

Université de Montréal

Calibration, réalisme et détermination d'étudiants et d'étudiantes
universitaires en fonction de l'expérience, de la distance temporelle,
du genre et du niveau de réussite.

Par

Catherine Cloutier

Département de psychologie

Faculté des études supérieures

Thèse présentée à la Faculté des études supérieures en vue de l'obtention
du grade de Philosophiae Doctor (Ph.D.) en psychologie

Juin, 2004

© Catherine Cloutier, 2004



BF

22

U54

2005

V. 004

Direction des bibliothèques

AVIS

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

NOTICE

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal
Faculté des études supérieures

Cette thèse intitulée:
Calibration, réalisme et détermination d'étudiants et d'étudiantes
universitaires en fonction de l'expérience, de la distance temporelle,
du genre et du niveau de réussite.

Présentée par:
Catherine Cloutier

a été évalué par un jury composé des personnes suivantes :

Judith Comeau, Présidente

Luc Lamarche, Directeur de recherche

Michel Pagé, Correcteur interne

Line Tremblay, Correctrice externe

Thèse acceptée le: 10 décembre 2004

RÉSUMÉ

Le niveau important d'attrition dans les différents programmes académiques nous a amenés à nous interroger sur le réalisme des étudiant-e-s inscrit-e-s au baccalauréat spécialisé de psychologie à l'Université de Montréal. Nous avons évalué leur réalisme à l'égard de divers événements de leur vie académique afin de vérifier l'hypothèse générale d'une augmentation de ce réalisme en fonction de leur expérience académique, de leur genre et de leur réussite. Nous avons eu recours, entre autres, à différents indices mettant en relation leurs prédictions et postdictions selon les approches transversale et longitudinale. Nous avons aussi fait des hypothèses à l'égard du réalisme des étudiant-e-s et de la distance temporelle séparant leurs prédictions de l'événement. Nos résultats indiquent un manque de réalisme important de la part des étudiant-e-s qui s'exprime surtout par de la surconfiance. Même si le manque de réalisme diminue en général avec l'expérience, il demeure à un niveau élevé. La diminution est plus forte au début. Pour ce qui est des examens, nous n'avons pas trouvé de lien avec le genre, mais le niveau de réussite semble une variable importante. Nous terminons en faisant des suggestions pour les recherches futures sur cette question.

Mots clés : Calibration - réalisme - détermination - étudiant-e universitaire - expérience - distance temporelle - genre - réussite.

SUMMARY

Low student retention rates in academic programs led us to inquire about the realism of male and female students enrolled in the bachelor's degree program in psychology at the University of Montreal. We assessed students' realism regarding their academic courses in order to verify the general hypothesis of an increase in realism in light of their actual academic experience, their gender and their academic performance. To this end, we examined different measures aimed at assessing the relationship between pre and post-event predictions through cross-section and longitudinal data analysis. We also made hypotheses regarding male and female students' realism in function of the temporal distance separating prediction from the event predicted. Our results indicate an important lack of realism in both male and female students, in the direction of overconfidence. Even though experience increases realism, a lack of realism persists. The decrease in realism is more pronounced in the beginning. Regarding predictions on test performance, we did not find a relationship to gender, but the level of achievement appears to be an important variable. We conclude by making suggestions for future research.

Key words: Calibration - realism - persistence - male/female university students - experience - temporal distance - gender - level of achievement.

TABLE DES MATIÈRES

IDENTIFICATION DU JURY.....	II
RÉSUMÉ.....	III
SUMMARY.....	IV
TABLE DES MATIÈRES.....	V
LISTE DES TABLEAUX.....	IX
LISTE DES FIGURES.....	X
REMERCIEMENTS.....	XI
INTRODUCTION.....	1
CHAPITRE I : CONTEXTE THÉORIQUE.....	4
Réalisme en général.....	6
Introduction.....	6
Types de jugements.....	8
Mesures du réalisme.....	9
Recherches.....	15
Réalisme des étudiants.....	25
Prédictions en fonction de l'expérience.....	25
À la suite d'un feedback donné par l'étude de la matière.....	27
À la suite du feedback donné par un examen (questions) de pratique.....	32
À la suite du feedback donné par l'examen.....	37
À la suite du feedback donné par les notes.....	48
Prédictions en fonction de la distance temporelle.....	56
HYPOTHÈSES ET QUESTIONS.....	64

CHAPITRE II: DESCRIPTION DE LA RECHERCHE.....	67
SUJETS.....	68
QUESTIONNAIRES.....	68
DÉROULEMENT.....	69
ANALYSE DES RÉSULTATS.....	70
Calibration.....	70
Calibration absolue.....	70
Calibration relative.....	72
Calibration de la certitude.....	72
Niveau de réalisme.....	72
Évolution de la confiance des prédictions et de la détermination.....	73
Niveau de confiance selon la distance temporelle.....	73
CHAPITRE III : RÉSULTATS.....	74
CALIBRATION.....	75
Calibration absolue.....	75
Calibration selon l'ordre des examens et le genre.....	75
Calibration et perception de l'examen.....	78
Calibration et note obtenue.....	79
Calibration de la certitude.....	78
Calibration de la certitude selon l'ordre des examens et le genre.....	78
Calibration de la certitude et perception de l'examen.....	81
Calibration de la certitude et note obtenue.....	81
Calibration relative.....	81

RÉALISME.....	85
Réalisme selon l'ordre des examens et le genre.....	86
Réalisme et perception de l'examen.....	88
Réalisme et note obtenue.....	88
Réalisme académique selon l'année d'inscription.....	88
Réalisation selon le niveau de confiance.....	91
Évolution du niveau de confiance des prédictions.....	94
Évolution du niveau de détermination.....	99
Niveau de confiance selon la distance temporelle.....	102
CHAPITRE IV : INTERPRÉTATION.....	107
CHAPITRE V : CONCLUSION.....	119
RÉFÉRENCES.....	124
APPENDICE A : QUESTIONNAIRES.....	XII
APPENDICE B : FIGURE 6.....	XXVI

LISTE DES TABLEAUX

Tableau 1 Moyennes et écart-types de la calibration selon l'examen et le genre.....	75
Tableau 2 Calibration, anova selon l'ordre des examens et le genre.....	77
Tableau 3 Moyennes et écart-types de la calibration de la certitude selon l'examen et le genre.....	79
Tableau 4 Calibration de la certitude, anova selon l'examen et le genre.....	80
Tableau 5 Proportion de réponses exactes en fonction des niveaux de confiance pour l'examen intra et l'examen final.....	83
Tableau 6 Moyennes et écart-types du réalisme selon l'examen et le genre.....	86
Tableau 7 Réalisme, anova selon l'examen et le genre.....	87
Tableau 8 Réalisme académique selon l'année de présentation et l'année d'inscription	89
Tableau 9 Réalisation selon le niveau de confiance.....	92
Tableau 10 Prédications académiques et professionnelles, évolution du niveau de confiance.....	95
Tableau 11 Évolution du niveau de détermination.....	100

LISTE DES FIGURES

Figure 1. Représentation graphique de la calibration.....	14
Figure 2. Représentation graphique de la calibration pour l'examen intra et l'examen final.....	82
Figure 3. Fréquences d'utilisation des niveaux de confiance pour l'examen intra et l'examen final.....	85
Figure 4. Niveau de confiance à trois moments en fonction de la distance temporelle.....	103
Figure 5. Niveau de confiance en fonction de la distance temporelle. Moyenne des trois mesures.....	105
Figure 6. Niveau de confiance en fonction de la distance temporelle à la suite du changement de la séquence.....	XXVII

REMERCIEMENTS

J'aimerais en tout premier lieu remercier mon directeur de thèse, le professeur Luc Lamarche, pour son support constant vers l'achèvement de cette thèse. Malgré les obstacles inévitables vers l'accomplissement de cette thèse, il a offert une grande disponibilité ainsi que démontré un dévouement hors du commun.

J'aimerais aussi remercier Miguel Chagnon, statisticien, qui m'a conseillée et aidée à faire les analyses statistiques. Son aide s'est révélée précieuse.

J'aimerais aussi remercier tous les étudiants et étudiantes qui ont bien voulu participer à cette étude.

J'aimerais aussi remercier les membres de ma famille, pour m'avoir insufflé ce goût de me dépasser et m'avoir transmis la détermination de persister dans l'effort vers l'atteinte d'un but. J'aimerais enfin remercier mon amie Patricia Clark qui m'a offert des conseils et des encouragements continus tout au long de l'accomplissement de cette thèse.

INTRODUCTION

Chaque année, une quantité imposante d'étudiant-e-s s'inscrivent dans divers programmes universitaires alors que leurs chances de réussite sont loin d'être acquises. Jusqu'à quel point sont-ils (elles) conscients de ce fait? Plus spécifiquement, sont-ils (elles) réalistes lorsqu'ils (elles) évaluent leur probabilité de réussir à un examen, à un trimestre, au baccalauréat ou lorsqu'ils (elles) se demandent s'ils (elles) seront accepté-e-s aux cycles supérieurs et pourront éventuellement exercer la profession qu'ils (elles) visent? Est-ce qu'ils (elles) deviennent plus réalistes avec le temps? Quelles sont les autres variables qui peuvent expliquer le niveau de leur réalisme?

Les enquêtes ICOPE (Pageau et Bujold, 2000) rapportent que la très grande majorité des étudiant-e-s s'inscrivent à un programme avec l'intention d'en obtenir le diplôme, que le choix de l'établissement est définitif, que 90% des étudiant-e-s à plein temps comptent cheminer dans le programme sans interruption, que près de 60% ne manifeste pas d'intérêt pour un autre programme d'étude, que plus de 90% s'inscrivent à leur programme dans le but d'accéder à une profession et que 75% aspirent atteindre un niveau plus élevé que le baccalauréat. Dans ce dernier cas, toutefois, lorsque les chercheurs soulignent les contraintes éventuelles qu'ils (elles) pourraient rencontrer (travail, temps, argent, famille, etc.), cette proportion baisse à 30%. Les mêmes enquêtes ont trouvé que parmi les caractéristiques du profil des étudiant-e-s qui ont les meilleures chances de réussite, il y a celle d'avoir une bonne connaissance du programme dans lequel on est inscrit. Ces données confortent l'hypothèse de l'importance du réalisme des étudiant-e-s pour la réussite de leur programme.

La présente recherche a été effectuée auprès d'étudiant-e-s inscrit-e-s au département de psychologie de l'Université de Montréal. Des données du Bureau de la recherche institutionnelle de cette université (2004) révèlent que, pour les cohortes de 1996, 1997 et 1998, l'attrition au programme initial après un an est déjà de 17.9% et l'attrition globale quatre ans après l'entrée

dans la cohorte (étudiant-e-s qui ont quitté l'Université de Montréal sans avoir complété leurs études dans le programme initial ou dans tout autre programme) est de 18.9%.

En 1994, il y a eu 145 diplômés au B.Sc. Cinquante et un d'entre eux ont répondu au sondage de Chouinard et Hamel (1995) révélant que 63% (32/51) se sont inscrits à la maîtrise à l'Université de Montréal. Le diplôme de maîtrise est d'ailleurs une condition d'admission à l'Ordre des psychologues du Québec. La poursuite des études demeure donc la voie privilégiée de la majorité des répondants de cette discipline, se répartissant comme suit: 33.3% (17/51) entreprennent une M.Ps., 18% (9/51) un Ph.D. et 7.8% (4/51) une M.Sc., le reste des étudiants (41.2%, 21/51) poursuivent leurs études dans un autre programme que la psychologie, alors que 26.8% (19/51) sont principalement sur le marché du travail. Pour la même année, il y a eu 88 diplômés d'une maîtrise. Le reste des répondants qui choisissent le marché du travail (18%), sont soit au chômage (3/39), soit inactifs par choix (2/39), mais presque tous ont un travail relié à leur formation (34/39) On note cependant que la plupart des emplois sont temporaires (50.4%) et plusieurs sont à temps partiel.

Notre recherche porte donc sur la relation, mesurée de diverses façons, entre les prédictions et les réalisations des étudiant-e-s de même que sur l'évolution de cette relation par rapport à différents aspects de leur vie académique. Elle a pour but, entre autres, d'aller au-delà des nombres rapportés ci-dessus afin d'examiner de plus près ce qui fonde la confiance des étudiant-e-s.

CHAPITRE I : CONTEXTE THÉORIQUE

« Mon ami avait vingt ans; j'en avais dix-huit : nous étions donc tous deux à cet âge où il est permis de confondre les rêves avec la réalité » (Lamartine, *Graziella, Épisode III*).

« Avoir normalement confiance en soi, c'est essentiel pour atteindre ses objectifs, qu'il s'agisse d'un diplôme universitaire, de monter une affaire qui marche, de bâtir une relation satisfaisante ou d'exercer l'esprit au bonheur. Le manque de confiance en soi inhibe nos efforts pour aller de l'avant, pour relever les défis, pour prendre certains risques nécessaires à la poursuite de nos objectifs. Inversement, l'excès de confiance sera tout aussi périlleux. Ceux qui s'exagèrent leurs propres mérites et leurs succès s'exposent en permanence à la frustration, à la déception et à la colère aussitôt que la réalité fait intrusion et que le monde extérieur ne valide pas cette image idéalisée qu'ils ont d'eux-mêmes » (Sa sainteté le Dalai-Lama et H. Cutler).

RÉALISME EN GÉNÉRAL

Introduction

Cette thèse porte sur le réalisme des croyances telles qu'exprimées dans des opinions qui sont le résultat d'un jugement. Comme les termes que nous employons dans la phrase qui précède sont sujets à différentes interprétations et qu'ils reviendront constamment tout au long de la thèse, il nous paraît essentiel d'en arrêter le sens que nous leur prêterons. À cet effet, nous allons recourir au dictionnaire *Le grand Robert de la langue française* (Rey et Robert, 2001).

Pour *jugement*, nous retenons la définition suivante : « faculté de l'esprit permettant de juger (plus ou moins bien) des choses qui ne font pas l'objet d'une connaissance immédiate certaine, ni d'une démonstration rigoureuse; l'exercice même de cette faculté ». D'où l'expression *être trompé par son jugement* et la citation retenue par *Le grand Robert* (Rey et Robert, 2001) « Tout le monde se plaint de sa mémoire, et personne ne se plaint de son jugement » (tiré de La Rochefoucauld, *Réflexions morales*).

Pour ce qui est de *croyance*, le même dictionnaire en donne comme premier sens « L'action, le fait de croire une chose vraie, vraisemblable ou possible ». La définition qu'il donne de *l'opinion* comme premier sens est « Manière de penser, de juger sur un sujet qu'il s'agisse d'une attitude de l'esprit qui tient pour vraie une assertion ou d'une assertion, d'un système d'assertions que l'esprit accepte ou rejette (généralement en admettant une possibilité d'erreur) ». Comme on le voit, ces deux dernières définitions sont très proches l'une de l'autre. Dans la suite de ce texte, la distinction que nous établirons entre ces concepts sera que l'opinion est l'expression verbale de la croyance bien qu'une telle distinction ne soit pas clairement affirmée dans *Le grand Robert*.

En ce qui concerne le *réalisme*, *Le grand Robert* (Rey et Robert, 2001), donne comme sens courant « Qui a le sens des réalités; qui témoigne du

sens des réalités » et, pour *réalité*, « Caractère de ce qui existe en fait (et qui n'est pas seulement une invention, une illusion, une apparence) ». D'où la possibilité de douter de la réalité d'un fait et celle d'avoir une vision réaliste. *Le grand Robert* ajoute comme définition courante de la réalité « La vie, l'existence réelle (en tant que distincte des désirs, des illusions, du rêve) ». Comme dans « Prendre ses désirs pour la réalité ».

Pour être plus complet par rapport à ce dont nous allons parler, il peut être utile d'enfiler les définitions suivantes tirées du *Cambridge Dictionary* : « *Real*. Existing in fact; not imaginary. *Realism*. A way of thinking and acting based on the facts of a situation and what appears to be possible, rather than on hopes for things which are unlikely to happen. *Realist*. Someone who hopes for or accepts only what seems possible or likely and does not have too high hopes or expectations. *Realistic*. Accepting things as they are in fact and not basing decisions on unlikely hopes for the future. *Reality*. State of things as they are, rather than as they are imagined to be ».

À défaut d'une connaissance immédiate certaine ou d'une démonstration rigoureuse, il faut, pour établir la probabilité d'une chose, s'en remettre à son jugement. Ce dernier, en effet, est inutile quand la connaissance d'une chose ne fait pas de doute, il est, au contraire, essentiel quand il y a incertitude, doute, et que l'on possède des informations ou des opinions qui peuvent être utiles pour générer une approximation de ce qui a été, de ce qui est ou de ce qui sera.

Comment, plus précisément, s'y prend-on alors? Il semble que l'utilisation de règles heuristiques permettant de réduire l'incertitude sans garantir une solution exacte et unique, contrairement aux algorithmes, soient un des moyens privilégiés (Kahneman, 1982). De nombreuses règles ont ainsi été identifiées. Elles peuvent parfois être à l'origine d'erreurs d'estimation importantes, comme lorsqu'elles conduisent à négliger certaines informations parce que l'on accorde trop de poids à d'autres informations plus

« parlantes ». Quoi qu'il en soit, le résultat final du jugement consistera à affirmer ou à nier avec plus ou moins de conviction l'existence d'un objet ou le rapport d'un objet à un attribut.

Se pose alors la question de la validité du jugement. Qu'entend-on par avoir un bon jugement? Peut-on mesurer la qualité du jugement d'une personne? La qualité du jugement varie-t-elle selon les situations, les circonstances, les objets auxquels il est appliqué? Nous aborderons ces questions dans la suite de ce texte, une fois que nous aurons distingué les types de jugements.

Types de jugements

Si l'on définit le résultat du jugement comme étant une croyance et une croyance comme étant la représentation d'un objet lié à un attribut, la qualité du jugement devient alors le degré d'adéquation de la force du lien à la réalité. La force du lien étant le niveau de confiance que la personne éprouve quant à la relation entre l'objet et l'attribut.

Ce niveau de confiance peut s'exprimer dans le discours courant par l'utilisation d'adverbes modalisateurs (Dubois et al., 1973). On peut aussi demander à la personne d'exprimer son niveau de confiance à l'aide d'un continuum de probabilités psychologiques celui-ci pouvant s'étendre, par exemple, de 0 à 1 (certitude d'absence d'une relation à certitude de présence d'une relation) en passant par .5 (incertitude maximale). Dans le cas d'un choix dichotomique, il n'est pas nécessaire d'avoir recours à une échelle de confiance de « pleine étendue » (0 à 1), une échelle de « demi-étendue » (.5 à 1) s'imposant alors.

L'objet d'une croyance peut être toute entité qu'il est possible de situer par rapport à une dimension du jugement alors que l'attribut est cette dimension du jugement. L'objet de la croyance peut être soi, comme par exemple, dans les jugements d'auto-efficacité. L'individu affirmant avec une plus ou moins grande conviction qu'il est capable d'obtenir tel ou tel résultat.

Dans le cas qui nous occupe dans cette thèse, l'objet est souvent une attente de résultat soit avant qu'une tâche soit accomplie (*prédiction*) soit après (*postdiction*). En fait, la croyance peut, bien sûr, se prononcer sur un objet autant par rapport à son état passé, présent ou à venir.

Étant donné ces considérations, comment peut-on évaluer le réalisme de quelqu'un? Cela suppose que le jugement porte sur des faits connus ou éventuellement connaissables. Autrement dit, qu'il y a un critère objectif permettant de dire si le jugement est vrai ou faux. Donc, s'interroger sur le réalisme d'une question dont on ne connaît pas la réponse, ni ne la connaîtra dans un avenir permettant l'estimation au moment voulu, n'a pas de sens. Il en est de même des jugements de valeurs. Ceci dit, il se peut que la personne croie qu'une opinion est vérifiable alors qu'elle ne l'est pas. La personne se trompe alors doublement : sur ce qu'elle pense savoir et sur ce qui est vérifiable (Lamarche, 1988a, 1994; Lamarche, Askévis-Leherpeux, et Déziel, 1996).

Dans les cas où une personne affirme être certaine d'une opinion, il suffit, pour voir si elle a raison, de confronter son opinion à la réalité. C'est plus compliqué quand le niveau de confiance est plus faible parce qu'une personne, par exemple, peut afficher un niveau de confiance de .8 sur un continuum de probabilité psychologique, se tromper quant à la réponse, tout en ayant raison quant à son niveau de confiance. Comment procéder alors pour déterminer le réalisme de la confiance? C'est ce que nous verrons dans ce qui suit.

Mesure du réalisme

De nombreuses formules ont été proposées pour établir le réalisme de la confiance. Certaines sont plus sophistiquées que d'autres, certaines sont plus souvent utilisées que d'autres, mais, toujours, il s'agit de mettre en relation, d'une façon ou d'une autre, un niveau de confiance et un degré d'exactitude. Notons que de nombreuses recherches se contentent d'établir

le degré d'exactitude comme quand on demande, par exemple, à un étudiant d'estimer en pourcentage la note qu'il obtiendra à un examen. À moins que l'étudiant en question soit tout à fait certain de son estimation, on ne peut se prononcer sur le réalisme de sa confiance. Cela est vrai même quand la confiance de réussir est mise en rapport avec la note en pourcentage, soit par une opération de soustraction ou de division. Cette remarque est tout aussi vraie quand le chercheur s'intéresse au comportement d'un agrégat d'individus et qu'il utilise la moyenne ou, encore, un indice de corrélation comme le r de Pearson. Dans l'étude du réalisme des étudiants, le Gamma de Goodman et Kruskal (γ) est souvent préféré comme indice d'association parce qu'il peut s'appliquer à des variables de niveau ordinal. Il varie lui aussi de 1 à -1, 0 signifiant que les deux variables sont indépendantes.

Souvent les chercheurs ont mis en relation un niveau de confiance et un résultat binaire (échec ou réussite, exact ou inexact), ils ont alors recours à la corrélation bisérielle de point qui est une mesure d'association entre une variable continue et une variable binaire. Soit X une variable continue (le niveau de confiance) et Y une variable nominale (échec ou réussite) qui peut prendre les valeurs 0 ou 1, la formule de la corrélation bisérielle de point est alors la suivante :

$$r_{pb} = \frac{(\bar{X}_1 - \bar{X}_0)\sqrt{p(1-p)}}{S_x}$$

où \bar{X}_0 est la moyenne de X quand $Y=0$, \bar{X}_1 est la moyenne de X quand $Y=1$, S_x est l'écart-type de X et p est la proportion de valeurs où $Y=1$

L'interprétation de la corrélation bisérielle de point est similaire à celle de la corrélation r de Pearson. La corrélation bisérielle de point est positive quand des valeurs élevées de X sont associées à $Y=1$ et que des valeurs faibles de X sont associées à $Y=0$.

D'autres indices ont été utilisés pour établir le réalisme de la confiance qui s'inspirent de l'indice d' de la théorie de la détection de signal. C'est le cas du

Confidence accuracy coefficient (CAQ) de (Shaughnessy, 1979) :

$$CAQ = \frac{\bar{X}_{cfe} - \bar{X}_{cfin}}{\sqrt{\sigma_{cf}}}$$

où \bar{X}_{cfe} est la moyenne du niveau de confiance pour les réponses exactes, \bar{X}_{cfin} est la moyenne du niveau de confiance pour les réponses inexactes et $\sqrt{\sigma_{cf}}$ est la racine carrée de l'écart-type de tous les jugements de confiance pour bonnes et mauvaises réponses. $CAQ = 0$, sauf si la personne peut distinguer entre les items où elle a bien répondu et ceux où elle a mal répondu.

L'indice le plus commun pour établir le réalisme de la confiance est le score de probabilité moyenne (PS) de Brier (1950) :

$$PS = \frac{1}{N} \sum_{i=1}^n (x_i - d_i)^2 \quad PS = \frac{1}{N} \sum_{i=1}^n (x_i - d_i)^2$$

Où N est le nombre d'évaluation de probabilités, x_i la probabilité psychologique attribuée à l'occasion i , et d_i est un indice de résultat qui prend la valeur 1 si l'événement se produit et 0 sinon. Le score est parfait (i.e. 0) pour une personne qui donne la probabilité 1 à tous les événements qui se produisent, et 0 à tous les événements qui ne se produisent pas.

Murphy (1973) a découpé ce score en trois composantes qui peuvent expliquer que la personne ne satisfait pas le critère de perfection :

$$PS = c(1-c) + \frac{1}{N} \sum_{i=1}^T n_i (x_i - c_i)^2 - \frac{1}{N} \sum_{i=1}^T n_i (c_i - c)^2$$

La première composante, $c(1 - c)$, exprime la connaissance de la personne, c étant la moyenne de l'indice de résultat d_i . Ce score varie de 0 (connaissance parfaite) à 0.25 (la personne devine pour tous les items).

La seconde composante mesure la calibration (i.e. réalisme de la confiance). Ici, N est le nombre de questions, T est le nombre de catégories de confiance utilisées, x_t est la réponse de confiance dans la catégorie t , c_t est la proportion de réponses exactes dans la catégorie t , et n_t est le nombre de fois où la catégorie de réponse t a été utilisée. Une personne est bien calibrée si, pour toutes les réponses où a été donnée une probabilité .XX, une proportion de XX% de celles-ci est exacte (i.e. si $x_t = c_t$). Cette composante est celle qui a reçu le plus d'attention dans les études psychologiques, probablement parce qu'elle est une expression directe de l'étendue de l'accord entre la probabilité psychologique et le résultat. Elle ne permet pas de distinguer entre les déviations qui sont dues à la surconfiance ou à la sous-confiance. C'est pourquoi cette mesure est souvent accompagnée d'une mesure de sur-/sous-confiance obtenue par la différence entre la moyenne générale de la probabilité psychologique et la moyenne générale de réussite, $x - c$.

La troisième composante mesure l'aptitude de la personne à discriminer les occasions où un événement cible se produit ($d_i = 1$) des occasions où il ne se produit pas ($d_i = 0$). Une aptitude qui s'appelle la résolution ou la discrimination. Une bonne discrimination se produit lorsque les catégories de la personne produisent des taux de base conditionnels qui s'éloignent du taux de base général par rapport à l'occurrence de l'événement.

Pour comprendre la distinction entre la calibration et la discrimination, Liberman et Tversky (1993) proposent l'exemple fictif suivant : Soit un médecin estimant le sexe d'un bébé avec la probabilité de .5. Si le médecin en question fait de façon répétée ce jugement, il sera parfaitement calibré. Cependant, les jugements du médecin ne discriminent pas entre les

situations où le bébé est un garçon et celles où c'est une fille. Inversement, le médecin en question pourrait faire des jugements qui aident à discriminer entre deux résultats possibles, même s'il est très surconfiant (ou sous-confiant). Voir figure 1.

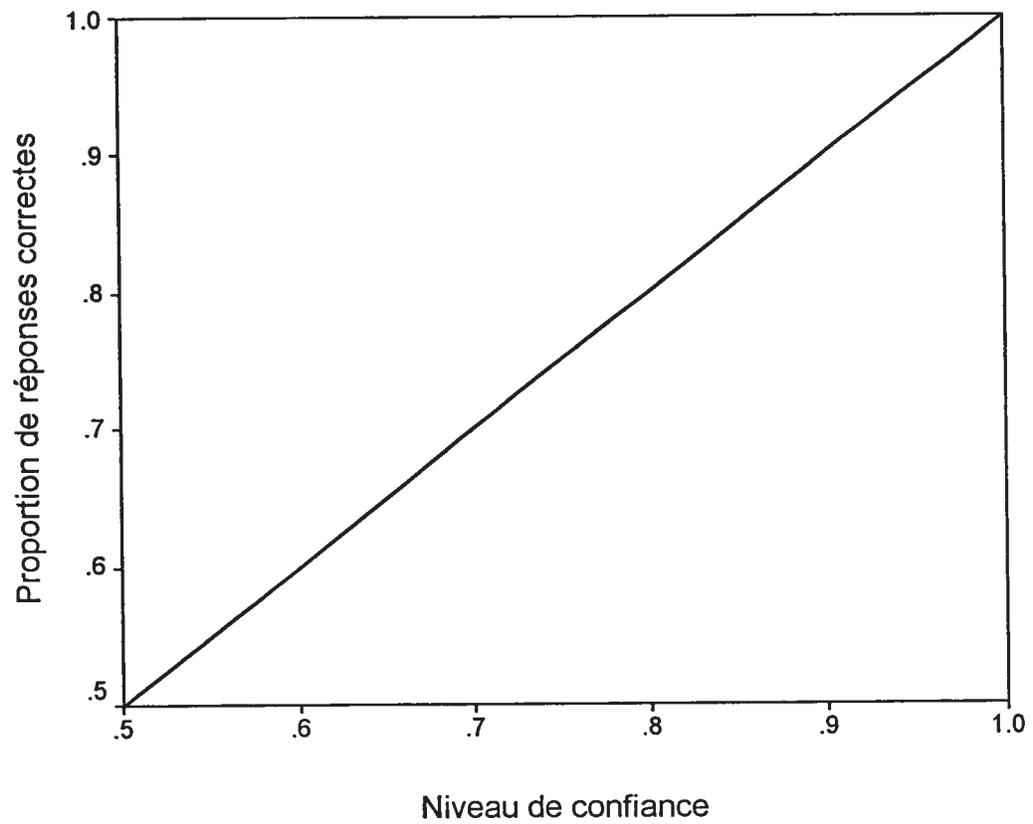


FIGURE 1. Représentation graphique de la calibration

Cette courbe de calibration est obtenue de la façon suivante :

- (a) On demande plusieurs évaluations de probabilité pour des items dont la réponse correcte est connue ou sera connue sous peu;
- (b) Pour chacune des catégories de probabilité, on calcule la proportion d'items pour lesquels la proposition est exacte;
- (c) Pour chaque catégorie (sur l'abscisse) on rapporte la proportion de réponses correctes (sur l'ordonnée).

Il y a calibration parfaite lorsque tous les points se retrouvent sur la ligne d'identité, sinon il y a surconfiance (sous la ligne d'identité) ou sous-confiance (au-dessus de la ligne d'identité).

Dans la partie qui suit, nous rapporterons les principales recherches portant sur le réalisme de la confiance en nous attardant plus particulièrement à celles concernant les étudiants.

Recherches

Tout jugement étant affecté d'un réalisme de la confiance plus ou moins grand, il n'est pas étonnant de constater que, les moyens de l'évaluer se raffinant, l'intérêt pour cette question soit allé en grandissant ces dernières années et que les chercheurs s'appliquent à le mesurer dans des situations très diverses.

La plus grande part des études s'est effectuée en laboratoire, mais, de plus en plus, en partie en réponse à des critiques portant sur la validité écologique de ces recherches, les chercheurs ont développé des moyens pour obtenir des réponses à leurs questions en se déplaçant dans des milieux plus naturels.

Il n'en demeure pas moins que la tâche la plus souvent utilisée pour établir le degré de calibration consiste à répondre à des questions de connaissances générales (*almanach questions*). Plus spécifiquement, on demande à une personne de choisir entre deux réponses proposées à la suite d'une question, comme dans l'exemple suivant : « Quel pays a la population la plus élevée? (a) France, (b) Indonésie? » La personne choisit une réponse et estime son niveau de confiance par rapport à cette réponse sur une échelle de probabilité en six points allant de .5 (choix au hasard) à 1 (certitude). Pour établir le réalisme de la confiance d'une personne, il faut poser de nombreuses questions semblables (plutôt 200 que 50).

De nombreux chercheurs se sont aussi penchés sur le réalisme de la confiance des jugements des professionnels dans leurs professions respectives (Fischhoff et MacGregor, 1981; Lichtenstein, Fischhoff et Phillips, 1982). Les premières recherches ont porté sur le réalisme de la confiance des météorologues et elles ont été effectuées dès 1906 par W.

Ernest Cooke. Par la suite, plusieurs autres (v.g. Murphy et Winkler, 1977; Williams, 1951), ont donné lieu à des observations permettant, entre autres choses, de mieux comprendre les conditions d'un bon réalisme de la confiance. Les météorologues posant des jugements répétitifs en termes de probabilités numériques par rapport à des événements bien définis, il n'est peut-être pas étonnant que les questions au sujet de la qualité de ces jugements aient été les premières à avoir été exprimées aussi clairement.

Pour des raisons semblables, mais aussi à cause des conséquences désastreuses potentielles d'une erreur, les diagnostics des médecins sont apparus à certains comme devant faire l'objet d'un examen attentif (v.g. Christensen-Szalanski, Beck, Christensen-Slzenski et Koepsell, 1983; Dawson, 2000; Detmer, Fryback et Gassner, 1978; Einhorn, 1972; Elstein, Shulman et Sprafka, 1978; Lusted, 1977). À ces deux professions, on pourrait ajouter les exemples de nombreuses autres dont les agents d'immeuble (Northcraft et Neale, 1987), les psychologues (Oskamp, 1965) et les négociateurs (Ansel et Touzard, 2002) pour n'en nommer que quelques-unes.

D'autres situations faisant appel à des conduites ayant les caractéristiques appropriées pour une étude de leur réalisme de la confiance les désignaient tout naturellement à l'attention des chercheurs. C'est le cas du pari, où le risque pris dépend d'une évaluation explicite des probabilités de la part du parieur, la surconfiance menant inexorablement à la faillite (v.g. Johnson et Bruce, 2001; Keren, 1987; Lichtenstein et Slovic, 1971; Wagenaar et Keren, 1985).

C'est aussi le cas du témoignage oculaire où, cette fois, ce qui est à craindre est l'erreur judiciaire (v.g. Busey, Tunnicliff, Loftus et Loftus, 2000; Olsson, 2000; Read, Lindsay et Nicholls, 1998). Enfin, le réalisme de la confiance des étudiants, l'objet de cette thèse, a aussi donné lieu à de nombreuses recherches avec des objectifs très variés (v.g. Beyer, 1990,

1998, 1998/1999, 1999; Beyer et Bowden, 1997; Gilovich, Kerr et Medvec, 1993; Hacker, Bol, Horgan et Rakow, 2000; Hackett et Betz, 1989; Hoch, 1985; Lundeberg et al., 1994; Orndorff, 1996; Prohaska, 1994; Schraw et Roedel, 1994; Shaughnessy, 1979; Sjostrom et Marks, 1994; Stone, 1994; Vallone, Griffin, Lin et Ross, 1990; Zakay et Glicksohn, 1992; Zeleznik et al., 1988). Ici, la surconfiance peut donner lieu à un gaspillage de temps, d'argent et d'effort, mais la sous-confiance est aussi à craindre dans la mesure où l'étudiant(e) n'exploite pas suffisamment son talent pour atteindre des objectifs qu'il valorise par ailleurs. Dans ce qui suit, nous verrons les principaux résultats obtenus dans les recherches portant sur le réalisme de la confiance pour mettre en perspective particulièrement ceux qui se rapportent spécifiquement aux étudiants.

Le principal résultat des études en laboratoire est que la plupart des individus font preuve d'un réalisme de la confiance déficient (Johnson et Bruce, 2001; Lichtenstein, Fischhoff et Phillips, 1982). L'écart le plus souvent observé entre les probabilités psychologiques et les probabilités objectives indique de la surconfiance. Il y a toutefois des différences individuelles systématiques (Alloy et Abramson, 1988; Dunning et Story, 1991; Klayman, 1999; Lichtenstein et al., 1982; Taylor et Brown, 1988).

Chez le même individu, le réalisme de la confiance peut aussi varier selon la tâche (Klayman, Soll, Gonzalez-Vallejo et Barlas, 1999; Lichtenstein, Fischhoff et Phillips, 1982). De sorte qu'il n'est pas possible de prédire avec certitude le réalisme de la confiance d'un individu d'une tâche à l'autre. Les questions dans certains domaines sont plus difficiles qu'elles n'apparaissent et, dans d'autres, plus faciles. Par exemple, Hoch (1985) trouve que les auto-prédictions par rapport à des événements personnels sont plus problématiques que la réponse à des questions de connaissance générale parce que susceptibles de biais à la fois cognitifs et motivationnels. Toutefois, Wright et Ayton (1988) observent que le réalisme de la confiance est très bon chez les parieurs. Il s'agit là, d'ailleurs, d'une situation qui a

toute la richesse, le dynamisme et l'incertitude d'un contexte naturel : problèmes mal définis; environnement incertain, dynamique; buts changeants, mal définis, compétitifs; boucle action-feedback; pression du temps; enjeux élevés.

La relation la plus étudiée est celle entre le réalisme de la confiance et la difficulté des questions : il y a surconfiance quand les questions sont difficiles et sous-confiance quand elles sont faciles (Baranski et Petrusic, 1995; Lichtenstein et Fischhoff, 1977; Suantak, 1996). Notons cependant que la qualité du réalisme de la confiance ne devrait pas être affectée par les questions très faciles comme, par exemple, celles portant sur l'identification de la personne : son nom, son âge, etc. Il en est de même des questions très difficiles où la personne ne se reconnaît pas compétente d'emblée pour avancer la moindre possibilité de répondre. En deçà de ces deux extrêmes du continuum *facile-difficile*, a été observée une adéquation à la réalité déficiente qui a donné lieu à des interprétations divergentes, mais qui, pour certaines, sont parfois présentées comme étant complémentaires.

Différentes raisons ont été avancées pour expliquer l'effet difficile-facile dans les jugements de probabilité psychologique, les principales sont : les biais dans le jugement, un manque d'ajustement adéquat des critères de décision et le choix biaisé des stimuli par les chercheurs.

Le biais cognitif est une façon inappropriée de traiter et d'évaluer l'information. L'utilisation d'un certain nombre d'heuristiques qui biaisent la recherche en mémoire ou la formation de croyances, produirait une insensibilité à la difficulté de la tâche soit par un manque d'attention par rapport à la qualité ou au poids des données (Griffin et Tversky, 1992); soit par un traitement insuffisant de l'information (Sniezek, Paese et Switzer, 1990); soit par manque d'attention par rapport aux données qui infirment la réponse privilégiée (Koriat, 1980). Dans ce dernier cas, la personne fouillerait d'abord sa mémoire pour trouver des informations pertinentes et

arriver à une première possibilité de réponse. Gardant cette réponse à l'esprit, la personne, rechercherait ensuite d'autres données. Les mécanismes de la mémoire associative favorisent le retrait de croyances consistantes avec la première impression et cette impression colore aussi l'interprétation des autres informations retrouvées de sorte que la personne perçoit une confirmation pour sa première réponse plus élevée que ce qui devrait être.

Dans plusieurs situations, des facteurs motivationnels peuvent exacerber ce biais. Les gens aiment penser qu'ils sont intelligents, qu'ils possèdent beaucoup de connaissances. Ils peuvent aussi avoir des raisons de vouloir qu'une réponse soit vraie plutôt qu'une autre (Klayman, Soll, Gonzalez-Vallejo et Barlas, 1999).

Le modèle de la partition de la variable de décision postule pour sa part, qu'une variable interne non échelonnée, le sentiment de confiance, est découpée de sorte qu'un ensemble de réponses numériques croissant de façon monotone peut lui être attaché. Le manque d'information par rapport à la difficulté de la tâche et ce qui est la meilleure répartition empêche la personne de changer suffisamment la partition quand la difficulté de la tâche augmente (Ferrell et McGoey, 1980).

La troisième explication retient que les individus intériorisent à partir de leur expérience la validité de certains indices permettant de produire des jugements dans des environnements particuliers. Dans les situations expérimentales, les individus recevraient des tâches artificielles et trompeuses qui impliquent des questions non représentatives de sorte que les indices habituels ne sont plus valides (Björkman, 1994; Gigerenzer, 1991; Juslin, 1993, 1994). Un changement mineur dans l'expérimentation peut faire la différence entre de bonnes ou de mauvaises réponses (Aytton et Wright, 1994; Beach, Christensen-Slzenski et Barnes, 1987). Cela revient à dire que les heuristiques que les gens utilisent peuvent produire de mauvais

jugements dans les tâches expérimentales conçues par les chercheurs, mais s'avérer tout à fait fonctionnelles dans l'environnement naturel. L'effet *facile-difficile* serait donc dû à la sélection des stimuli. À cet égard, le modèle écologique le plus explicite, la théorie des modèles mentaux probabilistes de Gigerenzer (1991), repose sur l'idée de base que les gens apprennent les contingences dans leur environnement naturel. Plus précisément, ils apprennent la validité des indices qu'ils utilisent pour parvenir à des jugements, c'est-à-dire les fréquences relatives de l'efficacité des indices et les valeurs numériques de ces validités.

Quand on leur présente des questions, les gens prennent en considération les indices disponibles pour choisir la réponse et lorsqu'un indice recouvert en mémoire est suffisant pour permettre un choix, la personne choisit en fonction de l'indice et rapporte la validité apprise de l'indice en tant que probabilité psychologique comme quoi la réponse est correcte. Si, en moyenne, les indices sont aussi efficaces dans l'ensemble de questions présentées qu'ils l'ont été dans l'expérience de la personne, alors les probabilités psychologiques de la personne s'avèreront comment étant bien calibrées.

Un exemple fréquemment cité de question de connaissance générale permet de mieux comprendre les prétentions du modèle écologique: « Quelle ville est la plus au nord, New York ou Rome? Choisissez une réponse et indiquez votre niveau de confiance. ». Un indice qu'une personne peut utiliser alors est la relation entre la latitude et le climat dans l'hémisphère Nord. Si la personne pense qu'il fait plus chaud en moyenne à Rome qu'à New York, elle répondra, faussement, New York. Puis elle donnera comme probabilité psychologique d'être correcte, la proportion approximative de fois où elle a trouvé cet indice comme conduisant à la bonne réponse.

La théorie des modèles mentaux probabilistes soutient donc que, dans plusieurs circonstances, les personnes ont une calibration raisonnable par

rapport à la structure de probabilité de leur environnement naturel. Il y aurait un ajustement cognitif, c'est-à-dire que les personnes auraient une expérience suffisante de leur environnement pour acquérir des représentations cognitives adéquates des probabilités écologiques.

Il reste que, malgré les critiques adressées aux premières recherches effectuées en laboratoire, même en contrôlant rigoureusement pour éviter les failles des recherches antérieures dénoncées par les modèles écologiques, l'exactitude des probabilités psychologiques demeure encore souvent déficiente (Brenner, Koehler, Liberman et Tversky, 1996; Griffin et Tversky, 1992; Keren, 1997).

Par exemple, dans une recherche, Dunning, Griffin, Milojkovic et Ross (1990) ont étudié la relation entre la confiance subjective et l'exactitude des prédictions sociales objectives dans des domaines qui sont très familiers à la personne. Il s'agissait, en effet, de prévoir le comportement d'un pair en se basant sur une base d'information riche qui est fournie à la personne. Les prédictions étaient comparées avec les réponses données par la cible. Les informations étaient obtenues soit par entrevue, soit par le fait de vivre depuis un certain temps avec la personne (compagnon de chambre). La mesure la plus simple que les chercheurs utilisent est la différence entre la confiance moyenne et l'exactitude moyenne des prédictions. Ils observent qu'en général les personnes sont très surconfiantes. Et cette surconfiance est d'autant plus grande que la confiance est grande. Selon les chercheurs, les raisons de la surconfiance sont multiples et s'additionnent les unes aux autres. Ils constatent, en particulier, que les personnes ne connaissent pas assez bien les contraintes imposées par les situations sur le comportement et qu'ils ne tiennent pas assez compte dans leurs inférences de l'incertitude reliée à ces contraintes.

Un autre facteur qui favorise l'inadéquation entre les probabilités subjectives et les probabilités objectives est la relative insensibilité des

personnes à la ligne de base (*base rate*). Il y a surestimation des probabilités avec les tâches qui ont des lignes de base faibles, et sous-estimation qui ont des lignes de base élevées, c'est ce que l'on appelle l'effet de la ligne de base (Ferrell, 1994; Lichtenstein et al., 1982). Même quand les personnes reçoivent de l'information explicite quant à la ligne de base, cela n'a pas d'effet sur leurs prédictions. Ces personnes sont aussi susceptibles que d'autres qui n'ont pas reçu cette information de croire que la cible vivra des événements rares et ce, avec une confiance aussi grande.

Par ailleurs, si la surconfiance est l'observation la plus courante quand on étudie le réalisme de la confiance, il semble que celle-ci s'accroît avec l'augmentation d'indices pour répondre à des questions. C'est ce qu'Oskamp (1965) a observé auprès de psychologues cliniciens auxquels il demandait de porter un diagnostic accompagné d'un niveau de confiance au fur et à mesure où il leur donnait des indices nouveaux à propos d'un cas réel. Le même phénomène de surconfiance croissante a été obtenu par Lamarche (1988b) avec des questions de connaissances générales.

Des recherches ont aussi montré que le degré de réalisme de la confiance variait systématiquement en fonction de la façon de poser les questions (Klayman, Soll, Gonzalez-Vallejo et Barlas, 1999; Lichtenstein, Fischhoff et Phillips, 1982). La surconfiance est moins grande avec les questions qui demandent un choix entre deux réponses (v.g. « Quel pays a l'espérance de vie la plus élevée : (A) Argentine, (B) Canada? »), que lorsque la personne doit donner un estimé numérique unique à une question comme la suivante : « Combien y a-t-il de calories dans une demie tasse de pouding chômeur? » On peut alors demander aux personnes d'estimer des fractiles (par exemple, donner une valeur telle qu'il y a 25% de chance que la réponse exacte soit plus élevée) ou de donner une étendue correspondant à un niveau de confiance donné (par exemple, fournir un estimé élevé et un estimé bas tel qu'il y ait 90% de chance que la bonne réponse tombe entre ces deux estimés.)

Des études montrent que, dans le cas du choix entre deux réponses, lorsque les personnes disent être sûres à un niveau de 70%, elles ont la bonne réponse moins de 60% du temps. Quand elles disent être sûres à 90%, elles atteignent un niveau d'exactitude autour de 75%. Dans le cas où les gens doivent donner une étendue, la surconfiance serait encore plus importante (Klayman, Soll, Gonzalez-Vallejo et Barlas, 1999).

La recherche montre qu'imaginer un événement futur (Anderson, 1983; Carroll, 1978; Gregory, Cialdini et Carpenter, 1982) ou expliquer pourquoi il peut se produire (Ross, Lepper, Strack et Steinmetz, 1977) peut le faire apparaître plus probable. Il n'en va pas autrement des attentes d'un individu par rapport à sa performance qui peuvent affecter son degré de réalisme (Stone, 1994). Ainsi, Trafimow et Sniezek (1994) amènent des individus à croire qu'ils auront une bonne ou une mauvaise performance en répondant à des questions de connaissances générales. Ceux qui croyaient qu'ils auraient une bonne performance n'ont pas mieux fait, mais étaient surconfiants. Les autres n'étaient ni surconfiants ni sous-confiants. Plus généralement, imaginer un avenir plein d'obstacles qui dépasse les capacités d'une personne peut l'amener à éviter la situation ou à mal performer. S'arrêter à ses forces personnelles et aux aspects abordables d'une tâche peut faire qu'une personne excellera (Cervone, 1989). Quand des individus jugent leur niveau d'auto-efficacité, ils doivent considérer l'interaction entre deux ensembles d'événements : difficultés et demandes de la situation, d'une part, et leurs propres aptitudes, d'autre part. Il peut y avoir beaucoup de facteurs de la situation et personnels à considérer et beaucoup d'incertitude associée. Cela est encore plus complexe quand la personne doit évaluer ses attentes de performance par rapport à celle de ses pairs (Campbell et Fairey, 1985; Sherman, Skov, Hervitz et Stock, 1981). Cela demande à la personne de non seulement évaluer sa propre performance future, mais aussi celle des autres.

De nombreuses recherches montrent qu'il y a un lien entre la désirabilité d'un événement et la surconfiance qu'il se produise. Ainsi, il y a surestimation de la probabilité d'événements favorables à soi et sous-estimation des événements défavorables à soi (Zakay, 1983). De plus, les gens ont tendance à exprimer un optimisme irréaliste à propos de leur possibilité d'obtenir des résultats désirables (v.g. un mariage heureux) et d'éviter des événements indésirables (v.g. un cancer des poumons) (Weinstein, 1980).

Les probabilités psychologiques ne sont donc pas indépendantes de l'utilité des résultats (Irwin, 1953; Marks, 1951). Les personnes évaluent les événements ayant un résultat positif comme étant plus probables (Slovic, 1966; Zakay, 1983). Les anglophones utilisent à cet égard l'expression de *wishful thinking*. Celui-ci s'exprimerait, entre autres, dans la plus grande confiance donnée aux jugements de performance quand il y a des récompenses importantes rattachées. Plus généralement, la plus grande confiance attribuée à des événements positifs peut être due à l'effet Pollyanna (héroïne du roman de Eleanor Porter de 1913). Selon cet effet, les gens ont tendance à exagérer ce qu'il y a de positif dans leur vie et à minimiser les aspects négatifs. Il y a d'ailleurs une documentation abondante datant des années 1940 montrant déjà une auto-évaluation trop positive (Festinger, 1942; Frank, 1953; Irwin, 1944). Des recherches ultérieures n'ont pas infirmé ces observations. Ainsi, Kuiper, Olinger, MacDonald et Shaw (1985) constatent que le rappel de traits de personnalité positifs plus facilement que négatifs. Pour leur part, Green et Gross (1979) rapportent une évaluation de soi plus positive que l'évaluation d'autrui.

Les personnes peuvent avoir plus grande confiance dans leurs jugements de résultats positifs parce qu'ils pensent qu'ils ont un plus grand contrôle sur ces événements. Langer (1975) explique que la surconfiance peut être plus grande dans les situations impliquant les habiletés. Souvent ces situations caractérisées par la présence de la compétition, du choix, de la pratique, de

l'effort peut amener l'individu à croire que tout dépend de l'effort et de ses aptitudes. Il y aurait une illusion de contrôle : des indices suggérant un rôle de l'aptitude plus grand que de la chance, il y aurait apparence de contrôlabilité. Weinstein (1980) obtient des données qui supportent cette interprétation de la surconfiance.

En revanche, Hoch (1985) constate que des personnes sont moins confiantes quant à leur perspective d'un futur emploi lorsqu'elles doivent donner des raisons qui pourraient faire que leur recherche d'emploi pourrait ne pas être fructueuse. Autrement dit, une analyse plus approfondie de la situation diminuerait la surconfiance. Cela pourrait expliquer en partie les résultats obtenus auprès des individus déprimés qui montrent qu'ils sont plus réalistes que ceux qui ne sont pas déprimés (Alloy et Abramson, 1988; Taylor et Brown, 1988). Ils seraient plus aptes à reconnaître qu'ils ont peu de contrôle sur la production des événements (Alloy, Abramson et Viscusi, 1981). Aussi, ils n'utiliseraient pas de biais de complaisance dans les attributions de succès et d'échecs (Campbell et Fairey, 1985). D'ailleurs, selon Bandura (1986), les meilleurs jugements d'auto-efficacité seraient ceux qui excèdent un peu ce qu'une personne peut accomplir, cette surestimation servant à accroître l'effort et la persévérance.

RÉALISME DES ÉTUDIANTS

Dans la section qui suit, nous rapportons les principaux résultats des recherches portant sur les prédictions académiques et professionnelles des étudiants en fonction de l'expérience qu'ils acquièrent et en fonction de la distance temporelle qui sépare leur prédiction de l'événement prédit.

Prédictions en fonction de l'expérience

Le réalisme des étudiants augmente-t-il avec l'expérience? Les examens périodiques auxquels sont soumis les étudiants constituent un objet approprié pour apporter une réponse à cette question. Les examens, en effet, par définition, comportent des réponses exactes à des questions. Si on

demande aux étudiants d'accompagner leurs réponses d'un jugement de confiance ou si on leur demande de prédire le nombre de leurs bonnes réponses, on peut se servir de la relation entre un critère objectif et leur jugement de confiance pour établir leur niveau de réalisme.

En fait, la mesure du réalisme des étudiants peut se faire à différents moments : avant l'apprentissage de la matière, pendant son apprentissage, lors de la recherche de la réponse et lors de l'émission de la réponse. Avant l'apprentissage, l'étudiant pourra, par exemple, baser son jugement sur la facilité qu'il a d'apprendre en général, durant l'apprentissage et lors de la phase de rétention, il fondera son jugement sur son estimation de connaissance de la matière, enfin, lors du recouvrement de l'information, il exprimera sa confiance à l'égard de la réponse donnée (Nelson et Narens, 1990). Plus précisément, les recherches sur le réalisme des étudiants se situent aux différents moments suivants : lors de l'inscription à un cours, pendant l'étude de la matière, une fois la matière apprise, mais avant l'examen, pendant l'examen (ou après l'examen quand il s'agit pour les étudiants de *postdire* combien de réponses sont correctes), après avoir reçu les notes à un examen partiel et avant un autre examen dans la même matière, après avoir reçu une note dans plusieurs examens et avant d'en prendre un nouveau, et ainsi de suite.

L'hypothèse générale est que l'étudiant, avec l'expérience, aura reçu une information de plus en plus riche. En principe, cela devrait favoriser une amélioration de la qualité de son jugement. Toutefois, comme nous le verrons dans ce qui suit, les recherches ne confirment pas nécessairement cette hypothèse. Ce n'est pas parce les étudiants ont reçu de l'information, même de qualité et abondante qu'ils l'utiliseront à bon escient. Ce qu'ils feront de cette information varie beaucoup en fonction de la personnalité et de la situation. Dans ce qui suit, nous présenterons les données des recherches en fonction de la nature du feedback qu'ils apportent à l'étudiant : à la suite de l'étude de la matière, à la suite d'un examen (questions) de

pratique, à la suite de l'examen, à la suite de la transmission de la note de l'examen. Toutefois, ces recherches que nous rapporterons couvrent parfois plus d'une de ces possibilités. Nous avons donc choisi de rapporter les recherches qui touchent à plus d'une possibilité en un seul endroit, c'est-à-dire là où elles portent sur le niveau d'information le plus riche (par exemple, dans la section « feedback à la suite d'un examen » plutôt qu'à la fois dans les sections « feedback à la suite d'un examen de pratique » et « feedback à la suite d'un examen »). Cela signifie que les conclusions tirées à la fin de chacune des sections, sauf pour la dernière, ne peuvent être que provisoires.

À la suite d'un feedback donné par l'étude de la matière

Avant même de se confronter à la matière à étudier, l'étudiant peut baser ses prédictions de résultats sur ce qu'il sait de la matière en général, sur la réputation du professeur qui l'enseigne, sur la perception de ses aptitudes en tant qu'étudiant, etc. (Butler et Winne, 1995; Hacker, Bol, Horgan et Rakow, 2000). Lors de l'apprentissage de la matière, les étudiants se livrent à un monitoring de cet apprentissage qui leur procure un feedback sur ce qui a été appris. Ce monitoring, un processus métacognitif, est réussi s'il mène l'étudiant à prédire avec une certaine exactitude son résultat à l'examen. Sinon, il y a « illusion de connaître » (Glenberg, Wilkinson et Epstein, 1982). Le feedback obtenu tout au long de son étude permet aussi à l'étudiant d'adapter ses stratégies d'apprentissage en fonction de la note visée. Par exemple, il s'allouera plus de temps pour étudier s'il perçoit qu'il n'a pas atteint le but recherché ou encore, il fera un meilleur choix de techniques mnémoniques (cf. Shaughnessy, 1979). Pour Maki et Berry, (1984), l'étudiant doit être capable de dire si la matière a été suffisamment apprise même si la nature exacte des questions n'est pas connue. Bien que souvent les étudiants rapportent avoir bien étudié, ils connaissent un échec. Pour ces auteurs, la prédiction du résultat à un examen fait appel à deux sortes de jugements : la métacompréhension, qui est un jugement sur la

compréhension de la matière et la prédiction quant à la rétention en mémoire de ce qui a été appris.

Afin de vérifier la qualité du monitoring des étudiants pendant l'étude, les chercheurs ont utilisé différentes approches : voir si les étudiants sont conscients de la plus grande efficacité d'une façon d'étudier qu'une autre, introduire des déformations dans un texte pour voir si cela allait affecter le niveau de confiance des étudiants, leur demander d'estimer leur possibilité d'appliquer un principe à propos duquel ils ont lu, leur fixer un critère de réussite à atteindre par leur étude, se demander si la confiance de réussite dépendait de la familiarité perçue avec un domaine d'étude.

Dans une première étude (Shaughnessy, 1981) 32 étudiants du premier cycle apprennent une liste de 40 paires de mots. Vingt de ces paires sont apprises par répétition et les 20 autres par imagerie interactive de la relation entre les deux mots de chaque paire. Les étudiants doivent estimer leur niveau de confiance quant à leur réussite à s'en rappeler (échelle de 0 à 100). Le niveau de confiance moyen pour ceux de la condition répétition est de 52.7, alors que ce niveau atteint 53.9 dans le cas où les étudiants apprennent les paires par imagerie. La différence entre les deux niveaux de confiance n'est pas significative. Par ailleurs, dans l'ensemble, les jugements sont surconfiants. Les étudiants ne se souviennent que d'environ la moitié des paires dont ils disent pouvoir se souvenir.

Dans une autre étude de la même recherche, Shaughnessy (1981) utilise les mêmes 40 paires de mots avec un premier groupe d'étudiants ($n = 24$) qui doivent dire de laquelle de deux paires ils se souviendront le mieux lorsque le chercheur présente 20 pairages de deux paires. Dix de ces 20 pairages sont mixtes, c'est-à-dire que l'une des paires a été apprise par répétition et l'autre par imagerie. Un deuxième groupe ($n = 24$), doit se livrer à un rappel libre après la présentation de la liste. Enfin, un troisième groupe ($n = 24$), doit identifier comment les paires ont été apprises, par répétition ou

par imagerie. Ce que le chercheur appelle le test du jugement de connaissance est le nombre de fois où les paires avec imagerie seront choisies comme plus probablement retenues dans les 10 pairages de paires qui sont mixtes par les sujets du premier groupe. Or, ce groupe choisit l'imagerie 5.7 fois, ce qui n'est pas significatif. Les résultats sont pourtant meilleurs de façon significative avec l'apprentissage par imagerie (6.25).

Pour leur part, Pratt, Luszcz, McKenzie-Keating et Manning (1982) demandent à 50 élèves inscrits à l'école secondaire jusqu'à quel point (échelle de 1 à 5), ils comprennent et se souviennent d'une histoire constituée de 15 propositions. Ce jugement est suivi d'un rappel immédiatement après ou sept jours plus tard. L'histoire en question est présentée de façon intacte ou avec une modification portant sur une partie plus ou moins importante. Cette partie étant présentée à un endroit inhabituel de l'histoire. Le jugement de connaissance de l'histoire est plus faible lorsque les déformations ont trait à des parties importantes de l'histoire. Les corrélations entre le rappel de la partie et l'estimation globale de la connaissance de l'histoire sont significatives pour trois parties de l'histoire sur cinq dans la condition délai et deux parties sur cinq dans la condition de rappel immédiat. Quand les parties déplacées sont moins essentielles, le jugement ne diffère pas de la condition intacte.

Glenberg et Epstein (1985) présentent à 85 sujets (38 femmes et 47 hommes), 15 textes d'un paragraphe traitant de toutes sortes de questions. À la suite de la présentation des 15 textes d'un paragraphe ou après la lecture de chaque texte, les chercheurs demandent aux sujets d'exprimer leur niveau de confiance (échelle de 1 à 6) quant à leur possibilité de vérifier éventuellement une inférence reliée à une proposition centrale du texte. Puis, les chercheurs présentent deux inférences, l'une correcte, l'autre fausse, entre lesquelles doivent choisir les sujets. La corrélation entre le niveau de confiance et la performance est $r_{pb} = .07$ dans la condition immédiate et $r_{pb} = .04$ dans la condition délai. Ces deux corrélations ne

diffèrent pas significativement entre elles ni de zéro. Dans l'ensemble la calibration est très faible. Le CAQ dans la condition immédiate est de .18 et dans la condition délai de .09. Encore là, cette différence n'est pas significative et n'est pas différente de 0.

Dans une recherche de Glenberg, Sanocki, Epstein et Morris (1987), deux études sont pertinentes par rapport à la question qui nous intéresse ici. Dans la première, 80 étudiants inscrits à l'université (niveau non spécifié) lisent 15 textes sur divers sujets. La moitié des étudiants se prononce sur leur aptitude à utiliser un principe illustré dans ces textes pour faire des inférences, mais l'autre moitié doit se prononcer sur leur capacité à reconnaître le mot à mot d'un principe lorsqu'ils doivent choisir entre ce principe et un autre énoncé qui en est une paraphrase. Il y a deux autres conditions. Dans la condition immédiate, la lecture de chaque texte est suivie de l'évaluation de confiance (1 à 6). Dans la condition avec délai, les textes sont présentés sur des pages consécutives et les 15 évaluations de la confiance se font à la fin du 15^e texte. La calibration est calculée à l'aide du coefficient de corrélation bisérielle de point r_{pb} et de la statistique Gamma (γ). Les résultats montrent qu'il n'y a pas de différence entre les conditions et que les sujets ne sont pas réalistes.

Dans une deuxième étude, 40 étudiants inscrits à un cours d'introduction en psychologie lisent des textes identiques à ceux de la première expérience. En plus, pour chaque texte, dans quatre problèmes, ils doivent choisir entre la paraphrase d'une idée dans le texte et une fausse réponse. Ils indiquent leur niveau de confiance de la même façon que dans la première étude. Il s'avère qu'une seule mesure de réalisme sur huit est significativement différente de zéro. Il s'agit de celle de la condition immédiate basée sur l'ensemble des quatre items.

Une troisième étude a pour but de vérifier si la familiarité avec un domaine prédit le niveau de confiance, alors que la performance, elle, ne

serait pas reliée à la confiance ni la familiarité à la performance. Les auteurs constituent trois groupes à partir de 88 sujets : les sujets du premier groupe évaluent la familiarité, expriment leur niveau de confiance et donnent une réponse pour chaque texte (on retient, en ce qui les concerne, le jugement de familiarité); les sujets du deuxième groupe se livrent à un rappel d'information pour chaque texte, expriment un niveau de confiance et donnent leur réponse (les chercheurs retiennent ici les données sur le rappel); les sujets du troisième groupe expriment un niveau de confiance et fournissent une réponse. Les données confirment les prédictions. La familiarité perçue est reliée à la confiance ($r = .57$), mais pas à la performance ($r = -.03$), d'où un mauvais réalisme.

Deux autres études permettent aux chercheurs de s'assurer que le facteur qui nuit à la calibration est bien la familiarité avec le domaine plutôt que la familiarité avec le texte. Pour ce faire, ils fournissent aux sujets un énoncé mot à mot du principe ou une paraphrase de celui-ci après avoir vérifié que cette manipulation affecte la familiarité avec le texte, ils constatent qu'elle n'affecte pas la confiance.

Snyder et Pressley (1989) présentent à des étudiants un texte de 12 pages à lire du début à la fin sans revenir en arrière. Puis, ils demandent d'étudier le texte à nouveau jusqu'à ce qu'ils puissent obtenir une note de 80%. Les étudiants, pour la plupart, utilisent le temps supplémentaire en étudiant le texte d'une façon non sélective (relecture du début à la fin). Pratiquement tous s'arrêtent avant de pouvoir répondre à 80% de la matière.

Il ressort des recherches rapportées jusqu'ici que les étudiants font preuve d'un monitoring de leur apprentissage qui est déficient. Le manque de réalisme et, en particulier, la surconfiance est la règle plutôt que l'exception. Nous verrons, dans ce qui suit, si les recherches qui se sont intéressées au réalisme des étudiants à des étapes ultérieures viennent conforter ces résultats ou non.

À la suite du feedback donné par un examen (questions) de pratique

Cette section contient peu d'études et elles ont pratiquement toutes comme origine les travaux de Glenberg et de ses collègues. Ces recherches n'avaient pas pour objectif de déterminer si la présentation d'un examen de pratique avant l'examen critère avait un effet sur le réalisme des étudiants par rapport à une condition où de tels examens n'étaient pas présentés. Elles ne comportent donc pas de comparaisons par rapport à ces deux conditions. Nous les avons quand même retenus parce qu'elles peuvent nous éclairer sur le réalisme des étudiants lorsqu'un examen de pratique est disponible.

Bol et Hacker (2001) rapportent qu'il y a habituellement amélioration de la performance avec des examens de pratique. C'est ce que l'on observe, par exemple, avec ceux qui se pratiquent pour réussir les tests d'admission américains que sont le *SAT* ou le *GRE (Graduate Record Examination)*. Ils attribuent cette meilleure performance à différents facteurs comme une plus grande familiarité avec la sorte d'items ; une plus grande confiance en soi ; une amélioration des aptitudes pour répondre à des examens ; etc. Cela, cependant ne nous éclaire pas sur l'amélioration du réalisme des étudiants à la suite d'un examen de pratique.

Glenberg, Wilkinson et Epstein (1982) utilisent une façon de mettre en lumière ce qu'ils appellent « l'illusion de connaître » des étudiants qui consiste à voir si une personne pourra discerner une contradiction dans un texte en percevant qu'elle n'a pas compris ce texte. La mémoire n'est pas en cause puisque le texte est accessible en tout temps à la personne. Dans une étude, ils présentent des textes de trois paragraphes portant sur trois questions. Le dernier paragraphe de ces textes est critique. Il y en a trois versions: (1) *contrôle*, où la dernière phrase est la continuation de la précédente; (2) *donnée*, où l'information de la dernière phrase contredit l'information de la phrase précédente (par exemple, remplacer « changement erratique » par « stabilité »); (3) *nouvelle*, où l'information nouvelle contredit

l'information de la phrase précédente. Les arguments des dernières phrases sont consistants de façon interne, mais contredisent les phrases qui précèdent. On dit aux sujets (94 étudiants inscrits à un cours d'introduction en psychologie) que les textes peuvent contenir des contradictions, d'y être attentifs. Après lecture, ils posent un jugement par rapport à leur compréhension (échelle de 1 à 4). Avant d'être exposés aux textes expérimentaux, on leur soumet un texte de pratique où on leur signale et explique la contradiction. Les auteurs de la recherche établissent qu'il y a illusion de connaître en calculant la proportion des cas où la confiance est élevée (3 ou 4) sans qu'il n'y ait détection de la contradiction. Les résultats montrent qu'il y a significativement plus d'illusions pour la troisième condition ($p = .511$) que pour la deuxième ($p = .234$). Et cela vaut pour les trois textes. Les auteurs de la recherche concluent que la confiance des lecteurs vient d'un postulat par défaut : lors de la lecture d'un texte, comprendre progresse bien à moins d'un signal d'erreur (par exemple, un terme non compris).

Dans la deuxième étude de Glenberg et Epstein (1985), les chercheurs lisent une série d'exposés d'un paragraphe sur différents thèmes. Pour chaque texte, les sujets évaluent leur confiance en leur capacité d'utiliser ce qui a été appris dans le texte pour produire des inférences correctes à propos d'un thème central du texte. En plus, pour chaque texte, une inférence dérivée du texte est présentée pour vérification. À partir de l'évaluation de la confiance et des réponses de vérification, trois mesures de calibration sont calculées. Les chercheurs présentent à 60 sujets (35 femmes et 25 hommes) l'une de deux conditions : (1) *familiarisation*, où il y a présentation de trois des 15 textes et des réponses; (2) *contrôle*, où les trois textes ne sont pas suivis des inférences. Par la suite, les sujets reçoivent les 12 textes suivants et expriment leur niveau de confiance (échelle de 1 à 6) après chaque texte. Les corrélations bisérielles de point entre la confiance et la performance sont $r_{pb} = .06$ et $r_{pb} = .12$ respectivement pour la condition contrôle et familiarisation. Les différences ne sont pas significatives par

rapport à zéro ou l'une par rapport à l'autre. Il en est de même des CAQ (.15 et .31). Le réalisme est donc mauvais.

Glenberg et al. (1987) rapportent une étude de Lovelace (1984) où des sujets peuvent prédire avec une certaine exactitude leur performance de reconnaissance de phrases ou de mots pairés mémorisés. Selon Glenberg et al., cela pourrait être le signe que les sujets testent spontanément leur mémoire durant l'apprentissage et utiliseraient ce feedback pour prédire leur performance. Les auteurs font l'hypothèse que le feedback à partir d'un examen préparatoire sera d'autant plus utile que les processus et les connaissances nécessaires pour répondre à l'examen lui-même sont semblables à ceux générés par le feedback. Dans l'étude 6 de leur recherche, ils veulent vérifier l'hypothèse selon laquelle un feedback à partir d'un examen de pratique peut prédire la performance à l'examen dans la mesure où les processus cognitifs et les connaissances nécessaires sont semblables dans les deux examens. Ils demandent donc à 48 étudiants inscrits à des cours d'introduction en psychologie de lire 16 textes suivis de 16 examens de pratique (problèmes de reconnaissance d'une idée). À la suite de leurs réponses, ils doivent donner un jugement de confiance par rapport à un examen à venir. Cet examen contient trois problèmes de reconnaissance d'idées : (1) *pareil*, même problème que dans l'examen de pratique; (2) *semblable*, problème qui est une paraphrase du même problème de l'examen de pratique; (3) *non relié*, problème tiré du texte, mais pas relié au même problème de l'examen de pratique. Les résultats montrent qu'il y a une certaine calibration pour la condition *pareil* ($r = .13$, $\gamma = .10$) et *semblable* ($r = .12$, $\gamma = .14$), mais pas pour *non relié* ($r = .08$, $\gamma = .04$). Cela ne constitue pas une confirmation solide de l'hypothèse.

Afin de mieux vérifier leur hypothèse, les auteurs réécrivent les problèmes moins bien réussis. Les problèmes sont présentés à 38 étudiants inscrits à un cours d'introduction en psychologie. La mesure de calibration

pour *pareil* donne une corrélation $r = .26$ (significatif par rapport à zéro) ($\gamma = .40$); pour *semblable* un $r = .07$ (significatif) ($\gamma = .08$); et pour *non relié* un $r = .04$ (non significatif) ($\gamma = .07$). L'hypothèse est donc solidement confirmée selon les auteurs.

Enfin, dans une dernière étude, les auteurs, au lieu de présenter l'idée verbatim dans les problèmes de reconnaissance d'idées, présentent cette idée de façon paraphrasée. Ils veulent ainsi éviter qu'il y ait une reconnaissance de surface de l'idée, ce qui n'est pas un signe de compréhension. Ils proposent ces problèmes à 37 étudiants qui ont les mêmes caractéristiques que les précédents. La mesure du réalisme pour *pareil* donne une corrélation de $r = .27$ (significatif par rapport à zéro) ($\gamma = .35$); pour *semblable* une corrélation de $r = .11$ (significatif) ($\gamma = .12$); et pour *non relié* une corrélation de $r = .18$ (non significatif) ($\gamma = .22$). La différence dans la calibration entre *pareil* et *semblable* est significative, même s'il y a calibration dans les deux cas. Ils expliquent la calibration pour *non reliée* par une erreur du Type I.

Leurs résultats montrent donc que le réalisme de la compréhension peut être amélioré par un examen préliminaire, mais que cet examen doit ressembler de très près à l'examen lui-même. Par ailleurs, les jugements qui sont basés sur des sentiments de familiarité non différenciés, même s'ils peuvent être subjectivement contraignants, ne permettent pas de prédire la performance demandant la connaissance d'un texte particulier.

Bol et Hacker (2001) retiennent comme sujets 59 étudiants du premier cycle inscrits à un cours de méthode de recherche en éducation. Ils doivent d'abord répondre à un examen de pratique dont les questions sont formulées pour être le plus semblables possible aux questions de l'examen réel. Puis les bonnes réponses de cet examen de pratique sont données en classe. L'examen réel est constitué de 25 questions à choix multiples et de cinq questions qui demandent un court développement. Les étudiants doivent

prédire leur résultat avant l'examen. Un groupe contrôle reçoit le même examen, mais qui n'est pas précédé d'un examen de pratique. Les chercheurs définissent l'exactitude par la valeur absolue de la différence entre les notes prédites et les notes obtenues. Pour ce qui est des questions à choix multiples, les étudiants qui ont droit à un examen de pratique s'avèrent moins exacts que les sujets du groupe contrôle dans leurs prédictions ($M = 9.87$ et $M = 6.14$ respectivement). Par ailleurs, il y a un effet du niveau de réussite : en fonction d'une division à la médiane, les meilleurs sont plus exacts que les moins bons ($M = 5.76$ et $M = 10.92$ respectivement).

Pour ce qui est des questions à court développement, il n'y a pas de différence en terme d'exactitude entre le groupe expérimental et le groupe contrôle. Lorsque le niveau de réussite est pris en compte, les résultats sont semblables à ceux des questions à choix multiples ($M = 3.77$ pour les meilleurs et $M = 7.55$ pour les autres). Les auteurs de la recherche concluent que la révision de la matière serait plus efficace qu'un examen de pratique pour favoriser la calibration. Il se peut aussi que les étudiants qui ont moins bien réussi ne basent pas leurs prédictions sur leur performance aux examens antérieurs, mais sur un concept de soi global en terme d'aptitude académique.

Nous pouvons résumer les recherches portant sur le réalisme des étudiants où ces derniers ont droit à un examen de pratique en affirmant que celui-ci n'augmente pas systématiquement le réalisme des étudiants. Souvent, l'information que donne cet examen n'a pas d'effet sur le réalisme. Des variables comme la forme des questions, le degré de ressemblance entre l'examen de pratique et l'examen critère, le niveau de réussite de l'étudiant, peuvent faire que le réalisme est plus ou moins grand. Il reste, toutefois, que, même dans les meilleures conditions, il n'atteint pas un niveau très élevé.

À la suite du feedback donné par l'examen

Nous verrons maintenant si le feedback donné par le fait de passer un examen (lorsque les étudiants se prononcent sur l'exactitude de chaque question) ou d'avoir passé l'examen (lorsqu'ils se prononcent à la fin de l'examen sur la note obtenue) rend les étudiants plus réalistes quant à leurs chances de réussir.

Dans une recherche de Bradley (1981), 92 hommes et 92 femmes inscrits à un cours d'introduction de psychologie au premier cycle évaluent leurs connaissances dans 12 domaines (par exemple, la logique, la littérature, la nutrition). Puis, le chercheur administre un examen de 36 questions (vrai/faux) sur le contenu de chacun des domaines où les étudiants doivent donner leur niveau de confiance (échelle 1 à 5). De ces questions, 15 sont très difficiles (une réponse correcte semble due à la chance). Les résultats montrent que l'admission d'incertitude complète (1) diminue avec l'accroissement de l'expertise quand la réponse est fautive. La certitude quand les étudiants ont de mauvaises réponses augmente avec le degré d'expertise et ce, autant pour les questions très difficiles que pour l'ensemble des questions. La certitude des moins experts lorsque la réponse est fautive est de .10, alors que celle des plus experts pour le même type de réponse est de .40. Il y a donc, un refus d'admettre leur ignorance et un refus d'admettre l'incertitude par les « experts », même s'il y a surconfiance chez les deux groupes de sujets.

Dans une étude où Pratt et al. (1982) se demandent si l'information dans certaines parties d'une histoire affecte les jugements métacognitifs de compréhension et la mémoire de l'histoire, ils présentent à 64 étudiants inscrits à un cours d'introduction à la psychologie une histoire constituée de 15 propositions sous deux formes : intacte et modifiée. La modification consiste à remplacer les noms par des pronoms ambigus dans trois des 15 propositions. Après avoir écouté l'histoire, les étudiants doivent dire à quel

point ils comprennent et se souviennent de l'histoire en utilisant une échelle de 1 à 5. Ce jugement est suivi d'un rappel de l'histoire. Ce déroulement est répété sept jours plus tard. La corrélation entre le rappel d'une partie de l'histoire et le jugement de connaissance de l'histoire dépend de l'importance de la partie de l'histoire en question. Il y a surconfiance quand la partie modifiée est moins importante.

Glenberg et Epstein, (1987) demandent à 55 étudiants qui ont suivi un bon nombre de cours en musique ou en physique de lire des textes à la suite de quoi ils devront répondre à un examen où ils choisissent entre deux inférences découlant de ces textes dont l'une est bonne. Ils indiquent aussi leur niveau de confiance à l'égard de la réponse sur une échelle en six points. Les résultats donnent une relation significative exprimée à l'aide de l'indice Gamma (γ) entre la confiance et la performance. Pour les étudiants en musique $\gamma = .42$ et pour les étudiants en physique $\gamma = .36$. Le réalisme n'est donc pas lié à l'expertise. Les auteurs concluent que les étudiants peuvent évaluer de façon assez exacte la qualité de leur performance dans un examen de vérification d'inférence et que la confiance des étudiants à l'égard de leur performance est basée sur l'auto-classification.

Dans une première étude où ils étudient le *perceived readiness for examination performance*, (Pressley, Snyder, Levin, Murray et Ghatala, 1987), 54 étudiants inscrits à un cours d'introduction à la psychologie lisent un chapitre d'un livre d'introduction. Cette lecture est suivie d'un examen de 50 items de questions à choix multiples. Trois questions sont présentées à titre d'exemple. Les étudiants doivent faire des prédictions avant de lire le texte, après sa lecture et une postdiction après l'examen (le nombre de questions sur 50 qu'ils pensent réussir ou qui ont été réussies). Les auteurs calculent la différence absolue entre les prédictions (postdiction) et les résultats. Ils constatent une différence significative entre la prédiction avant de lire le texte (9.11) et la postdiction à la suite de l'examen (2.83). En revanche, les différences entre la prédiction après la lecture du texte (6.33)

et la postdiction après l'examen, de même qu'entre la prédiction avant de lire le texte et après la lecture du texte ne sont pas significatives.

Dans la troisième étude, de la même recherche, Pressley et al. (1987), demandent à 162 étudiants (103 femmes, 59 hommes) inscrits à un cours d'introduction en psychologie, de lire un texte et de tenter de répondre à des questions de pratique (dispersées ou regroupées à la fin). Les étudiants doivent répondre à 30 questions où ils complètent une phrase. Ils font trois prédictions par rapport à leur performance (avant la lecture, après la lecture et après l'examen). Lorsqu'il n'y a pas de questions de pratique, les résultats sont semblables à ceux de l'étude précédente (4.44 avant la lecture, 3.67 après la lecture et 1.28 après l'examen), sauf que la différence entre la condition après la lecture et la condition après l'examen est significative, l'exactitude étant meilleure dans ce dernier cas. Lorsqu'il y a des questions de pratique, les différences d'exactitude après la lecture (4.17 pour les questions dispersées et 4.67 pour les questions regroupées) et après l'examen (1.78 pour les questions dispersées et 2.11 pour les questions regroupées) sont significatives par rapport à avant la lecture. Le niveau d'exactitude après l'examen est de 1.50 pour les questions dispersées et 2.72 pour les questions regroupées. La différence entre les questions regroupées et dispersées pour l'ensemble des autres conditions regroupées n'est pas significative.

Ghatala, Levin, Foorman et Pressley (1989) se demandent combien de temps il faudra pour que quelqu'un estime qu'il a assez étudié pour atteindre une quantité fixe d'information. Dans une étude, ils ont recours à trois conditions. Dans la première, les élèves lisent et écoutent un passage qui est suivi d'un examen. Ils peuvent répéter le cycle étude-examen jusqu'à ce qu'ils disent vouloir arrêter d'étudier. Les mêmes items reviennent dans chaque examen, mais dans un ordre différent. Les élèves doivent estimer combien d'items ils pensent avoir réussi après chaque examen. Dans la seconde condition, les élèves doivent aussi estimer le nombre d'items

réussis, mais ils n'ont pas à prendre de décision quant à la poursuite de l'étude. Dans la troisième condition, ils ne sont pas avertis qu'ils devront estimer le nombre de questions réussies et n'ont pas à prendre de décision quant à la poursuite de l'étude. Dans chaque condition, la moitié des élèves doit exprimer un niveau de confiance à chaque item (échelle de 1 à 4). Les auteurs de la recherche ont recours à 72 élèves de 3^e et 4^e années. Les données indiquent une inexactitude du monitoring importante (score moyen = 2.17). Pour 93.5% des inexactitudes, les élèves croient qu'ils ont plus d'items corrects qu'ils n'en ont. En ce qui concerne le niveau de confiance pour les items, dans la condition 1, il est de 3.15 pour les items réussis et de 2.57 pour les items manqués; dans la condition 2, il est de 3.39 pour les items réussis et de 2.69 pour les items manqués; dans la condition 3, il est de 3.41 pour les items réussis et de 2.78 pour les items manqués. Il n'y a pas d'effet selon les conditions. La comparaison des estimés et de la confiance par item montre que les élèves qui ont les meilleurs estimés sont aussi plus exacts en regard de la confiance.

Dans une autre étude, auprès de 42 élèves, les chercheurs opposent les examens qui exigent une courte réponse, aux examens faits de questions à choix multiples. Les questions sont les mêmes, mais au lieu de choisir entre des réponses possibles, les élèves doivent remplir un espace. Les résultats montrent une inexactitude du monitoring dans la condition qui correspond à la condition 1 ci-dessus (2.93, 11 élèves sur 14 surestiment); alors que le monitoring est plutôt exact dans les autres conditions qui correspondent à la 2^e et la 3^e de l'étude précédente (.79 et .43 respectivement). Les deux dernières conditions incitent donc les élèves à étudier les passages plus longtemps.

Beyer (1990) demande à 39 hommes et 34 femmes inscrits à un cours d'introduction en psychologie de répondre à cinq questionnaires comportant chacun 10 questions à choix multiples. Deux questionnaires posent des questions « masculines », elles portent sur la politique et le sport. Deux

questionnaires posent des questions « féminines », elles portent sur les vedettes et la mode. Un dernier questionnaire est composé de questions « neutres » (par exemple, des anagrammes). Après avoir répondu aux questionnaires, les étudiants doivent estimer le nombre de réponses réussies. L'exactitude est définie comme étant la différence entre la note obtenue et la postdiction. Les résultats montrent qu'il y a sous-estimation de la part des femmes pour la tâche « masculine » (-11.8) et pour une des tâches « neutres » (-4.8). Par ailleurs, il y a surestimation autant de la part des femmes (10.8) que des hommes (19.5) pour les mêmes catégories de questions « neutres ». En ce qui concerne les autres questionnaires, les hommes et les femmes se montrent exacts dans leurs postdictions. Enfin, la mise en relation de la prédiction et de la postdiction montre que la prédiction est ce qui explique la postdiction plutôt que le genre.

Une deuxième étude a pour but de confirmer avec plus de certitude les résultats obtenus dans la première étude. Elle s'adresse à un plus grand nombre de sujets ; 85 étudiants et 92 étudiantes inscrits à un cours d'introduction en psychologie. Le nombre d'items est augmenté à 40 et ces items sont plus difficiles. Il y a encore une fois trois types de tâches : « masculine » (vedettes sportives), « féminine » (vedettes du cinéma et de la télévision) et « neutre » (détection de caractères). La chercheuse leur demande encore combien de questions ils pensent avoir réussies. Les postdictions s'avèrent exactes pour les tâches féminines autant pour les étudiantes (inexactitude = 0) que pour les étudiants (inexactitude = -1). Il y a sous-confiance autant pour les étudiantes (inexactitude = -6) que pour les étudiants (inexactitude = -5) en ce qui concerne les tâches neutres. En revanche, il y a sous-confiance de la part des étudiantes (inexactitude = -5) et surconfiance de la part des étudiants (inexactitude = 4) en ce qui concerne les tâches masculines. Ces résultats confirment que les attentes prédisent en partie les différences d'exactitude de la confiance.

Pour Maki et Serra (1992a), il se peut que les résultats de Glenberg, Sanocki, Epstein et Morris (1987) ne montrant pas d'effet de pratique soient dus au fait de ne présenter qu'un seul item. Dans une première étude, ils reprennent donc les études 6 et 8 de Glenberg et al.. Ils utilisent les mêmes textes auprès de 72 étudiants, mais des conditions différentes. Pour la pratique, les étudiants ont droit à trois questions au lieu d'une seule. Les examens de pratique et les examens critères sont constitués de questions à choix multiples. Les questions ne sont pas les mêmes pour l'examen de pratique et l'examen critère, mais les idées sur lesquelles portent les deux examens sont liées. Dans une autre condition, il n'y a pas de pratique. Par ailleurs, les textes peuvent être lus sous formes d'unités d'idées ou sous forme de pages entières. Les étudiants lisent donc d'abord quatre textes parmi 12, puis ils répondent à une question de pratique pour chaque texte. Ils ne reçoivent pas de feedback à propos de leurs réponses. Les étudiants doivent mettre en rang les textes en ce qui concerne la performance prédite avant et après l'examen. Les données sont analysées à l'aide de la corrélation Gamma (γ). Dans le cas de la condition pratique, le γ pour la prédiction est de $\gamma = .220$ et de $\gamma = .348$ à la suite de l'examen. Pour ce qui est de la condition où il n'y a pas de pratique, le γ pour la prédiction est de $\gamma = .227$ et de $\gamma = .319$ à la suite de l'examen. Toutes ces valeurs sont significatives par rapport à zéro. Il n'y a pas d'effet significatif selon les conditions, mais il y a un effet entre la prédiction et la postdiction quand les valeurs des différentes conditions sont combinées. Les résultats sont donc les mêmes que ceux de Glenberg et al. (1987). Il n'y a pas d'augmentation de l'exactitude de la prédiction malgré la pratique.

Dans la deuxième étude, les chercheurs retiennent quatre conditions qu'ils présentent à 72 étudiants inscrits à un cours d'introduction en psychologie. Dans la première, les questions pour l'examen de pratique et pour l'examen critère sont les mêmes, dans la deuxième, elles sont semblables, dans la troisième, il n'y a pas de pratique, dans la quatrième, la

question est vue lors de la pratique, mais pas les choix de réponse. Il y a trois questions au lieu d'une pour l'examen de pratique. Il y a également trois questions pour l'examen critère. Dans la condition où il y a pratique, les étudiants indiquent leur niveau de confiance de 1 à 7 pour chaque question. Les textes sont les 12 textes de l'étude 1 étudiés quatre textes à la fois. Les étudiants doivent mettre les textes en rang en termes de compréhension avant et après l'examen critère. Les corrélations Gammas obtenues pour la prédiction et la postdiction sont toutes plus grandes que zéro (prédiction : examen identique $\gamma = .321$; examen semblable $\gamma = .105$; question sans les réponses possibles $\gamma = .235$; pas de pratique $\gamma = .275$; postdiction : examen identique $\gamma = .114$; examen semblable $\gamma = .078$; question sans les réponses possibles $\gamma = .094$; pas de pratique $\gamma = .055$). Ces différences par rapport à zéro sont toutes significatives, sauf pour la prédiction dans la condition semblable. En ce qui concerne les différences entre les conditions pour la prédiction, il n'y a pas de différences significatives sauf entre la condition semblable et la condition identique. Il y a donc une plus grande exactitude dans la condition identique que dans la condition semblable. Il n'y a pas de différence entre les conditions de pratique pour ce qui est de la postdiction. La seule différence significative se situe dans la comparaison entre la prédiction et la postdiction, la postdiction étant plus exacte. Les corrélations Gammas entre le niveau de confiance et la performance pour les conditions semblable et identique dans les examens de pratique (donnent une mesure du feedback autogénéré) ne montrent pas de différence. Le feedback est significativement différent de zéro dans les deux groupes et le niveau d'exactitude ne diffère pas entre les deux groupes. Cette étude montre aussi que les étudiants qui sont les plus exacts pour l'examen de pratique sont aussi les plus exacts pour l'examen critère.

Selon la quatrième hypothèse de la recherche de Sjostrom et Marks (1994), le réalisme est meilleur pour la postdiction que pour la prédiction. Les chercheurs demandent à 59 étudiantes et 31 étudiants de lire 12 textes

d'introduction à la psychologie présentés sur deux trimestres. Chaque texte est suivi d'un examen de 20 questions à choix multiples. Avant l'examen et après l'examen, les étudiants doivent indiquer leur niveau de confiance sur une échelle en sept points (échelle de 1 à 7) par rapport à la réussite de cet examen. Un test t entre les corrélations pour les prédictions et celles pour les postdictions montre que, dans l'ensemble, il n'y a pas de différence entre les deux. Les chercheurs concluent que les étudiants sont capables d'évaluer la difficulté prospective d'un examen. Ils observent aussi que la confiance augmente inversement avec la difficulté de l'examen. Le niveau de confiance des étudiants semble basé sur l'évaluation de leur aptitude générale par rapport à un cours ajustée pour la difficulté spécifique de l'unité. Selon eux, les prédictions pourraient être des effets de révélation.

Dans une première étude de Schraw, Dunkle, Bendixen et Roedel (1995), 59 hommes et 75 femmes inscrits à un cours d'introduction à la psychologie de l'éducation répondent à un questionnaire de questions à choix multiples portant sur sept domaines : La distance géographique entre des villes, les années où des présidents américains ont pris le pouvoir, la valeur calorique d'aliments, la vitesse maximum de course de certains animaux, des problèmes mathématiques de la vie de tous les jours, des jugements spatiaux et des connaissances générales. Il y a huit questions dans chaque domaine. Les étudiants doivent indiquer leur niveau de confiance en cochant un endroit sur un continuum de 0% à 100% de confiance. Ils utilisent une mesure de réalisme qui est la différence entre la confiance moyenne pour un examen et la performance moyenne dans cet examen. Cela permet de dire dans quelle mesure un individu est surconfiant ou sous-confiant pour chaque examen. La valeur de cet indice varie de -1 à $+1$. Une valeur plus grande que zéro est signe de surconfiance, alors qu'une valeur plus petite que zéro est signe de sous-confiance. Pour leur part, ils observent des valeurs qui varient de $-.09$ à $.25$. Ils concluent que le niveau de confiance est en général assez exact. Par ailleurs, les corrélations sont assez élevées d'un domaine à

l'autre, ce qui va dans le sens d'une aptitude à s'auto-monitorer qui n'est pas spécifique à un domaine. Une seconde étude, auprès de 35 hommes et 100 femmes du premier cycle où l'homogénéité entre les questionnaires pour les différents domaines est mieux assurée confirme cette hypothèse encore plus clairement.

Beyer (1998), demande à 275 étudiantes et 213 étudiants inscrits à un cours d'introduction en psychologie ou en psychologie du développement de répondre à un de trois questionnaires de 35 questions chacun portant sur des questions typées comme étant féminines (vedettes du monde du show business), masculines (sport) ou neutres (littérature et géographie). Les sujets doivent indiquer avant et après avoir répondu, le nombre de questions qu'ils pensent avoir réussies. De plus, un sous-groupe doit indiquer, pour chaque question, son niveau de confiance de l'avoir réussi sur une échelle en sept points. Dans aucune des analyses, le fait d'avoir eu à indiquer son niveau de confiance pour chaque question n'a un effet sur l'exactitude de l'auto-évaluation en terme de nombre d'items réussis. En ce qui concerne la prédiction, par rapport à la tâche féminine, l'inexactitude des étudiantes est de 4 et celle des étudiants de 0.1 ; par rapport à la tâche neutre l'inexactitude des étudiantes est de 0.6 et celle des étudiants de 4.0 ; par rapport à la tâche masculine, l'inexactitude des étudiantes est de 1.8 et celle des étudiants de 2.4. La chercheuse n'établit pas si ces résultats présentent des différences significatives. Pour ce qui est de la postdiction, par rapport aux tâches féminines et neutres, il n'y a pas de différence entre l'exactitude de l'évaluation des étudiantes et des étudiants (inexactitude de l'auto-évaluation = -1.5 autant pour les deux groupes en ce qui concerne la tâche féminine et inexactitude de l'évaluation de 2.2 et -1.0 respectivement pour les tâches neutres). Par rapport à la tâche masculine, les auto-évaluations des étudiants sont exactes (inexactitude = 0.1) alors que celles des étudiantes démontrent de la sous-confiance (2.6). Pour ce qui est de l'indice de calibration, les étudiants s'avèrent mieux calibrés que les étudiantes en

ce qui concerne les tâches masculines (.066 par opposition à .104), alors qu'il n'y a pas de différence pour les deux autres types de tâches (les indices de calibration se situant entre .060 et .082). À performance égale, il y a une relation positive entre les attentes et l'auto-évaluation : en général, lorsque les attentes sont faibles, il y a sous-estimation ou estimation exacte. Les attentes constituent un facteur explicatif plus important que le genre.

Dans une recherche de Bol et Hacker (2001), 59 étudiants du premier cycle inscrits à un cours de méthode de recherche en éducation répondent à un examen intra et à un examen final constitués de 25 questions à choix multiples et de cinq questions demandant une réponse à court développement. Les étudiants prédisent le nombre de bonnes réponses avant l'examen et après l'examen. L'exactitude est définie comme étant la valeur absolue de la différence entre la postdiction et la note. Une valeur plus petite indique donc une meilleure exactitude. Les auteurs constatent que pour les questions à choix multiples de l'examen intra, en général, la postdiction est plus exacte que la prédiction, 12.04 par opposition à 16.00. Ils n'observent pas, par ailleurs, de relation avec la performance. Pour ce qui est des questions demandant un court développement, l'exactitude de la prédiction est de 5,62 alors que celle de la postdiction est de 4.73.

Beyer (2002) demande à 289 étudiantes et 174 étudiants d'indiquer avant l'examen et après l'examen combien de questions ils ont réussi. Il y a trois examens, un examen de mathématique typé « masculin », un examen de syntaxe typé « féminin » et un examen « neutre » d'histoire et de géographie. Il y a 30 items par examen. Les étudiants doivent dire avant et après l'examen combien de questions ils pensent pouvoir réussir. De plus, la moitié des sujets doit indiquer leur niveau de confiance après chacune des questions. Comme prévu, il y a une relation entre le genre et l'exactitude de la prédiction pour les mathématiques, mais pas pour les autres matières. Dans le cours des mathématiques, les étudiants, à attentes semblables, surestiment plus que les étudiantes. Par ailleurs, les attentes des étudiantes

affectent plus leur auto-évaluation. Les sujets qui ont des attentes faibles surestiment moins ou sous-estiment plus que les sujets qui ont des attentes élevées. La chercheuse ne donne pas les données pour les différentes conditions. Elle présente plutôt des résultats en terme d'une équation de régression.

Y a-t-il réalisme de la confiance lorsque les étudiants posent leurs jugements à la suite de la connaissance de l'examen ? Deux recherches de Beyer rapportent de la sous-confiance chez certains étudiants en fonction de la nature de la tâche qu'elle leur présente (Beyer, 1990, 1998). Ces résultats sont exceptionnels, même dans les recherches de Beyer, et s'expliquent par les attentes peu élevées de certains étudiants à l'égard de certaines tâches. D'autres recherches ont trouvé que les étudiants pouvaient faire preuve de jugements exacts (Beyer, 1990, 1998; Ghatala, Levin, Foorman et Pressley, 1989; Glenberg et Epstein, 1987; Pratt, Luszcz, McKenzie-Keating et Manning, 1982; Schraw, Dunkle, Bendixen et DeBacker Roedel, 1995; Sjostrom et Marks, 1994). Toutefois, ces jugements exacts sont souvent moins fréquents que des jugements inexacts dans les mêmes recherches. En fait, les jugements surconfiants sont présents dans plusieurs des études des recherches citées (Beyer, 1990 ; Bradley, 1981 ; Ghatala et al., 1989 ; Maki et Serra, 1992a ; Pratt et al., 1982 ; Pressley, Snyder, Levin, Murray et Ghatala, 1987 ; Beyer, 1998; Bol et Hacker, 2001). Très peu de recherches rapportent des comparaisons entre l'étape où l'étudiant a répondu à l'examen et des étapes antérieures (Beyer, 1990 ; Maki et Serra, 1992b ; Pressley et al., 1987 ; Beyer, 2002; Bol et Hacker, 2001; Sjostrom et Marks, 1994). Quand c'est le cas, la surconfiance est généralement plus grande dans l'étape antérieure (par exemple, avant l'examen par opposition à après l'examen).

À la suite du feedback donné par les notes

Dans cette section, nous présentons les recherches qui peuvent nous éclairer sur le réalisme de la confiance des étudiants une fois qu'ils ont reçu une (des) note(s) soit à l'intérieur du même cours (plus d'un examen), soit dans des cours différents.

Shaughnessy (1979) propose à 47 étudiants inscrits à un cours d'introduction en psychologie de répondre à quatre examens de questions à choix multiples. Les étudiants doivent indiquer leur niveau de confiance sur une échelle en quatre points à la suite de chacune de leurs réponses. Shaughnessy obtient un CAQ de .95 pour le premier examen, ce qui indique que les étudiants sont capables de distinguer entre l'information connue et celle qui ne l'est pas. Au quatrième examen cet indice atteint .73, mais la différence n'est pas significative.

Dans une première étude de Maki et Berry (1984), 30 étudiants qui suivent un cours d'introduction en psychologie lisent la moitié d'un chapitre de leur manuel et prédisent leur performance pour chacune des sections de ce demi-chapitre (échelle : sûrement, probablement, peut-être correct, sûrement, probablement, peut-être incorrect) à un questionnaire à choix multiple (18 questions) auquel ils auront à répondre le lendemain. Le jour suivant cet examen, ils doivent lire l'autre demi-chapitre, prédire leur performance et, le lendemain, répondre à un nouvel examen de questions à choix multiples. Pour la moitié des sujets, un feedback est donné quant à la réussite du premier examen et de chacune de leurs questions. On leur rappelle aussi leur prédiction. Pour chaque étudiant, le jugement de confiance moyen pour les bonnes réponses et le jugement de confiance moyen pour les mauvaises réponses sont calculés. Les résultats montrent qu'il n'y a pas d'effet du feedback sur l'exactitude.

La troisième étude de Glenberg et Epstein, (1985) a pour objectif de voir l'effet de l'expérience acquise à la suite de la lecture de 15 textes suivis d'un

examen auprès de 39 étudiants. Après la lecture de chacun des textes : (1) l'étudiant exprime sa confiance de pouvoir vérifier l'inférence, (2) il y a un examen de vérification d'inférence, (3) l'étudiant exprime sa confiance par rapport à la réponse donnée, (4) il exprime sa confiance de pouvoir vérifier la nouvelle inférence sur le même texte, (5) il y a un nouvel examen de vérification d'inférence. Les auteurs de la recherche qualifient de « recalibration » la corrélation entre les données des étapes 4 et 5. Les résultats montrent que la première calibration, celle qui provient de la relation entre 1 et 2 (r_{pb} moyenne = .04, CAQ = .09) ne diffère pas de zéro. La deuxième calibration, celle qui provient de la relation entre 2 et 3, est différente de zéro et différente de la première calibration (r_{pb} = .23 et CAQ = .68). La troisième calibration, celle qui provient de la relation entre 4 et 5 est aussi différente de zéro et différente de la première calibration (r_{pb} = .19, différent de zéro, différent de première calibration et CAQ = .44). Les auteurs regroupent les résultats des cinq premiers textes, des cinq textes suivants et des cinq derniers. Ils constatent que les différences entre les trois types de calibration ne changent pas du début à la fin. De plus, les trois mesures ne changent pas du premier texte au quinzième texte. Selon les auteurs, la première inférence donnerait à l'étudiant la possibilité de détecter des déficiences dans sa représentation du texte ce qui lui permettrait de prédire sa performance pour la deuxième inférence. Il reste que cette amélioration est modeste.

Glenberg et Epstein (1987) veulent voir s'il y a un lien entre l'expertise et la calibration. L'expert dans un domaine sait, en effet, qu'il est compétent dans son domaine et moins dans d'autres. Toutefois, cette capacité à prédire une performance relative selon le domaine ne dit rien sur cette capacité à l'intérieur d'un même domaine. Leurs sujets, 50 étudiants qui ont une formation en physique ou en musique (ont suivi au moins deux cours dans la matière en question), lisent 32 textes dans ces deux domaines en deux sessions. Chaque texte illustre en un paragraphe un principe qui est énoncé

clairement dans le corps du texte. À la suite de la lecture de chaque texte, l'étudiant doit donner son niveau de confiance (échelle en six points) quant à sa possibilité de reconnaître la validité d'une inférence découlant du principe énoncé dans le texte. Les chercheurs utilisent le Gamma (γ) de Goodman-Kruskal comme indice de calibration. Le γ obtenu par les étudiants en musique est égal à .06; celui obtenu par les étudiants en physique est de .02. La différence entre les deux indices n'est pas significative et la différence entre les deux indices et 0 ne l'est pas non plus. Cela indique donc une absence de calibration. Par ailleurs, ils obtiennent que les étudiants qui n'ont pas suivi de cours de physique démontrent une calibration relativement bonne par rapport aux textes de physique. Cela pourrait signifier que l'expertise dans un domaine est inversement reliée au réalisme dans ce domaine, mais qu'elle est bonne lorsque les sujets évaluent leur performance d'un domaine à l'autre. Les auteurs concluent que les jugements de confiance sont basés sur l'auto-classification comme experts ou non dans un domaine plutôt que sur l'évaluation du degré de compréhension du texte. Les chercheurs signalent cependant que la mise en place de leur étude pourrait avoir favorisé une telle stratégie d'auto-classification : les sujets sont choisis pour leur expertise et les textes sont clairement reliés à un domaine ou l'autre.

La deuxième étude de Pressley et al. (1987) est effectuée auprès de 57 étudiants (23 hommes et 34 femmes) qui ont déjà beaucoup d'expérience avec le livre utilisé et les questions d'examens tirées de ce livre. Les auteurs ne précisent pas si les étudiants avaient reçu les notes à ces examens, mais il est probable que ce soit le cas. Les étudiants doivent lire un chapitre d'un livre d'introduction en psychologie et répondre à 26 questions en remplissant des blancs. Ils doivent aussi faire une prédiction avant la lecture, après la lecture et après l'examen du nombre de questions qu'ils pensent pouvoir réussir. La différence absolue entre le rappel prédit et le rappel réel est utilisée pour calculer l'inexactitude de la prédiction. Cette dernière est de

6,95 avant la lecture, de 5.05 après la lecture et de 3.58 après l'examen. La seule différence significative se situe entre la première et la dernière situation.

Ghatala et al. (1989) se posent la question suivante : Combien de temps une personne pense avoir besoin pour apprendre une quantité fixe d'information? Dans une étude auprès de 70 élèves de 4^e année, ceux-ci doivent atteindre un taux de réussite de 100% dans leurs réponses à un examen et estimer le nombre d'items qu'ils pensent pouvoir réussir. Ils sont soumis à l'une de quatre conditions : (1) ils peuvent lire un passage autant de fois qu'ils le veulent avant l'examen, (2) ils doivent lire le passage une fois avant l'examen, ils peuvent requérir autant de cycles lectures/examens qu'ils le veulent pour arriver à une maîtrise de 100% et ils ne reçoivent pas de feedback des notes, (3) comme la deuxième condition, sauf qu'ils doivent dire combien ils pensent avoir réussi d'items avant de continuer, (4) comme dans la troisième condition, sauf qu'ils sont informés quant au nombre d'items réussis après chaque examen. Le score d'inexactitude du monitoring est la différence absolue entre le nombre d'items réussis perçus et le nombre d'items réussis en fait. Les résultats à ce score selon les conditions sont les suivants : 1.67; 1.86; 2.22; 0. Les comparaisons entre les trois premières conditions montrent qu'il n'y a pas de différence entre la première et la deuxième condition ni entre la deuxième condition et la troisième. L'absence d'inexactitude dans la dernière condition montre que les élèves ont bien encodé le feedback qui leur a été donné.

Plusieurs auteurs attribuent à l'ancienneté la différence dans l'exactitude à s'auto-évaluer. Une méta-analyse de Falchikov et Boud, (1989) a pour but de vérifier si cela est exact. À la suite d'une comparaison des étudiants de première année avec ceux qui sont là depuis trois ans ou plus au premier cycle et ceux des cycles supérieurs, il n'y a pas une tendance claire qui se dégage. Toutefois, une comparaison entre les étudiants des cours d'introduction et ceux des cours avancés dans les études qui identifient

clairement le niveau du cours montre une capacité plus grande des étudiants dans les cours avancés à s'auto-évaluer. Par ailleurs, il semble que les étudiants plus avancés qui prennent des cours d'introduction ne semblent pas s'auto-évaluer mieux que les étudiants moins avancés. Les auteurs concluent donc que l'expertise dans un domaine est plus importante que la durée de l'inscription à l'université en ce qui concerne l'exactitude de l'auto-évaluation.

Maki et Serra (1992b) demandent à 36 étudiants inscrits à un cours d'introduction à la psychologie de lire quatre textes et de répondre ensuite à trois questions à choix multiples qui servent d'examen de pratique. Pour la moitié des étudiants, un feedback quant à leur réussite de la question leur est donné à la suite de chaque question. Les autres étudiants ne reçoivent pas de feedback. Après les examens de pratique, les étudiants doivent mettre en rang les textes en commençant par celui où ils s'attendent de mieux réussir jusqu'à celui où ils s'attendent de moins bien réussir. À la suite de cela, les étudiants répondent aux trois questions de l'examen critère pour chaque texte. Enfin, ils mettent en rang les textes en fonction de leur performance perçue. Pour ce qui est des prédictions, les étudiants qui n'ont pas reçu de feedback prédisent leur performance plus exactement. Pour ce qui est de la postdiction, lorsqu'il y a feedback, $\gamma = .212$; lorsqu'il n'y a pas de feedback, $\gamma = .374$. Les valeurs de Gamma sont significativement différentes de zéro dans les deux cas, mais elles ne le sont pas l'une de l'autre. Les auteurs de la recherche concluent que de donner un feedback peut réduire l'exactitude.

Lundeberg, Fox et Punéochar (1994) étudient les différences selon le genre dans les jugements de confiance par rapport à des items spécifiques d'examens. Les 254 sujets suivent trois cours de psychologie au premier ou au deuxième cycle. Les questionnaires sont constitués de différents types de questions : questions à choix multiples, vrai/faux, statistiques, etc. Les sujets doivent indiquer leur niveau de confiance après chaque réponse sur une

échelle en cinq points (réponse devinée « 1 » à très sûr « 5 »). Le réalisme de la confiance est calculé grâce à l'indice CAQ. Les résultats montrent qu'il y a surconfiance autant chez les étudiantes que chez les étudiants aux deux niveaux. La confiance est de neuf à 13 points plus élevée que le pourcentage de questions correctes. Aux deux niveaux, la confiance des étudiantes est un peu plus élevée. La différence est significative au deuxième cycle. La confiance des étudiantes au deuxième cycle, lorsqu'elle est exacte, est significativement différente de celle des étudiants. Les étudiantes sont aussi plus confiantes pour les réponses manquées. Trois des quatre groupes discriminent bien les réponses correctes des réponses incorrectes. Les étudiants du premier cycle semblent trop confiants lorsqu'ils se trompent relativement à lorsqu'ils ont la bonne réponse.

Dans la recherche de Sjoström et Marks (1994) dont les détails sont présentés plus haut, l'exactitude de l'auto-évaluation demeure pratiquement constante pour les 12 unités évaluées. Malgré une corrélation de .45 pour l'unité 3 et de -.10 pour l'unité 11, l'interaction entre les unités et les évaluations standardisées donnent un $R^2 = .012$. Les auteurs de la recherche croient que cela pourrait peut-être s'expliquer par l'exactitude initiale de l'auto-évaluation.

Pour Radhakrishnan, Arrow et Sniezek (1996) le feedback sur la performance devrait être d'autant plus utile pour l'exactitude de la prédiction que les standards sont clairs et que le feedback est abondant. Ce serait encore mieux si la personne pouvait comparer sa performance à celle des autres. Ils prévoient une augmentation de l'exactitude, mais qui ne serait pas parfaite parce qu'il y aurait persistance de la surévaluation due à un rehaussement de soi qui entraînerait une sélection du feedback. Ils prévoient que le gain marginal d'exactitude découlant de la répétition et de l'information qui s'accumule devrait diminuer avec le temps. Le changement dans l'exactitude avec la répétition des essais est basé sur trois postulats : (1) la tâche est de difficulté moyenne pour la personne de sorte que la

performance peut s'améliorer avec le temps, (2) le feedback s'accumule avec le temps, (3) les facteurs motivationnels vont continuer à générer des effets pré-post. Ils mènent une étude auprès de 23 étudiants d'un cours avancé de psychologie organisationnelle. Ces étudiants doivent évaluer leur performance 5 jours avant et immédiatement après avoir pris l'examen en donnant un estimé de leur note attendue. Ils subissent ainsi trois examens consécutifs. La différence moyenne absolue entre la note prédite et la note obtenue sert à mesurer l'exactitude. L'optimisme excessif et le pessimisme excessif sont mesurés par la différence directionnelle entre la note obtenue et la note attendue. Les étudiants estiment aussi la difficulté de l'examen à l'aide d'une échelle en neuf points. Enfin, ils fournissent un estimé de la moyenne de la classe. Les résultats montrent que l'exactitude de la postdiction est meilleure pour le premier examen que celle de la prédiction (l'erreur étant d'environ 12% moins grande pour la postdiction), mais il n'y a pas de différence pour les deux autres examens. Il y a aussi une amélioration de l'exactitude d'un questionnaire à l'autre et un changement dans le pattern d'amélioration, le plus grand gain se faisant dans les premiers questionnaires. L'inexactitude passe de 25.3 à 9.0 pour ce qui est de la prédiction du premier au troisième examen. Pour la postdiction, elle passe de 12.4 à 8.8 pour ces examens.

Hacker et al. (2000) demandent à 99 étudiants du premier cycle inscrits à un cours d'introduction à la psychologie de prédire avant chaque examen le pourcentage d'items qu'ils pensent pouvoir réussir et de postdire après chaque examen le pourcentage d'items qu'ils pensent avoir réussis. Les étudiants doivent répondre à trois examens répartis sur le trimestre : le premier a 63 items, le deuxième 96 et le troisième 133. Les auteurs de la recherche s'assurent qu'il y a la même proportion d'items de différentes difficultés dans chaque examen. Après le troisième cours, les étudiants ont reçu un feedback sur leurs prédictions, leurs postdictions et leurs performances pour deux examens, mais aussi pour les trois examens de

pratiques. Les analyses de régression linéaire donnent des R^2 croissants pour les prédictions (.08, .19, .24), ce qui montre un accroissement d'exactitude avec chaque examen. Mais les postdictions demeurent consistantes (.25, .25, .27). Les postdictions, sont meilleures que les prédictions. Le R^2 égal pour la prédiction et la postdiction au dernier examen, montrerait que l'exactitude de prédiction atteint un plafond. La prédiction ne pourrait pas dépasser la postdiction parce qu'il y a plus d'information lors de la postdiction. Pour les trois examens, ils observent que les étudiants qui réussissent le mieux (note se situant entre 70% et 100%) démontrent une bonne exactitude à la fois de prédiction et de postdiction des notes. Les étudiants qui réussissent modérément bien (note se situant entre 50% et 69%) font preuve d'une exactitude de prédiction moyenne, mais d'une exactitude de postdiction bonne. Enfin, ceux qui réussissent le moins bien font preuve de surconfiance élevée autant pour la prédiction que pour la postdiction. Pour le premier examen, ils constatent que l'optimisme de la prédiction contribue à l'optimisme de la postdiction. Les étudiants faibles ne démontrent pas d'amélioration entre le premier examen et le dernier en ce qui concerne la prédiction et la postdiction. Les étudiants plus forts produisent une prédiction et une postdiction faibles au premier examen, mais font montre d'une amélioration au troisième examen et ce, surtout pour la postdiction (.46). Donc les résultats ne montrant pas d'amélioration lors de la postdiction en général sont dus aux étudiants faibles. Ils relèvent, par ailleurs, que le temps passé à étudier ne contribue ni à la prédiction ni à la postdiction. Ils font l'hypothèse que les étudiants ne savaient pas quoi étudier au juste et jusqu'à quel point étudier en termes de temps et d'effort.

Dans la deuxième étude de la recherche de Beyer (2002) dont la méthode a été présentée plus haut, l'ordinateur donne la bonne réponse après chacune des réponses du sujet. Les analyses de régression multiple montrent que le feedback accroît de façon significative l'exactitude pour les trois examens. Par ailleurs, il n'y a pas d'interaction avec le genre.

Les recherches rapportées dans cette section ont en commun d'étudier le réalisme de la confiance des étudiants à la suite de l'information qu'ils ont reçue à partir de leurs notes aux examens. Toutefois, ce groupe de recherches est très hétérogène à plusieurs points de vue. Certaines recherches rapportent des résultats à plusieurs étapes du feedback donné à l'étudiant, alors que ce n'est pas le cas pour d'autres. Toutefois, ce qui retient surtout l'attention, c'est que, selon les recherches, il n'y a pas d'évolution du réalisme (v.g. Shaughnessy, 1979), une augmentation du réalisme (Glenberg et Epstein, 1985; Weinstein, 1980) ou même, une diminution du réalisme (Maki et Serra, 1992b). Par ailleurs, ici, encore, l'objectif premier de plusieurs de ces études n'est pas de vérifier l'effet de l'augmentation de l'information sur le réalisme, mais de mettre ce dernier en relation avec des variables comme le genre, l'aptitude ou l'expertise.

Prédictions en fonction de la distance temporelle

Dans cette section, nous abordons la question du réalisme de la confiance en fonction de la distance temporelle qui sépare le jugement de l'étudiant de l'événement qui le concerne (examen, réussite d'un trimestre, avenir professionnel, etc.).

Nisan (1972) obtient de 100 étudiants universitaires inscrits à un cours en introduction à l'économie de répondre à un test de raisonnement de 12 items dont la difficulté est connue soit immédiatement ou quatre semaines plus tard. Ce test est présenté comme servant à mesurer la possibilité de réussite à l'université. On leur demande d'indiquer quelles sont leurs chances de bien répondre sur une échelle de 0 à 100. Une nouvelle mesure est effectuée où on leur dit qu'ils auront à subir une autre version du test dans quatre semaines (pour le groupe qui doit prendre le test immédiatement) ou le même jour (pour le groupe qui doit prendre le test dans l'avenir). Là aussi ils doivent indiquer leurs chances de bien répondre. Les résultats montrent une tendance consistante à une attente plus grande de réussite pour le test à

prendre dans quatre semaines. Cette recherche ne permet pas de dire si la personne perçoit le même changement d'attentes chez une autre personne quand deux personnes sont en compétition. Selon l'auteur de la recherche, la distance temporelle donnerait un plus grand sens de contrôle d'où une évaluation plus grande de succès.

Nisan, (1973) effectue la même étude avec 47 étudiants noirs d'un milieu socio-économique inférieur. Dans leur cas, non seulement les attentes n'augmentent pas avec le temps, mais elles diminuent. L'auteur de la recherche interprète cela en proposant que cette population aurait une plus faible perception de son aptitude à contrôler les renforcements externes.

Gilovich, Kerr et Medvec (1993) critiquent la recherche de Nisan (1972) en soulignant le manque de crédibilité venant du fait que les étudiants ne reçoivent pas de raisons suffisantes pour expliquer le délai de quatre semaines du test. Ils proposent une recherche plus générale. Elle comporte quatre études. Dans la première, ils demandent à des étudiants inscrits à des cours de psychologie (leur nombre n'est pas donné) d'exprimer leur confiance de réussite (en percentiles) lors du premier jour de classe et juste avant l'examen. Les résultats montrent que la confiance est plus faible le jour de l'examen (82% par opposition à 67% pour un cours et 79% par opposition à 75% pour l'autre cours). Il se peut, cependant que les étudiants aient appris des choses sur le cours entre le premier jour et l'examen ou encore, que les étudiants aient des attentes non réalistes à propos des efforts de préparation. Afin d'éviter ces problèmes, les chercheurs se livrent à une deuxième étude où ils disent aux étudiants qu'un professeur prépare des tâches (mémoire, détection de mensonges, anagrammes, temps de réponse, persuasion) à utiliser dans une recherche ultérieure. On leur demande quel serait leur degré de réussite dans ces tâches de trois façons différentes : score en percentiles, échelle en 10 points dont les pôles sont « performance très pauvre » et « très bien », prédiction du montant qu'ils gagneraient s'ils étaient payés pour chaque bonne réponse. Puis, plus tard, certains

étudiants, individuellement, sont amenés en laboratoire où on leur pose la même question que la première fois. Les chercheurs notent une chute de la confiance très significative pour huit des neuf comparaisons (trois mesures pour chacune des trois tâches).

Dans une quatrième étude, 40 étudiants dressent une liste de leurs pensées à propos de la réussite ou de l'échec en fonction de la distance temporelle de la tâche. Les 20 étudiants qui doivent se prononcer par rapport à une tâche distante se focalisent plus sur les causes de succès et moins sur les causes d'échecs possibles. L'anticipation d'un feedback imminent pourrait amener une personne à être plus auto-critique et donc avoir plus de pensées liées à l'échec. Les quatre études soutiennent l'idée d'une relation systématique entre la distance temporelle et la confiance. Selon les chercheurs, elle pourrait s'expliquer de nombreuses façons: par une évaluation irréaliste d'efforts préparatoires futurs, par un pessimisme défensif qui serait une façon de faire face plus efficacement aux difficultés, par une auto-critique de prévention. Une plus grande activation chez la personne qui s'apprête à effectuer la tâche peut aussi être interprétée comme étant un manque de confiance quant à la réussite.

Deux cent soixante étudiants en administration (167 hommes et 93 femmes) retenus par Hoch (1985) doivent prévoir leurs résultats de recherche d'emploi qui aura lieu 9 mois plus tard ; la probabilité que leur salaire de départ excédera $n\$$ par année, la probabilité qu'ils recevront leur première offre d'emploi avant tel jour de tel mois de telle année, la probabilité qu'ils reçoivent plus que n offres d'emploi à la fin de l'année scolaire. Des réponses sont proposées et la personne doit exprimer une probabilité (de 0 à 1) par rapport à sa réponse. Tous ces événements impliquent des résultats positifs. Les événements sont construits de façon à faire varier leur ligne de base. Certains sujets doivent produire des raisons qui favorisent ou défavorisent la réalisation de l'événement avant de faire leurs prédictions. Quand la ligne de base est basse ou moyenne, l'exactitude s'accroît si les

sujets ont à produire une raison défavorable. Toutefois, quand la ligne de base est élevée, les raisons défavorables n'ont pas d'effet sur la réponse. Cela suggère que les sujets ont généré automatiquement des raisons pour supporter leurs prédictions comme produit accompagnant le processus de réponse. Une analyse montre que les sujets attribuent les raisons favorables à des facteurs internes et les raisons défavorables à des facteurs externes. Les sujets générant des raisons favorables internes sont moins exacts que ceux qui génèrent des raisons favorables externes ou des raisons défavorables des deux types.

Vallone, Griffin, Lin et Ross (1990) s'adressent à 98 étudiants universitaires inscrits en première année pour qu'ils produisent 20 prédictions pour le trimestre d'automne et 21 prédictions pour l'année académique. Ces prédictions portent sur une grande variété de questions académiques, sociales et de loisir. Les étudiants ont deux possibilités de réponse mutuellement exclusives pour chaque item. Elles sont accompagnées d'une échelle de confiance de demi-étendue. Les items académiques ont trait au nombre d'heures d'étude par jour, à la possibilité de l'abandon d'un cours, à un changement quant au cours préféré, à un changement quant au cours le moins préféré, à la possibilité d'étudier plus tard que 3:00 du matin. Les attentes par rapport à l'avenir ont trait au changement de l'objectif de carrière et à la planification des études postdoctorales. Les étudiants sont inscrits en première année d'université. À la fin de l'année, ils répondent à un questionnaire pour vérifier l'exactitude de leurs prédictions. Les résultats montrent un écart important entre la confiance générale moyenne et l'exactitude moyenne (82,3% par opposition à 68,2%). Cela vaut autant pour les prédictions portant sur le trimestre que sur l'année. Pour 29 items (environ 70%) la surconfiance est au minimum de 10%. Pour huit items, elle atteint plus de 20%. Des 12 prédictions d'un seul trimestre, 11 manifestent de la surconfiance. Les items concernant les absences en classe et la décision d'abandonner un cours démontrent la plus

grande surconfiance. La relation confiance et exactitude montre un écart de 0 lorsque la confiance se situe entre 50% et 69%, de 15 entre 70% et 89%, de 20 entre 90% et 100% et de 23 à 100%. L'exactitude augmente avec le niveau de confiance, mais l'écart (dans la direction de la surconfiance) augmente aussi. Les corrélations bisérielles de point pour les sujets individuels entre la confiance et l'exactitude ne dépassent pas 0,24. Quand les prédictions vont dans le sens des lignes de base, il y a peu de surconfiance. Ce qui n'est pas le cas des autres prédictions. Notons que les sujets ne reçoivent pas d'information explicite sur les lignes de base. Les stratégies et les croyances par rapport aux traits de personnalité qui défient les lignes de base sont particulièrement coûteuses. Les auteurs de la recherche soutiennent qu'il est particulièrement difficile de prédire sa conduite future et sa performance dans une situation donnée, même quand on connaît ses objectifs personnels, ses aptitudes et ses dispositions de même que les détails de sa façon de réagir aux situations à cause de la difficulté d'anticiper les facteurs inconnus et parfois inconnaissables qui feront partie de la situation quand ils devront l'affronter. Citons à titre d'exemples de ces facteurs l'influence de ses pairs, les demandes des événements qui se développeront alors, les changements dans les ressources et les opportunités.

Dunning et Story (1991) obtiennent de 164 étudiants en psychologie du premier cycle qui souffrent d'une dépression légère ou non de formuler des prédictions à propos de comportements et d'événements qui pourraient se produire dans le cours d'un trimestre. Par la suite, on leur demande ce qui s'est réellement passé. Les chercheurs font remarquer que par leurs comportements les étudiants en question peuvent s'assurer que leurs prédictions soient le plus exactes possible. Les étudiants doivent prédire 37 événements ou résultats susceptibles de se passer dans le cours d'un trimestre et exprimer un niveau de confiance sur une échelle de 50% à 100%. Les items sont tirés pour la plupart de Vallone et al. (1990). Certains

de ces items sont de nature académique. Ils sont catégorisés par les chercheurs en fonction de leur désirabilité et de leur contrôlabilité. Ceux qui sont désirables et contrôlables sont les suivants : avoir un « A » dans son cours préféré, être sur la liste du doyen pour le trimestre, choisir ou changer de majeur, participer à une expérience de psychologie. Un item se réfère à un événement désirable, mais non contrôlable, il s'agit d'avoir un « A » dans le cours présent. Ceux qui sont non désirables et contrôlables sont les suivants : voir un conseiller à cause de problèmes académiques, ne pas assister à sa classe la plus importante plus de 3 fois, étudier plus tard que 4 :00 AM à au moins une occasion. Ceux qui sont non désirables et non contrôlables sont les suivants : argumenter avec un professeur à propos d'une note, remettre en question une décision d'avoir choisi l'université où l'étudiant est inscrit, abandonner un cours après 5 semaines, manquer plus de deux jours de classe à cause de maladie, avoir une « déprime » académique pour plus de 2 semaines. Les résultats montrent qu'il n'y pas de relation dans les réponses avec le genre, le rang en classe et l'âge. Les étudiants démontrent une forte tendance à la surconfiance, la confiance moyenne étant de 83,5% alors que la réussite est de 76,9%.

Quatre-vingt-dix-huit étudiants (29 hommes et 69 femmes) du premier cycle en psychologie, doivent répondre à 87 items choisis en fonction de leur ligne de base et de leur dénouement positif ou négatif dans les domaines académique, social et personnel (Pulford et Colman, 1996). Les étudiants doivent prédire s'ils vivront l'événement durant la semaine qui suit et exprimer leur niveau de confiance sur une échelle de 50% à 100%. Une semaine plus tard, un questionnaire leur est présenté où ils doivent dire si ce qui avait été prédit s'est produit. Pour chaque item, les chercheurs calculent le pourcentage moyen de confiance, le pourcentage moyen d'exactitude et ils soustraient le premier du second pour évaluer le réalisme de la confiance. Pour l'ensemble des items, ils obtiennent une différence significative entre la confiance moyenne (84.42) et l'exactitude moyenne (82.10). Les résultats

indiquent une forte surconfiance et qui est plus grande pour les résultats positifs. Par ailleurs, les chercheurs constatent une relation en « U » inversé avec la ligne de base. La ligne de base étant la fréquence relative avec laquelle les membres de la population visée vivent l'événement en question sur une période de temps donnée. La confiance est plus grande aux niveaux intermédiaires de cette ligne de base alors qu'il y a sous-confiance aux deux extrêmes. Les auteurs rapprochent leurs données du phénomène de l'optimisme irréaliste à l'égard des événements heureux tel qu'étudié par (Weinstein, 1980). Cette surconfiance fait aussi penser à l'effet Pollyanna (héroïne du roman d'Eleanor Porter, 1913) qui consiste à exagérer les aspects positifs de son expérience et à minimiser les aspects négatifs (Boucher et Osgood, 1969). Dans la présente recherche, l'échelle de temps considérée est plus courte (une semaine) que dans les études sur l'optimisme irréaliste de Weinstein (toute la vie). Selon les auteurs de la recherche, il se peut que l'irréalisme de l'optimisme soit plus probable pour les prédictions impliquant une échelle de temps plus étendue à cause de problèmes cognitifs. Selon eux, il se peut aussi que l'irréalisme de l'optimisme concernant des événements positifs soit dû à des facteurs motivationnels.

Selon la *Construal Level Theory* (Lieberman, Sagristano et Trope, 2002; Nussbaum, Trope et Liberman, 2003; Sagristano, Trope et Liberman, 2002), les décisions à propos des événements futurs dépendent souvent de la distance qui nous en sépare. La distance temporelle changerait la réponse des gens par rapport aux événements futurs en changeant leur façon de construire ces événements. La construction mentale implique l'abstraction, et la distance temporelle est un des facteurs qui détermine le niveau d'abstraction. En construisant l'avenir lointain, les gens sont plus portés à utiliser des stéréotypes, des scripts généralisés plutôt que les détails plus concrets et non schématiques. La prédiction de son propre comportement se base sur une représentation abstraite et schématique qui néglige les facteurs

contextuels. Les auteurs citent comme exemple l'erreur de planification (*planning fallacy*) (Buehler, Griffin et Ross, 1994). La tendance à sous-estimer le temps pour compléter une tâche serait due à une représentation simplifiée des tâches futures. À une représentation qui ne tient pas compte de toute la complexité de ces tâches.

La *Construal Level Theory* propose que les gens construisent des représentations plus abstraites des événements futurs distants que des événements prochains. Les constructions d'un niveau élevé sont des représentations schématiques, hors contexte, qui tirent l'essentiel de l'information disponible. Les représentations d'un bas niveau sont plus concrètes, subordonnées, contextuelles et incidentes. Elles sont plus riches, mais moins structurées que les représentations de haut niveau. Des sujets (Liberman et Trope, 1998) décrivent des activités futures distantes en termes de descriptions superordonnées plus souvent que lorsque l'avenir considéré est proche. Par exemple, déménager dans un nouvel appartement, à un niveau élevé de représentation, peut être perçu comme étant l'occasion de commencer une nouvelle vie, alors qu'à un niveau de représentation plus concret, comme étant le besoin d'emballer ses effets.

Les actions visant un but (par exemple, étudier pour un examen) forment des hiérarchies. Les personnes font des prédictions sur leurs performances futures distantes qui sont plus influencées par leur compétence perçue (représentation abstraite) que par l'information concernant l'examen (un questionnaire à choix multiples ou à questions ouvertes). Ce serait l'inverse pour les prédictions à court terme.

La représentation d'une action à un niveau élevé ou bas aurait une incidence sur l'accent qui est mis sur les considérations concernant la désirabilité ou la faisabilité de cette action. La désirabilité concerne la valeur du résultat de l'action (le pourquoi), alors que la faisabilité se rapporte à la facilité ou à la difficulté d'accomplir l'action (le comment). Par exemple, en ce

qui concerne un cours, à un niveau élevé de la représentation, ce qui serait dominant serait l'intérêt de ce cours, alors qu'à un niveau de représentation plus bas, ce qui dominerait serait la difficulté de ce cours. Selon la *Construal Level Theory*, la désirabilité fait partie de la représentation à un niveau élevée. Elle influence plus les décisions dans le futur distant. En revanche, la faisabilité appartient plus à la représentation à un bas niveau et est moins importante pour le futur distant. Dans une étude (Liberman et Trope, 1998), les chercheurs demandent aux sujets d'indiquer leur intérêt pour assister à une conférence dans l'avenir proche ou lointain. La conférence est décrite comme étant intéressante ou non et comme donnée à un moment commode ou non. Ils trouvent que l'effet du niveau d'intérêt (désirabilité) sur la préférence s'accroît avec la distance temporelle, alors que l'effet de la commodité du moment (faisabilité), décroît avec la distance temporelle.

Les recherches rapportées ci-dessus ne mesurent pas toutes le réalisme de la confiance. Certaines se contentent de montrer l'évolution des attentes ou de la confiance en fonction de la distance temporelle. Par ailleurs, certaines font ressortir la difficulté de séparer l'effet de l'augmentation du niveau d'information de la distance temporelle. Les recherches faisant la comparaison du réalisme de la confiance à divers moments sont rares. Plusieurs de ces recherches s'intéressent avant tout à des facteurs explicatifs du niveau de réalisme de la confiance autres que la distance temporelle, cette dernière étant un prétexte pour étudier ces facteurs. La grandeur de la distance temporelle peut être très variable d'une recherche à l'autre. Il ressort surtout de ces recherches que, plus la distance temporelle est grande, plus la surconfiance est grande. La *Construal Level Theory* est l'approche qui donne l'explication la plus exhaustive de ce phénomène.

HYPOTHÈSES ET QUESTIONS

L'objectif de la présente recherche est donc de vérifier des hypothèses se rapportant aux prédictions académiques et professionnelles d'étudiants en

fonction de l'expérience et de la distance temporelle. Pour y arriver, nous avons recours à différents moyens de mesurer le réalisme des étudiants. Nous utilisons à la fois une approche transversale et une approche longitudinale et ce, auprès de grands groupes d'étudiants. Nous tenons compte de variables comme le genre et la performance. Les recherches rapportées ci-dessus n'ont abordé qu'une partie de cette problématique à la fois, et souvent pour répondre à des questions qui ne sont pas les nôtres.

Nos hypothèses se répartissent selon qu'elles concernent la calibration des probabilités psychologiques, le niveau de réalisme, l'évolution du niveau de confiance des prédictions, l'évolution du niveau de détermination et la distance temporelle entre le jugement et l'événement. La calibration des probabilités psychologiques, elle-même, est abordée de trois façons différentes : la calibration absolue, la calibration de la certitude et la calibration relative. La calibration absolue est celle que nous obtenons par le calcul de l'indice de Brier, alors que la calibration de la certitude se limite à la calibration mesurée pour le niveau de confiance « 10 ». Enfin, la calibration relative est déterminée par l'inspection visuelle des courbes de calibration obtenues comparées à la diagonale de la calibration parfaite.

Nos hypothèses et nos questions sont les suivantes :

La calibration des probabilités psychologiques est meilleure à un examen final qu'à un examen intra-trimestriel et ce, autant pour la calibration absolue que pour la calibration de la certitude.

La calibration des étudiantes est meilleure que celle des étudiants et ce, autant pour la calibration absolue que pour la calibration de la certitude.

Y a-t-il un lien entre la perception de la représentativité et de la difficulté de l'examen et la calibration autant de nature absolue que de la certitude ?

Il y a un lien direct entre la calibration autant de nature absolue que de la certitude et la note obtenue.

Les étudiant-e-s font preuve de calibration relative.

La calibration relative est meilleure pour l'examen final que pour l'examen intra-trimestriel.

La différence entre la note attendue et la note obtenue (réalisme de Type I) est plus grande pour l'examen intra-trimestriel que pour l'examen final.

Nous faisons, pour le réalisme de Type I, les mêmes hypothèses et nous nous posons les mêmes questions que pour la calibration, autant de nature absolue que de la certitude, en ce qui concerne l'effet de l'ordre des examens, du genre, de la perception de l'examen et de la note obtenue.

La relation entre le niveau de confiance de la prédiction d'un événement et sa réalisation (réalisme de Type II) est en relation directe avec l'année d'inscription des étudiant-e-s.

Il y a une baisse du niveau de confiance des prédictions chez les mêmes étudiant-e-s avec le passage du temps à trois moments différents sur une période de 12 mois.

Y a-t-il une évolution du niveau de détermination chez les mêmes étudiant-e-s avec le passage du temps à trois moments différents sur une période de 12 mois ?

Les étudiant-e-s sont moins confiants par rapport aux événements académiques proximaux que distants.

CHAPITRE II : DESCRIPTION DE LA RECHERCHE

SUJETS

Les sujets de cette recherche ont été recrutés au sein d'une population d'étudiant-e-s inscrits à dix cours de psychologie au département de psychologie de l'Université de Montréal en 1995, 1996 et 1998. Pour plus de clarté, nous présenterons leurs caractéristiques dans la section « déroulement de la recherche » étant donné qu'ils n'ont pas tous eu à répondre aux mêmes questionnaires.

QUESTIONNAIRES

Nous avons posé quatre grands types de questions aux étudiant-e-s : a) des questions servant à les identifier, b) des questions se rapportant à un examen, c) des questions se rapportant à la vie académique et d) des questions se rapportant à la vie professionnelle. Les trois questionnaires utilisés se retrouvent en Appendice A.

Les cinq questions portant sur l'identification des étudiant-e-s visaient à connaître le collège où ils avaient étudié, leur code permanent à l'université, le département où ils étaient inscrits, leur programme d'étude (mineur, majeur ou baccalauréat) et leur année d'étude (1^{ère}, 2^è ou 3^è).

À la suite d'un examen sous forme d'un questionnaire à choix multiples de 50 questions, les étudiant-e-s devaient répondre à quatre questions s'y rapportant et concernant : a) la difficulté de l'examen sur une échelle en 11 points (0 = facile à 10 = difficile), b) la représentativité de l'examen aussi sur une échelle en 11 points (0 = pas représentatif à 10 = représentatif), c) la note attendue exprimée sous forme d'une lettre (A+, A, A-, etc.) et d) la note souhaitée exprimée de la même façon.

Nous avons aussi demandé à certains étudiant-e-s d'indiquer leur niveau de confiance quant à l'exactitude de leur réponse pour chacune des 50 questions (0 = pas du tout sûr à 10 = tout à fait sûr). Ces étudiant-e-s

recevaient l'assurance que le niveau de confiance exprimée n'affecterait en rien la correction de l'examen.

Les étudiant-e-s ont répondu à 59 questions se rapportant à la vie académique se situant le long d'un horizon temporel plus ou moins éloigné. Ces questions avaient pour but de recueillir quatre types de données touchant : a) leur niveau de confiance exprimé sur une échelle en six points (de 50% à 100%) à l'égard de 22 prédictions (v.g. « Pensez-vous obtenir un « A » dans le cours que vous suivez en ce moment? »), b) leur niveau de détermination exprimé sur la même échelle pour 13 questions (v.g. « Tenez-vous à obtenir un « A » dans le cours que vous suivez en ce moment? »), c) leurs réalisations sous la forme de 18 questions (v.g. « Avez-vous obtenu un « A » dans le cours que vous suivez en ce moment? ») et d) leurs connaissances pour six questions (v.g. la question « Combien d'étudiants sont admis au baccalauréat à chaque année? » suivi d'un choix de réponses).

Les étudiant-e-s ont aussi répondu à neuf questions se rapportant à la vie professionnelle. Ces questions avaient pour but de recueillir trois types de données, leurs prédictions (trois questions), leur détermination (trois questions) et leurs connaissances (trois questions).

Matériellement, l'ensemble de ces questions était regroupé dans trois questionnaires : a) le questionnaire de prédictions, détermination et connaissances académiques et professionnelles (**QPDCAP**) contenant 55 questions, b) le questionnaire de la perception de l'examen contenant quatre questions (**QPE**) et c) le questionnaire des réalisations académiques (**QRA**) contenant pour sa part 18 questions.

DÉROULEMENT

Le déroulement de la recherche varie selon le groupe auquel sont inscrits les étudiant-e-s puisqu'ils n'ont pas tous eu à répondre aux mêmes questions. Les étudiant-e-s inscrits au cours de psychologie sociale à l'automne 1995 (*n*

= 135, 106 étudiantes et 29 étudiants) ont répondu au **QPDCAP** lors de la première rencontre. Une partie de ces étudiant-e-s ($n = 99$, c'est-à-dire 78 étudiantes et 21 étudiants) ont accepté d'indiquer leur niveau de confiance quant à l'exactitude de leur réponse pour chacune des 50 questions d'un questionnaire à choix multiples (trois choix par réponse), et ce pour l'examen intra et final. Ces mêmes étudiant-e-s ont répondu au **QPE**.

Pour leur part, des étudiant-e-s inscrits au cours de psychologie sociale à l'automne 1996 ($n = 136$, 117 étudiantes et 19 étudiants) et à l'automne 1998 ($n = 124$, 101 étudiantes et 23 étudiants) ont répondu au **QPDCAP** lors de la première rencontre. Lors des examens intra et final, une partie de ces étudiant-e-s, ont aussi répondu au **QPE**. À la fin du cours, ils ont aussi répondu au **QRA**.

Un quatrième groupe, constitué d'étudiant-e-s recruté-e-s à l'intérieur de trois cours différents de psychologie à l'automne 1996 ($n = 187$, 162 étudiantes et 25 étudiants) et un cinquième groupe, constitué d'étudiant-e-s recruté-e-s aussi à l'intérieur de trois cours différents de psychologie à l'automne 1998 ($n = 231$, 197 étudiantes et 34 étudiants) ont répondu au **QPDCAP** et au **QRA**.

ANALYSE DES RÉSULTATS

L'analyse des résultats implique l'utilisation de différentes méthodes dépendamment des variables concernées.

Calibration

Calibration absolue

Pour étudier la calibration absolue, nous avons eu recours à l'indice de Brier. Notons cependant que lorsqu'il y a deux options de réponse, on utilise une demi-échelle de probabilité, car s'il y avait une échelle pleine, cela reviendrait à dire que lorsque le sujet est sûr de la réponse « A » à 4 (échelle 0 à 10), il est sûr de la réponse « B » à 6 puisque les probabilités sont

complémentaires. Dans le cas de la présente recherche, l'étudiant a trois choix de réponses. L'attribution d'une probabilité à une option ne permet pas d'obtenir par défaut la probabilité de l'autre option. Ainsi, si quelqu'un attribue la probabilité de 6 à l'option « C », cela ne permet pas de conclure autre chose que la probabilité pour « A » et « B » ensemble est de 4. Pour connaître la probabilité psychologique de « A » et « B », il faudrait demander aux étudiant-e-s d'attribuer une probabilité à chacune des options et non seulement à la réponse retenue. Ainsi, un étudiant qui ne connaîtrait pas la réponse à une question devrait alors attribuer 3.33... à chacune des réponses, aucune ne lui apparaissant plus probable qu'une autre. (Dans le cas de la possibilité de deux réponses, la probabilité équivalente est de .5). Il nous est apparu qu'il serait trop fastidieux pour l'étudiant d'indiquer un niveau de confiance pour chacune des réponses possibles et que cela pourrait peut-être même nuire à sa performance à l'examen. Nous leur avons donc demandé d'indiquer la probabilité que pour la réponse retenue. Par ailleurs, nous avons utilisé une échelle en 11 points (0 à 10) pour exprimer le niveau de confiance, de sorte que l'étudiant pouvait exprimer un niveau de confiance nul pour la réponse choisie alors qu'il savait pertinemment que l'une des trois réponses était exacte étant donné la nature de la tâche.

Dans ces conditions, il va de soi que notre utilisation de l'indice de Brier est non orthodoxe et que les valeurs que nous avons obtenues ne peuvent en aucun cas être comparées à des valeurs semblables obtenues dans des recherches qui respectent plus strictement les conditions de l'utilisation de cet indice. Ceci dit, nous utilisons cet indice avec nos données pour établir des comparaisons entre les différents niveaux des variables. De la sorte, nous nous trouverons à comparer des valeurs comparables. C'est ainsi que nous utiliserons l'analyse de variance pour groupes indépendants pour voir si nous obtenons des effets des variables ordre de présentation des examens et genre. Les autres analyses portant sur la calibration absolue

(relation avec la perception de l'examen et la note obtenue) reposeront sur le coefficient de corrélation de Pearson r .

Calibration relative

La calibration relative est mise en évidence par l'utilisation de courbes tracées à partir de la relation entre la proportion de réponses exactes pour le niveau de confiance exprimé. Il y a calibration relative lorsque la proportion de réponses exactes est reliée, même de façon imparfaite, avec le niveau de confiance exprimé. L'analyse des statistiques descriptives de la fréquence d'utilisation des différents niveaux de confiance nous donnera aussi un aperçu de la compréhension par les étudiant-e-s de cette échelle.

Calibration de la certitude

À défaut d'avoir une mesure précise de la calibration absolue par l'indice de Brier étant donné ce qui a été dit plus haut, nous avons recouru à un indice de la calibration de la certitude. Il consiste à calculer la proportion de réponses exactes pour le niveau de confiance « 10 ». En effet, nous pouvons conclure qu'une proportion de réponses exactes pour ce niveau qui n'est pas égale à 1 indique nécessairement un défaut de calibration puisque l'étudiant affirme être tout à fait sûr de ces réponses. Nous appliquons les mêmes analyses à cet indice que celles que nous avons retenues pour la calibration absolue.

Niveau de réalisme

Le niveau de réalisme a été établi par la différence entre la note attendue et la note obtenue à la suite de la transformation de la note littérale en un nombre. Plus cette différence est grande (quelle qu'en soit la direction), moins le réalisme est bon. Il y a surconfiance quand la différence est positive et sous-confiance quand elle est négative.

En plus des analyses semblables à celles que nous avons utilisées pour la calibration absolue et la calibration de la certitude, nous avons calculé des

coefficients de corrélation bisérielle de point r_{pb} pour établir le réalisme académique en fonction de l'année d'inscription et de la distance temporelle des événements, d'une part, et nous avons procédé à des analyses descriptives des réalisations en fonction du niveau de confiance, d'autre part.

Évolution de la confiance des prédictions et de la détermination

L'étude de l'évolution de la confiance des prédictions et de la détermination a été effectuée auprès d'un groupe de 37 étudiant-e-s à partir d'analyses de variance pour mesures répétées des réponses données à trois moments différents : Septembre 1995, décembre 1995, septembre 1996.

Niveau de confiance selon la distance temporelle

Le même groupe de 37 étudiant-e-s a servi pour mesurer le réalisme académique selon la distance temporelle. Pour analyser la relation entre la distance temporelle et le niveau de confiance de prédictions d'événements souhaitables, nous avons situé les réponses des étudiant-e-s le long d'un continuum de distance temporelle à neuf niveaux : (1) les événements du trimestre actuel (huit items), (2) ceux du trimestre suivant (deux items), (3) de l'année suivante, (4) obtenir le baccalauréat, (5) être accepté à la maîtrise, (6) obtenir la maîtrise, (7) être accepté au doctorat, (8) obtenir le doctorat, (9) obtenir un emploi en psychologie. Cela nous permet de mesurer le niveau de confiance à la fois en fonction de la distance temporelle des événements entre eux (par exemple, entre obtenir le baccalauréat et obtenir un emploi en psychologie) et de l'évolution de la distance temporelle des mêmes événements puisque les mêmes étudiant-e-s répondent au questionnaire à trois moments différents. Dans ce dernier cas, cela s'applique aux six derniers événements puisque la distance temporelle ne varie pas selon le moment où a été prise la mesure pour les trois premiers (c'est-à-dire que, par exemple, le « trimestre suivant » n'est pas le même trimestre selon le moment où a été donnée la réponse).

CHAPITRE III : RÉSULTATS

Les résultats comportent quatre parties se rapportant respectivement à la calibration, au réalisme, à l'évolution du niveau de confiance des prédictions et à l'évolution du niveau de détermination.

CALIBRATION

La partie sur la calibration se subdivise en trois : l'analyse des résultats provenant des données sur l'indice de calibration absolue, celle se rapportant à la calibration de la certitude et celle que l'on obtient à partir de la représentation graphique de la calibration, qui permet de se faire une idée à propos de la calibration relative des étudiant-e-s.

Calibration absolue

Dans cette partie à propos de la calibration absolue, nous allons d'abord voir la question de la relation de la calibration avec l'ordre de présentation des examens et le genre. Nous allons ensuite tenter de voir s'il y a une relation entre la calibration absolue et la perception de l'examen. Enfin, nous verrons s'il y a une relation entre la calibration absolue de l'étudiant et sa note.

Calibration selon l'ordre des examens et le genre

Le Tableau 1 présente les moyennes et les écarts-types selon l'ordre de présentation de l'examen et le genre.

Tableau 1

Moyennes et écart-types de la calibration selon l'examen et le genre

Examen	Étudiantes		Étudiants	
	M	ET	M	ET
Intra	0.195	0.078	0.141	0.047
Final	0.184	0.098	0.135	0.055

Le Tableau 2 rapporte les résultats de l'anova quant à la calibration des 99 étudiant-e-s qui ont répondu en fonction de l'ordre de présentation de l'examen et du genre. Il n'y a pas d'effet d'interaction entre l'ordre de l'examen et le genre des répondants ($F(1, 97) = 0.03, p = .86$). Il n'y a pas, non plus, de différence dans la calibration selon qu'il s'agisse de l'examen intra ou final ($F(1, 97) = 0.48, p = .49$). En revanche, il y a un effet du genre ($F(1, 97) = 7.77, p = .01$). Les étudiants, en effet, démontrent une calibration légèrement supérieure à celle des étudiantes (0.14 par opposition à 0.19).

Tableau 2

Calibration, anova selon l'ordre des examens et le genre

Source	df	Carrés moyens	F	p
Intrasujets				
Examen	1	6.26 ⁻⁴	0.48	0.493
Examen X Genre	1	4.09 ⁻⁵	0.03	0.861
Erreur (Examen)	97	1.32 ⁻³		
Intersujets				
Genre	1	8.59 ⁻²	7.77	0.006
Erreur	97	1.11 ⁻²		

Calibration et perception de l'examen

Dans cette partie, nous rapportons les résultats quant à la relation entre la calibration absolue et la perception de l'examen pour l'examen intra et pour l'examen final. La corrélation de Pearson obtenue entre la calibration à l'examen intra et la difficulté perçue de l'examen est $r = .05$ ($n = 96$), celle avec la représentativité perçue est $r = -.10$ ($n = 97$). Ces corrélations ne sont pas statistiquement significatives.

La corrélation de Pearson obtenue entre la calibration à l'examen final et la difficulté perçue de l'examen est $r = -.05$ ($n = 97$), celle avec la représentativité perçue est $r = -.04$ ($n = 97$). Ces corrélations ne sont pas statistiquement significatives.

Calibration et note obtenue

La corrélation de Pearson entre la calibration et la note obtenue est significative pour l'examen intra ($r = .28$, $p < .01$: $n = 96$), mais pas pour l'examen final ($r = 0$; $n = 97$).

Calibration de la certitude

Les données à propos de la calibration de la certitude sont présentées dans le même ordre que celles de la section sur la calibration absolue.

Calibration de la certitude selon l'ordre des examens et le genre

Le Tableau 3 présente les moyennes et les écarts-types selon l'ordre de présentation de l'examen et le genre.

Tableau 3

Moyennes et écart-types de la calibration de la certitude selon l'examen et le genre

Examen	Étudiantes		Étudiants	
	<i>M</i>	<i>ET</i>	<i>M</i>	<i>ET</i>
Intra	0.262	0.184	0.327	0.225
Final	0.159	0.204	0.260	0.279

Le Tableau 4 rapporte les résultats de l'anova quant à la calibration pour le niveau de confiance «10» des étudiant-e-s en fonction de l'ordre de l'examen et du genre. Ils montrent qu'il n'y a pas d'effet d'interaction entre l'ordre de l'examen et le genre des répondants ($F(1, 97) = 0.898, p = .346$). Il y a, en revanche, une différence significative selon qu'il s'agisse de l'examen intra ou final ($F(1, 97) = 19.31, p < .001$). Il n'y a pas d'effet du genre ($F(1, 97) = 3.048, p = .084$). La calibration pour l'examen final ($M = 0.180$) est meilleure que pour l'examen intra ($M = 0.276$).

Tableau 4
Calibration de la certitude, anova selon l'examen et le genre

Source	dl	Carrés moyens	F	p
Intrasujets				
Examen	1	0.239	19.31	< .001
Examen X Genre	1	1.12 ²	0.898	0.346
Erreur (Examen)	97	1.24 ²		
Intersujets				
Genre	1	0.226	3.048	0.084
Erreur	97	7.423 ²		

Calibration de la certitude et perception de l'examen

Dans cette partie, nous rapportons les résultats quant à la relation entre la calibration de la certitude et la perception de l'examen pour l'examen intra et pour l'examen final. La relation entre la calibration pour le niveau de confiance «10» et la perception de la difficulté de l'examen est négative et significative pour l'examen intra ($r = -.26, p < .001; n = 99$), mais il n'y a pas de relation entre ces variables pour l'examen final. Par ailleurs, il y a une relation positive significative entre la calibration pour le niveau de confiance «10» et la perception de la représentativité de l'examen pour l'examen intra ($r = .25, p < .01; n = 99$), mais pas pour l'examen final.

Calibration de la certitude et note obtenue

Il y a une relation positive significative entre la calibration pour le niveau de confiance «10» et la note obtenue à l'examen intra ($r = .55, p < .01; n = 98$), mais pas pour l'examen final.

Calibration relative

La Figure 2 rapporte la proportion de réponses exactes en fonction du niveau de confiance pour l'examen intra et l'examen final. La diagonale représente la calibration parfaite. Cette figure permet de constater qu'il y a surconfiance pour l'examen intra du niveau 6 au niveau 10, alors qu'il y a sous-confiance à tous les niveaux pour l'examen final, sauf pour le niveau 10. Ces données révèlent que même si la calibration absolue est faible, il y a une calibration relative : plus le niveau de confiance est élevé, plus la proportion de réponses exactes est grande. Ici, encore, l'exception la plus notable est lors de l'utilisation du niveau de confiance 0 pour l'examen final.

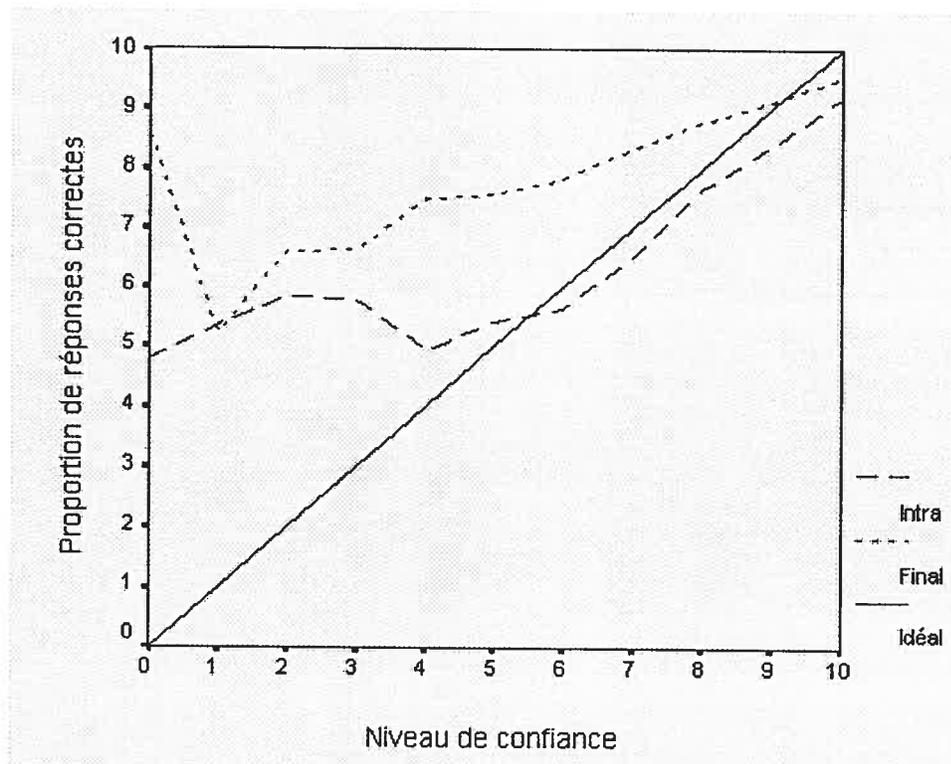


Figure 2. Représentation graphique de la calibration pour l'examen intra et l'examen final.

Le Tableau 5 ajoute aux informations de la figure la fréquence d'utilisation de chacun des niveaux de confiance de même que l'écart précis entre la proportion de réponses exactes et le niveau de confiance pour les examens intra et final.

Tableau 5

Proportion de réponses exactes en fonction des niveaux de confiance pour l'examen intra et l'examen final

Confiance	Intra			Final		
	PRE	Différence	Fréquence	PRE	Différence	Fréquence
10	9.17	0.83	1490	9.54	0.46	1527
9	8.36	0.64	432	9.10	-0.10	612
8	7.66	0.34	578	8.78	-0.78	728
7	6.50	0.50	500	8.32	-1.32	583
6	5.65	0.35	457	7.83	-1.83	419
5	5.41	-0.41	558	7.55	-2.55	457
4	4.94	-0.94	176	7.48	-3.48	131
3	5.79	-2.79	183	6.62	-3.62	157

Note. PRE : proportion de réponses exactes.

Tableau 5 (suite)

Proportion de réponses exactes en fonction des niveaux de confiance pour l'examen intra et l'examen final

2	5.85	-3.85	171	6.61	-4.61	109
1	5.36	-4.36	110	5.23	-4.23	65
0	4.80	-4.80	246	8.58	-8.58	141

Note. PRE : proportion de réponses exactes.

Notons que l'utilisation des niveaux de confiance se situant entre 5 et 10 constitue respectivement 82% et 88% de l'ensemble des utilisations pour l'examen intra et pour l'examen final.

La Figure 3 donne une illustration graphique de la répartition des fréquences d'utilisation des niveaux de confiance.

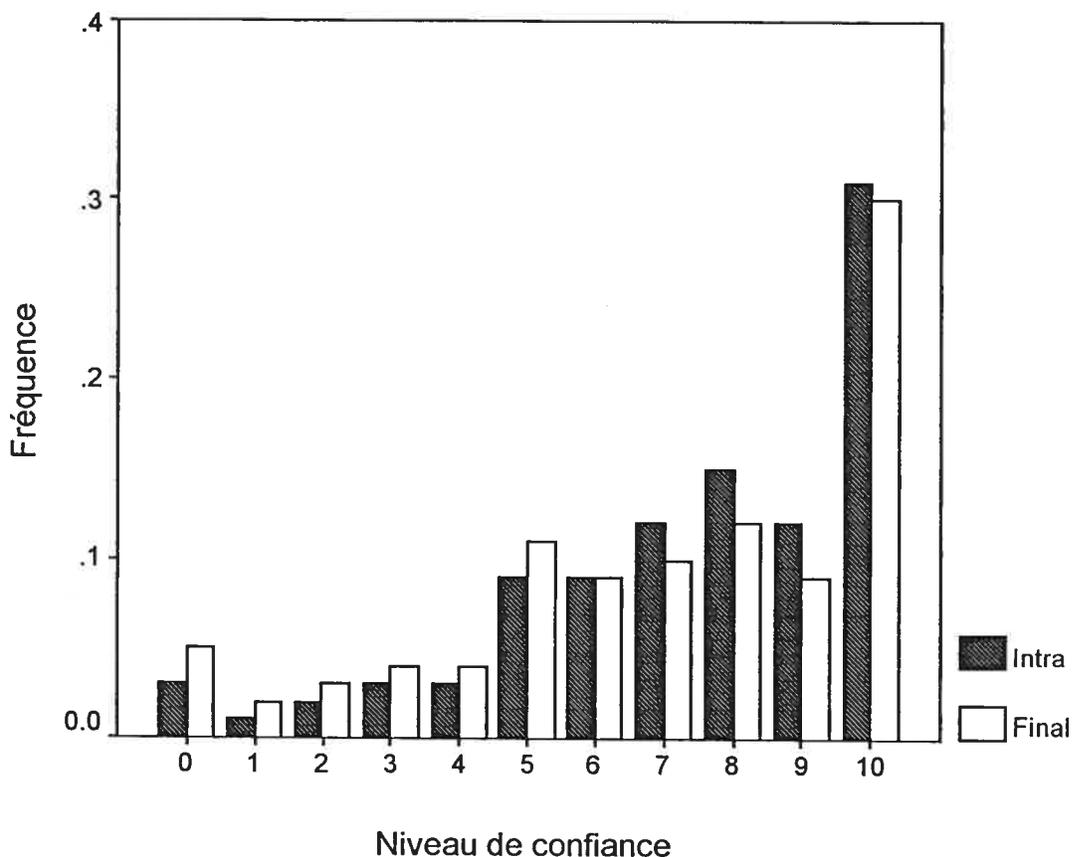


Figure 3. Fréquences d'utilisation des niveaux de confiance pour l'examen intra et l'examen final.

RÉALISME

Cette partie présente les résultats par rapport au réalisme. Elle contient les mêmes subdivisions que les parties qui précèdent : réalisme selon l'ordre des examens et le genre, en fonction de la perception de l'examen et de la note obtenue. En plus, elle aborde la question du réalisme académique selon l'année d'inscription et les réalisations en fonction du niveau de confiance.

Réalisme selon l'ordre des examens et le genre

Le Tableau 6 présente les moyennes et les écarts-types du réalisme selon l'ordre de présentation de l'examen et le genre.

Tableau 6

Moyennes et écart-types du réalisme selon l'examen et le genre

Examen	Étudiantes		Étudiants	
	<i>M</i>	<i>ET</i>	<i>M</i>	<i>ET</i>
Intra	0.31	2.375	-0.15	2.538
Final	-0.11	2.228	-0.46	2.388

Les résultats de l'anova en fonction de l'examen et du genre ne montrent aucun effet de ces variables sur le réalisme (Tableau 7).

Tableau 7
Réalisme, anova selon l'examen et le genre

Source	<i>dl</i>	Carrés moyens	<i>F</i>	<i>p</i>
Intrasujets				
Examen	1	12.609	2.90	0.089
Examen X Genre	1	0.291	0.07	0.796
Erreur (Examen)	337	4.342		
Intersujets				
Genre	1	15.993	2.45	0.118
Erreur	337	6.526		

Réalisme et perception de l'examen

Les corrélations entre le réalisme et la perception de l'examen sont faibles autant pour l'examen intra que pour l'examen final. Elles ne dépassent pas $r = -.14$ ($p < .01$; $n = 345$), cette dernière corrélation concerne la relation entre le réalisme et la représentativité perçue de l'examen final. La relation semblable en ce qui concerne l'examen intra n'est pas significative. Les autres corrélations significatives ont trait à la relation avec la difficulté de l'examen, $r = -.12$ ($p < .05$; $n = 340$) et $r = -.13$ ($p < .05$; $n = 344$) respectivement pour l'examen intra et l'examen final.

Réalisme et note obtenue

Les corrélations de Pearson entre le réalisme et la note obtenue sont significatives autant pour l'examen intra que pour l'examen final, $r = -.67$, ($p < .01$; $n = 346$) et $r = -.75$, ($p < .01$; $n = 348$) respectivement.

Réalisme académique selon l'année d'inscription

Les résultats présentés dans le Tableau 8 ont trait au réalisme des étudiant-e-s en fonction de l'année d'inscription pour les prédictions académiques. Elles ont été obtenues auprès de deux cohortes (1996 et 1998). Dans l'ensemble, les données sont assez semblables pour les deux cohortes. Ainsi, par exemple, les prédictions les plus réalistes sont celles se rapportant à l'obtention d'un prêt ou d'une bourse. Il y a quand même quelques exceptions notables à cette première observation. Comme, par exemple, la prédiction se rapportant à l'abandon d'un cours pour les étudiant-e-s inscrit-e-s au deuxième niveau. Nous traiterons donc les données des deux cohortes comme si elles n'en formaient qu'une et de façon globale plutôt que de nous attacher à des résultats spécifiques qui peuvent être trompeurs.

Tableau 8

Réalisme académique (r_{pb}) selon l'année de présentation et l'année d'inscription

Item	Année de présentation du questionnaire					
	1996			1998		
	Niveau d'inscription		3 ^{ème}	Niveau d'inscription		3 ^{ème}
	1 ^{ère}	2 ^{ème}	3 ^{ème}	1 ^{ère}	2 ^{ème}	3 ^{ème}
« A » pour le cours	0.21	-0.01	-0.21	0.26	0.33	-0.11
Étudier 3 heures	0.30	0.41	0.32	0.18	0.40	0.19
Notes au-dessus de la moyenne	0.12	0.02	0.28	0.24	0.28	0.14
« A » pour cours préféré	0.22	-0.06	-0.10	0.18	0.14	0.39
Manquer un jour de classe	0.08	0.54	0.18	-0.07	0.08	0.29
Abandonner un cours	0.23	0.20	0.29	0.04	0.25	0.45
Révision de note	0.03	-0.04	0.14	0.17	-0.04	-0.14
« A » et « A+ »	0.26	0.38	0.13	-0.04	-0.02	-0.03
Moyenne suffisante trimestre suivant	0.14	0.29	0.00	0.26	0.00	-0.15

Tableau 8 (suite)

Réalisme académique selon l'année de présentation et l'année d'inscription

Item	Année de présentation du questionnaire					
	1996			1998		
	Niveau d'inscription		Niveau d'inscription		Niveau d'inscription	
	1 ^{ère}	2 ^{ème}	3 ^{ème}	1 ^{ère}	2 ^{ème}	3 ^{ème}
Travail l'été suivant	0.34	0.49	0.53	-0.07	0.30	0.32
Moyenne voulue année	-0.10	0.32	0.15	0.30	0.33	0.14
Moyenne suffisante année suivante	0.11	0.07	0.28	0.22	0.00	0.24
Obtenir prêt	0.83	0.85	0.88	0.77	0.88	0.84
Obtenir bourse	0.64	0.58	0.69	0.55	0.59	0.65
Obtenir assistance recherche	0.23	0.29	-0.10	0.00	0.29	0.55
Participer à une recherche	0.30	0.18	0.68	0.25	0.25	0.59
Problèmes pour payer	0.56	0.36	0.64	0.58	0.26	0.41

Nous pouvons constater alors que les étudiant-e-s du troisième niveau d'inscription sont plus souvent réalistes que ceux du premier niveau, les étudiant-e-s du deuxième niveau se situant entre les deux autres à cet égard.

Il y a aussi plus de deux fois plus de prédictions dont le réalisme est plus grand au troisième niveau qu'au premier (24 par rapport à 11). La même différence est de 19 par rapport à 15 entre le deuxième et le premier niveau, et 18 et 16 entre le troisième et le deuxième niveau. Encore là, donc la grandeur de la différence entre le troisième niveau et le deuxième est plus petite que celle entre le troisième et le premier niveau.

Réalisation selon le niveau de confiance

Le Tableau 9 présente les niveaux de confiance moyens selon que la prédiction s'est avérée ou non pour les deux cohortes réunies. Ces données complètent celles présentées plus haut. Elles nous permettent de constater, qu'en général, les étudiant-e-s sont assez confiants que les événements heureux se produiront et que les événements malheureux ne se produiront pas quelle que soit l'issue. Les exceptions notables à cette règle sont les réponses aux items ayant trait aux conditions financières des étudiant-e-s où les différences du niveau de confiance sont grandes. Par ailleurs, les items portant sur la probabilité d'obtenir la note « A » pour le cours suivi et pour le cours préféré ont produit les plus petites différences entre ceux pour lesquels la prédiction s'est réalisée et les autres.

Tableau 9

Réalisation selon le niveau de confiance

Item	Année d'inscription					
	1ère		2e		3e	
	Oui	Non	Oui	Non	Oui	Non
« A » pour le cours	8.05	7.32	6.70	6.05	6.30	7.09
Étudier 3 heures	8.76	7.76	8.44	6.16	7.08	4.61
Notes au-dessus de la moyenne	7.86	7.19	7.10	6.43	7.32	6.33
« A » pour cours préféré	8.87	8.31	7.68	7.49	8.46	8.25
Manquer un jour de classe	3.66	3.64	4.91	2.91	5.89	4.26
Abandonner un cours	2.00	1.00	2.90	1.07	6.84	1.57
Révision de note	3.67	2.81	2.38	2.64	2.84	2.66
« A » et « A+ »	5.83	4.99	4.85	3.60	4.74	4.27
Moyenne suffisante trimestre suivant	9.17	7.92	9.25	7.50	9.20	9.50

Tableau 9 (suite)
Réalisation selon le niveau de confiance

Item	Année d'inscription					
	1 ^{ère}		2 ^e		3 ^e	
	Oui	Non	Oui	Non	Oui	Non
Travail l'été suivant	5.17	3.48	7.25	3.96	9.50	4.80
Moyenne voulue année	7.56	7.28	7.90	6.65	7.89	7.21
Moyenne suffisante année suivante	9.00	8.04	9.33	9.00	7.61	5.00
Obtenir prêt	9.14	2.00	9.18	1.26	8.99	0.98
Obtenir bourse	7.69	2.89	7.65	2.22	6.63	1.01
Obtenir assistance recherche	8.00	4.78	6.80	3.78	5.50	3.01
Participer à une recherche	7.66	6.33	8.80	7.82	9.13	5.87
Problèmes pour payer	6.82	1.97	5.17	2.04	6.57	1.50

Pour ce qui est de la comparaison des étudiant-e-s en fonction du niveau d'inscription, on observe que la différence moyenne entre ceux pour qui la prédiction s'est réalisée et les autres est d'autant plus grande que le niveau d'inscription est élevé : 1.84, 2.10 et 2.50 respectivement pour le premier, le deuxième et le troisième niveau d'inscription. Ce qui suggère une augmentation du niveau de réalisme avec l'avancement dans les études. Une comparaison plus détaillée entre le premier et le troisième niveau d'inscription permet de constater que la différence est particulièrement grande pour ce qui est des items se rapportant à l'abandon d'un cours et au travail d'été. En revanche, les différences les moins importantes ont trait aux items se rapportant aux problèmes pour payer, aux notes au-dessus de la moyenne et à l'obtention d'un « A » pour le cours préféré.

ÉVOLUTION DU NIVEAU DE CONFIANCE DES PRÉDICTIONS

Le Tableau 10 présente l'évolution du niveau de confiance des prédictions académiques et professionnelles des mêmes 37 étudiant-e-s à trois moments : septembre 1995, décembre 1995 et septembre 1996.

Tableau 10

Prédictions académiques et professionnelles, évolution du niveau de confiance

Item	Septembre 1995		Décembre 1995		Septembre 1996		p^{\ddagger}
	M	ÉT	M	ÉT	M	ÉT	
« A » pour le cours	7.22	2.04	3.68	3.14	5.81	2.31	< .001
Étudier 3 heures ce trimestre	8.19	1.98	7.78	2.64	7.49	2.67	n. s.
Notes au-dessus de la moyenne ce trimestre	7.38 _a	2.29	5.24	3.29	6.59 _a	2.43	.002
« A » pour cours préféré ce trimestre	8.35	1.72	6.54 _a	3.04	7.51 _a	2.10	.002
Manquer un jour de classe ce trimestre	3.16	2.95	4.11 _a	4.03	4.19 _a	3.33	.017
Abandonner un cours ce trimestre	0.95	1.78	0.92	2.40	1.65	2.08	n. s.
Révision de note ce trimestre	2.35	1.81	1.84	1.83	2.59	1.64	n. s.
« A » et « A+ » ce trimestre	6.97	1.59	2.73 _a	3.21	3.95 _a	2.94	< .001

Note. Les moyennes ayant le même indice ne sont pas significativement différentes à $p < .05$.

p^{\ddagger} valeur de p bicaudale obtenue à la suite de la comparaison des trois moments par une analyse de variance.

Tableau 10 (suite)

Prédictions académiques et professionnelles, évolution du niveau de confiance

Item	Septembre 1995			Septembre 1996		
	M	ÉT	M	M	ÉT	p [†]
Travail l'été suivant	4.84 _a	2.22	3.14	4.95 _a	2.74	< .001
Moyenne suffisante pour trimestre suivant	8.81	2.13	8.73	9.35	1.14	n. s.
Moyenne voulue cette année	7.78	1.46	3.38	6.86	1.46	< .001
Moyenne suffisante pour année suivante	9.11 _{ab}	1.10	8.62 _a	9.30 _b	1.02	.004
Obtenir le baccalauréat	8.92 _{ab}	1.67	8.35 _a	9.27 _b	0.96	.007
Obtenir un prêt baccalauréat	4.84	4.21	4.51	4.54	4.34	n. s.
Obtenir une bourse baccalauréat	4.19	3.57	3.53	3.31	3.61	n. s.
Obtenir une assistance de recherche baccalauréat	4.67	2.24	4.64	3.89	2.15	n. s.

Note. Les moyennes ayant le même indice ne sont pas significativement différentes à $p < .05$.

p[†] valeur de p bicaudale obtenue à la suite de la comparaison des trois moments par une analyse de variance.

Tableau 10 (suite)

Prédictions académiques et professionnelles, évolution du niveau de confiance

Item	Septembre 1995		Décembre 1995		Septembre 1996		p^*
	M	ÉT	M	ÉT	M	ÉT	
Participer à une recherche baccalauréat	7.92	1.66	8.73	1.56	8.32	2.15	n. s.
Problèmes pour payer baccalauréat	1.37	1.68	2.20	2.74	1.86	2.03	n. s.
Être accepté à la maîtrise	7.32	1.51	6.14 _a	1.58	6.03 _a	2.17	< .001
Obtenir la maîtrise	7.38	1.46	6.16 _a	1.55	5.95 _a	2.04	< .001
Être accepté au doctorat	6.08	1.95	5.00 _a	2.04	4.58 _a	2.35	< .001
Obtenir un doctorat	5.69	2.12	4.50 _a	2.27	4.39 _a	2.26	< .001
Obtenir un emploi en psychologie	8.38	1.04	7.32 _a	1.86	7.32 _a	1.78	.002
Être membre de l'O.P.Q.	8.32	1.36	6.78	1.96	6.57	1.97	< .001

Note. Les moyennes ayant le même indice ne sont pas significativement différentes à $p < .05$.

p^* valeur de p bicaudale obtenue à la suite de la comparaison des trois moments par une analyse de variance.

Il y a 24 items et trois moments différents. Cela donne donc 72 comparaisons possibles si l'on compare les résultats de chacun des moments avec chacun des autres. Il ressort que 41 de ces comparaisons aboutissent à une absence de différence. Cela est vrai 13 fois lorsque la comparaison est entre septembre 1995 et septembre 1996, 11 fois entre septembre 1995 et décembre 1995 et 17 fois entre décembre 1995 et septembre 1996. Il y neuf items (27 comparaisons possibles sur 41) où il n'y aucune différence significative quels que soient les moments comparés.

Par ailleurs, des 72 comparaisons possibles, il n'y en a que huit où le niveau de confiance augmente d'un moment à l'autre, ces augmentations se situent entre décembre 1995 et septembre 1996. Il y en a également une où le niveau de confiance augmente d'un moment à l'autre entre septembre 1995 et décembre 1995 et une où le niveau de confiance augmente entre septembre 1995 et septembre 1996.

Enfin, pour le tiers des comparaisons possibles, il y a diminution des niveaux de confiance : 10 fois entre septembre 1995 et septembre 1996, 13 fois entre septembre 1995 et décembre 1995, une fois entre décembre 1995 et septembre 1996.

Cinq des items où il n'y a pas de changement du niveau de confiance ont trait aux prédictions concernant le baccalauréat. Elles portent, plus spécifiquement sur les conditions matérielles des étudiant-e-s, leurs possibilités de financer leurs études.

Les quatre autres items où il n'y a pas de changement ont trait à des items se rapportant aux prédictions concernant le trimestre. Trois de ces items portent sur le comportement de l'étudiant à l'égard de décisions qu'il devra prendre éventuellement et par rapport auxquelles il exerce un certain contrôle (items 2, 6, 7).

Les six augmentations du niveau de confiance d'une période à l'autre ont trait pour cinq d'entre elles à la performance académique (items 1, 3, 10, 11, 12, 13). Les étudiant-e-s sont plus confiants de réussir en septembre 1996 qu'en décembre 1995. Pour les mêmes items, il n'y a pas de différence entre septembre 1995 et septembre 1996, sauf pour deux items (1, 11).

Pour ce qui est des 24 cas où il y a diminution du niveau de confiance, il y a un item (le 37) où cela diminue d'une période à l'autre; deux où il y a diminution entre la première et la deuxième périodes, la troisième et la deuxième périodes et la première et la troisième périodes (items 1 et 11); deux où il y a diminution de la première période à la deuxième alors que la troisième période est égale à la première période (items 3 et 10); enfin, il y a 16 cas où la diminution se situe entre la première et la deuxième période alors que la deuxième est égale à la troisième période (items 4, 5, 8, 19, 20, 21, 22, 36), excepté à l'item 5, où il y a une augmentation entre la première et la deuxième période.

L'analyse du contenu des items où il y a une chute de confiance montre qu'il y a une baisse de la confiance par rapport aux conditions de satisfaction des exigences pour être promu d'un niveau à l'autre. Cette baisse de la confiance s'effectue surtout entre la première et la deuxième période et, sauf exceptions, elle a tendance à se maintenir à la troisième période.

ÉVOLUTION DU NIVEAU DE DÉTERMINATION

Le Tableau 11 présente l'évolution du niveau de détermination académique et professionnelle des mêmes 37 étudiant-e-s à trois moments : septembre 1995, décembre 1995 et septembre 1996.

Tableau 11

Évolution du niveau de détermination

Item	Moment						ρ
	Sept. 1995		Déc. 1995		Sept. 1996		
	M	ÉT	M	ÉT	M	ÉT	
« A » pour le cours	4.28	0.45	3.70 _a	0.97	3.92 _a	0.68	.001
« A » pour cours préféré	4.36	0.48	4.19	0.52	4.24	0.60	n.s.
Notes au-dessus de la moyenne	4.36	0.48	4.22	0.63	4.19	0.62	n.s.
Travail l'été suivant	3.27	0.73	3.08	0.80	3.21	0.75	n.s.
Obtenir prêt	2.62	1.66	2.76	1.55	2.43	1.61	n.s.
Obtenir bourse	3.46	1.12	3.22	1.25	3.00	1.15	n.s.
Participer à une recherche	3.64 _a	0.65	3.59 _a	0.72	3.24	0.64	< .001
Obtenir assistance recherche	3.49 _a	0.77	3.08 _b	0.92	3.24 _{ab}	0.86	.024

Note. Les moyennes ayant le même indice ne sont pas significativement différentes à $p < .05$.

ρ^* valeur de p bicaudale obtenue à la suite de la comparaison des trois moments par une analyse de variance.

Tableau 11 (suite)
Évolution du niveau de détermination

Item	Moment						p
	Sept. 1995		Déc. 1995		Sept. 1996		
	M	ÉT	M	ÉT	M	ÉT	
Obtenir baccalauréat	4.78	0.42	4.46 _a	0.65	4.57 _a	0.55	.005
Être accepté à la maîtrise	4.47	0.50	4.24 _a	0.60	4.03 _a	0.80	.001
Obtenir la maîtrise	4.43	0.50	4.16 _a	0.60	4.00 _a	0.78	.001
Être accepté au doctorat	3.51	0.77	3.30	0.97	3.22	0.95	n.s.
Obtenir un doctorat	3.49	0.80	3.22	0.98	3.27	1.12	n.s.
Obtenir un emploi en psychologie	4.41	0.50	4.16 _a	0.60	4.08 _a	0.72	.004
Être membre de l'O.P.Q.	4.38	0.64	4.14	0.63	3.81	0.74	< .001

Note. Les moyennes ayant le même indice ne sont pas significativement différentes à $p < .05$.

p^+ valeur de p bicaudale obtenue à la suite de la comparaison des trois moments par une analyse de variance.

Il y a 15 items et trois moments différents. Cela donne donc 45 comparaisons possibles si l'on compare les résultats de chacun des moments avec chacun des autres. Il ressort que 29 de ces comparaisons aboutissent à une absence de différence. Cela est vrai huit fois lorsque la comparaison est entre septembre 1995 et septembre 1996, huit fois entre septembre 1995 et décembre 1995 et 13 fois entre décembre 1995 et septembre 1996. Il y sept items (21 comparaisons possibles sur 45) où il n'y a aucune différence significative quels que soient les moments comparés.

Par ailleurs, des 45 comparaisons possibles, il n'y en a aucune où le niveau de détermination augmente d'un moment à l'autre.

Il y a diminution des niveaux de détermination : sept fois entre septembre 1995 et septembre 1996, sept fois entre septembre 1995 et décembre 1995, deux fois entre décembre 1995 et septembre 1996.

La tendance générale va dans la direction d'une baisse de la détermination, baisse tantôt significative, tantôt non. La baisse la plus importante a trait au fait de devenir membre de l'Ordre des psychologues du Québec et la moins importante, au fait de trouver du travail l'été suivant.

NIVEAU DE CONFIANCE SELON LA DISTANCE TEMPORELLE

Les résultats qui suivent présentent l'évolution du niveau de confiance à trois moments différents des mêmes 37 étudiant-e-s par rapport à neuf événements situés plus ou moins loin dans le temps. La Figure 4 permet de comparer les résultats selon le moment où la prédiction a été faite.

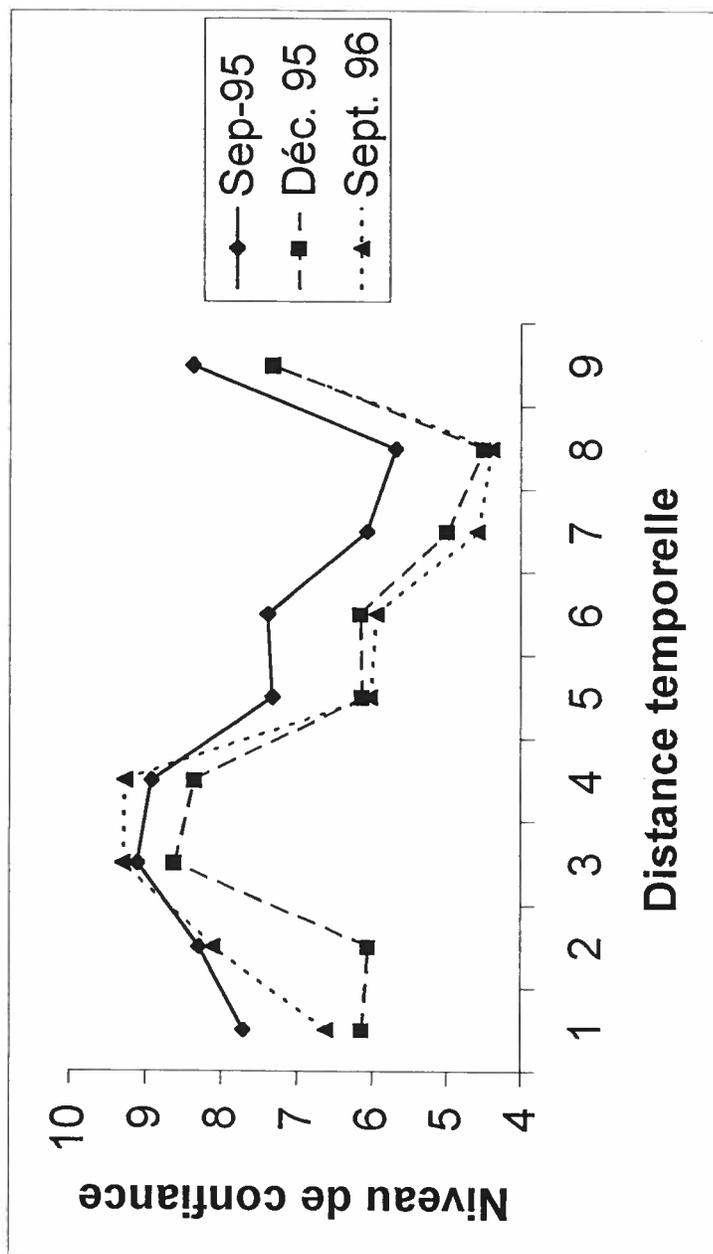


FIGURE 4. Niveau de confiance à trois moments en fonction de la distance temporelle.

Notons d'abord, qu'en général, le niveau de confiance est plus élevé lorsque les étudiant-e-s répondent en septembre 1995 qu'aux deux autres moments pour les six derniers événements de l'échelle de distance temporelle (niveaux de confiance moyens respectifs de 7.30, 6.25 et 6.26). La différence est encore plus importante si le niveau 4, obtenir le baccalauréat, n'est pas considéré (niveaux de confiance moyens respectifs de 6.97, 5.82 et 5.65).

Pour ce qui est de la comparaison des événements entre eux selon la distance temporelle (Figure 5), on observe une augmentation du niveau de confiance moyen selon que les items se rapportent au trimestre actuel (4.88) (items 1 à 8), le trimestre suivant (7.49) (items 10 et 11) ou l'année suivante (9.01) (item 12).

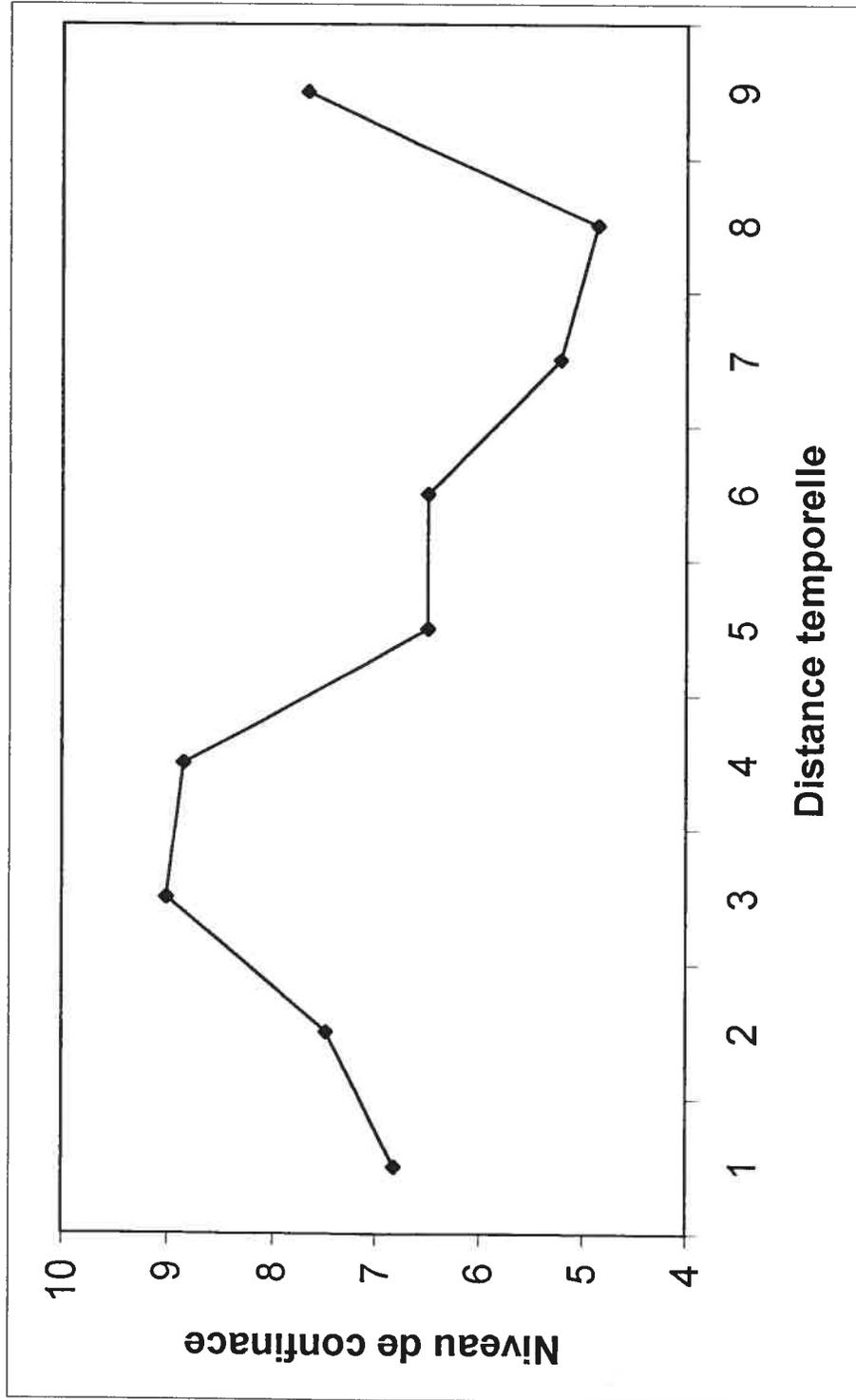


FIGURE 5. Niveau de confiance en fonction de la distance temporelle. Moyenne des trois mesures.

Cette augmentation est suivie d'une légère baisse à 8.85 (item 13) pour ce qui est du niveau de confiance par rapport à l'obtention du baccalauréat; d'une baisse plus importante pour ce qui est de l'acceptation à la maîtrise (6.50) (item 19) qui se maintient à l'égard de l'obtention de la maîtrise (item 20); la baisse se poursuit en ce qui concerne le doctorat (niveau de confiance à 5.22 (item 21) pour l'acceptation et 4.86 pour l'obtention (item 22)). Toutefois, il y a une augmentation importante à 7.67 (item pour ce qui est de l'obtention d'un emploi en psychologie (item 36).

CHAPITRE IV : INTERPRÉTATION

L'hypothèse affirmant qu'il y a une meilleure calibration pour l'examen final que pour l'examen intra en ce qui concerne la calibration absolue n'est pas confirmée. Les étudiant-e-s, pourtant, savent mieux à quoi s'en tenir par rapport à la matière et au professeur à cette étape du cours. Ils ont déjà subi un examen de même nature que l'examen final à choix multiples (trois réponses proposées dans les deux cas). Ils ont eu droit au feedback de la note de l'examen intra¹.

Plusieurs raisons pourraient expliquer ce résultat. Notons d'abord que les examens, tout en se ressemblant beaucoup, ne se ressemblaient pas autant que ceux pour lesquels Glenberg, Sanocki, Epstein, et Morris, (1987) ont observé une amélioration de la justesse de la confiance. Dans les examens que nous avons présentés aux étudiant-e-s, la principale différence concernait leur difficulté. Le deuxième examen étant plus facile que le premier. La réussite en pourcentage est meilleure, même si ce n'est pas le cas pour la note exprimée en lettre puisqu'elle prend en compte la dispersion des notes. Il se peut donc que le feedback du premier examen ait eu un effet sur la confiance des étudiant-e-s qui ne se répercute pas dans la mesure de la calibration. L'indice de calibration, en effet, ne distingue pas entre la surconfiance et la sous-confiance. Il repose sur un calcul qui fait la somme des écarts par rapport à la calibration à chacun des niveaux de confiance indépendamment de la direction de cet écart. Ainsi, un écart moyen dans une direction pour un examen n'est pas différencié d'un écart moyen dans la direction opposée pour l'autre examen par la mesure de calibration. La reproduction des écarts dans une figure permet de mieux voir ce qui en est. Nous y reviendrons donc quand nous parlerons de la calibration relative.

Par ailleurs, il se peut que l'échelle de confiance utilisée ait créé une confusion chez certains étudiant-e-s de sorte que la compréhension de cette

¹ Mais pas au feedback en classe des réponses.

échelle n'ait pas été uniforme. Après tout, encore une fois, si l'on est absolument sûr que la réponse que nous choisissons est la mauvaise, pourquoi la choisissons-nous plutôt que l'une des deux autres quand on sait pertinemment bien que l'une des réponses est bonne? Encore là, les résultats rapportés à l'égard de la calibration relative seront en mesure de mieux nous éclairer à ce propos. Les données se rapportant à la calibration de la certitude permettront aussi d'avoir une meilleure idée sur l'effet de l'expérience acquise par les étudiant-e-s puisque ces données sont moins équivoques.

Nous observons une différence significative en ce qui concerne la calibration absolue et le genre, mais, contrairement, à l'hypothèse que nous avons proposée, ce sont les étudiants qui ont la meilleure calibration absolue. Nous avons fait cette hypothèse même si les recherches antérieures ne démontraient pas un avantage clair des femmes sur les hommes (v.g. Beyer, 1998). Nous aurions obtenu une absence de différence que cela aurait été moins étonnant que les résultats que nous obtenons. Il reste que la différence observée, toute significative qu'elle soit, est à toutes fins pratiques minime, soit .05 point à peine. Notons que le *caveat* que nous avons exprimé à l'égard de l'indice de calibration absolue s'applique évidemment ici aussi et pourrait expliquer ce résultat inattendu. Nous verrons donc ce qui en est pour la calibration de la certitude.

Pour répondre à une des questions que nous nous posons, nous n'observons aucune corrélation significative entre la calibration absolue et la perception de l'examen. Cela est vrai que ce soit pour la perception de la représentativité de l'examen ou de sa difficulté, que ce soit pour l'examen intra ou pour l'examen final. Pour les raisons mentionnées plus haut, il sera intéressant de voir ce qui en est avec la calibration de la certitude. Notons que nous ne trouvons pas, non plus, de relation entre la perception de la

représentativité de l'examen et la perception de la difficulté, bien qu'on pourrait penser qu'un examen perçu comme étant moins représentatif serait perçu comme plus difficile. Par ailleurs, le deuxième examen est perçu, avec raison, comme plus facile.

L'hypothèse d'un lien entre la calibration absolue et la note obtenue est confirmée pour le premier examen, mais pas pour le deuxième. Comme nous l'avons vu, de nombreuses recherches (v.g. Hacker et al., 2000) établissent un lien entre le réalisme des étudiant-e-s à un examen et leur performance à cet examen ou en général. C'est aussi le cas ici, mais pour un seul examen. L'interprétation des résultats obtenus avec d'autres variables que la calibration absolue nous permettront de mieux comprendre ce qui a pu se passer.

Contrairement à ce que nous avons obtenu pour la calibration absolue, nous observons une relation entre la calibration de la certitude et l'ordre de l'examen. La calibration de la certitude est meilleure pour l'examen final conformément à l'hypothèse. La calibration de la certitude est plus facile à interpréter que la calibration absolue. Nous sommes aussi certains que faire se peut qu'il y a eu compréhension de la signification de ce point de l'échelle de confiance de la part des étudiant-e-s. Les étudiant-e-s semblent donc mieux capables de prévoir à l'examen final les bonnes réponses dont ils sont sûrs. Quand ils se trompent, autant dans le cas d'un examen que dans l'autre, ils le font forcément dans la direction de la surconfiance, des prédictions sous-confiantes étant impossibles dans le cas de la calibration de la certitude. Est-ce que les étudiant-e-s ont une meilleure calibration pour le niveau de confiance « 10 » à l'examen final qu'à l'examen intra parce qu'ils ont une information plus complète pour baser leurs prédictions? Ce n'est pas nécessairement le cas puisque la difficulté de ces deux examens n'a pas été contrôlée et que la difficulté d'une question, on l'a vu, est une des variables

le plus souvent évoquées pour expliquer la mauvaise calibration (Baranski et al., 1995; Lichtenstein et al., 1977; Suantak, 1996). Or, on l'a vu, la difficulté du deuxième examen est moins grande. Dans le contexte dans lequel a été réalisée cette recherche, il aurait été compliqué et, même à la limite, nuisible aux étudiant-e-s d'avoir tenté un contrôle de la difficulté. N'oublions pas, en effet, que cette recherche s'est faite dans le cadre d'une évaluation ayant des conséquences relativement importantes pour les étudiant-e-s, en milieu naturel, plutôt qu'en laboratoire. Dans ces circonstances, nous avons préféré intervenir le moins possible tout en espérant que le niveau de difficulté entre les deux examens s'avérerait, au pire, très peu différent. Étant donné que ce ne fut pas le cas, la variable difficulté ne peut être rejetée pour expliquer nos résultats. Les travaux antérieurs nous invitent, au contraire, à la considérer comme étant une interprétation rivale sérieuse.

Les étudiantes ont une moins bonne calibration de la certitude que les étudiants, mais cette différence n'atteint pas le niveau de signification. Cette hypothèse n'est donc pas confirmée. De toutes façons, elle prédisait une différence dans la direction contraire. Ces résultats vont dans le sens de ceux obtenus avec la calibration absolue, sauf qu'ici ils ne sont pas significatifs. Si des recherches ultérieures devaient confirmer cette tendance des étudiants à avoir une meilleure calibration que les étudiantes à des questionnaires à choix multiples semblables, la seule interprétation que nous voyons irait dans le sens des travaux de Beyer (1990). À savoir que les étudiantes auraient des attentes moins élevées de réussite que les étudiants pour la matière considérée et que cela affecterait éventuellement leur performance. La matière en question est ici la psychologie sociale. Rien, à priori, ne permet de penser que les étudiantes auraient des attentes moins élevées que les étudiants dans cette matière. Toutefois, la psychologie sociale est une discipline qui couvre un très large champ et le professeur du

cours où étaient inscrits les étudiant-e-s de notre recherche a choisi de mettre l'accent plus qu'habituellement sur une présentation se situant à un niveau d'abstraction se rapprochant parfois de la formalisation mathématique. Or, c'est précisément dans le cas de matières de cette nature que Beyer (1990) observe une moins bonne calibration et un moins bon réalisme, chez les étudiantes.

La perception de la représentativité et celle de la difficulté sont liées à la calibration de la certitude pour l'examen intra, mais pas pour l'examen final. La corrélation est directe pour la représentativité, mais inverse pour la difficulté. L'examen final étant plus facile, il est aussi perçu comme tel et la dispersion des jugements est moins grande dans le cas de cet examen, ce qui explique peut-être l'absence de corrélation entre la calibration de la certitude et la perception de la difficulté de cet examen. Il est plus difficile d'expliquer l'absence de relation avec la représentativité, surtout qu'il ne semble pas y avoir de lien entre la difficulté perçue et la représentativité perçue. Pour ce qui est de l'examen intra, on obtient une relation entre ce type de calibration et la difficulté perçue qui est semblable à celle à laquelle on peut s'attendre entre la calibration et la difficulté. Il resterait à se demander si la relation avec la difficulté perçue est aussi importante que celle avec la difficulté réelle et si elle s'explique de la même façon. Aussi, est-ce qu'un examen difficile non perçu comme tel entraîne le même effet sur la calibration?

L'hypothèse d'une corrélation entre la note obtenue et la calibration de la certitude pour l'examen intra est confirmée, mais ce n'est pas le cas pour l'examen final. Encore ici, la plus grande facilité de l'examen final est probablement la raison de l'absence de relation constatée.

Conformément à notre hypothèse, les étudiant-e-s font bien preuve d'une calibration relative autant pour l'examen intra que pour l'examen final. À part certaines déviations, en effet, plus le niveau de confiance est élevé, plus la proportion de réponses correctes est élevée. Notons que la fréquence d'utilisation des niveaux de confiance est inégale, de sorte que l'inspection visuelle de la figure rapportant la relation entre les niveaux de confiance et la proportion de réponses correctes est trompeuse. La calibration relative est en fait meilleure qu'il n'y paraît étant donné l'utilisation nettement moins importante des niveaux de confiance inférieurs.

Par ailleurs, l'hypothèse d'une calibration relative meilleure à l'examen final n'est pas confirmée. En effet, l'écart entre la courbe de la calibration de l'examen final et la diagonale est plus grand pour cet examen que pour l'examen intra. Si l'on fait exception des niveaux de confiance inférieurs qui, comme nous le disions, sont utilisés moins fréquemment, la courbe de calibration de l'examen intra démontre une légère surconfiance alors que celle de l'examen final suggère une sous-confiance importante. Ces résultats sont typiques de ceux habituellement obtenus en fonction de la difficulté de l'examen.

Contrairement à l'hypothèse, le réalisme de Type I n'est pas plus grand pour l'examen final que pour l'examen intra. Autrement dit, les étudiant-e-s ne sont pas meilleur-e-s pour prédire la note qu'ils auront à l'examen final qu'à l'examen intra. Là aussi nous serions tentés d'expliquer ces résultats par la différence de difficulté entre les deux examens. D'ailleurs, dans l'ensemble, les étudiant-e-s ont tendance à être légèrement surconfiant-e-s pour l'examen intra et légèrement sous-confiant-e-s pour l'examen final lorsque cet indice est utilisé.

Le réalisme de Type I n'est pas non plus meilleur pour les étudiantes que pour les étudiants. Les résultats indiquent même une légère tendance pour les étudiants à être sous-confiants par rapport aux étudiantes autant dans un examen que dans l'autre.

Les corrélations entre le réalisme de Type I et la perception des examens sont significatives, mais faibles, sauf pour ce qui est de la représentativité de l'examen pour l'examen intra où la corrélation n'est pas significative. Ces résultats obtenus avec un grand groupe d'étudiant-e-s renforcent l'hypothèse d'une relation entre la difficulté perçue et le réalisme déjà observé avec l'indice de la calibration de la certitude. On peut penser que si l'indice de la calibration absolue avait été plus fiable, on aurait, là aussi, observé cette relation. Pour ce qui est de la relation avec la représentativité perçue, le tableau est moins clair. Peut-être parce qu'il est plus facile pour un étudiant-e de porter un jugement sur la difficulté que sur la représentativité. Peut-être, aussi, parce que les étudiant-e-s n'ont pas tous-tous la même façon de déterminer si un examen est représentatif ou non. Par exemple, un critère important pourrait être pour un étudiant-e qu'il y ait autant de questions tirées des notes de cours que du livre obligatoire, alors qu'un-e autre étudiant-e pourrait ne même pas songer à cette possibilité.

Meilleure est la note autant à l'examen intra qu'à l'examen final, plus grande est la corrélation avec le réalisme de Type I, ce qui confirme notre hypothèse. Cela s'ajoute aussi aux relations semblables obtenues par d'autres mesures de réalisme (calibration absolue et calibration de la certitude).

La comparaison des données obtenues en fonction du niveau d'inscription en ce qui concerne le réalisme de Type II tend à confirmer l'hypothèse d'un plus grand réalisme en fonction du niveau d'inscription.

Elles ajoutent à cette hypothèse que la différence entre les deux derniers niveaux d'inscription est moins grande qu'avec le premier niveau. Il semblerait donc que les étudiant-e-s entretiendraient plus d'illusions au début de leur scolarité à l'université. Il semblerait aussi qu'une grande partie de ces illusions disparaîtrait dès le passage du premier au deuxième niveau puisque l'écart entre le deuxième et le troisième niveau est plus petit. Nos données à cet effet sont relativement claires. Toutefois, est-ce que cette différence selon les niveaux vient du fait que ceux qui se berçaient le plus d'illusions ont abandonné en cours de route ou est-ce qu'il y a eu une évolution chez les mêmes individus d'un niveau à l'autre? Ces données ne peuvent le dire étant donné la nature transversale des comparaisons. Nous verrons mieux ce qui en est avec les comparaisons longitudinales qui suivront plus bas.

La mise en rapport des niveaux de confiance et de la réalisation ou non des événements rapportée au Tableau 9 permet de constater que le réalisme en général est plutôt faible, sauf pour les conditions financières. Il faut dire, qu'à cet égard, l'information est claire contrairement aux autres cas. Il semble que les étudiant-e-s pour lesquels les événements heureux ne se manifesteront pas arrivent à se convaincre facilement du contraire. L'absence de différence ou les petites différences entre ceux pour lesquels l'événement souhaité se produira et ceux pour lesquels il ne se produira pas pourrait ne pas indiquer un manque de réalisme de la part de ces derniers si les autres démontraient une plus grande sous-confiance, mais ce n'est pas le cas. On doit en conclure qu'à moins d'un feedback de qualité, les étudiant-e-s arrivent à se convaincre relativement facilement que ce qu'ils souhaitent se produira. Cette tendance s'atténue avec le niveau d'inscription, mais elle ne disparaît pas complètement, loin de là. Ces résultats, encore une fois, ne nous permettent pas de dire pourquoi la différence s'atténue avec le niveau

d'inscription. Est-ce que les étudiant-e-s les moins réalistes au premier niveau ne poursuivent pas leurs études ou est-ce parce que les étudiant-e-s deviennent de plus en plus réalistes ?

L'étude longitudinale sur un an montre que, même si le niveau de confiance des étudiant-e-s est le plus souvent porté à se maintenir, notre hypothèse est partiellement confirmée puisque lorsqu'il y a changement du niveau de confiance des étudiant-e-s, il est très majoritairement dans le sens d'une diminution de la confiance. Une telle baisse de confiance n'est pas, non plus, rare, puisqu'elle implique le tiers des possibilités de changement. Les résultats dont nous avons parlé plus haut qui montraient un manque de réalisme selon le niveau d'inscription des étudiant-e-s pouvaient s'interpréter comme étant l'aboutissement de la disparition des étudiant-e-s les moins réalistes ou l'augmentation du réalisme selon le niveau d'inscription. Les résultats de l'étude longitudinale ne permettent pas de rejeter la première interprétation de façon définitive, mais rendent la deuxième plausible. Il se peut bien, qu'au bout du compte, les deux phénomènes jouent simultanément dans une mesure plus ou moins grande.

Pour ce qui est de l'évolution du niveau de détermination, les résultats sont semblables à ceux de l'évolution du niveau de confiance. Il y a, en général, une diminution de la détermination, même si cette diminution n'est pas toujours significative. Malgré, donc, une relative courte période de temps, un an, la hauteur des visées des étudiant-e-s baisse. Il y a un seul cas où le niveau de confiance de la réalisation d'un événement a tendance à augmenter alors que la détermination pour le réaliser diminue. Il s'agit de la situation où l'étudiant serait appelé à être sujet de recherche : il perçoit cet événement comme plus probable, mais il n'y tient pas... Nos résultats ne nous permettent pas d'établir une relation de cause à effet entre le niveau de

confiance de la réalisation d'un événement et la détermination de l'étudiant, mais cette relation a souvent été observée dans la documentation.

L'hypothèse d'une relation inverse entre le niveau de confiance et la distance temporelle est partiellement confirmée par nos résultats. En ce qui concerne les données se rapportant à la comparaison du niveau de confiance des étudiant-e-s à différents moments, on observe qu'il y a diminution du niveau de confiance avec le temps : Les étudiant-e-s sont plus confiant-e-s en septembre 1995 que plus tard, mais il n'y a pas de différence entre les deux derniers moments. La baisse de confiance se situerait donc surtout au premier trimestre, de septembre à décembre. L'information nouvelle qu'ils acquièrent par la suite n'aurait donc pas autant d'impact. Ces résultats impliquent que dans un tel contexte la baisse du niveau de confiance ne se ferait pas de façon continue, mais par saut. Une telle possibilité n'est pas évoquée dans les travaux que nous avons cités dans le contexte théorique. Elle mériterait d'être étudiée plus en profondeur et avec un plus grand groupe d'étudiant-e-s. Cette baisse de confiance veut-elle dire que les étudiant-e-s sont plus réalistes avec le temps? On peut le penser, mais nos résultats ne permettent pas de l'affirmer puisque nous n'avons pas mis le niveau de confiance en relation avec un critère objectif. Une recherche ultérieure pourrait permettre de répondre à cette question. Il serait aussi préférable de mener une pareille recherche sur une période plus longue. Cela, toutefois, nécessite au départ des groupes plus grands pour s'assurer de maintenir à terme un échantillon suffisant.

La comparaison selon la distance temporelle des événements confirme aussi en partie l'hypothèse et montre que la relation entre le niveau de confiance et la distance temporelle n'est pas, ici non plus, simple. Pour la séquence des trois premiers événements, il y a effectivement une augmentation du niveau de confiance : plus l'événement est distant, plus la

confiance est grande. En revanche, ce n'est pas le cas pour les cinq événements suivants où il y a plutôt baisse du niveau de confiance. Comment interpréter cette différence d'avec les trois premiers événements? Il se peut que les étudiant-e-s soient motivés par une perspective à relativement court terme. Ils sont prêts à mettre tous les efforts nécessaires pour atteindre leur objectif à l'intérieur de cette perspective et ils ont besoin d'y croire... On peut penser alors, qu'effectivement, moins l'événement positif est proche, moins ils perçoivent les embûches qui pourraient les empêcher d'atteindre leur objectif. Au-delà de la perspective temporelle en question, il se peut que leur détermination soit moins importante et, aussi, qu'ils perçoivent malgré tout la réalisation de ces événements comme dépendant moins de leur bonne volonté, comme échappant en plus grande partie à leur contrôle. La vérification de cette interprétation demanderait d'utiliser des moyens que nous n'avons pas prévus. Il va de soi aussi que la vérification de cette relation devrait être répliquée auprès d'un échantillon plus grand d'étudiant-e-s.

Reste à expliquer la hausse importante du niveau de confiance pour le dernier événement de la séquence : obtenir un emploi en psychologie. Cette divergence peut se comprendre facilement si l'on admet que, pour les étudiant-e-s, en fait, obtenir un emploi en psychologie suit le baccalauréat plutôt que le doctorat. Si l'on déplace cet événement du neuvième rang dans la séquence au cinquième, on obtient une courbe qui respecte sans doute mieux la séquence des événements telle que perçue par les étudiant-e-s. Elle a une forme qui comporte moins de points de brisures (cf. Appendice B). Ceci étant dit, connaissant le marché du travail, on ne peut s'empêcher de qualifier d'irréaliste un niveau de confiance de 7.67 par rapport à l'obtention d'un emploi dans le domaine de la psychologie avec en mains le seul baccalauréat.

CHAPITRE V : CONCLUSION

La principale hypothèse générale que nous faisons en était une de relation directe entre l'expérience des étudiant-e-s mesurée de différentes façons et à différents moments de leur parcours académique. En ce qui concerne les trois types de calibration (absolue, de certitude et relative) et le réalisme de type I que nous avons utilisés pour l'examen intra et l'examen final, nous n'obtenons à cet égard un effet qu'avec la calibration de la certitude. Il serait intéressant de voir si l'utilisation d'une mesure de la calibration absolue utilisant une démarche plus appropriée viendrait confirmer ce résultat. Quoi qu'il en soit, il faudrait aussi contrôler pour la difficulté des deux examens. Ou, au contraire, manipuler les niveaux de difficulté en même temps que les niveaux d'expérience. Cela nous permettrait de connaître l'effet de la difficulté en interaction avec l'expérience. À quel niveau de difficulté est-ce que le rôle de l'expérience est le plus important?

Pour ce qui est de la relation de l'expérience avec le réalisme d'événements académiques, autant les résultats de l'étude transversale que de l'étude longitudinale suggèrent qu'il y a augmentation du réalisme avec le temps. Toutefois, comme on l'a vu, d'une part, il y a des exceptions et, d'autre part, quelle est la validité de l'hypothèse rivale?

Les comparaisons opposant les genres en ce qui concerne les examens ne révèlent pas de différences claires entre les étudiants et les étudiantes quant au réalisme des uns et des autres.

Il en est tout autrement pour ce qui est de la note. Sur six possibilités, il y en a quatre où il y a une relation entre la note obtenue à l'examen et le réalisme. Les meilleurs étudiant-e-s sont aussi les plus réalistes. Comment rendre les étudiants qui réussissent moins bien plus réalistes? Plusieurs se sont posés cette question et ont fait des suggestions. Nous croyons à la pertinence d'une intervention à cet égard parce que certains n'ont pas les moyens de profiter de leur expérience et ont besoin d'aide.

Les données portant sur la perception des examens ne révèlent pas de relations importantes entre la calibration (absolue et de certitude) et la perception de la représentativité ou de la difficulté de l'examen. Toutefois, nous avons des raisons de croire qu'une mesure plus adéquate de la calibration absolue et un examen final moins facile pourrait donner des résultats différents étant donnée la relation qui existe pour la calibration de la certitude autant pour la représentativité que pour la difficulté à l'examen intra.

En résumé, les principaux résultats qui se dégagent de cette recherche sont donc les suivants :

- Le manque de réalisme des étudiant-e-s est la règle plutôt que l'exception.
- Le niveau de ce manque de réalisme est élevé.
- Le manque de réalisme s'exprime surtout par de la surconfiance.
- Le manque de réalisme diminue avec l'expérience, mais il reste à un niveau élevé.
- Le manque de réalisme ne diminue pas de façon monotone, la diminution est plus importante au début.
- Les étudiant-e-s sont capables de réalisme, mais il faut que l'information soit quasiment évidente comme dans le cas des conditions financières.
- Pour ce qui est du réalisme des étudiant-e-s à l'égard des examens en particulier :
 - La difficulté de l'examen s'est avérée une variable plus importante que l'expérience des étudiant-e-s.

- Nous n'avons pas trouvé de lien entre le réalisme et le genre.
- Nous n'avons pas trouvé de lien important entre le réalisme et la perception de l'examen, qu'il s'agisse de la représentativité ou de la difficulté.
- En revanche, il y a une relation entre le réalisme et la note obtenue.

Le manque de réalisme des étudiant-e-s que nous avons observé ne nous apparaît pas sain. S'il est vrai qu'une dose de surconfiance peut amener l'individu à se surpasser, il n'est pas vrai que cela sera efficace quelle que soit cette dose. Au-delà d'un certain point, l'étudiant-e risque de se dépenser inutilement. Et s'il doit s'en remettre à lui-même pour trouver ce seuil, il n'y arrivera peut-être jamais.

Dans les recherches à venir, à la mesure de la calibration, il faudrait adjoindre la mesure de la discrimination pour ajouter à notre compréhension du comportement des étudiant-e-s. Il va de soi, aussi, que des études longitudinales sur une plus longue période et avec un nombre plus important d'étudiant-e-s s'imposent. Il serait intéressant aussi de comparer des cohortes d'étudiant-e-s sur une plus grande période pour voir, entre autres, dans quelle mesure des variables contextuelles peuvent influencer le réalisme des étudiant-e-s. Il faudrait aussi être en mesure de pouvoir comparer le niveau de réalisme des étudiant-e-s en fonction de leur discipline. Enfin, il serait essentiel de se demander comment on pourrait aider les étudiant-e-s à établir un niveau de surconfiance adéquat : ni trop élevé ni trop faible. Pour cela, il nous faut d'abord, nous-mêmes, trouver le moyen d'établir quel est ce niveau, cette bande étroite en deçà et au-delà de laquelle la performance et le cheminement peuvent laisser à désirer. En attendant, il nous semble évident que les administrateurs de programmes

doivent démontrer une certaine retenue dans les informations officielles qu'ils diffusent afin de ne pas nourrir des illusions coûteuses. À cet égard, le site internet de l'Université McGill s'adressant aux étudiant-e-s intéressé-e-s par le baccalauréat en psychologie nous apparaît être un modèle intéressant.

RÉFÉRENCES

- Alloy, L. B., & Abramson, L. Y. (1988). Depressive realism: Four theoretical perspectives. In L. B. Alloy (Ed.), *Cognitive processes in depression*. (pp. 223-265). New York: Guilford Press.
- Alloy, L. B., Abramson, L. Y., & Viscusi, D. (1981). Induced mood and the illusion of control. *Journal of Personality and Social Psychology*, *52*, 1129-1140.
- Anderson, C. A. (1983). Imagination and expectation: The effects of imagining behavioral scripts on personal intentions. *Journal of Personality and Social Psychology*, *45*, 293-305.
- Ansel, D., & Touzard, H. (2002). Surconfiance et variables contextuelles: effets sur la performance du négociateur. *Cahiers Internationaux de Psychologie Sociale*(55), 26-38.
- Ayton, P., & Wright, G. (1994). Subjective probability: what should we believe? In G. Wright & P. Ayton (Eds.), *Subjective probability*. (pp. 163-183). Chichester: Wiley.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Baranski, J. V., & Petrusic, W. M. (1995). On the calibration of knowledge and perception. *Canadian Journal of Experimental Psychology*, *49*, 397-407.
- Beach, L. R., Christensen-Slzenski, J., & Barnes, V. (1987). Assessing human judgment: Has it been done, can it be done, should it be done? In G. Wright & P. Ayton (Eds.), *Judgmental forecasting*. (pp. 49-62). Chichester: Wiley.

- Beyer, S. (1990). Gender differences in the accuracy of self-evaluations of performance. *Journal of Personality and Social Psychology*, 59, 960-970.
- Beyer, S. (1998). Gender differences in self-perception and negative recall biases. *Sex Roles*, 38, 103-133.
- Beyer, S. (1998/1999). Gender differences in causal attributions by college students of performance on course examinations. *Current Psychology*, 17(346-358).
- Beyer, S. (1999). Gender differences in the accuracy of grade expectancies and evaluations. *Sex Roles*, 41, 279-296.
- Beyer, S. (2002). The effects of gender, dysphoria, and performance feedback on the accuracy of self-evaluations. *Sex Roles*, 47, 453-464.
- Beyer, S., & Bowden, E. M. (1997). Gender differences in self-perception: Convergent evidence from three measures of accuracy and bias. *Personality and Social Psychology Bulletin*, 23, 157-172.
- Björkman, M. (1994). Internal use theory: Calibration and resolution of confidence in general knowledge. *Organizational Behavior and Human Decision Processes*, 58, 386-405.
- Bol, L., & Hacker, D. J. (2001). A comparison of the effects of practice tests and traditional review on performance and calibration. *Journal of Experimental Education*, 69, 133-151.
- Boucher, J., & Osgood, C. E. (1969). The Pollyanna hypothesis. *Journal of Verbal Learning and Verbal Behavior*, 8, 1-8.
- Bradley, D. P. (1981). Overconfidence in ignorant experts. *Bulletin of the Psychonomic Society*, 17, 82-84.

- Brenner, L. A., Koehler, D. J., Liberman, V., & Tversky, A. (1996). Overconfidence in probability and frequency judgments. *Organizational Behavior and Human Decision Processes*, 65, 212-219.
- Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, 75, 1-3.
- Buehler, R., Griffin, D., & Ross, M. (1994). Exploring the "planning fallacy": Why people underestimate their task completion times. *Journal of Personality and Social Psychology*, 67, 366-381.
- Bureau de recherche institutionnelle (2004). *Cheminement universitaire des étudiants*. Montréal: Université de Montréal.
- Busey, T. A., Tunnicliff, J., Loftus, G. R., & Loftus, E. F. (2000). Accounts of the confidence-accuracy relation in recognition memory. *Psychonomic Bulletin & Review*, 7, 26-48.
- Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: a theoretical synthesis. *Review of Educational Research*, 65, 245-281.
- Campbell, J. D., & Fairey, P. J. (1985). Effects of self-esteem, hypothetical explanations, and verbalization of expectancies on future performance. *Journal of Personality and Social Psychology*, 48, 1097-1111.
- Carroll, J. S. (1978). The effect of imagining an event on expectations for the event: An interpretation in terms of the availability heuristic. *Journal of Experimental Social Psychology*, 14, 88-96.
- Cervone, D. (1989). Effects of envisioning future activities on self-efficacy judgments and motivation: An availability heuristic interpretation. *Cognitive Therapy and Research*, 13, 247-261.
- Chouinard, D., & Hamel, J. (1995). *Que sont-ils devenus?* (8 ed.). Montréal: Service d'orientation et de consultation psychologique.

- Christensen-Szalanski, J. J. J., Beck, D. E., Christensen-Slzenski, C. M., & Koepsell, T. D. (1983). The effect of journal coverage on physicians' perception of risk. *Journal of Applied Psychology, 68*, 278-284.
- Dawson, N. V. (2000). Physician judgment of uncertainty. In G. B. Chapman & F. A. Sonnenberg (Eds.), *Decision making in health care: Theory, psychology, and applications*. (pp. 211-252). New York: Cambridge University Press.
- Detmer, D. E., Fryback, D. G., & Gassner, K. (1978). Heuristics and biases in medical decision-making. *Journal of Medical Education, 53*, 682-683.
- Dubois, J., Giacomo, M., Guespin, L., Marcellesi, C., Marcellesi, J.-B., & Mével, J.-P. (1973). *Dictionnaire de linguistique*. Paris: Librairie Larousse.
- Dunning, D., Griffin, D. W., Milojkovic, J. D., & Ross, L. (1990). The overconfidence effect in social prediction. *Journal of Personality and Social Psychology, 58*, 568-581.
- Dunning, D., & Story, A. L. (1991). Depression, realism, and the overconfidence effect: Are the sadder wiser when predicting future actions and events? *Journal of Personality and Social Psychology, 61*, 521-532.
- Einhorn, H. J. (1972). Expert measurement and mechanical combination. *Organizational Behavior and Human Decision Processes, 7*, 86-106.
- Elstein, A. S., Shulman, L. S., & Sprafka, S. A. (1978). *Medicine problem solving: An analysis of clinical reasoning*. Cambridge, MA: Harvard University Press.
- Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research, 59*, 395-430.

- Ferrell, W. R. (1994). Discrete subjective probabilities and decision analysis. In G. W. P. Ayton (Ed.), *Subjective probability*. (pp. 411-451). New York: Wiley.
- Ferrell, W. R., & McGoey, P. J. (1980). A model of calibration for subjective probabilities. *Organizational Behavior and Human Decision Processes*, 26, 32-53.
- Festinger, L. (1942). A theoretical interpretation of shifts in level of aspiration. *Psychological Review*, 87, 215-251.
- Fischhoff, B., & MacGregor, D. (1981). *Selective confidence in forecasts*. (PTR-1092-81-12). Woodland Hills, CA: Perceptronics.
- Frank, J. D. (1953). Some psychological determinants of the level of aspiration. *American Journal of Psychology*, 47, 285-293.
- Ghatala, E. S., Levin, J. R., Foorman, B. R., & Pressley, M. (1989). Improving children's regulation of their reading PREP time. *Contemporary Educational Psychology*, 14, 49-66.
- Gigerenzer, G., Hoffrage, U., & Kleinbolting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, 98, 506-528.
- Gilovich, T., Kerr, M., & Medvec, V. H. (1993). Effect of temporal perspective on subjective confidence. *Journal of Personality and Social Psychology*, 64, 552-560.
- Glenberg, A. M., & Epstein, W. (1985). Calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 702-718.
- Glenberg, A. M., & Epstein, W. (1987). Inexpert calibration of comprehension. *Memory and Cognition*, 15, 84-93.

- Glenberg, A. M., Sanocki, T., Epstein, W., & Morris, C. (1987). Enhancing calibration of comprehension. *Journal of Experimental Psychology: General*, *116*, 119-136.
- Glenberg, A. M., Wilkinson, A. C., & Epstein, W. (1982). The illusion of knowing: failure in the self-assessment of comprehension. *Memory and Cognition*, *10*, 597-602.
- Green, S. K., & Gross, A. E. (1979). Self-serving biases in implicit evaluations. *Personality and Social Psychology Bulletin*, *5*, 214-217.
- Gregory, W. L., Cialdini, R. B., & Carpenter, K. M. (1982). Self-relevant scenarios as mediators of likelihood estimates and compliance. *Journal of Personality and Social Psychology*, *43*, 89-99.
- Griffin, D., & Tversky, A. (1992). The weighing of evidence and the determinants of confidence. *Cognitive Psychology*, *24*, 411-435.
- Hacker, D. J., Bol, L., Horgan, D., & Rakow, E. (2000). Test prediction and performance in a classroom context. *Journal of Educational Psychology*, *92*, 160-170.
- Hackett, G., & Betz, N. E. (1989). An exploration of the mathematics self-efficacy/mathematics performance correspondence. *Journal for Research in Mathematics Education*, *20*, 261-273.
- Hoch, S. J. (1985). Counterfactual reasoning and accuracy in predicting personal events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 719-731.
- Irwin, F. W. (1944). The realism of expectations. *Psychological Review*, *51*, 120-126.
- Irwin, F. W. (1953). Stated expectations as functions of probability and desirability outcomes. *Journal of Personality*, *21*, 329-335.

- Johnson, J. E. V., & Bruce, A. C. (2001). Calibration of subjective probability judgments in a naturalistic setting. *Organizational Behavior and Human Decision Processes*, 85, 265-290.
- Juslin, P. (1993). An explanation of the hard-easy effect in studies of realism of confidence in one's general knowledge. *European Journal of Cognitive Psychology*, 5, 55-71.
- Juslin, P. (1994). The overconfidence phenomenon as a consequence of informal experimenter-guided selection of almanach items. *Organizational Behavior and Human Decision Processes*, 57, 226-246.
- Kahneman, D., Slovic, P. & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press.
- Keren, G. (1987). Facing uncertainty in the game of bridge: a calibration study. *Organizational Behavior and Human Decision Processes*, 39, 98-114.
- Keren, G. (1997). On the calibration of probability judgements: some critical comments and alternative perspectives. *Journal of Behavioral Decision Making*, 10, 269-278.
- Klayman, J., Soll, J. B., Gonzalez-Vallejo, C., & Barlas, S. (1999). Overconfidence: It depends on how, what, and whom you ask. *Organizational Behavior and Human Decision Processes*, 79, 216-247.
- Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 107-118.
- Kuiper, N. A., Olinger, L. J., MacDonald, M. R., & Shaw, B. F. (1985). Self-schema processing of depressed and nondepressed content: The effects of vulnerability on depression. *Social Cognition*, 3, 77-93.

- Lamarche, L. (1988a). Confiance excessive: croyances et connaissances. *International Journal of Psychology*, 23, 165-180.
- Lamarche, L. (1988b). Évolution de la confiance excessive en fonction du nombre d'indices reçus pour répondre à une question. *Revue Canadienne des Sciences du Comportement*, 20, 265-272.
- Lamarche, L. (1994). "Overconfidence" in person perception. Paper presented at the Third European Small Group Meeting on Social Cognition., Vendôme, France.
- Lamarche, L., Askévis-Leherpeux, F., & Déziel, J. (1996). *Vérifiabilité perçue et temps de réponse*. Paper presented at the European Association of Experimental Social Psychology, Gmunden, Austria.
- Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology*, 32, 311-328.
- Liberman, N., Sagristano, M. D., & Trope, Y. (2002). The effect of temporal distance on level of construal. *Journal of Experimental Social Psychology*, 38, 523-534.
- Liberman, N., & Trope, Y. (1998). The role of feasibility and desirability considerations in near and distant future decisions: A test of temporal construal theory. *Journal of Personality and Social Psychology*, 75, 5-18.
- Liberman, V., & Tversky, A. (1993). On the evaluation of probability judgments: Calibration, resolution, and monotonicity. *Psychological Bulletin*, 114, 162-173.
- Lichtenstein, S., & Fischhoff, B. (1977). Do those who know also know more about how much they know? *Organizational Behavior and Human Decision Processes*, 20, 159-183.

- Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of probabilities: The state of the art to 1980. In P. S. A. T. P. Kahneman (Ed.), *Judgement under uncertainty: Heuristics and biases*. (pp. 306-334). Cambridge, U.K.: Cambridge University Press.
- Lichtenstein, S., & Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology*, *89*, 46-55.
- Lovelace, E. A. (1984). Metamemory: monitoring future recallability during study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 756-766.
- Lundeberg, M. A., Fox, P. W., & Punéochar, J. (1994). Highly confident but wrong: Gender differences and similarities in confidence judgments. *Journal of Educational Psychology*, *86*, 114-121.
- Lusted, L. B. (1977). *A study of the efficiency of diagnostic radiologic procedures: Final report on diagnostic efficiency*. Chicago: Efficiency Study Committee of the American College of Radiology.
- Maki, R. H., & Berry, S. (1984). Metacomprehension of text material. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 663-679.
- Maki, R. H., & Serra, M. (1992a). The basis of test predictions for text materials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 116-126.
- Maki, R. H., & Serra, M. (1992b). Role of practice tests in the accuracy of test predictions on text material. *Journal of Educational Psychology*, *84*, 200-210.
- Marks, R. (1951). The effect of probability, desirability, and "privilege" on the stated expectations of children. *Journal of Personality*, *19*, 332-351.

- Murphy, A. H. (1973). A new vector partition of the probability score. *Journal of Applied Meteorology*, 12, 595-600.
- Murphy, A. H., & Winkler, R. L. (1977). Can weather forecasters formulate reliable probability forecasts of precipitation and temperature? *National Weather Digest*, 2, 2-9.
- Nelson, T. O., & Narens, L. (1990). Metamemory: a theoretical framework and new findings. In G. Bower (Ed.), *The psychology of learning and motivation: advances in research and theory*. (Vol. 26, pp. 125-173). New York: Academic Press.
- Nisan, M. (1972). Dimension of time in relation to choice behavior and achievement orientation. *Journal of Personality and Social Psychology*, 21, 175-182.
- Nisan, M. (1973). Perception of time in lower-class black students. *International Journal of Psychology*, 8, 109-116
- Northcraft, G. B., & Neale, M. A. (1987). Experts, amateurs and real estate: an anchoring-and-adjustment perspective on property pricing decisions. *Organizational Behavior and Human Decision Processes*, 39, 84-97.
- Nussbaum, S., Trope, Y., & Liberman, N. (2003). Creeping dispositionism: the temporal dynamics of behavior prediction. *Journal of Personality and Social Psychology*, 84, 485-497.
- Olsson, N. (2000). A comparison of correlation, calibration, and diagnosticity as measures of the confidence-accuracy relationship in witness identification. *Journal of Applied Psychology*, 85, 504-511.
- Orndorff, R. M. (1996). A comparative study of declared and undeclared college students on career uncertainty and involvement in career

- development activities. *Journal of Counseling and Development*, 74, 632-639.
- Oskamp, S. (1965). Overconfidence in case-study judgments. *Journal of Consulting Psychology*, 29, 261-265.
- Pageau, D., & Bujold, J. (2000). *Dis-moi ce que tu veux et je te dirai jusqu'où tu iras: les caractéristiques des étudiantes et des étudiants à la rescousse de la compréhension de la persévérance aux études. Analyse des données des enquêtes ICOPE. 1er volet: les programmes de baccalauréat*. Montréal: Direction du recensement étudiant et de la recherche institutionnelle. Université du Québec.
- Pratt, M. W., Luszcz, M. A., McKenzie-Keating, S., & Manning, A. (1982). Thinking about stories: the story schema in metacognition. *Journal of Verbal Learning and Verbal Behavior*, 21, 493-505.
- Pressley, M., Snyder, B. L., Levin, J. R., Murray, H. G., & Ghatala, E. S. (1987). Perceived readiness for examination performance (PREP) produced by initial reading of text and text containing adjunct questions. *Reading Research Quarterly*, 22, 219-236.
- Prohaska, V. (1994). "I know I'll get an A": Confident overestimation of final course grades. *Teaching of Psychology*, 21, 141-143.
- Pulford, B. D., & Colman, A. M. (1996). Overconfidence, base rates and outcome positivity/negativity of predicted events. *British Journal of Psychology*, 87, 431-445.
- Radhakrishnan, P., Arrow, H., & Sniezek, J. A. (1996). Hoping, performing, learning, and predicting: Changes in the accuracy of self-evaluations of performance. *Human Performance*, 9, 23-49.
- Read, J. D., Lindsay, D. S., & Nicholls, T. (1998). The relationship between confidence and accuracy in eyewitness identification studies: Is the

- conclusion changing? In C. P. Thompson & D. J. Herrmann & J. D. Read & D. Bruce & D. G. Payne & M. P. Toglia (Eds.), *Eyewitness memory: Theoretical and applied perspectives*. (pp. 107-130). Mahwah NJ: Lawrence Erlbaum.
- Rey, A., & Robert, P. (Eds.). (2001). *Le grand Robert de la langue française* (2 éd.). Paris: Dictionnaires le Robert.
- Ross, L., Lepper, M. R., Strack, F., & Steinmetz, J. (1977). Social explanation and social expectation: Effects of real and hypothetical explanations on subjective likelihood. *Journal of Personality and Social Psychology*, *35*, 817-829.
- Sagristano, M. D., Trope, Y., & Liberman, N. (2002). Time-dependent gambling: odds now, money later. *Journal of Experimental Psychology: General*, *131*, 364-376.
- Schraw, G., & Roedel, T. D. (1994). Test difficulty and judgement bias. *Memory and Cognition*, *22*, 63-69.
- Schraw, G., Dunkle, M. E., Bendixen, L. D., & Roedel, T. D. (1995). Does a general monitoring skill exist? *Journal of Educational Psychology*, *87*, 433-444.
- Shaughnessy, J. J. (1979). Confidence-judgment accuracy as a predictor of test performance. *Journal of Research in Personality*, *13*, 505-514.
- Shaughnessy, J. J. (1981). Memory monitoring accuracy and modifications of rehearsal strategies. *Journal of Verbal Learning and Verbal Behavior*, *20*, 216-230.
- Sherman, S. J., Skov, R. B., Hervitz, E. F., & Stock, C. B. (1981). The effects of explaining hypothetical future events: From possibility to probability to actuality and beyond. *Journal of Experimental Social Psychology*, *17*, 142-158.

- Sjostrom, K. P., & Marks, A. (1994). Pretest and posttest confidence ratings in test performance by low-, medium-, and high-scoring students. *Teaching of Psychology, 21*, 12-16.
- Slovic, P. (1966). Value as a determiner of subjective probability. *IEEE Transactions of Human Factors in Electronics, HFE7*, 22-28.
- Snizek, J. A., Paese, P. W., & Switzer, F. S. I. (1990). The effect of choosing on confidence in choice. *Organizational Behavior and Human Decision Processes, 46*, 264-282.
- Snyder, B. L., & Pressley, M. (1989). *How do adult students use textbook restudy opportunities: they start at the beginning and reread*. Unpublished manuscript.
- Stone, D. N. (1994). Overconfidence in initial self-efficacy judgments: Effects on decision processes and performance. *Organizational Behavior and Human Decision Processes, 59*, 452-474.
- Suantak, L., Bolger, F., & Ferrell, W. R. (1996). The hard-easy effect in subjective probability calibration. *Organizational Behavior and Human Decision Processes, 67*, 201-221.
- Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103*, 193-210.
- Trafimow, D., & Snizek, J. A. (1994). Perceived expertise and its effect on confidence. *Organizational Behavior and Human Decision Processes, 57*, 290-302.
- Vallone, R., Griffin, D., Lin, S., & Ross, L. (1990). Overconfident prediction of future actions and outcomes by self and others. *Journal of Personality and Social Psychology, 58*, 582-592.

- Wagenaar, W. E., & Keren, G. (1985). Calibration of probability assessments by professional blackjack dealers, statistical experts, and lay people. *Organizational Behavior and Human Decision Processes*, 36, 406-416.
- Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, 39, 806-820.
- Wright, G., & Ayton, P. (1988). Immediate and short-term judgmental forecasting: Personologism, Situationism or Interactionism? *Personality and Individual Differences*, 9, 109-120.
- Zakay, D. (1983). The relationship between the probability assessor and the outcomes of an event as a determiner of subjective probability. *Acta Psychologica*, 53, 271-280.
- Zakay, D., & Glicksohn, J. (1992). Overconfidence in a multiple-choice test and its relationship to achievement. *Psychological Record*, 42, 519-524.
- Zelevnik, C., Hojat, M., Goepf, C. E., Amadio, P., Kowlessar, O. D., & Borenstein, B. (1988). Students' certainty during course test-taking and performance on clerkships and board exams. *Journal of Medical Education*, 63, 881-891.

APPENDICE A : QUESTIONNAIRES

Prédictions, détermination et connaissances

académiques et professionnelles

(QPDCAP)

Sigle du cours

(date)

Code permanent _____

Département _____

Collège où vous avez fait vos études _____

Encerclez le programme auquel vous êtes inscrit(e) :

Mineur

Majeur

Baccalauréat

En quelle année êtes-vous?

1^{ière} année 2^{ième} année 3^{ième} année

Consigne

Répondez par oui ou non aux questions qui suivent, puis indiquez votre niveau de confiance à l'aide de l'échelle suivante:

50%

60%

70%

80%

90%

100%

Si vous n'êtes pas du tout certain(e) de votre réponse utilisez le 50%. Cela veut dire que vous avez répondu "oui", mais que vous pensez que la réponse pourrait aussi bien être "non". Si vous êtes tout à fait certain(e) de votre réponse, vous utilisez le 100%. Par ailleurs, il se peut que vous soyez "plus ou moins certain(e)" de votre réponse, à ce moment-là utilisez les valeurs entre 60% et 90%

1. Pensez-vous obtenir un « A » dans le cours que vous suivez en ce moment?

		OUI	NON		
50%	60%	70%	80%	90%	100%

2. Pensez vous étudier au moins trois heures/cours par semaine durant le trimestre?

		OUI	NON		
50%	60%	70%	80%	90%	100%

3. Pensez-vous maintenir toutes vos notes au-dessus de la moyenne durant le trimestre?

		OUI	NON		
50%	60%	70%	80%	90%	100%

4. Pensez-vous obtenir au moins un « A » dans votre cours préféré à ce trimestre-ci?

		OUI	NON		
50%	60%	70%	80%	90%	100%

5. Pensez-vous manquer un jour de classe durant ce trimestre-ci?

		OUI	NON		
50%	60%	70%	80%	90%	100%

6. Pensez-vous abandonner un cours durant ce trimestre?

		OUI	NON		
50%	60%	70%	80%	90%	100%

7. Pensez-vous avoir à demander une révision de note parce qu'un professeur aurait été injuste durant ce trimestre?

		OUI	NON		
50%	60%	70%	80%	90%	100%

8. Pensez-vous avoir que des « A » et « A+ » dans votre bulletin à ce trimestre-ci?

		OUI	NON		
50%	60%	70%	80%	90%	100%

9. Pensez-vous avoir la moyenne suffisante pour poursuivre au trimestre suivant?

		OUI	NON		
50%	60%	70%	80%	90%	100%

10. Pensez-vous trouver un travail l'été prochain en psychologie?

		OUI	NON		
50%	60%	70%	80%	90%	100%

11. Pensez-vous obtenir la moyenne que vous voulez cette année?

		OUI	NON		
50%	60%	70%	80%	90%	100%

12. Pensez-vous avoir la moyenne suffisante pour poursuivre l'année prochaine?

		OUI	NON		
50%	60%	70%	80%	90%	100%

13. Pensez vous obtenir un baccalauréat en psychologie?

		OUI	NON		
50%	60%	70%	80%	90%	100%

14. Pensez-vous avoir un prêt (somme d'argent que vous devez rembourser) du Ministère de l'Éducation durant le baccalauréat?

		OUI	NON		
50%	60%	70%	80%	90%	100%

15. Pensez-vous avoir une bourse (somme d'argent que n'avez pas à rembourser) du Ministère de l'Éducation durant le baccalauréat?

		OUI	NON		
50%	60%	70%	80%	90%	100%

16. Pensez-vous obtenir une assistance de recherche (être payé pour aider un professeur dans sa recherche) pendant le baccalauréat?

		OUI	NON		
50%	60%	70%	80%	90%	100%

17. Pensez-vous participer à une recherche en psychologie en tant que sujet (autre que celle-ci) durant le baccalauréat?

		OUI	NON		
50%	60%	70%	80%	90%	100%

18. Pensez-vous avoir des problèmes pour payer vos frais de scolarité durant le baccalauréat?

		OUI	NON		
50%	60%	70%	80%	90%	100%

19. Pensez-vous être accepté(e) à la maîtrise en psychologie?

		OUI	NON		
50%	60%	70%	80%	90%	100%

20. Pensez-vous obtenir une maîtrise en psychologie?

		OUI	NON		
50%	60%	70%	80%	90%	100%

21. Pensez-vous être accepté(e) au doctorat en psychologie?

		OUI	NON		
50%	60%	70%	80%	90%	100%

22. Pensez-vous obtenir un doctorat en psychologie?

		OUI	NON		
50%	60%	70%	80%	90%	100%

23. Tenez-vous à obtenir au moins un « A » dans le cours que vous suivez en ce moment?

Pas du tout peu plus ou moins beaucoup à tout prix

24. Tenez-vous à obtenir au moins un « A » dans le cours que vous suivez en ce moment?

Pas du tout peu plus ou moins beaucoup à tout prix

25. Tenez-vous à maintenir toutes vos notes au-dessus de la moyenne pour ce trimestre?

Pas du tout peu plus ou moins beaucoup à tout prix

26. Tenez-vous à trouver un travail l'été prochain en psychologie?

Pas du tout peu plus ou moins beaucoup à tout prix

27. Tenez-vous à avoir un prêt (somme d'argent que vous devez rembourser) du Ministère de l'Éducation durant le baccalauréat?

Pas du tout peu plus ou moins beaucoup à tout prix

28. Tenez-vous à avoir une bourse (somme d'argent que n'avez pas à rembourser) du Ministère de l'Éducation durant le baccalauréat?

Pas du tout peu plus ou moins beaucoup à tout prix

29. Tenez-vous à participer à une recherche en psychologie en tant que sujet (autre que celle-ci durant le baccalauréat)?

Pas du tout peu plus ou moins beaucoup à tout prix

30. Tenez-vous à obtenir une assistance de recherche (être payé pour aider un professeur dans sa recherche) pendant le baccalauréat?

Pas du tout peu plus ou moins beaucoup à tout prix

31. Tenez-vous à obtenir un baccalauréat en psychologie?

Pas du tout peu plus ou moins beaucoup à tout prix

32. Tenez-vous à être accepté(e) à la maîtrise en psychologie?

Pas du tout peu plus ou moins beaucoup à tout prix

33. Tenez-vous à obtenir une maîtrise en psychologie?

Pas du tout peu plus ou moins beaucoup à tout prix

34. Tenez-vous à être accepté au doctorat en psychologie?

Pas du tout peu plus ou moins beaucoup à tout prix

35. Tenez-vous à obtenir un doctorat en psychologie?

Pas du tout peu plus ou moins beaucoup à tout prix

36. Pensez-vous obtenir un emploi en psychologie?

		OUI	NON		
50%	60%	70%	80%	90%	100%

37. Pensez-vous devenir membre de l'Ordre des psychologues du Québec?

		OUI	NON		
50%	60%	70%	80%	90%	100%

38. Quel revenu pensez-vous obtenir en tant que psychologue?

15 000 – 25 000	25 000 – 35 000	35 000 – 45 000
45 000 – 55 000	55 000 – 65 000	65 000 et plus

39. Tenez-vous à obtenir un emploi en psychologie?

Pas du tout peu plus ou moins beaucoup à tout prix

40. Tenez-vous à devenir membre de l'Ordre des psychologues du Québec?

Pas du tout peu plus ou moins beaucoup à tout prix

41. Quel revenu voudriez-vous obtenir en tant que psychologue?

15 000 – 25 000	25 000 – 35 000	35 000 – 45 000
45 000 – 55 000	55 000 – 65 000	65 000 et plus

42. Combien d'étudiants sont admis au baccalauréat à chaque année?

125 – 150 151 – 175 176 – 200 201 – 225

43. Combien d'étudiants sont admis à la maîtrise à chaque année?

5 – 25 26 – 45 46 – 65 66–85

44. Combien d'étudiants sont admis au doctorat à chaque année?

25 – 35 36 – 45 46 - 55 56 – 65

45. Combien d'étudiants environ obtiennent leur baccalauréat à chaque année?

75 – 100 100 – 125 125 – 150 150 – 175

46. Combien d'étudiants environ obtiennent leur maîtrise à chaque année?

25– 50 51 –75 76–100 101–125

47. Combien d'étudiants environ obtiennent leur doctorat à chaque année?

5 – 15 15 – 25 25 – 35 35 – 45

48. Combien de nouveaux membres sont admis à L'Ordre des psychologues du Québec à chaque année?

100 et moins 101 – 125 126 – 150 151 et plus

49. Combien de psychologues y a-t-il au Québec?

4500 – 5000 5001 – 5500 5501 – 6000 6001 – 6500

50. Quel est le salaire moyen d'un psychologue au Québec?

15 000 – 25 000 25 000 – 35 000 35 000 – 45 000

45 000 – 55 000 55 000 – 65 000 65 000 et plus

Commentaires (optionnel) :

Perception de l'examen

(QPE)

Sigle du cours

Date de l'examen intra-trimestriel ou final

Code permanent : _____

Code de l'examen : _____

Difficulté de l'examen :

0 1 2 3 4 5 6 7 8 9 10

Facile

Difficile

L'examen est représentatif de la matière du cours :

0 1 2 3 4 5 6 7 8 9 10

Pas représentatif

Représentatif

La note que je m'attends à avoir : _____ (lettre)

La note que je visais : _____ (lettre)

Réalisations académiques

(QRA)

Sigle du cours

(date)

Code permanent : _____

Consigne

Répondez par oui ou non aux questions qui suivent

1. Avez-vous obtenu un « A » dans le cours que vous suivez en ce moment?

OUI

NON

2. Avez-vous étudié au moins trois heures/cours par semaine durant le trimestre?

OUI

NON

3. Avez-vous maintenu toutes vos notes au-dessus de la moyenne durant le trimestre (pour les examens intra)?

OUI

NON

4. Avez-vous obtenu au moins un « A » (examen intra) dans votre cours préféré à ce trimestre-ci?

OUI NON

5. Avez-vous manqué un jour de classe durant ce trimestre-ci?

OUI NON

6. Avez-vous abandonné un cours durant ce trimestre?

OUI NON

7. Avez-vous eu à demander une révision de note parce qu'un professeur aurait été injuste durant ce trimestre?

OUI NON

8. Avez-vous eu que des « A » et « A+ » dans votre bulletin à ce trimestre-ci?

OUI NON

9. Avez-vous eu la moyenne suffisante pour poursuivre au trimestre suivant?

OUI NON

10. Avez-vous trouvé un travail l'été prochain en psychologie?

OUI NON

11. Avez-vous obtenu la moyenne que vous voulez cette année?

OUI

NON

12. Avez-vous eu la moyenne suffisante pour poursuivre l'année prochaine?

OUI

NON

13. Avez vous obtenu un baccalauréat en psychologie?

OUI

NON

14. Avez-vous eu un prêt (somme d'argent que vous devez rembourser) du Ministère de l'Éducation durant le baccalauréat?

OUI

NON

15. Avez-vous eu une bourse (somme d'argent que n'avez pas à rembourser) du Ministère de l'Éducation durant le baccalauréat?

OUI

NON

16. Avez-vous obtenu une assistance de recherche (être payé pour aider un professeur dans sa recherche) pendant le baccalauréat?

OUI

NON

17. Avez-vous participé à une recherche en psychologie en tant que sujet (autre que celle-ci) durant le baccalauréat?

OUI

NON

18. Avez-vous eu des problèmes pour payer vos frais de scolarité durant le baccalauréat?

OUI

NON

Commentaires (optionnel) :

APPENDICE B : *FIGURE 6*

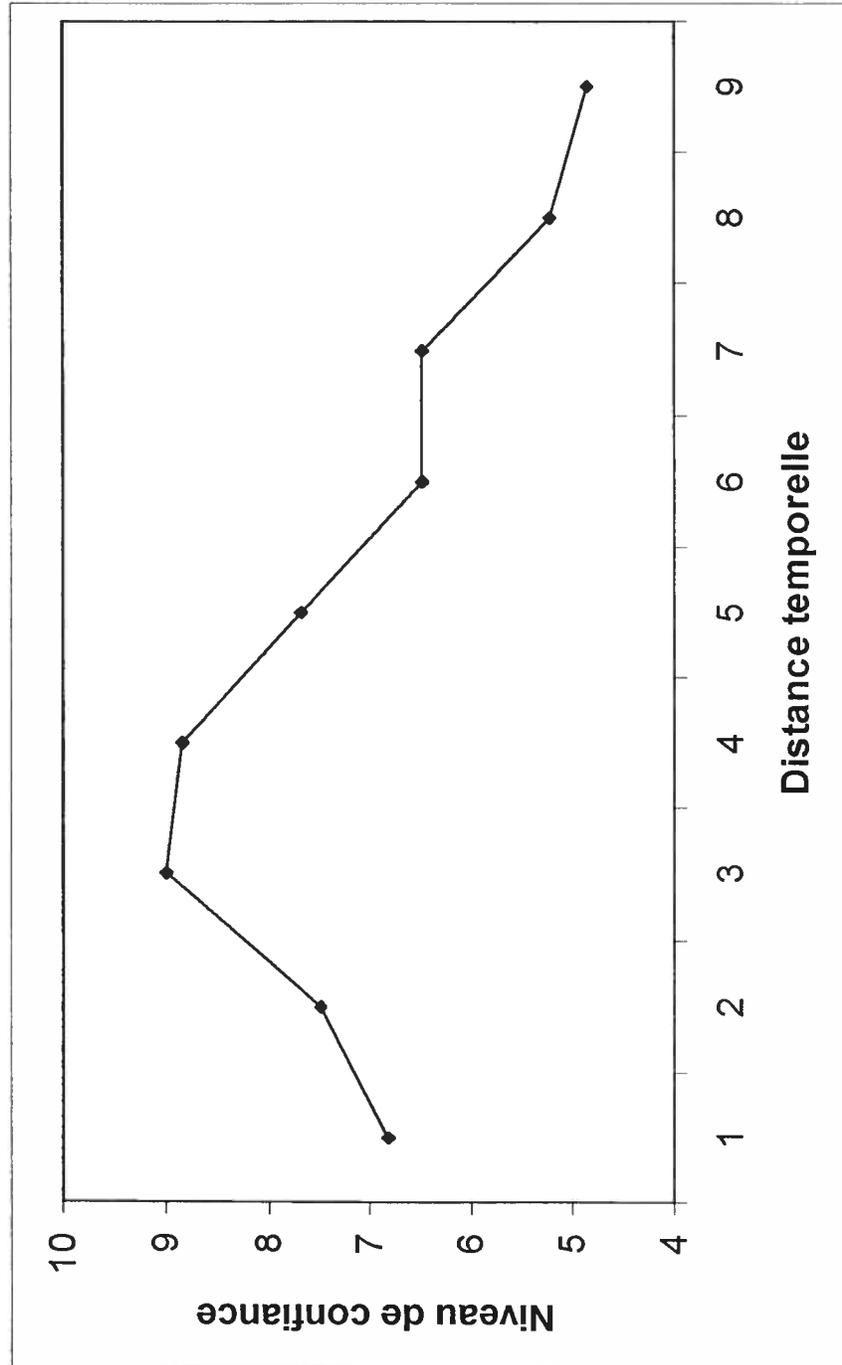


FIGURE 6. Niveau de confiance en fonction de la distance temporelle à la suite du changement de la séquence.