/A)-, quand A est le fait d'avoir un rhume et B celui d'avoir mal à la tête. Notons que cette stratégie est dénommée « representative thinking » (Dawes, 1986 ; Koehler, 1996b) mais aussi la stratégie de Fisher (Gigerenzer & Hoffrage, 1995), en rapport avec le test statistique éponyme de significativité. Pour Fisher, la règle de Bayes est inutile en science (Gigerenzer, Switjink, Porter, Datson, Beatty, & Krüger, 1989).

- La stratégie d’apparition conjointe

Les individus qui utilisent la stratégie d’apparition conjointe ne considèrent que les événements positifs et ne tiennent pas compte de l’information non p (Klayman & Ha, 1987), c’est-à-dire les faux positifs, pour utiliser une terminologie adaptée au paradigme de la mammographie. Dans le cas présent, cette stratégie revient à considérer l’intersection des deux événements : mal de tête et rhume (groupe d), et la quantifier par rapport au nombre total de personnes, en d’autres termes le groupe a de 100 personnes en fréquences : d / a. Ceci n’est pas valide, conformément à la règle de Bayes car il y a aussi des personnes qui ont mal à la tête parmi celles n’ayant pas de rhume, de même que certaines personnes ont des mammographies positives sans avoir réellement de cancer du sein.

- La stratégie d’évidence

Utiliser cette stratégie revient, en quelque sorte, à réviser de façon excessive son jugement. En effet, les individus se focalisent sur le nouvel événement et ne tiennent plus compte de l’information *a priori*. Ils répondent (d + f) / a, soit le nombre de personnes qui ont mal à la tête par rapport à l’ensemble de la population. Le fait d’avoir un rhume n’est pas intégré dans leur réponse.

- La stratégie de conservatisme

Le conservatisme est l’opposé de la stratégie d’évidence. Il s’agit de ne pas réviser son jugement, comme nous l’avons vu précédemment. Les individus ne tiennent pas compte du nouvel élément et donnent comme réponse le nombre de personnes ayant un rhume par rapport à l’ensemble de l’échantillon : b / a. Nous présentons cette stratégie en dernier, car c’est celle qui est la plus éloignée du concept de révision probabiliste. Cette distance est encore plus patente si le problème est présenté en deux temps. On demande dans une première question quelle est la probabilité de rencontrer quelqu’un avec un rhume. Ensuite on donne une nouvelle information : les proportions de maux de tête chez les enrhumés et les autres. Enfin on demande quelle est la probabilité de rencontrer quelqu’un avec un rhume parmi les personnes ayant des maux de tête. Des personnes ayant un biais de conservatisme vont donner la même réponse aux deux questions. Rappelons qu'ils peuvent ajuster cette réponse mais pas autant que nécessaire, comme l'ont décrit Tversky & Kahneman (1973) avec l'heuristique d'ancrage et ajustement.

Cette approche qualitative qui propose différentes stratégies, ou différents biais et sophismes, selon les écoles, permet une analyse fine des réponses considérées comme non bayésiennes. Cela permet, par exemple, de montrer que les individus ne répondant pas de façon bayésienne ne forment pas une seule et unique catégorie, mais peuvent renvoyer à différents groupes ou profils cognitifs. Dans cette perspective, il est intéressant de voir que les réponses non bayésiennes peuvent être classées. Ceci apporte beaucoup à une approche développementale.

f. Approche développementale des performances bayésiennes

- Le modèle en ondes qui se chevauchent de Siegler

Si les enfants ne semblent pas réviser leur jugement, de façon quantitative, on peut rendre compte de l'utilisation de différentes stratégies en fonction de leur âge. Est-ce qu'il y a une correspondance entre un âge et une stratégie, ou, comme le suggère Siegler (1999, 2000), avec son modèle évolutionniste d'ondes qui se chevauchent. Ce modèle précise que différentes stratégies son présentes à un même niveau de développement. A un même moment, les enfants disposent de différentes stratégies. Ces différentes stratégies sont en compétition les unes avec les autres, et progressivement, les stratégies les plus performantes deviennent plus fréquentes au détriment des plus basiques et des moins efficaces. Dans le même temps que certaines stratégies ne sont plus utilisées, d'autres plus élaborées vont apparaître et devenir dominantes. Ce modèle a été testé dans différents domaines, tels la lecture, l'orthographe, la résolution de problèmes, la mémoire et le calcul d'aire d'un rectangle (Crowley & Siegler, 1993 ; Siegler, 1996 ; Siegler & Shrager, 1984). Pour Siegler (1987 ; Gigerenzer & Richter, 1990) un même enfant peut varier dans le choix d'une stratégie d'un jour à l'autre ou d'un problème à un autre. Le choix parmi les différentes stratégies dépend de la difficulté du problème. Ainsi, pour un problème rapide et simple à traiter, les enfants choisiraient une stratégie rapide, comme la récupération dans le cas d'une addition simple et connue de l'enfant. Avec le temps et la familiarité, les enfants vont laisser de côté certaines stratégies peu efficaces, comme deviner, dans le cadre d'un problème d'addition de nombres (Siegler, 1987).

➢ Stratégies utilisées selon le développement

Le modèle de Siegler en ondes qui se chevauchent s'adapte-t-il au raisonnement probabiliste ? Comment évolue l'utilisation des stratégies non bayésiennes chez les enfants scolarisés ? Zhu & Gigerenzer (2006) ont analysé les performances bayésiennes de jeunes enfants chinois de cours moyen première et deuxième année, et de sixième, tout au moins les classes équivalentes dans le système chinois. Ces auteurs ont dans un premier temps montré que dans ces classes, les enfants ne répondent à aucun problème présenté en probabilités conditionnelles. Lorsque le problème est présenté en fréquences, les trois groupes scolaires répondent de façon bayésienne respectivement à 18,7 %, 39 % et 53,5 % des problèmes. Ces résultats montrent un effet plancher en probabilités et une amélioration des performances en fréquences. La stratégie bayésienne est donc de plus en plus utilisée au cours de ces trois niveaux scolaires. Mais que donne l'analyse des réponses non bayésiennes ? Zhu & Gigerenzer (2006) ont montré des modifications dans les taux d'utilisation des stratégies non bayésiennes. Plus précisément, le conservatisme est de moins en moins utilisé, l'évidence est davantage usitée en CM 2 qu'en CM 1 mais disparait en sixième. La fonction de vraisemblance n'est utilisée que par les enfants de CM 1. Lücking (2004) a montré un retard chez les enfants allemands, par rapport aux enfants chinois. Pour cet auteur, les enfants allemands ne sont pas capables si tôt que leurs homologues chinois de raisonner de façon probabiliste, même en fréquences naturelles. Les enfants chinois présentent souvent des performances mathématiques supérieures aux jeunes allemands, et semblent souvent plus motivés qu'eux (Artelt, Baumert, Klieme, Neubrand, Prenzel, Schiefele, Schneider, Schümer, Stanat, Tillmann, & Weiß, 2001 ; Stern, Rode, Ge, & Zhu, 2001).

➢ L'inférence bayésienne selon le développement

Ces résultats montrent que les enfants peuvent être bayésiens si le problème leur est présenté en fréquences naturelles. D'autres recherches vont encore plus loin. Par exemple, certaines études suggèrent que très jeunes les enfants ont des prédispositions pour traiter les nombres. Les enfants sont, par exemple, sensibles aux changements de nombre dans une collection d'objets visuels (Antell & Keating, 1983 ; Starkey & Cooper, 1980). Dès l'âge de

six mois, ils semblent capables de se rendre compte d'un ajout ou du retrait d'un élément dans une courte série d'objets. Le temps de fixation des objets en témoigne (Wynn, 1992). Les fractions sont néanmoins plus délicates à aborder pour les enfants (Brase, Cosmides & Tooby, 1998) car ils ne peuvent se les représenter sur leurs doigts (Butterworth, 1999 ; 2001). Nous pouvons donc nous interroger sur les compétences proto-bayésiennes des enfants. Peuvent-ils de façon précoce raisonner de façon bayésienne, ou émettre des inférences bayésiennes ?
Pour Zhu & Gigerenzer (2006) les enfants ont, en fréquences naturelles, de meilleures performances que les adultes en probabilités conditionnelles. Si ceci nous permet de défendre la thèse selon laquelle les individus sont, plus précocement que ce qui est avancé dans la littérature, capables de raisonner de façon bayésienne, ceci ne nous informe nullement sur des compétences d'inférences des enfants. Des études ont montré que les enfants étaient capables d'inférences causales entre des objets et des événements. Plus précisément, Sobel, Tenenbaum & Gopnik (2004) se sont intéressés aux liens entre deux événements que pouvaient reconnaître des enfants de trois et quatre ans. Ils ont proposé à des enfants des stimulus visuels et sonores. Les enfants sont installés dans une machine à stimulus et quand certains objets apparaissent, des sons ou des lumières apparaissent. Pour ces auteurs, le fait que les enfants retiennent le lien entre l'apparition des objets et celles du son ou de la lumière ne réfère pas à la capacité de faire des associations mais plutôt des inférences. La manipulation du délai entre l'apparition des objets et la survenue du stimulus et la différence entre les stimulus proposés encourage ces auteurs à penser que les enfants sont capables d'inférences causales qui s'appuient sur une structure d'apprentissage des mécanismes bayésiens. En effet, dans leur troisième expérimentation, les auteurs utilisent des objets plus ou moins ambigus et éloignés des objets de départ. Pour eux, si les enfants ne faisaient que des associations, ces dernières n'auraient plus lieux avec des stimulus différents. Par contre, s'ils utilisaient des inférences causales, ils élargiraient leurs inférences à d'autres éléments que ceux appris. Ces auteurs concluent à la capacité de faire des inférences causales à propos de données ambigües à l'âge de quatre ans mais pas à l'âge de trois ans. En limitant le nombre de stimulus associés, Beckers, Vandorpe, Debeys & De Houwer (2009) étendent ces résultats aux enfants âgés de trois ans.

Cette idée de structure bayésienne est également présente dans la littérature d'un autre domaine que le raisonnement : le langage. Yahya, Mahmod & Rahman Ramli (2010) ont proposé un modèle informatique bayésien rendant compte de la probabilité de rencontrer tel ou tel mot dans une phrase, selon les mots la constituant.

g. La théorie « Cognitive-Experiential Self-Theory » ou CEST de Epstein

Les modèles présentés ci-avant laissent entrevoir une précocité dans les performances de raisonnement, et dans le même ordre d'idées une structure intuitive permettant d'avoir des intuitions, ce qui nous différencie de l'ordinateur. Ces intuitions peuvent nous permettre de gagner du temps, ou d'économiser nos ressources cognitives, mais elles peuvent aussi nous induire en erreurs. Le modèle de la « Cognitive-Experiential Self-Theory » ou CEST d'Epstein est une explication possible de ces éléments. Le jugement, la prise de décision et le raisonnement sont envisagés selon un double processus : un mode de raisonnement logique rationnel, mathématique et un mode de raisonnement plus heuristique, lié à des approximations, à une simplification du problème (Epstein, Lipson, Holstein, & Huh, 1992 ; Evans et Over, 1996 ; Sloman, 1996 ; Stanovich & West, 2000 ; Stanovich & West 2008). Selon Epstein (1983, 1994, 1998), nous traitons l'information selon deux modes de traitement indépendants et fonctionnant conjointement : le mode *logique* ou rationnel, et le mode *heuristique* ou « expérientiel » qui opèrent selon des principes différents.

Le *mode logique* est délibéré et analytique, majoritairement verbal, conscient et pourrait fonctionner à partir de modèles mentaux et de l'application de règles de la logique formelle. Ce mode de traitement de l'information est coûteux en attention, lent, contrôlé et nécessite une analyse complexe, indépendante des émotions et entraînant une action (décision) différée. Il permet donc de juger et de raisonner logiquement sur des situations, mais est contraint par les limites de capacité de traitement du système cognitif humain.

Le *mode heuristique*, au contraire, fonctionne de façon automatique et préconsciente ; il est intuitif, rapide, associatif et holistique. Ce mode de traitement est basé sur des heuristiques et permet de se représenter les événements sous la forme d'exemples concrets construits selon les expériences passées avec connotations émotionnelles. C'est un mode de traitement rapide qui aboutit à une décision et qui permet l'action immédiate. Bien que ce mode de traitement expérientiel soit privilégié dans la vie quotidienne, il est possible de passer à un mode rationnel, plus analytique, quand le contexte le demande. Notre comportement est influencé conjointement par ces deux modes de traitement sur un continuum qui reflète leur influence mutuelle. Le mode de traitement heuristique est souvent dévalorisé par rapport au mode logique concernant son efficacité pour émettre des jugements et raisonnements probabilistes.

Ce modèle nous permet d'adopter une position différente par rapport aux théories précédentes et semble être une piste à propos d'un raisonnement bayésien intuitif.

h. Le fonctionnement exécutif

Au niveau développemental, la non utilisation du mode logique, le système 2, peut signifier deux choses : la non acquisition du raisonnement logique proprement dit -cette idée renvoie à une approche structuraliste de type piagétien- ou à une difficulté à mettre en œuvre des structures logiques existantes -ceci réfère davantage à une approche fonctionnaliste-. En effet, les résultats expérimentaux montrent que le développement cognitif n'est pas linéaire et que la performance de l'enfant n'est pas sous la seule dépendance de ses acquisitions logiques. Houdé (1995) explique les variations de performances intra individuelles, les décalages et « les apparentes régressions cognitives » par un déficit fonctionnel : le sujet qui prend en compte des facteurs non pertinents pour la tâche à résoudre va commettre un biais de raisonnement qui n'est pas dû à un défaut de logique ni de rationalité. Houdé & Joyces (1995) pensent que certaines erreurs de raisonnement ne témoignent pas forcément d'un défaut de rationalité mais d'un défaut de « programmation exécutive de l'inhibition ». La mise en œuvre du raisonnement logique correct est d'autant plus difficile que la tâche nécessite l'inhibition d'une donnée prégnante ou d'un automatisme cognitif. Cet automatisme cognitif, appelé « schème dangereux » par Houdé (1995), interfère avec le schème logique. Ceci peut engendrer une apparente irrationalité. Pour cet auteur, « se développer c'est apprendre à inhiber ». Un défaut de mise en œuvre des mécanismes inhibiteurs pourrait donc expliquer des différences entre un enfant précocement rationnel et un adulte irrationnel. Plus précisément, un enfant peut être précocement rationnel lorsque la tâche à laquelle il est soumis ne présente pas de schème non pertinent dangereux, et un adulte peut sembler irrationnel car se focalise sur un schème dangereux qu'il ne parvient pas à inhiber.

L'inhibition fait partie des processus exécutifs largement étudiés depuis une vingtaine d'années. Miyake, Friedman, Emerson, Witzki, Howerter, & Wager (2000) ont identifié auprès de jeunes adultes trois processus exécutifs de base : la flexibilité, la remise à jour en mémoire de travail et l'inhibition. Le processus de mise à jour en mémoire de travail est considéré par plusieurs auteurs comme une fonction exécutive importante (Miyake, Friedman, Emerson, Witzki, Howerter, & Wager, 2000 ; Shimamura, 2000). Cette fonction consiste à modifier continuellement le contenu de la mémoire de travail en fonction des nouvelles informations entrantes. La remise à jour intervient dans un grand nombre d'activités du quotidien, telles que l'apprentissage et l'organisation d'informations récemment acquises.

La flexibilité cognitive représente un aspect important du contrôle exécutif (Norman et Shallice, 1986), et peut être définie comme la capacité d'un sujet à changer de critère lors de l'analyse d'un stimulus, ou de point de vue lors de l'analyse d'un problème. La flexibilité attentionnelle, encore appelée « shifting » ou « switching », reflète selon Camus (1996) la capacité à changer entre les réponses alternatives afin d'assurer un contrôle cognitif flexible, et répond aux diverses réorientations de la focalisation attentionnelle. Ces réorientations continuelles peuvent présenter des effets bénéfiques pour la conduite, comme le fait de pouvoir faire face avec rapidité à une situation inattendue ou modifier un plan d'action lorsqu'il se révèle inadapté à la situation, mais peuvent aussi présenter des effets négatifs, telle la distraction, la désorganisation de la conduite provenant de l'incapacité à maintenir durablement la cohérence d'un plan d'action.

Le processus d'inhibition est également une fonction exécutive considérée comme importante par de nombreux auteurs (Hasher et Zacks, 1988 ; Baddeley, 1996). Même si les résultats semblent montrer qu'il est difficile de considérer ce mécanisme comme étant homogène puisqu'il existerait de multiples processus inhibiteurs. (Kok, 1999 ; Nigg, 2000), l'inhibition réfère couramment à la capacité de supprimer une réponse dominante, prégnante, mais réfère également au contrôle des interférences, au contrôle moteur et émotionnel.

Lehto, Juujaervi, Kooistra, & Pulkkinen (2003) ont observé la même structure factorielle chez des enfants âgés de 8 à 13 ans que celle de Miyake, Friedman, Emerson, Witzki, Howerter, & Wager (2000) chez de jeunes adultes.

La vitesse de traitement est souvent étudiée en parallèle des fonctions exécutives car elle est considérée comme une composante importante dans toute activité cognitive (Salthouse, 1996).

Dans une étude développementale récente effectuée auprès de 3400 sujets âgés de 5 ans à 93 ans, Salthouse & Davis (2006) observent que le pattern de corrélations entre des variables exécutives et des variables cognitives n'est pas le même selon l'âge du sujet. Par exemple, les performances à la tâche complexe du Wisconsin sont corrélées à l'intelligence fluide à tous les âges, mais sont aussi corrélées à l'intelligence cristallisée dans l'échantillon d'enfants et à la mémoire de travail dans l'échantillon d'adultes jeunes. Une même variable cognitive peut donc refléter différents aspects du fonctionnement cognitif selon l'âge. Les résultats expérimentaux de Senn, Espy, Kaufmann (2004) auprès d'enfants de 3 à 6 ans appuient cette conclusion.

Dans une récente revue de questions concernant le développement des fonctions exécutives, Best, Miller et Jones (2009) exposent les différents travaux expérimentaux qui montrent que les trajectoires développementales des fonctions exécutives sont différentes. En effet, la mémoire de travail se développe linéairement de l'âge de 4 ans à 15 ans tandis que le développement des capacités de flexibilité semble observer des paliers de croissance particulièrement entre 5 et 6 ans (Luciana & Nelson, 1998) ainsi qu'entre 11 et 15 ans (Huzinga, Dolan, Van der Molen, 2006). Les capacités d'inhibition augmentent dès le plus jeune âge avec une accélération entre 5 ans et 8 ans, même si l'inhibition continue à se développer pendant l'adolescence et même chez le jeune adulte (Romine & Reynolds, 2005)

Asato, Sweeney, Luna (2006) montrent que le développement de la résolution de problèmes complexes (Tower of London) entre l'âge de 8 ans et 30 ans est essentiellement lié aux capacités d'inhibition des sujets, même si la mémoire de travail est également corrélée aux performances. Welsh, Satterlee-Cartmell, and Stine (1999) observent également une corrélation modérée entre l'inhibition, la mémoire de travail et cette tâche complexe chez de jeunes adultes.

Demetriou & Kasi (2001) étudient les relations entre les processus cognitifs et métacognitifs. Ces auteurs suggèrent qu'il existe trois niveaux d'organisation reliés entre eux : Le premier niveau est celui des processus de base nécessaires à la résolution de la tâche ; par exemple, les capacités numériques et logiques pour une tâche d'arithmétique. Le second niveau concerne les processus cognitifs qui représentent et traitent l'information, ainsi que leurs interactions ; ce niveau concerne les structures de raisonnement, de mémoire, les processus attentionnels... Le troisième niveau ne réfère pas directement aux caractéristiques de l'organisation cognitive, mais plutôt aux capacités de gérer et d'utiliser de façon consciente et stratégiquement efficace les ressources cognitives disponibles ; ce niveau concerne le système métacognitif et les capacités exécutives.

Chapitre 3 - Problématique

Les recherches sur le raisonnement probabiliste se sont inscrites dans un contexte plus large. Leur but était de rendre compte de la rationalité / irrationalité de l'être humain. Avec l'émergence des sciences cognitives, la règle de Bayes s'est très vite imposée en tant que modèle normatif du raisonnement probabiliste. S'en écarter serait en quelque sorte un argument, ci ce n'est une preuve, en faveur de l'irrationalité humaine.

L'école *heuristiques et biais,* emmenée par Tversky & Kahneman, a marqué de ses nombreux travaux la littérature et tenté de mettre en exergue les nombreux écarts à la norme rencontrés en situation de raisonnement probabiliste. Pour ce faire, les auteurs de cette école ont proposé de nombreux paradigmes, plus proches de la vie quotidienne des participants et moins artificiels. Ces paradigmes, qu'aujourd'hui on qualifie d'écologiques en rapport à leur proximité avec la vie quotidienne, sont toujours usités actuellement. Pour cette école, l'être humain utilise des stratégies de simplification de son environnement, afin de faciliter le rapport qu'il entretient avec ce dernier. Ces stratégies intuitives visant à appréhender le monde de façon plus immédiate et spontanée sont appelées des heuristiques. Mais si ces heuristiques permettent sous certaines conditions de s'approcher de la réalité, elles faussent souvent la représentation que l'individu se fait du réel, et biaisent son rapport à l'environnement. Ces heuristiques ne sont pas l'exclusivité de l'homme de la rue, celui que l'on qualifie de naïf ayant une rationalité limitée. En effet, dans de nombreuses études, les auteurs rendent compte d'utilisations non adéquates des probabilités par des sujets censés être experts. Ceci tant dans leur utilisation que dans leur interprétation. Il est également fréquent de rencontrer dans la littérature des professionnels qui manipulent les probabilités sans savoir communiquer dessus.

Gigerenzer interpelle l'école *heuristiques et biais*, et évoque un paradoxe pour développer une théorie plus optimiste quant à la rationalité humaine. Le paradoxe qu'il énonce est le suivant : « Pourquoi des enfants de Genève raisonnement en accord avec les lois des probabilités, alors que des étudiants de Stanford semblent ne pas le faire ? » (Zhu & Gigerenzer, 2006, p. 287). Les enfants genevois sont cités en référence aux travaux de Piaget.

Pour ce dernier, dès l'âge de douze ans les enfants comprennent les probabilités et sont capables de les utiliser de façon combinatoire (Piaget & Inhelder, 1951 ; 1966 ; 1975). S'il pose la question, Gigerenzer propose dans le même temps une réponse. Les divergences entre les résultats quant aux performances bayésiennes rencontrées dans la littérature viennent du format de présentation. Dans les études, le raisonnement probabiliste est testé avec des probabilités, qu'il faut intégrer dans la règle de Bayes. Gigerenzer propose une alternative à ce format : les fréquences naturelles. Pour cet auteur, les fréquences naturelles facilitent le raisonnement probabiliste, et ce pour deux raisons essentielles. La première est le fait que dans l'environnement, l'être humain, comme l'animal d'ailleurs, rencontre des fréquences, et au cours de son évolution, il a appris à les évaluer (Humes, 1739 ; 1951). Pour Hascher & Zacks (1984) l'homme enregistre naturellement les fréquences des événements passés. De plus, selon Gould (1992), l'homme n'est pas fait pour traiter des probabilités. La seconde raison de cette facilitation par les fréquences du raisonnement probabiliste vient du format de fréquences elles-mêmes. Elles sont en effet plus riches en information et fournissent aux participants ce que les chercheurs nomment le taux de base. Ce dernier n'est pas présent dans un problème en probabilités. Ainsi, les fréquences simplifient la règle de Bayes et son utilisation devient plus aisée : il ne faut plus multiplier des probabilités entre elles, mais il suffit d'additionner deux fréquences. Gigerenzer apporte une précision : cette simplification computationnelle est possible grâce aux fréquences naturelles, et non avec les fréquences normalisées (Gigerenzer & Hoffrage, 1995 ; Hoffrage, Gigerenzer, Krauss, & Martignon, 2002).

Gigerenzer devient alors le représentant de l'approche fréquentiste. D'après cette rationalité écologique, les fréquences permettent à l'individu de raisonner de façon bayésienne. Quand ce n'est pas le cas, c'est-à-dire que les individus ne donnent pas la réponse attendue, il est possible en fréquences de classer les réponses non bayésiennes en différentes stratégies. Ceci permet d'affiner l'analyse du raisonnement bayésien dans le sens où l'on ne se contente pas d'une dichotomie réponses bayésiennes versus non bayésiennes. Ces différentes stratégies évoluent avec l'âge (Zhu & Gigenrezer, 2006). Il n'y a toutefois pas de correspondance terme à terme entre le niveau scolaire ou l'âge des enfants et l'utilisation de telle ou telle stratégie. L'utilisation des stratégies évolue davantage de manière continue. Comme l'évoque Siegler (1999, 2000) avec son modèle en ondes qui se chevauchent, plusieurs stratégies sont présentes à un même niveau de développement. Certains vont être de plus en plus utilisées avec l'âge, c'est le cas de la stratégie bayésienne et pré-bayésienne (Zhu

& Gigerenzer, 2006), tandis que d'autres vont ne plus l'être du tout, c'est le cas de la stratégie d'évidence. Lücking (2004) a montré que ces résultats étaient différents selon la culture. Ainsi, les enfants chinois sont plus précoces que les enfants allemands.

L'objectif de cette présente recherche est de rendre compte des performances en raisonnement bayésien des enfants français. Plus précisément, le but de ces travaux est, d'une part, de décrire l'émergence du raisonnement probabiliste sur une population de scolaires français, encore jamais étudiée, et d'autre part, d'appréhender les facteurs ayant un effet sur les perfomances. Ces travaux s'inscrivant dans la lignée des travaux de l'approche fréquentiste de Gigerenzer et se référant au modèle en ondes qui se chevauchent de Siegler, nous postulons que les fréquences naturelles vont faciliter le raisonnement bayésien. Ceci se manifeste de deux manières : une augmentation des réponses bayésiennes avec l'avancée en scolarité, et dans le même temps, des changements dans les stratégies utilisées.

Ces travaux se déclinent donc en cinq expérimentations successives.

L'objectif de la première expérimentation est de vérifier l'impact des fréquences naturelles sur les performances tant quantitatives que qualitatives en raisonnement bayésien au cours des années de collège et de lycée d'enfants et adolescents français. Nous postulons que plusieurs stratégies vont renvoyer à un même niveau scolaire. De plus, nous pensons que sur cette population, les performances vont s'améliorer avec l'avancée dans la scolarité. Plus précisément nous pensons pouvoir faire un double constat : les stratégies les plus éloignées de la règle de Bayes vont disparaître avec l'avancée scolaire, et en parallèle, nous allons rencontrer de plus en plus de réponses bayésiennes. Les problèmes proposés sont en fréquences naturelles et en probabilités conditionnelles, et les réponses attendues de la part des participants sont d'une forme correspondante au format de présentation : réponse en fréquences lorsque le problème est présenté en fréquences naturelles et réponse en pourcentages lorsque le problème est présenté en probabilités conditionnelles. Nous pensons obtenir des résultats plus proches des enfants allemands que de ceux des enfants chinois, en termes de précocité.

Puisque les fréquences simplifient l'utilisation de la règle Bayes, nous pensons que cette formule présente des difficultés mathématiques qui lui sont intrinsèques, et ce tant dans sa compréhension et dans sa représentation, que dans son utilisation. Nous postulons que le format de la réponse attendue a, comme le format de présentation, un impact sur les performances des participants. La seconde expérimentation propose donc aux participants de répondre aux mêmes problèmes bayésiens, toujours présentés sous les deux formats que sont les fréquences naturelles et les probabilités conditionnelles, mais cette fois la réponse qu'ils doivent donner n'est plus chiffrée directement. En effet, les participants doivent se situer sur un continuum libre, en portant une croix. Nous postulons que l'effet facilitateur des fréquences sur les probabilités sera maintenu. Ceci signifierait que les fréquences ne sont pas facilitatrices uniquement au niveau de l'utilisation de la règle de Bayes, mais qu'en plus de la simplification computationnelle, elles permettent une représentation plus aisée des problèmes bayésiens. Ceci est en accord avec l'idée de la théorie de Epstein : les individus sont capables de traiter les problèmes de raisonnement bayésien selon un mode heuristique, sans utiliser directement la règle de bayes, comme ce serait le cas avec le mode logique. Nous supposons, que pour les deux formats, les performances s'amélioreront avec l'avancée scolaire.

La troisième expérimentation va plus loin dans l'abstraction du chiffre. En effet, les participants doivent à nouveau répondre en se situant sur un continuum. Cette fois, par contre les problèmes présentés ne sont ni en fréquences naturelles ni en probabilités conditionnelles. Les problèmes conservent leur structure, toutefois, les nombres sont supprimés et remplacés par des quantités exprimées de façon littérale. Chaque probabilité/fréquence est donc remplacé par un adjectif numérique : soit peu soit beaucoup. Cette méthodologie a pour objectif d'appréhender le rôle accordé par les participants à chacune des données fournies dans le problème bayésien énoncé. Nous postulons que l'avancée du niveau scolaire aura un impact sur le poids attribué par les individus aux différentes données de l'énoncé. Plus précisément, nous pensons que les profils cognitifs que nous trouverons seront différents selon le niveau scolaire. Par exemple, nous pensons rencontrer des profils disjonctifs chez les participants en début de collège. Ces derniers peuvent se focaliser sur la première donnée de l'énoncé et ne tenir compte que de celle-ci pour estimer la probabilité demandée. Ceci est à mettre en relation avec les stratégies -ici la stratégie d'évidence d'après l'exemple- décrites dans la première expérimentation.

A la suite de cette abstraction en trois temps du nombre dans l'étude du raisonnement bayésien, la quatrième expérimentation vise à appréhender le poids du niveau scolaire dans les performances bayésiennes. Ainsi, les notes en mathématiques sont utilisées comme prédicteurs des performances des participants en raisonnement probabiliste. Les fonctions exécutives de bas niveau et la vitesse de traitement le sont aussi, comme cela a déjà été fait pour le raisonnement déductif et le raisonnement inductif. Nous postulons que le niveau en mathématiques est le prédicteur le plus précis, au regard de la difficulté mathématique intrinsèque de la formule, et des performances intuitives décrites dans les expérimentations 2 et 3. En d'autres termes, le raisonnement bayésien serait à aborder dans une perspective structuraliste.

Enfin, la cinquième expérimentation est une étude longitudinale des performances bayésiennes. Ainsi, les participants de la première expérimentation sont rencontrés dix-neuf mois plus tard, et sont soumis aux mêmes problèmes, présentés dans les deux formats : fréquences naturelles et probabilités conditionnelles. Nous postulons que cette approche longitudinale va mettre en avant une évolution dans l'utilisation des stratégies des enfants et adolescents rencontrés précédemment. Cette analyse longitudinale, couplée à l'analyse transversale de la première expérimentation va nous permettre d'avoir une description plus fine de l'évolution des performances bayésiennes des collégiens et lycéens français.

Deuxième partie

Partie expérimentale

Chapitre 4 - Impact du format de présentation sur les performances bayésiennes : analyse quantitative et qualitative selon le niveau scolaire

I. Introduction

Le but de cette première expérimentation est de vérifier la thèse de l'approche fréquentiste selon laquelle les fréquences naturelles sont facilitatrices par rapport aux probabilités conditionnelles (Gigerenzer & Hoffrage, 1995, par exemple), sur une population de jeunes français scolarisés en collège et en lycée. Zhu & Gigerenzer (2006) ont montré que des enfants chinois étaient capables de raisonner de façon probabiliste dès la classe de CM1, et que leurs performances augmentaient de manière significative en CM2 et encore davantage en sixième. Lücking (2004) a trouvé des résultats similaires sur une population de jeunes enfants allemands mais avec toutefois un décalage : les enfants chinois sont plus précocement aptes au raisonnement bayésien, tant quantitativement -nombre de bonnes réponses- que qualitativement -utilisation de stratégies de plus en plus proches de la stratégie bayésienne-. Le décalage qu'elle trouve est d'environ une année scolaire. Notons que plusieurs stratégies sont présentes chez des enfants d'une même classe, conformément au modèle d'ondes qui se chevauchent de Siegler (1999 ; 2000).

Aucuns travaux portant sur l'effet du format de présentation des problèmes bayésiens, à notre connaissance, n'ont été menés sur une population de jeunes français. De plus, les recherches menées par Zhu & Gigerenzer (2006) comparent des enfants de cours moyen et de sixième à des adultes. Or leurs résultats montrent un effet plancher des performances de leurs jeunes participants avec le format probabiliste.

Notre étude a donc plusieurs objectifs.

Le premier est de rendre compte des performances de jeunes français. En effet, la littérature indique un effet inter culturel sur le développement des performances bayésienns. Nous postulons que nos résultats seront plus proches de ceux trouvés auprès d'enfants allemands que de ceux recueillis auprès d'enfants chinois, notre système éducatif se rapprochant davantage de celui allemand que de celui chinois.

Le second objectif est d'appréhender les performances bayésiennes chez des adolescents scolarisés en collège et en lycée, afin de connaître à quel moment cette population commence à traiter de façon bayésienne des problèmes présentés en probabilités conditionnelles. Nous faisons l'hypothèse que cela arrive au début des années de lycée, en accord avec les programmes scolaires qui ne les abordent qu'en classe de seconde. Ceci apportera des précisions quant à l'effet plancher décrit par Zhu & Gigenrenzer (2006).

Le troisième objectif est de décrire sur une plus large étendue les changements dans les choix de stratégies. Nous pensons que certaines stratégies -les moins élaborées- seront absentes chez les lycéens, et qu'ainsi, les variabilités intra individuelle et interindividuelle vont diminuer avec l'avancée dans la scolarité.

II. Méthode

A. Participants

120 participants ont participé à cette expérience. Ils sont répartis en six groupes de 20. Les participants sont des élèves d'un collège et d'un lycée publics de la région Centre Val de Loire. Ces deux établissements sont considérés comme représentatifs des établissements de l'académie, au regard de critère objectifs. Les principaux des deux établissements nous ont en effet procuré les résultats des évaluations de leur établissement respectif. Le collège se situe légèrement au dessus de la moyenne concernant les tests de connaissances passés à l'entrée en sixième et le taux de réussite au brevet. Le lycée est, de la même manière, un peu au dessus des moyennes académique et nationale concernant le taux de réussite au baccalauréat.
Le fait de recruter les participants directement dans les établissements présente différents avantages : cela nous permet de contrôler le type d'établissement fréquenté, d'avoir accès aux notes en mathématiques (cf. chapitre 7) et de savoir où retrouver les participants plusieurs mois plus tard (cf. chapitre 8).

Pour tous les participants, nous avons obtenu un accord parental.

Les 120 participants se répartissent en six groupes 20 par niveau scolaire, soit 20 élèves de sixième, 20 de cinquième, 20 de quatrième, 20 de troisième, 20 de seconde générale et 20 de première économique et sociale (cette section nous semblant être la plus générale et la plus neutre par rapport aux mathématiques, en comparaison aux premières littéraires et scientifiques). Précisons que ces élèves ont été rencontrés en septembre, soit au début de leur année scolaire. Leur niveau est donc davantage celui de l'année précédente révolue que celle de l'année en cours, au regard des programmes scolaires.

Les participants sont recrutés sur la base de volontariat pendant des heures libres dans leur emploi du temps.

Pour chaque groupe, la parité est respectée. La variable genre est ainsi contrôlée, même si la littérature ne fait état du genre des participants.

Le tableau 4.1, à suivre, récapitule les différentes caractéristiques des groupes de cette expérimentation.

	Classes					
	Sixième	Cinquième	Quatrième	Troisième	Seconde	Première
Effectifs (Garçons ; Filles)	20 (10 ; 10)	20 (11 ; 9)	20 (10 ; 10)	20 (10 ; 10)	20 (10 ; 10)	20 (8 ; 12)
Regroupements de classes	6ème et 5ème		4ème et 3ème		2nde et 1ère	

Tableau 4.1. Répartition des effectifs selon le niveau scolaire.

Dans la suite des travaux, nous regrouperons les classes par deux. En effet, nous n'avons pas obtenus de différences significatives en termes de performances entre les élèves de sixième et ceux de cinquième, entre ceux de quatrième et de troisième, et entre ceux du lycée. Pour cette raison, nous avons procédé à des regroupements de classes et n'avons plus que trois groupes différents selon le niveau scolaire (cf. tableau 4.1) : première moitié du collège (sixième + cinquième), seconde moitié du collège (quatrième + troisième), et lycée (seconde + première). Ces regroupements semblent en accord avec les contenus des programmes scolaires : les fractions sont censées être maîtrisées dès la moitié du collège et les probabilités sont abordées en classe de seconde.

B. Matériel

Le matériel utilisé dans cette expérience est la série des dix problèmes issus de Zhu & Gigerenzer (2006) traduits en français par nos soins (cf. annexe 1). Le problème des maux de tête présenté dans le chapitre 2 en est un exemple. Les traductions ont été validées par une personne ayant l'anglais comme langue maternelle.

Les dix problèmes ont été, de façon aléatoire, répartis en deux séries (A et B) de cinq problèmes, ceci dans le but d'avoir deux versions équivalentes de problèmes. Ces deux versions ont été présentées aux participants en deux formats différents : fréquences naturelles et probabilités conditionnelles. Les deux versions ont été contrebalancées. Ainsi, la moitié des participants a eu la première série (A) en fréquences et la seconde moitié des participants la seconde série (B) en fréquences. Les séries en fréquences traitées par les participants, nous avons proposé la seconde série (B) en probabilités à la première moitié des participants et la première série (A) à l'autre moitié des participants en probabilités. Ce contre-balancement a été réalisé dans le but de palier des différences de difficultés entre les problèmes : un problème plus difficile avec un taux de base très élevé, par exemple.

Chaque participant est donc confronté à, dans un premier temps, cinq problèmes en probabilités conditionnelles, et, dans un second temps, cinq problèmes équivalents en fréquences naturelles.

C. Procédure

Les passations ont lieu dans une des salles de cours des établissements ayant accueilli cette recherche. Ceci est vrai pour l'ensemble des cinq expérimentations.

Les passations ont eu lieu en petits groupes : cinq élèves pour un expérimentateur. Dans un premier temps, les expérimentateurs ont présenté le but de l'expérience comme étant une comparaison entre les différents niveaux de classes de l'établissement sur des problèmes mathématiques. A cette occasion, nous avons remercié les élèves pour leur participation et les avons rassurés quant à leur réussite ou non réussite aux épreuves présentées. Au regard de la littérature, nous savions que pour beaucoup de participants, le format probabiliste est difficile. Aussi, nous avons fait preuve de bienveillance afin de ne pas frustrer les participants.

La série des cinq problèmes en probabilités conditionnelles a été présentée en premier lieu aux participants. Dans un second temps, la série des cinq problèmes équivalents mais en fréquences naturelles a été proposée. Cet ordre est essentiel. De fait, les fréquences permettant une facilitation quant à la présentation et la représentation des problèmes, il était impératif de ne pas induire cette simplification possible chez les participants en leur proposant la version fréquentiste avant la version probabiliste. Le fait de faire passer la version la plus difficile avant est courant et cet ordre classique, pour s'assurer que les stratégies ou les processus cognitifs sollicités sont bien auto initiés et non par la tâche précédente.

Les passations ont eu lieu en temps libre. Les participants décidaient eux-mêmes de rendre leur copie. Ils avaient à leur disposition une page par problème et un crayon. La calculatrice n'était pas autorisée. Les expérimentateurs ont pris soin de préciser aux participants qu'ils pouvaient exploiter comme ils le désiraient la page présente sous le problème et leur ont demandé de ne pas effacer leurs calculs, ni leurs éventuels schémas.

La consigne donnée aux participants était la suivante :

> « Vous devez répondre à la question qui vous est posée à propos de l'histoire qui vous est décrite. Soulignez les nombres qui vous semblent importants dans l'énoncé et laisser apparents tous vos

calculs, brouillons ou dessins. Vous avez le temps qu'il vous faut et vous décidez quand vous voulez rendre le document. »

Le fait de leur demander de souligner les nombres importants dans l'énoncé et de laisser apparents leurs calculs a pour but, d'une part, de recueillir davantage d'informations sur leur façon de traiter l'énoncé, et d'autre part, de quotter plus aisément la stratégie utilisée. En effet, si un participant nous écrit le calcul qu'il souhaite réaliser et se trompe dans sa réalisation, nous décidons de quotter la réponse par la stratégie qui renvoie au calcul souhaité, sans tenir compte de l'erreur de calcul. Par exemple, un enfant indiquant vouloir calculer 8/17 + 2/17 et qui répond 11/17 aura une réponse que nous quotterons comme 11/17. En effet, son erreur de calcul n'est pas à considérer dans l'analyse de son niveau de raisonnement bayésien, a fortiori dans la perspective intuitive de raisonnement probabiliste, notamment bayésien, que nous tentons de valider par ces travaux.

III. Résultats

Comme précisé précédemment, nos résultats ne nous montrent pas de différences significatives entre certaines classes, et sont donc en référence à trois groupes expérimentaux (les regroupements de classes) selon le niveau scolaire.

A. Impact du format selon le niveau scolaire

Une analyse de variance a été menée pour rendre compte de l'impact du format -fréquences versus probabilités- selon le niveau scolaire. Toutefois, les problèmes en probabilités conditionnelles n'ont jamais été réussis, par aucun participant. Cet effet plancher nous a fait réaliser une analyse de variance simple à un facteur. L'effet du niveau scolaire est significatif [F(2 ; 117) = 16,54, p<.001] pour le nombre de problèmes réussis en fréquences naturelles (cf. figure 4.1).

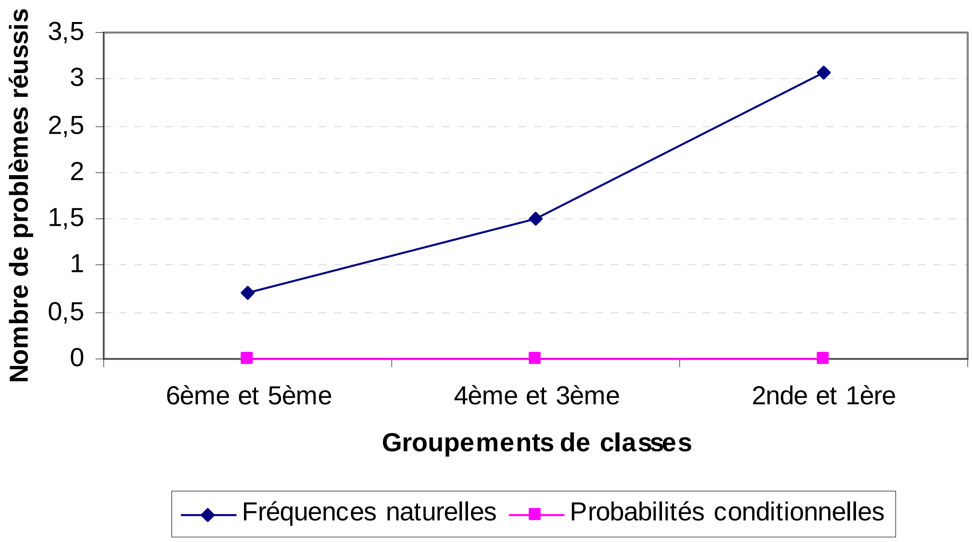


Figure 4.1. Nombres de problèmes réussis selon le format et le niveau scolaire

Des comparaisons planifiées *post hoc* (cf. tableau 4.2) nous apportent des précisions sur cet effet global : le groupements de classe n° 3 (classes de seconde et de première) ont réussi un nombre significativement plus élevé que le deuxième groupement [F(1 ; 117) = 14,05, p<.001] et que le premier groupement [F(1 ; 117) = 31,95, p<.001]. Ces analyses nous suggèrent que les nombres de problèmes réussis par les deux premiers groupements tendent à être différents significativement, mais ne le sont pas au seuil courant de p.05 [F(1 ; 117) = 3,62, p=.059].

	4ème et 3ème	2nde et 1ère
6ème et 5ème	F(1 ; 117) = 3,62, p=.059	F(1 ; 117) = 31,95, p<.001
4ème et 3ème		F(1 ; 117) = 14,05, p<.001

Tableau 4.2. Comparaisons post hoc entre les nombres de problèmes réussis en fréquences naturelles selon les groupements de classes

B. Analyse qualitative des différentes stratégies rencontrées en fréquences naturelles

Les stratégies sont décrites dans le tableau 2.1., et selon les critères établis dans le second chapitre. Nous apportons toutefois une précision sur deux nouvelles étiquettes que nous avons utilisées lors du traitement de nos résultats. Certains participants n'ont pas répondu à tous les problèmes. Nous avons nommé ceci « Non réponse ». Il nous restait ensuite diverses réponses qui ne correspondaient pas à celles décrites dans la littérature. Ces réponses n'étant pas présentes en nombre suffisant, il ne nous a pas semblé nécessaire de rendre compte d'une nouvelle stratégie. Pour cette raison, nous avons une rubrique « Aléatoire » qui correspond à un choix par défaut. Se trouvent étiquetées sous ce label toutes les réponses minoritaires et disparates. Les différentes stratégies rencontrées selon les groupements scolaires sont présentées dans le tableau 4.3. Des effectifs observés étant inférieurs à la condition n > 5, nous n'avons pas pu effectuer un test de comparaison de fréquences (Chi^2).

Stratégies	Groupements de classes		
	6ème et 5ème	4ème et 3ème	2nde et 1ère
Non réponse	12	1	8
Aléatoire	66	37	35
Conservatisme	31	4	3
Evidence	14	29	6
Apparition conjointe	6	12	2
Fonction de vraisemblance	13	6	11
Pré-Bayésienne	29	51	12
Bayesienne	29	60	123

Tableau 4.3. Utilisations des différentes stratégies selon les groupements de classes

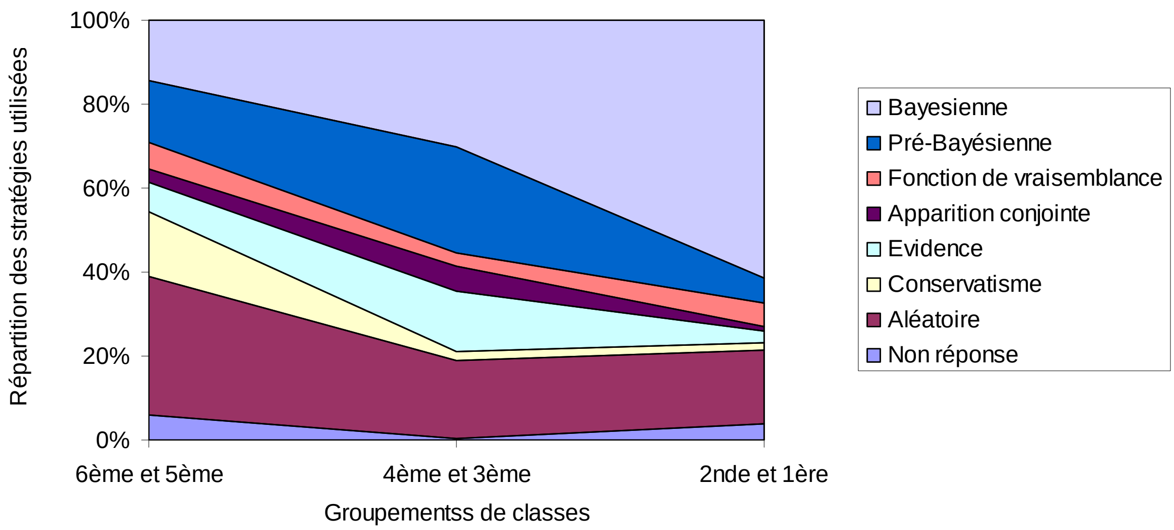


Figure 4.2. Répartition des différentes stratégies selon les groupements de classe

C. Analyse différentielle des choix de stratégies selon les différents groupements de classes

Les différentes stratégies sont utilisées par les participants des trois groupements de classes. La fréquence de leur utilisation varie d'un groupement de classes à un autre. Par exemple, la stratégie bayésienne est de plus en plus utilisée, ce qui a déjà été mis en évidence dans avec la figure 4.1. Ces résultats ne nous ont pas semblés suffisants. En effet, ces données réfèrent aux nombres d'utilisation de chaque stratégie par groupement de classes. Chaque participant étant soumis à cinq problèmes, deux hypothèses s'offrent à nous : lorsqu'une stratégie est plus utilisée (par exemple la stratégie bayésienne), cela signifie que tous les participants utilisent cette stratégie mais pas pour l'ensemble des problèmes, soit cela signifie que seulement certains participants utilisent toujours cette stratégie et ce pour les cinq problèmes. Une approche différentielle nous permet de préciser ces données. Nous constatons une diminution de la dispersion intra individuelle dans l'utilisation de stratégies avec l'avancée en classes, et une plus forte dispersion interindividuelle dans le deuxième groupement, c'est-à-dire celui concernant la seconde moitié du collège (cf. tableau 4.4.). De plus, le nombre de participants utilisant de façon constante une seule et unique stratégie pour l'ensemble des problèmes augment avec l'avancée en classes (cf. tableau 4.5.). Cette constance dans le choix d'une stratégie augment notamment avec le passage au lycée.

	6ème et 5ème	4ème et 3ème	2nde et 1ère
Moyennes	2,05	2	1,4
Ecarts tpes	0,9	0,96	0,71

Tableau 4.4. Stratégies différentes utilisées selon les groupements de classes

	6ème et 5ème	4ème et 3ème	2nde et 1ère
Effectifs	12	15	28

Tableau 4.5. Nombre de participants utilisant une seule stratégie selon les groupements de classes

IV. Discussion

Cette expérimentation est la première de nos travaux dont l'objectif principal est de décrire les performances bayésiennes de collégiens et lycéens français. Lors de cette première expérimentation, nous avons testé l'impact du format de présentation du problème bayésien, et du niveau scolaire des participants, sur leurs réponses données aux problèmes. Les problèmes sont présentés de façon quantitative -en nombre- et la réponse demandée aux participants l'est également (en fréquences naturelles ou en probabilités conditionnelles selon le format).

Cette expérimentation présente des données sur une population de collégiens et lycéens français. Ceci a un double intérêt : d'une part la littérature ne fait nullement référence à des travaux sur une population de nationalité française, et d'autre part, il n'existe pas plus, à notre connaissance, de recherches développementales menées sur des collégiens et des lycéens. L'impact du format de présentation est soit testé sur des professionnels, donc des adultes, utilisant les probabilités dans leur quotidien, soit sur des adultes étudiants, ou encore sur des enfants plus jeunes.

D'une manière générale, nos résultats confirment l'hypothèse centrale de l'approche fréquentiste (Gigerenzer & Hoffrage, 1995, par exemple) : les fréquences naturelles sont facilitatrices par rapport aux probabilités conditionnelles. En effet, nos participants des trois groupements de classes ont répondu de façon correcte à de nombreux problèmes présentés en fréquences naturelles, alors qu'aucun problème sous le format probabiliste n'a été correctement analysé. Un de nos objectifs était de préciser les résultats apportés par Zhu & Gigenrenzer (2006). En effet, dans leurs travaux, ces chercheurs ont obtenu un effet plancher quand le problème est présenté en probabilités conditionnelles, auprès d'élèves chinois scolarisés dans des classes équivalentes à nos classes de CM1, CM2 et sixième. En choisissant une population sur une étendue plus grande, à savoir de la classe de sixième à la classe de première, nous pensions pouvoir décrire l'émergence de réponses bayésiennes correctes à des problèmes présentés en probabilités conditionnelles. Toutefois, nous sommes également confrontés à un effet plancher. Il semblerait que jusqu'en début de classe de première, les élèves français sont incapables de manier les probabilités conformément à la règle de Bayes. Pouvons-nous pour autant dire que cette population de collégiens et lycéens français est non bayésienne ? La réponse est non. Les collégiens et lycéens français ne sont ni non bayésiens, ni irrationnels. En effet, Lorsque le problème est présenté en fréquences naturelles, les plus jeunes d'entres eux sont capables de raisonner de façon bayésienne.

Nos participants présentent un profil plus proche de celui des enfants allemands décrit par Lücking (2004) que celui des enfants chinois décrit par Zhu & Gigerenzer (2006). Il nous parait difficile de savoir si nos participants sont, comme cela est avancé dans certaines études (Artelt, Baumert, Klieme, Neubrand, Prenzel, Schiefele, Schneider, Schümer, Stanat, Tillmann, & Weiß, 2001 ; Stern, Rode, Ge, & Zhu, 2001), moins motivés que les enfants chinois. Nos participants nous ont semblés motivés et emplis de bonnes intentions et de volonté de réussir, même si nous savons que cette analyse rapide et intuitive n'a pas grande valeur, et ne permet pas d'appréhender réellement la volonté des participants.

Les différences entre nos participants français et ceux chinois de Zhu & Gigerenzer (2006) ne sont pas que quantitatives. En effet, leurs enfants testés répondent à davantage de problèmes de façon correcte, au regard de la règle de Bayes, et ce de façon plus précoce, en termes d'avancée dans le cursus scolaire, mais les stratégies qu'ils utilisent lorsqu'ils échouent à ces problèmes sont également différentes de celles utilisées par nos participants. Par exemple, Zhu & Gigerenzer (2006) n'ont pas rencontré la stratégie d'apparition conjointe dans leurs travaux. Nos participants ont, quant à eux, utilisé cette stratégie quelques fois. Nous avons également rencontré quelques non réponse et une proportion non négligeable (33 % pour les -6ème / 5ème et environ 17 % pour les deux autres groupements de classes) de réponses diverses que nous n'avons pas pu identifier de façon précise. Ces réponses étant elles-mêmes différentes entre elles, nous avons fait le choix d'utiliser une catégorie par défaut. Nous pensons qu'elle reflète une réelle non compréhension du problème posé et que les participants ont répondu au hasard une combinaison des données présentes dans l'énoncé, et pour d'autres une combinaison de nombres qui nous ont semblé bien loin de ceux donnés dans l'énoncé.

Ces deux nouveaux types de réponses non bayésiennes sont présentes dans les trois groupements de classes, de même que les autres décrites par Zhu & Gigerenzer (2006). Notre hypothèse est donc partiellement vérifiée. En effet, nous avons bien une évolution dans le choix des stratégies avec l'avancée dans le cursus scolaire, toutefois, toutes les stratégies restent utilisées par nos groupements de classes. Les stratégies les moins efficaces ne sont pas abandonnées totalement, même par les lycéens, comme on s'y attendait, d'après le modèle d'ondes qui se chevauchent de Siegler (1999 ; 2000).

Une approche différentielle nous permet de nuancer ces résultats et d'apporter de plus amples informations. Chaque participant ayant répondu à plusieurs problèmes équivalents, seule une analyse de la variabilité dans les choix de stratégies d'un problème à un autre par un même participant nous permet de choisir entre deux hypothèses explicatives : les participants de 2nde et 1ère répondent tous à quelques problèmes de façon bayésienne, ou certains de ces participants répondent à tous les problèmes de façon bayésienne alors que le reste du groupe ne répond à aucun problème de façon bayésienne. Nos résultats nous ont permis de montrer que la variabilité inter individuelle dans le troisième groupement était plus faible que dans les deux autres. Il en est de même pour la variabilité intra individuelle. Ceci signifie qu'au lycée, les participants ont tendance à utiliser une seule et unique stratégie pour l'ensemble des problèmes et que, dans le même temps, ces participants ont une propension à utiliser la même stratégie. Ceci peut s'expliquer en partie par le fait que pour ce groupement de classes, le nombre de réponses bayésiennes est conséquent. Le fait que dans le troisième groupement vingt-huit participants sur les quarante répondent de manière bayésienne aux cinq problèmes proposés nous donne un groupe plus homogène. Nous ne sommes tout de même pas confrontés à un effet plafond.

Afin de pouvoir appréhender à quel moment dans la scolarité les jeunes français sont capables de répondre à des problèmes bayésiens qui leurs sont présentés en probabilités conditionnelles, il serait intéressant de faire passer ces problèmes à des élèves de fin de première, et non de début comme pour cette expérience, des élèves de terminale, et des étudiants post baccalauréat.

Chapitre 5 - Impact du format de présentation et du niveau scolaire sur les performances bayésiennes : analyse de la représentation du problème et des performances intuitives

I. Introduction

Notre première expérience a confirmé la thèse fréquentiste de Gigerenzer, selon laquelle les fréquences naturelles sont facilitatrices par rapport aux probabilités conditionnelles et permettent de meilleures performances en raisonnement bayésien. Le choix de notre population, jusqu'à maintenant non étudiée, a permis d'élargir les connaissances théoriques à de jeunes collégiens et lycéens français, et ainsi de montrer que le format fréquentiste est réellement facilitateur, tant pour des collégiens que des lycéens. Une des raisons avancée par Gigerenzer et ses collègues pour expliquer cet effet facilitateur, est que les fréquences simplifient considérablement l'utilisation de la règle de Bayes (Gigerenzer & Hoffrage, 1995). En d'autres termes, pour ces auteurs, les fréquences naturelles sont facilitatrices dans le sens où, lorsqu'on les insère dans la règle de Bayes, on obtient une formule réduite par rapport à l'insertion, dans cette même règle, de probabilités conditionnelles (cf. figure 2.3.) (Gigerenzer & Hoffrage, 1995). Il est en effet plus simple d'utiliser la formule de Bayes avec des fréquences qu'avec des probabilités, d'un point de vue purement mathématique.

Pour Piaget & Inhelder (1951 ; 1975) les enfants maîtrisent le concept de nombre et celui de probabilités dès le troisième stade : celui de la genèse des opérations formelles (Hoeman & Ross, 1971 ; Kreitler & Kreitler, 1986), soit dès l'âge de onze / douze ans. Ils sont alors capables de combiner des probabilités entre elles. Ces travaux sont remis en question dans de nombreuses études qui plaident en faveur de capacités plus précoces.

Cependant, la littérature nous fournit de nombreuses recherches menées dans le cadre de l'approche fréquentiste et qui ont montré que les probabilités n'étaient maîtrisées, dans un contexte bayésien, ni par les enfants, ni par beaucoup d'adultes, professionnels utilisant les probabilités dans leur fonction ou non.

Le but de cette deuxième expérimentation est de tester dans quelle mesure les fréquences naturelles favorisent davantage un raisonnement baysésien, plus précisément pour quelles raisons. Certes elles simplifient la règle de Bayes et aboutissent à une formule plus simple à calculer, toutefois, nous postulons qu'elles sont également facilitatrices par la représentation qu'elles suscitent chez les participants. L'objectif est de tester des problèmes bayésiens sous les deux formats -probabiliste et fréquentiste- sans que les réponses demandées aux participants nécessite la règle de Bayes de façon explicite. Si les participants traitent les problèmes bayésiens selon le mode de raisonnement heuristique et intuitif (Epstein, 1983, 1994, 1998), le système 1 (Kahneman, 2003), nous pensons que les fréquences naturelles sont également facilitatrices lorsqu'il ne leur est pas demandé de fournir ni un calcul précis, ni une réponse chiffrée. Ceci signifierait que les fréquences naturelles sont facilitatrices, certes grâce à la simplification qu'elles permettent de la règle de Bayes, mais également au regard de la compréhension et de la représentation du problème bayésien chez les participants.

II. Méthode

A. Participants

120 participants ont été sélectionnés, selon les mêmes critères que pour la première expérimentation. Le genre a, de nouveau, été contrôlé (cf. tableau 5.1.).

	Classes					
	Sixième	Cinquième	Quatrième	Troisième	Seconde	Première
Effectifs (Garçons ; Filles)	20 (10 ; 10)	20 (10 ; 10)	20 (10 ; 10)	20 (10 ; 10)	20 (10 ; 10)	20 (9 ; 11)
Regroupements de classes	6ème et 5ème		4ème et 3ème		2nde et 1ère	

Tableau 5.1. Répartition des effectifs selon le niveau scolaire

B. Matériel

Le matériel présenté aux participants est toujours le set des dix problèmes bayésiens élaborés par Zhu & Gigerenzer (2006) et traduits par nos soins. Toutefois, si les énoncés restent identiques dans les deux conditions -fréquentiste et probabiliste-, la question posée également, le type de réponse attendu diffère. En effet, nous avons proposé aux participants des continuums longs de quinze centimètres. Ces continuums sont présentés avec deux pôles distincts et sous chacune des bornes est inscrite la probabilité équivalente : « Aucun » à gauche ou « Tous » à droite pour la version fréquentiste, et « N'a pas ... » à gauche ou « A … » à droite pour la version probabiliste (cf. figure 5.1.).

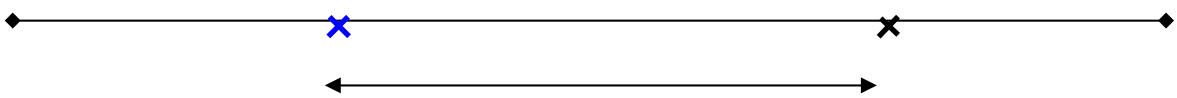
Continuum pour la version probabiliste :

N'a pas de rhume A un rhume

Continuum pour la version fréquentiste :

<u>Figure 5.1.</u> Continuums proposés pour le problème des maux de tête

La borne de gauche signifie une fréquence naturelle de zéro sur un nombre quelconque ou une probabilité de zéro pourcent, selon le format utilisé dans l'énoncé. La borne de droite signifie une fréquence de x/x (numérateur et dénominateur égaux), ou une probabilité de cent pourcent, selon que l'énoncé est fréquentiste ou probabiliste. Une croix au milieu équivaut à une probabilité de 50 %. Le continuum est utilisé comme une variable continue. La réponse bayésienne est calculée par nos soins et placée sur un continuum repère qui n'est pas présenté aux participants. La distance entre la croix théorique et la croix placée par les participants est mesurée (cf. figure 5.2.). Cette distance est utilisée comme mesure de la performance bayésienne intuitive des participants. Ainsi, plus la distance est petite, plus le participant s'approche d'un raisonnement bayésien.



Croix placée par le participant : ×

Croix théorique calculée selon la règle de Bayes, puis placée à la proportionnelle sur le continuum : ×

<u>Figure 5.2.</u> Mesure de la distance entre la croix théorique et la croix du participant

Par exemple, au problème des maux de tête (cf. figure 2.5.), la réponse attendue est une fréquence de 42/54, soit une probabilité de 0,78. Cette probabilité est rapportée par un calcul de proportionnalité sur le continuum mesurant quinze centimètres. Ceci nous donne 11,7. La croix théorique se trouve donc à 11,7 cm de la borne de gauche du continuum.

La distance entre les deux croix est mesurée et utilisée comme performance bayésienne intuitive des participants.

Notons que les distances sont utilisées en valeur absolue afin de ne pas avoir de compensations entre les surestimations et les sous-estimations au sein d'un même groupement de classes. Ainsi, que la croix du participant soit à gauche ou à droite de la croix théorique, la distance est mesurée et est utilisée de façon identique.

Sont proposées à chaque participant deux séries de cinq problèmes : cinq en probabilités conditionnelles, et les cinq suivants en fréquences naturelles. Les problèmes sont contrebalancés au sein des deux listes. Chaque problème est donc présenté sous les deux formats, afin d'éviter un biais relatif à un problème spécifique.

C. Procédure

Les passations ont, comme les précédentes, eu lieu dans une salle de cours des établissements visités. Elles ont été collectives, à raison de cinq enfants et/ou adolescents de la même classe par session. Les participants ont, dans un premier temps, répondu à la version probabiliste des problèmes, et, dans un second temps, à la version fréquentiste. Le temps de passation est libre. Les participants rendent leurs documents quand bon leur semble.

III. Résultats

Une analyse de variance a été réalisée pour tester l'effet de nos deux variables indépendantes : le format de présentation de l'énoncé et le niveau scolaire (groupements de classe) sur le raisonnement bayésien sans calcul explicite. Les performances mesurées ici sont les distances moyennes entre les croix des participants et la croix théorique placée à la proportionnelle sur le continuum, d'après la règle de Bayes. Plus précisément, ces distances sont rapportées en valeur absolue. Il ne nous a pas semblé pertinent de tenir compte du placement qualitatif de la croix du participant par rapport à la croix théorique (à gauche ou à droite). En effet, le but de cette expérimentation est de tester les performances bayésiennes sans calcul des participants. Notre objectif n'est pas de vérifier si les participants surestiment ou sous estiment la probabilité *a posteriori*, mais de combien ils s'en approchent. Seule la distance est donc importante.

Notre analyse de variance souligne un effet significatif du format de présentation [$F(1 ; 117) = 43{,}31$, $p<.001$] (cf. figure 5.3.). Les distances moyennes entre les croix des participants et les croix théoriques pour les cinq continuums sont plus faibles en condition fréquentiste qu'en condition probabiliste. Notre analyse montre un effet de la variable niveau scolaire [$F(1 ; 117) = 10{,}32$, $p<.001$]. Toutefois, notre analyse ne montre pas d'interaction significative entre ces deux variables [$F(1 ; 117) = 1{,}30$, $p=.2776$].

Afin d'avoir des informations supplémentaires sur l'évolution des performances bayésiennes sur continuum, nous avons réalisé des analyses post hoc qui ont pour but de comparer les moyennes des différents groupes. L'effet d'interaction entre les deux variables principales n'étant pas significatif, nous avons effectué des comparaisons intra facteur. En d'autres termes, nous avons testé les différences entre les groupements de classe, et ce, condition de présentation par condition de présentation, dans le but de vérifier si l'évolution des performances entre les groupements de classes est la même quand l'énoncé est présenté en probabilités conditionnelles (cf. tableau 5.2) et quand l'énoncé est présenté en fréquences naturelles (cf. tableau 5.3).

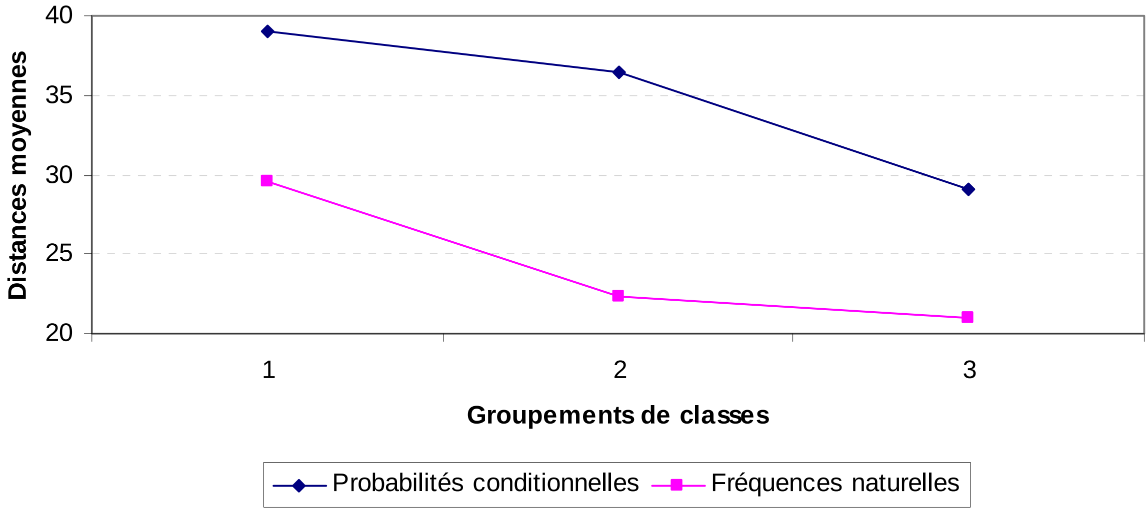


Figure 5.3. Distances moyennes entre les croix des participants et la croix théorique selon les groupements de classes

	4ème et 3ème	2nde et 1ère
6ème et 5ème	F(1 ; 117) = 0,69, p=.41	F(1 ; 117) = 10,69, p<.01
4ème et 3ème		F(1 ; 117) = 5,93, p<.05

Tableau 5.2. Comparaisons post hoc entre les distances moyennes, en valeur absolue, séparant les croix théoriques des croix des participants quand l'énoncé est probabiliste

	4ème et 3ème	2nde et 1ère
6ème et 5ème	F(1 ; 117) = 7,59, p<.01	F(1 ; 117) = 10,56, p<.01
4ème et 3ème		F(1 ; 117) = 0,24, p=.62

Tableau 5.3. Comparaisons post hoc entre les distances moyennes, en valeur absolue, séparant les croix théoriques des croix des participants quand l'énoncé est fréquentiste

L'étude de ces comparaisons planifiées *a posteriori* nous permettent, d'une part, de voir que lorsque l'énoncé est présenté en probabilités conditionnelles, les performances du troisième groupement de classes (2^{nde} et $1^{ère}$) sont significativement différentes des deux autres groupements de classes [F(1 ; 117) = 10,69, p<.01 et F(1 ; 117) = 5,93, p<.05 respectivement par rapport au groupement de $6^{ème}$ / $5^{ème}$ et $4^{ème}$ / $3^{ème}$]. Les performances entre les deux groupements de classes issus du collège n'apparaissent pas comme significativement différents [F(1 ; 117) = 0,69, p=.41].

D'autre part, quand l'énoncé est sous un format fréquentiste, les performances entre les le premier groupement des classes de $6^{ème}$ et $5^{ème}$ sont significativement différentes de celles du groupement $4^{ème}$ / $3^{ème}$ [F(1 ; 117) = 7,59, p<.01] et de celles du groupement 2^{nde} / $1^{ère}$ [F(1 ; 117) = 10,56, p<.01]. Toutefois, nous n'avons pas pu montrer que les performances entre le groupement de classes de fin de collège et celui de classes de lycée étaient différentes significativement [F(1 ; 117) = 0,24, p=.62].

Les performances des participants s'améliorent entre le début et la fin du collège quand le problème présenté est fréquentiste, et elles suivent un palier avec le passage au lycée. Par contre, quand le problème est présenté en probabilités conditionnelles, les performances sont stables durant les années collèges et ne s'améliorent que chez les participants de lycée.

IV. Discussion

Cette recherche visait à rendre compte des performances bayésiennes que nous qualifions d'intuitives, d'heuristiques, de collégiens et lycéens français. Plus précisément le but de cette deuxième expérimentation était d'aller plus loin que l'approche fréquentiste, en proposant un double effet facilitateur des fréquences naturelles par rapport aux probabilités conditionnelles. Dans un premier temps, cet effet facilitateur concerne le mode logique, rationnel et qui renvoie au calcul réel d'une probabilité conditionnelle pour le format probabiliste ou d'une fréquence naturelle, lorsque le problème est présenté dans un format fréquentiste. Ce premier aspect est l'argument cité par l'approche fréquentiste, en faveur d'une rationalité écologique. Dans un second temps, l'aspect qui nous intéresse plus particulièrement ici est l'aspect facilitateur quant à la compréhension et la représentation du problème que les fréquences naturelles permettent. Cette distinction fait référence aux théories du double processus dans le sens où le premier aspect renverrait davantage au système 2, logique et rationnel, et le second aspect, que nous proposons d'étudier, renverrait dans de plus grandes proportions au système 1, intuitif et heuristique (Epstein, 1983, 1994, 1998).

Nos résultats nous montrent que lorsque le problème est présenté en fréquences naturelles ou en probabilités conditionnelles, les participants n'estiment pas de façon équivalente la probabilité ou la fréquence qui est évoquée. Une analyse de variance a mis en exergue un effet significatif du format de présentation des données. Les probabilités conditionnelles sont toujours estimées moins précisément que les fréquences naturelles, et ce, pour tous les groupements de classes. Dans le même temps, cette analyse statistique nous a également permis de montrer un effet significatif des groupements de classes. Avec l'avancée dans le cursus scolaire, les participants fournissent une réponse de moins en moins éloignée de la réponse calculée avec la règle de Bayes. Remarquons que cette amélioration des performances intuitives, ne se fait pas de la même manière selon le format de présentation. Quand le problème proposé est fréquentiste, les performances des participants sont meilleures dès le deuxième groupement des classes de 4ème et 5ème. Ces performances ne semblent pas s'améliorer par la suite, avec l'arrivée au lycée. Lorsque le problème présenté est probabiliste, les performances des participants ne semblent pas différentes entre les deux groupements de classes de collège. Par contre, les performances s'améliorent avec le passage au lycée. Le fait d'obtenir deux profils différents à propos des performances intuitives selon que le problème est fréquentiste ou probabiliste nous conforte dans l'idée d'une sollicitation, de la part des participants, d'un mode de raisonnement heuristique. Ce mode, aussi appelé système 1 est qualifié d'intuitif et réfère à un mode de fonctionnement plus quotidien. Nous pouvons imaginer que les participants lorsqu'ils sont confrontés à des problèmes de type bayésien ne passent pas par un calcul explicite conformément à la règle de Bayes, mais qu'ils utilisent davantage une approche intuitive. Avec ce mode de raisonnement, les fréquences sont également facilitatrices. Les élèves du troisième groupement de classes donnent une réponse moyenne, avec un énoncé en fréquences naturelles, qui est éloignée d'environ 20 millimètres sur le continuum. Ceci équivaut à une erreur d'environ 13 %. Lorsque l'énoncé est en probabilités conditionnelles, ils se trompent d'environ 25 %. Ces résultats sont un argument en faveur d'une rationalité humaine expérientielle et d'une structure bayésienne précoce qui permet de raisonner de façon plus ou moins adaptée à notre environnement et plus ou moins précise selon les conditions. Nous adoptons également une vision optimiste de l'être humain, en suggérant que les individus sont capables, bien avant de pouvoir comprendre et utiliser la règle de Bayes, d'adopter un raisonnement de type bayésien et de réviser leur jugement, de faire des inférences probabilistes. Nous pensons que pour la plupart des activités quotidiennes, qui ne requièrent bien souvent qu'une estimation, nos participants ne sont pas trop éloignés de la réponse bayésienne et semblent ainsi capables de s'adapter.

Chapitre 6 - Impact des données présentées dans des problèmes bayésiens sur les représentations des individus et leurs réponses intuitives

I. Introduction

L'approche fréquentiste accorde aux fréquences naturelles un effet facilitateur, et ce avec deux arguments : le premier est le fait que contrairement aux probabilités conditionnelles elles informent de façon explicite sur le taux de base, et le second argument est qu'elles simplifient la règle de Bayes.

Nos deux expérimentations précédentes accordent aux fréquences naturelles, en plus des deux arguments cités par Gigerenzer, un effet facilitateur quant à la compréhension et la représentation du problème. Elles ne simplifient pas uniquement un raisonnement qui s'appuie sur le mode analytique et logique, mais également un raisonnement qui réfère à un mode davantage heuristique et intuitif. Les performances bayésiennes intuitives sont donc autant que les performances bayésiennes du système 2, améliorées par les fréquences naturelles de l'énoncé.

Evans, Handley, Perham, Over, & Thompson (2000) ont montré que le format de la question posée à la suite du problème bayésien a son importance. Pour ces auteurs, une question posée sous la forme d'une proportion amène à des performances intermédiaires entre celles obtenues quand la question est posée en fréquences naturelles et en probabilités conditionnelles.

Le but de notre troisième expérimentation est de tester l'effet des trois informations fournies de façon qualitative dans un problème bayésien selon le niveau scolaire et le format de la question posée. Rappelons que les trois données essentielles sont les données B, D et F (cf. figure 2.5.) La donnée B est la probabilité de l'événement E1. La donnée D renvoie à la probabilité de rencontrer en même temps les événements E1 et E2. Enfin, la donnée F réfère aux faux positifs : rencontrer l'événement E2 sans l'événement E1.

De façon à faire abstraction du nombre, comme suite aux deux précédentes expériences, nous faisons varier les trois données en valeur qualitative, c'est-à-dire avec une information de quantité qui est littérale : des adjectifs numériques tels *peu* et *beaucoup*. Un tel protocole expérimental va nous permettre d'appréhender l'attention accordée par les participants à chacune des trois données. Cette approche est utilisée par Anderson (1979, 1981, 1996) pour décrire sa théorie de l'intégration de l'information. Pour lui, la significativité des effets simples et des effets d'interaction renvoie à différents profils cognitifs d'intégration de l'information. Plus précisément, lorsque seuls les effets simples sont significatifs, nous parlons de profils additifs, et lorsque sont significatifs à la foi les effets simples et les effets d'interaction, nous avons de profils multiplicatifs. Enfin, quand un seul effet simple est significatif, cela équivaut à un modèle disjonctif. Dans le cadre bayésien qui est notre propos ici, ces différents modèles renverraient à différentes façons d'intégrer l'information présentée dans les problèmes bayésiens. Ainsi, un modèle additif signifie que les participants tiennent compte des différentes données mais ne les combinent pas entre eux. Ils ne révisent pas leur jugement. Un modèle multiplicatif signifie que les participants intègrent dans leur raisonnement les différentes données en les relativisant les unes par rapport aux autres. Ils révisent leur jugement. Enfin, un modèle disjonctif renvoie à des participants qui ne s'intéressent qu'à une seule probabilité, comme c'est le cas avec la stratégie de conservatisme. Notons qu'une méthodologie assez semblable a été utilisée par Oaksford & Hahn (2004). Ces derniers ont fait varier les arguments (négatifs et positifs), les croyances antérieures (faibles ou fortes) et le preuves (1 ou 50) dans des dialogues argumentatifs. Les participants devaient donner une réponse chiffrée et non se situer sur un axe comme nous le demandons ici.

Nous postulons que dans des énoncés qualitatifs sans nombre, l'effet du format de la question -fréquentiste ou probabiliste- sur les réponses des participants n'est pas significatif. Nous supposons, dans le même temps, que les profils cognitifs des participants ne seront pas différents selon que la question est en fréquences naturelles ou en probabilités.

Par contre, d'après l'évolution des performances bayésiennes décrites par l'approche fréquentiste et par nos deux précédentes expérimentations, nous supposons que les profils cognitifs selon le niveau scolaire des participants seront différents. Par exemple, nous pensons que le premier groupement de classes sera disjonctif, le deuxième additif et le troisième, celui des lycées, présentera un profil multiplicatif qui signifie que les participants révisent leur jugement en tenant compte des différentes données de l'énoncé.

II. Méthode

A. Participants

120 participants ont été sélectionnés, selon les mêmes critères que pour la première et la seconde expérimentation, à savoir six groupes de 20 élèves de chaque classe de collège et des deux première classes de lycée : seconde et première. Le genre a également été contrôlé (cf. tableau 6.1.).

	Classes					
	Sixième	Cinquième	Quatrième	Troisième	Seconde	Première
Effectifs (Garçons ; Filles)	20 (10 ; 10)	20 (10 ; 10)	20 (10 ; 10)	20 (10 ; 10)	20 (10 ; 10)	20 (10 ; 10)
Regroupements de classes	6ème et 5ème		4ème et 3ème		2nde et 1ère	

Tableau 6.1. Répartition des effectifs selon le niveau scolaire

B. Matériel

Le matériel utilisé est l'ensemble des problèmes bayésiens extrait de Zhu & Gigerenzer (2006) que nous avons traduits. Nous les avons également modifiés de façon à ne plus avoir de données quantitatives dans l'énoncé. Nous avons donc substitué aux trois données en nombres présentes dans l'énoncé des adjectifs en lettres (cf. figure 6.1.). Les trois données nécessaires à la résolution du problème sont, indépendamment du format la donnée B -le taux de base des personnes ayant un rhume-, la donnée D -le nombre de personnes ayant un rhume et ayant des maux de tête- et la donnée F -les personnes n'ayant pas de rhume mais souffrant malgré tout de maux de tête-. Les autres données n'ont pas besoin d'êtres données de façon explicite dans l'énoncé, elles peuvent être retrouvées par un calcul de complémentarité. Par exemple, on soustrait à 100 -l'échantillon entier- le nombre de personnes b ayant un rhume, pour connaître le nombre c de personnes n'ayant pas de rhume. Selon le même principe et en prenant comme référent le nombre de personnes ayant ou non un rhume, il est aisé par soustraction de trouver le nombre de personnes n'ayant pas de maux de tête parmi celles ayant un rhume -donnée E- et le nombre de personnes n'ayant pas de maux de tête parmi celles n'ayant pas de rhume -donnée G-.

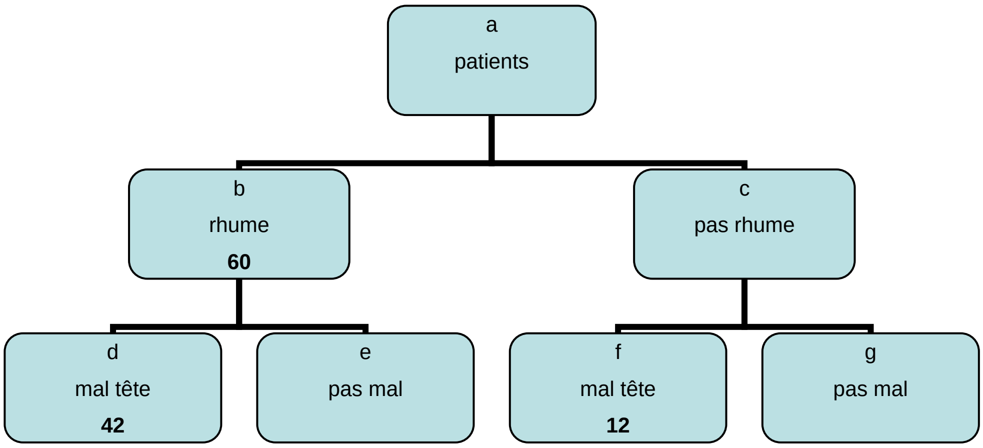


Figure 6.1. Données essentielles de l'énoncé nécessaires pour la résolution du problème bayésien

Ces trois données sont remplacées par des adjectifs numériques qui rendent compte de deux quantités : PEU versus BEAUCOUP. En croisant les trois données dans les deux conditions, nous obtenons 2^3 conditions, soit huit conditions différentes (cf. tableau 6.2).

Versions	Probabilité B	Probabilité D	Probabilité F
vA	Peu	Peu	Peu
vB	Peu	Peu	Beaucoup
vC	Peu	Beaucoup	Peu
vD	Peu	Beaucoup	Beaucoup
vE	Beaucoup	Peu	Peu
vF	Beaucoup	Peu	Beaucoup
vG	Beaucoup	Beaucoup	Peu
vH	Beaucoup	Beaucoup	Beaucoup

Tableau 6.2. Huit versions des problèmes bayésiens selon le croisement des trois données de l'énoncé

Problème des maux de tête dans la version vA :

« Peu de personnes ont un rhume. Parmi ces quelques personnes ayant un rhume, peu ont un mal de tête. Parmi les personnes n'ayant pas de rhume, peu ont tout de même mal à la tête. »

Chaque participant est soumis à six problèmes équivalents présentés dans les huit conditions décrites dans le tableau 6.2. ci-dessus. Trois d'entre eux sont présentés avec une question en probabilités et trois d'entre eux sont présentés avec une question en fréquences naturelles :

> Question posée dans la condition probabiliste :
> Tu rencontres les patients avec maux de tête, quelle est la probabilité qu'une personne ait un rhume ?
>
> Question posée dans la condition fréquentiste :
> Tu rencontres les patients avec maux de tête, combien d'entre eux ont un rhume ?

Comme pour la deuxième expérimentation, les participants doivent répondre à la question posée en plaçant une croix sur le continuum. Cette fois, il ne nous est pas possible de placer une croix théorique calculée via la règle de Bayes. Ce que nous mesurons est donc la distance en millimètres entre la borne de gauche (Aucun n'a de rhume pour la version fréquentiste / N'a pas de rhume pour la version probabiliste) et la croix du participant.

C. Procédure

Les conditions de passations sont similaires aux précédentes en ce qui concerne le lieu, le nombre de participants et l'accord de leurs parents.

Les expérimentateurs présentent dans un premier temps la version des trois problèmes déclinés en huit conditions et avec une question en probabilités conditionnelles aux participants. Dans un second temps, ils leur proposent la version des trois problèmes avec la question posée en fréquences naturelles et déclinés dans les huit conditions.

L'ordre des huit conditions pour chaque problème est randomisé, de façon à ne pas influencer les participants. Cet ordre aléatoire nous semble plus approprié que l'ordre méthodique que nous avons utilisé pour décliner les huit conditions (cf. tableau 6.2.). En effet, telles que nous les avons présentées, ces huit conditions sont ordonnées et les quatre qui renvoient à une probabilité b faible sont à la suite. Dans le but de tester l'impact de chacune des trois données sur les réponses bayésiennes de nos participants, nous avons préféré mélanger les huit conditions.

III. Résultats

Dans un premier temps, nous avons effectué des moyennes entre les trois mesures par participant et par format de question, et ce dans chacune des huit conditions. Ainsi, chaque participant a répondu a trois problèmes dans la version vA, soit avec des données B, D et F, qualifiées de « PEU » toutes les trois (cf. tableau 6.2.) (« Peu de personnes ont un rhume. Parmi ces quelques personnes ayant un rhume, peu ont un mal de tête. Parmi les personnes n'ayant pas de rhume, peu ont tout de même mal à la tête »). Il en est de même pour les autres versions de vB à vH. Pour chacune des versions, les trois réponses ont été moyennées par participant. Chaque participant a donc deux séries de huit scores : les huit moyennes des trois problèmes avec une question en fréquences naturelles et les huit moyennes des trois problèmes avec une question posée en termes de probabilités conditionnelles.

Dans un second temps, nous avons effectué des analyses de variance à mesures répétées afin de vérifier un effet du format de question. Nos analyses ne nous permettent pas prétendre à un effet du format de la question. Par exemple, il n'y a pas de différences significatives entre la moyenne des trois problèmes avec une question en fréquences et la moyenne des trois problèmes avec un question probabiliste, dans la version vA [F(1 ; 119) = 1,09, p=.29].

Pour cette raison, nous avons effectué une nouvelle moyenne entre les réponses des participants aux six problèmes de chaque version.

Les données que nous utilisons dans la suite de ces résultats sont donc les valeurs moyennes des participants aux six problèmes (trois avec une question fréquentiste et trois avec une question probabiliste), selon leur appartenance à un des trois groupements de classes utilisés dans les expérimentations précédentes.

Une analyse de variances avec un plan de mesures répétées 2*2*2 nous a permis de tester, d'une part, l'effet de chacune des données de l'énoncé, et l'interaction entre ces données, d'autre part, sur la réponse des participants. Pour les trois groupements de classes, nous obtenons un effet simple des trois données B, D et F de l'énoncé (cf. tableau 6.3) qui est significatif. Ainsi, les élèves des classes de sixième et cinquième, par exemple, prennent en compte la donnée B [F(1 ; 39) = 7,85, p<.01], la donnée D [F(1 ; 39) = 27,28, p<.001] et la donnée F [F(1 ; 39) = 36,53, p<.001] indépendamment les unes des autres pour répondre à la question qui leur est posée.

	6ème et 5ème	4ème et 3ème	2nde et 1ère
B	***F(1 ; 39) = 7,85, p<.01***	***F(1 ; 39) = 38,0, p<.001***	***F(1 ; 39) = 14,3, p<.001***
D	***F(1 ; 39) = 27,28, p<.001***	***F(1 ; 39) = 269,8, p<.001***	***F(1 ; 39) = 116, p<.001***
F	***F(1 ; 39) = 36,53, p<.001***	***F(1 ; 39) = 15,9, p<.001***	***F(1 ; 39) = 8,8, p<.01***
B*D	F(1 ; 39) = 1,33, p=.26	F(1 ; 39) = 1,3, p=.26	***F(1 ; 39) = 13, p<.01***
B*F	F(1 ; 39) = 0,96, p=.33	F(1 ; 39) = 0,6, p=.45	***F(1 ; 39) = 6,3, p<.05***
D*F	F(1 ; 39) = 2,50, p=.12	F(1 ; 39) = 1,7, p=.20	F(1 ; 39) = 0,0, p=.84
B*D*F	F(1 ; 39) = 2,82, p=.10	F(1 ; 39) = 1,9, p=.18	F(1 ; 39) = 0,0, p=.94

Les effets en **gras** et *italique* sont significatifs

Tableau 6.3. Effets simples et d'interactions des/entre les données B, D et F de l'énoncé sur les réponses des participants selon le niveau scolaire

Notre analyse nous a permis de mettre en avant des effets simples concernant les trois données de l'énoncé pour les trois groupements de classes (cf. figure 6.2. ; 6.3. et 6.4.). Par exemple, les participants des classes de sixième et cinquième modulent leur probabilité conditionnelle en réponse au problème posé selon que la donnée B varie [F(1 ; 39) = 27,28, p<.001].

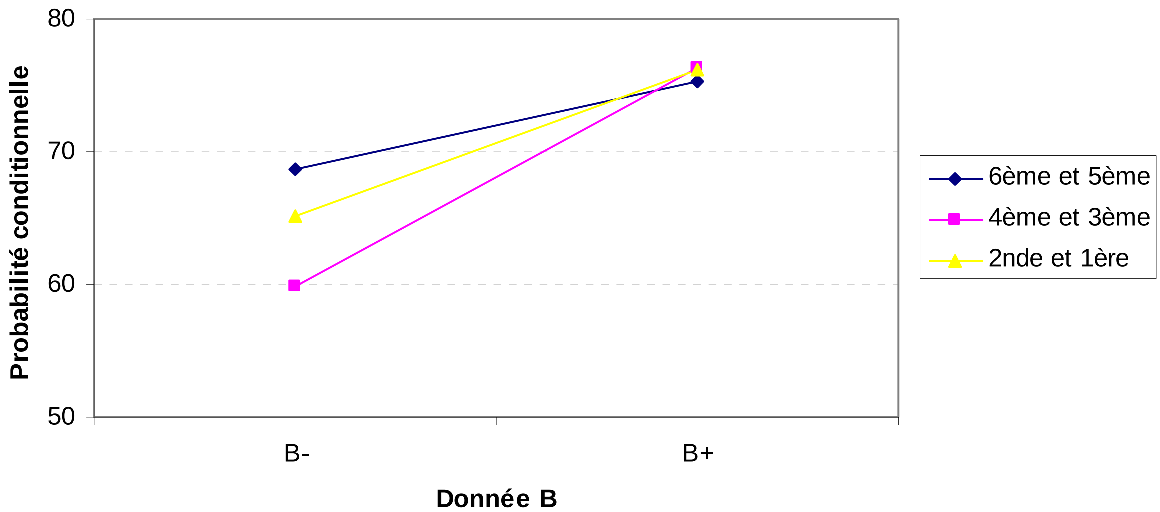


Figure 6.2. Intégration de la donnée B par les trois groupements de classe

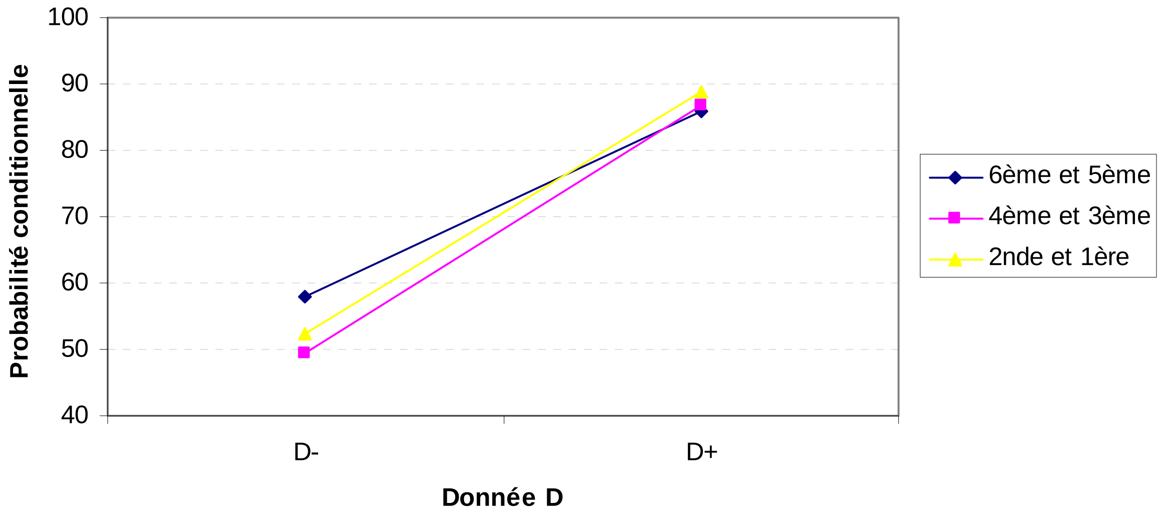


Figure 6.3. Intégration de la donnée D par les trois groupements de classe

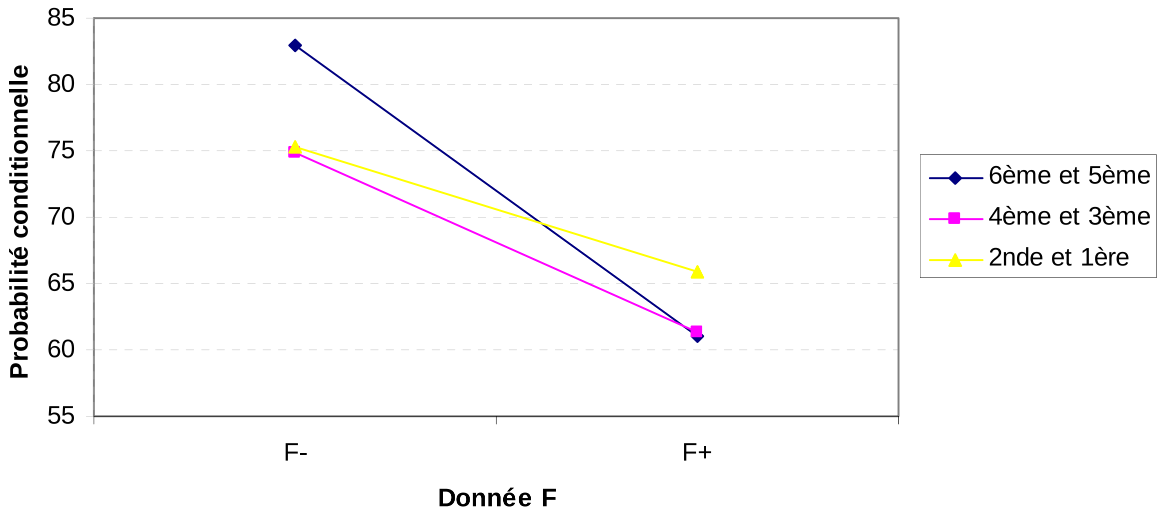


Figure 6.4. Intégration de la donnée F par les trois groupements de classe

Notre analyse nous a permis de montrer que l'interaction entre les effets de la donnée B et de la donnée D n'est pas significative pour les élèves des classes de sixième et cinquième (cf. figure 6.5.) ni pour les élèves de quatrième et troisième (cf. figure 6.6.). Toutefois, elle l'est pour les élèves de seconde et de première (cf. figure 6.7.). L'absence d'interaction est caractérisé par des lignes parallèles tandis qu'un effet d'interaction significatif est représenté par des lignes en éventail.

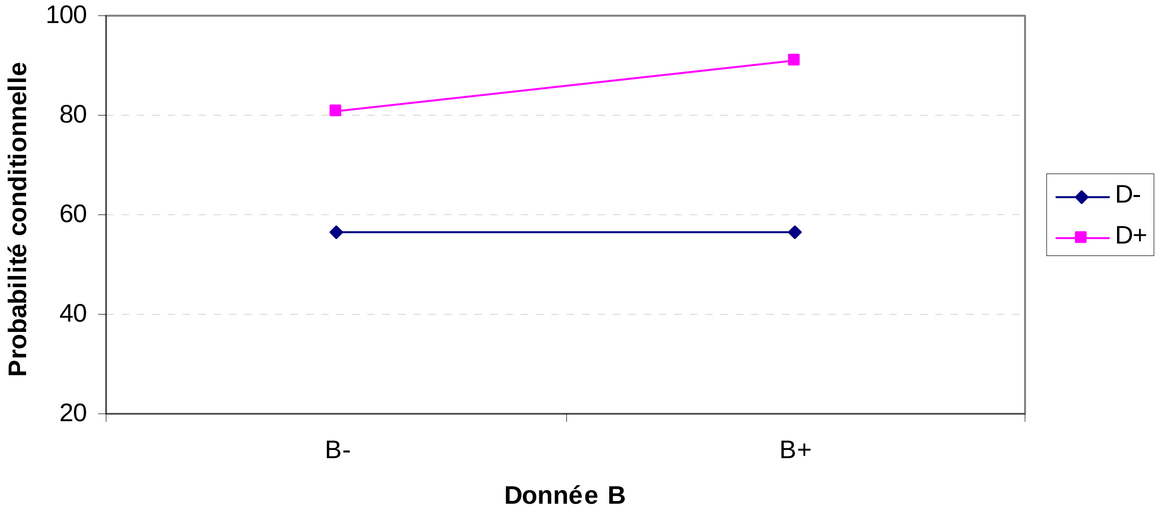


Figure 6.5. Intégration des données B et D par les élèves de 6ème et 5ème

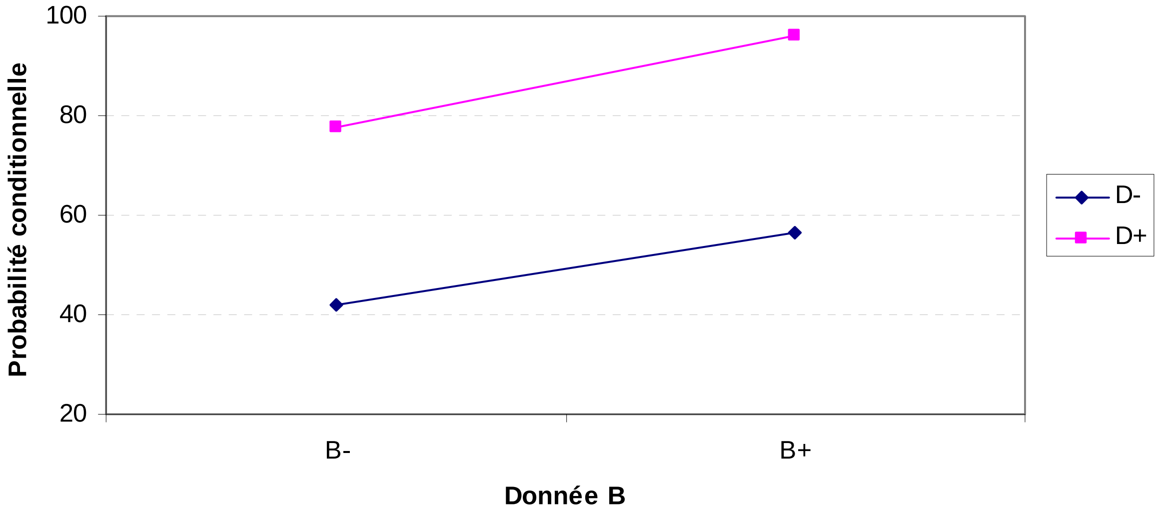


Figure 6.6. Intégration des données B et D par les élèves de 4ème et 3ème

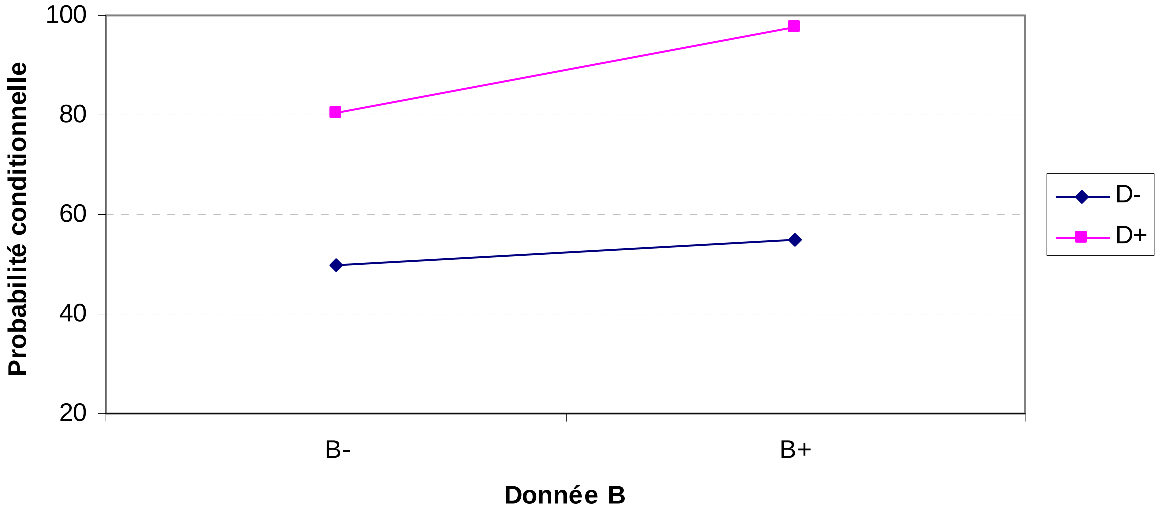


Figure 6.7. Intégration des données B et D par les élèves de 2nde et 1ère

Notre analyse nous a également permis de montrer que l'interaction entre les effets de la donnée B et de la donnée F n'est pas significative pour les élèves des classes de sixième et cinquième (cf. figure 6.8.) ni pour les élèves de quatrième et troisième (cf. figure 6.9.). Toutefois, elle l'est pour les élèves de seconde et de première (cf. figure 6.10.).

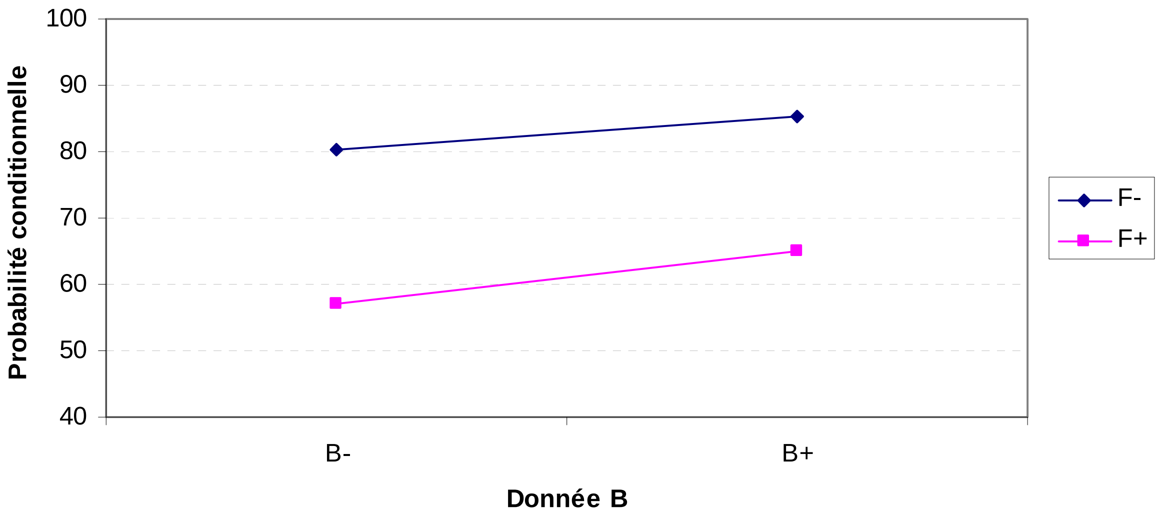


Figure 6.8. Intégration des données B et F par les élèves de 6ème et 5ème

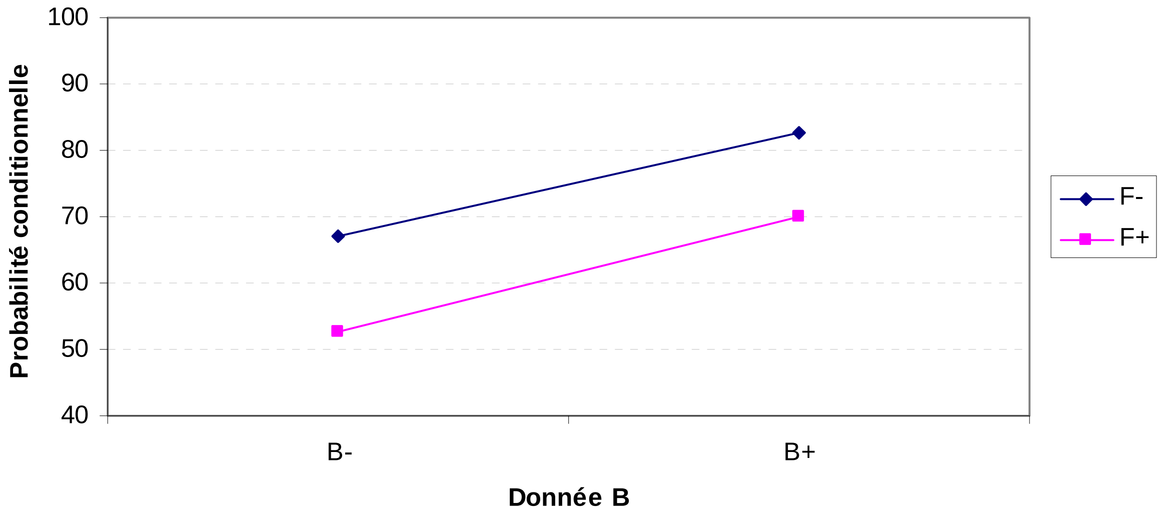


Figure 6.9. Intégration des données B et F par les élèves de 4ème et 3ème

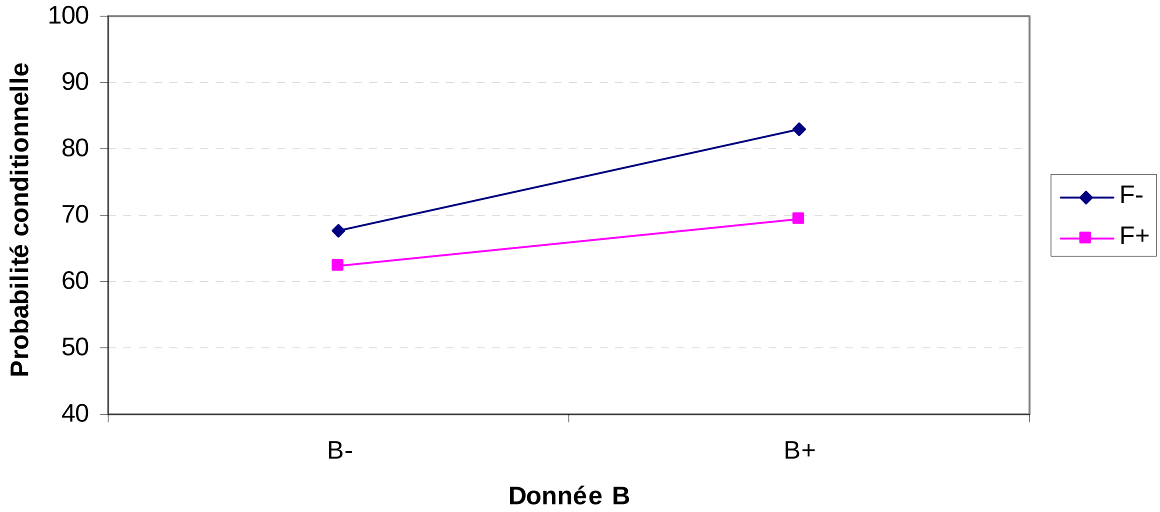


Figure 6.10. Intégration des données B et F par les élèves de 2nde et 1ère

L'effet d'interaction entre les données D et F n'est par contre significatif pour aucun des trois groupements de classes (cf. figure 6.11. ; 6.12. et 6.13.).

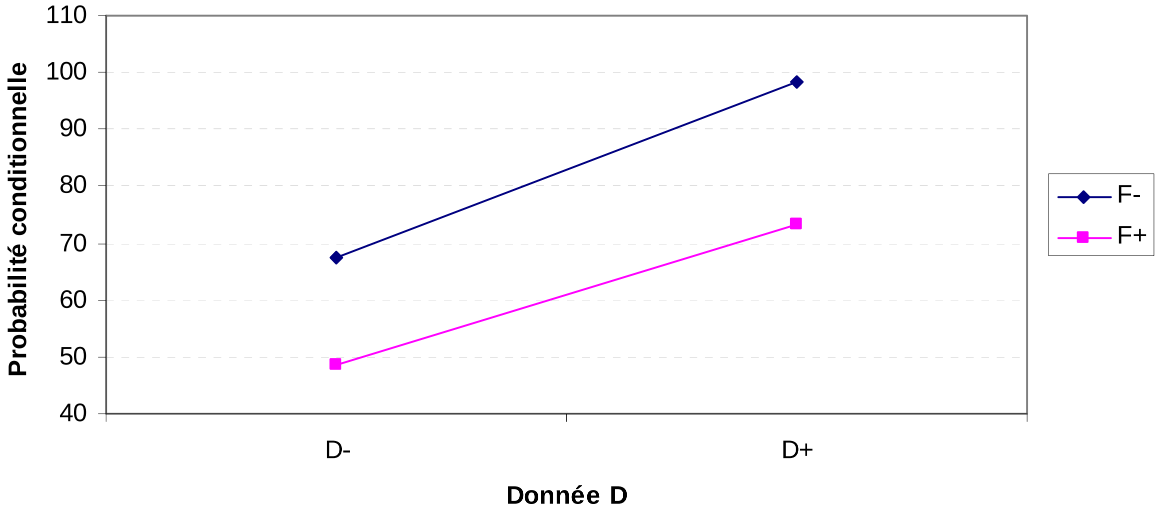


Figure 6.11. Intégration des données D et F par les élèves de 6ème et 5ème

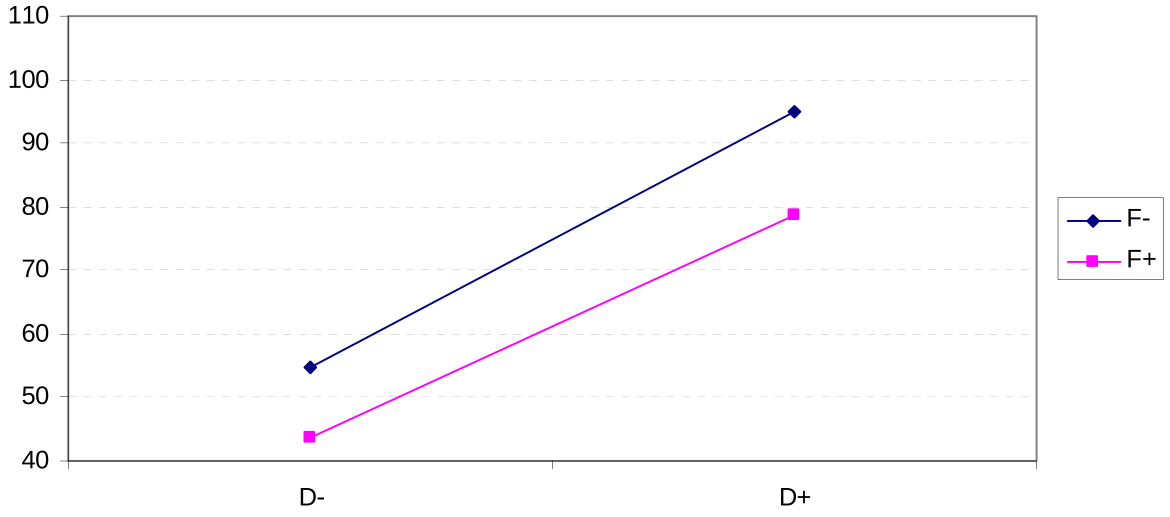


Figure 6.12. Intégration des données D et F par les élèves de 4ème et 3ème

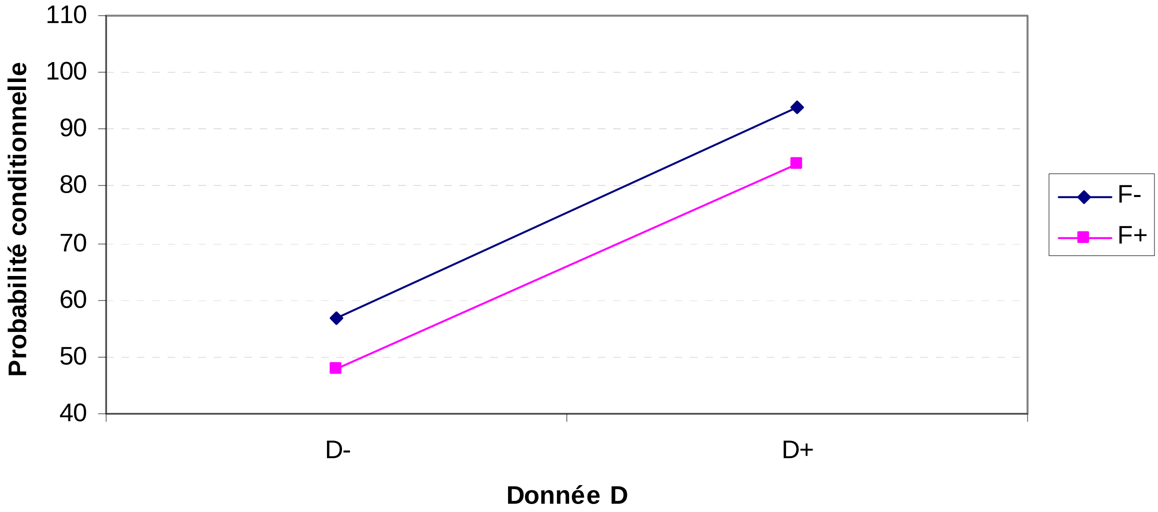


Figure 6.13. Intégration des données D et F par les élèves de 2nde et 1ère

Aucun des groupements de classes ne semble raisonner en prenant en considération les trois données B, D et F de l'énoncé ensemble.

Enfin, nous avons effectué une analyse dimensionnelle pour chacune des huit versions de problèmes bayésiens correspondant aux combinaisons des trois données binaires présentes dans l'énoncé. Ceci nous permet d'appréhender l'ordre de grandeur de la réponse à donner selon la version (cf. tableau 6.4.). Par exemple, quand les données b, d et f sont faibles - version A-, la réponse à donner l'est également.

	Valeurs de la réponse
A	Faible
B	Faible
C	Moyenne
D	Faible
E	Elevée
F	Moyenne
G	Elevée
H	Elevée

Tableau 6.4. Ordre de grandeur des probabilités révisées selon la version de problème

Une comparaison entre les moyennes marginales par version et par groupements de classes avec les résultats de notre analyse dimensionnelle précise que les participants des trois groupements qui tiennent compte des trois données de l'énoncé le font de la manière adéquate. Le sens de l'effet est cohérent. Plus précisément, les trois effets simples mis en avant pour les trois groupements de classes vont dans le bon sens. En d'autres termes, les participants ont appréhendé de façon correcte l'effet de chacune des trois données : quand la donnée B et la donnée D sont élevées, la probabilité conditionnelle l'est également. L'effet de la donnée F est inverse, c'est-à-dire que lorsque la donnée F est élevée, la probabilité conditionnelle est plus faible car il y a plus de chance de rencontrer des faux positifs. Les deux effets d'interaction mis en avant pour le groupement des classes de seconde et de première sont également cohérents avec notre analyse dimensionnelle. En résumé, quand les participants prennent en considération une probabilité, ils le font judicieusement.

IV. Discussion

Le premier objectif était de vérifier si, à lui seul, le format de la question pouvait influencer les réponses intuitives des participants. Plus précisément, lorsque les notions de probabilité et de fréquence sont exemptes de l'énoncé, une question fréquentiste ou une question probabiliste donnent-elles des réponses différentes de la part des individus ? La réponse est non. Nos résultats nous ont, en effet, montré que les différences entre les réponses des participants, selon que la question qui leur posée soit en fréquences naturelles ou en probabilités conditionnelles, ne semblent pas différentes. Ceci a été vérifié pour nos huit variantes de problème. Comme l'ont évoqué Evans, Handley, Perham, Over, & Thompson (2000), il est important que la question soit présentée sous le même format que l'énoncé. Nous complèterons cette remarque en apportant une précision : la question doit être cohérente avec l'énoncé, en termes de format de présentation -fréquentiste ou probabiliste-, lorsque l'énoncé est proposé de façon quantitative. Nous entendons par de façon quantitative, un énoncé avec des nombres. Toutefois, lorsque l'énoncé est qualitatif et que la seule référence à la notion de quantité qui y est faite est qualitative -avec des adjectifs numériques-, le fait que la question soit posée de façon fréquentiste ou probabiliste n'a plus de réelle importance. L'hypothèse explicative que nous proposons pour rendre compte de ce premier résultat réfère à la différence entre les fréquences et les probabilités. Plus précisément, les fréquences sont facilitatrices car, d'une part, leur insertion dans la règle de Bayes est plus aisée, et d'autre part, elles présentent une information numérique plus conséquente avec une présentation explicite du taux de base (Gigerenzer & Hoffrage, 1995). Notre recherche précédente nous a permis d'ajouter un autre argument en faveur de l'aspect facilitateur des fréquences naturelles : elles simplifient également la représentation et la compréhension du problème bayésien. Mais notons, qu'à lui seul, le format de la question ne permet pas de faciliter la démarche des participants. La représentation que se font les participants, et la compréhension qu'ils ont, des énoncés, sont équivalentes avec des données numériques qualitatives.

Le second objectif de cette expérience était de mettre en avant des profils cognitifs d'intégration de l'information statistique, et plus particulièrement bayésienne, par les participants, selon leur niveau scolaire et le format de la question. Le format de la question posée à la suite du problème ne présentant pas un effet significatif, nous discuterons uniquement cette notion de profils, sans tenir compte du format de la question. Nos analyses statistiques nous ont permis de faire ressortir un effet simple de chacune des trois données importantes dans les énoncés de problèmes bayésiens : la première -B- est la probabilité *a priori* de rencontrer l'événement E1. La deuxième -D- est la probabilité de rencontrer les deux événements E1 et E2 en même temps, soit la probabilité conditionnelle de rencontrer E1 sachant E2. La troisième donnée -F- est, en quelques sortes, le nombre de faux positifs, soit la probabilité d'avoir l'événement E2 sans l'événement E1. Sur le modèle d'Anderson (1979, 1981, 1996) et de sa théorie d'intégration de l'information, nous avons réalisé des analyses statistiques dans le but de décrire des profils cognitifs, sous forme d'algèbre cognitive. Les trois groupements de classes prennent en compte les trois données de l'énoncé. Toutefois, leur prise en considération n'est pas la même selon le niveau scolaire. En effet, pour les participants de collège, nous n'avons pas obtenu d'interaction significative. Ceci signifie que ces participants prennent en compte les trois données de l'énoncé mais de façon indépendante l'une par rapport à l'autre. Ces participants présentent un profil que nous qualifions d'additif, et non disjonctif comme nous l'avions supposé. En ce qui concerne les participants de lycée, nos analyses statistiques ont fait ressortir deux interactions comme significatives : l'interaction entre la donnée B et la donnée D, dans un premier temps, et l'interaction entre la donnée B et la donnée F, dans un second temps. Ces effets d'interaction sont la manifestation d'un profil multiplicatif. Ceci signifie que les participants intègrent dans leur raisonnement deux données en même temps. L'interaction entre la donnée B et la donnée D nous permet de savoir que les participants de lycée accordent de l'importance à l'apparition conjointe des deux événements. Pour savoir si la proportion de personnes présentant des céphalées parmi celles ayant un rhume est importante, les participants de lycée orientent leur attention sur les personnes ayant un rhume et celles ayant dans le même temps des maux de tête. Comme l'ont montré Gigerenzer, Hoffrage & Ebert (1998), certains professionnels négligent le taux de faux positifs. Il semblerait que nos participants de lycée commettent le même biais car l'interaction entre les trois données n'est pas significative. En ce qui concerne l'interaction entre la donnée B et la donnée F, ceci signifie que les participants intègrent dans leur raisonnement à la fois la donnée B et la donnée F. Au regard de la formule de Bayes, cela signifie qu'ils tiennent compte de la probabilité *a priori* et du nombre de faux positifs. Ceci nuance notre remarque précédente.

En conclusion, si les collégiens prennent en considération les trois données de l'énoncé mais seulement de façon additive, les lycéens semblent traiter de façon multiplicative les données de l'énoncé en intégrant dans leur raisonnement soit la probabilité *a priori* et l'apparition conjointe, soit la probabilité *a priori* et le nombre de faux positifs.

Précisons, grâce à notre analyse dimensionnelle, que les participants ont une représentation du problème qui peut être éloignée de celle qui découle de la règle de bayes, mais qui va dans le même sens. En effet, les participants donnent une réponse plus élevée lorsque la donnée B est « beaucoup » que lorsqu'elle est « peu ». En d'autres termes, ils ont l'intuition que lorsque la probabilité B est élevée, la probabilité *a posteriori* qui est demandée le sera aussi. Les participants semblent donc raisonner dans la bonne direction.

Chapitre 7 - Effet des performances en mathématiques, des fonctions exécutives de bas niveau et de la vitesse de traitement sur les performances bayésiennes, selon le niveau scolaire

I. Introduction

Selon Epstein (1998), le modèle de la « Cognitive-Experiential Self-Theory » permet de distinguer deux systèmes d'opérations mentales bien distinctes les unes des autres. ». La théorie du double processus évoque, d'une part le système 1, qui suit un mode de raisonnement, heuristique et plutôt intuitif, et d'autre part, le système 2, qui réfère à un mode plus logique et rationnel. Le système 1 est contraint par les limites de capacité de traitement du système cognitif humain. Pour Demetriou & Kasi (2001), trois niveaux d'organisation sont en interaction à propos des liens entre la cognition et la métacognition. Pour ces auteurs, le premier niveau équivaut à l'ensemble des processus de base nécessaires à la résolution de la tâche. Nous pensons que dans le domaine du raisonnement bayésien, ce système 2 et ce premier niveau correspondent aux structures logiques, aux capacités à traiter les probabilités. Rappelons que pour Piaget et Inhelder (1966), dès l'âge de 12 ans les enfants possèdent les compétences nécessaires au traitement des probabilités selon leurs règles propres. La littérature nous montre que les enfants ne sont pourtant pas bayésiens. Nous pensons que le fonctionnement exécutif est une explication possible de ce constat. De fait, Kahneman (2003) considère le système 2 comme étant exécutif. En parallèle, Demetriou & Kasi (2001) pensent que le premier niveau est en interaction étroite avec, d'une part, le deuxième niveau qui concerne les structures de raisonnement, de mémoire et les processus attentionnels, et d'autre part, le troisième niveau qui renvoie aux capacités cognitives des participants. Les fonctions exécutives peuvent donc être un facteur explicatif des erreurs de raisonnement, comme le suggère Houdé (2007).

L'approche fréquentiste propose dans le cadre du raisonnement bayésien une explication des erreurs de raisonnement : les individus ne sont pas faits pour traiter les probabilités conditionnelles mais plutôt des fréquences naturelles. Pour les auteurs de ce courant, les fréquences sont facilitatrices car elles simplifient la règle de Bayes, à savoir la formule mathématique.

Nos précédentes expérimentations ont montré que les individus peuvent s'approcher d'un raisonnement bayésien lorsque le matériel fait abstraction des difficultés mathématiques inhérentes à la règle de Bayes, et enlève le nombre de l'énoncé.

L'objectif de notre quatrième expérimentation est de tester l'impact des fonctions exécutives et du niveau scolaire en mathématiques sur les performances à des problèmes bayésiens qui nécessitent un calcul de la part des participants. Le niveau en mathématiques réfère à un modèle d'acquisition des schèmes logiques, tandis que le fonctionnement exécutif renvoie davantage à un développement fonctionnel.

Réviser son jugement c'est passer d'une probabilité à une autre, et donc ne pas se focaliser sur une probabilité particulière de l'énoncé, comme c'est le cas avec la stratégie de conservatisme.

Ainsi, nous postulons que pour résoudre des problèmes bayésiens en fréquences naturelles, les participants doivent mettre en œuvre des stratégies d'inhibition et de flexibilité. Les fréquences naturelles fournissent de façon explicite les taux de base. Ceci n'est pas le cas des probabilités conditionnelles. Nous supposons que chez les plus jeunes, cette différence dans la quantité d'information peut avoir un effet dans la différence de performances selon le format. Nous supposons que la mémoire de travail a un rôle dans les performances bayésiennes, et que ce rôle est différent selon le format (Sprenger & Dougherty, 2006). Enfin, nous pensons que le niveau académique en mathématique peut médier l'implication de fonctions exécutives.

II. Méthode

A. Participants

Les participants de cette quatrième expérience sont les mêmes que ceux de notre première expérience.

B. Matériel

Le raisonnement bayésien est évalué par deux séries de cinq problèmes équivalents présentés dans le format fréquentiste et dans le format probabiliste (cf. chapitre 4).

Quatre tests sont proposés à chaque participant afin d'évaluer leur vitesse de traitement, leur capacité d'inhibition, leur remise à jour en mémoire de travail et leur flexibilité.

1. Le test X-O, mesure de la vitesse de traitement

Pour ce test, le sujet doit, en trente secondes, comparer le plus possible de paires de lettres X et O, et cocher si elles sont composées de deux lettres identiques (OO ; XX) ou de deux lettres différentes (XO ; OX) (cf. annexe 2).
Plus le nombre de bonnes réponses est grand, plus le sujet traite rapidement l'information.

2. Le Stroop, mesure de l'inhibition

Le sujet passe pour ce test trois planches (cf. annexe 3). Pour les trois, il doit donner le plus grand nombre d'items en 45 secondes.

La première planche contrôle la vitesse de lecture du sujet : il doit lire le plus de noms de couleurs tous écrits en noir.

La seconde permet de tenir compte de la rapidité avec laquelle la personne dénomme les couleurs : il doit donner à haute voix la couleur avec laquelle sont dactylographiées des séries de croix.

La troisième planche teste la capacité du sujet à résister à l'automatisme qui consiste à lire le mot pour se focaliser sur la couleur de l'encre ; il n'y a jamais congruence entre ces deux informations.

La confrontation de ces trois nombres d'items réalisés en 45 secondes par planche - respectivement S_1, S_2 et S_3-, nous permet d'obtenir un score d'interférence :

$$\text{Score d'interférence} = S3 - \left(\frac{S_1 * S_2}{S_1 + S_2} \right)$$

Figure 7.1. Formule de calcul du score d'interférence dans le cadre du Stroop test

Meilleures sont les capacités du sujet à résister à l'interférence et plus son score s'éloigne positivement de zéro (Stroop, 1935).

3. Le N-Back test, mesure la remise à jour en mémoire de travail

Pour ce test, l'expérimentateur présente oralement une séquence de 30 lettres au sujet (cf. annexe 4). A chaque nouvelle lettre, ce dernier doit dire si oui ou non cette lettre est la même que l'avant dernière énoncée.

Le score du sujet correspond à la somme du nombre de bonnes réponses, un score de 28 étant le maximum (les deux premières lettres présentées ne demandent aucune réponse de la part du sujet). Plus ce score est élevé, plus le participant est capable de remettre à jour l'information traitée.

4. Le Plus-Minus test, mesure de la flexibilité

Ce test, adapté de Jersild (1927) et Spector & Biederman (1976), est composé de 3 planches de trente-deux nombres chacune, que le sujet doit remplir le plus rapidement possible (cf. annexe 5).

Pour la première liste le sujet doit ajouter 3 à chacun des nombres et soustraire 3 à ceux de la seconde. Pour la troisième, il lui faut alterner les opérateurs : il commence par ajouter 3 au premier nombre, puis le soustraire au second et ainsi de suite pour tous les nombres de la liste.

Le participant inscrit le résultat de son calcul à côté du nombre donné. L'expérimentateur chronomètre le temps nécessaire et informe en temps réel le sujet d'une éventuelle erreur qu'il doit alors corriger.

Le temps de réalisation à chacune des listes est noté respectivement T_1, T_2 et T_3. De la manière suivante, nous calculons un coût de shifting (cf. figure 7.2).

$$\text{Coût de shifting} = T_3 - \left(\frac{T_1 + T_2}{2}\right)$$

Figure 7.2. Formule de calcul du coût de shifting dans le cadre du Plus-Minus test

Plus le coût de shifting est faible, plus le sujet est flexible.

C. Procédure

Nos participants ont été rencontrés en deux temps : dans une première session, ils ont répondu aux problèmes bayésiens décrits dans la première expérience (cf. chapitre 4) et dans une seconde session qui a eu lieu une ou deux semaines après, selon les participants, ils ont été soumis aux trois tests de fonctions exécutives de bas niveau et au test de vitesse de traitement. La moitié de chaque groupement de classes a été rencontrée la semaine qui a directement suivi la première session et l'autre moitié la semaine d'encore après.

Pour chaque participant, la deuxième session est individuelle et a lieu dans le même ordre. Les participants répondent dans un premier temps au test X-O mesurant la vitesse de traitement. Ensuite, ils répondent au Stroop, le test d'inhibition. Puis le N-Back test leur est proposé afin de mesurer leur remise à jour en mémoire de travail. Enfin, Le Plus-Minus test leur est proposé pour évaluer leur flexibilité.

Cinq minutes de pause sont intercalées entre chaque test pour l'ensemble des participants.

III. Résultats

A. Evaluation des performances en mathématiques

Le niveau académique en mathématiques des participants est mesuré grâce à leur moyenne en mathématiques pour le semestre en cours. Afin de contrôler autant que possible d'éventuels biais de la part des enseignants -noter plus ou moins sévèrement par exemple, la moyenne brute des évaluations en mathématiques de chaque participant est standardisée. Nous utilisons pour ce faire une loi Z centrée réduite, en utilisant la moyenne et l'écart type de chaque classe pour obtenir une note centrée réduite en mathématiques par élève.

B. Performances bayésiennes

Rappelons dans un premier temps les résultats obtenus à la première expérience concernant le raisonnement bayésien. Aucun problème présenté en probabilités conditionnelles n'a été correctement résolu, et ceci est vrai pour les trois groupements de classes (cf. figure 4.1.). Des analyses *a posteriori* nous ont permis de préciser l'effet principal obtenu en fréquences naturelles selon le niveau scolaire des participants (le concept de niveau scolaire renvoie, comme c'est le cas depuis le début de ces travaux, à la classe dans laquelle se trouvent les participants. Il ne faut pas confondre ce concept avec celui de performances ou niveau en mathématiques). Ces analyses *post hoc* nous ont permis de mettre en évidence une différence significative entre le nombre de réponses bayésiennes données en fréquences naturelles par le troisième groupement de classes et celles données par le deuxième groupement de classes [$F(1 ; 117) = 14,05$, $p<.001$]. et une différence tout aussi significative entre les performances du troisième groupement et celles du premier [$F(1 ; 117) = 31,95$, $p<.001$]. Par contre, le nombre de problèmes résolus par les élèves de $6^{ème}$ et $5^{ème}$ n'est pas significativement différent du nombre de problèmes résolus par les élèves de $4^{ème}$ et $3^{ème}$ [$F(1 ; 117) = 3,62$, $p=.059$].

C. Fonctions exécutive et vitesse de traitement

Au moyen d'analyses de variance, nous avons mis en avant une amélioration des compétences exécutives avec l'avancée dans le cursus scolaire. Plus précisément, nos analyses montrent des différences en flexibilité [$F(2 ; 117) = 19,05$, $p<.01$], en inhibition

[F(2 ; 117) = 4,22, p<.05], en remise à jour en mémoire de travail [F(2 ; 117) = 7,52, p<.01] et en vitesse de traitement [F(2 ; 117) = 8,97, p<.01] (cf. tableau 7.1).

Fonctions exécutives	F (2 ; 117)	P
Flexibilité	19,05	<.001
Remise à jour	7,52	<.01
Inhibition	4,22	<.05
Vitesse de traitement	8,97	<.01

Tableau 7.1. Effet des groupements de classes sur les performances exécutives des participants

Des comparaisons *post hoc* nous apportent davantage de précisions concernant l'évolution des performances exécutives des participants selon leur appartenance à un groupement de classes (cf. tableau 7.2.). La flexibilité est la seule fonction exécutive qui diffère des trois autres de par son profil à travers les trois groupements de classes. En effet, les performances de flexibilité augmentent entre avec le passage entre le premier et le deuxième groupement de classes. Elles n'évoluent pas de façon significative entre la fin de collège et le lycée. Les autres fonctions exécutives ne semblent pas évoluer durant les années de collège, mais augmentent de façon significative avec le passage au lycée.

	4ème et 3ème	2nde et 1ère
6ème et 5ème	***F(1 ; 117) = 29,68, p<.001 (Flexibilité)*** F(1 ; 117) = 0,005, p=.94 (Remise à j) F(1 ; 117) = 0,02, p=.88 (Inhibition) F(1 ; 117) = 1,78, p=.19 (Vitesse)	***F(1 ; 117) = 27,42, p<.001 (Flexibilité)*** ***F(1 ; 117) = 11,05, p<.001 (Remise à j)*** ***F(1 ; 117) = 6,71, p<.05 (Inhibition)*** ***F(1 ; 117) = 17,21, p<.001 (Vitesse)***
4ème et 3ème		F(1 ; 117) = 0,04, p=.83 (Flexibilité) ***F(1 ; 117) = 11,52, p<.001 (Remise à j)*** ***F(1 ; 117) = 5,92, p<.05 (Inhibition)*** ***F(1 ; 117) = 7,93, p<.01 (Vitesse)***

Les effets en **gras** et *italique* sont significatifs.

Tableau 7.2. Comparaisons *post hoc* des performances exécutives selon les groupements de classes

D. Prédictions des performances bayésiennes par les fonctions exécutives et la vitesse de traitement

Nous avons ensuite effectué des analyses de régression multiple, avec comme prédicteurs des performances bayésiennes des participants, leurs scores des trois fonctions exécutives, celui de vitesse de traitement et leur moyenne centrée réduite en mathématiques, par groupements de classes. Les performances bayésiennes sont en réalité le nombre de problèmes réussis en fréquences naturelles, puisque nous avons eu un effet plancher en probabilités conditionnelles.

En accord avec ces analyses, les performances bayésiennes en fréquence ne semblent pas être prédites par les mêmes prédicteurs selon le groupement de classes (cf. tableau 7.3.). Seule la vitesse de traitement ressort comme prédicteur du nombre de problèmes fréquentistes résolus par les participants des classes de sixième et cinquième (β = .39 ; R^2 = .1894). L'inhibition est le seul prédicteur significatif du nombre de problèmes correctement résolus en fréquence naturelles par les élèves des classes de quatrième et cinquième (β = .36 ; R^2 = .2135). Enfin, les performances bayésiennes en fréquences naturelles des élèves de lycée peuvent être prédites par leurs aptitudes mathématiques (β = .42 ; R^2 = .2375). La vitesse de traitement, l'inhibition et le niveau en mathématiques expliquent respectivement 18,95 %, 21, 35 % et 23,75 % des performances bayésiennes en fréquences naturelles des élèves de $6^{ème}$ et $5^{ème}$, de $4^{ème}$ et $3^{ème}$, et de 2^{nde} et $1^{ère}$.

	Prédicteurs significatifs	Beta	R^2
6ème et 5ème	Vitesse de traitement	.39	.1894
4ème et 3ème	Inhibition	.36	.2135
2nde et 1ère	Niveau en mathématiques	.42	.2375

Tableau 7.3. Prédicteurs significatifs des performances bayésiennes en fréquences naturelles selon le groupement de classes

IV. Discussion

L'objectif de notre étude était de tester le poids des fonctions exécutives de bas niveau et de la vitesse de traitement dans les capacités de raisonnement probabiliste, plus particulièrement le raisonnement bayésien. Au regard de nos précédents résultats, il nous a semblé qu'une difficulté propre à ce type de raisonnement réside dans les contraintes mathématiques intrinsèques à la règle de Bayes. Le niveau en mathématiques nous a donc paru être un prédicteur important des performances bayésiennes quantitatives. Ces dernières réfèrent aux problèmes bayésiens pour lesquels il est nécessaire de calculer. Face à un effet plancher des performances des participants dans un format probabiliste, nous n'avons pas pu étudier les différences en termes de poids de l'implication du fonctionnement exécutif dans le raisonnement bayésien selon que les problèmes sont présentés en probabilités conditionnelles ou en fréquences naturelles.

Dans un premier temps, nous avons mis en avant un profil développemental des fonctions exécutives selon la classe des participants. Pour les trois fonctions exécutives de bas niveau et pour la vitesse de traitement, nous avons rencontré un développement par paliers. La flexibilité semble se développer au cours du collège et stagner entre la fin des années de collège et le début des années de lycée. Les trois autres fonctions évoluent de manière opposée. Ainsi, la remise à jour en mémoire de travail, l'inhibition et la vitesse de traitement ne semblent pas se développer au cours des années de collège. Par la suite, lors de l'arrivée au lycée, les participants semblent plus efficients dans ces capacités de remise à jour, d'inhiber l'information non pertinente et de traiter l'information de façon rapide. Nos résultats diffèrent de ceux décrits par Best, Miller et Jones (2009). Nous relativisons la portée de nos données dans le sens où nous n'avons utilisé qu'un seul test par fonction mesurée. De plus, le contact de terrain avec les participants nous a informé d'un fait important dont les chercheurs doivent tenir compte à l'avenir quant à la familiarité avec certains tests couramment utilisés. Le Stroop test est devenu très connu depuis son utilisation dans des softwares visant divers entraînements cérébraux sur certaines consoles très populaires. Nous avons donc trouvé dans ces jeux éducatifs de développement cognitif personnel la troisième planche du Stroop test et le Trail Making Test. Les participants sont donc, entre autres, habitués à inhiber la lecture du mot pour se focaliser sur la couleur de l'encre et à switcher entre une lettre et un chiffre.

Concernant nos analyses de régression, nos résultats nous suggèrent des implications modérées des fonctions exécutives et du niveau de performances en mathématiques. De plus, ces implications sont différentes selon la classe d'appartenance des élèves testés. La vitesse de traitement semble pouvoir expliquer 19 % des variations comportementales quant au raisonnement bayésien chez des élèves de sixième et de cinquième. Les autres fonctions n'interviennent pas de façon significative dans ce type de raisonnement auprès de cette population. Pour les participants du deuxième groupement de classes -quatrième et troisième-, l'inhibition joue un rôle important. Elle permet d'expliquer 21,35 % des variations des performances bayésiennes des participants. Nous avons constaté, que le fait marquant entre les participants du premier groupement et du deuxième groupement de classes est une diminution du nombre de réponses données avec une utilisation de la stratégie de conservatisme. Nous rappelons que cette stratégie réfère à la focalisation excessive de la part du participant sur la probabilité *a priori* de l'énoncé. La réponse donnée n'intègre aucune nouvelle donnée et le jugement n'est pas révisé. Nous pensons que les participants de ce groupement de classes de quatrième et de troisième sollicitent donc leurs capacités d'inhibition de façon à se détacher de cette probabilité *a priori* jusqu'alors prégnante. Pour les participants des classes de seconde et de première, ni les fonctions exécutives ni la vitesse de traitement ne semblent participer au raisonnement bayésien. L'hypothèse explicative que nous proposons s'articule autour de la distinction entre les deux modes de raisonnement. Nous pensons qu'à ce niveau scolaire, pour répondre à des problèmes en fréquences, même s'ils calculent, les participants opèrent davantage selon le système 1 qui est le mode plus intuitif. Certes la réponse à donner nécessite un calcul, toutefois, pour se représenter le problème, les participants utiliseraient davantage un mode heuristique. Quant au calcul, les fréquences naturelles le simplifiant, nous pensons que le système 2 n'est pas nécessaire pour l'addition de deux nombres au lycée. C'est pourquoi nous obtenons comme prédicteur significatif des performances bayésiennes fréquentistes de ces participants le niveau en mathématiques. Les élèves les plus à l'aise en mathématiques utiliseraient un mode heuristique, le système 1, pour répondre se représenter le problème dont la structure peut leur être familière. Il serait intéressant de vérifier cette hypothèse auprès de participants qui commenceraient à savoir répondre aux problèmes en probabilités.

Chapitre 8 - Analyse longitudinale des performances bayésiennes, selon le niveau scolaire et le format de présentation

I. Introduction

Nos précédentes expérimentations ont confirmé dans de larges mesures la thèse de l'approche fréquentiste selon laquelle les fréquences naturelles avaient un effet facilitateur par rapport aux probabilités conditionnelles sur les performances bayésiennes des individus (Gigerenzer & Hoffrage, 1995). Nous avons également confirmé le modèle développemental en ondes qui se chevauchent de Siegler (1999, 2000). En effet, comme Zhu & Gigerenzer (2006) l'ont fait sur des enfants chinois, nous avons montré que le développement des performances bayésiennes est davantage continu qu'il n'y parait sur des enfants et adolescents français. Plus précisément, nous avons montré que les participants émettent un raisonnement bayésien de plus en plus proche du raisonnement bayésien optimal, en référence à la règle de Bayes, avec l'avancée scolaire. Une analyse des stratégies nous a permis de ne pas nous contenter d'une vision dichotomique des réponses bayésiennes quantitatives -échecs versus bonnes réponses- en comptant le nombre de problèmes réussis. Cette analyse par stratégie nous a permis d'adopter une vision qualitative plus fine, et qui rend compte d'une hiérarchie des stratégies. Ainsi, les erreurs ne sont pas toutes équivalentes et les participants peuvent être en progrès même lorsqu'ils n'augmentent pas leur nombre de problèmes résolus. D'après Siegler (1999, 2000), les stratégies les moins efficaces disparaissent au profit de stratégies plus élaborées et davantage adaptées. Nous avons vu que c'était en partie le cas avec une analyse transversale, en utilisant des cohortes de vingt participants de différentes classes (cf. chapitre 4). Nous pensons qu'il serait intéressant de croiser à cette méthodologie une approche longitudinale, dans le sens où, il n'est pas dit que l'ensemble des participants opèrent au même rythme des changements dans l'utilisation de leurs stratégies. La littérature ne fait pas référence, à notre connaissance d'étude longitudinale dans cette perspective bayésienne. Dans le but de tester plus précisément le modèle de Siegler quant aux stratégies utilisées en raisonnement bayésien dans un format fréquentiste et d'apporter des informations complémentaires aux travaux de l'approche fréquentiste, nous réalisons une étude longitudinale auprès de participants ayant répondu aux deux premières expérimentations. L'intervalle entre les deux sessions de recueil de données est de dix-neuf mois. Les participants qui étaient en début d'une année scolaire *n* à la première session seront tous en fin d'année scolaire *n+1* lors de la seconde session. Nous postulons que la majorité des participants améliorera ses performances bayésiennes quantitatives dans un contexte fréquentiste. En d'autres termes, ils répondront de façon bayésienne à davantage de problèmes présentés en fréquences naturelles. Ces participants présenteraient une amélioration quantitative de leurs performances. Dans un second temps, nous pensons voir apparaître des réponses correctes en conditions probabiliste auprès de certains des anciens élèves de début de seconde et de première lors de la première sessions, et devenus respectivement des élèves de fin de première et de terminale. Cette hypothèse s'appuie sur les programmes scolaires d'une part, et sur les résultats de Zhu & Gigerenzer (2006), d'autre part, qui montrent que dans la population adulte, 47 % des problèmes sont correctement traités par les adultes. Notre troisième hypothèse est relative aux stratégies utilisées par les participants en condition fréquentiste. Nous pensons que les participants qui ne répondraient pas à davantage de problèmes à cette seconde session auraient remis à jour leurs stratégies utilisées, en écartant les stratégies les moins adaptées en faveur de stratégies plus proches de celles de la règle de Bayes. Ces participants présenteraient une amélioration qualitative de leurs performances. Cette étude longitudinale est l'étude 1bis, en référence à la première expérience. Enfin, nous pensons qu'en parallèle à cette augmentation des performances en raisonnement probabiliste selon un mode logique et rationnel, nous pouvons constater une amélioration des performances heuristiques et intuitives dans le même type de raisonnement. Notre dernière hypothèse porte donc sur les réponses intuitives des participants sur un continuum. Plus précisément, et au regard des résultats de notre deuxième expérience, nous postulons que les participants donneront une réponse plus proche de celle calculée avec la règle de Bayes. Cette amélioration des performances heuristiques serait d'autant plus marquée en fréquences naturelles pour les collégiens et en probabilités pour les lycéens. Cette étude longitudinale est l'étude 2bis, en référence à la deuxième expérimentation.

II. Méthode

A. Participants

Les participants de cette expérience sont une partie de ceux de la première et de la deuxième expérimentation, dix-neuf mois après. Cet intervalle entre les deux recueils de données correspond à presque deux années scolaires. Plus précisément, les participants qui étaient en début de sixième à la première session -expérience numéro 1, cf. chapitre 4- sont en fin de classe de cinquième lors de la seconde session expérimentale. Il en est de même pour tous les participants.

La difficulté d'un recueil selon une méthode longitudinale réside dans la possibilité de retrouver les participants quelques temps plus tard. Le fait d'avoir recruté les participants dans des établissements scolaires présente l'avantage de pouvoir plus aisément les retrouver. Les participants qui ont quitté le collège pour aller au lycée entre les deux sessions sont pour la plupart aller dans le même établissement. Nous avons donc pu retrouver une partie de nos participants. En résumé, nous avons réussi à rencontrer à nouveau dix participants de chaque classe pour chacune des deux premières expérimentations. Notons que les participants qui étaient en sixième sont en cinquième et que ceux qui étaient en première sont en terminale. Nous avons donc, pour chaque expérience, dix participants par classe, de la cinquième à la terminale (cf. tableau 8.1. et tableau 8.2.). La parité a pu être respectée.

	Classes					
	Cinquième	Quatrième	Troisième	Seconde	Première	Terminale
Effectifs (Garçons ; Filles)	10 (5 ; 5)	10 (5 ; 5)	10 (5 ; 5)	10 (5 ; 5)	10 (5 ; 5)	10 (4 ; 6)

Tableau 8.1. Répartition des sujets de l'étude longitudinale 1bis par classe

	Classes					
	Cinquième	Quatrième	Troisième	Seconde	Première	Terminale
Effectifs (Garçons ; Filles)	10 (5 ; 5)	10 (5 ; 5)	10 (5 ; 5)	10 (5 ; 5)	10 (5 ; 5)	10 (5 ; 5)

Tableau 8.2. Répartition des sujets de l'étude longitudinale 2bis par classe

B. Matériel

Les mêmes problèmes que ceux utilisés dans la première expérimentation (cf. chapitre 4) sont proposés aux participants de l'étude longitudinale 1bis.
Les participants de l'étude longitudinale 2bis sont soumis aux problèmes avec continuum utilisés dans la deuxième expérience (cf. chapitre 5).

C. Procédure

1. Etude longitudinale 1bis

La procédure est la même que pour la première expérience. Les soixante participants doivent répondre à cinq problèmes présentés de la condition fréquentiste pour lesquels ils doivent donner une réponse sous forme de fréquence naturelle.

Les vingt participants des classes de première et de terminale lors de cette seconde session doivent dans un premier temps répondre aux cinq problèmes présentés dans la condition probabiliste pour lesquels ils doivent fournir une réponse en pourcentages. Seuls les participants des classes de première et de terminale ont été testés dans cette seconde session sur les problèmes probabilistes, de façon à ne pas frustrer les participants au regard de l'effet plancher obtenu au cours de la première expérimentation.

2. Etude longitudinale 2bis

La procédure est la même que celle appliquée lors de la deuxième expérience. Les participants répondent dans un premier temps aux problèmes avec un énoncé probabiliste, puis à ceux présentés en fréquences naturelles.

Pour les deux études longitudinales, les sessions ont lieu en groupe de dix participants. Le temps est libre, et les participants n'ont pour seul matériel que le document distribué et un crayon.

III. Résultats

A. Performances quantitatives

1. Contexte probabiliste

Les performances quantitatives sont testées lors de l'étude longitudinale 1bis.

Lors de la première session (cf. chapitre 4), aucun participant n'a résolu, conformément à la règle de Bayes, de problèmes bayésiens présentés avec un contexte probabiliste. Lors de cette seconde session, quatre participants ont correctement résolus les cinq problèmes proposés en probabilités conditionnelles. Un d'entre eux est un élève de première, et les trois autres sont des élèves de terminale. Ces quatre participants étaient tous bayésiens lors de la session 1 lorsque les problèmes étaient présentés en fréquences naturelles. En résumé, à la session 1 quatre élèves de début de classes de seconde et de première ne résolvaient aucun problème bayésien présentés en probabilités conditionnelles mais résolvaient ceux proposés en fréquences naturelles. A la session 2, ces mêmes élèves respectivement en fin de classes de première et de terminale résolvent tous les problèmes indépendamment du format de présentation de l'énoncé.

2. Contexte fréquentiste

Nous n'avons pas réalisé d'analyses statistiques quantitatives telles des analyses de variance à mesures répétées pour tester l'effet développemental entre les deux sessions, et ce pour deux raisons. La première est que pour certains groupes, nous n'avons pas de variance. Par exemple, aucun des élèves de sixième de la première session que nous avons retrouvés à la session 2 ne résolvaient de problèmes en fréquences en début de leur classe de sixième. De la même manière, tous les élèves en classe de première lors de la session 1 ont résolu tous les problèmes. La seconde raison est qu'il ne nous a pas semblé pertinent d'effectuer ce type d'analyses sur des groupes de seulement dix participants. Nous pensons qu'une analyse qualitative est donc plus appropriée.

Nous avons comparé le nombre de problèmes fréquentistes résolus par les participants aux deux sessions selon leur classe d'appartenance lors de la première session (cf. tableau 8.3.). Ainsi, nous constatons que pour les participants qui étaient en début de sixième à la session 1 aucun problème n'a été résolu. Lors de la session 2, cinq problèmes ont été résolus. Ces cinq problèmes ont été résolus par le même participant.

	Session 1	Session 2
6ème	0	5
5ème	0	10
4ème	3	17
3ème	11	38
2nde	25	50
1ère	50	50

Tableau 8.3. Nombre de problèmes fréquentistes répondus de façon bayésienne selon la classe des participants à la session 1

Concernant les élèves de classe de cinquième à la session 1, ils sont passés de zéro problème à 10 problèmes résolus. Nous remarquons que ces dix problèmes ont été résolus par deux d'entre eux qui ont donc utilisé pour les cinq problèmes proposés une stratégie bayésienne.

Les élèves de quatrième, troisième et seconde ont également amélioré leurs performances intra groupe. L'analyse des stratégies utilisées qui suit va nous apporter des précisions quant à l'augmentation de ces performances.

B. Performances qualitatives

Dans un deuxième temps, nous nous intéressons aux stratégies utilisées pour répondre aux problèmes fréquentistes par les participants selon leur classe lors de la première session. Nous constatons des changements dans l'utilisation des stratégies. Par exemple, entre les sessions 1 et les sessions 2, les participants de chaque classe semblent délaisser les stratégies les plus basses sur le graphique -les stratégies les plus basses sont les plus éloignées de la stratégie bayésienne- pour utiliser dans de plus grandes proportions la stratégie bayésienne ou diverses stratégies s'en approchant (cf. figure 8.1.).

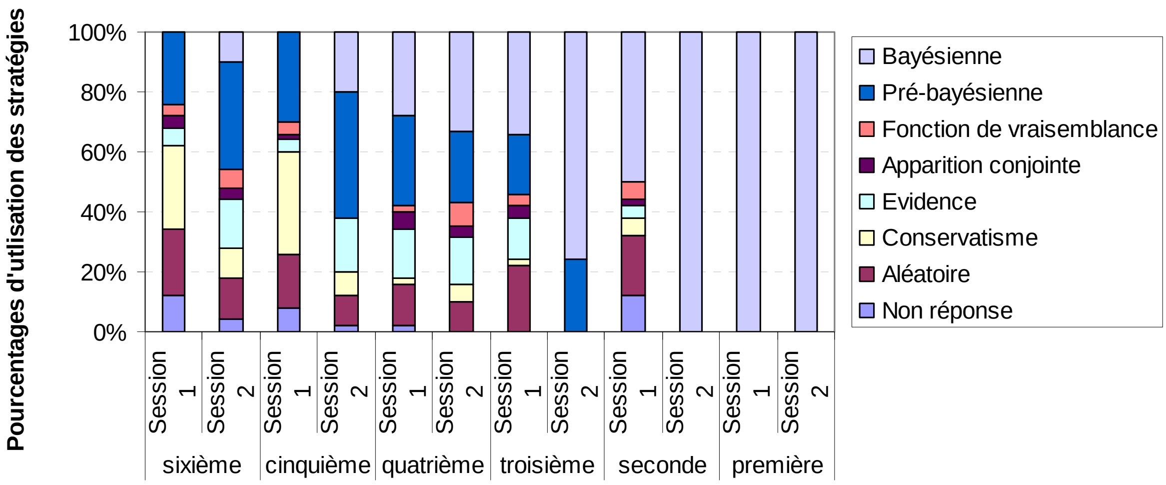
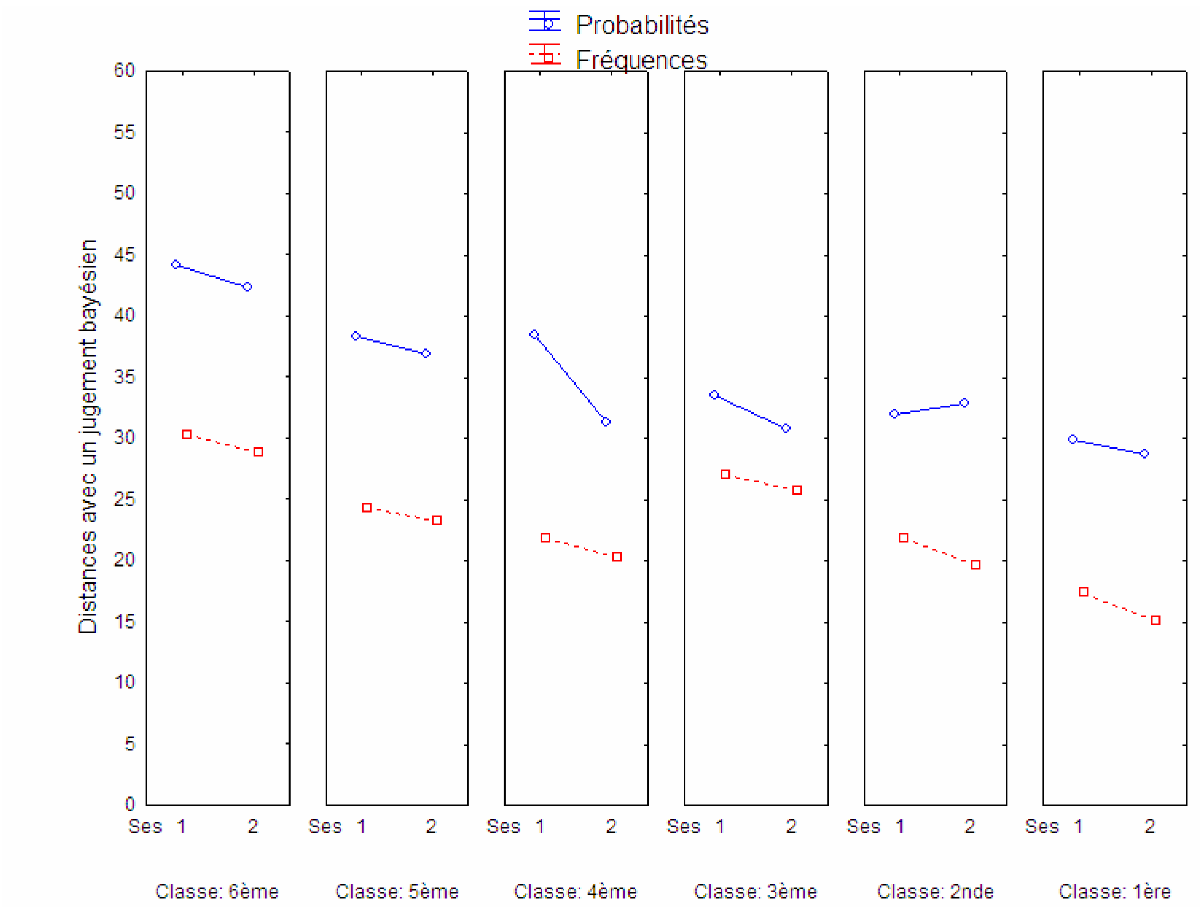


Figure 8.1. Pourcentages d'utilisation des différentes stratégies pour répondre à des problèmes fréquentistes selon la classe et les sessions

C. Performances heuristiques

Les performances heuristiques et intuitives ont été testées lors de la deuxième expérience de nos travaux en tant que première session (cf. chapitre 5). Elles réfèrent aux écarts entre les réponses bayésiennes calculées par la formule et les réponses intuitives des participants sur un continuum. Nous effectuons des comparaisons classiques entre les performances des participants par groupe entre les deux sessions (cf. figure 8.2.). Il ya donc un effet du format -les fréquences sont facilitatrices par rapport aux probabilités conditionnelles-, un effet de la classe -avec l'avancée dans le cursus scolaire les participants estiment de plus en plus finement- et un effet de développement -les participants sont meilleurs à la seconde session qu'à la première session-. Une analyse de variance en mesures répétées, à interpréter avec prévoyance au regard de considérations méthodologiques déjà discutées telles la taille des échantillons, semble nous confirmer cette interaction triple entre les trois variables [$F(5 ; 54) = 3{,}31$, $p<.05$)]. Nous remarquons toutefois des valeurs moyennes plus élevées pour les participants qui étaient en classe de troisième lors de la session 1. Ceci s'explique par le fait que deux participants répondent systématiquement de façon très éloignées de la réponse correcte. Deux participants sur un échantillon de dix suffisent à orienter significativement la moyenne du groupe.

Nous remarquons (cf. figure 8.2.) que l'amélioration la plus conséquente entre les deux sessions renvoie aux participants en classe de quatrième dans la condition probabiliste.



Ses 1 : Session 1 ; Ses 2 : Session 2

<u>Figure 8.2.</u> Analyse longitudinale entre les deux sessions selon le format et la classe des participants lors de la première session

IV. Discussion

Le but de cette dernière expérience était de rendre compte de l'évolution de trois types de performances bayésiennes chez des collégiens et lycéens français : les performances quantitatives -combien de problèmes résolus de façon bayésienne selon le format-, les performances qualitatives -évolution des stratégies- et performances heuristiques -estimation de la réponse sur un continuum-.

L'approche utilisée pour cette expérimentation était de rencontrer des participants déjà testés dix-neuf mois plus tôt. Cet intervalle de dix-neuf mois correspond à un laps de temps écoulé entre le début d'une année scolaire et la fin de la suivante. Par exemple, les participants étant en début de classe de sixième lors de la session 1 ont été testés en fin de cinquième lors de la session 2. Cette approche avait pour but d'apporter des informations complémentaires à l'approche transversale utilisée dans la littérature et dans nos précédentes expériences.

Concernant les performances bayésiennes testées dans la littérature, c'est-à-dire le nombre de problèmes bayésiens résolus, nous avons montré qu'il augmente lorsque les participants passent des classes de sixième et cinquième aux classes de quatrième et de troisième, d'une part, et lorsqu'ils qu'ils quittent le collège et arrivent au lycée dans un contexte fréquentiste. Lorsque les énoncés sont présentés en probabilités conditionnelles, seuls quelques élèves de fin de lycée, tout au moins dans les filières économiques et sociales, sont en mesure de répondre de façon bayésienne.

Concernant l'évolution des stratégies utilisées par les participants, les résultats de cette étude longitudinale vont dans le sens des travaux de Siegler (1999 ; 2000). Nous constatons, en effet, que certaines stratégies éloignées de la stratégie bayésienne disparaissent des outils des participants avec l'avancée scolaire, au profit de l'apparition, ou d'une utilisation plus constante plus massive, de stratégies plus proches de la stratégie bayésienne. Ces stratégies seraient davantage adaptatives que les précédentes.

Enfin, notre approche longitudinale nous a permis de vérifier l'évolution des performances bayésiennes intuitives ou heuristiques des participants selon leur niveau scolaire et le format de présentation des problèmes bayésiens. Nous avons constaté que les participants entre les deux sessions expérimentales affinaient leur représentation de problèmes, et ceci, semble-t-il, tant lorsque l'énoncé est en fréquences naturelles que lorsqu'il est en probabilités conditionnelles. L'amélioration qui semble être la plus flagrante (cf. figure 8.1.) est celles des participants initialement en début de classe de quatrième et qui sont donc lors de la session 2 en fin de classe de troisième dans un contexte probabiliste. Ceci signifie qu'au cours des deux années de quatrième et de troisième, les participants arrivent à mieux se représenter les probabilités, même s'ils ne sont toujours pas capables de les traiter correctement au regard de la formule bayésienne. Tout semble indiquer que leur raisonnement probabiliste devient bayésien lors de ces deux années, sans pour autant réussir à le mettre en place. Nous pensons que cet exemple décrit une amélioration des compétences mais pas encore des performances bayésiennes, dans le sens où ces participants semblent être capables de s'approcher de la réponse, sans pour autant l'atteindre.

Troisième partie

Discussion générale

Chapitre 9 - Discussion générale

L'objectif principal de ce travail de recherche était d'appréhender les performances bayésiennes de collégiens et de lycéens français. Ce travail s'inscrit dans une démarche développementale. Le critère que nous avons retenu pour rendre compte du développement n'est pas l'âge mais le niveau scolaire, plus précisément les performances bayésiennes sont étudiées selon la classe des participants. Un deuxième facteur a été central dans ces travaux. Il s'agit du format de présentation des énoncés. Nous avons tenu compte de la distinction classique entre les problèmes probabilistes et les problèmes fréquentistes. Les premiers sont présentés sous forme de probabilités conditionnelles et les seconds ont un énoncé en fréquences naturelles. La question posée quant à la probabilité révisée est en accord avec le format de l'énoncé, comme l'ont suggéré Evans, Handley, Perham, Over, & Thompson (2000). Nous apportons néanmoins une précision : la question à elle seule ne permet pas de modifier la représentation des individus lorsque l'énoncé ne contient pas de nombre. En effet, nous avons montré que les profils cognitifs sont les mêmes lorsque la notion de grandeur présente dans l'énoncé n'est évoquée que par des adjectifs de quantités tels Peu ou Beaucoup, sans donner de nombre. Le nombre implique une unité, et réfère donc à une probabilité, c'est-à-dire un pourcentage, soit une proportion, c'est-à-dire une fréquence. Conformément aux précisions apportées par Hoffrage, Gigerenzer, Krauss, & Martignon (2002), les fréquences que nous avons étudiées sont des fréquences naturelles et non des fréquences normalisées.

Cette discussion générale va s'articuler autour de trois axes d'interprétation et de réflexions. Le premier renvoie aux performances bayésiennes quantitatives. Le deuxième réfère aux performances bayésiennes qualitatives. Enfin, le troisième concerne les performances bayésiennes heuristiques.

I. Les performances bayésiennes quantitatives

La notion de *performances bayésiennes quantitatives* renvoie tout simplement au nombre de problèmes bayésiens dont la réponse est correcte. Ce sont celles qui sont étudiées dans la littérature dans un nombre croissant de publication. Par réponse correcte, nous entendons que la probabilité conditionnelle donnée par le participant est celle obtenue par la règle de Bayes (cf. figure 2.2.). Une réponse correcte équivalente est une fréquence naturelle conforme à celle calculée par la règle de Bayes simplifiée (cf. figure 2.3.). La distinction entre ces deux types de réponses a été popularisée par l'approche fréquentiste initiée par Gigerenzer, Hell, & Blank (1988). Cet auteur reprend l'idée chère aux associationnistes (Hume, 1739, 1951) selon laquelle l'homme, tout comme d'autres espèces, a au niveau phylogénétique une habitude des fréquences naturelles. En d'autres termes, au cours de son évolution l'homme a développé une familiarité à l'égard des fréquences naturelles. Par exemple, l'être humain enregistre naturellement les fréquences d'événements (par exemple, « parmi les cinq dernières fois où je suis allé pêcher à cet endroit, j'ai réussi à prendre du poisson par trois fois » ; un exemple plus contemporain pour lequel l'homme n'a pas à attraper sa nourriture pourrait être « les deux fois où je suis allé à ce restaurant j'ai été déçu »). Grâce aux fréquences naturelles, l'homme s'adapte à son environnement (Brunswik, 1939 ; Gallistel, 1990 ; Shanks, 1991). Ceci nous pousse à croire que l'être humain raisonnait de façon bayésienne bien avant que Bayes ne propose sa formule mathématique. Les événements étant présents dans la nature sous un format fréquentiste -les probabilités sont une invention humaine même si selon Gould (1992) l'être humain n'est pas fait pour raisonner avec les règles des probabilités-, les problèmes posés en fréquences naturelles sont donc davantage en adéquation avec la réalité (Gigerenzer & Murray, 1987). L'isomorphisme entre l'environnement et le problème est une condition, certes non suffisante, mais nécessaire à une analyse rationnelle de la part de l'être humain (Brunswik, 1939 ; Anderson, 1991).

Le fait que les fréquences soient davantage familières aux individus, peut être un argument en faveur de l'approche fréquentiste. Pour les auteurs qui s'inscrivent dans cette démarche, les fréquences sont facilitatrices car elles résonnent davantage dans l'esprit des individus, qui, ainsi, raisonnent de façon plus éclairée, et plus adaptée. D'après nous, cette familiarité est à rapprocher des théories du double processus et de la théorie « Cognitive-Experiential Self-Theory » -ou CEST- de Epstein (1983, 1994, 1998). Pour Epstein et de nombreux chercheurs, l'esprit humain serait divisé en deux. Cette séparation n'est bien évidemment pas localisationniste ni phrénologique, mais fonctionnelle. Plus précisément, les individus auraient la possibilité de traiter l'information et donc de raisonner de deux manières différentes, qui renvoient chacune à un système, ou à un mode, différent. Le système 1 est un mode de fonctionnement intuitif. Il a lieu de façon plus ou moins consciente. Il est intuitif, rapide, associatif et holistique, et s'appuie sur l'expérience du sujet. Ce dernier traite souvent les problèmes en les simplifiant (Epstein, Lipson, Holstein, & Huh, 1992 ; Evans et Over, 1996 ; Sloman, 1996 ; Stanovich & West, 2000 ; Stanovich & West, 2008). Le système 2, quant à lui, est un mode de fonctionnement analytique. Il est sous tendu par des connaissances théoriques des individus et renvoie à l'application de règles de la logique formelle.

Dans le cadre du raisonnement probabiliste, la règle de Bayes est considérée comme un modèle normatif. Nous pensons qu'un individu qui raisonne conformément à la règle de Bayes sollicite le système 2, soit le mode analytique. Pour ce faire, l'individu a besoin de connaissances théoriques concernant les probabilités. Selon les travaux pionniers de Piaget & Szeminska (1941), les enfants acquièrent la notion du nombre grâce à la construction simultanée du système ordinal et du système cardinal. Pour Piaget et Inhelder (1951, 1966, 1975), les enfants acquièrent au cours de leur développement, et à la suite du concept de nombre, le concept de probabilité. Cette acquisition se fait en trois étapes. Les enfants passent d'un stade ou le déductif et le non déductif sont indifférenciés à un troisième stade à la fin duquel ils comprennent la notion de probabilité et sont capables de les manier, et ce autour de douze ans. Ce stade est qualifié de genèse des opérations formelles (Hoemann & Ross, 1971 ; Kreitler & Kreitler, 1986).

Nous sommes confrontés à un paradoxe dont l'enjeu est, ni plus ni moins, un argument en faveur d'une rationalité ou d'une irrationalité humaine. Ce paradoxe est le suivant : comment peut-on expliquer que des enfants soient capables de raisonner de façon probabiliste, en accord avec les travaux de Piaget et ceux de Zhu & Gigerenzer (2006), alors que l'homme de la rue ne semble pas l'être, pour cause d'une rationalité limitée (Simon, 1956). D'après Murphy, Lichtenstein, Fischhoff & Winkler (1980) cet être naïf est désorienté lorsqu'il s'agit d'interpréter des probabilités, par exemple diffusées par les médias. L'homme expert ne semble pas en restes lorsqu'il s'agit d'utiliser, ou de communiquer, de façon maladroite les probabilités (Budescu & Wallsten, 1995 ; Gigerenzer, Hoffrage & Ebert, 1998 ; Gigerenzer, Hoffrage & Kleinbölting, 1991). Par exemple, des psychiatres ou des médecins semblent très maladroits pour informer leurs patients des risques qu'ils encourent en suivant un traitement, et même, dans certains cas semblent ne pas interpréter eux-mêmes les risques réels que la simple règle de Bayes nous quantifie dès que l'on y insère correctement les informations relatives aux probabilités des événements.

L'adjectif simple ne signifie pas ici facile à utiliser. En effet, multiplier et diviser des probabilités entre elles, n'apparaît pas être une tâche aisée, au regard de la littérature. L'approche fréquentiste préconise donc une utilisation de fréquences naturelles et non de probabilités conditionnelles dans le sens où, elles permettent l'utilisation d'une règle de Bayes simplifiée.

Un autre argument en faveur de l'aspect facilitateur des fréquences naturelles est qu'elles sont plus informatives que leurs cousines les probabilités conditionnelles. En effet, elles donnent de façon explicite le taux de base, ce qui n'est pas le cas des probabilités conditionnelles d'après Hoffrage, Gigerenzer, Krauss, & Martignon (2002).

Ces différents arguments servent à expliquer les différences de performances quantitatives de nos participants selon que les problèmes que nous leur avons proposés sont en probabilités conditionnelles ou en fréquences naturelles (cf. expérience 1 et 5). Les collégiens et lycéens que nous avons rencontrés, qui semblent représentatifs de tous les collégiens et lycéens qui suivent le système éducatif français, ne sont pas en mesure de répondre à des problèmes bayésiens lorsque ceux-ci sont présentés en probabilités conditionnelles. Le fait qu'un élève de fin de première et trois élèves de fin de terminale parmi dix soient capables de répondre de façon bayésienne aux problèmes présentés dans un format probabiliste, nous laisse suggérer qu'au cours de ces deux années scolaires, les participants acquièrent des connaissances théoriques relatives à l'utilisation des probabilités ou développent des compétences de base requises pour mettre en place ce type de raisonnement. Notre quatrième expérimentation ne nous a pas permis de trancher entre cette première approche structuraliste et cette seconde approche fonctionnaliste. En effet, devant l'effet plancher des performances obtenues auprès des participants aux problèmes bayésiens présentés en probabilités conditionnelles, il ne nous a pas été possible de tester un éventuel effet des compétences exécutives dans le but d'expliquer le développement des performances de raisonnement bayésien dans ce format. En ce qui concerne les problèmes présentés en fréquences naturelles, nous constatons un effet développemental patent. Plus précisément, les performances de collégiens en classes de sixième et de cinquième sont plus faibles que celles de collégiens de classes de quatrième et de troisième, qui sont elles-mêmes plus faibles que celles de lycéens. Ces résultats obtenus par une méthode transversale (cf. expérimentation 1) sont confirmés par une méthode longitudinale (cf. expérimentation 5). Ces deux constats - effet plancher des performances bayésiennes en probabilités conditionnelles et amélioration des performances en fréquences naturelles- laissent à penser que les progrès, qu'ils soient relatifs à des connaissances théoriques ou à des processus sous jacents permettant la mise en place de raisonnement de haut niveau- ont lieu de façon différentielle selon le format de présentation des énoncés. Notre quatrième expérience nous indique que les performances quantitatives en fréquences naturelles sont, pour une certaine part, dépendantes du fonctionnement exécutif ou du niveau d'aptitudes en mathématiques. Rappelons que la vitesse de traitement semble impliquée dans la résolution de problèmes en fréquences naturelles chez des participants de classes de sixième et de cinquième, que l'inhibition intervient dans le raisonnement bayésien fréquentiste des élèves de quatrième et de troisième, et que le niveau en mathématiques des élèves de seconde et de première peut expliquer une partie des différences observées en termes de nombres de problèmes résolus. Nous pensons que la vitesse de traitement est un indicateur du niveau cognitif global, ce qui peut expliquer le lien entre cette capacité à traiter rapidement l'information et le raisonnement bayésien en fréquences naturelles. Ces participants de sixième sont pour la plupart âgés de douze ans. C'est l'âge auquel selon Piaget et Inhelder (1951, 1966, 1975) les enfants ont acquis les compétences relatives aux concepts de probabilités. Toutefois, en accord avec les programmes scolaires, ils n'ont pas encore reçu d'enseignement à ce sujet. Toujours en référence aux programmes scolaires, ce n'est qu'en classes de seconde que les élèves sont sensibilisés aux notions de probabilités, plus précisément aux probabilités conditionnelles, et aux axiomes de Kolmogorov les régissant. Les performances bayésiennes quantitatives étant mesurées par une utilisation d'une formule mathématique, il nous semble indispensable de tenir compte du niveau en mathématiques des participants, tout au moins en discussion. Ce niveau en mathématiques est un élément culturel et est relatif aux contenus nationaux. Nos participants sont plus proches des participants allemands que des participants chinois qui semblent plus précoces que leurs homologues européens (Lücking, 2004). Pour Artelt, Baumert, Klieme, Neubrand, Prenzel, Schiefele, Schneider, Schümer, Stanat, Tillmann, & Weiß (2001) et Stern, Rode, Ge, & Zhu (2001), la motivation des enfants chinois suffit à expliquer leurs performances mathématiques plus précoces. Une autre hypothèse serait la plus grande régularité de leur système numérique. Quatre grands types éducatifs sont décrits en Europe : le modèle anglo-saxon, le modèle latin, le modèle germanique et le modèle scandinave (le programme PISA qui définit ces modèles, constate des différents niveaux mathématiques entre les élèves selon leur pays). Dans le but de tester les différences de performances bayésiennes, selon le modèle éducatif, nous sommes en train de monter une collaboration avec une équipe irlandaise.

En résumé, le format de présentation des énoncés et des questions posées a un effet différentiel sur les performances bayésiennes quantitatives des participants. Plus précisément, la majorité des participants répond à de plus en plus de problèmes présentés en fréquences naturelles entre les classes de sixième et de première, alors que seuls quelques élèves de fin de lycée semblent capables de répondre aux énoncés probabilistes. Comme l'évoque Feynman (1967), les fréquences naturelles sont équivalentes aux probabilités conditionnelles d'un point de vue mathématique, mais non d'un point de vue psychologique, dans le sens où elles n'amènent pas les mêmes performances.

II. Les performances bayésiennes qualitatives

Pendant de nombreuses années, seules les performances bayésiennes quantitatives ont été étudiées. Gigerenzer, porte étendard de l'approche fréquentiste a proposé une analyse qualitative des réponses. Pour cet auteur, il est possible de référer à différentes stratégies en raisonnement bayésien selon les données utilisées par les participants pour répondre à la question posée, et selon la façon dont ils ont utilisé ces données (cf. figure 2.5. et tableau 2.1.). Zhu & Gigerenzer (2006) décrivent différentes stratégies et montrent qu'elles sont toutes utilisées par des participants d'un même niveau scolaire. Il n'y a donc pas de correspondance terme à terme entre les la classe des élèves et les stratégies qu'ils utilisent, mais plutôt une concurrence entre les stratégies au sein d'un même groupe d'élèves, et parfois même au niveau intra individuel. Ce constat est en accord avec le modèle en ondes qui se chevauchent de Siegler (1999, 2000). Pour cet auteur, les participants peuvent fluctuer dans l'utilisation des stratégies, et d'un point de vue développemental, les stratégies les plus archaïques sont laissées de côté au profit de stratégies plus évoluées et plus adaptatives. Nos résultats vont dans ce sens. En effet, une analyse transversale (expérience 1) et une analyse longitudinale(expérience 5) confirment ce modèle. Au sein d'une même classe, les stratégies sont toutes utilisées. Ce qui est caractéristiques d'une classe, c'est donc la fréquence d'utilisation d'une stratégie plus qu'une autre. Par exemple, la stratégie de conservatisme est beaucoup moins utilisée en quatrième et en troisième qu'en sixième et en cinquième. Cette stratégie de conservatisme est une illustration de l'heuristique d'ancrage et ajustement selon Tversky & Kahneman (1973). Pour ces auteurs, les participants n'arrivent pas à se détacher de la première information. En contexte bayésien, ils se focalisent sur la probabilité *a priori* et ne révisent pas leur jugement. Cette stratégie est à opposer à la stratégie d'évidence, qui est une révision excessive de la probabilité première. La stratégie de conservatisme peut être considérée comme un défaut d'inhibition, dans le sens où, pour intégrer les nouvelles informations, les individus doivent inhiber la probabilité *a priori* qui agit comme une ancre du jugement. Nos résultats (expérimentation 5) nous suggèrent que la stratégie qui est la plus écartée entre les classes de sixième/cinquième et celles de quatrième/troisième est cette stratégie de conservatisme. L'inhibition est la fonction exécutive qui intervient dans la résolution de problèmes fréquentistes chez les élèves de quatrième et de troisième. Nous pensons que ces participants sollicitent leur capacité d'inhibition pour se décentrer par rapport à la probabilité *a priori* qui agit chez les participants de sixième et de cinquième tel un schème peu pertinent et dangereux. Cette analyse qualitative permet d'avoir une vision plus précise des performances bayésiennes dans le sens où elle nous permet d'adopter un point de vue plus large que la dichotomie réponses bayésiennes versus non bayésiennes. Ainsi, elle permet de mettre en évidence des progrès continus. En effet, conformément à nos travaux, nous pouvons dire que les individus ne progressent pas uniquement lorsqu'ils répondent à des problèmes bayésiens de façon correcte alors qu'avant ce n'était pas le cas. Deux autres cas de figure réfèrent à des progrès. Le premier est la diminution des variations intra individuelles. Dans ce premier cas, les individus vont niveler par le bas leur utilisation de stratégies. En d'autres termes, ils ne vont plus utiliser les stratégies les plus éloignées de la stratégie bayésienne qu'ils utilisaient au préalable. Le second cas est une ascension dans la hiérarchie des stratégies : ils vont commencer à utiliser de nouvelles stratégies plus élaborées et plus proches de la stratégie bayésienne. Ces deux sortes de progrès s'accordent parfaitement avec le modèle de Siegler. Les stratégies pré-bayésienne et bayésienne correspondent à un jugement révisé qui prend en compte les trois probabilités de l'énoncé en même temps. Notre troisième expérience nous a permis de dégager une algèbre cognitive (Anderson, 1979, 1981, 1996) quant à l'intégration des trois données de l'énoncé. Plus précisément, nous avons remarqué que seuls les participants de seconde et de première présentaient un profil multiplicatif. Certes ils ne multiplient pas les trois données entre elles, mais ceci est plus élaboré que le profil additif des participants des classes de sixième à troisième. Le profil additif de ces élèves signifie qu'ils prennent en compte chacune des données sans les nuancer l'une par rapport à l'autre. Plus précisément, ils savent que, d'une part, la probabilité *a priori* est importante et que, d'autre part le taux de faux positifs l'est également. Ils intègrent le fait que si, par exemple, la probabilité *a priori* est plus grande, la probabilité *a posteriori* le sera aussi. Dans le même temps ils savent que plus le taux de faux positifs augmente, plus la probabilité *a posteriori* diminue. Toutefois, au regard de leur profil additif, ils ne savent pas nuancer l'effet de l'un par rapport à celui de l'autre. Par exemple, l'effet d'un nombre élevé de faux positifs est le même pour eux que la probabilité *a priori* soit élevée ou faible, et réciproquement. Répondre de façon bayésienne à un problème nécessite avec un énoncé probabiliste de multiplier les probabilités entre elles, tandis qu'avec un énoncé fréquentiste, répondre de façon bayésienne ne requiert que de savoir additionner la donnée D et la donnée F (cf. figure 2.5.). L'adéquation entre le profil cognitif et les opérations mathématiques à utiliser -addition ou multiplication- peut être une piste explicative. Nous pensons que les participants qui présentent un profil multiplicatif seront davantage en mesure de multiplier des probabilités que des individus additifs. Puisque les participants de seconde et de première présentent un profil additif et qu'ils utilisent beaucoup la stratégie bayésienne, nous pouvons voir une relation. Cette relation concernant la stratégie bayésienne, elle est également vraie pour les performances quantitatives.

III. Les performances bayésiennes heuristiques

En accord avec la littérature, il est possible de raisonner selon deux modes de fonctionnement : le système 1 qui est intuitif et expérientiel et le système 2 qui est analytique. Les performances bayésiennes quantitatives concernent un raisonnement logique et analytique qui nécessite des connaissances théoriques. Les connaissances théoriques étant plus simples en fréquences qu'en probabilités, le format fréquentiste est facilitateur. Sobel, Tenenbaum & Gopnik (2004) suggèrent que les enfants très jeunes présentent une structure bayésienne et sont capables d'émettre des inférences bayésiennes. Nous avons tenté de rendre compte de compétences bayésiennes qui ne nécessitent ni calcul ni formule mathématique. Notre protocole en continuum (expériences 2, 3 et 5) répond à cette contrainte. Nos résultats permettent de rendre compte de l'effet facilitateur des fréquences naturelles par rapport aux probabilités conditionnelles. Cette fois-ci, n'ayant pas de calcul à faire les participants ne font que se représenter le problème. Nous supposons qu'ils sollicitent le système 1, intuitif et heuristique. Pour l'école *heuristiques et biais*, dont les représentants les plus connus sont Kahneman et Tversky, les heuristiques sont des stratégies intuitives qui permettent une simplification des problèmes rencontrés. Ces heuristiques sont pour ces auteurs souvent décrites comme des écarts à un modèle normatif, plus logique et rationnel. Leur utilisation est donc une preuve de non rationalité humaine. Nous adoptons la position plus optimiste de Gigerenzer (1996). Pour cet auteur, les heuristiques sont des outils adaptatifs que l'être humain a à sa disposition. En effet, le nombre de situations quotidiennes pour lesquelles l'homme de tous les jours calcule réellement une probabilité ou une fréquence nous semble très faible. Ceci ne l'empêche pas d'être adapté à son environnement. Nous pensons qu'il en est de même pour le raisonnement bayésien. Certes les heuristiques de représentativité, de disponibilité et d'ancrage et ajustement peuvent faire que les individus n'évaluent correctement la probabilité conditionnelle. Néanmoins, nous pensons que ces heuristiques leur permettent de s'en approcher. Nos résultats nous suggèrent que les performances heuristiques et intuitives des élèves s'améliorent avec l'avancée dans le cursus scolaire. Nous pensons, que le mode heuristique, le système 1, qui s'appuie sur un aspect expérientiel de l'individu, devient de plus en plus performant à force d'être sollicité. Nos résultats nous montrent que la distance entre la réponse heuristique des participants et la réponse théorique diminue au fil des classes. En classe de première, la distance moyenne des participants par rapport à la réponse théorique est d'environ 30 mm en probabilités conditionnelles et de seulement 15 mm en fréquences naturelles. Rapportés en pourcentages, ces distances signifient que les élèves de cette classe commettent une erreur d'environ 20 % quand le problème est présenté en probabilités conditionnelles et de seulement environ 10 % lorsque le problème est proposé dans le format fréquentiste. Nous pensons que dans la majorité des situations quotidiennes, une erreur de 10 % est acceptable. Rappelons que le seuil considéré comme significatif en recherche est souvent de 5 %. Pour une estimation heuristique, intuitive et sans calcul, les 10 % nous semblent tout à fait acceptables et témoigner d'un raisonnement bayésien adapté. Les individus présentent donc des intuitions bayésiennes sans connaître la règle de Bayes. Lorsqu'il leur est demandé d'utiliser cette formule mathématique, les individus peuvent échouer. Nous pensons qu'ils présentent donc des compétences bayésiennes qui sont en décalage par rapport à leurs performances.

En résumé, l'ensemble de ces travaux nous a permis de rendre compte d'un profil développemental des performances bayésiennes, tant quantitative, qualitatives, qu'heuristiques. Nous avons confirmé l'effet facilitateur des fréquences naturelles, et avons élargit cet effet à un mode heuristique qui peut paraître inapproprié dans un contexte de laboratoire mais semble adapté à la vie quotidienne. L'approche fréquentiste de Gigerenzer nous parait très défendable. Nous pensons qu'il est intéressant d'enseigner les probabilités à partir des fréquences, au niveau pédagogique et didactique. En ce qui concerne le raisonnement bayésien, nous pensons que les individus peuvent très jeunes s'en approcher, bien avant de connaître ou de comprendre la règle de Bayes et les probabilités. Nous pensons que parmi les individus qui ne maîtrisent ni les probabilités ni la règle de Bayes, il y a beaucoup « de bayésiens qui s'ignorent », pour reprendre l'expression de Lecoutre (2005).

IV. Perspectives de recherches

Au regard de nos résultats, il nous semble intéressant de mener une étude interculturelle du raisonnement bayésien auprès de jeunes enfants scolarisés. Lücking (2004) a testé des enfants allemands dont les performances sont moins précoces que celles des enfants chinois testés par Zhu & Gigerenzer (2006). Les élèves que nous avons rencontrés présentent un profil assez similaire à celui des enfants allemands. Le programme d'études PISA (*Programme for International Student Assessment)* révèle des différences en termes de performances mathématiques selon les pays européens. A l'heure du parachèvement de ce travail, nous sommes en contacts avec une équipe de recherche irlandaise. Nous envisageons une collaboration avec cette équipe afin de tester les performances bayésiennes de leur population de scolaires.

La réalisation de ces travaux nous a interpellé quant à un constat que nous trouvons fâcheux. En effet, il nous semble très dommageable que des professionnels de la santé qui utilisent les probabilités dans leur activité quotidienne puissent prendre des décisions qui s'appuieraient sur des heuristiques. Si nous avons déclaré que se tromper de 10 % pouvait être une erreur minime qui n'entravait en rien une adaptation dans la majorité des cas, cette même marge d'erreur nous parait énorme et lourde de conséquences dans le domaine de la santé. Nous pensons qu'il serait intéressant de mener une étude auprès de médecins, afin d'appréhender leurs critères de décisions et la façon dont ils révisent leur jugement en fonction de la survenue séquentielle de nouveaux symptômes. Les protocoles expérimentaux que nous avons décrits dans ce travail de thèse pourraient servir à décrire la façon dont les médecins intègrent l'information. En fonction de leur algèbre cognitive et de leur sensibilité aux probabilités, les médecins sont-ils bayésiens et utilisent-ils davantage le système 1, intuitif et heuristique, ou le système 2, logique et rationnel ? Nous pensons que selon son expérience, l'aspect familier des symptômes -maladie très fréquente versus maladie rare- et le contexte -période de grippe saisonnière par exemple- le médecin sollicitera usera davantage d'un mode de raisonnement ou d'un autre. Sachant que le contexte est très important pour définir des probabilités, la façon dont le patient décrit ses symptômes, et l'objectivité avec laquelle il le fait, sont également des critères importants. Si l'homme de la rue est peut être un bayésien qui s'ignore, il nous semble plus grave que certains médecins soient des non bayésiens qui s'ignorent.

Bibliographie

Anderson, N.H. (1979). Algebraic rules in psychological measurement. *American Scientist, 67,* 555-563.

Anderson, N.H. (1981). *Foundations of information integration theory.* New York: Academic Press.

Anderson, N.H. (1991). *Information Integration Theory :* Vol. I et II. *Cognition.* Hillsdale : Lawrence Erlbaum.

Anderson, N.H. (1996). *A functional theory of cognition.* Hillsdale, NJ: Erlbaum.

Antell, S.E., & Keating, D.P. (1983). Perception of numerical invariance in neonates. *Child Development, 54,* 695-701.

Aristote. (1983). *Organon III. Les premiers analytiques*. Traduction nouvelle et notes par J. Tricot. Paris : Vrin.

Arkes, H.R. (1981). Impediments to accurate clinical judgment and possible ways to minimize their impact. *Journal of Consulting & Clinical Psychology*, *49(3),* 323-330.

Artelt, C., Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Schümer, G., Stanat, P., Tillmann, K.-J., & Weiß, M. (Eds.). (2001). *PISA 2000. Zusammenfassung zentraler Befunde* [Summary of key results]. Berlin: Max Planck Institute for Human Development.

Asato, M. R., Sweeney, J. A., Luna, B. (2006). Cognitive processes in the development of TOL performance. *Neuropsychologia, 44*, 2259-2269.

Baddeley A. (1996) Exploring the central executive, *The Quaterly Journal of Experimental Psychology, 49A,* 5-28.

Baratgin, J. (2002). Le jugement probabiliste. In G. Politzer, *Le raisonnement humain.* (pp. 241-269). Hermès.

Baratgin, J., & Noveck, I.A. (2000). Not only base rates are neglected on the Engineer-Lawyer problem: An investigation of reasoners' underutilization of complementarity. *Memory & Cognition, 29(1),* 79-91.

Bar-Hillel, M. (1973). On the subjective probability of compound events. *Organizational Behavior and Human Performance, 9,* 396-406.

Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. *Acta Psychologica, 44,* 211-233.

Bar-Hillel, M. (1990). Back to base rates. In R. M. Hogarth (Ed.), *Insights in decision making* (pp. 309-330). Chicago: University of Chicago Press.

Bayes, T. (1763). An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London, 53,* 370-418.

Beckers, T., Vandorpe, S., Debeys, I., & De Houwer, J. (2009). Three-Year-Olds' Retrospective Revaluation in the Blicket Detector Task: Backward Blocking or Recovery From Overshadowing? *Experimental Psychology, 56(1),* 27-32.

Begg, I., & Harris, G. (1982). On the interpretation of syllogisms. *Journal of Verbal Learning and Verbal Behavior, 21,* 595-620.

Best, J. R., Miller, P. H., Jones, L. L. (2009). Executive functions after 5: changes and correlates. *Developmental Review, 29*, 180-200.

Beyth-Marom, R., & Fischhoff, B. (1983). Diagnosticity and Pseudodiagnosticity. *Journal of Personality and Social Psychology, 45(6),* 1185-1195.

Bideaud, J. (1997). Du bébé à l'enfant de Piaget : quelle construction du nombre ?, *Psychologie Française, 42 (1),* p.45-56.

Bideaud, J., Meljac, C., & Fisher, J.-P. (1991). *Les chemins du nombre.* Lille : Presses Universitaires de Lille.

Binet, A. (1902). *La psychologie du raisonnement.* (3e édition). Paris : Alcan.

Birnbaum, M.H. (1983). Base rates in Bayesian inference: Signal detection analysis of the cab problem. *American Journal of Psychology, 96,* 85-94.

Blanché, R. (1973). *Le raisonnement*. Paris : Presses Universitaires de France.

Brase, G.B. (2002). Ecological and evolutionary validity: Comment on Johnson-Laird, Legrenzi, Girotto, Legrenzi, and Caverni's (1999) mental model theory of extensional reasoning. *Psychological Review, 109,* 722-728.

Brase, G.L., Cosmides, L., & Tooby, J. (1998). Individuation, counting, and statistical inference: The role of frequency and whole object representations in judgment under uncertainty. *Journal of Experimental Psychology: General, 127,* 3-21.

Bresson, F. (1965). Les décisions. In Fraisse & Piaget (dir.), *Traité de psychologie*. *Langage, Communication et Décision* (Tome VIII, Chap. XXIX, pp. 221-306).

Bruner, J.S. (1974). Going beyond the information given. In J.S. Bruner (Ed.), *Beyond the information given* (pp. 218-234). Londres : Allen & Unwin. (Edition originale 1957).

Brunswik, E. (1939). Probability as a determiner of rat behavior. *Journal of Experimental Psychology, 25,* 175-197.

Budescu, D.V., & Wallsten, T.S. (1995). Processing linguistic probabilities: General principles and empirical evidence. In J.R. Busemeyer, R. Hastie, & D. Medin (Eds.), *Decision Making from the Perspective of Cognitive Psychology: The Psychology of Learning and Motivation* (pp. 275-318). New York: Academic Press.

Butterworth, B. (1999). *What counts: How every brain is hardwired for math.* New York: Free Press.

Butterworth, B. (2001). Statistics: What seems natural? *Science, 292,* 853-854.

Campos, L.M., & Romero, A.E. (2008). Bayesian network models for hierarchical test classification from a thesaurus. *International Journal of Approximate Reasoning, 50(7),* 932-944.

Camus, J.F. (1996). *La psychologie cognitive de l'attention*. Paris, Armand Colin.

Casscells, W., Schoenberger, A., & Graboys, T. (1978). Interpretation by physicians of clinical laboratory results. *New England Journal of Medicine, 299,* 999-1000.

Chapamn, L. J., & Chapman, J. P. (1959). Atmosphere effect re-examined. *Journal of Experimental Psychology, 58*(220), 220-226.

Chater, N. & Oaksford, M. (1999). The probability heuristics model of syllogistic reasoning. *Cognitive Psychology, 38,* 191-258.

Cohen, J. (1963). *Hasard, Adresse et Chance: La psychologie du pari et du jeu*. Paris: PUF. (Traduit par E. Grin).

Cohen, L.J. (1981). Can human irrationality be experimentally demonstrated? *The Behavioral and Brain Sciences, 4,* 317-331 et 359-370.

Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition, 58,* 1-73.

Crowley, K. & Siegler, R.S. (1993). Flexible strategy use in young children's tic-tac-toe. *Cognitive Science, 17(4),* 531-561.

Daston, L. (1988). *Classical probability in the Enlightenment.* Princeton, NJ: Princeton University Press.

Dawes, R.M. (1986). Representative thinking in clinical judgment. *Clinical Psychology Review, 6,* 425-441.

Dawes, R.M., Mirels, H.L., Gold, E., & Donahue, E. (1993). Equating inverse probabilities in implicit personality judgments. *Psychological Science, 4,* 396-400.

De Finetti, B. (1937). La prévision, ses lois logiques, ses sources subjectives. *Annales de l'institut Henri Poincaré, VII,* 1-68.

Demetriou, A. & Kasi, S. (2001). *Unity and modularity in the mind and the self.* London: Routledge.

Domenici, R., Toni, C., Spinetti, I., Rocchi, A., & Presciuttini, S. (2007). When Bayesian reasoning helps in directing investigations : A solved casework of a double infanticide. *Forensic Science International: Genetics Supplement Series, 1(1),* 413-414.

Dowek, G. (1995). *La logique.* Evreux : Flammarion.

Dulany, D.L., & Hilton, D.J. (1991). Conversational implicature, conscious representation and the conjunction fallacy. *Social Cognition, 9,* 85-100.

Eddy, D.M. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 249-267). Cambridge, England: Cambridge University Press.

Edwards, W. (1954). The theory of decision making. *Psychological Bulletin, 51,* 380-417.

Edwards, W., Lindman, H., & Savage, L.J. (1963). Bayesian statistical inference for psychological research. *Psychological Review, 70,* 193-242.

Epstein, S. (1983). The stability of confusion: A reply to Mischel and Peake. *Psychological Review, 90(2),* 179-184.

Epstein, S (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist, 49,* 709-724.

Epstein, S. (1998). *Cognitive-experiential self theory: A dual-process personality theory with implications for diagnosis and psychotherapy. In* R. F. Bornstein, & J. M. Masling, Empirical perspectives on the psychoanalytic unconscious (pp. 99-140). Washington, DC: American Psychological Association.

Epstein, S., Lipson, A., Holstein, C., & Huh, E. (1992). Irrational reactions to negative outcomes: evidence for two conceptual systems. *Journal of Personality and Social Psychology, 62,* 328–339.

Evans, J.St.B.T. (1989). *Bias in human reasoning: Causes and consequences.* London: Erlbaum.

Evans, J. St. B. T. & Bradshaw, H. (1986) Estimating sample-size requirements in research design: A study of intuitive statistical judgment. *Current Psychological Research and Reviews 5,* 10-19.

Evans, J.S.B.T., Handley, S.J., Perham, N., Over, D.E., & Thompson, V.A. (2000). Frequency versus probability formats in statistical word problems. *Cognition, 77,* 197-213.

Evans, J.St.B.T., & Over, D.E. (1996). *Rationality and reasoning.* Hove: Psychology Press.

Evans, J.St.B.T., & Over, D.E. (1997). Rationality in reasoning: The problem of deductive competence. *Cahiers de Psychologie Cognitive, 16*, 3-38.

Fagley, N. S. (1988) Judgmental heuristics: Implications for the decision making of school psychologists. *School Psychology Review 17,* 311-321.

Falk, R. & Konold, C. (1997). Making Sense of Randomness: Implicit Encoding as a Basis for Judgment. *Psychological Review, 104(2),* 301-318.

Feynman, R.P. (1967). *The character of physical law.* Cambridge, MA: MIT Press.

Fiedler, K., Brinkmann, B., Betsch, T., & Wild, B. (2000). A sampling approach to biases in conditional probability judgments: beyond base rate neglect and statistical format. *Journal of Experimental Psychology: General, 129,* 399-418.

Fischhoff, B., & Beyth-Marom, R. (1983). Hypothesis evaluation from a Bayesian perspective. *Psychological Review, 90,* 239-260.

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1979). Subjective sensitivity analysis. *Organizational Behavior and Human Performance, 23,* 339-359.

Fong, G.T. & Nisbett, R.E. (1991). Immediate and Delayed Transfer of Training Effects in Statistical Reasoning. *Journal of Experimental Psychology: General, 120(1),* 34-45.

Fortin, M.F. (1996). Le processus de recherche, de la conception à la réalisation. Ed Décarie.

Gallistel, C.R. (1990). *The organization of learning.* Cambridge, MA: MIT Press.

George, C. (1997). *Polymorphisme du raisonnement humain*. Paris : Presses universitaires de France.

Gigerenzer, G. (1991a). From tools to theories: A heuristic of discovery in cognitive psychology. *Psychological Rexiew, 98*, 254-267.

Gigerenzer, G. (1991b). How to make cognitive illusions disappear. Beyond “heuristics and biases.” In W. Stroebe & M. Hewstone (Eds.), *European Review of Social Psychology, 2,* 83-115.

Gigerenzer, G. (1993). The bounded rationality of probabilistic mental models. In K.I. Manktelow & D.E. Over (Eds.), *Rationality: Psychological and philosophical perspectives* (pp. 284-313). London: Routledge.

Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky. *Psychological Review, 103,* 592-596.

Gigerenzer, G. (1998). Ecological intelligence: an adaptation for frequencies. In D. Cummins & C. Allan (Eds), *The Evolution of Mind* (pp. 9-29). New York: Oxford University Press.

Gigerenzer, G. (2002). *Calculated risks: How to know when numbers deceive you.* New York: Simon & Schuster (UK version: *Reckoning with risk: Learning to live with uncertainty.* London: Penguin).

Gigerenzer, G., Hell, W., & Blank, H. (1988). Presentation and content: The use of base rates as a continuous variable. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 513-525.

Gigerenzer, G., Hertwig, R., Van den Broek, E., Fasolo, B., & Katsikopoulos, K.V. (2005). “A 30% Chance of Rain Tomorrow”: How Does the Public Understand ProbabilisticWeather Forecasts?. *Risk Analysis, 25(3),* 623-629.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review, 102,* 684-704.

Gigerenzer, G., Hoffrage, U., & Ebert, A. (1998). AIDS counseling for low-risk clients. *AIDS CARE, 10,* 197-211.

Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, *98*, 506-528.

Gigerenzer, G., & Murray, D.J. (1987). *Cognition as intuitive statistics.* Hillsdale. NJ: Erlbaum.

Gigerenzer, G., & Richter, H.R. (1990). Context effects and their interaction with development: Area judgments.*Cognitive Development, 5,* 235-264.

Gigerenzer, G., Switjink, Z., Porter, T., Daston, L., Beatty, J., & Krüger, L. (1989). *The empire of chance: How probability changed science and everyday life.* Cambridge, UK: Cambridge University Press.

Gigerenzer, G., Todd, P.M., and the ABC Research Group (1999). *Simple Heuristics that Make us Smart.* New York: Oxford University Press.

Ginossar, Z., & Trope, Y. (1987). Problem solving in judgment under uncertainty. *Journal of Personality and Social Psychology*, *52*, 464-479.

Girotto, V., & Gonzalez, M. (2001). Solving probabilistic and statistical problems: A matter of information structure and question form. *Cognition, 78,* 247-276.

Gochet, P., & Gribomont, P. (1990). *Logique : méthodes pour l'informatique fondamentale,* (vol. 1). Paris : Hermès.

Gould, S.J. (1992). *Bully for brontosaurus: Further reflections in natural history.* New York: Penguin Books.

Hamm, R.M. (1993). Explanation for common responses to the blue/green cab probabilistic inference word problem. *Psychological Reports, 72,* 219-242.

Hammond, K.R. (1955). Probabilistic functionalism and the clinical method. *Psychological Review, 62,* 255-262.

Hammond, K.R. (1996). *Human Judgment and Social Policy, Irreductible uncertainty, Inevitable Error, Unavoidable Injustice.* New York: Oxford University Press.

Hammond, K.R., McClelland, G.H., & Mumpower, J. (1980). *Human Judgment and Decision Making. Theories, Methods, and Procedures.* Praeger: Hemisphere Publishing Corporation.

Hasher, L., & Zacks, R.T. (1984). Automatic processing of fundamental information. The case of frequency of occurrence. *American Psychologist, 39,* 1327-1388.

Hasher L., Zacks R.T. (1988) Working memory, comprehension, and aging : a review and a new view, *in* G.H. Bower (Edit.), *The Psychology of Learning and Motivation, 22*, 193-225. San Diego, CA: Academic Press.

Heider, F. (1958). *The psychology of interpersonal relations.* New York: Wiley.

Hilton, D.J. (1995). The Social Context of Reasoning: Conversational Inference and Rational Judgment. *Psychology Bulletin, 118(2),* 248-271.

Hoemann, H.W. & Ross, B.M. (1971). Children's understanding of probability concepts. *Child Development, 42,* 221-236.

Hoffrage, U., & Gigerenzer, G. (1998). Using natural frequencies to improve diagnostic inferences. *Academic Medicine, 73,* 538-540.

Hoffrage, U., Gigerenzer, G., Krauss, S., & Martignon, L. (2002). Representation facilitates reasoning: What natural frequencies are and what they are not. *Cognition, 84*, 343-352.

Hogarth, R.M., & Einhorn, H.J. (1992). Order Effects in Belief Updating: The Belief-Adjustement Model. *Cognitive Psychology, 24,* 1-55.

Houdé, O. (1995). *Rationalité, développement et inhibition.* Paris : Presses Universitaires de France.

Houdé, O. (2007). First insight on « neuropedagogy of reasoning ». *Thinking and reasoning, 13,* 81-89.

Houdé, O. & Joyce, C. (1995). Développement logico-mathématique, cortex préfrontal et inhibition : l'exemple de la catégorisation. *Revue de Neuropsychologie, 5,* 281-307.

Huizinga, M., Dolan, C. V., Van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia, 44*, 2017-2036.

Hume, D. (1951). *A treatise of human nature* (L.A. Selby-Bigge, Ed.). Oxford. England: Clarendon Press. (Original work published 1739).

Jersild, A. T. (1927). Mental set and shift. Archieves of Psychology, Whole No. 89.

Johnson-Laird, P.N. (1983). *Mentals models: Towards a cognitive science of language, inference and consciousness.* Cambridge, Grande-Bretagne: Cambridge University Press.

Johnson-Laird, P.N., & Byrne, R.M.J. (1991). *Deduction.* London: Erlbaum.

Johnson-Laird, P.N., Legrenzi, P., Girotto, V., Legrenzi, M.S., & Caverni, J.-P. (1999). Naïve probability: A mental model theory of extensional reasoning. *Psychological Review, 106,* 62-88.

Jones, S.K., Taylor, K., & Frisch, D. (1995). Biases of Probability Assessment: A Comparison of Frequency and Single-Case Judgments. *Organizational Behavior and Human Decision Processes, 61(2),* 109-122.

Josephson, J.R., & Josephson, S.G. (1994). *Abductive inference.* Cambridge, Grande-Bretagne: Cambridge University Press.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist, 58,* 697-720.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment Under Uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3,* 430-450.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80,* 237-251.

Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review*, *103*, 582-591.

Klayman, J., & Ha, Y. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review, 94,* 211-228.

Kleiter, G.D. (1994). Natural sampling: Rationality without base rates. In G.H. Fischer & D. Laming (Eds.), *Contributions to mathematical psychology, psychometrics, and methodology* (pp. 375-388). New York: Springer.

Kleiter, G.D., Krebs, M., Doherty, M.E., Garavan, H., & Chadwick, B. (1997). Do Subjects Understand Base Rates. *Organizational Behavior and Human Decision Process, 72(1),* 25-61.

Koehler, J.J. (1996a). The base rate fallacy reconsidered: Descriptive, normative, and methodological challenges. *Behavioral and Brain Sciences, 19,* 1-53.

Koehler, J.J. (1996b). On conveying the probative value of DNA evidence: Frequencies, likelihood ratios, and error rates. *University of Colorado Law Review, 67,* 859-886.

Kok, A. (1999). Varieties of inhibition : manifestations in cognition, event-related potentials and aging. *Acta Psychologica, 101*, 129-158.

Kolmogorov, A. (1950). *Foundations of probability theory* (N. Morrison, Trans.). New York: Chelsea.

Kreitler, S. & Kreitler, H. (1986). Development of probability thinking in children 5 to 12 years old. *Cognitive Development, 1,* 365-390.

Kroznick, J.A., Li, F., & Lehman, D.R. (1990). Conversational conventions, order of information acquisition, and the effect of base rates and individuating information on judgments. *Journal of Personality and Social Psychology, 59,* 1140-1152.

Kurzenhäuser, S. & Hoffrage, U. (2002). Teaching Bayesian reasoning : An evaluation of a classroom tutorial for medical students. *Medical teacher, 24(5), 516-521.*

Laplace, P.S. (1774) Mémoire sur la probabilité des causes par les évènemens. *Mémoires de l'Académie royale des sciences presentés par divers savans*
6, 621-56.

Laplace, R.-S. (1951). *A philosophical essay on probabilities* (F.W. Truscott & F.L. Emory, Trans.). New York: Dover. (Original work published 1814).

Lear, J. (1980). *Aristotle and logical theory*. Cambridge: Cambridge University Press.

Lecoutre, B. (2005). Et si vous étiez un bayésien qui s'ignore ? *Modulad, 32,* 92-105.

Lewis, C., & Keren, G. (1999). On the difficulties underlying Bayesian reasoning: comment on Gigerenzer and Hoffrage. Psychological Review, 106, 411-416.

Luciana, M., & Nelson, C. A. (1998). The functional emergence of prefrontally-guided working memory systems in four- to eight-year-old children. *Neuropsychologia*, *36*(3), 273-293.

Lücking, A. (2004). *The development of Bayesian reasoning in children.* Unpublished diploma thesis, Free University, Berlin.

Lukasiewick, J. (1972). *La syllogistique d'Aristote.* Paris : Armand Colin.

Lyon, D., & Slovic, P. (1976). Dominance of accuracy information and neglect of base rates in probability estimation. *Acta Psychologica, 40,* 287-298.

Lynch, J.G. & Ofir, C. (1989). Effects of Cue Consistency and Value on Base-Rate Utilization. *Journal of Personality and Social Psychology*, *56(2),* 170-181.

Macchi, L. (1995). Pragmatic aspects of the base rate fallacy. *The Quarterly Journal of Experimental Psychology, 48A,* 188-207.

Macchi, L. (2000). Partitive formulation of information in probabilistic problems: beyond heuristics and frequency format explanations. *Organizational Behavior & Human Decision Processes, 82,* 217-236.

Macchi, L., & Girotto, V. (1994). Probabilistic reasoning with conditional probabilities: the three boxes paradox. *Paper presented at the Society for Judgement and Decision Making annual meeting*, St. Louis, MO.

Macchi, L., & Mosconi, G. (1998). Computational features vs frequentist phrasing in the base-rate fallacy. *Swiss Journal of Psychology, 57,* 79-85.

Manis, M., Dovalina, I., Avis, N.E. (1980). Base rates can affect individual predictions. *Journal of Personality and Social Psychology, 38(2),* 231-248.

Martignon, L., Vitouch, O., Takezawa, M., & Forster, M.R. (2003). Naive and yet enlightened: From natural frequencies to fast and frugal decision trees. In D. Hardman & L. Macchi (Eds.), *Thinking: Psychological perspectives on reasoning, judgment and decision making* (pp. 189-211). Chichester, UK: Wiley.

Meehl, P.E. (1954). *Clinical Versus Statistical Prediction : A Theoretical Analysis and a Review of the Evidence.* Minneapolis: University of Minnesota Press.

Michalski, R.S. (1991). Toward a unified theory of learning: An outline of basics ideas. *Conference on the fundamentals of Artificial Intelligence.* Paris, 1-5 juillet.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex “Frontal Lobe” Tasks: A Latent Variable Analysis. *Cognitive psychology, 41,* 49-100.

Monahan, J. & Steadman, H.J. (1996). Violent Storms and Violent People How Meteorology Can Inform Risk Communication in Mental Health Law. *American Psychologist, 51(9),* 931-938.

Murphy, A.H., Lichtenstein, S., Fischhoff, B., & Winkler, R.L. (1980). Misinterpretations of precipitation probability forecasts. *Bulletin of the American Meteorological Society*, *61*, 695-701.

Murphy, A.H., & Winkler, R.L. (1971). Forecasters and probability forecasts: Some current problems. *Bulletin American Meteorological Society*, *52*, 239-247.

Newstead, S.E. (1989). Interpretational errors in syllogistic reasoning. *Journal of Memory and Language, 28,* 78-91.

Nickerson, S. (1996). Ambiguities and unstated assumptions in probabilistic reasoning. *Psychological Bulletin, 120,* 410-433.

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology : Views from cognitive and personality psychology and working inhibition taxonomy. *Psychological Bulletin, 126*, 220-246.

Nisbett, R.E. (1993). *Rules For Reasoning.* Hillsdale, New Jersey : Lawrence Erlbaum.

Nisbett, R.E. & Borgida, E. (1975). Attribution and the psychology of Prediction. *Journal of Personality and Social Psychology, 32(5),* 932-943.

Norman, D. A. and Shallice, T. (1986). Attention to action: Willed and automatic control of behaviour. In Davidson, R. J., Schwartz, G. E., and Shapiro, D., editors, Consciousness and Self-Regulation: Advances in Research and Theory. Plenum Press.

Oaksford, M. & Chater, N. (2001). The probabilistic approach to human reasoning. *Cognitive Sciences, 5(8),* 349-357.

Oaksford, M. & Hahn, U. (2004). A Bayesian Approach to the Argument from Ignorance. *Canadian Journal of Experimental Psychology, 58(2),* 75-85.

Oléron, P. (1989). *Le raisonnement.* Paris : Presses Universitaires de France.

Paulos, J.A. (1988). *Innumeracy: Mathematical illiteracy and its consequences.* New York: Vintage Books.

Peirce, C.S. (1974). *Collected papers of Charles Sanders Peirce.* Cambridge, Masachusetts : The Belknap Press of Harvard University Press.

Peterson, C.R., & Beach, L.R. (1967). Man as an intuitive statistician. *Psychological Bulletin, 68,* 29-46.

Piaget, J. (1972). *Essai de logique opératoire.* Revue établie par J.-B. Grize. Paris : Dunod.

Piaget, J., & Inhelder, B. (1951). *La genèse de l'idée de hasard chez l'enfant*. Paris : Presses Universitaires de France. (Réédité en 1975).

Piaget, J., & Inhelder, B. (1966). *La Psychologie de l'enfant.* Paris: Presses Universitaires de France.

Piaget, J., & Inhelder, B. (1975). *The origin of the idea of chance in children.* New York: Norton. (Original work published 1951).

Piaget, J., & Szeminska, A. (1941). *La genèse du nombre chez l'enfant.* Neuchâtel et Paris, Delachaux et Niestlé.

Piattelli-Palmarini, M. (1995). *La Réforme du jugement ou comment ne plus se tromper.* Paris : Odile Jacob.

Ploger, D. & Wilson, M. (1991). Statistical Reasoning: What Is the Role of Inferential Rule Training? Comment on Fong and Nisbett. *Journal of Experimental Psychology: General 120(2),* 213-214.

Politzer, G. (Ed.) (2002). *Le raisonnement humain.* Paris: Hermès.

Politzer, G., & Braine, M. D. S. (1991). Responses to inconsistent premises cannot count as suppression of valid inference. *Cognition, 38,* 103-108.

Politzer, G., & Noveck, I. (1991). Are conjunction rule violations the result of conversational rule violations? *Journal of Psycholinguistic Research, 20,* 83-103.

Pollard, P., & Evans, J.St.B.T. (1983). The role of representativeness in statistical inference. In J.St.B.T. Evans (Ed.), *Thinking & reasoning* (pp. 309-330). London: Routledge & Kegan Paul.

Pólya, G. (1954). *Mathematics and plausible reasoning.* Princeton University Press. (Edition française : *Les mathématiques et le raisonnement "plausible"*. Paris : Gauthier-Villars, 1958).

Poulton, E.C. (1994). *Behavioral Decision Theory. A New Approach.* Cambridge: Cambridge University Press.

Real, L.A. (1991). Animal choice behavior and the evolution of cognitive architecture. *Science, 253,* 980-986.

Real, L.A., & Caraco, T. (1986). Risk and foraging in stochastic environments: Theory and evidence. *Annual Review of Ecological Systems, 17,* 371-390.

Reeves, L.M. & Weisberg, R.W. (1994). The Role of Content and Abstract Information in Analogical Transfer. *Psychological Bulletin, 115(3),* 381-400.

Revlis, R. (1975). Two models of syllogistic reasoning: Feature selection and conversion. *Journal of Verbal Learning and Verbal Behavior, 14*(2), 180-195.

Romine, C. B., & Reynolds, C. R. (2005). A model of the development of frontal lobe function: Findings from a meta-analysis. *Applied Neuropsychology, 12*, 190-201.

Rouanet, H. (1961). Etudes de décisions expérimentales et calcul de la probabilité. In *La Décision* (pp. 33-44). Paris: Editions du CRNS.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in Cognition. *Psychology and Aging, 103,* 403-428.

Salthouse, T. A., Davis, H. P. (2006). Organization of cognitive abilities and neuropsychological variables across the lifespan. *Developmental Review, 26*, 31-54.

Schwartz, N., Strack, F., Hilton, D., & Naderer, G. (1991). Base rates, representativeness, and the logic of conversation: The contextual relevance of "irrelevant" information. *Social Cognition, 9,* 67-84.

Sedlmeier, P. (1999). *Improving statistical reasoning: Theoretical models and practical applications.* Mahwah, NJ: Erlbaum.

Sedlmeier, P., & Gigerenzer, G. (2001). Teaching Bayesian reasoning in less than two hours. *Journal of Experimental Psychology: General, 130,* 380-400.

Sells, S. B. (1936). The atmosphere effect: An experimental study of reasoning. *Archives of Psychology*, 3-72.

Senn, T. E., Espy, K. A., & Kaufmann, P. M. (2004). Using path analysis to understand executive function organization in preschool children. *Developmental Neuropsychology, 26*, 445–464.

Shafir, E., Smith, E.E., & Osherson, D.N. (1994). Typicality and reasoning fallacies. *Theory and Decision, 36,* 103-129.

Shanks, D. (1991). A connectionist account of base-rate biases in categorization. *Connection Science, 5* (2), 143-162.

Shimamura, A. P. (2000). The role of the prefrontal cortex in dynamic filtering. *Psychobiology*, *28*, 207-218.

Siegler, R.S. (1987). The perils of averaging data over strategies: an example from children's addition. *Journal of Experimental Psychology: General, 116,* 250-264.

Siegler, R.S. (1996). *Emerging minds: the process of change in children's thinking.* New York: Oxford University Press. (Trad. Fr. L'*Emergence de l'Esprit,* Bruxelles : De Boeck Université).

Siegler, R.S. (1999). Strategic development. *Trends in Cognitive Science, 3,* 430-435.

Siegler, R.S. (2000). The rebirth of children's learning. *Child Development, 71,* 26-35.

Siegler, R.S. & Shrager, J. (1984). Strategy choices in addition and substraction: how do children know what to do ?. In C. Sophian (Ed.), *Origins of Cognitive skills* (pp. 229-293). Hillsdale, NJ: Erlbaum.

Simon, H.A. (1956). Rational choice and the structure of the environment. *Psychological Review, 63,* 129-138.

Sloman, S.A. (1996). The empirical case for two systems of reasoning. *Psychological Review, 1,* 3-32.

Slovic, P., & Lichtenstein, S. (1971). Comparison of Bayesian and Regression Approaches to the Study of Information Processing in Judgment, *Organizational Behavior and Human Performance 6*, 649-744.

Sobel, D.M., Tenenbaum, J.B., & Gopnik, A. (2004). Children's causal inferencies from indirect evidence: Bakwards blocking and Bayesian reasoning in preschoolers. *Cognitive Science, 28(3),* 303-333.

Sprenger, A. & Dougherty, M.R. (2006). Differences between probability and frequency judgments: The role of individual differences in working memory. *Organizational BNehavior and Human Decision Process, 99,* 202-211.

Spector, A., & Biederman, I. (1976). Mental set and mental shift revisited. *American Journal of Psychology, 89,* 643–679.

Stanovich, K.E., & West, R.F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences, 23,* 645-665.

Stanovich, K.E., & West, R.F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology, 94,* 672-695.

Starkey, P., & Cooper, R.G.J. (1980). Perception of numbers by human infants. *Science, 210,* 1033-1035.

Stern, E., Rode, C., Ge, F., & Zhu, L. (2001). *More than just numbers: Active diagrammatic competencies in Chinese and German secondary school students.* Poster session presented at the Biennial Meeting of the Society for Research in Child Development, Minneapolis, MI.

Störring, G. (1908). Experimentelle untersuchungen über einfache Schlussprozesse. *Archiv. für die Gesante Psychologie, 11*, 1-127. (Traduction française par R. Jamet, R., In : D. Déret, *Raisonnement catégorique : Etude génétique des jugements de compatibilité.* Thèse de Doctorat, Université de Paris-8, 1995).

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643-662.

Thorndike, E. L. (1922). The effect of changed data upon reasoning. *Journal of Experimental Psychology, 5*, 33-8.

Todd, P.M., and Gigerenzer, G. (2000). Precis of Simple Heuristic that make us Smart. *Behavioral and Brain Sciences, 23,*727-741.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 4,* 207-232.

Tversky, A., & Kahneman, D. (1974). Judgement under uncertainty: Heuristics and biases. *Science, 185*, 1124-1131.

Tversky, A., & Kahneman, D. (1980). Causal schemata in judgments under uncertainty. In M. Fishbein (Ed.), *Progress in social psychology: Vol. 1.* (pp. 49-72). Hillsdale: Erlbaum.

Tversky, A., & Kahneman, D. (1982). Judgment of and by representativeness. In D. Kahneman, P. Slovic & A. Tversky (Eds.), *Judgement under uncertainty: Heuristic and biases* (pp. 84-100). Cambridge: Cambridge University Press.

Villejoubert, G. & Mandel, D. (2002). The inverse fallacy: An account of deviations from Bayes's theorem and the additivity principle. *Memory and Cognition (30)2,* 171-178.

Von Winterfeldt, D., Edwards, W. (1986). *Decision Analysis and Behavioral Research.* Cambridge: Cambridge University Press.

Wells, G.L. & Harvey, J.H. (1978). Naïve attributors' attributions and predictions: What is informative and when is an effect? *Journal of Personality and Social Psychology 36,* 483-490.

Welsh, M. C., Satterlee-Cartmell, T., and Stine, M. (1999). Towers of Hanoi and London: contribution of working memory and inhibition performance. *Brain and Cognition, 41,* 231-242.

Wilkins, M. C. (1928). The effect of changed material on the ability to do formal syllogistic reasoning. *Archives of Psychology, 16*, 5-83.

Wolfe, C.R. (1995). Information seeking on Bayesian conditional probability problems: A fuzzy-trace theory account. *Journal of Behavioral Decision Making, 8,* 85-108.

Woodworth, R. S., & Sells, S. B. (1935). An atmosphere effect in formal syllogistic reasoning. *Journal of Experimental Psychology, 18*, 451-460.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature, 358,* 749-750.

Yahya, A.A., Mahmod, R., & Rahman Ramli, A. (2010). Dynamic Bayesian networks and variable length genetic algorithm for designing cue-based model for dialogue act recognition. *Computer Speech & Language, 24(2),* 190-218.

Zhu, L. & Gigerenzer, G. (2006). Children can solve Bayesian problems: The role of representation in mental computation. *Cognition, 98*, 287-308.

Index des Figures et des Tableaux

Annexes

Annexe 1

10 problèmes en probabilités conditionnelles

On place devant toi un grand paquet de gâteaux salés ou sucrés qui ont différentes formes. Dans le paquet, la probabilité que tu prennes un gâteau salé est de 20 %. Si le gâteau est salé, la probabilité qu'il soit rond est de 70 %. Si le gâteau n'est pas salé, la probabilité qu'il soit aussi rond est de 30 %. Imagine que tu prennes un gâteau rond, Quelle est la probabilité que ce gâteau soit salé ?

…………..%

Un groupe d'enfants joue aux cartes. Les enfants qui ont une carte avec un chat dessiné sur une face gagnent un bonbon. La probabilité qu'une carte ait un chat dessiné sur une face est de 30 %. Si une carte a un chat dessiné sur une face, la probabilité qu'elle soit rouge sur son autre face est de 40 %. Si une carte n'a pas de chat dessiné sur une face, la probabilité qu'elle soit aussi rouge sur son autre face est de 50 %. Imagine qu'un enfant pioche une carte rouge sur une face, quelle est la probabilité que cette carte ait un chat dessiné sur son autre face ?

…………..%

Dans une ville, en hiver, la probabilité qu'une personne ait les mains abîmées par le froid est de 40 %. Si une personne a les mains abîmées par le froid, la probabilité qu'elle porte des gants est de 90 %. Si une personne n'a pas les mains abîmées par le froid, la probabilité qu'elle porte aussi des gants est de 50 %. Suppose que tu rencontres dans cette ville une personne qui porte des gants, quelle est la probabilité que cette personne ait les mains abîmées par le froid ?

…………..%

Dans un hôpital, la probabilité qu'un patient ait un rhume est de 40 %. Si un patient a un rhume, la probabilité qu'il ait des maux de tête est de 30 %. Si un patient n'a pas de rhume, la probabilité qu'il ait aussi des maux de tête est de 50 %. Suppose que tu rencontres dans cet hôpital un patient qui a des maux de tête. Quelle est la probabilité que ce patient ait un rhume ?

…………..%

Dans une école, la probabilité qu'un enfant mange à la cantine est de 50 %. Si un enfant mange à la cantine, la probabilité qu'il porte des lunettes est de 70 %. Si un enfant ne mange pas à la cantine, la probabilité qu'il porte des lunettes est de 40 %. Suppose que tu rencontres dans l'école un enfant à lunettes, quelle est la probabilité que cet enfant à lunettes mange à la cantine ?

…………..%

Dans la ville de Dongdong, la probabilité qu'un enfant soit obèse est de 10 %. Si un enfant est obèse, la probabilité qu'il ait une mère obèse est de 30 %. Si un enfant a un poids normal, la probabilité qu'il ait aussi une mère obèse est de 20 %. Suppose que tu rencontres une mère obèse dans cette ville, quelle est la probabilité que cette mère ait un enfant obèse ?

…………..%

Les mères recommandent toujours à leurs enfants de ne pas trop regarder la télévision pour ne pas s'abîmer les yeux. Supposons que tu veuilles tester cette croyance et que tu possèdes les informations suivantes : la probabilité qu'un enfant soit myope est de 30 %. Si un enfant est myope, la probabilité qu'il regarde trop la télévision est de 70 %. Si un enfant n'est pas myope, la probabilité qu'il regarde aussi trop la télévision est de 40 %. Suppose que tu rencontres un enfant qui regarde trop la télévision, quelle est la probabilité que cet enfant soit myope ?

…………..%

Dans une école, la probabilité qu'un enfant ait les dents abîmées est de 20 %. Si un enfant a les dents abîmées, la probabilité qu'il aime manger sucré est de 50 %. Si un enfant n'a pas les dents abîmées, la probabilité qu'il aime aussi manger sucré est de 30 %. On te présente un enfant de cette école qui aime manger sucré, quelle est la probabilité que cet enfant ait les dents abîmées ?

…………..%

Le directeur d'une école explique le nouveau règlement aux élèves rassemblés dans la cour. Le directeur dit : « ceux qui ont compris ce que j'ai dit, levez la main s'il vous plaît ». La probabilité qu'un élève ait compris est de 70 %. Si un élève a compris, la probabilité qu'il lève la main est de 90 %. Si un élève n'a pas compris la probabilité qu'il lève quand même la

main est de 30 %. Imagine qu'un élève lève la main, quelle est la probabilité que cet élève ait compris le directeur ?
…………..%

Pingping va dans un petit village pour demander son chemin. Dans ce village, la probabilité qu'il rencontre un menteur est de 10 %. Si une personne ment, la probabilité qu'elle ait un nez rouge est de 80 %. Si une personne ne ment pas, la probabilité qu'elle ait aussi un nez rouge est de 10 %. Imagine que Pingping rencontre dans ce village une personne qui a un nez rouge. Quelle est la probabilité que cette personne mente ?
…………..%

<u>10 problèmes en fréquences naturelles</u>

On place devant toi un grand paquet de gâteaux salés ou sucrés qui ont différentes formes. Dans le paquet, sur 100 gâteaux, 20 sont salés. Parmi ces 20 gâteaux salés, 14 sont ronds. Parmi les 80 autres gâteaux sucrés, 24 sont aussi ronds. Imagine que tu prennes une pile de gâteaux ronds, combien d'entre eux sont des gâteaux salés ?
………. sur ……….

Un groupe d'enfants joue aux cartes. Les enfants qui ont une carte avec un chat dessiné sur une face gagnent un bonbon. Sur 100 cartes, 30 ont un chat dessiné sur une face. Parmi ces 30 cartes, 12 sont rouges sur leur autre face. Sur les 70 cartes restantes qui n'ont pas de chat dessiné sur une face, 35 sont quand même rouges sur leur autre face. Imagine qu'un enfant pioche un groupe de cartes rouges sur une face, combien d'entre elles ont un chat dessiné sur leur autre face ?
………. sur ……….

Dans une ville, en hiver, sur 100 personnes 40 ont les mains abîmées par le froid. Parmi ces 40 personnes qui ont les mains abîmées, 36 portent des gants. Sur les 60 personnes qui n'ont pas les mains abîmées, 30 portent aussi des gants. Suppose que tu rencontres dans cette ville un groupe de personnes qui portent des gants, combien d'entre elles ont les mains abîmées par le froid ?
………. sur ……….

Dans un hôpital, sur 100 patients 40 ont un rhume. Parmi ces 40 patients qui ont un rhume, 12 ont des maux de tête. Sur les 60 patients restant qui n'ont pas de rhume, 30 ont aussi des maux de tête. Suppose que tu rencontres dans cet hôpital un groupe de patients qui ont des maux de tête. Combien d'entres eux ont un rhume ?

………. sur ……….

Dans une école, sur 100 enfants 50 mangent à la cantine. Parmi ces 50 enfants, 35 ont des lunettes. Sur les 50 autres enfants qui ne mangent pas à la cantine, 20 portent des lunettes. Suppose que tu rencontres dans cette école un groupe d'enfants à lunettes, combien d'entre eux mangent à la cantine ?

………. sur ……….

Dans la ville de Dongdong, sur 100 enfants 10 sont obèses. Parmi ces 10 enfants obèses, 3 ont des mères obèses. Sur les 90 enfants restant qui ont un poids normal, 18 ont quand même des mères obèses. Suppose que tu rencontres un groupe de mères obèses dans cette ville, combien d'entres elles ont des enfants obèses ?

………. sur ……….

Les mères recommandent toujours à leurs enfants de ne pas trop regarder la télévision pour ne pas s'abîmer les yeux. Supposons que tu veuilles tester cette croyance et que tu possèdes les informations suivantes : sur 100 enfants 30 sont myopes. Parmi ces 30 enfants myopes, 21 regardent trop la télévision. Parmi les 70 autres enfants qui voient bien, 28 regardent aussi trop la télévision. Suppose que tu rencontres un groupe d'enfants qui regardent trop la télévision, combien d'entre eux sont myopes ?

………. sur ……….

Dans une école, sur 100 enfants 20 ont les dents abîmées. Parmi ces 20 enfants qui ont les dents abîmées, 10 aiment manger sucré. Parmi les 80 enfants qui n'ont pas les dents abîmées, 24 aiment aussi manger sucré. On te présente un groupe d'enfants qui aime manger sucré, combien d'entre eux ont les dents abîmées ?

………. sur ……….

Le directeur d'une école explique le nouveau règlement aux élèves rassemblés dans la cour. Le directeur dit : « ceux qui ont compris ce que j'ai dit, levez la main s'il vous plaît ». Sur 100 élèves, 70 ont compris. Parmi ces 70 élèves qui ont compris, 63 lèvent la main. Parmi les 30 élèves restant qui n'ont pas compris, 9 lèvent la main. Imagine qu'un groupe d'élèves lève la main, combien d'entre eux ont compris le directeur ?

………. sur ……….

Pingping va dans un petit village pour demander son chemin. Dans ce village, sur 100 personnes 10 vont mentir. Parmi ces 10 personnes qui mentent, 8 ont un nez rouge. Sur les 90 personnes restantes qui ne mentent pas, 9 ont aussi un nez rouge. Imagine que Pingping rencontre dans ce village un groupe de personnes qui ont des nez rouges, combien d'entres eux vont mentir ?

………. sur ……….

Annexe 2

COMPARAISONS XO

"Voici une série de paires de lettres, vous devez les comparer et cocher dans l'une des 2 cases Identique ou différent le plus rapidement possible durant 30 secondes."

		IDENTIQUE	DIFFERENT			IDENTIQUE	DIFFERENT
X	O	☐	☐	O	X	☐	☐
X	X	☐	☐	X	X	☐	☐
O	X	☐	☐	O	O	☐	☐
O	O	☐	☐	O	X	☐	☐
X	O	☐	☐	X	O	☐	☐
O	X	☐	☐	X	X	☐	☐
X	X	☐	☐	O	X	☐	☐
O	O	☐	☐	O	O	☐	☐
O	X	☐	☐	X	O	☐	☐
X	O	☐	☐	X	X	☐	☐
X	X	☐	☐	X	O	☐	☐
X	O	☐	☐	O	O	☐	☐
O	O	☐	☐	X	X	☐	☐
O	O	☐	☐	X	O	☐	☐
X	O	☐	☐	X	O	☐	☐
O	O	☐	☐	X	X	☐	☐
X	X	☐	☐	O	X	☐	☐
O	X	☐	☐	O	O	☐	☐
X	X	☐	☐	X	O	☐	☐
O	O	☐	☐	X	X	☐	☐
O	X	☐	☐	X	O	☐	☐
X	O	☐	☐	X	O	☐	☐
X	O	☐	☐	X	X	☐	☐
X	X	☐	☐	O	X	☐	☐
O	X	☐	☐	O	X	☐	☐
O	O	☐	☐	X	X	☐	☐
X	O	☐	☐	O	O	☐	☐

Annexe 3

L E   S T R O O P

Consigne : « Je vais vous présenter trois planches. Sur la première vous devrez lire les noms de couleur écrits en noir ; sur la seconde les couleurs des croix. Sur la troisième, il vous faudra nommer les couleurs de l'encre et non lire les mots. Pour les trois planches vous devez aller le plus vite possible. »

ROUGE	BLEU	VERT	ROUGE	BLEU
VERT	VERT	ROUGE	BLEU	VERT
BLEU	ROUGE	BLEU	VERT	ROUGE
VERT	BLEU	ROUGE	ROUGE	BLEU
ROUGE	ROUGE	VERT	BLEU	VERT
BLEU	VERT	BLEU	VERT	ROUGE
ROUGE	BLEU	VERT	BLEU	VERT
BLEU	VERT	ROUGE	VERT	ROUGE
VERT	ROUGE	BLEU	ROUGE	BLEU
BLEU	VERT	VERT	BLEU	VERT
VERT	ROUGE	BLEU	ROUGE	ROUGE
ROUGE	BLEU	ROUGE	VERT	BLEU
VERT	ROUGE	BLEU	ROUGE	VERT
BLEU	BLEU	ROUGE	VERT	ROUGE
ROUGE	VERT	VERT	BLEU	BLEU
BLEU	BLEU	ROUGE	VERT	ROUGE
ROUGE	VERT	BLEU	ROUGE	VERT
VERT	ROUGE	VERT	BLEU	BLEU
ROUGE	BLEU	ROUGE	VERT	ROUGE
VERT	ROUGE	VERT	BLEU	VERT

Planche A

XXXX	XXXX	XXXX	XXXX	XXXX
XXXX	XXXX	XXXX	XXXX	XXXX
XXXX	XXXX	XXXX	XXXX	XXXX
XXXX	XXXX	XXXX	XXXX	XXXX
XXXX	XXXX	XXXX	XXXX	XXXX
XXXX	XXXX	XXXX	XXXX	XXXX
XXXX	XXXX	XXXX	XXXX	XXXX
XXXX	XXXX	XXXX	XXXX	XXXX
XXXX	XXXX	XXXX	XXXX	XXXX
XXXX	XXXX	XXXX	XXXX	XXXX
XXXX	XXXX	XXXX	XXXX	XXXX
XXXX	XXXX	XXXX	XXXX	XXXX
XXXX	XXXX	XXXX	XXXX	XXXX
XXXX	XXXX	XXXX	XXXX	XXXX
XXXX	XXXX	XXXX	XXXX	XXXX
XXXX	XXXX	XXXX	XXXX	XXXX
XXXX	XXXX	XXXX	XXXX	XXXX
XXXX	XXXX	XXXX	XXXX	XXXX
XXXX	XXXX	XXXX	XXXX	XXXX
XXXX	XXXX	XXXX	XXXX	XXXX

Planche B

ROUGE	BLEU	VERT	ROUGE	BLEU
VERT	VERT	ROUGE	BLEU	VERT
BLEU	ROUGE	BLEU	VERT	ROUGE
VERT	BLEU	ROUGE	ROUGE	BLEU
ROUGE	ROUGE	VERT	BLEU	VERT
BLEU	VERT	BLEU	VERT	ROUGE
ROUGE	BLEU	VERT	BLEU	VERT
BLEU	VERT	ROUGE	VERT	ROUGE
VERT	ROUGE	BLEU	ROUGE	BLEU
BLEU	VERT	VERT	BLEU	VERT
VERT	ROUGE	BLEU	ROUGE	ROUGE
ROUGE	BLEU	ROUGE	VERT	BLEU
VERT	ROUGE	BLEU	ROUGE	VERT
BLEU	BLEU	ROUGE	VERT	ROUGE
ROUGE	VERT	VERT	BLEU	BLEU
BLEU	BLEU	ROUGE	VERT	ROUGE
ROUGE	VERT	BLEU	ROUGE	VERT
VERT	ROUGE	VERT	BLEU	BLEU
ROUGE	BLEU	ROUGE	VERT	ROUGE
VERT	ROUGE	VERT	BLEU	VERT

Planche C

Annexe 4

N - B A C K   T E S T

Consigne: « Je vais vous dire plusieurs lettres et il faudra me dire si la dernière lettre dite est la même que l'avant dernière. Puis, à chaque nouvelle lettre, il faudra me dire par « oui » ou par « non » si elle est la même que l'avant dernière.

Afin de mieux comprendre la consigne nous allons donner un exemple :

Si je vous dit « Z », « R », puis « K », est-ce que « K » correspond à l'avant dernière lettre ? (NON). Si je vous dis ensuite « R » (OUI). Si je vous dit « G » (NON) » (voir schéma ci-dessous).

Exemple :

Z – R – K	Z – R – K – R
NON	OUI

Séquence de lettres :

F	L	F	L	S	L	R	L	S	R	S	B	P
B	D	G	D	G	L	M	F	M	N	M	N	L
N	L	M	L									

Annexe 5

L E   P L U S   /   M I N U S

Voici des listes de nombres et vous devez :

- Pour la liste A, ajouter le chiffre 3 (+3) à chacun des nombres
- Pour la liste B, soustraire le chiffre 3 (-3) à chacun des nombres
- Pour la liste C, **alternativement** : ajouter 3 pour le premier nombre, soustraire 3 pour le second, ajouter 3 pour le troisième,… et ainsi de suite pour tous les nombres de la liste.

Vous devez compléter chaque liste le plus rapidement et le plus justement possible, et sans revenir en arrière.

Inscrivez vos réponses sur les pointillés à côté des nombres présentés.

Essayez tout d'abord de compléter l'exemple :

+ 3		**- 3**		**+ 3, - 3, …**	
18	…..	52	…..	27	…..
67	…..	94	…..	16	…..
31	…..	37	…..	40	…..
56	…..	65	…..	82	…..
87	…..	77	…..	64	…..
72	…..	33	…..	31	…..

LISTE A

49	…..	13	…..
13	…..	61	…..
75	…..	92	…..
38	…..	40	…..
21	…..	88	…..
80	…..	54	…..
45	…..	35	…..
32	…..	63	…..
59	…..	26	…..
77	…..	39	…..
85	…..	47	…..
94	…..	73	…..
16	…..	58	…..
67	…..	95	…..
24	…..	17	…..
81	…..	42	…..

LISTE B

97	…..	35	…..
19	…..	73	…..
53	…..	20	…..
77	…..	91	…..
45	…..	52	…..
10	…..	74	…..
24	…..	31	…..
36	…..	66	…..
62	…..	43	…..
79	…..	28	…..
55	…..	11	…..
23	…..	89	…..
87	…..	15	…..
41	…..	37	…..
96	…..	64	…..
17	…..	55	…..

LISTE C

52	…..	19	…..
81	…..	28	…..
27	…..	94	…..
30	…..	61	…..
65	…..	70	…..
93	…..	39	…..
49	…..	86	…..
16	…..	43	…..
58	…..	95	…..
25	…..	78	…..
82	…..	12	…..
76	…..	84	…..
31	…..	45	…..
97	…..	51	…..
53	…..	26	…..
44	…..	67	…..

Olivier SOREL

APPROCHE DEVELOPPEMENTALE DU RAISONNEMENT BAYESIEN

-Analyse quantitative et qualitative selon le format de présentation, et le niveau scolaire-

Résumé

Emettre des inférences bayésiennes, c'est-à-dire réviser son jugement quant à l'apparition d'un événement, et passer d'une probabilité *a priori* à une probabilité *a posteriori* est une activité quotidienne qui, en théorie requiert la formule de Bayes. Toutefois, en pratique, l'être humain est sensible à différents biais et n'utilise pas toujours à bon escient les informations dont il dispose. Zhu & Gigerenzer (2006) ont montré qu'un contexte en fréquences naturelles favorisait davantage une révision de jugement qu'un contexte présenté en probabilités conditionnelles, les fréquences rendant explicite le taux de base (Hoffrage, Gigerenzer, Krauss & Martignon, 2002). Le but de ce travail est de préciser dans quelle mesure le format de présentation et les exigences quant au jugement attendu influencent les performances bayésiennes de collégiens et lycéens français. Des cohortes de 20 participants de classe de sixième à première ont été soumises à des problèmes bayésiens. Les trois premières expériences font de façon progressive abstraction du nombre. Les résultats confirment l'effet facilitateur des fréquences sur les probabilités. L'analyse des stratégies utilisées par les participants suggère qu'il ne faut pas se contenter d'une cotation dichotomique : réponse bayésienne versus non bayésienne. En effet, nos résultats précisent qu'avec l'avancée scolaire les participants commettent des erreurs quantitatives de moins en moins éloignées de la réponse théorique, et estiment qualitativement de plus en plus finement l'occurrence du dit événement. La quatrième expérience tente de faire le lien entre les fonctions exécutives de bas niveau, la vitesse de traitement et le niveau scolaire en mathématiques, d'une part, et les performances bayésiennes, d'autre part. Les résultats montrent que la vitesse de traitement et l'inhibition prédisent modérément les performances bayésiennes des collégiens, mais pas celles des lycéens. La dernière expérience est une analyse longitudinale des performances des participants testés à dix-neuf mois d'intervalle. Les résultats étayent ceux de l'analyse transversale, puisque la majorité des participants présente des performances accrues.

Mots clés : inférences bayésiennes, fréquences naturelles, probabilités conditionnelles, raisonnement probabiliste.

Résumé en anglais

Emit bayesian inferences, in other words revise his judgment about the appearance of an event, and change from prior probability to posterior probability, is a daily activity which, in theory, requires the Bayes' rule. However, in practice, human is sensible to different bias and he doesn't use systematically wisely all information at his disposal. Zhu & Gigerenzer (2006) showed that when data are presented in natural frequencies, bayesian performances increase. The reason is that base rate information is contained in natural frequencies (Hoffrage, Gigerenzer, Krauss, & Martignon, 2002). The aim of this work is to specify in what measure the format of presentation and the experimenter's request about the judgment to product influence the bayesian performances of French schoolchild from the beginning to the end of secondary school. Groups of 20 participants from sixth grader to eleventh grade were tested on bayesian problems. The first three experiments progressively disregard number. Results confirm the easier effect of frequencies on the probabilities. The analyse of strategies used by participants indicates that we don't have to be content with dichotomous quotation: bayesian versus no bayesian answer. In fact, our results specify that with the school advance participants make quantitative mistakes which are less and less distant of the theoretic answer, and estimate qualitatively more and more precisely the event occurrence. The fourth experience tries to make the link between the executive functions of low level, the processing speed and the school level in mathematical, on the one hand, and the bayesian performances, on the other hand. Results show that processing speed and inhibition predict moderately bayesian performances during the first middle of secondary school, but don't during the second. The last experience is a longitudinal analyse of participants' performances tested nineteen months later. Results support the transversal analyse ones, since the majority of participants produce better performances.

Key words: bayesian inferences, natural frequencies, conditional probabilities, probabilistic reasoning.