

MANUEL TECHNIQUE & GUIDE UTILISATEUR

VOCATION

Ce manuel a été réalisé par les psychologues du département de Science et Innovation de Central Test.

Le test VOCATION et le logo Central Test sont des marques déposées.

Tout droit de reproduction, modification ou réutilisation doit être accordé par écrit par l'éditeur Central Test.

*Copyright © 2020 par Central Test - Empowering Talent.
Tous droits réservés.*

SOMMAIRE

Introduction	4
1. Description générale et applicabilité	5
1.1. Domaines d'application principale	6
1.2. Structure des questions	6
1.3. Mode de passation, langues et besoins particuliers	7
1.4. Désirabilité sociale	8
1.5. Conditions requises	8
1.6. Questions techniques et de sécurité	9
1.6.1. Respect de l'éthique et de la déontologie	9
1.6.2. Sécurité et confidentialité des données utilisateurs	10
1.6.3. Architecture logicielle et disponibilité du serveur	10
1.6.4. Rapports informatisés	10
1.7. Commercialisation, conditions et disponibilité	11
2. Développement conceptuel et pratique de l'échelle	12
2.1. Fondement théorique	12
2.2. Développement du test	14
2.2.1. Construction des items et des facteurs	14
2.2.2. Prétests	15
2.2.3. Définitions des facteurs	15
2.2.4. Profils combinés	19
2.2.5. Format de réponses « ipsatives »	19
2.2.6. Théorie de réponse à l'item (IRT) thurstonienne et les réponses ipsatives	20
2.2.6.a. Aptitude de l'item (Item Facility)	20
2.2.6.b. Difficulté de l'item et pouvoir discriminant	21

2.2.7. Etude pilote	21
3. Description de l'échantillon	22
4. Normes	24
5. Fidélité	33
5.1. Fidélité interne	33
5.2. Test-retest	34
6. Validité	36
6.1. Validité concourante	36
6.2. Validité de construit	38
6.2.1. Corrélations inter-facteurs	39
6.2.2. Analyse factorielle confirmatoire (AFC)	40
6.2.3. Validité conséquentielle	42
7. VOCATION et égalité des chances	43
7.1. Mesure de l'équivalence	43
7.1.1. Fonctionnement différentiel d'un item (DIF – Differential Item Functioning)	44
7.1.1.a. Analyse DIF : Genre	45
7.1.1.b. Analyse DIF : Âge	46
7.1.1.c. ANOVAs et tests-t	47
8. Informations et considérations supplémentaires	52
8.1. Interpréter les scores	52
8.2. Besoin organisationnel : Choix de carrière, Recrutement, Développement	53
Références	55

Introduction

VOCATION est un outil d'évaluation des intérêts professionnels conçu pour analyser les préférences professionnelles des individus. Basé sur la théorie du choix **VOCATIONnel** de John Holland (1959), **VOCATION** fournit à l'utilisateur des informations détaillées sur les domaines professionnels, les activités et les environnements de travail qui correspondent aux intérêts de l'évalué.

VOCATION est une mise à jour de l'Inventaire des Intérêts Professionnels-Révisé de *Central Test*. Cette nouvelle version a l'avantage de proposer :

- Des questions plus modernes tenant compte des réalités professionnelles actuelles
- Un questionnaire court et aisément compréhensible
- Plus de questions situationnelles
- Une durée moyenne de passation de 8 minutes
- Une échelle ipsative permettant un meilleur contrôle de la désirabilité sociale et de la gestion des impressions
- Un développement qui tient compte des avancées récentes en matière de psychométrie telles que la théorie de réponse à l'item (IRT) Thurstonienne
- Des rapports pratiques et aisément compréhensibles
- Des profils contemporains proposant un code RIASEC à 3 facteurs.

Dans ce manuel, vous trouverez des informations utiles sur le développement du questionnaire, les fondements théoriques, les propriétés psychométriques et les applications pratiques.

Les chapitres 1, 2, 3 et 8 sont spécifiquement adaptés aux praticiens des ressources humaines, tandis que les chapitres 4, 5, 6 et 7 comprennent des informations techniques et psychométriques détaillées.

1. Description générale et applicabilité

VOCATION permet de visualiser les intérêts professionnels des participants à ce test à travers six grands domaines :

Réaliste

Les personnes de type « réaliste » ont tendance à aimer les activités physiques et/ou manuelles. Guidées par le souci du résultat, elles sont attirées par des tâches nécessitant une grande dextérité et une endurance physique et, par conséquent, sont susceptibles de préférer les activités se déroulant en plein air ou celles qui consistent à travailler manuellement avec leurs mains.

Investigateur

Ce domaine correspond aux individus qui apprécient les activités à caractère intellectuel ou l'étude des connaissances diverses. Ils souhaitent s'impliquer dans des tâches permettant, entre autre, de résoudre des problèmes, de mener des enquêtes, d'analyser des informations, ou encore de conduire des recherches afin de mieux comprendre un sujet ou l'environnement autour d'eux.

Artiste

Ceux qui s'intéressent aux activités « artistiques » sont susceptibles de vouloir s'exprimer à travers des moyens créatifs tels que la musique, le dessin, l'art, mais également d'utiliser leur créativité et de matérialiser leurs idées dans différents projets professionnels. Ils ont tendance à s'intéresser à des activités qui leur permettent d'être intuitifs et imaginatifs et d'utiliser pleinement leur liberté de pensée et d'expression.

Social

Les personnes qui s'intéressent aux activités « sociales » aiment être entourées, communiquer et aider les autres par, notamment, la transmission de connaissance, le soin ou l'accompagnement. Elles peuvent avoir une préférence pour les activités qui leur permettent de rencontrer de nouvelles personnes et de déployer leur sens de la communauté ou de faire partie d'un groupe.

Entreprenant

Les personnes qui s'intéressent aux activités du type « entreprenant » ont tendance à jouer des rôles qui impliquent une prise de risque, la possibilité d'influencer autrui, de faire de la vente ou des affaires. Elles apprécient de persuader les autres de leurs idées et/ou de diriger, de gérer des personnes ou des projets. Elles sont également susceptibles d'être dynamiques et de prendre des initiatives.

Conventionnel

Les personnes qui s'intéressent aux activités « conventionnelles » préfèrent mener des tâches de manière ordonnée et méticuleuse. Elles sont susceptibles de vouloir adhérer à une certaine structure dans les activités qu'elles effectuent et à organiser leur environnement. Elles sont susceptibles d'apprécier des tâches nécessitant une précision et une attention particulière aux détails et/ou traitant des données numériques.

1.1. Domaines d'application principale

VOCATION a été conçu pour les professionnels des ressources humaines, les conseillers d'orientation, les psychologues du travail et les spécialistes du coaching du monde entier. Il peut être utilisé pour l'orientation/la mobilité, les bilans de compétences/la formation et le recrutement. Le tableau 1 détaille les différentes utilisations de **VOCATION** :

Table 1. Les différentes utilisations de **VOCATION**.

Orientation et Mobilité	Bilans de compétences et Formation	Recrutement
Pour accompagner les étudiants et les salariés tout au long de leur carrière dans leurs choix et leur développement professionnel.	Pour susciter un dialogue constructif avec le bénéficiaire sur ses intérêts, ses zones de frustration et ses moteurs d'épanouissement.	Pour sécuriser une décision d'embauche et favoriser la fidélisation des futurs employés grâce à des informations révélatrices sur leurs sources de motivation.

VOCATION apporte un riche éclairage sur les domaines d'intérêts d'un individu. Ceci dit, et selon le contexte d'utilisation, il peut être couplé à d'autres outils comme des questionnaires de personnalité ou des tests d'aptitudes pour une vision globale des potentialités d'une personne. **VOCATION** ouvre la voie à une réflexion autour de ce qui motive la personne au travail en termes d'environnements et de tâches préférés mais aussi de style d'apprentissage et de manière de faire.

Table 2. Les différentes utilisations de **VOCATION**.

Tests complémentaires	Objectifs	Avantages
Profil Pro 2	Orientation professionnelle	Des informations approfondies sur les intérêts et la personnalité adaptées à une carrière spécifique.
Profil Vente R	Développement de carrière	Intérêt pour les emplois liés à la vente avec les compétences requises.
CTPI-R	Évolution managériale	Cerner les intérêts pour la gestion de projets ou d'individus et les environnements les plus adaptés.

1.2. Structure des questions

Format

Contrairement aux inventaires d'intérêts traditionnels, les questions dans **VOCATION** sont présentées dans un format de choix forcé où les répondants choisissent la réponse qui représente le mieux leur préférence suivant deux options listées. Les formats de choix forcé réduisent le biais de

réponse (Chan, 2003; dix Holt, van Duijn & Boomsma, 2010), de gestion des impressions (Baron, 1996; Bartram, 2007) et sont généralement plus difficiles à falsifier (Brown & Maydeu-Olivares, 2011). Les formats de choix forcé rendent impossible de fournir la même réponse sur tous les éléments, ce qui entraîne une plus grande différenciation des scores dans un profil (Cruise, 2011/2012). Les recruteurs, les conseillers professionnels et les praticiens des ressources humaines peuvent donc être convaincus que les réponses sont des représentations plus précises des préférences professionnelles individuelles.

Questions

Le questionnaire est composé de 60 questions qui mesurent 6 traits et 12 facettes. Chaque question doit être répondue afin que les répondants puissent passer à l'étape suivante de l'évaluation. Voici un type de question.

Vous préférez les activités pour lesquelles il faut :

- ☐ Suivre attentivement des procédures.
- ☐ Prendre des risques et mettre en œuvre des changements.

1.3. Mode de passation, langues et besoins particuliers

VOCATION est passé en ligne et fonctionne sur l'ensemble des systèmes d'exploitation moderne (Windows, MAC-OS et Linux). Une version PDF du questionnaire est disponible sur demande.

Comme tous les questionnaires modernes, l'évaluation peut être passée à distance et sans supervision. Si besoin, des évaluations de groupes standardisées peuvent être organisées. Les rapports sont générés en ligne. Par conséquent, si le test a été réalisé en version papier, un administrateur doit saisir manuellement les réponses du candidat sur la plate-forme en ligne du questionnaire afin de générer le rapport correspondant.

Le questionnaire est actuellement disponible en français, en anglais et en allemand. Pour en savoir plus sur la disponibilité du test dans une langue en particulier, veuillez contacter le service Consulting de *Central Test* (consulting@centraltest.com).

VOCATION a été conçu pour être utilisé en toute simplicité par des individus aux parcours divers, ainsi que les personnes ayant des besoins particuliers, quel qu'ils soient. Tous les candidats au test peuvent ajuster l'affichage à l'écran en fonction de leurs besoins. Par exemple, les répondants atteints de déficiences visuelles ou ayant besoin d'une taille de police plus large pour plus de confort peuvent en augmenter la taille en gérant l'affichage à l'écran (en appuyant sur les touche « CTRL » et « + »). De même, les répondants peuvent accéder aux consignes du test tout au long de l'évaluation. Il est également possible de naviguer dans les questions ou de changer ses réponses comme on le souhaite en sélectionnant les invites à l'écran. Pour les candidats non-voyants ou atteints d'un handicap visuel sévère, les items peuvent également être lus par un administrateur qualifié. Les personnes atteintes de déficiences auditives peuvent elles aussi passer le questionnaire, puisque l'intégralité des contenus est présentée sous forme visuelles.

Bien que le temps moyen de passation soit de 8 minutes, **VOCATION** n'impose pas de limite de temps. Les personnes ayant des troubles de l'apprentissage peuvent donc prendre tout le temps dont elles ont besoin. Si un candidat a eu besoin d'un temps beaucoup plus important pour

compléter le questionnaire, il est recommandé aux administrateurs de se demander si cette personne pouvait avoir des besoins particuliers. Veuillez contacter l'éditeur du test pour davantage d'informations sur l'administration du questionnaire à des personnes ou à des groupes ayant des besoins particuliers.

1.4. Désirabilité sociale

La désirabilité sociale est la tendance d'une personne à répondre d'une façon qui sera vue comme favorable par les autres. Lorsque cela n'est pas délibéré, on parle d'autosuggestion. En revanche, lorsque cela est délibéré, on parle de gestion de l'impression (Tett & Simonet, 2011). Étant des êtres sociaux par nature, les individus sont enclins à rechercher un certain degré d'acceptation sociale.

Contrairement à ce qui se passe pour les tests d'aptitude, la gestion de l'impression a une incidence sur les résultats d'inventaires d'intérêts tels que le **VOCATION**. Le fait qu'un individu tente de se présenter sous un jour favorable a une incidence négative sur les résultats, que le questionnaire soit effectué en ligne ou sur papier (Dodou & de Winter, 2014).

La désirabilité sociale est particulièrement problématique dans le cadre du recrutement, puisque le candidat est généralement tenté de répondre de façon à impressionner le recruteur, même s'il n'y a pas de « bonnes » ou de « mauvaises » réponses. Les candidats peuvent donner des réponses en fonction de ce qu'ils pensent que les recruteurs attendent d'eux pour un poste donné. La désirabilité sociale représente un risque important pour la validité des évaluations (Mueller-Hanson, Heggstad & Thornton, 2003). Si la désirabilité sociale était équivalente chez tous les candidats, les administrateurs de tests pourraient en tenir compte dans l'établissement des scores et l'interprétation des résultats.

VOCATION offre une véritable garantie contre la désirabilité sociale grâce au format ipsatif du questionnaire et au questionnaire intégré de vérification des réponses. Des travaux de recherches ont démontré que les mesures ipsatives résistaient mieux à la distorsion des réponses que les échelles de Likert plus communément utilisées (Bowen, Martin, & Hunt, 2002; Kluger, Reilly, & Russell, 1991). Non seulement le format ipsatif limite la possibilité pour les candidats d'améliorer leurs scores, mais les items sont combinés de façon à ce qu'il soit difficile pour les candidats de concevoir une stratégie de réponse, puisque les deux choix de réponses aux items sont aussi attirants l'un que l'autre.

1.5. Conditions requises

Dans la plupart des cas, l'évaluation sera effectuée à distance via l'accès en ligne et sans supervision. Toutefois, dans le cas de l'évaluation dans un cadre supervisé, les directives suivantes devraient être appliquées.

Bien que les instructions pour remplir **VOCATION** soient présentées à l'écran, il est important que les administrateurs communiquent avec les candidats. Avant de commencer l'évaluation, les instructions devraient être lues à haute voix par l'administrateur et les candidats devraient avoir la possibilité de poser des questions avant de procéder à l'évaluation. Les administrateurs doivent vérifier que les candidats comprennent ce qu'ils doivent faire avant de commencer l'évaluation.

Les instructions d'essai sont les mêmes pour chaque personne. Les instructions normalisées garantissent que tous les participants sont exposés à des conditions similaires pendant le processus d'évaluation, afin de s'assurer qu'il soit juste et objectif, et que les conditions d'évaluation ne donnent pas, directement ou indirectement, un avantage à un individu ou à un groupe.

Autres conseils :

- Si les candidats ont des questions après l'ouverture de la séance d'évaluation, l'administrateur devrait essayer de répondre sans déranger les autres candidats.
- Si un candidat ne comprend pas une question, l'administrateur doit répondre attentivement et éviter d'expliquer la signification des mots ou des questions dans ses propres mots afin de ne pas influencer le candidat.
- Si un administrateur soupçonne que le manque de compréhension d'un candidat est dû à son niveau d'éducation, l'évaluation devrait être interrompue.
- Si un administrateur soupçonne que le manque de compréhension d'un candidat est dû à une compétence linguistique, l'administrateur devrait demander au candidat de compléter l'évaluation dans la langue avec laquelle il est le plus à l'aise.
- Si un administrateur soupçonne qu'un candidat est prudent lorsqu'il répond, le candidat devrait être encouragé à approuver sa réponse la plus sincère.

Les candidats reçoivent une notification pour passer le questionnaire par courrier électronique avec un lien Web intégré et des détails de connexion. Les liens vers le questionnaire sont protégés par mot de passe et sont exclusifs au compte du candidat. Il est possible de définir les dates d'expiration afin de s'assurer que les candidats remplissent le questionnaire dans les délais impartis. Il est possible de configurer l'invitation de manière à ce que l'évaluation soit complétée à la même adresse IP une fois que le questionnaire a été consulté à partir d'une adresse IP particulière.

VOCATION n'a pas de restriction de temps. Cependant, les candidats prennent généralement environ 8 minutes pour remplir le questionnaire. À des fins de surveillance, les temps d'achèvement sont enregistrés pour permettre aux administrateurs de voir si un candidat s'est précipité dans le questionnaire ou a eu besoin d'un délai inhabituel pour compléter le test. Les candidats sont encouragés à répondre spontanément aux questions, sans se précipiter, afin d'éviter de passer trop de temps à réfléchir à chaque déclaration; il n'y a ni bonnes ni mauvaises réponses.

1.6. Questions techniques et de sécurité

Les points suivants constituent des recommandations pour administrer **VOCATION** de manière sécurisée et s'assurer que le test soit administré de manière objective et équitable. Les règles de sécurité sont également importantes pour protéger les informations du candidat pendant et après le test.

1.6.1. Respect de l'éthique et de la déontologie

En tant qu'éditeur d'outils d'évaluation en ligne, *Central Test* respecte les conventions légales du droit du travail français et les dispositions réglementaires issues de la loi du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés : Déclaration N°1834067.

Les résultats du test et les données personnelles ne peuvent être transmis à des tiers sans le consentement explicite du candidat. Les candidats ont le droit d'accéder à leurs résultats et l'évaluateur doit leur accorder cet accès s'ils le demandent. En outre, si un répondant l'exige, toutes

les données devront être définitivement supprimées. Il devra pour ce faire contacter l'entreprise ou *Central Test* directement.

1.6.2. Sécurité et confidentialité des données utilisateurs

Central Test garantit la sécurité et la confidentialité des données générées durant l'administration des tests en ligne : les codes d'accès sont cryptés dans la base de données et ne sont accessibles par personne. Les mots de passe sont générés par ordinateur et ne sont pas visibles par les entreprises. Un module de réécriture d'URL (générateur de liens sécurisés) est installé et assure un traitement crypté et sécurisé des informations.

La navigation et l'échange de données via l'application *Central Test* peuvent être sécurisés et représentés par une icône en forme de cadenas dans le navigateur Internet. L'application de *Central Test* dispose d'un certificat « GeoTrust SSL TRUE BusinessID », une signature numérique émise et certifiée par GeoTrust®. Ce certificat prouve que nous prenons la sécurité au sérieux et il protège toutes les évaluations de *Central Test* contre l'hameçonnage.

Les invitations au test sont envoyées par des liens sécurisés d'utilisation limitée (en nombre et en durée). La durée du test et les adresses IP sont enregistrées.

1.6.3. Architecture logicielle et disponibilité du serveur

Central Test possède un réseau de 10 serveurs qui assure un service de grande qualité. Chaque serveur est hébergé dans un centre de données spécialisé caractérisé par :

- des lignes dédiées ;
- une bande passante garantie de 100 Mbps (par serveur) ;
- des alimentations redondantes ;
- un temps de rétablissement des équipements de 4 heures ;
- un temps d'intervention de 2 heures en cas d'incident ;
- un suivi et une assistance 24/7/365.

L'accès à l'administration des serveurs est restreint à une seule adresse IP et sécurisé par mot de passe. L'échange de données est sécurisé. L'application de *Central Test* a été testée sur les navigateurs suivants : Internet Explorer 8 et plus, Firefox 21 et plus, Google Chrome 30 et plus et Safari 7 et plus.

1.6.4. Rapports informatisés

Les rapports peuvent être générés quelques minutes après l'achèvement du test par les candidats. Le rapport de **VOCATION** est convivial, personnalisable et disponible en plusieurs langues. Le contenu du rapport peut être adapté en fonction des besoins de l'administrateur. Il est possible de configurer des rapports distincts pour le candidat et pour l'évaluateur. Les rapports sont disponibles au format PDF.

1.7. Commercialisation, conditions et disponibilité

Les modalités de commercialisation de **VOCATION** varient d'un pays à l'autre. Toutefois, les points suivants s'appliquent habituellement :

VOCATION, est uniquement disponible pour les psychologues, les professionnels des RH et les conseillers professionnels.

Une formation doit obligatoirement être suivie avant de pouvoir utiliser et diffuser ce test.

Le présent manuel est accessible aux personnes qui ont été formées à l'utilisation de **VOCATION**.

VOCATION peut être commandé sur Internet ou en contactant le représentant du service client *Central Test* ou un partenaire local dans le monde (customer@centraltest.com).

Pour en savoir plus sur les tarifs de **VOCATION** et les services associés, veuillez contacter le représentant du service client de *Central Test*.

Pour en savoir plus sur les formations, veuillez contacter le service client de *Central Test*.

2. Développement conceptuel et pratique de l'échelle

2.1. Fondement théorique

Modèle de John Holland

La typologie de Holland, également connue sous le nom de Modèle RIASEC, a été développée par l'états-unien John Holland (1959) et repose sur sa théorie selon laquelle l'intérêt ou le choix de carrière professionnel d'une personne est déterminé par des « types de personnalité » sous-jacents. Le modèle est populaire en raison de son application pratique, de sa structure simple et de sa disponibilité dans l'évaluation des intérêts professionnels.

Bien que chaque individu soit complexe, Holland (1977) croit que tout le monde peut être facilement classé en 6 types de personnalité de base ou une combinaison de ces 6 types - Manuels, Penseurs, Créatifs, Aidants, Persuasifs et Organiseurs. Holland (1997) a fait l'hypothèse que le choix de carrière d'un individu est une expression de sa personnalité et que le modèle RIASEC peut également être utilisé pour identifier les environnements de travail appropriés à chacun (Gottfredson, 1999).

Holland (1959) a développé sa théorie tout au long de sa carrière, en commençant dès les années 1960 lors de son passage dans l'armée en tant que psychologue assistant. Nauta (2010) a proposé que bon nombre des idées de Holland pour sa théorie sont le résultat de sa propre formation et des frustrations qu'il a expérimenté lorsqu'il a recouru à des services d'orientation professionnelle. Nauta (2010) déclare également que lorsque Holland servait dans l'armée, il s'est rendu compte que les gens appartenaient à des types de personnalités rangés par catégorie et de nature **VOCATION**nelle. Holland a observé que les recrues de l'armée étaient classées en fonction des résultats de leur entretien, et c'est de là que lui est venue l'idée de trier les individus en types psychologiques définis (Nauta, 2010).

Holland, étant expérimenté dans le domaine de l'orientation professionnelle, a commencé à remarquer les revers des tests de carrière disponibles à ce moment-là. La plupart de ces tests utilisaient des notations assez longue des items et il n'y avait aucune liste définie de professions à laquelle pouvait correspondre les résultats des évaluations. Cela a mené Holland (1985) à concevoir l'inventaire des préférences **VOCATION**nelles (VPI).

Holland a continué à travailler sur sa théorie dans les années qui ont suivies. Il a également publié la « Self-Directed Search » en 1970, qui a depuis fait l'objet de quatre révisions. Holland a également publié le « Occupations Finder » (1977), qui est un système de classification professionnelle. Bien que Holland ait pris sa retraite en 1980, il a continué à travailler sur sa théorie. En 1997, il l'a révisé et lui a ajouté un aspect d'identité environnementale (Staff, 2008).

Sur la base de six types de personnalité, Holland a développé le modèle RIASEC. RIASEC est l'acronyme de Réaliste (les manuels), Investigateur (les penseurs), Artiste (les créatifs), Social (les aidants), Entrepreneur (les persuasifs) et Conventionnel (les organisateurs). Chaque type est caractérisé par une constellation d'intérêts, d'activités préférées, de croyances, d'aptitudes potentielles, de valeurs et de caractéristiques personnelles (Nauta, 2010).

Selon Nauta (2010), Holland a insisté sur le fait que la plupart des gens recherchent un environnement de travail qui leur permet d'utiliser les compétences et les capacités qu'ils possèdent et de prendre des rôles avec lesquels ils sont à l'aise.

La théorie de Holland est largement acceptée et joue un rôle essentiel dans la réalisation d'un changement de paradigme radical dans la mesure des décisions liées à la carrière, de la satisfaction et de la stabilité du travail (Bullock, Andrews et Braud, 2010).

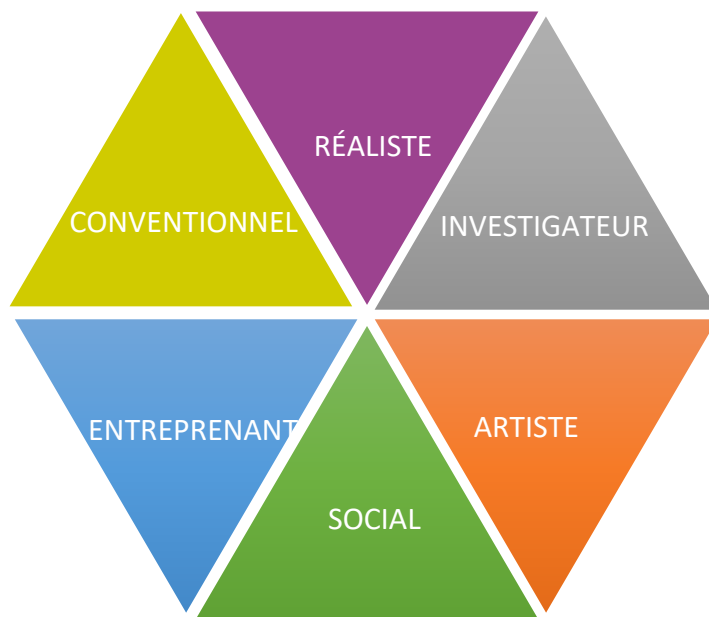


Figure 1. Le cadre structurel du modèle RIASEC (1977) de Holland représenté en hexagone.

Une quantité importante de recherches a été menée sur la relation possible entre les dimensions RIASEC et les traits de personnalité. Larson, Rottinghaus et Borgen (2002) ont mené une méta-analyse sur la relation entre les dimensions RIASEC de la théorie de Holland et les facteurs du Big Five (ouverture, conscienciosité, extraversion, agréabilité et névrosisme) (Costa & McCrae, 1992). Ils ont constaté qu'il existait des corrélations significatives pour les hommes et les femmes entre la dimension Artiste et « ouverture » du Big Five, entre la dimension Entreprenant et « extraversion », entre la dimension Social et « extraversion », entre Investigateur et « ouverture », ainsi qu'entre la dimension Social et l'« agréabilité ». Barrick et al (2003) ont également trouvé une corrélation ainsi qu'un autre chevauchement entre Conventionnel et « conscienciosité » et Social et « agréabilité ». Cependant, la recherche ne semble pas montrer de corrélation entre les dimensions du modèle RIASEC et « névrosisme » (Schinka, Dye, & Curtiss, 1997).

Cette recherche suggère qu'il est possible de déduire les intérêts professionnels des traits de personnalité d'un individu et, par conséquent, qu'il soit possible de faire ressortir des traits de personnalité à partir de ses intérêts professionnels. Cependant, la recherche suggère que les intérêts professionnels sont plus prédictifs du choix de l'emploi que la personnalité (Fruyt & Mervielde, 1999), ce qui peut être un prédicteur précis du succès sur le lieu de travail (Iddekinge, Putka, Roth et Lanivich, 2011).

Subdivision des facteurs RIASEC

La recherche a révélé qu'il existe des différences considérables entre les aspects de chaque dimension du modèle RIASEC néerlandais (1958). Les six facteurs : Réaliste, Investigateur, Artiste, Social, Entreprenant et Conventionnel sont trop larges et compte tenu de la complexité du monde organisationnel actuel, ces facteurs auraient avantage à être simplifiés ou subdivisés afin de représenter une plus grande diversité de différences individuelles. Dingemanse, Van Amstel, De Fruyt

et Wille (2007) ont exploré l'idée qu'il devrait y avoir plus de séparation dans les types RIASEC plutôt que juste entre eux. Ils ont créé le « Career Insight Questionnaire (CIQ) », dans lequel ils ont construit 15 « blocs de construction des types Holland » (Dingemanse, Van Amstel, De Fruyt & Wille, 2015, p237) à partir du modèle RIASEC à six facteurs. Par exemple, le réalisme a été subdivisé en « Artisanat », « Extérieur » et « Mécanique/Construction », et Investigateur a été subdivisé en « Science » et « Théorie ». Puis, en comparant leurs facteurs avec les traits de personnalité du Big Five de Costa et McCrae, Dingemanse et al (2015) ont révélé que les corrélations entre les traits de personnalité du Big Five n'étaient pas cohérentes entre les aspects de chacun des facteurs RIASEC lorsqu'ils étaient subdivisés. Par exemple, une corrélation négative a été trouvée entre l'agréabilité et le facteur Investigation de Holland mais seulement pour la composante scientifique de ce facteur, ce qui suggère qu'un individu qui obtient un haut score d'agréabilité peut ne s'intéresser qu'à une partie de la dimension « I » de Holland.

Sur la base de cette recherche, il a été décidé que le fractionnement de chacun des facteurs RIASEC en deux parties pour **VOCATION** donnerait une représentation plus précise de la gamme des intérêts professionnels existants. Vous trouverez plus de détails sur ces subdivisions au chapitre 2.2.3.

2.2. Développement du test

VOCATION a été développé à partir d'un processus approfondi à plusieurs étapes. Le précurseur du questionnaire (Inventaire des Intérêts Professionnels-R) a été développé pour la première fois en 2006. Les éléments et les facteurs du questionnaire ont été mis à jour afin d'assurer : a) que les éléments reflètent les méthodes de travail contemporaines; b) la prévention efficace du faux et du biais social; c) que les praticiens aient un outil qui améliore la précision des décisions d'évaluation; et d) que les individus soient clairement guidés lorsqu'ils évaluent l'adéquation avec les rôles et les carrières désirés. En plus de ces aspects, tout au long de la révision de cette évaluation, il a été assuré que la structure des facteurs et des éléments satisfaisaient aux normes de fiabilité et de validité les plus élevées.

2.2.1. Construction des items et des facteurs

VOCATION a été développé à partir d'une base d'articles utilisés dans le questionnaire original de l'Inventaire des Intérêts Professionnels-R. Le questionnaire original a été examiné en trois grandes étapes :

1. un retour sur le questionnaire a été demandé et recueilli à partir d'entretiens avec les clients et dans des ateliers de travail et d'échange afin de pouvoir s'assurer que le questionnaire et les facteurs mesurés répondaient à leurs besoins;
2. un groupe d'experts composé de psychométricien(ne)s et de chercheur(e)s a ensuite examiné les éléments à des fins de validité apparente et pour donner leur avis d'expert;
3. le questionnaire a été testé à six reprises avec un échantillon important et représentatif de personnes. Les résultats ont été analysés et des items ont été supprimés et/ou modifiés en fonction de leurs qualités psychométriques, afin d'obtenir la meilleure fiabilité et validité.

Tous les changements ont été apportés afin de s'assurer que les aspects de **VOCATION** mesurent avec précision les domaines et intérêts professionnels contemporains. Les items ont été analysés de manière à tenir compte des éventuels biais culturels, de leur clarté et de leur précision. Deux exemples d'éléments modifiés résultant de ce processus sont décrits dans le tableau 3 ci-dessous.

Table 3. Exemple d'éléments modifiés résultants de l'examen du feedback.

Facteur, Facette	Item original	VOCATION
Artistique, A1	Aider dans l'écriture du script.	Exprimez-vous à travers l'écriture d'un script.
Entreprenant, E2	Manager de proximité, s'assurant d'une gestion fluide de la chaîne de TV.	S'assurer que les équipes fonctionnent efficacement.

2.2.2. Prétests

Des prétests ont été conduits auprès d'une gamme de professionnels hommes et femmes à travers le monde afin de s'assurer que les items étaient exacts, pertinents, clairs, concis, sans biais culturels et qu'ils reflétaient les comportements de travail ciblés. Après six prétests en français, le questionnaire final a été mis à l'essai sur 277 professionnels francophones européens. Les échantillons prétestés comprenaient des francophones de toute l'Europe afin d'établir le questionnaire en tant que mesure européenne des intérêts professionnels.

Lors de la collecte des données des prétests, la fidélité composite (composites reliabilities (CR), définie au point 5.1.) pour chaque dimension a été calculée à partir des saturations générées des items sur le facteur. Les facteurs avec une fiabilité en dessous de .55 ont été signalés. Les facteurs qui n'ont pas convergés ont été inspectés afin de trouver les moins bons items qui ont ensuite été revus, modifiés ou changés puis soumis à un autre cycle de prétest.

2.2.3. Définitions des facteurs

Au vue des constatations selon lesquelles les facteurs RIASEC sont plutôt larges et devraient être séparés, comme cela a été mentionné au chapitre 2.1., **VOCATION** comprend les six facteurs originaux, mais chacun est divisé en deux facettes.

Sur le rapport de l'utilisateur du test, chaque facteur est présenté avec une courte indication des traits de personnalité, des valeurs, des aptitudes potentielles, du style d'apprentissage et de l'environnement qui peuvent être approuvés pour l'individu, en fonction de son intérêt significatif pour le facteur. Les aspects de la personnalité ont été développés à l'aide de la recherche de Holland (1985), ainsi que de recherches complémentaires sur la relation entre les traits de personnalité et le modèle RIASEC, comme cela a été décrit au chapitre 2.1. Les valeurs associées à chaque facteur signifient les valeurs qui peuvent être importantes pour l'individu. Les aptitudes potentielles sont les compétences possibles que l'individu peut détenir. Le style d'apprentissage se rapporte à la méthode d'apprentissage qui peut être la plus avantageuse pour l'individu lors de la saisie et la compréhension de nouvelles informations. L'environnement se rapporte au type d'environnement de travail susceptible de convenir à l'individu, en fonction de ses intérêts.

Les définitions de chaque facteur ainsi que les définitions de chaque facette et les informations supplémentaires liées aux aspects de la personnalité, aux valeurs, aux aptitudes potentielles, au style d'apprentissage et à l'environnement, sont présentées ci-dessous.

Réaliste

Les personnes avec des intérêts « réalistes » ont tendance à apprécier les activités physiques et/ou manuelles. Motivées par des résultats concrets, elles exercent des tâches qui nécessitent une dextérité importante et une bonne endurance physique, et sont par conséquent susceptibles d'apprécier les activités se déroulant en extérieur ou celles qui impliquent des tâches manuelles. Ce critère est divisé en R1 et R2 comme suit :

(R1) Activités physiques et en extérieur : Marque une volonté de s'impliquer dans des activités physiques et des tâches en dehors d'un bureau et/ou d'un espace confiné.

(R2) Intérêts manuels et techniques : Marque une volonté de réaliser des objets concrets et de s'impliquer dans des activités qui comportent des tâches manuelles et/ou l'utilisation d'outils, de machines.

Personnalité : Solide, authentique, naturelle et pratique.

Valeurs : Bon sens, pragmatisme et sens de l'effort.

Aptitudes potentielles : Dextérité, à l'aise avec les tâches techniques, intelligence mécanique et endurance physique.

Style d'apprentissage : Une approche concrète et pratique, où les informations sont acquises en appliquant les exemples.

Environnement : Elles seront plus à l'aise dans des environnements qui leur permettent d'obtenir des résultats concrets, d'utiliser des équipements techniques ou de travailler sur site. Elles ont tendance à ne pas apprécier d'être confiné dans un espace intérieur, à l'image d'un bureau.

Investigateur

Cette dimension correspond à des personnes qui apprécient l'exploration des connaissances théoriques. Elles aiment résoudre les problèmes, mener des enquêtes et rechercher des informations afin de mieux comprendre un sujet ou l'environnement qui les entoure. Ce critère inclut I1 et I2 comme suit :

(I1) Curiosité intellectuelle et apprentissage : Marque une préférence pour des activités qui nécessitent un esprit critique et la volonté d'acquérir des connaissances.

(I2) Science et technologie : Marque un intérêt pour la science et les nouvelles technologies, ainsi que la volonté d'explorer et de mener des recherches sur ces thèmes.

Personnalité : Curieux de tout et volonté permanente d'apprendre et de comprendre leur environnement. Ces personnes sont souvent précises, intellectuelles et ont une mentalité scientifique.

Valeurs : Curiosité intellectuelle, esprit critique et logique.

Aptitudes potentielles : Recherche, raisonnement logique, capacités d'apprentissage et d'analyse.

Style d'apprentissage : Personne particulièrement adaptée à une approche qui utilise la recherche théorique et systématique.

Environnement : Particulièrement à l'aise dans des environnements axés sur la pensée, qui favorisent l'expertise plutôt que la productivité.

Artiste

Les personnes intéressées par les activités « artistiques » chercheront probablement à s'exprimer par des moyens créatifs, tels que la musique, le dessin et l'art, mais également utiliser leur créativité et matérialiser leurs idées dans différents projets professionnels. Elles pourraient avoir besoin d'exprimer leurs idées et leurs sentiments de cette manière et souhaitent disposer de l'indépendance pour y parvenir. Elles ont tendance à être intéressées par des activités qui leur permettent d'être intuitives et imaginatives et n'ont pas nécessairement le désir de faire partie de projets qui sont trop conventionnels, avec des méthodes strictes qui ne laissent que peu de place à l'improvisation. Ce critère est divisé en A1 et A2 comme suit :

(A1) Sens esthétique et expression : Marque un intérêt pour des activités qui exigent une sensibilité artistique et l'expression de ses émotions de différentes façons.

(A2) Créativité et design : Marque une préférence pour la conception, l'imagination et l'innovation.

Personnalité : Intuitive, créative, idéaliste et indépendante. Sensible et réceptive à l'art et à l'esthétique.

Valeurs : Sens de la beauté, de l'originalité, de l'imagination et de la liberté.

Aptitudes potentielles : Originalité, imagination et capacité à penser différemment/à innover.

Style d'apprentissage : Personne particulièrement adaptée à une méthode qui comprend une certaine créativité, intuition, une communication par l'art, les mots, le langage et des concepts visuels/idées conceptuelles.

Environnement : À l'aise dans des environnements qui encouragent l'expression des idées et la créativité, et au sein desquels les méthodes ne sont pas définies ou structurées de manière excessive.

Social

Les personnes avec un intérêt pour le domaine « social » apprécient d'être entourées d'autres personnes, de communiquer avec elles et/ou de les aider. Elles ont tendance à apprécier les nouvelles rencontres et ont un sens développé de la coopération ou de l'appartenance à un groupe. En outre, elles sont susceptibles de vouloir s'exprimer elles-mêmes, mais aussi d'écouter et d'être attentives aux besoins des autres. Ce critère inclut S1 et S2 comme suit :

(S1) Dévouement aux autres : Marque un intérêt à aider les autres, écouter, informer, former, et à être utile à la communauté.

(S2) Relations personnelles : Marque une préférence à être en contact avec d'autres personnes, à expliquer, conseiller et communiquer avec des clients ou des collègues.

Personnalité : Patient, compréhensif et attentif aux autres. Elles ont tendance à être aimables, chaleureuses et accueillantes.

Valeurs : Altruisme, coopération et générosité.

Aptitudes potentielles : Empathie, écoute et communication.

Style d'apprentissage : Apprendre en partageant avec les autres et en travaillant en équipe.

Environnement : À l'aise dans les environnements sociaux qui favorisent le contact personnel, la collaboration et la communication.

Entreprenant

Les personnes intéressées par les activités du domaine « entreprenant » ont tendance à apprécier les tâches qui impliquent un aspect de prise de risque ou d'influence des autres. Elles apprécieront le fait de vendre et de rallier autrui à leurs idées et/ou de diriger ou de mener les autres. Elles seront également susceptibles d'apprécier la communication avec autrui et la prise d'initiatives. Ce critère est divisé en E1 et E2 comme suit :

(E1) Entreprenant : Marque une préférence à développer de nouvelles activités, à matérialiser des projets, à prendre des risques et à gérer des situations difficiles.

(E2) Leadership : Marque une volonté à encourager, motiver et gérer ou mener une équipe/d'autres personnes.

Personnalité : Ouverte, dynamique, affirmée et entrepreneuriale.

Valeurs : Prise de risques, statut et compétition.

Aptitudes potentielles : Prise d'initiatives, gérer et motiver les autres.

Style d'apprentissage : Apprendre en faisant et en prenant part concrètement au type de situation qu'elles étudient.

Environnement : À l'aise dans des environnements concurrentiels, où elles peuvent prendre des risques et aller au-delà de leur zone de confort afin d'évoluer rapidement.

Conventionnel

Les personnes intéressées par les activités « conventionnelles » préfèrent effectuer leur travail de manière méthodique et méticuleuse. Elles sont enclines à vouloir s'en tenir à une certaine structure dans les tâches qu'elles effectuent, à être organisées et à bien gérer leur temps. Elles sont susceptibles de porter une grande attention aux détails et à être attirées par des projets qui englobent des aspects numériques et exigent une approche précise. Ce critère est divisé en C1 et C2 comme suit :

(C1) Méthodique : Marque une préférence pour organiser, planifier les tâches, appliquer les méthodes et règles de façon rigoureuse et contrôler la qualité ou la structure.

(C2) Intérêts pour les données et les chiffres : Marque l'intérêt d'une personne à raisonner avec des chiffres, à analyser des données statistiques et informatiques et à comprendre les lois des mathématiques.

Personnalité : Méticuleuse, autonome, prudente et ordonnée.

Valeurs : Précision, stabilité et efficacité.

Aptitudes potentielles : Organisation, analyse des données, attention portée aux détails et aptitude à travailler avec les chiffres.

Style d'apprentissage : Apprendre en suivant les règles, les instructions et les procédures établies.

Environnement : À l'aise dans des environnements structurés et hiérarchisés, où les règles et les procédures sont clairement définies.

2.2.4. *Profils combinés*

À ce jour, la plupart des outils RIASEC délivrent les profils d'intérêts professionnels des individus en identifiant leurs deux ou trois principaux intérêts sur les 6 facteurs RIASEC et en les combinant pour expliquer quel domaine ou profession est susceptible d'intéresser la personne. Holland lui-même a noté l'importance d'utiliser des combinaisons de facteurs, en indiquant dans son outil de recherche autodidacte (1979) qui utilise une combinaison à trois facteurs, qu'il est essentiel pour l'utilisateur de tenir compte de toutes les variations possibles de la combinaison de trois facteurs. Cela suggère que le profil des intérêts professionnels d'un individu ne doit pas être déduit de son niveau d'intérêt sur l'un des facteurs RIASEC seul.

À l'appui de l'utilisation de la combinaison de plusieurs facteurs RIASEC pour dériver le profil des intérêts professionnels d'un individu, la recherche a révélé qu'il existe un grand chevauchement entre les six facteurs RIASEC et la personnalité. Par exemple, et comme illustré précédemment, Barrick et al (2003) et Larson, Rottinghaus et Borgen (2002) ont trouvé une relation significative entre la personnalité extravertie et les facteurs entreprenants et sociaux du modèle RIASEC néerlandais. Cela suggère que, en raison de la complexité de la personnalité, les individus sont susceptibles d'être intéressés par de multiples facteurs RIASEC.

À la lumière de cela, les rapports **VOCATION** tirent le profil des intérêts professionnels d'un individu en combinant leurs deux intérêts les plus élevés (calculés à un niveau de signification égal) et leur troisième intérêt le plus élevé (calculé à un niveau de signification légèrement inférieur aux deux premiers). Et, à leur tour, présentent une liste des métiers les plus appropriés possibles. L'utilisation d'une combinaison des trois facteurs principaux permet une mesure plus précise des professions susceptibles d'intéresser l'individu. En outre, garder le troisième intérêt le plus élevé à un niveau de signification légèrement inférieur permet une vision plus réaliste des profils professionnels de l'individu et laisse le loisir de se concentrer sur leurs deux intérêts principaux, au cas où l'individu ne s'identifie pas bien à son troisième intérêt.

2.2.5. *Format de réponses « ipsatives »*

Les options de réponse pour les questions dans **VOCATION** sont écrites en format binaire forcé (BF). Les éléments du format BF demandent aux répondants de faire des jugements comparatifs en approuvant les énoncés comme « le plus comme moi » ou « le moins comme moi », et ont été montrés pour surmonter plusieurs des limitations associées aux questions à échelle Likert (Brown & Maydeu-Olivares, 2011a; Chan, 2003; Saville & Willson, 1991; dix Holt, van Duijn et Boomsma, 2010). Les avantages de l'utilisation des éléments BF comprennent :

- une suppression des biais de réponse, comme la tendance à l'accord avec tous les éléments (biais d'acquiescement);
- une gestion réduite des impressions et falsification;
- des comparaisons intra-personnelles plus puissantes.

Malgré l'avantage d'utiliser des items à choix forcés dans les questionnaires, ils ont été sous-utilisés à cause de la nature des données produites par ce type de questionnaire. Ces données sont « ipsatives », ce qui signifie que la somme des scores des individus reste constante. Historiquement, la théorie des tests classiques (TTC) a été jugée insuffisante pour la notation et la validation des données ipsatives (Cornwell & Dunlap, 1994; Loo, 1999; Meade, 2004), car :

- la nature relative des scores signifie que les comparaisons entre les individus peuvent être difficiles;
- la dépendance linéaire des scores déforme la validité de construction des échelles;
- le manque de variance déforme les corrélations avec des critères externes;
- les hypothèses requises pour les estimations de fiabilité traditionnelles sont violées.

2.2.6. Théorie de réponse à l'item (IRT) thurstonienne et les réponses ipsatives

Les techniques d'analyses contemporaines, développées par Brown et Maydeu-Olivares (2011, 2013), ont permis l'utilisation fiable des items BF dans le développement du questionnaire. En utilisant une combinaison de méthodologies de l'IRT et de modèles de comparaison thurstonienne (Thurstone, 1927; 1931), Brown & Maydeu-Olivares ont montré que les problèmes précédents dans l'interprétation des données ipsatives pouvaient être surmontés, par exemple :

- des valeurs « absolues » peuvent être obtenues, ce qui permet des comparaisons significatives entre les individus;
- les distorsions qui entraînent des problèmes de validité peuvent être corrigées.

Malgré leurs avantages potentiels, les modèles thurstoniens n'ont pas été précédemment utilisés dans l'analyse des données ipsatives pour les raisons suivantes (Brown, 2010) :

- il n'y avait aucune méthode d'estimation efficace disponible;
- la technologie informatique n'était pas assez puissante pour faciliter l'analyse;
- les programmes statistiques antérieurs n'ont pas facilité le calcul des scores car la variance de l'erreur pour chaque question était de 0.

Pour **VOCATION**, l'IRT thurstonienne a été utilisée pour réviser, réécrire et valider à la fois les items et les facteurs. La figure 2 montre le modèle de l'IRT thurstonienne utilisé pour analyser les items et les facteurs du questionnaire.



2.2.6.a. Aptitude de l'item (Item Facility)

L'item se réfère à la capacité d'un élément à différencier les répondants et à identifier les effets de sol ou de plafond. Un élément a des effets de plancher ou plafond lorsque la probabilité d'endosser une option de réponse est beaucoup plus faible ou supérieure à la probabilité d'endosser l'option alternative. Étant donné que les items de **VOCATION** étaient binaires, un item était examiné à partir de la proportion de réponses pour chaque option. Plus équivalente est la proportion, plus l'item est équilibré. Les items de **VOCATION** ayant des effets de plancher et de plafond ont été supprimés ou

modifiés. L'aptitude des items modifiés ou nouveaux a ensuite été examinée à l'aide de nouveaux prétests.

2.2.6.b. Difficulté de l'item et pouvoir discriminant

Ici, un « item » se réfère à la(aux) réponse(s) possible(s) pour chaque question sur **VOCATION**. La difficulté de l'item est le premier paramètre d'un modèle IRT. Il est défini comme l'emplacement sur le continuum de la capacité pour lequel la probabilité d'endosser l'item est égale à 50% des participants. Dans une évaluation basée sur l'intérêt, le paramètre de difficulté reflète le niveau d'intérêt nécessaire pour que le participant ait 50% de chances de choisir l'item. Un item est dit « facile » lorsque la probabilité que 50% des personnes qui l'ont adopté ne disposent seulement que d'un faible niveau sur l'intérêt mesuré par l'item. (Par exemple, un item "facile" mesurant Entreprenant pourrait être choisi par 50% des participants ayant seulement un faible niveau de trait pour Entreprenant). Au contraire, un item est dit "difficile", lorsque la probabilité que 50% des personnes qui l'ont approuvé aient un haut niveau de trait (par exemple, un item "difficile" mesurant Réaliste pourrait atteindre la probabilité d'être choisi de 50% des participants seulement s'ils présentent un haut niveau sur le trait Réaliste).

2.2.7. Etude pilote

Le questionnaire final a été conduit sur 277 professionnels francophones européens. Les fidélités composites pour chaque facteur ont été calculées à partir des saturations générées par facteur pour chaque item. Les facteurs ont été signalés si leurs fidélités étaient inférieures à .55. Les facteurs qui n'étaient pas convergents ont été inspectés afin d'identifier les moins bons items qui ont ensuite été revus, modifiés ou changés puis soumis à un autre cycle de prétest. La fidélité composite à l'échelle moyenne pour l'étude finale était de .89.

3. Description de l'échantillon

L'échantillon comprenait 277 professionnels francophones européens (90 hommes, 187 femmes) âgés de 15 à 70 ans ($\mu = 40.08$; $\sigma = 10.87$). Tous les participants représentaient une gamme de catégories professionnelles des professions techniques, administratives, de service, de gestion et non managériales, et une gamme de métiers de l'industrie, y compris les services bancaires et professionnels, le marketing et le secteur des services, la fabrication et le développement de produits. Les données ont été recueillies auprès du Centre d'études de *Central Test*. Pour assurer la qualité des données obtenues, les participants dont les réponses étaient jugées peu fiables ou qui ont complété le questionnaire en moins de 6 minutes ont été exclus des analyses. Les tableaux 4 à 7 fournissent un aperçu des descriptifs d'échantillons.

Table 4. Pays d'origine.

Pays du candidat	Pourcentage (%)	N
France	91.33	253
Belgique	5.42	15
Suisse	1.44	4
Autre pays européen	1.81	5

Table 5. Genre.

Sexe	Pourcentage (%)	N
Homme	32.49	90
Femme	67.51	187

Table 6. Groupe d'âge.

Age (années)	Pourcentage (%)	N
Moins de 25	7.58	21
26 – 33	22.02	61
34 – 41	25.63	71

42 – 49	24.55	68
Plus de 50	20.22	56

Table 7. Niveau d'étude.

Niveau d'étude	Pourcentage (%)	N
En cours d'études	3.25	9
Niveau 1 : Master ou doctorat	36.82	102
Niveau 2 : Licence	20.94	58
Niveau 3 : Bac +2	15.16	42
Niveau 4 : Baccalauréat	16.97	47
Niveau 5 : CAP, BEP	6.86	19

4. Normes

Les normes permettent de comparer le score d'un individu aux scores d'autres individus ayant passé le même test. Les individus qui participent au test au même moment sont représentatifs de la population et sont utilisés comme normes de groupe. Plus l'échantillon est représentatif de la population, plus la distribution servant à classer les individus, sera précise.

Avec les normes, les développeurs de tests et les psychologues sont en mesure d'établir des comparaisons entre les individus et de prendre des décisions plus objectives. Les groupes de normes nous permettent d'attribuer un sens aux scores du test et à quantifier la performance d'un individu, ou le niveau d'un attribut, d'un intérêt ou d'un trait par rapport à ses pairs.

Les tableaux 8 et 13 présentent la moyenne, l'écart-type (ET), les scores minimum et maximum pour chaque facteur de **VOCATION**, organisés en catégories basés sur l'âge et le genre.

Table 8. Normes par facteur pour tout l'échantillon.

Facteur	Moy	ET	Min	Max	Etendu	SE
R	4.05	1.80	0.00	9.50	9.50	0.12
I	5.81	1.73	1.00	9.50	8.50	0.10
A	5.38	2.28	0.50	10.00	9.50	0.13
S	5.75	1.97	0.50	10.00	9.50	0.12
E	4.80	1.80	0.50	10.00	9.50	0.09
C	4.22	2.12	0.00	10.00	10.00	0.13

Table 9. Normes par facette pour tout l'échantillon.

Facette	Moy	ET	Min	Max	Etendu	SE
R1	4.20	2.16	0.00	10.00	10.00	0.13
R2	3.90	2.22	0.00	10.00	10.00	0.13
I1	6.40	1.97	2.00	10.00	8.00	0.12
I2	5.23	2.17	0.00	10.00	10.00	0.13
A1	4.92	2.47	0.00	10.00	10.00	0.15
A2	5.83	2.55	0.00	10.00	10.00	0.15

S1	5.99	2.16	0.00	10.00	10.00	0.13
S2	5.51	2.37	0.00	10.00	10.00	0.14
E1	4.74	2.00	0.00	10.00	10.00	0.12
E2	4.86	2.24	0.00	10.00	10.00	0.13
C1	4.84	2.41	0.00	10.00	10.00	0.14
C2	3.59	2.59	0.00	10.00	10.00	0.16

Table 10. Normes des facteurs par genre.

Facteur	Moy	ET	Min	Max	Etendu	SE
Femme						
R	3.90	1.69	0.00	9.00	9.00	0.12
I	5.69	1.69	1.50	9.50	8.00	0.12
A	5.65	2.23	0.50	9.50	9.00	0.16
S	5.91	2.03	0.50	10.0	9.50	0.15
E	4.69	1.77	0.50	10.0	9.50	0.13
C	4.17	2.08	0.00	10.0	10.0	0.15

Facteur	Moy	ET	Min	Max	Etendu	SE
Homme						
R	4.37	1.98	0.50	9.50	9.00	0.21
I	6.06	1.81	1.00	9.50	8.50	0.19
A	4.81	2.29	0.50	10.0	9.50	0.24
S	5.41	1.81	0.50	9.00	8.50	0.19
E	5.03	1.85	1.00	9.50	8.50	0.19
C	4.32	2.19	0.50	9.50	9.00	0.23

Table 11. Normes des facettes par genre.

Facette	Moy	ET	Min	Max	Etendu	SE
Femme						
R1	4.14	2.08	0.00	10.00	10.00	0.15
R2	3.65	2.06	0.00	10.00	10.00	0.15
I1	6.36	1.92	2.00	10.00	8.00	0.14
I2	5.02	2.05	0.00	10.00	10.00	0.15
A1	5.26	2.41	0.00	10.00	10.00	0.18
A2	6.04	2.51	0.00	10.00	10.00	0.18
S1	6.25	2.17	0.00	10.00	10.00	0.16
S2	5.57	2.38	0.00	10.00	10.00	0.17
E1	4.65	1.94	0.00	10.00	10.00	0.14
E2	4.73	2.26	0.00	10.00	10.00	0.17
C1	4.96	2.38	0.00	10.00	10.00	0.17
C2	3.37	2.56	0.00	10.00	10.00	0.19

Facette	Moy	ET	Min	Max	Etendu	SE
Homme						
R1	4.32	2.33	0.00	10.00	10.00	0.25
R2	4.41	2.47	0.00	10.00	10.00	0.26
I1	6.47	2.07	2.00	10.00	8.00	0.22
I2	5.66	2.37	0.00	10.00	10.00	0.25
A1	4.20	2.46	0.00	10.00	10.00	0.26
A2	5.41	2.59	0.00	10.00	10.00	0.27
S1	5.44	2.05	0.00	10.00	10.00	0.22

S2	5.38	2.37	1.00	10.00	9.00	0.25
E1	4.93	2.12	1.00	10.00	10.00	0.22
E2	5.13	2.19	0.00	10.00	10.00	0.23
C1	4.60	2.47	0.00	10.00	10.00	0.26
C2	4.04	2.61	0.00	10.00	10.00	0.28

Table 12. Normes des facteurs par âge.

Facteur	Moy	Et	Min	Max	Etendu	Es
Moins de 25						
R	3.76	1.60	1.50	6.50	5.00	0.35
I	5.98	1.65	3.50	9.00	5.50	0.36
A	5.43	2.37	1.00	9.00	8.00	0.52
S	5.74	1.93	2.00	8.50	6.50	0.42
E	4.67	2.00	1.50	10.00	8.50	0.44
C	4.43	1.58	2.00	7.50	5.50	0.34

Facteur	Moy	Et	Min	Max	Etendu	Es
26 – 33						
R	3.80	2.02	0.50	8.50	8.00	0.26
I	5.50	1.82	2.00	9.50	7.50	0.23
A	4.76	2.15	0.50	9.00	8.50	0.28
S	5.72	2.08	1.50	9.50	8.00	0.27
E	5.07	1.74	1.00	8.50	7.50	0.22
C	5.14	2.31	0.00	9.50	9.50	0.30

Facteur	Moy	Et	Min	Max	Etendu	Es
34 – 41						
R	4.15	1.67	0.50	7.50	7.00	0.20
I	5.97	1.60	2.50	9.50	7.00	0.19
A	5.77	2.38	1.00	9.50	8.50	0.28
S	5.71	2.01	0.50	10.0	9.50	0.24
E	4.70	1.84	1.00	9.50	8.50	0.22
C	3.69	1.84	0.00	7.50	7.50	0.22
Facteur	Moy	Et	Min	Max	Etendu	Es
42 – 49						
R	4.27	2.01	0.00	9.50	9.50	0.24
I	5.90	1.89	1.00	9.00	8.00	0.23
A	5.34	2.31	0.50	9.00	8.50	0.28
S	5.56	1.72	2.00	9.00	7.00	0.21
E	4.82	1.96	1.00	9.50	8.50	0.24
C	4.10	2.17	0.50	8.00	7.50	0.26
Facteur	Moy	Et	Min	Max	Etendu	Es
Plus de 50						
R	4.03	1.51	0.50	7.50	7.00	0.20
I	5.78	1.66	2.00	9.00	7.00	0.22

A	5.56	2.16	0.50	10.0	9.50	0.29
S	6.04	2.15	0.50	10.0	9.50	0.29
E	4.64	1.53	0.50	7.50	7.00	0.20
C	3.95	2.08	0.50	10.0	9.50	0.28

Table 13. Normes des sous-facteurs par âge.

Facette	Moy	ET	Min	Max	Etendu	SE
Moins de 25						
R1	4.05	2.06	1.00	8.00	7.00	0.45
R2	3.48	2.23	0.00	8.00	8.00	0.49
I1	6.05	1.69	3.00	9.00	6.00	0.37
I2	5.90	2.07	3.00	10.00	7.00	0.45
A1	5.05	2.82	0.00	9.00	9.00	0.62
A2	5.81	2.54	1.00	10.00	9.00	0.55
S1	6.10	2.59	0.00	9.00	9.00	0.56
S2	5.38	1.91	3.00	9.00	6.00	0.42
E1	4.05	2.25	1.00	10.00	9.00	0.49
E2	5.29	2.53	1.00	10.00	9.00	0.55
C1	5.48	1.97	2.00	9.00	7.00	0.43
C2	3.38	1.83	0.00	7.00	7.00	0.40

Facette	Moy	ET	Min	Max	Etendu	SE
26-33						
R1	4.20	2.31	0.00	9.00	9.00	0.30

R2	3.41	2.36	0.00	10.00	10.00	0.30
I1	5.85	2.11	2.00	10.00	8.00	0.27
I2	5.15	2.07	1.00	10.00	9.00	0.27
A1	4.21	2.17	0.00	9.00	9.00	0.28
A2	5.31	2.57	0.00	10.00	10.00	0.33
S1	6.03	2.07	2.00	10.00	8.00	0.27
S2	5.41	2.57	0.00	10.00	10.00	0.33
E1	5.05	1.84	1.00	9.00	8.00	0.24
E2	5.10	2.20	0.00	9.00	9.00	0.28
C1	5.66	2.34	0.00	10.00	10.00	0.30
C2	4.62	3.04	0.00	10.00	10.00	0.39

Facette	Moy	ET	Min	Max	Etendu	SE
----------------	------------	-----------	------------	------------	---------------	-----------

34-41

R1	4.30	2.24	0.00	10.00	10.00	0.27
R2	4.00	1.90	0.00	8.00	8.00	0.23
I1	6.66	1.82	3.00	10.00	7.00	0.22
I2	5.28	2.11	0.00	10.00	10.00	0.25
A1	5.38	2.64	0.00	10.00	10.00	0.31
A2	6.17	2.57	0.00	10.00	10.00	0.31
S1	5.93	2.09	0.00	10.00	10.00	0.25
S2	5.49	2.42	1.00	10.00	9.00	0.29
E1	4.75	1.99	1.00	10.00	9.00	0.24
E2	4.66	2.35	0.00	10.00	10.00	0.28
C1	4.24	2.23	0.00	9.00	9.00	0.26

C2	3.14	2.23	0.00	9.00	9.00	0.26
-----------	------	------	------	------	------	------

Facette	Moy	ET	Min	Max	Etendu	SE
----------------	------------	-----------	------------	------------	---------------	-----------

42-49

R1	4.37	2.25	0.00	10.00	10.00	0.27
R2	4.18	2.55	0.00	10.00	10.00	0.31
I1	6.47	2.12	2.00	10.00	8.00	0.26
I2	5.34	2.26	0.00	9.00	9.00	0.27
A1	4.69	2.51	0.00	9.00	9.00	0.30
A2	5.99	2.59	0.00	10.00	10.00	0.31
S1	5.84	1.97	1.00	9.00	8.00	0.24
S2	5.28	2.20	0.00	10.00	10.00	0.27
E1	4.75	2.16	0.00	10.00	10.00	0.26
E2	4.90	2.29	0.00	10.00	10.00	0.28
C1	4.74	2.54	0.00	10.00	10.00	0.31
C2	3.47	2.62	0.00	9.00	9.00	0.32

Facette	Moy	ET	Min	Max	Etendu	SE
----------------	------------	-----------	------------	------------	---------------	-----------

Plus de 50

R1	3.95	1.86	0.00	8.00	8.00	0.25
R2	4.11	1.99	1.00	9.00	8.00	0.27
I1	6.70	1.82	2.00	10.00	8.00	0.24
I2	4.86	2.28	0.00	9.00	9.00	0.30
A1	5.32	2.23	1.00	10.00	9.00	0.30
A2	5.80	2.44	0.00	10.00	10.00	0.33

S1	6.14	2.43	0.00	10.00	10.00	0.32
S2	5.95	2.48	1.00	10.00	9.00	0.33
E1	4.64	1.88	0.00	9.00	9.00	0.25
E2	4.64	1.98	1.00	9.00	8.00	0.26
C1	4.63	2.48	0.00	10.00	10.00	0.33
C2	3.27	2.48	0.00	10.00	10.00	0.33

5. Fidélité

La fidélité indique dans quelle mesure une évaluation ne comprend pas d'erreur et le degré avec lequel les items d'une échelle mesurent le même construit. Par exemple, tous les items qui évaluent l'Investigation doivent être cohérents et refléter la définition de ce construit. Une première préoccupation, lorsque l'on élabore et utilise des évaluations, est de savoir à quel point les erreurs de la mesure aléatoire influencent la performance du test. Si un test n'est pas fiable, il est impossible de le valider car l'exactitude des résultats (validité) dépend de la cohérence (fidélité) avec laquelle ces mêmes résultats sont produits. La fidélité permet aux utilisateurs du test d'être assurés que les résultats reflètent précisément les attitudes, aptitudes et comportements au travail des répondants. Dans le contexte d'une évaluation au travail, la fidélité garantit aux utilisateurs que les scores obtenus reflètent la manière dont un candidat à un poste est susceptible de se comporter professionnellement, d'interagir avec les autres et d'effectuer des tâches spécifiques. Lorsqu'une évaluation a une très bonne fidélité, elle contribue également à empêcher les biais susceptibles d'être provoqués par des erreurs de jugement durant le processus de recrutement et de sélection.

5.1. Fidélité interne

La fidélité interne ou inter-items indique dans quelle mesure les items d'une échelle sont homogènes dans leur contenu et mesurent tous le même construit. La fidélité composite est une technique robuste utilisée pour évaluer la cohérence interne des items d'une échelle. Elle est mathématiquement définie comme la quantité de variance du score de l'échelle attribuable à tous les facteurs sous-jacents. L'alpha de Cronbach, qui est une méthode populaire de détermination de la fidélité interne d'une mesure, a une faiblesse majeure en ce sens qu'elle suppose que tous les items d'une échelle sont tous aussi importants et les items ont alors tous le même poids. Par conséquent, s'il y a augmentation du nombre d'items, la fidélité augmenterait artificiellement même si les saturations peuvent être faibles et non significatives.

Par conséquent, la fiabilité composite est une alternative recommandée car elle considère à la fois l'erreur de saturation et de mesure (Dunn, Baguley et Brunsden, 2014; Raykov, 1997; Schmitt, 1996). Les valeurs internes de fiabilité varient de 0 à 1, avec des valeurs plus élevées plus souhaitables. En règle générale, une fiabilité de .70 ou plus est recommandée (Nunnally, 1978), avec des valeurs supérieures à .9 (excellent); .7 à .9 (bon); .6 à .7 (acceptable); .5 à .6 (médiocre).

Nous choisissons d'utiliser l'indice « *global lower bound* » (GLB) qui est un indice qui se rapproche de l'alpha de Cronbach.

La fidélité composite de **VOCATION** a été calculée sur l'échantillon européen francophone standardisé (N = 309). Le tableau 14 décrit la fidélité interne de **VOCATION**. Tous les facteurs ont montré une fiabilité GLB supérieure à .7 avec des valeurs bonnes (I = .81) à excellente (A = .90). En plus de cela, **VOCATION** a un indice de fidélité globale égale à .83, ce qui indique que l'échelle a une très bonne cohérence interne et est très fiable.

Table 14. Fidélité composite par facteurs.

Facteur	GLB
R	0.84
I	0.81
A	0.90
S	0.87
E	0.84
C	0.88

5.2. Test-retest

La fidélité test-retest implique d'administrer le test deux fois à un même participant dans les mêmes conditions et de comparer les résultats pour chaque individu pour déterminer la stabilité du test.

Plus le score de la première administration (Temps 1) est proche du score de la seconde administration (Temps 2), plus le test est fidèle. La fidélité test-retest nous permet de déterminer la cohérence du score d'un test au fil du temps. Le coefficient de corrélation de Pearson nous donne une corrélation entre les deux administrations et un coefficient entre 0 et 1. Plus la corrélation est proche de 1, plus la fidélité est bonne.

La formule du coefficient de corrélation de Pearson est :

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

dans laquelle [x] et [y] sont les moyennes de l'échantillon et X et Y sont les valeurs associés à ces deux variables.

La fidélité test-retest de **VOCATION** a été calculée sur un échantillon de professionnels francophones européens (n=95), à deux moments avec un intervalle de 3 à 6 semaines. L'âge des participants s'étend de 17 à 70 ans ($\mu = 41.89$, $\sigma = 10.29$). La fiabilité test-retest a été calculée à partir des scores individuels pour tous les facteurs et en utilisant le coefficient de corrélation de Pearson. La Table 15 présente les coefficients de corrélation entre le temps 1 et le temps 2.

Table 15. Fidélité test-retest.

Facteur	R
R	.77
I	.75
A	.89
S	.79
E	.73
C	.86

Note : Entre le test et le retest, certains items des dimensions S (Social) et E (Entrepreneur) ont été modifiés afin d'améliorer la fidélité de ces dimensions.

Les corrélations test-retest pour **VOCATION** montrent d'excellents résultats, la corrélation la plus faible étant de .73 et la moyenne étant de .80. Par conséquent, on peut conclure que **VOCATION** est une mesure très fiable dans le temps puisque qu'une personne passant ce test à deux moments différents aura des résultats qui seront constamment les mêmes.

6. Validité

La validité est une mesure du lien entre le test et le construit sous-jacent, dans le cas présent les intérêts. En termes de validité psychométrique, cela se réfère au degré avec lequel les faits et la théorie scientifique étayent l'interprétation des scores aux tests. Bien que la fiabilité d'une évaluation se rapporte à la mesure dans laquelle les résultats sont consistants dans le temps, la validité nous indique dans quelle mesure les résultats d'une évaluation sont exacts et nous permettent de faire des prédictions sur un résultat ciblé. Par conséquent, un test peut être fiable mais pas valide (on peut toujours obtenir une mauvaise réponse), mais il ne peut pas être valide sans être fiable, car la confiance dans la précision des résultats dépend de la cohérence avec laquelle nous obtenons les mêmes résultats. La validité établie protège les praticiens contre les décisions défectueuses, coûteuses, contraires à l'éthique et juridiquement risquées.

La validité est dérivée du mot latin “validus”, qui signifie fort. Selon les normes pour les tests éducatifs et psychologiques (Normes : APA, AERA et NCME, 2014), la validation est un processus qui fournit des preuves de la base scientifique de l'interprétation des notes d'évaluation. Le processus de validation dépend donc des preuves soutenant : a) le contenu du test; b) les processus de réponse; c) la structure interne; d) les relations avec d'autres variables; et e) les conséquences de l'évaluation. **VOCATION** a été validé à la suite d'un processus approfondi en plusieurs étapes qui a permis de s'assurer que les items et les facteurs étaient exacts, représentatifs, cohérents, clairs, équitables et non culturellement biaisés.

La validité de façade est une mesure qui permet de savoir si l'évaluation semble mesurer ce qu'elle prétend mesurer, par ceux qui passent cette évaluation (Nevo, 1985). C'est un aspect important à prendre en compte pour obtenir la coopération et l'engagement du candidat. La validité de façade et les attitudes à l'égard des tests influent sur la motivation, la concentration et l'anxiété des candidats (McCarthy & Goffin, 2003). Par conséquent, la question « Est-ce que **VOCATION** semble, pour les candidats potentiels, mesurer les préférences et les aspirations professionnelles ? » a été abordée lors de la création de l'évaluation, afin de s'assurer que des aspects tels que les questions situationnelles n'étaient pas trompeurs ou confus pour l'utilisateur. Les sections suivantes présenteront plus en détail d'autres aspects de la validité qui ont été pris en compte pour **VOCATION**.

6.1. Validité concourante

La validité concourante permet de valider une nouvelle méthode par rapport à une méthode déjà éprouvée. Si deux tests censés mesurer des attributs étroitement liés sont fortement corrélés, alors la nouvelle méthode peut être qualifiée de valide. La validité concourante consiste à réaliser deux tests en même temps, ou peu de temps l'un après l'autre, afin que la corrélation ne soit pas altérée par des modifications dans le temps. Les tables 17 à 19 montrent les résultats détaillés de la validité concourante de **VOCATION** établie à partir de notre échantillon anglophone (N = 302).

Le groupe international d'items de personnalité (*International Personality Item Pool*, IPIP) est une collection de 2413 items mesurant la personnalité internationalement reconnue. Le projet a été lancé en 1996 et est disponible dans plus de 25 langues.

Le questionnaire IPIP pour cette étude a été extrait du site Internet de l'IPIP (Goldberg, 1999). Hors des échelles de l'IPIP, les échelles d'intérêts professionnels de l'Oregon (ORVIS) ont été choisies pour être l'échelle la plus adéquate pour la validité concourante, car elle mesure les mêmes construits que

le RIASEC de Holland (Pozzebon, 2010). Le tableau 16 ci-dessous présente la correspondance entre les facteurs de **VOCATION** et de l'ORVIS.

Table 16. VOCATION et facteurs ORVIS correspondants.

VOCATION	RIASEC
R	Production
I	Analyse
A	Créativité
S	Altruisme
E	Leadership
C	Organisation

Les participants ont été informés de la consigne suivante : « Sans tenir compte du salaire, veuillez évaluer le degré auquel vous souhaitez faire les activités suivantes, en utilisant l'échelle suivante ». Les participants ont été invités à évaluer chaque déclaration sur une échelle de Likert à 5 points de Fortement déteste (1) à Fortement aimer (5).

Table 17. Corrélation entre les facteurs de VOCATION et de l'ORVIS (IPIP).

	Production	Analyse	Créativité	Altruisme	Leadership	Organisation
R	.50***	.04	-.03	-.04	.06	-.11
I	.12	.65***	-.09	-.16	.46***	.31***
A	.20**	-.08	.67***	.17*	.28***	-.14
S	.04	-.23**	.13	.57***	.43***	.00
E	-.04	.04	-.01	.18*	.78***	.42***
C	-.25**	.20**	-.39***	-.29***	.09	.59***

Note : * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 18. Corrélation entre les facettes de **VOCATION** et l'ORVIS (IPIP).

	Production	Analyse	Créativité	Altruisme	Leadership	Organisation
R1	.47***	-.18*	.09	.04	-.28***	-.40***
R2	.51***	.19*	.00	-.25***	-.34***	-.26***
I1	.05	.38***	-.07	-.27***	-.37***	-.12
I2	.24**	.43***	.10	-.18*	-.28***	-.26***
A1	.25***	-.16*	.68***	.18*	.01	-.31***
A2	.16*	.03	.47***	-.09	-.18*	-.28***
S1	.11	-.31***	.18*	.60***	-.21**	-.37***
S2	.11	-.26***	.23**	.50***	.24***	-.19**
E1	.28*	.06	.54***	.08	.38**	.10
E2	.04	-.01	.13	.28***	.76***	.29**
C1	-.19*	-.13	-.30***	-.34***	-.52***	.04
C2	-.11	.29***	-.26***	-.35***	-.23***	.37***

Note : * $p < .05$, ** $p < .01$, *** $p < .001$.

Les corrélations entre les facteurs **VOCATION** et l'échelle ORVIS (IPIP) présentent d'excellents résultats. Les corrélations entre les facteurs correspondants variaient de .50 (entre R et Production) à .78 (entre E et Leadership), affichant ainsi un excellent niveau de validité concurrente, ce qui indique que **VOCATION** est cohérent avec les modèles existants d'intérêts professionnels. Pour un examen approfondi de la relation entre **VOCATION** et l'ORVIS, les corrélations entre les facettes ont été analysées. Les résultats étaient en grande partie compatibles avec les scores des facteurs, mais il y avait quelques différences. Par exemple, E1 est fortement corrélé avec la créativité, cela est probablement dû au fait que E1 mesure l'état d'esprit entreprenant, le fait d'être novateur et de penser loin, ce qui nécessite beaucoup de pensée créative.

6.2. Validité de construit

La validité de construit détermine dans quelle mesure un test évalue ce qu'il prétend évaluer. Elle est essentielle à la validité globale (perçue et réelle) d'une échelle. Lorsqu'un construit n'est pas clairement défini, les inférences directes ne peuvent être faites entre ce construit et les autres comportements (Cronbach et Meehl, 1955). Souvent, il n'existe pas d'unité de mesure acceptée pour

les construits, et même les plus connues, comme l'intelligence, font l'objet de débats. La validité de construit d'une échelle peut par conséquent être menacée si un construit est : a) trop large ou trop étroit; b) désigné à tort comme une dimension ou un élément spécifique du construit; c) faussé par d'autres construits liés qui masquent l'effet du construit principal; ou d) élaboré après des recherches insuffisantes. Par conséquent, pour élaborer une échelle qui mesure les intérêts liés au travail, le construit doit refléter des théories connues qui définissent des intérêts, en identifiant ses composants, et souligner les comportements positifs et négatifs associés à chaque dimension.

Pour les évaluations basées sur les intérêts, la validité du construit est importante car elle : a) ajoute de la légitimité et de l'acceptabilité à la mesure normalisée des intérêts des individus; b) facilite la compréhension des implications organisationnelles des résultats; et c) permet aux organisations et aux individus d'adapter les interventions pour répondre aux besoins cibles. La validité de construit répond à la question : « Est-ce que **VOCATION** définit et classe les préférences et aspirations professionnelles d'une manière qui représente une compréhension théorique et pratique ? » Pour **VOCATION**, la validité de construit a été établie en utilisant deux approches : les corrélations inter-facteurs et l'analyse factorielle confirmatoire, comme détaillé dans les sections ci-dessous.

6.2.1. Corrélation inter-facteurs

Les facteurs de **VOCATION** mesurent les préférences professionnelles, toutefois, ils mesurent chacun des aspects différents de celle-ci. En d'autres termes, si une personne obtient un score très élevé pour un facteur cela ne veut pas nécessairement dire qu'elle obtiendra un score élevé pour tous les autres facteurs. La table 19 présente les corrélations entre les facteurs.

Table 19. Corrélation inter-facteurs.

	R	I	A	S	E	C
R	1					
I	-.06	1				
A	-.06	.11*	1			
S	-.23***	-.48***	-.26***	1		
E	-.37***	-.36***	-.39***	.14**	1	
C	-.20***	-.14*	-.55***	-.18**	.05	1

Note : * $p < .05$, ** $p < .01$, *** $p < .001$.

Les corrélations présentées dans la table 19 décrivent les relations entre les différents aspects mesurés par **VOCATION**. Comme on peut le voir dans la table 19, les deux facteurs ayant la plus forte corrélation sont S et E à .66. De plus, A et C s'opposent avec une corrélation négative forte à -.55. Et, conformément à la théorie de Holland, les facteurs S et E corrélaient positivement ($r = .14$), car en effet, deux facteurs proches dans l'hexagone sont supposés corrélés positivement. Cependant, nous avons observé des différences avec le modèle, par exemple, S et A sont proches l'un de l'autre dans l'hexagone, mais on les considère comme corrélés négativement dans notre modèle. Peut-être cela reflète-t-il le fait que l'artiste aura une sensibilité tournée vers les belles choses du monde et l'accomplissement de soi à travers l'art alors que le social trouvera le beau dans les relations humaines et s'accomplira à travers son interaction avec les autres.

6.2.2. Analyse factorielle confirmatoire (AFC)

Pour étudier les construits théoriques, ou facteurs, pouvant être représentés par un ensemble d'items, ainsi que la qualité des items individuels, une analyse factorielle exploratoire (AFE) et une analyse factorielle confirmatoire (AFC) sont généralement réalisées. L'AFE et l'AFC sont deux formes d'analyse factorielle. Une analyse factorielle analyse les facteurs (construits hypothétiques) d'un test qui sont mis au point pour interpréter des ensembles d'inter-corrélations fortes entre des items dans une matrice de corrélation. L'analyse factorielle vise à simplifier un ensemble d'items en quelques variables communes, celles-ci étant identifiées en extrayant et en regroupant les items fortement corrélés à un même facteur.

L'objectif de l'AFE est d'identifier et de maximiser l'importance de chaque dimension (pourcentage de variance), tandis que l'AFC est une approche fondée sur des hypothèses qui déterminent le degré auquel les indicateurs observés reflètent le construit sous-jacent en spécifiant le nombre de dimensions attendues, puis en classant la valeur des relations attendues (Schreiber, Nora, Stage, Barlow et King, 2006). Ainsi, cette analyse fournit une estimation chiffrée de l'adéquation de la structure factorielle proposée au phénomène dans le monde réel. L'AFC a donc étudié la cohérence des mesures d'un construit avec la compréhension établie et les arguments théoriques de ce construit.

Une Analyse Factorielle Confirmatoire (AFC) avec indicateurs factoriels continus a été réalisée afin de vérifier la convergence des résultats avec le modèle de l'évaluation. L'AFC est une forme d'analyse factorielle utilisée pour déterminer si un construit est cohérent avec sa signification prévue. Deux grandes approches permettent de confirmer la structure factorielle d'une échelle. Une approche confirmatoire stricte consiste à tester un modèle à l'aide de tests d'adéquation des modèles d'équations structurelles (SEM), afin de déterminer si le modèle spécifié par la recherche est cohérent avec les données. L'autre approche permet de tester deux ou trois modèles pour identifier le plus approprié. Dans notre cas, une approche confirmatoire stricte a été adoptée car il n'existe qu'un seul modèle formulé pour **VOCATION**.

L'AFC étant utilisée pour déterminer l'adéquation du modèle aux données, l'évaluation des indices d'adéquation est essentielle. Plusieurs tests statistiques sont utilisés pour déterminer si le modèle est adapté aux données de l'AFC et du SEM. Plus les indices d'adéquation répondent aux exigences statistiques, meilleur est le modèle. Le χ^2 (khi-deux ou khi carré) est considéré comme le premier indice d'adéquation. Toutefois, en raison de problèmes liés à l'utilisation du χ^2 (sensibilité à la taille de l'échantillon et inadéquation lorsque les hypothèses de distribution ne sont pas respectées), il est fortement recommandé de mesurer l'adéquation relative. Le χ^2 est une mesure de l'adéquation absolue qui fournit une indication fondamentale du degré d'adéquation de la théorie proposée aux données, et qui ne se base pas sur une comparaison avec un modèle de référence (Jöreskog et Sörbom, 1993).

Pour la validation de **VOCATION**, l'erreur quadratique moyenne de l'approximation (RMSEA, Steiger et Lind, 1980), l'indice d'ajustement comparatif de Bentler (CFI) et l'indice Tucker-Lewis ont été choisis comme mesures supplémentaires de l'adéquation, car ils étaient les moins susceptibles d'être affectés par la taille de l'échantillon (Fan, Thompson et Wang, 1999). En outre, l'utilisation d'indices différents est nécessaire car ils reflètent différents aspects de l'adéquation au modèle (Hooper, Coughlan et Mullen, 2008). Comme le χ^2 , le RMSEA est une mesure de l'adéquation absolue qui évalue la quantité d'erreur associée à un modèle proposé. Plus le RMSEA est bas, plus le modèle est robuste. Des indices de 0,005-0,075 indiquent une bonne adéquation (Hu et Bentler, 1999). C'est parce que le RMSEA concerne les résidus du modèle et qu'il évalue dans quelle mesure le modèle est inexact. Le CFI, tout comme le TLI, sont des indices de l'adéquation relative ou incrémentielle qui

évaluent l'amélioration globale d'un modèle proposé par rapport à un modèle indépendant où les variables observées ne sont pas corrélées (Hooper et al. 2008, Holtzman et Vezzu, 2011). Si les valeurs supérieures du CFI et du TLI (0,9 ou plus) indiquent une bonne adéquation, cela ne correspond pas forcément à un RMSEA de 0,05 ou moins (Hu et Bentler, 1999).

Deux analyses AFC ont été effectuées sur le modèle en 6 facteurs et 12 facettes. Les Tables 20 et 21 présentent les résultats de la qualité du modèle.

IRT Thurstonienne

« La théorie de la réponse à l'item (IRT) réfère à un ensemble de modèles mathématiques qui décrivent, en termes probabilistes, la relation entre la réponse d'une personne à une question d'enquête et son niveau de « variable latente » mesuré par l'échelle » (Hays & Fayers, 2005, p.55). Dans les tests ipsatifs, un répondant doit faire un jugement comparatif entre deux ou plusieurs items pour donner une réponse. À l'inverse, dans les tests normatifs, le jugement est mesuré en termes absolus pour chaque attribut évalué et ne compare pas les préférences entre les différents attributs. La notation de test classique ne reflète pas ce processus de comparaison psychologique, mais ignore le fait que les répondants indiquent une préférence par un jugement absolu. La loi de Thurstone sur les jugements comparatifs (Thurstone, 1927, 1931) génère des résultats basés sur les jugements comparatifs des individus. Afin d'analyser les données à l'aide de la modélisation Thurstonienne de l'IRT, les données brutes ont été converties en résultats binaires suivant la procédure décrite par Brown et Maydeu-Olivares (2011) pour le modèle de réponse à l'item de comparaison de paires et de classement des données. Pour **VOCATION**, les 60 questions se composaient de blocs de 2 items (options de réponse). Les réponses observées aux comparaisons appariées ont été codées en utilisant la formule ci-dessous, où « *l* » indique la paire {a, b};

$$y_l = \{1 \text{ si l'item } a \text{ est préféré à l'item } b \quad 0 \text{ si l'item } b \text{ est préféré à l'item } a$$

Le modèle Thurstonien de l'IRT a été adapté aux résultats binaires des comparaisons jumelées avec une structure de 6 et 12 facteurs latents. Au total, 60 résultats binaires ont été générés à partir de 60 blocs de deux items. La version Mplus 7.1 (Muthén & Muthén, 2014) a été utilisée pour estimer les paramètres des items à partir des corrélations tétrachoriques en utilisant l'algorithme des moindres carrés non pondérés (ulsmv).

Table 20. Analyse factorielle confirmatoire en utilisant l'IRT Thurstonienne pour le modèle à 6 facteurs.

N	X2	df	X2/df	p	RMSEA	P <=.05	CFI	TLI
277	2043.017	1635	1.250	.000	.030	1.000	.868	.857

Table 21. Analyse factorielle confirmatoire en utilisant l'IRT Thurstonienne pour le modèle à 12 facettes.

N	X ²	df	X ² /df	p	RMSEA	P ≤ .05	CFI	TLI
277	1834.158	1584	1.158	.000	.024	1.000	.919	.909

Les tableaux 20 et 21 montrent les résultats du niveau d'ajustement du modèle aux données en fonction des deux modèles (6 et 12 facteurs). En raison de la taille de l'échantillon plus robustes des estimateurs, le RMSEA et CFI / TLI ont été utilisés pour évaluer le niveau d'ajustement du modèle. À partir de ces indicateurs, on constate que les deux modèles sont bien adaptés et que le modèle à facteur 12 s'harmonise mieux que le 6. Ceci confirme que **VOCATION** est un modèle valide. Le fait que le facteur 12 correspond mieux que les 6 donne plus de crédibilité à l'idée que chaque facteur du modèle RIASEC puisse être divisé en deux facettes.

6.2.3. Validité conséquentielle

La validité conséquentielle a émergé des préoccupations éthiques liées aux ramifications sociétales des évaluations. Elle peut être décrite comme une sorte de preuve de validité qui concerne les conséquences intentionnelles et non intentionnelles, positives et négatives, liées à l'interprétation et à l'utilisation des évaluations (Messick, 1995). Messick (1980, p. 1012) a soutenu que « non seulement les tests devraient être évalués en fonction de leurs propriétés de mesure, mais que les applications d'essai devraient être évaluées en fonction de leurs conséquences sociales potentielles ». La validité consécutive est considérée comme une composante importante de l'évaluation de la validité, plutôt que d'une conceptualisation isolée (Hubley & Zumbo, 2011). Étant donné les coûts financiers liés à l'embauche d'un candidat, ainsi que la législation sur l'emploi régissant la diversité et l'inclusion de groupes démographiques spécifiques, la validité conséquentielle est importante pour **VOCATION**. Les détails de la validité consécutive mesurée sont présentés ci-dessous, au chapitre 7.

7. VOCATION et égalité des chances

Les organisations sont de plus en plus pressées de s'assurer que les outils utilisés pour le recrutement, la sélection, le développement, la promotion et les décisions de mesure du rendement des employés sont équitables, objectifs et précis. Cette pression provient d'un groupe de demandeurs progressivement bien informé, d'une main-d'œuvre plus sophistiquée et d'une législation internationale du travail et des normes de tests actuelles. L'égalité des chances représente donc une considération critique pour toute évaluation du travail moderne et une évaluation n'est considérée comme véritablement valide que si elle est appropriée pour les différents groupes. **VOCATION** est une mesure juste et précise des préférences et aspirations professionnelles, et les praticiens des ressources humaines et les psychologues peuvent avoir confiance dans son application variée.

Dans le cadre de l'évaluation, les mesures utilisées ne doivent pas avoir d'impact négatif lorsqu'elles s'appliquent à différents groupes. L'impact indésirable se produit lorsqu'il existe une différence significative dans le taux de sélection et les décisions d'emploi pour un groupe protégé par rapport à un autre. Ceci est particulièrement important pour les organisations, car il existe un risque de choisir le mauvais candidat pour un emploi, ainsi que les coûts juridiques et financiers associés, si la justice est saisie pour discrimination à l'emploi. L'impact indésirable est classé comme discriminatoire lorsqu'un candidat croit (et que les preuves de validité ne montrent pas autrement), que le fait de ne pas avoir eu un poste est lié à une appartenance à un groupe démographiquement déterminé. Pour la plupart des pays d'Europe et d'Amérique du Nord, ainsi que l'Australie, l'équité dans le processus de sélection est assujettie à la législation. Plus précisément, pour les États-Unis, la Loi de 1971 sur l'évaluation par test des employés et autres procédures (Département du travail des États-Unis) protège les candidats contre les procédures d'évaluation déloyales, alors que l'évaluation équitable du travail au Royaume-Uni est protégée par ce qui suit :

- Loi des relations inter-race (1976)
- Loi sur la discrimination du genre (1975;1986)
- Loi sur l'invalidité (1998)
- La régulation égalitaire de l'emploi (2003)
 - Orientation sexuelle
 - Religion et croyance
- Discrimination à l'égard de l'âge (2006)

En France, les articles L1132-1, L1132-2, L1132-3 et L1142-1 du code du travail et 225-1, 225-2, 225-3, 225-4, 432-7 du code pénal régissent la législation en matière de discrimination au travail tant en la définissant qu'en exposant les sanctions encourues par les organisations en cas de non-respect de ces articles.

7.1. Mesure de l'équivalence

L'équivalence de mesure (ME) se réfère à la cohérence de la mesure pour certains groupes spécifiés en raison de la nationalité, l'appartenance ethnique et du genre de ses membres. L'objectif de la ME est de s'assurer que les individus de différents groupes avec le même statut sur un attribut, ont le même ensemble d'observations sur une évaluation (Drasgow, 1987; McLeary & Cruise, 2012). La compréhension de ME est importante dans les évaluations afin d'établir la validité, la confiance et l'équivalence entre les groupes. L'équivalence de la mesure est obtenue lorsque des candidats de

différents milieux culturels, groupes d'âge, groupes ethniques, genres et groupes professionnels ayant le même niveau sur un trait ont une chance égale d'être choisis pour un rôle qui requiert le niveau désiré sur ce trait. La ME aide à déterminer dans quelle mesure les répondants interprètent les items de manière similaire et que les différences identifiées entre les groupes reflètent les différences réelles dans le trait mesuré (Drasgow, 1984). Une approche pour évaluer l'équivalence de mesure est l'analyse du fonctionnement différentiel d'item.

7.1.1. Fonctionnement différentiel d'un item (DIF – Differential Item Functioning)

Un item fonctionne différemment (DIF) si les répondants de différents groupes ont des probabilités différentes de répondre à l'item, malgré le même niveau d'habileté au trait. Lorsqu'une probabilité différente de répondre à l'item reste constante à différents niveaux du trait, le DIF est dit uniforme (Figure 3). Ici, l'item d'intérêt donne constamment à un groupe un avantage à tous les niveaux du trait, même lorsque les groupes ont un certain niveau pour la capacité. Inversement, lorsqu'une probabilité différente de répondre à l'item varie selon le niveau du trait, le DIF est caractérisé comme non uniforme (Clauser & Mazor, 1998 ; Hanson, 1998). Plutôt qu'un avantage constant donné au groupe de référence à travers le continuum de capacité pour le trait, il s'agit ici d'un avantage pour ce groupe qui va varier en fonction des différents niveaux du trait. Par exemple, un item peut donner au groupe de référence un avantage à l'extrémité supérieure du continuum et un désavantage à l'extrémité inférieure de même trait (Figure 4)

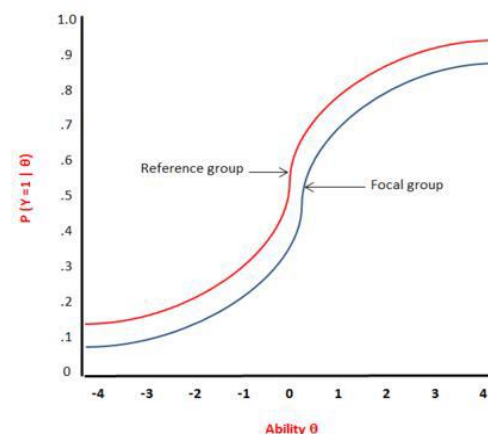


Figure 3. Courbe caractéristique de l'item avec une DIF uniforme.

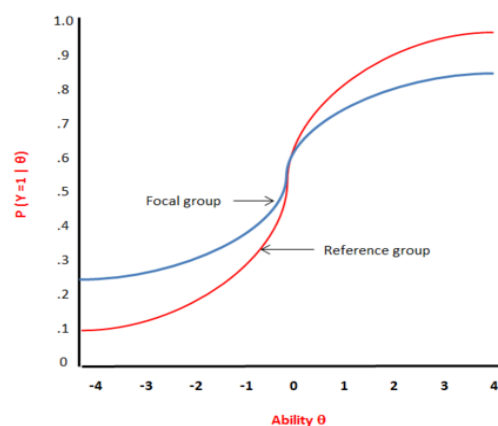


Figure 4. Courbe caractéristique de l'item avec une DIF non-uniforme.

L'analyse DIF fournit des informations sur le fait qu'un item est perçu de manière quel que soit le genre, l'âge et la culture. S'il n'y a pas de DIF, cela signifie qu'il n'y a pas de biais qui favoriserait un groupe par rapport à l'autre. Par exemple, deux individus de différents pays ayant un intérêt absolu similaire dans le domaine Artiste auront la même probabilité d'endosser les items mesurant cette dimension. Cependant, de vraies différences entre les groupes peuvent exister (par exemple, certaines cultures mettront en avant davantage le côté artistique des individus, d'autres moins). De telles différences réelles peuvent s'observer dans les normes (voir chapitre 4).

7.1.1.a. Analyse DIF : Genre

Ces analyses ont été effectuées pour s'assurer que tous les items étaient interprétés de la même façon par les participants de l'un ou l'autre sexe et qu'il n'y avait pas de biais à l'égard des hommes ou des femmes, pour pouvoir conclure que **VOCATION** était un test juste pour les deux genres. Pour cette analyse, nous avons utilisé notre échantillon de francophones européens se composant de 187 femmes (67,51% ; M = 38.42, ET = 10.76) et 90 hommes (32.49% ; M = 41.44, ET = 11.86).

Table 22. Coefficients pour les Items avec une DIF significative et leur taille d'effet.

Question	R ²	Taille d'effet
R1_3	<0.13	Négligeable
R1_9	<0.13	Négligeable
R2_6	<0.13	Négligeable
I2_6	<0.13	Négligeable
S1_6	<0.13	Négligeable
S2_2	<0.13	Négligeable
S2_4	<0.13	Négligeable
S2_8	<0.13	Négligeable
E1_7	<0.13	Négligeable
E2_1	<0.13	Négligeable
E2_5	<0.13	Négligeable
E2_6	<0.13	Négligeable
E2_10	<0.13	Négligeable
C1_2	<0.13	Négligeable

L'analyse a identifié 14 items sur 120 avec un effet DIF pour le genre. Le tableau 22 montre le coefficient DIF ainsi que la taille de l'effet R^2 . Selon le classement de la taille de l'effet de Zumbo et Thomas (1997), tout R^2 sous la valeur .13 est considéré comme négligeable. Sur les 14 items identifiés comme ayant un DIF, aucun n'a affiché une taille d'effet de .13 ou plus, permettant ainsi de considérer que **VOCATION** ne contient aucun item qui fonctionne différemment selon les genres. Deux items (Figure 5 et Figure 6) ont été choisis pour illustrer un DIF uniforme et non uniforme. La figure 6 montre le DIF uniforme pour l'item 6 de la facette I2 (Curiosité intellectuelle et apprentissage).

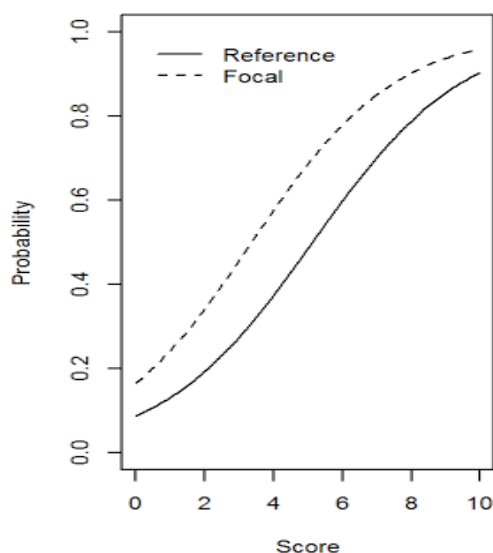


Figure 6. DIF uniforme pour la facette I2.

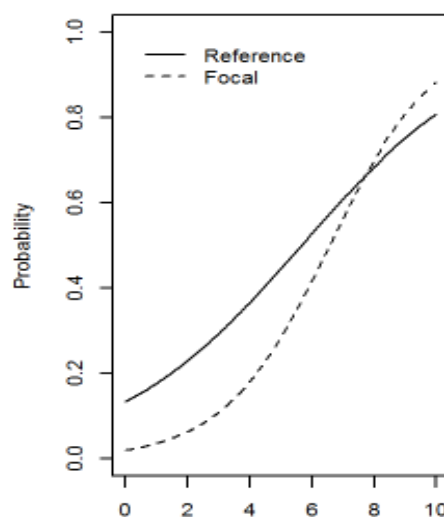


Figure 7. DIF non-uniforme pour la facette A2.

7.1.1.b. Analyse DIF : Âge

Des analyses DIF ont été menées en se basant sur l'âge, dans le but de s'assurer que **VOCATION** était équitable quel que soit l'âge des individus. Des analyses ont été réalisées avec un total de 207 participants, tous issus de l'échantillon normé d'européens francophones. Les participants ont été séparés en deux groupes selon leur âge. Nous avons tenté de diviser en deux selon la médiane (Med = 40), ainsi 105 participants étaient âgés de moins de 40 ans (ET = 6.57) et 102 participants étaient âgés de plus de 40 ans (ET = 6.45).

Table 23. Coefficients pour les items avec une DIF significative et leur taille d'effet.

Question	R^2	Taille d'effet
R1_1	<0.13	Négligeable
R1_8	<0.13	Négligeable

L'analyse a identifié 2 items sur 120 avec un effet DIF pour l'âge. Cependant nous pouvons le voir avec le tableau 23, les tailles d'effet pour chacun de ces items sont négligeables. Nous pouvons alors conclure que ces items sont non discriminatoires en regard de l'âge des individus.

7.1.1.c. ANOVAs et Tests-t

Les analyses du DIF montrent que l'échelle de **VOCATION** a une validité conséquentielle élevée et, bien que quelques items présentent un DIF, ils ont eu un impact négligeable sur l'évaluation globale, ce qui signifie que le questionnaire n'a pas de biais en ce qui concerne le genre et l'âge. Par conséquent, les différences significatives entre les genres ou les groupes d'âge sont dues à des différences réelles et non à des biais du questionnaire. Pour explorer davantage les différences entre les groupes, les tests t et ANOVA ont été effectués entre hommes et femmes et les groupes d'âges (voir Table 24-29).

Groupes d'âge par facteur

Les participants ont été divisés en cinq groupes en fonction de leur âge : Moins de 25 (n = 21), 26-33 (n = 61), 34-41 (n = 71), 42-49 (n = 68), plus de 50 (n = 56).

Table 24. ANOVA pour les facteurs par groupe d'âge.

Facteur	Moins de 25	26 – 33	34 – 41	42 – 49	Plus de 50	F	df	p
R	3.76	3.80	4.15	4.27	4.03	0.731	4	.271
I	5.98	5.50	5.97	5.90	5.78	0.742	4	.564
A	5.43	4.76	5.77	5.34	5.56	1.765	4	.136
S	5.74	5.72	5.71	5.56	6.04	0.478	4	.752
E	4.67	5.07	4.70	4.82	4.64	0.540	4	.707
C	4.43	5.14	3.69	4.10	3.95	4.549	4	.001**

Note : * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 25. Post-hoc (test-t) pour les facteurs par groupes d'âge.

Age	Moins de 25	26 – 33	34 – 41	42 – 49
C				
26 – 33	1.000			
34 – 41	1.000	.001**		

42 – 49	1.000	.047*	1.000	
Plus de 50	1.000	.020*	1.000	1.00

Note : * $p < .05$, ** $p < .01$, *** $p < .001$.

L'ANOVA nous permet d'identifier des différences significatives pour le facteur C. Une analyse post-hoc a été effectuée en utilisant des tests t *pairwise* (correction de Bonferroni) afin de voir où ces différences se sont produites. Ce test permet de comparer les différentes tranches d'âge deux à deux (par paire).

Après cette analyse post-hoc, nous avons constaté pour le facteur C que le groupe 26-33 ($M = 5.14$) a un score significativement plus grand que le groupe 34-41 ($M = 3.69$, $p = .001$), le groupe 42-49 ($M = 4.10$, $p = .047$) et plus de 50 ($M = 3.95$, $p = .020$). Une des explications qui pourrait être donnée serait celle de l'état du marché de l'emploi actuel. En effet, leur entrée dans le marché du travail, les 26–33 ans, fraîchement diplômés d'études supérieures longues, ont un intérêt pour tout ce qui est conventionnel, notamment, car dans le monde d'aujourd'hui, les jeunes actifs seraient plus à la recherche de stabilité et de sécurité (e.g. un contrat à durée indéterminée à la clé). Ils vont avoir ainsi un intérêt plus marqué pour les emplois conventionnels, dont le nombre d'emplois ne faiblit pas et est assuré sur le long terme.

Groupe d'âge par facettes

Table 26. ANOVA pour les facettes par groupe d'âge.

Facette	Moins de 25	26 – 33	34 – 41	42 – 49	Plus de 50	F	df	p
R1	4.05	4.20	4.30	4.37	3.95	0.352	4	.843
R2	3.48	3.41	4.00	4.18	4.11	1.357	4	.249
I1	6.05	5.85	6.66	6.47	6.70	2.030	4	.090
I2	5.90	5.15	5.28	5.34	4.86	0.994	4	.411
A1	5.05	4.21	5.38	4.69	5.32	2.443	4	.047*
A2	5.81	5.31	6.17	5.99	5.80	1.009	4	.403
S1	6.10	6.03	5.93	5.84	6.14	0.184	4	.947
S2	5.38	5.41	5.49	5.28	5.95	0.674	4	.611
E1	4.05	5.05	4.75	4.75	4.64	1.026	4	.394
E2	5.29	5.10	4.66	4.90	4.64	0.634	4	.639
C1	5.48	5.66	4.24	4.74	4.63	3.481	4	.009**

C2	3.38	4.62	3.14	3.47	3.27	3.351	4	.011*
-----------	------	------	------	------	------	-------	---	-------

Note : * $p < .05$, ** $p < .01$, *** $p < .001$.

Pour un examen approfondi des différences entre les groupes d'âge pour les facteurs RIASEC, une ANOVA a été menée sur les facettes. Les résultats sont présentés dans le tableau 26 ci-dessus. Les différences trouvées dans C se sont révélées exister pour C1 et pour C2, alors que pour A, qui avait précédemment une ANOVA non significative, nous avons maintenant des différences qui existent entre les différents groupes d'âges pour A1. Ces résultats indiquent en outre que les facteurs RIASEC devraient être divisés en facettes pour être correctement interprétés.

Le test t post-hoc pour A1 ne montre aucune différence entre les groupes d'âge comparés deux à deux.

Le test-t post hoc pour C1 montre une fois encore une différence entre les 26-33 ($M = 5.66$) mais uniquement avec les 34-41 ($M = 4.24$, $p = .007$), qui serait donc plus tournés vers des activités nécessitant d'être méticuleux et organisés que leurs aînés. Le test t post-hoc pour C2 indique lui toujours une opposition entre les 26-33 ($M = 4.62$) et les 34-41 ($M = 3.47$, $p = .001$), mais aussi les plus de 50 ans ($M = 3.47$, $p = .044$). Avec l'avènement du big data et des nouvelles technologies ces dernières années, il n'est pas étonnant de voir que les plus jeunes générations ont un goût plus prononcé pour les chiffres qui gouvernent les technologies du monde d'aujourd'hui.

Table 27. Post-hoc (T-test) pour les facettes par groupes d'âges.

Age	18 – 25	26 – 33	34 – 41	42 – 49
A1				
26 – 33	1.000			
34 – 41	1.000	.067		
42 – 49	1.000	1.000	.979	
50 et +	1.000	.150	1.000	1.000
C1				
26 – 33	1.000			
34 – 41	.364	.007**		
42 – 49	1.000	.283	1.000	
50 et +	1.000	.194	1.000	1.000
C2				

26 – 33	.552			
34 – 41	1.000	.001***		
42 – 49	1.000	.109	1.000	
50 et +	1.000	.044*	1.000	1.000

Note : * $p < .05$, ** $p < .01$, *** $p < .001$.

Genre

Afin de vérifier s'il existe des différences entre les scores des hommes et femmes, un test t pour échantillon indépendant a été mené afin de comparer tous les facteurs **VOCATION** entre les femmes (n = 144) et les hommes (n = 63).

Table 28. Tests-t pour les facteurs par genre.

Facteur	Femme	Homme	T	df	p
R	3.90	4.37	-2.05	275	0.04*
I	5.69	6.06	-1.66	275	0.10
A	5.65	4.81	2.92	275	0.00***
S	5.91	5.41	1.97	275	0.05
E	4.69	5.03	-1.51	275	0.13
C	4.17	4.32	-0.57	275	0.57

Note : * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 29. Tests-t pour les facettes par genre.

Facette	Femme	Homme	T	df	p
R1	4.12	4.38	-0.93	268	0.35
R2	3.64	4.45	-2.83	268	0.00***
I1	6.40	6.43	-0.13	268	0.90

I2	5.07	5.59	-1.86	268	0.06
A1	5.24	4.23	3.19	268	0.00***
A2	6.04	5.44	1.82	268	0.07
S1	6.23	5.49	2.64	268	0.01**
S2	5.53	5.41	0.38	268	0.70
E1	4.64	4.87	-0.90	268	0.37
E2	4.72	5.08	-1.24	268	0.22
C1	4.99	4.62	1.20	268	0.23
C2	3.38	4.00	-1.86	268	0.06

*Note : *p < .05, ** p < .01, ***p < .001.*

Les tests t ont révélé des différences significatives entre les hommes et les femmes sur les facteurs de R et A et tendancielle pour S. Les hommes ont été significativement plus intéressés par R (homme = 4.37 ; femme = 3.90) ($t(275) = -2.05, p = .04$). Nous retrouvons le même résultat pour R2 (homme = 4.45, femme = 3.64 ; $t(268) = -2.83, p = 0.00$), mais pas pour R1, $t(268) = -.93, p = .35$. Cela pourrait être dû à l'idée que les hommes sont généralement censés être ceux qui prennent des emplois de main-d'œuvre manuelle, comme la mécanique ou la construction, etc.

Pour A, nous observons cette fois que les femmes ont un score plus élevé que les hommes (homme = 4.81 ; femme = 5.65) ($t(275) = 2.92, p = .00$). De même que pour E, lorsque l'on s'intéresse aux sous-facteurs, nous retrouvons une différence significative pour A1 (homme = 4.23, femme = 5.24 ; $t(268) = 3.19, p = 0.00$), mais pas pour A2 ($t(268) = 1.82, p = .07$), bien que la valeur du p soit tendancielle. La différence A1 peut être due à une tendance pour les femmes à être plus en contact avec leurs émotions et sont donc peut-être plus disposées à exprimer leurs émotions à travers des activités artistiques.

Si les femmes sont plus S de manière tendancielle que les hommes, (homme = 5.41 ; femme = 5.91) ($t(275) = 1.97, p = .05$), cette différence existe clairement et significativement pour S1 (homme = 5.49, femme = 6.23 ; $t(268) = 2.64, p = 0.01$). Cette différence peut exister en raison du fait que la société attendrait des femmes de détenir des professions orientées vers le social (infirmières, enseignantes, service à la clientèle, etc.), qui nécessitent plus de dévouement envers les autres.

8. Informations et considérations supplémentaires

VOCATION a été développé pour avoir plusieurs usages et utilités dans le monde organisationnel. Il est important que des normes, des scores et des interprétations appropriés soient utilisés afin de guider les décisions en matière de ressources humaines pour le recrutement et la sélection, l'élaboration de carrière, le coaching et pour un management performant. Une interprétation précise des tests psychométriques aide les administrateurs et les candidats à : a) mettre l'évaluation dans un contexte; b) comparer efficacement les candidats pour des capacités bien ciblées ; b) identifier avec précision les zones de développement ; d) maximiser l'adéquation entre les exigences professionnelles et les aptitudes et intérêts des candidats ; et e) promouvoir l'organisation en tant que lieu de travail juste et objectif. Le chapitre actuel donne des directives pour que les praticiens des ressources humaines comprennent et interprètent avec succès les scores des candidats à **VOCATION** pour différents objectifs organisationnels.

Le logiciel de génération de rapports pour **VOCATION** fournit des graphiques et des commentaires personnalisés pour les 6 facteurs et 12 facettes qui se rapportent au score du candidat. Le rapport ne doit être interprété que par des professionnels qui ont terminé la formation accréditée de *Central Test* et connaissent ce manuel. La formation et la connaissance du processus d'évaluation sont essentielles car les scores devraient être interprétés à la lumière des informations concernant la fidélité, la validité, l'erreur de mesure standard et la connaissance de tous les aménagements qui ont été mis en place lors des évaluations et des séances d'évaluation qui ont été réalisées. Il est donc recommandé de présenter les scores dans un contexte d'évaluation et d'une manière claire, significative et appropriée pour le public cible.

8.1. Interpréter les scores

Il est important de noter que toutes les évaluations conduites par *Central Test* ont été développées uniquement à des fins de ressources humaines. Par conséquent, aucune conclusion clinique ne peut être tirée des résultats - pas même des scores extrêmes. Il existe six niveaux différents de commentaires au total. Les niveaux sont déterminés par le score d'un candidat. Le tableau 30 fournit un exemple d'indicateurs utilisé pour décrire différents niveaux d'un facteur. Les scores extrêmes sont définis comme les valeurs les plus faibles et les plus élevées sur un facteur. Pour **VOCATION**, le score minimum est fixé à 0 et le score maximum est fixé à 10. Compte tenu du format ipsatif de l'évaluation, il est impossible d'obtenir des scores extrêmes sur tous les facteurs.

Chaque facteur se compose de deux facettes. Des scores élevés sur un facteur indiquent qu'il existe une forte probabilité que la personne évaluée puisse être caractérisée comme ayant un intérêt pour ce facteur. Les scores intermédiaires (i.e. 4-7) indiquent une forte congruence avec la norme du groupe et un peu moins une préférence individuelle pour ce facteur. Cependant, tous les commentaires doivent être considérés comme des hypothèses qui doivent être confirmées avec d'autres sources de données, telles que des entrevues ou d'autres évaluations, ou par le contenu du CV et les références des candidats.

Table 30. Exemple décrivant les différents niveaux pour un trait dans le rapport de **VOCATION**.

Niveau 1 (score 0 – 2)	Niveau 2 (score 2.1 – 3.5)	Niveau 3 (score 3.6 – 5.0)	Niveau 4 (score 5.1 – 6.4)	Niveau 5 (score 6.5 – 7.8)	Niveau 6 (score 7.9 – 10)
Ne semble pas indiquer d'intérêt	Indique un intérêt faible	Indique un faible intérêt, mais possibilité d'apprécier	Indique un intérêt, mais pas une préférence	Indique un intérêt	Indique un fort intérêt

8.2. Besoin organisationnel : Choix de carrière, Recrutement, Développement

Grâce à l'identification des intérêts du travail du XX^e siècle, **VOCATION** est un outil efficace pour, d'une part, l'employeur, dans la sélection des candidats à l'emploi sur la base des domaines professionnels qui correspondent aux intérêts professionnels d'un candidat, et, de l'autre, le candidat à l'emploi, dans la sélection réussie d'un projet de carrière.

Choisir une carrière appropriée est une tâche difficile, et avec le nombre accablant de champs de travail qui existent dans le monde professionnel moderne, il peut être difficile de savoir vers où s'orienter. **VOCATION** aide à identifier le travail le plus adéquat à un candidat en raison de ses intérêts. En choisissant une carrière sur la base de ses intérêts, un individu peut trouver une plus grande satisfaction au travail.

Cela conduit à parler de l'importance de l'évaluation des intérêts dans le recrutement. Afin d'embaucher avec succès un candidat qui se sentira satisfait de son travail et, par conséquent, souhaite rester avec l'entreprise, la satisfaction professionnelle de l'employé devient importante pour le recruteur. En utilisant **VOCATION**, un employeur sera en mesure d'inférer à quel point un emploi est approprié pour un candidat en fonction d'une évaluation objective de ses intérêts. Si les intérêts associés à ce candidat ne s'harmonisent pas avec ceux du rôle de travail en question, il pourrait y avoir un problème de concordance entre les deux, et cela pourrait générer des frustrations, un taux d'absentéisme plus élevé, ou encore une baisse de productivité dû à une motivation réduite pour les tâches à accomplir.

Le recrutement et la sélection sont une fonction organisationnelle importante qui peut prendre beaucoup de temps et coûter cher. En outre, les mauvaises décisions de recrutement ont des effets supplémentaires sur la productivité, la capacité et la durabilité de l'organisation. Si, par exemple, quelqu'un travaillant dans l'administration financière et juridique se sent plus à l'aise lorsqu'il fait face à des tâches qui impliquent le suivi de règles strictes, une personne travaillant dans le marketing se sentira lui plus à l'aise dans des tâches plus imaginatives et créatives. Il est recommandé qu'avant d'administrer **VOCATION**, les recruteurs aient une description détaillée du poste et des spécifications personnelles issue d'une analyse de travail.

VOCATION aide les recruteurs à personnaliser les profils d'emploi dans leur compte client en spécifiant les scores souhaités pour chaque facteur. L'évaluation permet au recruteur de comparer facilement le profil cible avec le profil d'un candidat.

Par ailleurs, **VOCATION** peut être utilisé dans le cadre de l'évaluation des performances, qui est un élément clé du développement et de la réussite des employés. Les examens du rendement devraient être des évaluations équitables et équilibrées des réalisations et des capacités d'un employé et

devraient donner l'occasion de discuter des attentes et des réalisations. **VOCATION** soutient les praticiens des ressources humaines avec des examens du rendement qui peuvent aider à discuter de la performance au travail en relation avec les intérêts et aspirations professionnelles ; si un individu ne performe pas autant qu'il le pourrait au travail, il se peut que ce soit dû à un manque d'intérêt pour le domaine particulier sur lequel il travaille. L'identification de cela peut aider le recruteur à déterminer si l'employé doit ou non être transféré dans un autre secteur de l'entreprise ou à un aspect différent du projet auquel il est assigné.

Bien que **VOCATION** soit efficace pour identifier les domaines de développement, il ne devrait pas être utilisé pour suivre directement les progrès de la performance d'un employé. Nous recommandons donc l'utilisation d'évaluations objectives du rendement ou d'autres évaluations conjointement avec **VOCATION**.

Références

- Baron, H. (1996). Strengths and limitations of ipsative measurement. *Journal of Occupational and Organizational Psychology*, 69(1), 49-56.
- Bartram, D. (2007). Increasing Validity with Forced-Choice Criterion Measurement Formats. *International Journal of Selection and Assessment*, 263-272.
- Barrick, M. R., Mount, M. K., & Gupta, R. (2003). Meta-analysis of the relationship between the five-factor model of personality and Holland's occupational types. *Personnel Psychology*, 56, 45-74. <http://dx.doi.org/10.1111/j.1744-6570.2003.tb00143.x>
- Bullock, E. E., Andrews, L. M., Braud, J., & Reardon, R. C. (2009-2010, Winter). Holland's theory in an international context: Applicability of RIASEC structure and assessments. *Career Planning and Adult Development Journal*, 25(4), 29-58.
- Brown, A., & Maydeu-Olivares, A. (2011, 06). Item Response Modeling of Forced-Choice Questionnaires. *Educational and Psychological Measurement*, 71(3), 460-502. doi:10.1177/0013164410375112
- Brown, A., & Maydeu-Olivares, A. (2013). How IRT can solve problems of ipsative data in forced-choice questionnaires. *Psychological Methods*, 18(1), 36-52. <http://doi.org/10.1037/a0030641>
- Chan, W. (2003). Analyzing ipsative data in psychological research. *Behaviormetrika*. <http://doi.org/10.2333/bhmk.30.99>, 99-121.
- Christiansen, N., Sliter, M., & Frost, C. T. (2014b). What employees dislike about their jobs: Relationship between personality-based fit and work satisfaction. *Personality and Individual Differences*, 71, 25-29. <http://doi.org/10.1016/j.paid.2014.07.013>
- Clauser, B. E., & Mazor, K. M. (1998). Using Statistical Procedures to Identify Differentially Functioning Test Items. *Educational Measurement: Issues and Practice*, 17(1), 31-44. <http://doi.org/10.1111/j.1745-3992.1998.tb00619.x>
- Costa, P. T., & McCrae, R. R. (1992). *Revised NEO personality inventory (NEO-PI-R) and NEO five-factor (NEO-FFI) inventory professional manual*. Odessa, FL: Psychological Assessment Resources.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281-302. <http://doi.org/10.1037/h0040957>
- Cruise, P. A. (2011). *The Role of Culture in Personnel Selection: Does culture affect hiring the right person for the job?* Germany: Lambert Academic Publishing.
- Cruise, P. A. (2012). The stability and organisational value of personality assessment. *Psyche Magazine Vol. 64, Spring*, pg 14-15. UK: Psychometrics Forum. Retrieved from <http://www.psychometricsforum.org/uploads/Psyche/Psyche%2064%20-%20Spring%202012.pdf>
- De Fruyt, F., & Mervielde, I. (1999). RIASEC types and big five traits as predictors of employment status and nature of employment. *Personnel Psychology*, 52(3), 701-727. <http://doi.org/10.1111/j.1744-6570.1999.tb00177.x>

- Dingemanse, S. A., De Fruyt, F., Vergauwe, J., & Wille, B. (2015). A closer look at the psychological diversity within Holland interest types: Construct validation of the Career Insight. *Consulting Psychology Journal: Practice and Research*, 67(3), 234-257.
- Dingemanse, S. A., Van Amstel, B., De Fruyt, F., & Wille, B. (2007). *Loopbaan Inzicht Vragenlijst. Handleiding [Career Insight Questionnaire. Manual]*. Amsterdam: Harcourt Assessment.
- Drasgow, F. (1984). Scrutinizing psychological tests: Measurement equivalence and equivalent relations with external variables are the central issues. *Psychological Bulletin*, 95(1), 134-135. <http://doi.org/10.1037/0033-2909.95.1.134>
- Drasgow, F. (1987). Study of the measurement bias of two standardized psychological tests. *Journal of Applied Psychology*, 72(1), 19-29. <http://doi.org/10.1037/0021-9010.72.1.19>
- Dodou, D., & de Winter, J. C. F. (2014). Social desirability is the same in offline, online, and paper surveys: A meta-analysis. *Computers in Human Behavior*, 36, 487-495. <http://doi.org/10.1016/j.chb.2014.04.005>
- Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, 105(3), 399-412. <http://doi.org/10.1111/bjop.12046>
- Fan, X., Thompson, B., & Wang, L. (1999). Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 56-83. <http://doi.org/10.1080/10705519909540119>
- Gilliland, S. W. (1993). The Perceived Fairness of Selection Systems: An Organizational Justice Perspective. *Academy of Management Review*, 18(4), 694-734. <http://doi.org/10.5465/AMR.1993.9402210155>
- Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality Psychology in Europe, Vol. 7* (pp. 7-28). Tilburg, The Netherlands: Tilburg University Press.
- Gottfredson, G. D. (1999). John L. Holland's contributions to **VOCATIONAL** psychology: A review and evaluation. *Journal of VOCATIONAL Behaviour*, 55, 15-40.
- Harmon, L. W., & Zytowski, D. G. (1980). Reliability of Holland Codes Across Interest Measures for Adult Females. *Journal of Counseling Psychology*, 27 (5), 478-483.
- Harvill, L. M. (1991). Standard Error of Measurement. *Educational Measurement: Issues and Practice*, 10(2), 33-41. <http://doi.org/10.1111/j.1745-3992.1991.tb00195.x>
- Hanson, B. A. (1998). Uniform DIF and DIF Defined by Differences in Item Response Functions. *Journal of Educational and Behavioral Statistics*, 23(3), 244-253. <http://doi.org/10.3102/10769986023003244>
- Hausknecht, J. P., Day, D. V., & Thomas, S. C. (2004). Applicant Reactions to Selection Procedures: An Updated Model and Meta-Analysis. *Personnel Psychology*, 57(3), 639-683. <http://doi.org/10.1111/j.1744-6570.2004.00003.x>

- Hays, R. D., & Fayers, P. (2005). Evaluating multi-item scales. *Assessing quality of life in clinical trials*, 41–54.
- Holland, J. L. (1958). A personality Inventory Employing Occupational Titles. *Journal of Applied Psychology*, 42(5), 336 – 342.
- Holland, J.L. and Holland, J. E. (1977). **VOCATIONAL** indecision: More evidence and speculation. *Journal of Counseling Psychology*, 24(5), 404-424.
- Holland, J. L. (1985). **VOCATIONAL** preference inventory (VPI): *Professional manual*. Odessa, Florida, US: Psychological Assessment Resources.
- Holland, J. L. (1997). *Making VOCATIONAL choices: A theory of VOCATIONAL personalities and work environments (3rd ed.)*. Odessa, Florida, US: Psychological Assessment Resources.
- Holt, J. C., Duijn, M. A., & Boomsma, A. (2010). Scale construction and evaluation in practice: A review of factor analysis versus item response theory applications. *Psychological Test and Assessment Modeling*, 52(3), 272-297
- Holtzman, S., & Vezzu, S. (2011). *Confirmatory Factor Analysis and Structural Equation Modeling of Noncognitive Assessments using PROC CALIS*. Retrieved from <http://www.lexjansen.com/nesug/nesug11/sa/sa07.pdf>
- Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modelling: Guidelines for determining model fit. *Articles*, 2.
- Hubley, A. M., & Zumbo, B. D. (2011). Validity and the consequences of test interpretation and use. *Social Indicators Research*, 103(2), 219-230. <http://doi.org/10.1007/s11205-011-9843-4>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55. <http://doi.org/10.1080/10705519909540118>
- Iddekinge, C. H., Putka, D. J., Roth, P. L., & Lanivich, S. E. (2011). A meta-analysis of relations between **VOCATIONAL** interests and employee performance and turnover. *Journal of Applied Psychology*, 1167-1194.
- Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language*. Scientific Software International.
- Kaplan, R. M., & Saccuzzo, D. P. (2008). *Psychological Testing: Principles, Applications, and Issues (7th ed.)*. Belmont, CA: Wadsworth Publishing Co Inc.
- Larson, L. M., Rottinghaus, P. J., & Borgen, F. H. (2002). Meta-analyses of Big Six Interests and Big Five Personality Factors. *Journal of VOCATIONAL Behavior*, 61, 217-239.
- McCarthy, J. M., & Goffin, R. D. (2003). Is the Test Attitude Survey Psychometrically Sound? *Educational and Psychological Measurement*, 63(3), 446-464. <http://doi.org/10.1177/0013164403063003007>
- McLeary, C., & Cruise, P. (2012). Modeling Organisational Trust: Addressing Challenges of Measurement Invariance. In *Proceedings of the 11th European Conference on Research Methods (p. 470)*. Academic Conferences Limited.

- Messick, S. (1980). *Test validity and the ethics of assessment*. *American Psychologist*, 35(11), 1012-1027.
<http://doi.org/10.1037/0003-066X.35.11.1012>
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741-749.
<http://doi.org/10.1037/0003-066X.50.9.741>
- Muthén, L. K., & Muthén, B. O. (2014). *Mplus User's Guide (7th ed.)*. Los Angeles, CA: Muthén & Muthén.
- Nevo, B. (1985). Face validity revisited. *Journal of Educational Measurement*, 22(4), 287-293.
<http://doi.org/10.1111/j.1745-3984.1985.tb01065.x>
- Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw-Hill.
- Nauta, M. M. (2010). The Development, Evolution, and Status of Holland's Theory of **VOCATIONAL** Personalities Reflections and Future Directions for Counseling Psychology. *Journal of Counseling Psychology*, 11-22.
- Pozzebon, J. A., Visser, B. A., Ashton, M. C., Lee, K., & Goldberg, L. R. (2010). Psychometric characteristics of a public-domain self-report measure of **VOCATIONAL** interests: The Oregon **VOCATIONAL** Interest Scales. *Journal of Personality Assessment*, 92, 168-174.
- Saville, P., & Willson, E. (1991). The reliability and validity of normative and ipsative approaches in the measurement of personality. *Journal of Occupational Psychology*, 64(3), 219-238.
<http://doi.org/10.1111/j.2044-8325.1991.tb00556.x>
- Schinka, J. A., Dye, D. A., & Curtiss, G. (1997). Correspondence Between Five-Factor and RIASEC Models of Personality. *Journal of Personality Assessment*, 355-368.
- Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review. *The Journal of Educational Research*, 99(6), 323-338.
<http://doi.org/10.3200/JOER.99.6.323-338>
- Steiger, J. H., & Lind, J. C. (1980). Statistically based tests for the number of common factors. *In annual meeting of the Psychometric Society, Iowa City, IA (Vol. 758)*.
- Tett, R. P., & Simonet, D. V. (2011). Faking in Personality Assessment: A « Multisaturation » Perspective on Faking as Performance. *Human Performance*, 24(4), 302-321.
- Thurstone, L. L. (1927). The method of paired comparisons for social values. *The Journal of Abnormal and Social Psychology*, 21(4), 384-400.
<http://doi.org/10.1037/h0065439>
- Thurstone, L. L. (1931). Multiple factor analysis. *Psychological Review*, 38(5), 406-427.
<http://doi.org/10.1037/h0069792>
- Zumbo, B. D., & Thomas, D. R. (1997). *A measure of effect size for a model-based approach for studying DIF*. Prince George, Canada: Edgeworth Laboratory for Quantitative Behavioral Science, University of Northern British Columbia.

NOUS CONTACTER

Service Client : **service.client@centraltest.com**

Formation et Consulting : **consulting@centraltest.com**



www.centraltest.com