

UNIVERSITÉ DE SHERBROOKE

Faculté d'éducation

Étude de l'équivalence des modalités Internet et papier-crayon pour l'évaluation
psychométrique de la personnalité

par

Véronique Gingras

Mémoire présenté à la Faculté d'éducation

en vue de l'obtention du grade de

Maître ès sciences (M.Sc.)

Maîtrise en orientation

Août 2014

© Véronique Gingras, 2014

UNIVERSITÉ DE SHERBROOKE

Faculté d'éducation

Étude de l'équivalence des modalités Internet et papier-crayon pour l'évaluation
psychométrique de la personnalité

Véronique Gingras

A été évalué par un jury composé des personnes suivantes :

Monsieur Sylvain Bourdon	Président du jury
Monsieur Yann Le Corff	Directeur de recherche
Monsieur Éric Yergeau	Autre membre du jury

Mémoire accepté le 14 août 2014

SOMMAIRE

De nos jours, les conseillères et conseillers d'orientation sont nombreux dans leurs différents secteurs d'emplois à utiliser la psychométrie comme outil de travail complémentaire à leurs activités de counseling. L'arrivée d'Internet et de ses nombreuses possibilités a considérablement modifié l'usage traditionnel de la psychométrie en format papier-crayon sous supervision d'un administrateur. En effet, Internet a été très rapidement accueilli comme modalité d'administration par les professionnels de l'orientation en raison de ses multiples avantages économiques et pratiques. Des risques sur le plan de la qualité de la mesure sont cependant associés à cette pratique, puisque la plupart des tests sont simplement transférés de papier-crayon à Internet, sans validation préalable. De plus, Internet permet une administration à distance sans supervision ce qui contrevient à plusieurs normes de pratique établies afin d'assurer un usage valide des tests.

Un nombre important d'études empiriques ont été réalisées afin de documenter l'impact d'un changement de modalité d'administration pour un test psychométrique. Leur recension a toutefois mis en lumière des résultats contradictoires. De plus, plusieurs d'entre elles présentaient des faiblesses méthodologiques suffisantes pour justifier la réalisation d'une nouvelle étude utilisant un devis à mesures répétées. Dans le cadre d'une recherche source sur l'évaluation de la personnalité à l'ère d'Internet, nous avons étudié l'équivalence des distributions de scores et de la fidélité. L'Inventaire de personnalité Le Corff a été administré à 407 personnes étudiantes de l'Université de Sherbrooke en format Internet et papier-crayon en inversant l'ordre de passation pour la moitié de l'échantillon.

Pour le premier objectif de recherche portant sur l'équivalence des distributions de scores, un test t pour échantillons appariés a été réalisé afin de détecter la présence de différences significatives pour les cinq échelles de l'inventaire. Quatre échelles ont obtenu une différence significative, mais toutes ont eu une taille d'effet jugée « faible ». Les indices d'asymétrie et d'aplatissement n'ont obtenu aucune différence significative. Le deuxième

objectif de recherche sur la fidélité a été complété à l'aide du test de Feldt (1980) pour échantillons dépendants. Le calcul d'un score t pour chaque paire d'échelles a permis de détecter empiriquement des différences significatives pour deux échelles avec un intervalle de confiance de 95%.

Les résultats ont permis de répondre à l'objectif général de recherche sur la vérification de l'équivalence d'un test, suite à un changement de modalité d'administration. La présence de faibles différences significatives entre les distributions de scores soutient l'utilisation interchangeable des formats papier-crayon ou Internet d'un instrument mesurant la personnalité. Le devis à mesures répétées croisé est une force du projet. En contrepartie, l'âge et la composition de l'échantillon restreignent la généralisabilité des résultats.

TABLE DES MATIÈRES

SOMMAIRE	4
LISTE DES TABLEAUX.....	8
REMERCIEMENTS	9
INTRODUCTION	10
PREMIER CHAPITRE – PROBLÉMATIQUE.....	11
1. UTILISATION DE LA PSYCHOMÉTRIE EN ORIENTATION.....	11
1.1 Définition de la psychométrie	13
1.2 Définition des tests psychométriques	15
2. ARRIVÉE D’INTERNET ET CHANGEMENT DE MODALITÉ D’ADMINISTRATION	16
2.1 Avantages de l’administration par Internet	17
2.2 Avantages liés à la réduction des coûts.....	18
2.3 Avantages liés à l’économie de temps	19
2.4 Avantages liés à l’aspect pratique	19
3. RISQUES DE L’ADMINISTRATION DE TESTS PAR INTERNET.....	20
3.1 Risques liés à la validité.....	21
3.2 Risques liés aux normes.....	21
3.3 Risques liés à la fidélité	22
3.4 Augmentation du testing internet non supervisé et ses risques	23
4. RÉGLEMENTATION ENCADRANT LE TESTING INTERNET	24
4.1 Normes de pratique en matière de testing.....	25
4.2 Recommandations de la Commission internationale des tests	25
4.3 Implication des différentes réglementations	26
5. SYNTHÈSE	26
6. QUESTION GÉNÉRALE DE RECHERCHE	27
DEUXIÈME CHAPITRE– RECENSION DES ÉCRITS	30
1. MÉTHODOLOGIE DE LA RECENSION	30
2. RÉSULTATS DE LA RECENSION.....	31
2.1 Changement de modalité d’administration de papier-crayon à Internet sous supervision.....	32
2.2 Changement de modalité d’administration de papier-crayon à Internet sans supervision.....	36
3. SYNTHÈSE	47
4. IDENTIFICATION DU BESOIN DE RECHERCHE.....	49

5. OBJECTIFS DE RECHERCHE.....	50
TROISIÈME CHAPITRE– MÉTHODOLOGIE.....	51
1. DEVIS DE RECHERCHE.....	51
2. ÉCHANTILLON	51
3. INSTRUMENTS DE MESURE.....	52
4. PROCÉDURE.....	53
5. PLAN D’ANALYSE	54
6. CONSIDÉRATIONS ÉTHIQUES	55
QUATRIÈME CHAPITRE– PRÉSENTATION DES RÉSULTATS	56
1. DESCRIPTION DE L’ÉCHANTILLON	56
2. ÉQUIVALENCE DES DISTRIBUTIONS DES SCORES.....	57
3. VÉRIFICATION DE LA FIDÉLITÉ	60
CINQUIÈME CHAPITRE – DISCUSSION DES RÉSULTATS.....	62
1. RÉPONSES AUX OBJECTIFS DE RECHERCHE	62
1.1 Premier objectif de recherche	62
1.2 Deuxième objectif de recherche	63
2. PISTES THÉORIQUES EXPLICATIVES	63
2.1 L’influence du format	63
2.2 Le niveau de supervision	64
2.3 La perception de la confidentialité	64
3. COMPARAISON DES RÉSULTATS AUX ÉTUDES RECENSÉES.....	64
4. IMPLICATIONS DES RÉSULTATS	66
5. FORCES	66
6. LIMITES.....	67
CONCLUSION.....	69
RÉFÉRENCES BIBLIOGRAPHIQUES.....	70
ANNEXE A - QUESTIONNAIRES UTILISÉS DANS LE CADRE DU	
MÉMOIRE	75
ANNEXE B - FORMULAIRE DE CONSENTEMENT- PAPIER	82
ANNEXE C - FORMULAIRE DE CONSENTEMENT - INTERNET	85
ANNEXE D - ATTESTATION DE CONFORMITÉ ÉTHIQUE	87

LISTE DES TABLEAUX

Tableau 1- Composition de l'échantillon	57
Tableau 2 - Comparaison des scores moyens des échelles de l'IPLC	58
Tableau 3 - Comparaison des indices d'asymétrie et d'aplatissement.....	59
Tableau 4 – Coefficients de cohérence interne	61

REMERCIEMENTS

Parce que j'ai eu la chance de bénéficier d'une tonne de soutien, je tiens à manifester ma reconnaissance envers les personnes concernées. Tout d'abord, un merci à Yann Le Corff qui a été un directeur de recherche efficace et patient tout au long des dédales de ce projet. Merci d'avoir facilité par bien des façons, la réalisation de ce mémoire. Un merci également aux professeurs Sylvain Bourdon et Éric Yergeau pour avoir accepté de faire partie de mon jury et pour leurs judicieux conseils et encouragements.

Je désire souligner l'appui financier du Conseil de recherche en sciences humaines et le Fonds québécois de recherche-Société et culture qui m'a permis de consacrer davantage de temps et d'énergie à ce projet.

Merci à des personnes si formidables qu'il m'est difficile de les qualifier. Audrey, Claude-Marie et Marie-Michèle, merci mille fois pour votre folie, complicité et le bonheur que ça procure. Une mention spéciale à mon amie Geneviève dont l'objectivité et la vivacité d'esprit ne connaissent pas la distance!

Johanne et Robert, merci d'avoir suivi patiemment chacune des étapes de ce parcours universitaire et d'avoir été inspirants par votre ardeur et votre énergie. Maxime, Christine et Marie-Pier, avec une telle fratrie, il est difficile de ne pas mettre pas la barre haute, merci d'être les personnes que vous êtes.

INTRODUCTION

La psychométrie a subi des changements à de nombreuses reprises, que ce soit en ce qui a trait à son contexte d'utilisation, ses instruments, les construits mesurés ou le type de professionnels en faisant usage. Depuis quelques années, les conseillères et conseillers d'orientation administrent de plus en plus les tests par Internet pour plusieurs raisons d'ordre pratique et économique. Ce changement enfreint des normes de pratiques instaurées, sans qu'il existe un consensus scientifique sur les effets d'un changement de modalité d'administration, de papier-crayon à Internet. La psychométrie étant une mesure perfectible d'un construit psychologique, il y a nécessité d'en assurer les propriétés telles la validité et la fidélité. Ce mémoire s'intéresse au changement de modalité d'administration, de papier-crayon à Internet, des tests psychométriques par les conseillers et conseillères d'orientation.

Le premier chapitre présentera la situation actuelle du testing, ses avantages et ses risques associés, les particularités du testing non supervisé et l'implication des différentes réglementations régissant l'usage de la psychométrie. La recension d'écrits constituant le deuxième chapitre dressera le portrait de l'état des connaissances actuelles. Ceci conduira à l'énoncé des objectifs de recherche. Les aspects méthodologiques et éthiques seront présentés dans le troisième chapitre. Enfin, les résultats issus des analyses statistiques effectuées et leur interprétation en lien avec la question de recherche posée constitueront respectivement les quatrième et cinquième chapitres.

PREMIER CHAPITRE – PROBLÉMATIQUE

Le présent chapitre discute de l'administration d'outils psychométriques par Internet. On y expose tout d'abord la place de la psychométrie dans la pratique des professionnels qui l'utilisent, particulièrement les conseillers et conseillères d'orientation. Avant de retracer l'historique de l'arrivée d'Internet dans le domaine de l'évaluation psychométrique, les termes et concepts centraux de la discipline sont définis. Ensuite, le phénomène du changement de modalité d'administration de tests, de papier-crayon à Internet, est décortiqué de manière à en faire ressortir les avantages, de même que les risques qui y sont associés sur le plan de la qualité de la mesure. Enfin, il est question des diverses instances qui règlementent et encadrent les pratiques de testing, incluant l'utilisation des technologies. Cette présentation permet à la fin du chapitre d'énoncer la question générale de recherche. L'étude de cette problématique permettra de documenter les effets du changement de modalité d'administration sur les propriétés des tests.

1. UTILISATION DE LA PSYCHOMÉTRIE EN ORIENTATION

Les conseillers d'orientation sont, avec les psychologues, les professionnels les plus habilités à la psychométrie dans le cadre de leur pratique professionnelle. Leur formation initiale leur fournit les connaissances requises pour l'administration, la comptabilisation et l'interprétation de tests psychométriques relativement à leurs interventions et leur mandat: « [l]e conseiller d'orientation détient des connaissances portant sur les théories psychologiques (développement normal et psychopathologie), ainsi que sur la psychométrie, l'évaluation des personnes et les instruments requis » (Gouvernement du Québec, 2005, p.41).

Dans le cadre de leurs fonctions, la psychométrie leur est utile entre autres, afin d'évaluer le fonctionnement psychologique de leurs clients, d'aider à la

connaissance et la compréhension de soi et de faciliter la communication et l'expression de soi (Guédon, Savard, Le Corff et Yergeau, 2011). L'évaluation au moyen de la psychométrie ne sert donc plus seulement à valider un choix de formation ou de carrière, comme ce fut le cas lors des premières décennies d'utilisation en orientation. La plupart des contextes de counseling de carrière peuvent désormais bénéficier de la psychométrie en raison d'un rôle plus large qui comprend notamment la connaissance de soi, le développement personnel et l'étude du réseau social (Anastasi, 1994).

Dans le *Profil des compétences générales des conseillers d'orientation* adopté en 2004 par l'Ordre des conseillers et conseillères d'orientation et psychoéducateurs et psychoéducatrices du Québec (OCCOPPQ)¹, la psychométrie fait partie de la gamme d'outils et d'interventions mentionnés pour mettre en place la compétence générale « Évaluer la situation de manière rigoureuse ». L'apport de la psychométrie en rapport à cette compétence est également reconnu dans le *Guide d'évaluation en orientation* (OCCOQ, 2010) qui sert de référence en la matière pour les membres de l'Ordre. La capacité à contribuer significativement à la compréhension d'une situation donnée en fournissant un caractère objectif au processus évaluatif est soulignée en tant qu'attribut distinctif de la psychométrie (*Ibid.*). Le recours à ce genre d'outils serait également un des seuls moyens disponibles pour avoir accès à certaines informations chez la personne (*Ibid.*). Guédon *et al.* (2011) mentionnent, par exemple, que les résultats de test peuvent fournir des informations sur le client dont lui-même n'a pas conscience.

Compte tenu des nombreux contextes d'utilisation possibles et de la reconnaissance attribuée par l'OCCOQ dans ses documents institutionnels, il n'est pas étonnant de constater l'utilisation répandue de la psychométrie par les conseillers et conseillères d'orientation québécois. Selon l'*Enquête sur les*

¹ Depuis le 8 décembre 2010, les psychoéducateurs et les conseillers d'orientation font partie de deux ordres distincts. L'Ordre des conseillers et conseillères d'orientation du Québec (OCCOQ) a toutefois conservé tel quel le profil général des compétences de même que d'autres documents institutionnels adoptés précédemment.

pratiques en matière d'évaluation psychométrique par les conseillers et conseillères d'orientation (Le Corff, Yergeau et Savard, 2011), 32 % de ces professionnels rapportent utiliser « souvent » un test psychométrique au cours d'un processus d'évaluation alors que 15 % disent l'utiliser de manière systématique (« à chaque fois»). Seulement 7 % affirment ne jamais avoir recours à la psychométrie dans le cadre de leur pratique professionnelle (*Ibid.*). Ces résultats corroborent l'affirmation, tirée d'un article d'une édition du magazine *En pratique* portant sur l'évaluation, qui soulignait que « l'utilisation des tests psychométriques, après avoir suscité de nombreuses réactions au cours des différentes périodes de l'histoire, semble généralement reconnue aujourd'hui comme partie intégrante de la pratique de conseiller d'orientation » (Delisle-Laberge et Locas, 2004, p.5).

En somme, la psychométrie fait partie des outils dont disposent les conseillers et conseillères d'orientation afin de réaliser leurs activités d'évaluation qui se situent au cœur de leur pratique professionnelle, notamment en étant appropriée pour favoriser la connaissance, l'expression et la compréhension de soi chez le client. Selon les données disponibles, ils sont nombreux à y avoir recours de manière régulière.

1.1 Définition de la psychométrie

Avant d'étayer davantage la problématique, il convient de définir tout d'abord ce qu'est la psychométrie, soit la science de la mesure des concepts psychologiques (Furr et Bacharach, 2008). Ceci nous introduit au principe de mesure en psychologie qui diffère du concept de mesure scientifique traditionnelle. Bernaud (2007) explique cette différence inévitable avec les sciences naturelles en raison des sujets respectifs de chaque discipline : « En effet, mesurer revient fondamentalement à évaluer une grandeur. Si cette opération est objectivable dans le cadre d'une distance ou d'une température, elle prend un sens différent dans le domaine des sciences de l'homme » (p.10). Étant donné que plusieurs concepts psychologiques n'ont pas d'unité de mesure proprement dite, la

mesure de ceux-ci est relative. Ainsi, la psychométrie n'a pas recours à une unité de mesure fixe, mais plutôt à un étalonnage d'indicateurs qui permet de situer l'individu en rapport à un groupe de personnes composant l'échantillon représentatif (*Ibid.*). Fortin (2010) désigne cette première dimension de la mesure « normative » en raison de l'échantillon qui permet d'établir des normes auxquelles les scores peuvent être comparés et interprétés.

Le caractère singulier de la mesure en psychométrie introduit une deuxième dimension qui est appelée « critériée » (*Ibid.*). Celle-ci fait appel à la notion de critère pour mesurer un construit psychologique. C'est l'observation d'indicateurs (ex. : comportements, épreuves), sur la base théorique et non par rapport à une norme observée, qui rendent compte d'un concept sous-jacent, tel que les intérêts professionnels ou la tendance à l'anxiété. Pour cette raison, Laveault et Grégoire (2002) affirment que la mesure, en éducation comme en psychologie, repose sur des estimations.

Les auteurs sont nombreux à fournir leur définition personnelle de la psychométrie. Anastasi (1994) décrit cette discipline comme la « mesure objective et standardisée d'un échantillon de comportements » (p.26) dans le but d'évaluer un concept psychologique chez une personne. Des caractéristiques d'objectivité et de standardisation sont établies comme conditions nécessaires par l'auteure pour distinguer ce qui relève ou non du domaine de la psychométrie. Celles-ci sont réitérées par différents auteurs (Bernaud, 2007; Domino et Domino, 2006). L'objectivité réfère à la caractéristique d'un test de dépendre uniquement de la performance du répondant et de ne pas être influencé par des variables externes, telles que l'examineur (Anastasi, 1994). Bernaud (2007) l'associe à l'approche quantitative de l'investigation psychologique pour qui ce serait gage d'objectivité. La standardisation réfère plutôt au principe d'uniformité de la mesure, lequel peut être obtenu par l'établissement de procédures d'administration et d'interprétation précises et rigoureuses (Anastasi, 1994).

D'autres énoncent que le but fondamental de ce type d'évaluation est de mesurer les différences entre les individus, en se basant sur le concept de la représentation théorique qui stipule que les différences entre les individus d'un échantillon sont représentatives de la variation interindividuelle d'une population (Furr et Bacharach, 2008; Kaplan et Saccuzzo, 1997). Bernaud (2007) ajoute, aux caractéristiques d'objectivité et de standardisation, la nécessité d'une référence à un cadre théorique et la démonstration de la qualité de la mesure par le biais de propriétés psychométriques, telles la validité, la fidélité et la sensibilité et la spécificité, pour qu'une évaluation soit considérée comme relevant de la psychométrie.

1.2 Définition des tests psychométriques

La définition d'un test psychométrique² de Cronbach (1990), une figure importante du développement de la discipline, est toujours considérée d'actualité : « *A test is a systematic procedure for observing behavior and describing it with the aid of numerical scales or fixed categories* » (p.32). On retrouve dans cet énoncé les attributs de tests psychométriques énumérés dans la section précédente.

Les *Normes de pratique de testing en psychologie et en éducation* (American Educational Research Association, American Psychological Association, National Council on Measurement and Education, 2003) présentent une description qui reprend essentiellement les critères énoncés dans la section précédente au sujet de la psychométrie : « Un test est un instrument ou un processus d'évaluation dans lequel un échantillon de comportement d'une personne, dans un domaine précis, est obtenu et, par la suite, évalué et corrigé en utilisant un processus standardisé » (p.3). On mentionne également les étapes subséquentes à toute passation, soit la comptabilisation et l'interprétation des scores. Sans celles-ci, les scores sont dénués de sens, car on ne leur a pas attribué

² Le terme « test psychométrique » a été préféré à celui d'« outil psychométrique » tout au long du texte par souci de clarté. Il prend toutefois un sens générique et désigne l'ensemble des méthodes disponibles d'évaluation psychométrique (inventaire, épreuve, questionnaire, échelle, etc.)

une valeur et mis en contexte. C'est au professionnel habileté à le faire que revient cette tâche. Enfin, le terme « testing » est utilisé en référence au processus général d'administration d'un test à quelqu'un : on parle alors d'une pratique de testing (Furr et Bacharach, 2008).

2. ARRIVÉE D'INTERNET ET CHANGEMENT DE MODALITÉ D'ADMINISTRATION

Depuis sa création, Internet a entraîné un grand nombre de changements et plusieurs sphères d'activité en ont bénéficié en raison de ses multiples possibilités. Le domaine de l'évaluation psychométrique n'y fait pas exception en tirant profit de cette évolution technologique. En fait, la progression d'Internet et ses conséquences seraient telles que cela aurait modifié de manière radicale le monde du testing appliqué aux personnes (Jones et Dages, 2003). Traditionnellement, la grande majorité des tests psychométriques étaient conçus de manière à être administrés dans un format papier-crayon (Domino et Domino, 2006). Mais rapidement, les capacités technologiques se sont raffinées et les concepteurs de test ont commencé à s'intéresser aux possibilités de l'administration de tests psychométriques par ordinateur (Anastasi, 1994). Quelques tests ont d'abord été transposés dans un format informatisé; les répondants devaient entrer leurs réponses sur l'ordinateur qui les comptabilisait ensuite. On parlait alors de testing informatisé (*computer-based testing*). Puis, le développement informatique s'est poursuivi et l'arrivée d'Internet a alimenté un intérêt toujours croissant pour le virage technologique.

Les éditeurs de tests sont désormais nombreux à offrir la passation de tests reconnus par le biais d'Internet. L'utilisation d'Internet, en matière d'évaluation psychométrique, tend à prendre de plus en plus de place dans la pratique professionnelle (Grégoire, 2010). À cet effet, l'enquête effectuée auprès des membres de l'OCCOQ révèle que 68 % d'entre eux utilisent Internet pour l'administration de tests psychométriques dans le cadre de leurs fonctions (Le Corff *et al.*, 2011). Lors d'une étude similaire auprès d'une partie de cette

population en 2004³, la proportion de répondants rapportant ce genre de pratiques était alors seulement de 11 %. L'échantillon de conseillers et conseillères d'orientation rapporte également une préférence pour ce mode d'administration (40 % d'entre eux), alors que 27 % affirment préférer l'administration traditionnelle par papier crayon et 32 % sont indifférents (*Ibid.*). Ces données suggèrent un important changement, sur une courte période de temps, en ce qui a trait à la modalité d'administration utilisée et privilégiée par les utilisateurs de tests psychométriques.

Selon les écrits consultés (Barak, 1999; Hirai, Vernon, Clum, et Skidmore, 2011; Ployhart, Weekley, Holtz et Kemp, 2003; Potosky et Bobko, 2004), il semble que le phénomène soit généralisé dans les autres pays utilisant le testing, où on remarque sensiblement la même tendance constante et marquée vers l'administration par Internet. Les professionnels œuvrant en sélection de personnel font aussi grand usage du testing Internet. En fait, la prédiction de Buchanan et Smith (1999)⁴ quant à l'essor imminent du testing Internet aux fins de sélection de personnel s'avéra juste.

Ainsi, chez les conseillers et conseillères d'orientation québécois, de même que chez les autres utilisateurs d'évaluation psychométrique ailleurs dans le monde, cette pratique semble avoir été adoptée par une majorité de professionnels, car on remarque une augmentation constante et significative de l'utilisation d'Internet. Les tests sont toujours traditionnellement développés dans un format papier-crayon. Cependant, on peut penser que cette tendance se renversera en faveur de tests initialement mis sur pied sur Internet.

2.1 Avantages de l'administration par Internet

³ OCCOPPQ. (2004). Questionnaire sur l'inventaire des pratiques des conseillers et conseillères d'orientation en employabilité et en réadaptation, Document inédit.

OCCOPPQ. (2004). Questionnaire sur l'inventaire des pratiques des conseillers et conseillères d'orientation en éducation. Document inédit.

⁴ « The next ten years will probably be the decade of the Internet in personnel selection » (p.194).

Ce changement remarquable dans les pratiques d'administration de tests peut s'expliquer en raison des nombreux avantages qui sont rattachés à l'utilisation d'Internet (Grégoire, 2010). En fait, toutes les parties impliquées dans le processus de testing, les professionnels qui l'utilisent, les gestionnaires d'entreprises, les éditeurs de tests et même les clients, peuvent profiter des possibilités qu'offre l'administration à partir d'Internet. Cette particularité a probablement à voir avec le développement remarquablement rapide de cette sphère d'activité. Les avantages peuvent être regroupés selon trois grandes catégories, soit la réduction des coûts, l'économie de temps et leur aspect pratique.

2.2 Avantages liés à la réduction des coûts

En utilisant Internet comme modalité d'administration de tests, les entreprises offrant des services professionnels d'évaluation psychométrique économisent de l'argent en réduisant le matériel de test à acheter. Bien sûr, la passation Internet demeure payante, mais les coûts seraient moins élevés qu'en format papier-crayon (Guédon *et al.*, 2011). De plus, puisque le test peut être complété par le client chez lui, un administrateur n'est plus requis et des économies dans les ressources humaines peuvent être réalisées. Plusieurs auteurs n'hésitent pas à citer la réduction des coûts comme avantage contribuant à la popularité du testing Internet (Barak, 1999; Hirai *et al.*, 2011; Kaplan et Saccuzzo, 1997; Nezami et Butcher, 2000; Ployhart *et al.*, 2003). Dans le champ de la sélection de personnel, compte tenu de leur contexte particulier et de leurs objectifs ciblés qui diffèrent des motifs d'utilisation en counseling de carrière, la réduction des coûts est un attrait important du recours au testing Internet.

2.3 Avantages liés à l'économie de temps

L'un des bénéfices importants associés au testing Internet est l'économie de temps que cela procure à l'administrateur. Ce dernier n'a pas à assurer la surveillance et l'assistance de la passation, le répondant pouvant compléter le test chez lui ou sur place, mais sans la nécessité d'un administrateur qui assurerait la standardisation de la passation. Par ailleurs, il y a une économie de temps dû à la comptabilisation qui se fait de manière informatisée. Dès que le test est complété, le conseiller peut avoir accès aux résultats via un serveur informatique. Les délais de cotation peuvent alors être beaucoup plus courts qu'en format papier-crayon où un envoi à l'éditeur peut être nécessaire et prendre quelques jours (Hirai *et al.*, 2011). « La rapidité avec laquelle les analyses de données et la correction de tests pouvaient être effectuées » a d'ailleurs été un des premiers motifs pour l'utilisation de l'ordinateur en testing, bien avant qu'Internet ne soit développé (Anastasi, 1994, p.85).

2.4 Avantages liés à l'aspect pratique

Un autre avantage recensé dans la littérature consultée est le pragmatisme du testing Internet. En plus de la cotation, certains tests offrent des rapports d'interprétation informatisés. Selon les réponses fournies par le client, l'ordinateur les compile et leur associe des éléments de profil pour constituer une interprétation textuelle du test. Non seulement la pondération des résultats peut être rapide, mais les rétroactions peuvent être immédiates (Hirai *et al.*, 2011).

Comparativement aux méthodes traditionnelles, certains auteurs soulignent que le testing par ordinateur permet une meilleure standardisation d'administration (Kaplan et Saccuzo, 1997; Nezami et Butcher, 2000; Tippins, 2009). Toutefois, paradoxalement, d'autres indiquent que le testing Internet procure l'avantage d'offrir une plus grande flexibilité d'administration puisque le test peut être complété selon le lieu et le moment désiré (Ployhart *et al.*, 2003). Les limites géographiques pourraient donc être contrées par le testing Internet qui

permet de joindre des répondants plus éloignés (Grégoire, 2010; Hirai *et al.*, 2011; Tippins, 2009).

Enfin, un avantage singulier signalé par certains auteurs concerne la façon de répondre aux items. Ce type de passation pourrait accroître le niveau d'autodévoilement des personnes, ce qui se solderait en des interprétations plus fidèles aux personnes évaluées (Hirai *et al.*, 2011; Kaplan et Saccuzzo, 1997).

3. RISQUES DE L'ADMINISTRATION DE TESTS PAR INTERNET

Malgré tous les avantages reliés à l'administration de tests sur Internet énumérés ci-dessus, certaines précautions demeurent à considérer et à étudier davantage, en raison des préjudices qu'un test mal construit et présentant des lacunes métrologiques pourrait causer envers les répondants. En effet, les conséquences possibles d'un rapport d'interprétation erroné sont variées, étant donné que « les risques liés aux outils en ligne dépendent en partie du motif pour lequel ils sont utilisés, et du contexte à l'intérieur duquel ils sont administrés (Grégoire, 2010, p.15). Actuellement, la vaste majorité des tests psychométriques que l'on retrouve sous format Internet sont simplement transférés de façon informatisée sans que l'on s'assure de l'équivalence du test, « les résultats du test par Internet [étant] traités et interprétés comme s'il s'agissait de la version originale papier-crayon » (Guédon *et al.*, 2011, p.112).

Toutefois, puisque la modalité d'administration est différente de celle pour laquelle le test a été développé, il est nécessaire de procéder à une nouvelle validation de l'outil. On ne peut présumer de l'équivalence du test lorsqu'il y a un changement de modalité d'administration en raison des disparités possibles entre le format papier-crayon et Internet (Buchanan, 2000; Domino et Domino, 2006). Parmi les aspects à considérer, figurent la validité, l'équivalence des normes et la fidélité.

3.1 Risques liés à la validité

De façon très globale, le concept de validité vise à s'assurer qu'un test mesure effectivement le concept qu'il prétend mesurer. Pour la plupart, les auteurs s'entendent sur le caractère déterminant de la validité sur le plan de l'évaluation d'un test. Laveault et Grégoire (2002) déclarent qu'il s'agit d'un facteur crucial dans la détermination de la qualité d'un instrument psychométrique donné. Toutefois, Bernaud (2007) mentionne qu'un débat a toujours cours au sujet de « la question de seuils suffisants pour considérer la qualité d'une méthode (à propos de la *validité*) » (p.12). Ceci rappelle que la validité ne se définit pas de manière dichotomique (le test est valide ou ne l'est pas), mais plutôt, que sa valeur se construit de manière progressive à partir d'inférences (Furr et Bacharach, 2008).

Le risque relatif à la validité théorique lors d'un transfert papier-crayon à Internet réside en une modification potentielle du construit que l'outil est censé mesurer. Concrètement, cela veut dire qu'en transférant un test, on n'est plus assuré que le test mesure toujours le même concept psychologique (Buchanan et Smith, 1999; Domino et Domino, 2006). Si tel est le cas, une interprétation des scores n'est alors plus possible, car celle-ci serait basée sur un modèle théorique différent de celui à partir duquel le test original a été conçu et validé empiriquement. Buchanan et Smith (1999) proposent quelques causes possibles de cette perte de validité, soit le changement de format de présentation, le type de participants, les conditions et les motivations de passation.

3.2 Risques liés aux normes

L'équivalence des normes constitue un autre enjeu important de l'administration de tests psychométriques à partir d'Internet. Bien souvent, les tests utilisés ont recours à un échantillon normatif servant de référence pour l'interprétation des scores (Kaplan et Saccuzzo, 1997). Sa représentativité est importante, car c'est à partir de celui-ci que les scores d'un individu vont être interprétés. Les normes doivent être constituées de manière à représenter la

population définie, de la même façon que l'on doit s'assurer que les personnes à qui on administre le test font partie de cette même population (Grégoire, 2010). L'enjeu est complexe et il est admis qu'on ne peut être entièrement assuré de la représentativité d'un échantillon, bien que certaines précautions, telle l'utilisation d'un procédé aléatoire puisse en augmenter grandement la probabilité (Furr et Bacharach, 2008).

Dans le contexte de normes ayant été comptabilisées via papier-crayon, donc dans un contexte différent que celui dans lequel le test sera administré, il y a lieu de vérifier leur validité en contexte de testing en ligne. Sans validation de l'équivalence des normes dans cette autre modalité d'administration, il existe un risque que l'interprétation du test s'appuie sur des éléments de références erronés. En effet, des études font état de différences significatives entre les résultats des formats Internet et papier-crayon d'un même test, ce qui laisse penser que de nouvelles normes pourraient être nécessaires lors d'un transfert de format (Carstairs et Myors, 2009). Ceci introduit la possibilité que les personnes puissent répondre différemment selon la modalité d'administration, d'où le besoin de créer de nouvelles normes, à partir du même format que celui dans lequel le test sera administré.

En somme, il existe un risque que le changement de modalité d'administration entraîne un biais suffisant pour entraver la représentativité et l'équivalence des normes entraînant des répercussions importantes « sur l'interprétation, puisque la pondération des résultats (ex. : rang centile) s'en trouve affectée » (Guédon *et al.*, 2011, p.113).

3.3 Risques liés à la fidélité

La fidélité constitue une propriété essentielle d'un test psychométrique servant à définir la précision d'un instrument de mesure qui concerne deux volets distincts. Il s'agit premièrement d'une façon de mesurer la constance des résultats obtenus. Aussi, comme aucun test n'est parfaitement précis, il existe toujours une

part de variance du résultat à un test qui n'est pas attribuable à la variance du concept mesuré. Cette part est nommée l'erreur de mesure. La fidélité sert aussi à fournir une estimation de l'erreur de mesure d'un test. Anastasi (1994) résume la situation en énonçant que « la fidélité d'un test permet de déterminer dans quelle mesure les différences individuelles dans les scores aux tests peuvent être attribuées à de « vraies » différences [...] et dans quelle mesure elles doivent être attribuées à des erreurs dues au hasard » (p.97). La standardisation des conditions de passation, des consignes et de l'environnement contribue à restreindre l'erreur de mesure et augmente le degré de fidélité d'un test (*Ibid.*).

Lorsqu'il y a une modification de la modalité d'administration, de papier-crayon à Internet, ceci peut entraîner des passations largement différentes de celles dans lesquelles la fidélité du test a été mesurée (ex. : présence ou non d'un administrateur, conditions propices au testing ou non, matériel du test), pouvant avoir des conséquences sur les indices de fidélité du test.

3.4 Augmentation du testing internet non supervisé et ses risques

L'usage grandissant du testing Internet a entraîné un nouveau mouvement vers le testing non supervisé. Dans cette situation, le répondant complète le test au lieu et moment désiré sans l'assistance d'un administrateur. Précisément, la définition du testing non supervisé de Tippins (2009) réfère « *to Internet-based testing completed by a candidate without a traditional human proctor* » (p.2). Tous les avantages du testing Internet peuvent alors être exploités : réduction des ressources humaines, coûts moins élevés, possibilité de joindre des sujets plus éloignés, facilité à modifier les paramètres du test, possibilité d'enregistrer les données (délai de réponse, niveau d'items réussis), absence d'erreurs dans la cotation, (*Ibid.*).

Toutefois, plusieurs risques, en plus de tous ceux énoncés précédemment, subsistent toujours et leur étendue n'a toujours pas été mesurée. Dans son article, Pearlman (2009) cible des enjeux pragmatiques, légaux et éthiques à examiner

davantage. Tout d'abord, en n'étant pas supervisé, la possibilité de tricherie serait plus grande et la validité serait alors compromise, les résultats ne représentant plus nécessairement les individus concernés. Par ailleurs, sans nécessairement être une tentative de tricherie délibérée, on peut se demander dans quelle proportion les personnes répondant à un questionnaire autoadministré auraient recours à leurs proches, entravant ainsi la validité de l'outil. De plus, Tippins (2009) évoque un doute quant à l'environnement choisi par les répondants qui pourrait ne pas être adéquat au contexte de testing (présence de sources de distractions, environnement non anonyme) et biaiser les résultats. Rappelons que la standardisation fait partie des deux principaux attributs des tests psychométriques. En contexte de testing non supervisé, tout le processus d'administration du test est amputé d'un aspect important lié à la standardisation, soit l'énoncé des instructions, les réponses aux questions du répondant, le contrôle de l'environnement extérieur et du temps alloué, etc. Étant donné ces conditions qui peuvent différer grandement de celles dans lesquelles le test a été standardisé et dans lesquelles les normes ont été établies, les risques liés à la qualité de la mesure sont importants et il devient très périlleux de défendre la validité de ce type de testing (Pearlman, 2009). Dans la section suivante, l'aspect éthique et légal du testing Internet sera discuté de manière à mieux illustrer les tensions existantes entre les avantages et les risques du testing Internet.

4. RÉGLEMENTATION ENCADRANT LE TESTING INTERNET

Compte tenu de l'impact potentiel d'une évaluation psychométrique sur une personne, différentes réglementations régissent l'usage des tests. Les professionnels qui y ont recours, soit majoritairement des conseillers et conseillères d'orientation et des psychologues, doivent prendre connaissance des divers documents législatifs et déontologiques, selon leur contexte culturel de pratique, et s'y soumettre afin de les utiliser de manière éthique et responsable et éviter tout préjudice envers les personnes. De par leur statut, ils sont imputables des conséquences de l'ensemble du processus de testing et doivent donc veiller à ce qu'il soit bénéfique pour le client (Guédon *et al.*, 2011).

4.1 Normes de pratique en matière de testing

En 1966, l'American Educational Research Association, l'American Psychological Association et le National Council on Measurement and Education ont publié conjointement un ouvrage de référence commun quant à l'utilisation de tests dans leurs domaines respectifs (*Standards for Educational and Psychological Tests and Manuals*). Laveault et Grégoire attestent qu'il s'agit d'«une référence incontournable pour les spécialistes de la mesure en psychologie et en éducation » (2002, p.163). Les *Normes de pratique de testing en psychologie et en éducation* (OCCOPPQ et Sarrazin, 2003) est l'adaptation française de la dernière édition américaine publiée en 1999. Celle-ci s'adresse à toutes les parties impliquées dans l'utilisation de tests psychométriques et veille à la diffusion de principes guidant l'adoption de pratiques professionnelles valides et éthiques (*Ibid.*). Considérant l'article 43⁵ du Code de déontologie des conseillers et conseillères d'orientation du Québec⁶ auquel doivent se conformer ses membres, l'ouvrage de Sarrazin (2003) fait office de référence pour ces professionnels.

4.2 Recommandations de la Commission internationale des tests

La Commission internationale des tests est destinée à promouvoir l'utilisation éthique et équitable des tests psychométriques par le biais de la publication de documents de référence issus de la concertation de plusieurs acteurs internationaux. Selon Bartram (2009), la Commission répond à la volonté fort présente d'uniformisation des pratiques entre les différents états du monde. En 2001, l'organisme a diffusé les lignes de conduite générales liées à l'utilisation de tests (*International guidelines for test use*). Ultérieurement, cette publication a entraîné l'émission d'autres recommandations spécifiques au testing Internet (*International guidelines on computer-based and Internet delivered testing*, 2006).

⁵ « [l]e membre doit exercer sa profession en respectant les règles de l'art et en tenant compte des normes de pratique généralement reconnues dans sa profession. » (OCCOQ, article 43, p.4).

⁶ Le Code de déontologie comprend les devoirs et obligations que les membres de l'OCCOQ doivent remplir, en vertu du Code des professions qui régit légalement tout le système professionnel québécois (OCCOQ, 2010).

4.3 Implication des différentes réglementations

Les différentes réglementations présentées servent à encadrer la pratique du testing, de manière à éviter les préjudices possibles pour le public et à maintenir le niveau de qualité psychométrique. Anastasi (1994) justifie le contrôle des tests par le caractère scientifique des différents instruments psychométriques exigeant ainsi des attentions particulières quant à son emploi, sous peine de causer des dommages importants. Alors que le testing papier-crayon semble s'accommoder assez bien des normes éthiques, le testing Internet, phénomène relativement récent en comparaison à son prédécesseur, semble soulever davantage de questionnements. En effet, certains principes mis de l'avant par les organismes présentés précédemment ne sont pas respectés par l'utilisation d'Internet comme modalité d'administration de tests et cela représente un risque pour la validité.

Bernaud (2007) soutient la présence d'un consensus de la part de la communauté scientifique statuant que les mêmes normes de pratique du testing papier-crayon devraient s'appliquer au testing complété de manière informatisée. Toutefois, le testing Internet non supervisé contrevient explicitement à des normes de pratique concernant l'administration et est tout de même employé massivement. Dans le même ordre d'idées, il est recommandé de considérer les mêmes critères de sélection lors de la sélection d'un instrument psychométrique (Guédon *et al.*, 2011). Or, comme les conséquences d'un transfert de modalité d'administration, de papier-crayon à Internet, sont inconnues, on ne peut affirmer que les critères de sélection exigeant l'examen des propriétés psychométriques sont respectés. À ce sujet, la Commission internationale des tests a été claire à l'effet que les propriétés psychométriques doivent avoir été évaluées dans les mêmes conditions que dans lesquelles le test sera utilisé (Bartram, 2009). Le testing Internet, supervisé ou non, va à l'encontre de cette norme mise de l'avant par plusieurs des organismes réglementaires.

5. SYNTHÈSE

Les tests psychométriques présentent des propriétés spécifiques (validité, fidélité) qui en assurent la qualité. Le changement de modalité d'administration pourrait être suffisant pour les modifier et on ne peut dès lors, être assuré de l'équivalence des formats du test (Barak et English, 2002, dans Grégoire, 2010). Il en va de même concernant les normes, recueillies initialement par des tests administrés de façon papier-crayon. Domino et Domino (2000) sont du même avis : « *When we take a p-p test and administer it through a computer, we haven't simply changed the medium, but we have potentially changed the test. Thus, we cannot assume that existing norms and evidence of reliability and validity of the p-p format apply automatically to the cf [computer format]* » (p.463). Des normes explicites et rigoureuses ont été adoptées afin d'encadrer la pratique du testing. À celles-ci, se sont ajoutées des nouvelles visant explicitement le testing Internet. En dépit des actions entreprises, le testing Internet enfreint plusieurs règles éthiques et déontologiques.

Malgré les risques potentiels soulignés par les chercheurs et une série de principes guidant une meilleure utilisation du testing Internet, la pratique reste très répandue, sans qu'elle soit légitimée par des études scientifiques rigoureuses. Les avantages inhérents et le fait que l'établissement de nouvelles normes et le processus de validation soient des processus longs et coûteux peuvent expliquer pourquoi le testing Internet demeure néanmoins largement utilisé. À la lumière des nouvelles informations présentées dans cette troisième section, on constate toutefois que certains des risques liés à l'usage d'Internet en testing sont en contradiction avec certains avantages précédemment mentionnés. Une meilleure connaissance des risques associés au transfert de modalité d'administration des tests permettrait de faire un choix plus éclairé entre le testing traditionnel papier-crayon et le testing Internet, en fonction des avantages et particularités de chaque méthode.

6. QUESTION GÉNÉRALE DE RECHERCHE

L'intérêt envers Internet comme modalité d'administration des tests psychométriques par les conseillères et conseillers d'orientation est grandissant depuis plusieurs années. Cela s'explique facilement par les nombreux avantages qui y sont rattachés pour le professionnel et pour l'entreprise qui l'emploie. Des gains sur les plans de la réduction des coûts, de l'optimisation des ressources humaines et matérielles et de ses multiples possibilités ont été répertoriés en ce sens. Cependant, bien que le phénomène soit considérablement implanté, des risques demeurent sur le plan de la qualité de la mesure des concepts psychologiques. La validité, la fidélité et l'équivalence des normes sont spécifiquement identifiées comme des enjeux importants en raison de leur incidence sur la qualité de l'interprétation des scores. Il n'est pas possible de connaître, sans étude préalable, l'impact du changement de modalité d'administration sur un test psychométrique. De plus, l'arrivée du testing en ligne encourage de plus en plus la pratique du testing Internet non supervisé. Celle-ci va cependant à l'encontre de plusieurs normes et principes éthiques établis par différents organismes importants du domaine psychométrique. Considérant les nombreux avantages offerts par le testing Internet, les conseillers et conseillères d'orientation y ont recours malgré la violation de certaines lignes directrices professionnelles mises de l'avant de manière à limiter les dangers potentiels.

À ce jour, les conséquences réelles d'un tel changement, de papier-crayon à Internet, sont peu documentées. Guédon *et al.* (2011) soulignent, qu'en raison des conséquences potentielles du testing Internet sur l'interprétation et l'intervention avec les clients, de nouvelles études devraient être conduites afin de valider les propriétés du test. De plus, il est justifié de préciser que les conséquences d'une interprétation erronée peuvent être importantes selon les situations. En contexte de sélection de personnel, cela peut fausser les résultats et diminuer l'efficacité du testing dans le choix du meilleur candidat. En counseling de carrière, des individus vulnérables peuvent s'appuyer sur un profil d'intérêts imprécis pour se réorienter. Ces situations méritent qu'on s'attarde à la question de l'équivalence des formats papier-crayon et Internet considérant l'ampleur de leur utilisation. Une meilleure connaissance des conséquences d'un changement

de modalité d'administration pourrait préciser les normes de pratique en fonction des nouveaux éléments connus et à mieux assister les professionnels dans le domaine de la psychométrie.

Ainsi, en raison de l'ambivalence des chercheurs quant à l'équivalence des modalités d'administration papier-crayon et Internet et des préjudices possibles envers la clientèle des conseillers et conseillères d'orientation se soumettant à la passation de tests dans un format Internet, la question générale de recherche se pose comme suit : « Quels sont les effets d'un changement de modalité d'administration, de papier-crayon à Internet, sur un test psychométrique? ».

DEUXIÈME CHAPITRE– RECENSION DES ÉCRITS

Ce chapitre présente la recension d'écrits effectuée afin de présenter l'état actuel des connaissances quant à la question de l'équivalence d'un test psychométrique en format papier-crayon et en format Internet. La méthodologie de recension employée sera décrite et un rappel quant aux différences entre le testing informatisé et Internet sera fait en fonction de leur incidence sur la recension. Les différentes études retenues en raison de leur pertinence au projet de recherche constitueront la majeure partie du chapitre. Une synthèse permettra de faire le point sur l'état de la recherche actuelle, en plus d'identifier le manque de connaissances sur le sujet. Le chapitre se conclura avec les objectifs de la recherche.

1. MÉTHODOLOGIE DE LA RECENSION

À partir des banques de données pertinentes, une recherche portant sur les effets d'un changement de modalité d'administration, de papier-crayon à Internet, a été effectuée à l'aide des mots-clés « testing », « Internet », « web », « en ligne », « évaluation », « papier-crayon », « test » et « équivalence ». En raison du grand volume d'études en lien avec le sujet de recherche écrites dans la langue anglaise, les mots-clés ont aussi été utilisés dans cette langue, soit « testing », « Internet », « web-based », « online », « assessment », « paper-and-pencil », « test » et « equivalence ». Les banques de données sollicitées se situent dans le domaine de la psychologie : *PsychINFO*, *ERIC*, *Academic Search Complete* et *Psychological and Behavioral Collection*. Afin de compléter la recension et maximiser le volume d'information traitant de l'équivalence des modalités d'administration papier-crayon et informatique, certains numéros spéciaux de périodiques portant sur le thème du testing Internet ont été consultés. De plus, les listes de références bibliographiques des textes retenus ont aussi été consultées.

La recension a permis de répertorier différentes études qui traitaient directement de la question de recherche sur les effets d'un changement de modalité d'administration sur les tests psychométriques. Pour être considérés dans la recension, les tests étudiés devaient satisfaire aux standards psychométriques. Ceci a éliminé les études évaluant le potentiel d'Internet comme moyen de mener des recherches en psychologie. Par exemple, l'étude de Lewis, Watson et White (2009) a administré un sondage sur les habitudes de consommation d'alcool et de conduite automobile de façon papier-crayon et Internet. Comme le sondage ne répondait pas aux critères psychométriques, l'étude n'a pas été retenue dans le cadre de la recension d'écrits. Les études dont le devis utilisait des épreuves scolaires administrées en format papier-crayon et Internet ont aussi été exclues.

2. RÉSULTATS DE LA RECENSION

Comme cela a été spécifié, les risques liés à l'utilisation d'Internet comme modalité d'administration des tests dépendent en grande partie des motifs et des conditions d'usage qui y sont associés (Grégoire, 2010). Cela fait notamment référence à l'absence de supervision qui est courante lors de testing Internet. Avant l'existence d'Internet, on parlait de testing informatisé, car le test était souvent accessible via un ordinateur désigné ou à partir d'un disque compact. Dans les deux cas, cela favorisait la présence d'un administrateur, ce que recommandait déjà Cronbach (1990) : « *Even though the disk is intended to be complete in itself, an informed person should be standing by the test taker to call upon if he needs help* » (p. 61). De nos jours, Internet a entraîné un virage important sur le plan des risques potentiels, en raison de la suppression massive de supervision durant le testing. Whitaker (2007) offre une très pertinente revue des différences entre le testing informatisé et Internet. Il énonce entre autres qu'Internet étant accessible en tout temps à partir de presque n'importe quel endroit, il est beaucoup moins fréquent, pour des raisons de commodité et de coût, que les répondants soient assistés d'un administrateur. Ils seraient donc moins sujets à lire les instructions adéquatement et plus sujets à ignorer les mises en

garde que lors d'une passation supervisée (Barak et English, 2002). Il est noté qu'un plus grand nombre de recherches devraient étudier l'équivalence du testing Internet et informatique comme modalité d'administration afin d'explorer l'enjeu de l'administration sous supervision ou non.

Ce faisant, il a été décidé de reprendre cette indication lors de la recension d'écrits en regroupant les études portant sur les effets d'un changement de modalité d'administration selon si la passation Internet (ou informatique) était supervisée ou non. Sauf exception, toutes les passations des formats papier-crayon dans les études recensées ont été supervisées. Les différences entre les études étaient alors plus faciles à répertorier et à comparer entre elles. La plupart des études ont pour principe que s'il y a équivalence, on peut en déduire que le transfert de modalité d'administration n'entraîne pas de changement dans les propriétés psychométriques du test et que l'administration à partir d'Internet est valide.

2.1 Changement de modalité d'administration de papier-crayon à Internet sous supervision

L'étude de Salgado et Moscoso (2003) prend en compte plusieurs mesures différentes pour évaluer l'équivalence des formats d'un instrument de mesure (moyennes, écarts-types, coefficients de fidélité, structures factorielles, degré de congruence entre les facteurs et coefficients d'équivalence et de stabilité respectifs). Les formats papier-crayon et Internet du *Personality Inventory of Five Factor* ont été administrés à 162 étudiants universitaires avec un intervalle de deux à trois semaines entre chaque passation. Pour diminuer l'effet de présentation, l'ordre des versions a été inversé pour la moitié de l'échantillon. Malgré certaines différences non significatives, notamment l'écart-type qui était légèrement plus élevé dans la version Internet pour les cinq dimensions, l'hypothèse d'équivalence entre les deux versions du questionnaire a été soutenue (moyennes, écarts-types, coefficients de fidélité, structures factorielles, degré de congruence entre les facteurs et coefficients d'équivalence et de stabilité

respectifs). En dépit d'un échantillon limité, le devis à mesures répétées confère néanmoins à cette étude une pertinence certaine.

Ployhart, Weekley, Holtz et Kemp (2003) ont mené une étude visant à évaluer l'équivalence de tests supervisés administrés à partir d'Internet comparativement à leurs versions papier-crayon, dans un contexte de sélection de personnel. Une batterie d'instruments comprenant la mesure de trois traits du modèle de personnalité à cinq facteurs (Consciencieux, Agréabilité, Stabilité émotionnelle), un formulaire d'informations biographiques et un test de jugement situationnel ont été utilisés. L'échantillon était composé de 425 personnes employées de l'entreprise, 2544 postulants ayant complété la version papier-crayon et 2356 postulants ayant complété la version Internet. Le contexte de l'étude était réel et toutes les personnes postulaient pour le même emploi au sein de l'entreprise. Les résultats présentent des distributions significativement plus normales, des moyennes plus faibles, des indices de cohérence interne plus élevés, plus de variance et des intercorrélations plus élevées pour les versions Internet. Les auteurs mentionnent que les groupes n'étant pas équivalents, il est impossible d'affirmer que les différences ne sont pas dues à des différences dans l'échantillon, plutôt qu'au mode d'administration. Bien qu'ils ne possèdent pas les données pour appuyer leurs propos, Ployhart *et al.* (2003) suggèrent des niveaux de motivation différents, la désirabilité sociale, le manque de confiance envers Internet et la confidentialité des scores et des facteurs technologiques pour expliquer les différences entre les versions du test. Comme mentionné par les auteurs, malgré le caractère réel de l'expérimentation, nous ne sommes pas en mesure d'associer les différences aux modes d'administration. Celles-ci nous indiquent cependant des pistes intéressantes de recherche.

Dans l'étude de Potosky et Bobko (2004) portant sur l'utilisation d'Internet pour évaluer des construits cognitifs dans un contexte de sélection de personnel, 65 participants ont complété un test d'aptitudes et un test de jugement situationnel en format papier-crayon et Internet. Le contexte n'était toutefois pas aussi réel que celle de Ployhart *et al.* (2003), car aucune offre d'emploi n'y était

associée. Toutes les étapes (passation Internet et papier-crayon) étaient supervisées et l'ordre de passation a été inversé entre deux sous-groupes avec un intervalle de trente à quarante minutes. L'indice d'intercorrélations pour le test d'aptitudes est acceptable (0,60), mais insuffisant pour assurer le même résultat de choix de candidats. Il est plus élevé pour le test de jugement (0,84). Les moyennes des scores Internet du test d'aptitudes sont significativement plus basses que celles de la version papier-crayon. Pour la seconde partie de l'étude, différents biais potentiels limitant la validité du testing Internet d'habiletés ont été examinés. Le temps de téléchargement, la plus grande difficulté à revenir en arrière et à compléter les sections à leur guise, l'importance des instructions et le design matériel du test ont été identifiés par les auteurs comme les plus importants. Cette étude présente des failles méthodologiques importantes (petit échantillon, intervalle trop court entre les passations) qui nuisent à la validité des résultats.

L'étude de Meade, Michels et Lautenschlager (2007) a utilisé le concept d'invariance de mesure qui consiste en l'étude de l'équivalence conceptuelle selon différents sous-groupes, afin d'examiner l'équivalence des versions papier-crayon et Internet d'un inventaire de personnalité. Pour ce faire, 727 étudiants universitaires ont complété le test dans un des deux formats. Dans le but d'explorer l'influence du choix du format complété, certains avaient le choix, alors que d'autres y étaient assignés, ce qui a abouti en quatre sous-groupes. L'étude des versions de 11 des 32 échelles de l'*Occupational Personality Questionnaire* (OPQ-32) a résulté en des scores comparables pour les deux formats. Cependant, l'invariance de mesure entre les personnes ayant choisi le format et ceux ayant été assignés n'a pas été observée pour trois échelles (*Outgoing*, *Optimism* et *Evaluative*). Indépendamment du choix ou non du format, l'ajustement à la structure factorielle de la version Internet était meilleur que celui de la version papier-crayon. Toutefois, seulement sept des onze échelles ont démontré un degré d'ajustement suffisant, allant d'élevé à modéré. Pour expliquer l'absence d'invariance de mesure, les auteurs ont identifié des facteurs différents selon les échelles : le format du test lui-même (*Outgoing*, *Optimism*), l'assignation à un format (*Affiliative*, *Variety Seeking*) et les caractéristiques des personnes les

amenant à choisir un format plutôt que l'autre (*Evaluative*). Le choix de la version Internet était associé à plus d'heures passées sur Internet et à jouer à des jeux informatiques. L'apport unique de cette étude consiste aux différences obtenues dans les résultats des questionnaires dont les personnes avaient choisi le format. Ceci indique qu'une précaution supplémentaire concernant l'équivalence des modalités doit être prise lorsqu'un choix est offert.

Par le biais de leur devis incluant quatre conditions distinctes d'administration, Cronk et West (2002) ont examiné à la fois l'effet du format d'administration et l'effet de la supervision. Un sondage composé de 30 items tirés de l'outil psychométrique *Visions of Morality Scale* a été complété par 196 collégiens répartis proportionnellement dans les quatre conditions. Les moyennes et les écarts-types des différents sous-groupes n'ont présenté aucune différence significative. Toutefois, le taux de réponse a été significativement moins élevé pour la version Internet à compléter à la maison ce que les auteurs attribuent à des facteurs, telles la convenance, l'accessibilité et l'aisance à utiliser l'informatique. L'étude est limitée en termes de résultats.

Dans le cadre d'un devis de recherche à mesures répétées avec ordres de passation inversés, Naus, Phillip, Samsi (2008) ont administré le *Short Form Health Survey*, le *Beck Depression Index II* et le *NEO-Five Factor Inventory* en version Internet et papier-crayon à 76 étudiantes universitaires. Les analyses effectuées n'ont révélé aucun effet d'ordre de passation. Les corrélations et les intercorrélations entre les différents construits ont supporté l'hypothèse d'équivalence entre les formats. En contrepartie, des différences significatives dans les moyennes ont été observées sur trois échelles du NEO (Ouverture, Agréabilité et Consciencieux) selon le mode d'administration. Le sentiment d'anonymat et la désirabilité sociale sont énoncés comme piste d'explication des différences sur ces échelles. L'échantillon exclusivement féminin nuit à la généralisation des résultats, de même que sa petite taille.

2.2 Changement de modalité d'administration de papier-crayon à Internet sans supervision

Une des premières recherches d'importance sur la validité d'Internet en recherche a été celle de Buchanan et Smith (1999) qui ont conduit une étude utilisant le *Self-Monitoring Scale-Revised* administré à 1187 participants. Parmi eux, 968 ont complété la version Internet et 224 volontaires l'ont rempli de façon papier-crayon. Les analyses statistiques descriptives effectuées (moyenne, écart-type, aplatissement) n'ont identifié aucune différence significative entre les deux versions. Le coefficient de fidélité de la version Internet (0,75) a été comparable à celui de la version originale papier-crayon (0,70) et celui de l'étude (0,73). Des analyses factorielles confirmatoires ont permis d'identifier un modèle à trois facteurs avec un haut degré d'ajustement entre les deux versions. Les saturations d'items sur le premier facteur sont similaires. Les résultats issus de la modalité Internet font toutefois état de corrélations plus fortes entre le premier facteur et le score total de l'échelle, ce qui amène les auteurs à suggérer que la version Internet soit une mesure plus efficace. À titre d'explications possibles, les auteurs mentionnent un possible niveau plus élevé d'honnêteté et d'autodévoilement et l'hétérogénéité de l'échantillon. En effet, il est avancé que la version Internet ayant été ouverte à tous, le niveau d'hétérogénéité ait été plus grand que la version papier-crayon restreinte à une population étudiante exclusivement. L'utilisation d'un échantillon comparatif et le moment de l'étude relativisent la pertinence et la portée des résultats, bien qu'elle offre une analyse étoffée de l'équivalence des structures factorielles.

Afin d'évaluer la possibilité d'utiliser Internet pour collecter les données liées aux concepts d'attachement parental et d'ajustement émotionnel, Fouladi, McCarthy et Moller (2001) ont administré trois tests psychométriques (l'*Inventory of Parent and Peer Attachment*, le *Negative Mood Scale* et le *Trait Meta-Mood Scale* (TMMS)). L'échantillon ayant complété les versions papier-crayon (n = 164) et les versions Internet (n = 234) de ces instruments était composé d'étudiants fréquentant la même université. Des différences ont été notées entre les résultats des différentes versions (moyennes, écarts-types), mais celles-ci

étaient très petites. Les analyses factorielles font état d'un bon degré d'ajustement entre les structures factorielles des versions respectives. Toutefois, les modèles issus de ces deux formats diffèrent du modèle original des tests, à l'exception du TMMS. Selon les auteurs, cela soulève quelques questions, mais ne remet pas en question l'équivalence des versions ni leur usage. On peut néanmoins remettre en question cette conclusion hâtive, en raison de l'absence d'adéquation au modèle théorique qui peut représenter un manque de validité important. Il est aussi impossible de généraliser les résultats, compte tenu de l'homogénéité de l'échantillon (ethnie, âge, niveau d'étude).

Bartram et Brown (2004) ont exploré l'équivalence des versions Internet et papier-crayon à l'aide du test *Occupational Personality Questionnaire* (OPQ32i). L'administration non supervisée de la version Internet (n = 768) et supervisée de la version papier-crayon (n = 1127) a permis de constater la comparabilité de la fidélité de chaque format (0,78 - 0,80) et des relations entre les échelles. Les effets de taille des différences de moyennes intergroupes (Internet, papier-crayon) suggèrent des petites différences significatives, insuffisantes selon les auteurs pour requérir de nouvelles normes. L'étude du sous-échantillon le plus comparable composé de cadres et de professionnels (n = 610, Internet; n = 154, papier-crayon) relève que la supervision ou le format papier-crayon affecterait six échelles par plus ou moins un quart d'un écart-type, ce qui aurait un impact minime sur l'interprétation du profil. La promptitude des auteurs à déclarer qu'il y a équivalence est à souligner et il convient de faire preuve de plus de prudence dans l'interprétation de ces résultats afin de s'assurer de la validité de l'instrument suite à un changement de modalité d'administration.

Joubert et Kriek (2009) ont mené une recherche portant sur l'équivalence du même test (OPQ32i) dans les mêmes conditions de supervision que l'étude de Bartram et Brown (2004), mais dans un contexte réel de sélection de personnel. Les auteurs ont conduit deux études parallèles ayant le même devis, à l'exception de la composition majoritaire des participants (étude 1 : étudiants diplômés, étude 2 : employés-cadres). Ils ont recensé des indices d'asymétrie et

d'aplatissement, des écarts-types et des intercorrélations similaires de même qu'un bon degré d'ajustement entre les échelles. L'analyse statistique des écarts-types combinés ont indiqué des effets de taille allant de très petit (0,01) à moyen (0,54) quant aux différences significatives entre les moyennes d'échelles. Les auteurs ont attribué les effets de taille les plus élevés à des différences d'échantillonnage et non au mode d'administration. Des coefficients alpha plus élevés ont été remarqués dans le cas des passations Internet. Considérant les effets de tailles les plus élevés, les conclusions des auteurs, à l'effet que les différences les plus grandes soient dues à des biais d'échantillonnage, semblent injustifiées. Ceci est d'autant plus important compte tenu de l'utilisation de deux échantillons indépendants.

La recherche réalisée par Coyne, Warszta, Beadle et Sheehan (2005) comprend deux études. La première a eu recours à 86 employés qui ont complété un test d'habiletés en format Internet non supervisé et papier-crayon supervisé, avec un ordre de passation inversé selon deux sous-groupes. Des coefficients d'équivalence et de stabilité de niveau moyen (0,53 à 0,67, après correction : 0,65 à 0,81) ont été obtenus aux différentes échelles (verbale, numérique, spatiale); l'échelle verbale étant particulièrement basse (0,53, corrigé à 0,65). Les effets de taille des moyennes ont varié entre 0,04 et 0,49, indiquant des scores significativement plus élevés en version papier-crayon. La deuxième étude a utilisé deux échantillons, dont l'un de 64 étudiants (version Internet et papier-crayon) et l'autre de 54 étudiants (version informatique et papier-crayon). Contrairement à la version Internet, la passation de la version informatique était supervisée dans un laboratoire. Les coefficients d'équivalence et de stabilité les plus faibles étaient encore ceux des échelles d'habiletés en version Internet (0,47-0,69, après correction : 0,72 - 0,80) et en version informatisée (0,61 - 0,80, après correction : 0,60 - 0,82). Pour les moyennes, des effets de taille entre 0,22 et 0,90 ont été répertoriés pour ces mêmes échelles d'habiletés. Des analyses subséquentes ont démontré l'absence de différences significatives entre ces deux modes de passation et leur différence commune avec la passation papier-crayon. Cela suggère donc l'attribution des différences au format du test et non au fait

d'être supervisé ou non. Les coefficients des échelles d'habiletés ont supporté l'équivalence de toutes les versions. Le devis à mesures répétées utilisé confère davantage d'intérêt aux résultats obtenus, selon lequel deux des trois grandes échelles du test d'habileté sont inéquivalentes lorsque la modalité d'administration est modifiée (mode informatisé ou Internet). Cependant, on dispose de peu d'informations sur l'intervalle de temps entre les deux passations, à l'exception qu'il a été d'au minimum une semaine.

Afin de rendre compte de la faisabilité d'utiliser Internet comme modalité d'administration pour l'évaluation des troubles d'anxiété sociale, Hedman *et al.* (2010) ont administré six questionnaires en format papier-crayon ($n = 64$) et Internet ($n = 57$). Les alphas de Cronbach et l'étude de la distribution (moyennes et écarts-types) soutiennent l'hypothèse d'équivalence dans les deux formats, à l'exception du *Quality of Life Inventory* qui obtient des différences significatives dans les deux cas. La validité théorique a été soutenue par onze des quinze intercorrélations attendues, suggérant ainsi une bonne validité convergente. Une différence selon le mode d'administration, dans les coefficients de corrélations entre les questionnaires sur la dépression et l'anxiété en général et les questionnaires sur l'anxiété sociale, laisse sous-entendre une meilleure discrimination des différents états émotionnels par le format papier-crayon. Selon les auteurs, cela pourrait être attribuable à une meilleure distinction des différents blocs de questions dans la version papier-crayon qu'à l'ordinateur. La petite taille de l'échantillon et le devis à mesures non répétées contribuent à relativiser la validité externe et interne de l'étude.

Avec les mêmes objectifs que Hedman *et al.* (2010), Hirai *et al.* (2011) ont administré les mêmes questionnaires sur l'anxiété en version Internet ($n = 262$) et en version papier-crayon ($n = 260$). L'étude de la distribution (aplatissement, asymétrie), des moyennes, des écarts-types et des indices de cohérence interne n'a permis de repérer aucune différence significative. Les analyses factorielles confirmatoires ont identifié des structures factorielles avec un degré d'ajustement acceptable. Les fortes corrélations entre les deux questionnaires (SIAS: 0,81, SPS:

0,81) indiquent une évidence pour la validité convergente entre les différentes modalités d'administration. Cependant, les différences entre les saturations factorielles suggèrent que lorsqu'ils sont administrés dans des contextes et formats différents, les items du test évaluent l'anxiété sociale différemment. Ainsi dit, une même personne n'obtiendrait pas nécessairement le même score s'il passait le test dans deux formats différents. La validité théorique de l'outil, qui correspond au degré de certitude que le test mesure le construit à l'étude (Anastasi, 1994), est donc compromise. Il est noté que les personnes souffrant de ce type d'anxiété seraient potentiellement plus sujettes à être influencées par la modalité d'administration, particulièrement lorsque cela est lié à des aspects sociaux (ex. : supervision ou non). En conclusion, la comparaison des résultats obtenus via des modalités différentes est déconseillée. L'emploi d'Internet comme seule modalité d'administration est soutenu, ce qui implique la création de nouvelles normes. Cette étude démontre bien les différentes étapes nécessaires à l'étude de l'équivalence des modalités d'administration par les nombreuses analyses statistiques effectuées. Les auteurs présentent également une discussion nuancée et contextualisée sur l'impact de tests n'ayant pas démontré l'invariance de mesure.

Toujours dans l'évaluation de l'anxiété, mais cette fois, spécifique aux ordinateurs, Barbeite et Weiss (2004) ont administré plusieurs mesures⁷, originellement développées en format papier-crayon, dans un format Internet. L'échantillon, composé de 226 participants issus d'une banque de volontaires aux fins de recherche, a complété cinq échelles mesurant soit l'anxiété ou le sentiment d'efficacité personnelle relié aux ordinateurs sur Internet. Les résultats ont ensuite été comparés à des études de validation déjà produites. Les analyses factorielles exploratoires et confirmatoires n'ont pas obtenu des seuils suffisants aux différents indices pour appuyer l'utilisation de ces mesures dans un format Internet. De plus, seulement un des cinq tests a retenu le même nombre de facteurs en format papier-crayon et en format Internet. Ces résultats suggèrent

⁷ La *Computer Anxiety Scale*, la *Computer Anxiety Rating Scale*, la *Computer Self-efficacy Scale*, la sous-échelle mesurant l'anxiété reliée aux ordinateurs et la sous-échelle mesurant la confiance reliée aux ordinateurs issue de la *Computer Attitude Scale*.

l'absence d'équivalence de mesure entre les deux modalités d'administration. La principale limite est qu'on ne peut être assuré de la cause des différences en raison des deux échantillons distincts.

Aluja, Rossier, Zuckerman (2007) ont étudié les moyennes, les écarts-types, les distributions et les structures factorielles des versions Internet (non supervisé) et papier-crayon (supervisé) d'un inventaire de personnalité basé sur le modèle *Big Five*. Des échantillons suisse (Internet, $n = 1403$; p-c, $n = 742$) et espagnol (Internet, $n = 1941$; p-c, $n = 1886$) ont été utilisés. L'échelle *Sociability* a obtenu des différences significatives avec un effet de taille moyenne pour l'échantillon espagnol (-,46) et petite pour l'échantillon suisse (-0,24). L'échelle *Impulse Sensation Seeking* a obtenu un effet de taille minime dans les deux échantillons ($< 0,20$). Aucun coefficient de fidélité n'a été significativement différent entre les formats. Les analyses factorielles exploratoires et confirmatoires ont démontré un bon degré d'ajustement à la structure factorielle de la version originale. Encore une fois, les résultats peuvent être biaisés en fonction du devis intersujets.

Avec un objectif similaire à l'étude de Ployhart *et al.*, l'équivalence de l'administration Internet et papier-crayon du testing cognitif a été étudiée dans un contexte réel comportant des enjeux importants (*high-stakes*) dans l'étude menée par Carstairs et Myers (2009). Deux cohortes (p-c : $n = 159$, Internet : $n = 143$) et différents niveaux de supervision ont mené à l'observation de différences significatives avec un très grand effet de taille ($d = 2,00$ et $1,21$) suggérant que l'absence de supervision aurait un impact bien plus grand que le changement de modalité d'administration. Malgré le recours à deux échantillons, l'étude permet d'éclairer l'enjeu de la supervision énoncé dans la problématique en observant des différences très élevées.

Même si leur objectif de recherche tenait davantage de la validation d'Internet comme moyen de recherche, l'étude de Meyerson et Tryon (2003) a été retenue, puisqu'elle utilisait un instrument psychométrique (le *Sexual Boredom*

Scale). Un échantillon composé de 533 participants a complété la version Internet. Leurs résultats ont été comparés à ceux d'une recherche antérieure ayant sollicité 253 participants⁸. Les coefficients de fidélité présentent selon le format de faibles différences (0,00 - 0,15) : le plus élevé étant parfois relié à la version en ligne, parfois à la version papier-crayon. Les indices classiques d'ajustement suggèrent un très bon degré d'ajustement entre les corrélations respectives des deux groupes. Les deux groupes ont été présentés comme équivalents, mais il ressort quelques différences significatives, quoique petites, soit plus d'hommes et de diversité ethnique que dans le format papier-crayon. Comme la fidélité initiale est moyenne pour quatre sous-échelles (0,56-0,76) selon la recherche antérieure, il est difficile d'utiliser avec confiance les résultats qui y sont associés.

Coles, Cook et Blake (2007) ont étudié l'équivalence de deux tests associés à l'évaluation des troubles compulsifs (*Obsessive Compulsive Inventory* et *Obsessive Beliefs Questionnaire*) en format Internet et papier-crayon. Le premier sous-groupe de l'échantillon (n = 55) les a complétés en format Internet et l'autre, en format papier-crayon (n = 50). Les tests n'ont révélé aucune différence significative reliée à l'ordre de passation. Les comparaisons des moyennes, des écarts-types, des intercorrélations, des indices de fidélité et des corrélations reliées au critère externe n'ont pas permis d'observer des différences significatives selon le mode d'administration. D'autres analyses portant sur l'influence potentielle de l'expérience avec les ordinateurs sur les scores ont conclu que cette hypothèse n'était pas appuyée. Considérant que cette étude a utilisé des échantillons différents et de petite taille pour comparer les deux modalités d'administration, celle-ci demeure une étude exploratoire pertinente.

Dans le but d'évaluer Internet comme méthode de recherche concernant les attitudes envers les personnes âgées, Helmes et Campbell (2010) ont utilisé un devis à mesures répétées avec un intervalle minimum d'une semaine entre les deux passations. Les résultats de 60 étudiants universitaires ayant complété les tests en format Internet (non supervisé) et en format papier-crayon (supervisé) des

⁸ Watt, J.D. et Ewing, J.E. (1996). Toward the development and validation of a measure of sexual boredom. *Journal of Sex Research*, 33, 57-66.

cinq mesures⁹ ont pu être interprétés. L'ordre de passation a été inversé pour la moitié de l'échantillon. Pour quatre instruments, des différences significatives ont été observées entre les scores, avec des effets de taille allant de moyen (0,37) à élevés (1,94). Pour chacune d'entre elles, plus d'attitudes négatives ont été exprimées dans le format Internet. Les coefficients de fidélité ont été similaires et suffisants, à l'exception d'une mesure dont le coefficient a été insuffisant dans les deux formats, révélant alors un manque de précision relatif à l'outil et non au mode d'administration. Le fait qu'il était le dernier test à compléter a pu diminuer sa fidélité selon les auteurs. Ceux-ci ont bâti une étude cohérente qui a permis d'explorer de manière intéressante l'effet d'Internet sur l'autodévoilement des répondants.

L'étude de Carlbring *et al.* (2007) s'est penchée sur l'équivalence d'Internet comme modalité d'administration pour l'évaluation de l'agoraphobie et du trouble panique¹⁰. Selon les divers tests, un échantillon entre 344 et 350 participants a été recueilli en raison de passations incomplètes. Les tests en format Internet et papier-crayon ont été complétés sans supervision, à l'endroit choisi par les participants, avec l'ordre de passation inversé pour la moitié de l'échantillon. Un intervalle d'une journée était suggéré aux participants. Les indices de cohérence interne et les corrélations étaient similaires pour les deux modalités d'administration. Cependant, des différences significatives ont été notées dans les moyennes obtenues. Pour deux tests, les résultats issus de l'administration Internet étaient plus élevés; les effets de taille étaient toutefois petits. À l'opposé, des scores significativement plus élevés ont été observés pour trois tests dans le format papier-crayon. Pour un d'entre eux, le *Beck Anxiety Inventory*, l'effet de taille était si large que l'utilisation des mêmes normes serait inappropriée. L'intervalle d'une journée seulement entre les passations pourrait constituer un biais ayant pu influencer les résultats, un délai plus long est généralement

⁹ Le *Beliefs about Older Workers Questionnaire*, le *Attitude Toward the Employment of Older People Scale*, une échelle mesurant les perceptions quant au développement, à la capacité de performance, à la stabilité et aux habiletés interpersonnelles.

¹⁰ Les tests utilisés étaient le *Body sensations questionnaire*, l'*Agoraphobic cognitions questionnaire*, le *Mobility Inventory*, le *Beck anxiety inventory*, le *Beck depression inventory*, le *Quality of life inventory* et le *Montgomery Åsberg depression rating scale (MADRS)*.

recommandé. Aussi, étant donné le mode de recrutement volontaire par Internet, on dispose de peu d'informations quant à leur représentativité et leurs caractéristiques (ex. : population clinique).

Dans le but d'évaluer l'effet de la supervision lors de testing Internet, les chercheurs Chuah, Drasgow et Roberts (2006) ont administré les mêmes tests sous trois conditions différentes (papier-crayon supervisé, Internet non supervisé, Internet supervisé). Un échantillon composé de 728 étudiants universitaires ayant été affectés aléatoirement à une condition d'administration a complété un test formé d'adjectifs servant à dresser le profil de la personnalité basé sur le modèle à cinq facteurs. Le facteur « ouverture » n'a cependant pu être étudié, des tests ayant révélé sa multidimensionnalité. Des analyses factorielles ont révélé une structure factorielle similaire pour les trois passations. Des tests de validation en référence à un critère externe ont aussi soutenu l'équivalence des modalités d'administration, à l'exception d'une seule corrélation. Dix-neuf items ont été identifiés comme présentant un fonctionnement différent lorsque comparés entre les modes d'administration. Plus d'items de ce type ont été répertoriés dans la comparaison papier-ordinateur (12) que dans celle papier-internet (5), amenant les auteurs à expliquer ces différences par la présence de faux positifs et à soutenir l'équivalence des formats traditionnels et informatiques. L'étude des moyennes a mené à l'observation d'aucune différence significative, avec de petits effets de taille (-0,26-0,07). Les conclusions émises par les auteurs expliquent peu les disparités inattendues observées entre les formats informatique et Internet. De plus, l'abandon d'une dimension diminue la valeur des résultats d'analyse factorielle.

Dans leur étude, Hardré, Crowson et Xie (2010) ont affecté aléatoirement un nombre équivalent d'étudiants universitaires aux deux conditions de passation (papier-crayon et Internet), pour un total de 326 participants. Dans les deux cas, ils disposaient d'une semaine pour remplir le formulaire au moment et au lieu choisis. Plusieurs échelles ($n = 17$) évaluant différents construits liés à l'éducation et aux sciences sociales composaient le questionnaire. Les analyses

font état d'un nombre plus important de pertes de données en format Internet pour plusieurs mesures, mais les différences sont petites. Une tendance vers des moyennes plus élevées en format papier-crayon est aussi observée, mais les différences ne sont pas significatives statistiquement. Six mesures ont présenté significativement plus de variance en format papier-crayon. Les alphas de Cronbach étaient similaires pour les deux conditions d'administration. Le temps de passation a été significativement plus long et le sentiment positif significativement plus élevé en format papier-crayon. Dans le premier cas, compte tenu de la différence de taille d'échantillon selon les formats, cette information doit être interprétée avec prudence. Les données qualitatives récoltées font état d'un pourcentage suffisamment important de participants ayant rempli les questionnaires dans des conditions différentes de celles préconisées lors d'administration de tests psychométriques; 46 % l'ont fait en compagnie d'autres personnes et entre 30 % et 40 % en participant à d'autres activités (télévision, conversation avec d'autres, navigation Internet). Le devis de recherche mixte de cette étude soulève des questionnements importants quant à la précision et à la validité des données obtenues, spécialement lors d'absence de supervision, en vertu des résultats concernant le contexte de passation choisi. De façon générale, l'étude divulgue de l'information pertinente quant aux différents formats d'un même test et ses propriétés, en plus de documenter les effets d'absence de supervision en contexte Internet.

Dans un contexte réel de recherche d'emploi rappelant celui de Ployhart *et al.*, l'étude de Robie et Brown (2006) a comparé l'administration Internet d'un test de personnalité à celle papier-crayon effectuée dans les locaux de l'employeur auprès d'un très large échantillon de postulants ($n = 370\ 122$). L'instrument était formé d'items partagés en trois échelles (*Agreeability*, *Emotional Stability* et *Conscientiousness*). Ce devis intersujets a indiqué une différence significative dans les scores de l'échelle *Emotional Stability* avec un petit effet de taille ($d = 0,24$). Les intercorrélations et les coefficients alpha ont été très similaires. Des analyses statistiques portant sur la race et le genre ont permis d'indiquer quelques différences supplémentaires. Ainsi, un nombre plus élevé d'hommes et une plus

faible proportion de personnes de race blanche et d'origine asiatique ont choisi l'administration Internet. Un plus grand nombre de personnes ont accepté de révéler leur origine ethnique dans ce format, mais l'écart est minime. L'expression du genre n'a indiqué aucune différence. Les auteurs de l'étude concluent que les mesures de personnalité semblent équivalentes de manière générale entre les modalités, à l'exception d'échelles s'apparentant au concept de « Stabilité émotionnelle » où une hausse des scores est possible. Cette étude aborde des aspects intéressants liés au concept d'équivalence (ex. : autodévoilement, différence entre les différentes populations), mais la divulgation de peu d'informations concernant le devis, le mode papier-crayon utilisé et le degré de supervision constitue une limite importante.

L'étude de Herrero et Meneses (2006) a utilisé deux versions courtes de tests mesurant respectivement le stress (*Perceived Stress Scale*) et la dépression (*Center for Epidemiological Studies-Depression Scale*) afin de comparer leur administration Internet à celle papier-crayon. Un total de 530 étudiants universitaires, tous à l'aise avec Internet, a complété les deux instruments dans l'un ou l'autre des modes d'administration. Les indices de cohérence interne n'ont révélé aucune différence significative. Les moyennes se sont elles aussi révélées plutôt similaires, à l'exception d'un item mesurant le stress significativement plus élevé en format Internet. Les analyses factorielles et en composantes principales ont également soutenu l'équivalence des formats dans les deux modalités avec des facteurs et des saturations factorielles similaires. Les auteurs suggèrent que pour les personnes à l'aise avec Internet, la modalité d'administration n'a pas d'incidence sur leurs réponses. Cette affirmation est cohérente avec les résultats obtenus, mais doit être nuancée compte tenu des faibles alphas de Cronbach du test sur le stress (Internet= 0,72, papier-crayon= 0,68). De plus, le devis utilisant deux groupes non équivalents induit une validité interne moindre qu'un devis à mesures répétées.

3. SYNTHÈSE

À la suite de la recension d'écrits, certains constats se dégagent des résultats répertoriés par les études. Tout d'abord, la première observation concerne un des aspects importants de la validité d'un test, soit l'équivalence des résultats entre les modalités d'administration. Plusieurs études ont obtenu des différences significatives entre les moyennes des formats du même test psychométrique. Les scores étaient parfois plus élevés en format papier-crayon et parfois plus élevés en format Internet. Cette variation peut indiquer des effets différents du changement de modalité d'administration selon divers aspects, tels que le type de construit mesuré, la procédure d'administration ou l'ordre de passation. D'autres résultats d'études ont supporté l'hypothèse d'équivalence des normes avec des moyennes similaires. Toutefois, seulement une d'entre elles (Salgado et Moscoso, 2003) avait utilisé un devis à mesures répétées qui élimine le biais que des différences trouvées soient attribuables à des différences intergroupes plutôt qu'à la modalité d'administration.

En ce qui concerne la fidélité, certaines études appuient l'hypothèse d'équivalence des formats, alors que d'autres, en plus faible proportion comparativement à l'équivalence des moyennes, présentent des différences assez élevées pour la rejeter. C'est donc dire que des tests administrés dans un format différent de leur conception initiale pourraient avoir un niveau de précision significativement différent. Toutefois, il n'y a eu aucune différence en lien avec la fidélité des tests lorsque la passation était supervisée, suggérant ainsi que la supervision entraîne une fidélité comparable à celle obtenue dans le format original.

Enfin, un certain nombre d'études ayant procédé à des analyses factorielles ont aussi observé des différences entre les structures factorielles des formats papier-crayon et Internet. Certaines ont obtenu des saturations factorielles différentes (Meade *et al.*, 2007; Hirai *et al.*, 2011) et d'autres ont carrément obtenu un modèle factoriel différent avec un nombre non équivalent de facteurs (Fouladi, Moller et McCarthy, 2001; Barbeite et Weiss, 2004). Ces observations

indiquent différents impacts majeurs possibles d'un changement de modalité d'administration sur un test psychométrique. Nous avons remarqué que parmi ces études, seulement une d'entre elles avait une taille d'échantillon supérieur à 270 participants. De plus, elles avaient toutes utilisées un devis à mesures non répétées, induisant un biais méthodologique important où on ne peut être assuré de la cause des différences en raison de deux échantillons distincts. Toutefois, en étudiant les devis des études où les solutions factorielles étaient équivalentes, il a été noté que la majorité d'entre elles avaient également eu recours à ce type d'échantillonnage. Concernant le choix de distinguer dans la recension d'écrits, les études avec la passation Internet supervisée et celles ne l'ayant pas supervisée, des différences similaires concernant les moyennes et les structures factorielles ont été observées indépendamment du mode de supervision. Enfin, il est important de noter que toutes les études ayant obtenu des saturations factorielles différentes ou un nombre non équivalent de facteurs ont utilisé un test psychométrique mesurant soit la personnalité ou un aspect lié à la personnalité (ex. : l'anxiété). Ceci suggère que l'étude de la personnalité serait davantage affectée par le changement de modalité d'administration que d'autres concepts, telles les aptitudes.

Quelques limites importantes se retrouvent dans plusieurs études nuisant ainsi de façon importante à la généralisation des résultats. Premièrement, peu d'études empiriques ont utilisé un même échantillon afin de s'assurer que les différences trouvées, s'il y a lieu, se rapportent au changement de modalité d'administration et non à des différences entre les personnes composant les échantillons. Toutefois, il est à noter que l'utilisation d'un même échantillon entraîne un autre biais, soit celui que le participant se souvienne de ses réponses d'une passation à l'autre. Deuxièmement, la majorité d'entre elles n'a pas utilisé des groupes équivalents, ce qui aurait pu nuancer cette faiblesse méthodologique. Enfin, quelques-unes ont utilisé un échantillon de petite taille avec une faible puissance statistique. Dans ce contexte, malgré le nombre relativement important de recherches ayant été conduites, il demeure souvent hasardeux de se fier aux conclusions qui sont émises.

De façon connexe, il a été remarqué chez certains auteurs l'emploi d'une grande souplesse dans l'analyse et l'interprétation des résultats, tout particulièrement dans la perspective d'équivalence qui était étudiée. Ce manque de prudence et de rigueur en lien avec les principes statistiques utilisés est préoccupant, considérant tous les risques potentiels qui peuvent être prématurément mis de côté, sans avoir été dûment évalués.

4. IDENTIFICATION DU BESOIN DE RECHERCHE

La recension d'écrits a répertorié plusieurs études ayant obtenu des résultats contradictoires quant à l'équivalence des formats papier-crayon et Internet des tests psychométriques, notamment les différences entre les moyennes et les structures factorielles. Ces différences pourraient être cruciales sur la valeur des tests. Par exemple, lorsqu'on ne retrouve pas une équivalence dans les structures factorielles, cela signifie que les deux formats de l'outil psychométrique ne mesurent pas exactement le même concept et ne peuvent donc pas être utilisés de manière interchangeable, comme c'est le cas majoritairement actuellement. Les interprétations, provenant de tests administrés par Internet n'ayant pas été préalablement validés empiriquement sous ce format, peuvent donc être erronées et avoir des répercussions sur les parcours de vie des personnes répondantes.

Au-delà des résultats contradictoires, ce qui ressort dans l'ensemble, ce sont le faible nombre d'études portant sur le sujet ayant utilisé un devis de recherche limitant les biais méthodologiques connus. Considérant la popularité de ce type d'administration chez les professionnels de l'orientation, il convient de développer une recherche visant à vérifier l'équivalence des résultats et le maintien des propriétés psychométriques des tests lors d'un changement de modalité d'administration en palliant certaines lacunes des études recensées. Par exemple, l'utilisation d'un devis à mesures répétées permettrait d'être plus confiant que les différences, s'il y a lieu, ne sont pas dues à des différences

d'échantillonnage. Ce biais méthodologique était fréquent dans la recension effectuée. De plus, une taille échantillonnale plus grande permettrait d'obtenir un niveau de puissance statistique suffisant pour détecter des différences significatives qui peuvent avoir échappé aux études menées sur des échantillons plus petits.

5. OBJECTIFS DE RECHERCHE

L'objectif général de la recherche est la vérification de l'équivalence des résultats et du maintien de certaines propriétés psychométriques d'un test, suite à un changement de modalité d'administration. Les objectifs spécifiques de recherche sont les suivants :

1. Vérifier l'équivalence des distributions des scores (moyenne, écart-type, asymétrie et aplatissement) entre les formats papier-crayon et Internet d'un inventaire de personnalité.
2. Vérifier l'équivalence de la fidélité entre les formats papier-crayon et Internet d'un inventaire de personnalité.

TROISIÈME CHAPITRE– MÉTHODOLOGIE

Le troisième chapitre détaille les aspects méthodologiques liés à l'opérationnalisation du projet de recherche. Le devis de recherche est d'abord présenté, suivi de la description de l'échantillon et des instruments de mesure choisis. Les étapes de la collecte et l'analyse des données sont également précisées. Le recours à un échantillon humain a exigé une attention particulière quant aux considérations éthiques qui sont explicitées dans le présent chapitre.

1. DEVIS DE RECHERCHE

Compte tenu de la recension de la littérature effectuée qui nous a amenés à préciser les objectifs de recherche, le devis de recherche quasi expérimental croisé vise à vérifier l'effet du changement de modalité d'administration sur les normes et la fidélité d'un test de personnalité. La présente étude s'insère au sein d'un projet de recherche plus large dirigé par le professeur Le Corff (L'évaluation de la personnalité à l'ère d'Internet : étude de son équivalence entre son évaluation psychométrique par Internet et sous forme traditionnelle papier-crayon).

2. ÉCHANTILLON

Le mode d'échantillonnage est de type non probabiliste volontaire. L'échantillon est composé de 407 étudiantes ($n = 336$) et étudiants ($n = 71$) de deux facultés (Éducation, Lettres et sciences humaines) de l'Université de Sherbrooke. Avec cette taille d'échantillon, il est possible de détecter des petites tailles d'effet ($d = 0,20$) à un seuil alpha de 0,05 et une puissance de 0,80. Le choix de faire appel à des étudiants universitaires pour composer notre échantillon est justifié en fonction de la proximité du chercheur principal et des assistantes de recherche avec le campus universitaire. Outre ce facteur de convenance, il est aussi probable que la majorité de l'échantillon ait un degré de familiarité assez élevé avec Internet. En réduisant l'écart possible entre les participants, on diminue

l'effet d'une variable extérieure (la connaissance de l'informatique et d'Internet) qui aurait pu avoir une influence sur les résultats.

3. INSTRUMENTS DE MESURE

Un premier questionnaire de six questions a servi à obtenir les données descriptives concernant l'âge, le sexe, la faculté d'appartenance, le cycle d'études et l'origine ethnique des participants.

Le deuxième questionnaire présenté était l'Inventaire de personnalité Le Corff (IPLC; Le Corff, 2014). Il s'agit d'un questionnaire contenant 104 items auto-rapportés de type Likert en cinq points (fortement en désaccord, plutôt en désaccord, ni en accord ni en désaccord, plutôt en accord, fortement d'accord) (annexe A). Il est basé sur le modèle théorique des cinq facteurs, communément nommé le « Big Five » (Goldberg, 1990). Ce modèle, empiriquement reconnu et très populaire, représente la personnalité en un construit psychologique décomposé en cinq grandes dimensions, nommé facteurs. Il s'agit du modèle de la personnalité le plus validé empiriquement (Mervielde, De Clercq, De Fruyt et Van Leeuwen, 2005; Saucier et Goldberg, 2001).

L'échelle Ouverture est composée de 21 items. Dans l'échantillon normatif, son coefficient alpha de Cronbach est 0,72 et son coefficient de stabilité test-retest sur une à trois semaines est de 0,82. L'échelle Consciencieux est composée de 20 items. Son coefficient alpha de Cronbach est 0,88 et son coefficient de stabilité test-retest est de 0,89. L'échelle Extraversion comporte 22 items. Son coefficient alpha de Cronbach est 0,88 et son coefficient de stabilité test-retest est de 0,85. L'échelle Agréabilité composée de 21 items a obtenu un coefficient alpha de Cronbach de 0,73 et un coefficient de stabilité test-retest de 0,74. L'échelle Stabilité émotionnelle composée de 22 items a obtenu un coefficient alpha de Cronbach de 0,90 et un coefficient de stabilité test-retest de 0,88 (Le Corff, 2014).

Sur le plan de la validité, un bon niveau de convergence a été obtenu entre les échelles de l'IPLC et celles du NEO-PI-3. Les coefficients de convergence sont respectivement de 0,79 pour l'échelle Ouverture, 0,83 pour l'échelle Conscientieux, 0,80 pour l'échelle Extraversion, 0,62 pour l'échelle Agréabilité et -0,88 pour l'échelle Stabilité émotionnelle (Névrosisme). La convergence plus faible entre les échelles Agréabilité de l'IPLC et du NEO-PI-3 s'expliquerait par une conceptualisation théorique quelque peu différente (Le Corff, 2014).

Nous avons choisi l'IPLC au lieu du NEO-PI-3, outil fréquemment utilisé en recherche sur la personnalité, pour des raisons pratiques. Il n'était pas possible d'utiliser le NEO-PI-3 sur Internet en y annexant d'autres questionnaires, ce qui aurait rendu la collecte de données complexe en plus d'ajouter un biais important sur le plan de la présentation des questionnaires. De plus, l'utilisation du NEO-PI-3 aurait entraîné des frais supplémentaires importants alors que l'utilisation de l'IPLC ne faisait encourir aucun frais. Enfin, l'IPLC est un instrument québécois alors que le NEO-PI-3 est un instrument traduit de l'américain et dont les propriétés psychométriques de la version traduite n'ont pas été étudiées.

4. PROCÉDURE

Chaque questionnaire a été transféré sur papier et sur Internet, via le logiciel d'interface LimeSurvey, en reproduisant le plus fidèlement possible les modes de réponse et de présentation de manière à diminuer les biais d'instrumentation possibles. Compte tenu de la recherche plus large dans laquelle s'intègre le présent projet, d'autres instruments, soit un questionnaire mesurant le niveau d'autodévoilement, la version X2 de l'Échelle de désirabilité sociale Marlowe-Crowne et un questionnaire sur les perceptions et réactions envers les tests Internet, ont aussi été administrés aux participants. Cependant, les données de ces autres questionnaires n'ont pas été utilisées dans ce mémoire. La durée totale de la passation des questionnaires était d'environ trente minutes.

Les assistantes de recherche rattachées au projet de recherche sur l'évaluation de la personnalité par Internet se sont rendues dans les classes de l'Université de Sherbrooke pour administrer le questionnaire aux participants volontaires en format papier-crayon durant des périodes déterminées et supervisées. Les participants étaient invités par courriel à compléter le test en format Internet qui était téléaccessible au moment et lieu désiré via le logiciel LimeSurvey. Le lien Internet leur était transmis par courriel. Environ la moitié de l'échantillon était aléatoirement affectée à compléter en format Internet d'abord, et en format papier-crayon ensuite. La procédure était inversée pour la seconde moitié de l'échantillon afin de contrôler l'effet possible de l'ordre d'administration. Le délai entre les deux passations allait d'une à trois semaines selon la disponibilité des participants. Chaque participant a reçu un léger dédommagement financier de vingt dollars suite à la réalisation des deux passations. Un code d'identification constitué des six derniers chiffres du matricule étudiant permet de jumeler les deux passations de chaque participant pour procéder aux analyses statistiques.

Le choix d'un devis à mesures répétées (*within between subjects*) présente des avantages importants, car il permet « à la fois de réduire la variabilité globale en utilisant un pool commun de sujets pour tous les traitements, et d'enlever de notre terme d'erreur les différences liées au sujet » (Howell, 1998, p.508). Comparativement au devis intersujet, cela permet de comparer la performance des participants entre les deux conditions d'expérimentation (Evans, 2007).

5. PLAN D'ANALYSE

Les données descriptives du premier questionnaire composé de variables nominales et ordinales ont été analysées afin de rendre compte de la composition de l'échantillon. Pour répondre aux deux premiers objectifs de recherche, nous avons comparé les moyennes, les écarts-types, les indices d'asymétrie et d'aplatissement, ainsi que les indices de fidélité de chacune des cinq échelles de l'IPLC. La technique de Feldt (Charter et Feldt, 1996) été retenue pour tester la

signification de l'écart entre les coefficients de fidélité issues des deux modes d'administration. Celle-ci permet une meilleure puissance statistique. Nous avons utilisé le logiciel SPSS pour conduire les analyses statistiques descriptives et comparatives.

6. CONSIDÉRATIONS ÉTHIQUES

Un formulaire de consentement libre et éclairé, joint en annexe B et C, a été présenté aux participants avant la passation, spécifiant leur droit de se retirer en tout temps sans préjudice et le caractère volontaire et confidentiel de leur participation. En outre, le document expliquait le projet de recherche, les bénéfices attendus, les risques et inconvénients liés à la participation, la gestion des données collectées, le respect de la confidentialité, les sources de financement obtenues et les coordonnées des personnes responsables. La confidentialité des données a été assurée par l'utilisation d'un code d'identification chiffré pour chaque participant. Ces données sont enregistrées sur un serveur informatique sécurisé et seules les deux assistantes de recherche les utilisant dans le cadre de leur mémoire et le professeur-chercheur y ont accès par le biais d'un mot de passe. Toutes les données brutes seront détruites cinq ans après la fin de la publication des résultats. Le projet de recherche a reçu l'approbation de certification éthique du comité d'éthique de la recherche Éducation et sciences sociales de l'Université de Sherbrooke (Annexe D).

QUATRIÈME CHAPITRE– PRÉSENTATION DES RÉSULTATS

Ce chapitre décrit les résultats des analyses statistiques effectuées afin de répondre aux objectifs de recherche. En premier lieu sont présentés les caractéristiques de l'échantillon, puis les données portant sur l'équivalence des distributions de scores et des coefficients de cohérence interne des deux modalités d'administration.

1. DESCRIPTION DE L'ÉCHANTILLON

L'échantillon est composé de 407 personnes étudiantes au niveau universitaire ayant rempli les questionnaires en format papier-crayon et Internet. L'âge moyen est de 21,8 ans avec un écart-type de 3,2 ans. L'échantillon compte 82,6 % de femmes (n = 336) et 17,4 % d'hommes (n = 71) (Tableau 1). La grande majorité de l'échantillon total (97,5 %) a le français comme langue maternelle. Considérant que la totalité des participants étudiait à l'Université de Sherbrooke dans un programme francophone, nous avons choisi de conserver les participants n'ayant pas le français comme langue maternelle. Une forte proportion de l'échantillon étudie à la Faculté d'éducation (85,5 %), les autres proviennent de la Faculté des lettres et sciences humaines (14,5 %) (Tableau 1). Enfin, 90,2 % de l'échantillon étudiait dans un programme de premier cycle alors que 9,8 % étudiaient au deuxième cycle.

Tableau 1
Composition de l'échantillon (%)

Sexe	Femme	82,6
	Homme	17,4
Faculté	Éducation	85,5
	Lettres et sciences humaines	14,5

2. ÉQUIVALENCE DES DISTRIBUTIONS DES SCORES

Les résultats présentés dans cette section répondent au premier objectif du mémoire, soit la vérification de l'équivalence des distributions des scores entre les formats papier-crayon et Internet d'un inventaire de personnalité. Pour ce faire, les résultats des moyennes, des écarts-types et des indices d'asymétrie et d'aplatissement sont comparés.

La comparaison des moyennes et des écarts-types a été réalisée à l'aide du test t pour échantillons pairs dont les résultats sont présentés au Tableau 2. Les scores moyens des cinq échelles ont été comparés selon les modalités d'administration Internet et papier-crayon. Des différences statistiquement significatives ont été observées sur quatre des cinq échelles (Ouverture, Conscience, Extraversion et Agréabilité). L'échelle Ouverture a obtenu des scores plus élevés [$t(406) = -2,21, p < 0,05$] en format papier-crayon ($M = 54,88$) qu'en format Internet ($M = 54,44$), ce qui a été le cas aussi pour l'échelle Conscience ($M = 52,45$ contre $51,57$) [$t(406) = -4,42, p < 0,05$]. L'échelle Extraversion a également obtenu des résultats significativement plus élevés en format papier-crayon ($M = 59,74$ contre $58,69$) [$t(406) = -5,29, p < 0,05$], tout comme l'échelle Agréabilité ($M = 55,98$ contre $55,13$) [$t(406) = -4,54, p < 0,05$]. Enfin, les résultats concernant l'échelle Stabilité émotionnelle n'étaient pas statistiquement significatifs ($M = 44,38$ contre $44,51$) [$t(406) = 0,48; p < 0,05$].

Afin de rendre compte de l'ampleur des écarts de moyenne observés sur les échelles, les coefficients d de Cohen ont été calculés en utilisant la formule de Morris et DeShon's (2002) pour devis intrasujets. Les échelles Ouverture ($d = 0,11$) et Stabilité émotionnelle ($d = 0,02$) ont obtenu des coefficients inférieurs à 0,20 indiquant ainsi des très petites tailles d'effet, conformément au barème généralement reconnu (Cohen, 1988). Les échelles Conscience ($d = 0,22$), Extraversion ($d = 0,26$) et Agréabilité ($d = 0,23$) ont obtenu des coefficients légèrement supérieurs au seuil de 0,20, différences pouvant alors être qualifiées de faibles (Cohen, 1988).

Tableau 2

Comparaison des scores moyens des échelles de l'IPLC selon le format d'administration utilisé (Internet ou papier-crayon)

Échelles de l'IPLC	Format Internet		Format papier-crayon		t	d
	M	ET	M	ET		
Ouverture	54,44	(8,21)	54,88	(8,15)	-2,21*	0,11
Conscience	51,57	(7,37)	52,45	(7,75)	-4,41*	0,22
Extraversion	58,69	(9,85)	59,74	(10,19)	-5,29*	0,26
Agréabilité	55,13	(6,91)	55,98	(6,75)	-4,54*	0,23
Stabilité émotionnelle	44,51	(11,93)	44,39	(11,85)	0,48	0,02

N = 407; * $p < 0,05$

La vérification de l'équivalence des indices d'asymétrie et d'aplatissement permet de savoir si les distributions des scores diffèrent significativement ou non selon le format utilisé. Nous avons calculé les intervalles de confiance à 95 % autour des scores d'asymétrie et d'aplatissement afin de vérifier leur similitude entre les deux modalités d'administration. Ainsi, si les intervalles de confiance se

chevauchent, on ne peut pas affirmer que les deux scores sont statistiquement différents, alors que si les intervalles de confiance ne se chevauchent pas, la différence entre les scores est statistiquement significative au seuil alpha de 0,05. Tel que présenté au Tableau 3, aucune différence significative n'a été observée.

Tableau 3
Comparaison des indices d'asymétrie et d'aplatissement

Asymétrie					
Échelles de l'IPLC	Format Internet	Intervalle de confiance (95%)	ES	Format p.-crayon	Intervalle de confiance (95%)
Ouverture	0,11	[-0,13; 0,35]	(0,12)	0,07	[-0,17; 0,31]
Conscience	-0,28	[-0,52; -0,04]	(0,12)	-0,35	[-0,59; -0,12]
Extraversion	-0,06	[-0,79; -0,31]	(0,12)	-0,28	[-0,51; -0,04]
Agréabilité	-0,76	[-0,99; -0,52]	(0,12)	-0,66	[-0,89; -0,42]
Stabilité émotionnelle	-0,44	[-0,67; -0,20]	(0,12)	-0,36	[-0,60; -0,13]
Aplatissement					
Ouverture	-0,30	[-0,76; 0,17]	(0,24)	-0,38	[-0,85; 0,09]
Conscience	0,18	[-0,30; 0,65]	(0,24)	-0,06	[-0,53; 0,41]
Extraversion	-0,23	[-0,70; 0,24]	(0,24)	-0,17	[-0,64; 0,30]
Agréabilité	2,07	[1,60; 2,54]	(0,24)	1,54	[1,07; 2,01]
Stabilité émotionnelle	-0,02	[-0,49; 0,45]	(0,24)	-0,28	[-0,75; 0,19]

ES = Erreur standard

3. VÉRIFICATION DE LA FIDÉLITÉ

Nous avons tout d'abord obtenu les coefficients de cohérence interne de chacune des échelles, dans les deux modalités d'administration. Tel que présenté au Tableau 4, les coefficients de cohérence interne en format papier-crayon varient entre 0,71 et 0,87. En format Internet, l'étendue est similaire, de 0,72 à 0,85. Les coefficients de corrélation de Pearson inter-échelles varient entre 0,85 et 0,92, indiquant une stabilité test-retest très élevée entre les deux formats. Nous avons utilisé la technique de Charter et Feldt (1996) qui permet la comparaison de coefficients de cohérence interne dépendants afin de vérifier l'équivalence de formes similaires d'un même test. En voici l'équation :

$$t_{N-2} = \frac{|\hat{\alpha}_1 - \hat{\alpha}_2| \sqrt{N-2}}{\sqrt{4(1-\hat{\alpha}_1)(1-\hat{\alpha}_2)(1-r_{12}^2)}}$$

Dans l'équation, le $\hat{\alpha}$ correspond aux alphas de Cronbach, le r au coefficient de corrélation entre les scores totaux aux tests et le N au nombre total de participants. Les scores t obtenus pour les cinq échelles ont été comparés aux valeurs critiques de la distribution t de Student afin de déterminer la présence de différences significatives.

La procédure a permis d'observer des différences significatives pour deux des cinq échelles, soit Agréabilité ($t = 2,38, p < 0,02$) et Stabilité émotionnelle ($t = 4,10, p < 0,001$). Il importe de noter que ce test apparaît très sensible, ayant détecté comme significative des différences de 0,02 et 0,03 entre les coefficients alphas. Il est également intéressant de remarquer que plus la taille de l'échantillon (N), plus la corrélation entre les scores (r) et plus les coefficients alphas ($\hat{\alpha}$) seront élevés, plus le score t obtenu sera élevé. C'est donc dire qu'un même écart absolu de fidélité pourrait être significatif pour une échelle ayant une fidélité et une intercorrélations élevées et non-significatif pour une échelle dont ces indices seraient plus faibles. Cela pourrait notamment expliquer que la différence de 0,02 pour l'échelle de Stabilité émotionnelle soit significative alors qu'elle ne l'est pas pour

l'échelle Ouverture; la première affiche une fidélité et une intercorrélations supérieures à la seconde.

Tableau 4
Coefficients de cohérence interne

Échelles de l'IPLC	Format Internet	Format papier-crayon	<i>t</i>	<i>r</i>
	α	α		
Ouverture	0,78	0,76	1,82	0,88
Conscience	0,73	0,73	0,00	0,86
Extraversion	0,86	0,86	0,00	0,92
Agréabilité	0,78	0,75	2,38*	0,85
Stabilité émotionnelle	0,89	0,87	4,10**	0,91

* $p < 0,02$, ** $p < 0,001$

CINQUIÈME CHAPITRE – DISCUSSION DES RÉSULTATS

Le cinquième chapitre fournit des éléments d'explication aux résultats précédemment présentés à l'aide de pistes théoriques. Les résultats sont également éclairés par les informations issues de la recension d'écrits ayant mené à la question de recherche. Au terme de la discussion, les limites de l'étude et des pistes de recherches futures sont proposées.

1. RÉPONSES AUX OBJECTIFS DE RECHERCHE

Le présent travail avait pour objectif de recherche la vérification de l'équivalence des distributions des scores et des coefficients de cohérence interne entre les formats papier-crayon et Internet d'un inventaire de personnalité. La recension d'écrits indiquait des résultats contradictoires, certaines études avançant que le changement de modalité d'administration pouvait modifier les propriétés psychométriques et les distributions de scores d'un test, d'autres ne suggérant aucun effet. Les analyses statistiques effectuées ont permis de répondre aux objectifs en analysant les résultats.

1.1 Premier objectif de recherche

Le devis à mesures répétées a permis de comparer les scores d'un vaste échantillon de 407 personnes dans les deux modalités d'administration afin de vérifier l'équivalence des distributions de scores. Le test t pour échantillons pairés a permis d'observer des différences significatives sur quatre échelles (Ouverture, Conscience, Extraversion et Agréabilité) entre les moyennes. Cependant, les tailles d'effet ont toutes été petites. Tous les indices d'asymétrie et d'aplatissement ont été similaires entre les deux modalités d'administration. Ces résultats supportent l'équivalence des distributions et donc, que les mêmes normes soient utilisées pour les deux formats.

1.2 Deuxième objectif de recherche

Les coefficients de cohérence interne obtenus pour les deux modalités d'administration ont été analysés statistiquement en utilisant la formule de Feldt pour échantillons dépendants (1980). Deux différences significatives ont été observées pour les échelles Agréabilité et Stabilité émotionnelle, mais ces différences nous apparaissent être de faible taille. Les résultats obtenus lors de la présente recherche, ne faisant état d'aucune différence significative de moyenne ou grande taille, suggèrent que les deux formats peuvent être utilisés de façon interchangeable. La fidélité fait partie des propriétés psychométriques faisant d'un test un instrument de mesure rigoureux. En procédant à la vérification de la fidélité entre les modalités d'administration papier-crayon et Internet, on s'assure du maintien de cette propriété, considérée comme le niveau de précision de l'outil.

2. PISTES THÉORIQUES EXPLICATIVES

La recension d'écrits nous avait pistés sur quelques hypothèses afin d'expliquer les possibles différences entre les formats d'un même test suite à un changement de modalité d'administration. L'influence du format, le niveau de supervision, la perception de la confidentialité composaient les pistes explicatives les plus fréquentes. En raison des faibles différences observées, ces hypothèses n'ont pas été supportées par les résultats.

2.1 L'influence du format

L'interprétation d'Hirai *et al.* (2011) suggérait que l'administration par Internet était un changement suffisant pour modifier la façon de répondre des participants au test sur l'anxiété sociale. Le manque de confiance envers Internet était aussi nommé comme une explication plausible des différences observées (Ployhart *et al.*, 2003). Les différences observées dans la présente étude ne supportent pas cette hypothèse. Il se peut que la familiarité de l'échantillon avec Internet et l'informatique aient réduit l'effet possible de l'influence du format.

2.2 Le niveau de supervision

Le niveau de supervision est une variable fréquemment discutée en lien avec la validité du testing Internet. L'absence de standardisation de l'administration soulève particulièrement des questions (Pearlman, 2009; Tippins, 2009). Nous l'avons d'ailleurs considérée dans la recension d'écrits en distinguant les études selon si la passation Internet était supervisée. En fonction des résultats de leur étude, Carstairs et Myors (2009) stipulaient que la supervision était un facteur beaucoup plus important que la modalité d'administration (papier-crayon ou Internet). Malgré le fait que le devis utilisé ne permettait pas d'étudier l'effet de la supervision, les résultats ne soutiennent pas cette hypothèse. Cela sera plus amplement discuté dans la section suivante.

2.3 La perception de la confidentialité

Ployhart et al. (2003) avançaient le niveau perçu de la confidentialité comme hypothèse pouvant expliquer les différences significatives observées auprès d'un vaste échantillon. L'absence de grandes différences significatives suggère que les personnes étant assez familières avec l'informatique, tel que c'est le cas dans notre échantillon, ne perçoivent pas le niveau de confidentialité différemment lorsqu'ils fournissent des données de façon papier-crayon ou face à un ordinateur.

3. COMPARAISON DES RÉSULTATS AUX ÉTUDES RECENSÉES

La recension d'écrits avait permis de mettre en évidence le faible nombre d'études empiriques ayant utilisé un devis à mesures répétées qui permet « un contrôle rigoureux de l'influence potentielle des différences individuelles » (Fortin, 2010, p.343). D'autres, ayant utilisé un devis intrasujets, n'avaient pas utilisé un devis croisé afin d'éviter l'effet de séquence. Les choix méthodologiques que nous avons effectués ont permis de pallier ces faiblesses. Nous avons également fait l'usage de la formule de Feldt afin de comparer

statistiquement les coefficients de fidélité, ce qui n'avait été fait dans aucune des études recensées.

En fait, nos résultats corroborent ceux de Salgado et Moscoso (2003) qui avaient utilisé un devis à mesures répétées croisé afin de vérifier l'équivalence des scores entre les modalités papier-crayon et Internet d'un test de personnalité. Ces derniers n'avaient également pas obtenu de différence significative et les effets de taille étaient tous inférieurs au seuil de 0,20. À la lumière des résultats de Salgado et Moscoso (2003) et ceux de la présente étude, on peut suggérer que les écarts entre les modalités d'administration répertoriées dans la recension sont attribuables aux limites méthodologiques des études et non à de réelles différences relatives au format.

Dans la section synthèse de la recension d'écrits, nous remarquons parmi toutes les études retenues que les différences significatives au niveau de la fidélité étaient moins fréquentes que les différences significatives relatives aux moyennes. De plus, aucune différence n'avait été relevée lorsque la passation Internet était supervisée. Les résultats de notre étude sont cohérents avec ce constat, avec deux différences significatives de faible taille.

Conformément au devis de recherche, l'effet de la supervision n'a pu être documenté, car la passation papier-crayon était supervisée, alors que la passation Internet ne l'était pas. Cependant, au terme de la recension d'écrits, il avait été remarqué qu'indépendamment que la passation Internet ait été supervisée ou non, des différences significatives avaient été trouvées. Dans notre étude, en présence de seulement quelques différences significatives mais de petits effets de taille, ces résultats indiquent que le fait d'être supervisé ou non n'a pas entraîné de différences sur les scores à l'inventaire de personnalité. Cette observation est importante, considérant que l'administration de tests par Internet sans supervision est très fréquente dans la pratique des conseillères et conseillers d'orientation. L'absence de différences notables soutient également que l'administration en

contexte de groupe versus individuel n'a pas eu d'influence au niveau des distributions des scores et de la fidélité.

4. IMPLICATIONS DES RÉSULTATS

Lors de la présentation de l'état de la situation dans le premier chapitre, il était question de la nécessité d'obtenir plus de données empiriques afin de vérifier l'influence d'un changement de format sur les normes et les propriétés psychométriques des instruments psychométriques. L'usage grandissant des tests en format Internet par les conseillers d'orientation était documenté. Pour la pratique professionnelle, les résultats de la présente étude supportent l'utilisation interchangeable des deux formats pour l'évaluation psychométrique de la personnalité. Parmi les différences trouvées, toutes ont obtenu un petit effet de taille et n'auraient pas d'impact notable sur l'interprétation de résultats. Ces donc dire que les mêmes normes pourraient être utilisées pour les deux formats, que celles-ci aient été établies en format papier-crayon ou en format Internet. L'équivalence des indices de fidélité supporte également l'équivalence des formats papier-crayon et Internet.

Ces résultats fournissent des données empiriques supplémentaires quant à l'application de lignes directrices concernant le testing Internet qui est de plus en plus utilisé. De nouvelles normes de pratique tenant compte des enjeux, des risques et des avantages pourraient être élaborées de manière à uniformiser les pratiques et à guider les professionnels faisant usage de la psychométrie.

5. FORCES

Tel que déjà mentionné, le devis à mesures répétées (intrasujet) constitue une force importante du projet, car il a permis de s'assurer que les différences observées ne soient pas dues à des différences propres à l'échantillon. De plus, l'inversion de l'ordre de passation pour une partie de l'échantillon a empêché un effet de séquence possible. La grande taille de l'échantillon ($n = 407$) augmente également la puissance statistique des analyses effectuées. Pour plusieurs des

études colligées, la petite taille de l'échantillon était une faiblesse méthodologique.

6. LIMITES

Le fait que l'échantillon soit entièrement composé d'étudiants universitaires et majoritairement féminin et provenant des domaines de l'éducation et des lettres et sciences humaines et sociales constitue une limite de la recherche. En effet, en raison de ces caractéristiques, il convient de faire preuve de prudence dans la généralisabilité des résultats. Toutefois, il est intéressant de relever que Whitaker (2007) n'avait pas observé de différence entre les genres dans son étude sur le sujet.

Il est notamment possible de supposer que les étudiants universitaires ont une plus grande familiarité avec l'utilisation d'Internet que la population générale. D'autres recherches devraient être menées pour vérifier l'équivalence auprès de personnes moins familières avec l'informatique et Internet.

Il convient également de rappeler que l'instrument de mesure utilisé évaluait la personnalité. Il a été mentionné dans la recension d'écrits que le type de construit mesuré pouvait avoir un impact différent sur le changement de modalité d'administration. Par exemple, l'évaluation des aptitudes par Internet induit davantage la question du testing non-supervisé et la possibilité de tricher. Il s'agit d'un enjeu beaucoup moins important en évaluation de la personnalité. Il est donc possible que les tests d'aptitudes ou de psychopathologie, par exemple soient plus sensibles à un changement de modalité d'administration.

Dans le même ordre d'idées, il faut rappeler que les personnes ayant accepté de participer à l'étude le faisaient dans un contexte comportant un enjeu relativement faible. Cette caractéristique pourrait influencer les résultats et donc l'équivalence des formats, particulièrement lorsque celle-ci est non supervisée, comme l'avaient relevé Carstairs et Myors (2009) dans leur étude en lien avec la sélection de personnel. Donc, bien que les résultats de la présente étude

soutiennent l'usage des deux formats de façon interchangeable, avec ou sans supervision, le choix de la modalité d'administration devrait toujours se faire en analysant rigoureusement les risques potentiels, le contexte et le construit mesuré. Ces aspects constituent des pistes de recherche intéressantes.

CONCLUSION

La recension des écrits portant sur les effets d'un changement de modalité d'administration d'un test psychométrique, faisait état de résultats contradictoires, suggérant parfois des différences notables des propriétés psychométriques (normes, fidélité, validité de contenu). Cependant, le faible nombre d'études recensées utilisant un devis méthodologique à mesures répétées a surtout permis d'identifier cette lacune importante de la documentation scientifique sur la problématique à l'étude. À la lumière des informations, l'objectif de recherche retenu a été la vérification de l'équivalence des distributions de scores et des coefficients de cohérence interne, suite à un changement de modalité d'administration de papier-crayon à Internet. Pour ce faire, un devis à mesures répétées croisé a été utilisé auprès d'un échantillon d'environ quatre cents étudiants universitaires.

Les résultats des analyses des scores obtenus de façon papier-crayon et par Internet font part d'un faible nombre de différences significatives (moyennes, écarts-types, indices d'asymétrie et d'aplatissement). Celles-ci ont toutes été de petite taille. Concernant le deuxième objectif de vérification des coefficients de cohérence interne, deux échelles ont obtenu des différences relativement faibles.

Ce mémoire a permis de fournir des résultats empiriques quant à une pratique grandissante chez les conseillères et conseillers d'orientation québécois, en palliant certaines faiblesses répertoriées au terme de la recension des écrits. La taille d'échantillon constitue une force du projet, alors que l'âge, les caractéristiques de l'échantillon et le type de construit mesuré restreignent la généralisation des résultats. Des études sur le même thème avec un échantillon plus hétérogène, des personnes moins familières à l'usage d'Internet ou dans un contexte avec un enjeu réel important seraient des pistes de recherche pertinentes.

RÉFÉRENCES BIBLIOGRAPHIQUES

- Aluja, A., Rossier, J. et Zuckerman, M. (2007). Equivalence of paper and pencil vs Internet forms of the ZKPQ-50-CC in Spanish and French samples. *Personality and Individual Differences*, 43, 2022-2032.
- American Educational Research Association, American Psychological Association et National Council on Measurement in Education. (2003). *Normes de pratique de testing en psychologie et en éducation* (Trad. par G. Sarrazin). Montréal : Institut de recherches psychologiques.
- Anastasi, A. (1994). *Introduction à la psychométrie* (Trad. par F. Gagné.) Montréal : Guérin Universitaire.
- Barak, A. (1999). Psychological applications on the Internet: A discipline on the threshold of a new millennium. *Applied and Preventive Psychology*, 8, 231-245.
- Barak, A. et English, N. (2002). Prospects and limitations of psychological testing on the Internet. *Journal of Technology in Human Services*, 19(2/3), 65-89.
- Barbeite, F.G. et Weiss, E.M. (2004). Computer self-efficacy and anxiety scales for an Internet sample: testing measurement equivalence of existing measures and development of new scales. *Computers in Human Behavior*, 20, 1-15.
- Bartram, D. (2009). Commentaries : The International Test Commission guidelines on computer-based and Internet-delivered testing. *Industrial and Organizational Psychology*, 2, 11-13.
- Bartram, D. et Brown, A. (2004). Online Testing : Mode of Administration and the stability of OPQ-32i Scores. *International Journal of Selection and Assessment*, 12(3), 278-284.
- Bernaudo, J.-L. (2007). *Introduction à la psychométrie*. Paris : Dunod.
- Buchanan, T. (2000). Potential of the Internet for personality research. In M.H. Birnbaum (dir.), *Psychological experiments on the Internet* (p.121-140). San Diego, CA : Academic Press.
- Buchanan, T. et Smith, J.L. (1999). Using the Internet for psychological research : Personality testing on the World Wide Web. *British Journal of Psychology*, 90(1), 125-144.
- Carlbring, P., Brunt, S., Bohman, S., Austin, D., Richards, J., Öst, L.-G. et Andersson, G. (2007). Internet vs paper and pencil administration of

- questionnaires commonly used in panic/agoraphobic research. *Computers in Human Behavior*, 23, 1421-1434.
- Carstairs, J. et Myers, B. (2009). Internet testing : A natural experiment reveals test score inflation on a high-stakes, unproctored cognitive test. *Computers in Human Behavior*, 29(3), 738-742.
- Charter, R.A. et Feldt, L.S. (1996). Testing the equality of two alpha coefficients. *Perceptual and motor skills*, 82, 763-768.
- Chuah, S.C., Drasgow, F. et Roberts, B.W. (2006). Personality assessment: Does the medium matter? No. *Journal of Research in Personality*, 40, 359-376.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2^e éd.). New York : Academy Press.
- Coles, M.E., Cook, L.M. et Blake, T.R. (2006). Assessing obsessive compulsive symptoms and cognitions on the Internet: Evidence for the comparability of paper and Internet administration. *Behaviour Research and Therapy*, 45, 2232-2240.
- Coyne, I., Warszta, T., Beadle, S. et Sheehan, N. (2005). The impact of mode of administration on the equivalence of a test battery: A quasi-experimental design. *International Journal of Selection and Assessment*, 13(3), 220-224.
- Cronbach, L.J. (1990). *Essentials of psychological testing* (5^e éd.). New York, NY : Harper Collins (1^{re} éd. 1949).
- Cronk, B.C. et West, J.L. (2002). Personality research on the Internet : A comparison of Web-based and traditional instruments in take-home an in-class settings. *Behavior Research Methods, Instruments, & Computers*, 34(2), 177-180.
- Delisle-Laberge, L. et Locas, R. (2004). D'hier à demain: Historique et problématiques de l'évaluation en orientation. *En pratique*, (2) 4-6.
- Domino, G. et Domino, M.L. (2006). *Psychological Testing: An introduction* (2^e éd.) Upper Saddle River, NJ: Prentice Hall (1^{re} éd. 2000).
- Evans, J. (2007). *Your psychology project: the essential guide*. London : SAGE Publications.
- Feldt, L.S. (1980). A test of the hypothesis that Cronbach's alpha reliability coefficient is the same for two tests administered to the same sample. *Psychometrika*, 45, 99-105.

- Fortin, M.-F. (2010). *Fondements et étapes du processus de recherche: Méthodes quantitatives et qualitatives* (2^e éd.). Montréal: Chenelière éducation (1^{re} éd. 2006).
- Fouladi, R.T., McCarthy, C.J. et Moller, N.P. (2001) *Paper-and-pencil or online? Evaluating coping and attachment measures*. Communication présentée à la 109^e rencontre annuelle de l'American Psychological Association, San Francisco, CA, 24-28 mai.
- Furr, R. M. et Bacharach, V. R. (2008). *Psychometrics: An introduction*. CA: SAGE Publications.
- Gouvernement du Québec (2005). *Rapport du comité d'experts de la modernisation de la pratique professionnelle en santé mentale et en relations humaines*. Québec : Office des professions.
- Grégoire, S. (2010). Pour un usage éthique et professionnel des outils psychométriques en ligne. *En pratique*, 13, 15-18.
- Guédon, M.-C., Savard, R., Le Corff, Y. et Yergeau, É. (2011). *Tests à l'appui: pour une intervention intégrée de la psychométrie en counseling de carrière* (2^e éd.). Québec: Septembre (1^{re} éd. 2000).
- Hardré, P.L., Crowson, H.M. et Xie, K. (2010). Differential effects of Web-based and paper-based administration of questionnaire research instruments in authentic contexts-of-use. *Journal of educational computing research*, 42(1), 103-133.
- Hedman, E., Ljótsson, B., Rück, C., Furmark, T., Carlbring, P., Lindefors, N. et Andersson, G. (2010). Internet administration of self-report measures commonly used in research on social anxiety disorder : A psychometric evaluation. *Computers in Human Behavior*, 26, 736-740.
- Helmes, E. et Campbell, A. (2010). Differential sensitivity to administration format of measures of attitudes toward older adults. *The Gerontologist*, 50(1), 60-65.
- Hirai, M., Vernon, L.L., Clum, G.A. et Skidmore, S.T. (2011). Psychometric properties and administration measurement invariance of Social Phobia Symptom Measures: Paper-pencil vs. Internet administrations. *Journal of Psychopathology and Behavioral Assessment*. 33(4), 470-479.
- International Test Commission (2001). International guidelines for test use. *International Journal of Testing*, 1, 93-114.
- International Test Commission (2006). International guidelines on computer-based and Internet delivered testing. *International Journal of Testing*, 6, 143-172.

- Jones, J.W. et Dages, K.D. (2003). Technology trends in staffing and assessment: A practice note. *International Journal of Selection and Assessment*, 11(2/3), 247-252.
- Joubert, T. et Kriek, H.J. (2009). Psychometric comparison of paper-and-pencil and online personality assessments in a selection setting. *Journal of Industrial Psychology*, 35(1), 78-88.
- Kaplan, R.M. et Saccuzzo, D.P. (1997). *Psychological testing: Principles, Applications and Issues* (4^e éd.). Pacific Grove: Brooks/Cole Publishing (1^{re} éd. 1982).
- Laveault, D. et Grégoire, J. (2002). *Introduction aux théories des tests* (2^e éd.). Bruxelles : De Boeck Université (1^{re} éd. 1997).
- Le Corff, Y. (2014). Inventaire de personnalité Le Corff (IPLC). Manuel professionnel. Montréal : Institut de recherches psychologiques de Montréal.
- Le Corff, Y., Yergeau, E. et Savard, R. (2011). Évaluation psychométrique : enquête sur les pratiques des conseillers d'orientation. *L'Orientation*, 1(1), 18-20.
- Lewis, I., Watson, B. et White, K.M. (2009). Internet versus paper-and-pencil survey methods in psychological experiments: Equivalence testing of participant responses to health-related messages. *Australian Journal of Psychology*, 61(2), 107-116.
- Meade, A.W., Michels, L.C. et Lautenschlager, G.J. (2007). Are Internet and paper-and-pencil personality tests truly comparable? An experimental design measurement invariance study. *Organizational Research Methods*, 10, 322-346.
- Meyerson, P. et Tryon, W.W. (2003). Validating Internet research: A test of the psychometric equivalence of Internet and in-person samples. *Behavior Research Methods, Instruments, & Computers*, 35(4), 614-620.
- Naus, M.J., Philipp, L.M. et Samsi, M. (2009) From paper to pixels: A comparison of paper and computer formats in psychological assessment. *Computers in Human Behavior*, 25, 1-7.
- Nezami, E. et Butcher, J.N. (2000). Objective personality assessment. In G.Goldstein et M. Hersen (dir.) *Handbook of psychological assessment* (2^e éd.). New York, NY: Pergamon Press (1^{re} éd. 1990).

- Ordre des conseillers et conseillères d'orientation du Québec (2010). *Site web de l'OCCOQ*. Site téléaccessible à l'adresse <www.orientation.qc.ca>. Consulté le 16 avril 2012.
- Ordre des conseillers et conseillères d'orientation du Québec (s.d.) *Code de déontologie des membres de l'Ordre des conseillers et conseillères d'orientation du Québec*. Montréal : OCCOQ.
- Ordre des conseillers et conseillères d'orientation et des psychoéducateurs et psychoéducatrices du Québec (OCCOPPQ) (2004). *Le profil des compétences générales des conseillers d'orientation*. Montréal : OCCOPPQ.
- Pearlman, K. (2009). Unproctored Internet testing: Practical, legal, and ethical concerns. *Industrial and Organizational Psychology*, 2(1), 14-19.
- Ployhart, R.E., Weekley, J.A., Holtz, B.C. et Kemp, C. (2003). Web-based and paper-and-pencil testing of applicants in a proctored setting: are personality, biodata and situational judgment tests comparable? *Personnel Psychology*, 56(3), 733-752.
- Potosky, D. et Bobko, P. (2004). Selection testing via the Internet: Practical considerations and exploratory empirical findings. *Personnel Psychology*, 57(4), 1003-1034.
- Tippins, N.T. (2009). Internet alternatives to traditional proctored testing: Where are we now? *Industrial and Organizational Psychology*. 2(1), 2-10.
- Université de Sherbrooke. (2011). *Politique en matière d'éthique de la recherche avec des êtres humains*. Sherbrooke: Université de Sherbrooke.
- Whitaker, B.G. (2007). Internet-based attitude assessment: does gender affect measurement equivalence? *Computers in Human Behavior*, 23(3), 1183-1194.

ANNEXE A - QUESTIONNAIRES UTILISÉS DANS LE CADRE DU MÉMOIRE

QUESTIONNAIRE 1 : DONNÉES DESCRIPTIVES

Veillez indiquer les **6 derniers chiffres** de votre matricule étudiant : _____

(Cette information ne sera utilisée que pour associer votre questionnaire Internet à votre questionnaire papier, mais ne permettra pas de vous identifier personnellement.)

Quel âge avez-vous (en années) ? _____ ans

Quel est votre sexe?

1. Homme
2. Femme

Dans quelle faculté étudiez-vous?

1. Administration
2. Droit
3. Éducation
4. Éducation physique et sportive
5. Génie
6. Lettres et sciences humaines
7. Médecine et sciences de la santé
8. Sciences
9. Théologie et études religieuses

Quel est votre cycle d'études?

1. 1^{er} cycle
2. 2^e cycle
3. 3^e cycle

QUESTIONNAIRE 2

Ce questionnaire présente 121 énoncés qui peuvent permettre de décrire votre personnalité. Décrivez-vous honnêtement, en pensant à la façon dont vous êtes actuellement, pas à la façon dont vous aimeriez être dans le futur. Il n'y a pas de bonnes ou de mauvaises réponses. Répondez à toutes les questions.

Veillez indiquer dans quelle mesure vous êtes en accord ou en désaccord avec chaque énoncé, c'est-à-dire à quel point vous pensez que l'énoncé vous représente bien, selon le choix de réponses suivant :

FD = Fortement en désaccord

D = En désaccord

N = Neutre

A = D'accord

FA = Fortement d'accord

1	Je suis heureux à faire du travail qui ne demande pas de réflexion.	FD D N A FA
2	Je réussirai à atteindre les objectifs que je me suis fixés.	FD D N A FA
3	Je suis un leader naturel.	FD D N A FA
4	Je renonce souvent à des choses pour le bien d'autrui.	FD D N A FA
5	Je trouve que ça en prend beaucoup pour que je sois frustré ou irrité.	FD D N A FA
6	Je reconnais généralement les émotions qui sont représentées dans les différentes formes d'art.	FD D N A FA
7	Je me désengage parfois à la dernière minute.	FD D N A FA
8	J'ai du plaisir à être en compagnie d'autres personnes.	FD D N A FA
9	Je suis gentil avec les personnes contre qui je devrais être fâché.	FD D N A FA
10.	Je me sens souvent dépassé par les événements.	FD D N A FA
11.	J'aime proposer des idées ou façons de faire nouvelles et différentes.	FD D N A FA
12.	J'obtiens ce que je veux parce que je travaille fort pour l'obtenir.	FD D N A FA

13.	J'ai de la difficulté à amener les gens à travailler ensemble.	FD D N A FA
14.	Je serais mal à l'aise que mon succès se fasse aux dépens d'autrui.	FD D N A FA
15.	Je suis une personne dont l'humeur varie facilement.	FD D N A FA
16.	J'apprécie différentes formes d'art.	FD D N A FA
17.	Je fais les choses strictement selon les règles.	FD D N A FA
18.	Je deviens agité lorsque je passe trop de temps à la maison.	FD D N A FA
19.	Je fais mon possible pour éviter les disputes.	FD D N A FA
20.	Je surmonte facilement les revers de la vie.	FD D N A FA
21.	Je peux envisager différents points de vue.	FD D N A FA
22.	Je n'en fais pas plus que ce qui est attendu de moi.	FD D N A FA
23.	Je cherche à influencer les autres.	FD D N A FA
24.	Je donne une chance à tous.	FD D N A FA
25.	Je garde mes émotions sous contrôle.	FD D N A FA
26.	J'ai souvent conscience de la façon dont la couleur et l'éclairage d'une pièce affectent mon humeur.	FD D N A FA
27.	Je ne fais pas de listes de tâches.	FD D N A FA
28.	Je travaille mieux lorsque je suis seul.	FD D N A FA
29.	Je m'inquiète souvent à propos de choses qui s'avèrent finalement sans importance.	FD D N A FA
30.	Je me sens facilement blessé.	FD D N A FA
31.	Je ne suis pas intéressé par les idées abstraites.	FD D N A FA
32.	Je préfère les tâches difficiles aux tâches faciles.	FD D N A FA
33.	J'ai une habileté hors du commun à persuader les autres.	FD D N A FA
34.	Je deviens impatient quand les autres me parlent de leurs problèmes.	FD D N A FA
35.	Je ressens les émotions intensément.	FD D N A FA
36.	J'ai besoin d'un espace pour exprimer ma créativité.	FD D N A FA
37.	Je ne suis pas dérangé par les personnes désordonnées.	FD D N A FA
38.	Je me fais facilement des amis.	FD D N A FA

39.	Je crois qu'il est préférable de pardonner et d'oublier.	FD D N A FA
40.	Je ne donne à personne l'occasion de me blesser une seconde fois.	FD D N A FA
41.	J'aime résoudre des problèmes complexes.	FD D N A FA
42.	Je travaille très fort pour réussir.	FD D N A FA
43.	Je n'exprime pas mes opinions.	FD D N A FA
44.	Je ressens de la culpabilité lorsque je crois avoir fait de la peine à quelqu'un.	FD D N A FA
45.	Je me sens rarement en colère contre les gens.	FD D N A FA
46.	J'aime me perdre dans mes pensées.	FD D N A FA
47.	Je porte attention aux détails.	FD D N A FA
48.	Je n'apprécierais pas un emploi qui nécessite beaucoup d'interaction sociale.	FD D N A FA
49.	Je crois qu'une trop grande modestie cause des ennuis.	FD D N A FA
50.	Je me sens rarement triste pendant plus d'une heure.	FD D N A FA
51.	Je ne me sens pas à l'aise d'avoir à gérer une grande quantité d'information.	FD D N A FA
52.	J'essaie de surpasser les accomplissements des autres.	FD D N A FA
53.	Je prends l'initiative.	FD D N A FA
54.	Je m'informe rarement du bien-être d'autrui.	FD D N A FA
55.	Je réagis fortement à la critique.	FD D N A FA
56.	J'aime les choses claires et simples.	FD D N A FA
57.	Je prends des décisions rapidement.	FD D N A FA
58.	Je parle à plusieurs personnes différentes dans les fêtes ou événements sociaux.	FD D N A FA
59.	J'évite d'avoir l'air supérieur aux autres.	FD D N A FA
60.	Je ne m'inquiète jamais à propos de choses que j'ai dites ou faites.	FD D N A FA
61.	Je suis intéressé par une variété de choses.	FD D N A FA

62.	Je peux rester concentré sur mes tâches même lorsque je suis heureux et excité par un événement à venir.	FD D N A FA
63.	J'aime être en charge d'un groupe ou d'une équipe.	FD D N A FA
64.	Je ne suis jamais trop occupé pour venir en aide à un ami ou à un collègue.	FD D N A FA
65.	Je deviens agressif lorsque je me sens blessé.	FD D N A FA
66.	J'analyse rarement mes émotions.	FD D N A FA
67.	Je prends le temps de peser le pour et le contre lorsque j'ai à faire un choix.	FD D N A FA
68.	Je me dévoile peu aux autres.	FD D N A FA
69.	Je préfère laisser les autres être au centre de l'attention.	FD D N A FA
70.	Je me rappelle davantage de mes échecs que de mes succès.	FD D N A FA
71.	Je suis ouvert aux nouvelles expériences.	FD D N A FA
72.	J'ai tendance à perdre intérêt envers les tâches que j'entreprends.	FD D N A FA
73.	J'aime impliquer les gens dans un projet.	FD D N A FA
74.	J'essaie de faire en sorte que les autres se sentent bienvenus.	FD D N A FA
75.	Je demeure calme sous la pression.	FD D N A FA
76.	J'écoute mes sentiments et intuitions lorsque je prends des décisions importantes.	FD D N A FA
77.	Je ne suis pas toujours à l'heure.	FD D N A FA
78.	J'exprime mes émotions positives telles que la joie et l'enthousiasme.	FD D N A FA
79.	Je ne laisse pas les autres prendre le crédit pour le travail que j'ai fait.	FD D N A FA
80.	Je crois qu'en fin de compte tout ira bien.	FD D N A FA
81.	Je n'aime pas l'inconnu.	FD D N A FA

82.	Je n'aime pas avoir la responsabilité de prendre des décisions.	FD D N A FA
83.	Je n'aime pas avoir de l'autorité sur les autres.	FD D N A FA
84.	Je crois que la plupart des gens sont prêts à mentir pour leur avancement.	FD D N A FA
85.	J'agis impulsivement lorsque je suis sous l'emprise des émotions.	FD D N A FA
86.	Je préfère la routine plutôt que la variété.	FD D N A FA
87.	J'aime agir avec spontanéité.	FD D N A FA
88.	Je ne suis pas facilement amusé.	FD D N A FA
89.	Je porte moins de respect à mes supérieurs si je suis en désaccord avec eux.	FD D N A FA
90.	Je pense souvent à la possibilité de conséquences négatives.	FD D N A FA
91.	Je suis très content lorsque j'apprends de nouvelles choses.	FD D N A FA
92.	Je suis facilement découragé dans les situations difficiles.	FD D N A FA
93.	J'ai une personnalité forte.	FD D N A FA
94.	Je crois que la plupart des gens sont dignes de confiance.	FD D N A FA
95.	J'ai tendance à réagir impulsivement.	FD D N A FA
96.	Je préfère discuter avec les gens de leurs activités quotidiennes plutôt que de leurs sentiments.	FD D N A FA
97.	J'aime être bien préparé.	FD D N A FA
98.	Je suis une personne optimiste.	FD D N A FA
99.	Je ne me laisserai pas dominer par autrui.	FD D N A FA
1 0 0	Je me retiens souvent de faire des choses par peur d'être embarrassé.	FD D N A FA
1 0 1	Je m'adapte facilement aux nouvelles situations.	FD D N A FA
1 0 2	Je peux passer des heures à ne rien faire.	FD D N A FA
1 0 3	Je préfère rester en arrière-plan.	FD D N A FA

1 0 4	Je tiens compte des besoins ou intérêts d'autrui lorsque je prends une décision.	FD D N A FA
1 0 5	Je résiste facilement aux tentations.	FD D N A FA
1 0 6	Je suis une personne terre à terre.	FD D N A FA
1 0 7	Je crois que planifier à l'avance fait que les choses tournent mieux.	FD D N A FA
1 0 8	Je suis passionné par tout ce dans quoi je suis impliqué.	FD D N A FA
1 0 9	Je préfère la compétition à la coopération.	FD D N A FA
1 1 0	J'ai souvent l'impression que les autres sont contre moi.	FD D N A FA
1 1 1	Je préfère m'en tenir aux choses qui ont fait leurs preuves.	FD D N A FA
1 1 2	Je trouve difficile de me mettre au travail.	FD D N A FA
1 1 3	J'essaie d'éviter de parler en public.	FD D N A FA
1 1 4	J'essaie de faire en sorte que chaque membre d'un groupe se sente inclus.	FD D N A FA
1 1 5	Je cède à mes désirs.	FD D N A FA
1 1 6	Je crois que l'esthétisme n'est pas moins important que l'utilité.	FD D N A FA
1 1 7	Je suis apprécié par mes amis pour ma fiabilité.	FD D N A FA
1 1 8	Je suis habituellement actif et énergique.	FD D N A FA
1 1 9	Je pense que les leaders devraient laisser chaque membre du groupe avoir son mot à dire dans les décisions.	FD D N A FA
1 2 0	Je m'inquiète de ce que les autres pensent de moi.	FD D N A FA
1 2 1	Je réponds aléatoirement aux sondages et questionnaires.	FD D N A FA

ANNEXE B - FORMULAIRE DE CONSENTEMENT- PAPIER

Formulaire de consentement –Format papier

Projet de recherche : *L'évaluation de la personnalité à l'ère d'Internet : étude de l'équivalence entre son évaluation psychométrique par Internet et sous forme traditionnelle papier-crayon.*

Chercheur principal : Yann Le Corff, Ph.D., c.o., professeur, Département d'orientation professionnelle, Université de Sherbrooke.

Collaboratrices : Kim Desroches et Véronique Gingras, candidates à la M.Sc. en orientation.

Madame,
Monsieur,

Nous vous invitons à participer au projet de recherche en titre. Son objectif principal est d'étudier l'impact de l'utilisation d'Internet comme modalité d'évaluation psychologique en comparaison avec la modalité traditionnelle papier-crayon.

En quoi consiste la participation au projet?

Votre participation à ce projet de recherche consiste à répondre à un questionnaire portant sur la personnalité. Il vous sera demandé de répondre à ce questionnaire une fois par Internet et une fois sur papier. La passation par Internet pourra avoir lieu à l'endroit et au moment qui vous convient alors que la passation papier aura lieu en groupe dans un local de classe de l'université. Le seul inconvénient lié à votre participation est le temps consacré à la recherche, soit environ une demi-heure par passation.

Qu'est-ce que l'équipe de recherche fera avec les données recueillies?

Pour éviter votre identification comme personne participante à cette recherche, les données recueillies dans cette enquête seront traitées de manière **entièrement confidentielle**. Les résultats de cette recherche ne permettront pas d'identifier les personnes participantes. Les six derniers chiffres de votre numéro de matricule étudiant sont demandés afin de permettre d'associer votre questionnaire papier à votre questionnaire Internet. Les chercheurs n'auront pas accès aux noms associés aux matricules (dont seulement 6 des 8 chiffres sont demandés), ainsi aucune identification personnelle ne sera possible.

Les résultats de cette recherche seront diffusés sous forme de publications scientifiques et professionnelles. Les données seront utilisées par des étudiantes de maîtrise dans le cadre de leur mémoire. Les données informatisées seront conservées sur un serveur sécurisé de l'Université de Sherbrooke, alors que les questionnaires papiers seront conservés sous clef dans le laboratoire de l'équipe de recherche. Les seules personnes qui y auront accès sont

les membres de l'équipe de recherche. Les données ne seront pas utilisées à d'autres fins que celles décrites dans le présent document et elles seront détruites cinq ans après la fin du projet.

Est-il obligatoire de participer?

La participation à cette étude se fait sur une base volontaire. Vous êtes entièrement **libre de participer ou non**, et de vous retirer en tout temps sans avoir à motiver votre décision ni à subir de préjudice de quelque nature que ce soit.

Y a-t-il des risques, inconvénients ou bénéfiques?

Au-delà du temps consacré à votre participation, le chercheur considère que les inconvénients possibles sont pratiquement inexistantes. Par contre, votre participation nous aidera à mieux connaître les impacts de l'utilisation d'Internet comme modalité d'évaluation de la personnalité. Un montant symbolique de 10\$ par passation (20\$ au total) vous sera offert en guise de compensation pour le temps que vous consacrerez à participer à ce projet.

Qui finance cette recherche?

Ce projet de recherche est financé par le Fonds québécois de recherche sur la société et la culture (FQRSC).

Que faire si j'ai des questions concernant le projet?

Si vous avez des questions concernant ce projet de recherche, n'hésitez pas à communiquer avec le professeur Yann Le Corff, Ph.D., c.o., au 1-800-267-8337, poste xxxxx ou par courriel à .

Date : _____

Yann Le Corff, Ph.D., c.o., professeur adjoint
Chercheur responsable du projet de recherche

J'ai lu et compris le document d'information au sujet du projet *L'évaluation de la personnalité à l'ère d'Internet : étude de l'équivalence entre son évaluation psychométrique par Internet et sous forme traditionnelle papier-crayon*. J'ai compris les conditions, les risques et les bienfaits de ma participation. J'ai obtenu des réponses aux questions que je me posais au sujet de ce projet. **En acceptant de répondre au questionnaire, je consens librement à participer à ce projet de recherche.**

Ce projet a été revu et approuvé par le comité d'éthique de la recherche Éducation et sciences sociales, de l'Université de Sherbrooke. Cette démarche vise à assurer la protection des participantes et participants. Si vous avez des questions sur les aspects éthiques de ce projet (consentement à participer, confidentialité, etc.), n'hésitez pas à communiquer avec M. Eric Yergeau, président de ce comité, par l'intermédiaire de son secrétariat au numéro suivant : 819-821-8000 poste xxxxx, ou par courriel à: ethique.ess@usherbrooke.ca.

ANNEXE C - FORMULAIRE DE CONSENTEMENT - INTERNET

Formulaire de consentement – Format Internet

Projet de recherche : *L'évaluation de la personnalité à l'ère d'Internet : étude de l'équivalence entre son évaluation psychométrique par Internet et sous forme traditionnelle papier-crayon.*

Chercheur principal : Yann Le Corff, Ph.D., c.o., professeur, Département d'orientation professionnelle, Université de Sherbrooke.

Collaboratrices : Kim Desroches et Véronique Gingras, candidates à la M.Sc. en orientation.

Madame,
Monsieur,

Nous vous invitons à participer au projet de recherche en titre. Son objectif principal est d'étudier l'impact de l'utilisation d'Internet comme modalité d'évaluation psychologique en comparaison avec la modalité traditionnelle papier-crayon.

En quoi consiste la participation au projet?

Votre participation à ce projet de recherche consiste à répondre à un questionnaire portant sur la personnalité. Il vous sera demandé de répondre à ce questionnaire une fois par Internet et une fois sur papier. La passation par Internet pourra avoir lieu à l'endroit et au moment qui vous convient alors que la passation papier aura lieu en groupe dans un local de classe de l'université. Le seul inconvénient lié à votre participation est le temps consacré à la recherche, soit environ une demi-heure par passation.

Qu'est-ce que l'équipe de recherche fera avec les données recueillies?

Pour éviter votre identification comme personne participante à cette recherche, les données recueillies dans cette enquête seront traitées de manière **entièrement confidentielle**. Les résultats de cette recherche ne permettront pas d'identifier les personnes participantes. Les six derniers chiffres de votre numéro de matricule étudiant sont demandés afin de permettre d'associer votre questionnaire papier à votre questionnaire Internet. Les chercheurs n'auront pas accès aux noms associés aux matricules (dont seulement 6 des 8 chiffres sont demandés), ainsi aucune identification personnelle ne sera possible.

Les résultats de cette recherche seront diffusés sous forme de publications scientifiques et professionnelles. Les données seront utilisées par des étudiantes de maîtrise dans le cadre de leur mémoire. Les données informatisées seront conservées sur un serveur sécurisé de l'Université de Sherbrooke, alors que les questionnaires papier seront conservés sous clef dans le laboratoire de l'équipe de recherche. Les seules personnes qui y auront accès sont les membres de l'équipe de recherche. Les données ne seront pas utilisées à d'autres fins que celles décrites dans le présent document et elles seront détruites cinq ans après la fin du projet.

Est-il obligatoire de participer?

La participation à cette étude se fait sur une base volontaire. Vous êtes entièrement **libre de participer ou non**, et de vous retirer en tout temps sans avoir à motiver votre décision ni à subir de préjudice de quelque nature que ce soit.

Y a-t-il des risques, inconvénients ou bénéfiques?

Au-delà du temps consacré à votre participation, le chercheur considère que les inconvénients possibles sont pratiquement inexistantes. Par contre, votre participation nous aidera à mieux connaître les impacts de l'utilisation d'Internet comme modalité d'évaluation de la personnalité. Un montant symbolique de 10\$ par passation (20\$ au total) vous sera offert en guise de compensation pour le temps que vous consacrerez à participer à ce projet.

Qui finance cette recherche?

Ce projet de recherche est financé par le Fonds québécois de recherche sur la société et la culture (FQRSC).

Que faire si j'ai des questions concernant le projet?

Si vous avez des questions concernant ce projet de recherche, n'hésitez pas à communiquer avec le professeur Yann Le Corff, Ph.D., c.o., au 1-800-267-8337, poste xxxxx ou par courriel à .

Ce projet a été revu et approuvé par le comité d'éthique de la recherche Éducation et sciences sociales, de l'Université de Sherbrooke. Cette démarche vise à assurer la protection des participantes et participants. Si vous avez des questions sur les aspects éthiques de ce projet (consentement à participer, confidentialité, etc.), n'hésitez pas à communiquer avec M. Eric Yergeau, président de ce comité, par l'intermédiaire de son secrétariat au numéro suivant : 819-821-8000 poste xxxxx, ou par courriel à : ethique.ess@usherbrooke.ca.

Yann Le Corff, Ph.D., c.o., professeur adjoint
Chercheur responsable du projet de recherche

J'ai lu et compris le formulaire de consentement du projet *L'évaluation de la personnalité à l'ère d'Internet : étude de l'équivalence entre son évaluation psychométrique par Internet et sous forme traditionnelle papier-crayon*. J'ai compris les conditions, les risques et les bienfaits de ma participation. J'ai obtenu des réponses aux questions que je me posais au sujet de ce projet. J'accepte librement de participer à ce projet de recherche.

J'accepte de participer [Bouton Web]

Je refuse de participer [Bouton Web]

ANNEXE D - ATTESTATION DE CONFORMITÉ ÉTHIQUE



UNIVERSITÉ DE
SHERBROOKE

Comité d'éthique de la recherche
Éducation et sciences sociales

Attestation de conformité

Le comité d'éthique de la recherche Éducation et sciences sociales de l'Université de Sherbrooke certifie avoir examiné la proposition de recherche suivante :

L'évaluation de la personnalité à l'ère d'Internet : étude de l'équivalence entre son évaluation psychométrique par Internet et sous forme traditionnelle papier-crayon

Yann Le Corff

Professeur, Département d'orientation professionnelle, Faculté d'éducation

Projet financé par le FQRSC – Établissement de nouveaux professeurs-chercheurs

Le comité estime que la recherche proposée est conforme aux principes éthiques énoncés dans la *Politique institutionnelle en matière d'éthique de la recherche avec les êtres humains*.

Membres du comité

Eric Yergeau, président du comité, professeur à la Faculté d'éducation, Département d'orientation professionnelle

Chantale Beaucher, professeure à la Faculté d'éducation, Département de pédagogie

France Beauregard, professeure à la Faculté d'éducation, Département d'enseignement au préscolaire et primaire

Mélanie Lapalme, professeure à la Faculté d'éducation, Département de psychoéducation

Carlo Spallanzani, professeur à la Faculté d'éducation physique et sportive

Christina St-Onge, professeure à la Faculté de médecine et des sciences de la santé, Département de médecine

Serge Striganuk, professeur à la Faculté d'éducation, Département de gestion de l'éducation et de la formation

Vincent Beaucher, membre versé en éthique

France Dupuis, membre représentante du public

Le présent certificat est valide pour la durée de la recherche, à condition que la personne responsable du projet fournisse au comité un rapport de suivi annuel, faute de quoi le certificat peut être révoqué.

Le président du comité,

Eric Yergeau, 29 septembre 2011

