

MANUAL TÉCNICO & GUÍA DE USO VOCATION

Este manual ha sido elaborado por los psicólogos/as del departamento de I+D de Central Test.

El test VOCATION y el logo de Central Test son marca registrada.

Queda prohibida la reproducción, modificación o adaptación sin el consentimiento expreso por escrito del editor Central Test.

Índice

Introducción	4
1. Descripción y uso	5
1.1. Factores	5
1.2. Aplicación y combinación	6
1.3. Estructura del cuestionario	7
1.3.1. Formato	7
1.3.2. Preguntas	7
1.4. Administración, idiomas y adaptaciones para necesidades especiales	8
1.5. Deseabilidad social	8
1.6. Instrucciones para la administración	8
1.7. Información técnica y deontológica	9
1.7.1. Aplicación del código deontológico	9
1.7.2. Privacidad y seguridad de los datos del usuario/a	10
1.7.3. Características del sistema informático y disponibilidad del servidor	10
1.7.4. Generación de los informes	11
1.8. Condiciones de suministro y disponibilidad	11
2. Marco conceptual y desarrollo	12
2.1. Marco teórico	12
2.2. Desarrollo del test	14
2.2.1. Factores y elaboración de ítems	15
2.2.2. Pruebas preliminares	15
2.2.3. Definiciones de factor	16
2.2.4. Perfiles combinados	20
2.2.5. Formato de respuesta ipsativa	20
2.2.6. Teoría de respuesta a los ítems de thurstone (tri) y las respuestas ipsativas	21
3. Puntuación	25
3.1. Relación entre puntuaciones brutas y baremadas	25
3.2. Conversión de las puntuaciones directas a puntuaciones típicas	27
3.3. Error típico de medida	28
3.4. Características del grupo normativo adecuado	29
3.5. Generación de parámetros para una población concreta	29
3.6. Puntuación en los informes	30
	2

4. Descripción de la muestra	31
5. Baremos	33
6. Fiabilidad	46
6.1. Consistencia interna	46
6.2. Fiabilidad test-retest	48
7. Validez	50
7.1. Validez concurrente	50
7.1.1. La escala de intereses vocacionales oregón de ipip	51
7.2. Validez de constructo	53
7.2.1. Análisis factorial y correlaciones entre los factores	54
7.2.2. Análisis factorial confirmatorio y exploratorio	54
7.2.3. Validez consecuencial	57
8. Vocation y la igualdad de oportunidades	58
8.1. T-tests y anovas	58
9. Información y consideraciones adicionales	69
9.1. Interpretación de las puntuaciones	69
9.2. Necesidades organizacionales: elección de carrera, contratación y desarrollo	70
Referencias	72

Introducción

VOCATION es un test para evaluar los intereses profesionales y la preferencia ocupacional de las personas. Está basado en el modelo tipológico de intereses vocacionales de John Holland (1959) y su objetivo es proporcionar información sobre las áreas ocupacionales que encajan con sus intereses profesionales de la persona.

VOCATION es una actualización del **TEST DE INTERESES VOCACIONALES-R**, sus ventajas son:

- Ítems actualizados, adaptados al sector empresarial y organizacional
- Cuestionario accesible y breve, con un tiempo medio de realización de 15 minutos
- Control de los sesgos de respuesta y la deseabilidad social (formato ipsativo)
- Uso de técnicas psicométricas establecidas como la TRI de Thurstone
- Informes funcionales, claros y personalizables
- Perfil personal donde se combinan los tres factores más importantes

Este manual técnico y guía de uso explica el modelo teórico de **VOCATION**, los factores y subfactores que lo componen, el diseño y la estructura del cuestionario, las características psicométricas, la investigación del test, baremos, administración, interpretación y su uso en diferentes procesos en la organización.

1. Descripción y uso

VOCATION establece la relación entre los intereses profesionales de la persona y la personalidad. Está basado en el modelo RIASEC con sus 6 factores principales: *Realista, Investigador, Artístico, Social, Emprendedor y Convencional*. Cada uno de estos factores principales se ha subdividido en dos para obtener información más detallada del perfil de la persona. A continuación, se proporciona una descripción de los factores.

1.1. Factores

El modelo RIASEC utilizado en **VOCATION** está compuesto por los siguientes seis factores:

Realista: Las personas con puntuación alta en este factor tienen preferencia por actividades físicas o manuales. Estarán motivadas por tareas de tipo práctico, se sentirán atraídas por actividades que requieran resistencia física o destreza y disfrutarán en actividades al aire libre o que impliquen trabajar con las manos, por ejemplo, arreglando o ensamblando objetos.

Investigador: Las personas con puntuación elevada en “investigador/a” tienen preferencia por investigar o desarrollar conocimiento teórico. Disfrutan solucionando problemas, realizando estudios o investigaciones para entender mejor un tema o el entorno que les rodea.

Artístico: Las personas con puntuación elevada en “artístico” tienen preferencia por expresarse de forma creativa, mediante la música, la pintura o el arte. Les gustarán las actividades que les permitan usar su intuición e imaginación y disponer de margen para improvisar.

Social: Las personas con puntuación alta en “social” tendrán interés en actividades en las que tengan que comunicarse, relacionarse y/o ayudar a los demás. Preferirán actividades que les permitan conocer a personas nuevas y en las que puedan sentirse parte de un grupo o comunidad.

Emprendedor: Las personas con puntuación alta en “emprendedor/a” mostrarán tendencia hacia actividades que impliquen iniciar proyectos, tener influencia sobre otras personas y/o asumir riesgos. Es probable que disfruten intentando convencer a otros de sus ideas o en tareas de coordinación o dirección. Probablemente serán personas dinámicas y con capacidad para tomar la iniciativa.

Convencional: Las personas con puntuación alta en “convencional” tendrán preferencia por actividades que conlleven realizar las tareas de forma metódica y ordenada. Es probable que prefieran ajustarse a las reglas o normativas y que las actividades estén estructuradas y organizadas. También es probable que disfruten con tareas que requieran precisión, atención al detalle o trabajar con números.

1.2. Aplicación y combinación

VOCATION se diseñó para su uso en organizaciones, centros educativos y/o empresas. Se puede utilizar en centros educativos como herramienta de apoyo para los programas de orientación laboral y también es útil en el ámbito de RRHH. Por ejemplo, en selección de personal, movilidad o desarrollo profesional.

Además, al ser una herramienta versátil y completa, puede resultar de utilidad en los programas de orientación y desarrollo profesional, desarrollo personal, mentoring y coaching. La siguiente tabla describe algunas de las posibles aplicaciones del test:

Tabla 1. Aplicaciones de VOCATION.

Orientación Profesional	Desarrollo Profesional	Mentoring / Coaching
Ayudar a estudiantes, becarios/as y jóvenes profesionales a decidir su trayectoria profesional	Ayudar a profesionales a explorar nuevas oportunidades y aumentar sus competencias	Entender los intereses vocacionales de la persona y dar orientación laboral

VOCATION puede utilizarse en combinación con otros cuestionarios para obtener un perfil más completo de la persona. Algunos test que pueden ser especialmente útiles son **PERFIL PRO 2** y **TEST DE VENTAS-R**.

La siguiente tabla muestra dos ejemplos de cómo **VOCATION** puede combinarse, aunque la elección de los tests para completar el perfil obtenido depende de las necesidades en cada caso.

Tabla 2. Ejemplos de combinaciones.

Test adicional	Objetivos	Ventajas
PERFIL PRO 2	Orientación profesional	Información sobre el ajuste entre personalidad e intereses para una profesión determinada.

1.3. Estructura del cuestionario

En **VOCATION** se ha cuidado que el cuestionario fuera sencillo y rápido, pero con el número de preguntas necesario para poder medir los intereses de forma precisa y un formato accesible para los usuarios/as.

1.3.1. Formato

VOCATION utiliza un formato de elección forzada en sus preguntas. Este formato implica que se debe elegir la opción que más se ajuste con la realidad o preferencia de la persona en cada una de las preguntas.

El formato de elección forzada se utiliza para reducir los sesgos en la respuesta (Chan, 2003; ten Holt, van Duijn & Boomsma, 2010), permiten reducir la posibilidad de responder de forma que se distorsione la imagen o mostrarse de forma diferente a cómo se es en realidad (Baron, 1996; Bartram, 2007) y son más difíciles de falsificar (Brown & Maydeu-Olivares, 2011). Este formato dificulta responder a los ítems de la misma forma y permite mayor variabilidad en las respuestas (Cruise, 2011/2012). Este tipo de formato permite tener un nivel mayor de confianza en que las respuestas representen con mayor exactitud las preferencias profesionales de la persona.

1.3.2. Preguntas

El cuestionario está compuesto por 60 preguntas, con dos opciones cada una, que evalúan los 6 factores; cada factor está asimismo dividido en dos subfactores, por lo que hay 12 subfactores en total. Se deben contestar todas las preguntas de forma obligatoria y en el orden establecido. En la imagen se puede ver un ejemplo de pregunta.

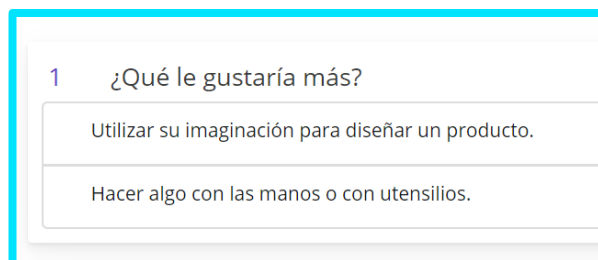
A screenshot of a questionnaire question, labeled '1'. The question is '¿Qué le gustaría más?'. Below the question, there are two options: 'Utilizar su imaginación para diseñar un producto.' and 'Hacer algo con las manos o con utensilios.' The entire question and options are enclosed in a light gray box with a thin blue border.

Figura 1. Ejemplo de pregunta.

1.4. Administración, idiomas y adaptaciones para necesidades especiales

VOCATION se administra mediante internet y es compatible con cualquier sistema operativo actual incluyendo Windows, Mac-OS y Linux. Al igual que la mayoría de los test modernos, la evaluación se puede realizar a distancia y sin necesidad de supervisión. Sin embargo, existe la posibilidad de realizar evaluaciones en grupo bajo condiciones controladas si fuera necesario. Los informes se generan online solamente.

VOCATION está disponible en español, francés, inglés, alemán, holandés y árabe. Para más información sobre la disponibilidad de idiomas puede contactar con el servicio de consultoría en *consulting@centraltest.com*.

VOCATION se diseñó para su uso por personas de todos los grupos sociales y/o con dificultades perceptivas. La presentación del test en la pantalla puede ajustarse, según las necesidades individuales. Por ejemplo, si hay dificultad visual, o se necesita un tamaño de fuente mayor, se puede aumentar mediante "CTRL" y "+". Las instrucciones del test están disponibles durante todo el tiempo. Se puede contactar con *Central Test*, editor del test para cualquier asesoramiento respecto a la administración.

1.5. Deseabilidad social

La deseabilidad social es la tendencia de las personas a presentarse de manera positiva, se debe a la propia naturaleza del ser humano a vivir en sociedad y tener la necesidad de buscar la aceptación social. Si se hace de forma no intencionada, se considera autoengaño; pero, si es intencional se considera un intento de manipular la imagen que perciben los demás (Tett & Simonet, 2011). Al contrario que en los test de habilidad cognitiva, los cuestionarios como **VOCATION** son vulnerables a este tipo de falsificación que puede suponer una fuente de error en los resultados (Dodou & de Winter, 2014).

VOCATION previene la deseabilidad social de forma activa, la medida más eficaz es el uso de un formato de respuesta forzada o ipsativo. La investigación indica que los formatos ipsativos son menos sensibles a la distorsión que otros formatos como las escalas Likert (Bowen, Martin, & Hunt, 2002; Kluger, Reilly, & Russell, 1991). El formato ipsativo reduce la habilidad de los candidatos/as para distorsionar las puntuaciones y hace más difícil que puedan poner en práctica estrategias de respuesta. Para prevenir en mayor medida los sesgos de respuesta y la deseabilidad social, los ítems se han emparejado de forma que las dos opciones de respuesta sean igual de atractivas.

1.6. Instrucciones para la administración

VOCATION está diseñado para su administrarse online, mediante la plataforma virtual de *Central Test*, por lo que habitualmente no hay supervisión. En el caso de que se realice una administración supervisada, se pueden tener en cuenta las siguientes

recomendaciones.

Aunque las instrucciones para realizar **VOCATION** se pueden leer en la pantalla, puede facilitarse la administración explicándoselas a los candidatos/as y resolviendo las dudas que puedan surgir.

- Si alguien tiene dudas después de que la prueba haya comenzado, se intentará responder sin molestar a los demás.
- Si alguien no entiende una pregunta, se responderá evitando influir la respuesta.
- Si se sospecha que hay dificultades debido al nivel educativo, se valorará no continuar con el test.
- Si hay dificultades debido al idioma, se valorará realizar la evaluación en un idioma que resulte más sencillo para la persona.
- Si se cree que alguien está intentando manipular la respuesta se le orientará hacia la honestidad en las respuestas.

Una vez se ha invitado a una persona a hacer el test, esta recibirá la invitación mediante correo electrónico, con un enlace a la página web y las claves de acceso. El enlace al cuestionario está protegido con una contraseña única para la cuenta de cada persona. Es posible fijar una fecha tope para la realización del test. Además, es posible realizar la invitación de forma que la evaluación se deba completar en una dirección IP determinada.

VOCATION no tiene límite de tiempo, el tiempo medio de administración es de 15 minutos. Aunque no es un test cronometrado, el tiempo de realización se registra por motivos de supervisión y seguimiento. Permite saber si se ha realizado demasiado rápido o se ha necesitado un tiempo excesivo para explorar posibles errores en las puntuaciones. Es preferible que se conteste de forma espontánea, sin prisa, evitando pensar demasiado las respuestas. Al igual que otros test psicométricos para evaluar características como la personalidad o la motivación, no hay respuestas correctas o incorrectas, el objetivo es obtener el perfil correspondiente a la persona.

1.7. Información técnica y deontológica

En este apartado se especifican las condiciones de administración y éticas en la administración de **VOCATION**. Estas directrices permiten que la administración de **VOCATION** sea objetiva e imparcial, y resulte una experiencia agradable para la persona que lo realiza. Las condiciones de seguridad permiten proteger la información del candidato/a, durante y después de realizarlo.

1.7.1. Aplicación del código deontológico

Como empresa editorial de soluciones de evaluación *online*, *Central Test* está sujeto a las regulaciones legales de la Ley Francesa del Trabajo y a las directrices de la ley del 6 de enero de 1978, para el uso y almacenamiento de información tecnológica, archivos y

libertades: Declaración N°1834067. Además, se ajusta a las regulaciones del Parlamento y el Consejo Europeo que han entrado en vigor el 25 de mayo de 2018, RGPD o Reglamento General de Protección de Datos.

Los resultados obtenidos en los test no deben compartirse con terceras personas sin el consentimiento explícito del usuario/a. Los usuarios/as tienen derecho a acceder a sus resultados. El administrador/a debe permitir el acceso automáticamente o cuando la persona lo solicite. El usuario/a tiene el derecho a solicitar que sus datos personales se eliminen de forma permanente. El procedimiento para la eliminación de los datos consiste en contactar a la organización que realiza la evaluación o a *Central Test* de forma directa para que proceda a eliminar los datos.

1.7.2. Privacidad y seguridad de los datos del usuario/a

Central Test garantiza la seguridad y privacidad de los datos que se generen durante la administración de las pruebas online. Los códigos de acceso están codificados en una base de datos a la que ninguna persona tiene acceso. Las contraseñas se generan por ordenador y las empresas no pueden visualizarlas. Se instala un módulo de reescritura de URL (que genera un enlace seguro y el acceso).

La plataforma de *Central Test* utiliza un sistema de seguridad y codificación SSL para garantizar la navegación e intercambio de información segura. La navegación segura está garantizada por el protocolo https que se puede identificar en el icono en forma de candado en el navegador de internet. La aplicación de *Central Test* tiene una firma digital "GeoTrust SSL True BusinessID", emitida y certificada por GeoTrust©. Esta certificación prueba nuestro compromiso con la seguridad y la protección de las evaluaciones de *Central Test* respecto a ataques para suplantar la identidad. Las invitaciones para realizar los test se realizan mediante enlaces seguros y con un tiempo de uso limitado (en cantidad y duración). Se registra el tiempo que se tarda en hacer el test y se hace un seguimiento de la dirección IP. El acceso a las pruebas puede limitarse por cantidad y por tiempo.

1.7.3. Características del sistema informático y disponibilidad del servidor

El hardware de la plataforma informática de *Central Test* consiste en 10 servidores físicos que proporcionan un servicio de alta calidad. Cada servidor se aloja en un centro de dato especializado y dispone de:

- Conexiones privadas
- Un ancho de banda de 100 Mbps (por servidor)
- Fuentes de alimentación adicionales
- 4 horas de tiempo de reiniciación del equipo
- 2 horas de tiempo de intervención en caso de incidentes
- Supervisión y asistencia continuada 24/7/365

El acceso al servidor de administración está restringido a una dirección de IP protegida con contraseña. El intercambio de datos es seguro. Las aplicaciones de *Central Test* han sido probadas y son compatibles con los siguientes navegadores: Internet Explorer 8 o superior, Firefox 21 o superior, Google Chrome 30 o superior y Safari 7 o superior.

1.7.4. Generación de los informes

Los informes se generan de forma automática, en pocos minutos, una vez que los usuarios/as han terminado el test, es posible obtener una copia en PDF. Los informes de **VOCATION** son funcionales y claros. El contenido del informe se puede adaptar fácilmente a las necesidades. El informe de **VOCATION**, al igual que todos los informes de *Central Test*, se puede personalizar.

1.8. Condiciones de suministro y disponibilidad

Las condiciones de suministro y disponibilidad para **VOCATION** varían en función del país. Sin embargo, existen una serie de condiciones generales.

- **VOCATION** está disponible para profesionales con suficientes conocimientos en psicometría.
- Es recomendable realizar una formación previa en el test para utilizarlo e interpretar los resultados de forma correcta.
- **VOCATION** puede solicitarse *online* o contactar con un comercial o distribuidor/a de *Central Test* en cualquier parte del mundo (customer@centraltest.com).
- Para obtener información sobre el coste de **VOCATION**, sus servicios y productos asociados puede contactar con cualquier comercial de *Central Test* o distribuidor/a en cualquier parte del mundo.
- Para obtener más información sobre la formación de la plataforma o acreditación puede contactar con el servicio de atención al cliente de *Central Test* o sus distribuidores/as.

2. Marco conceptual y desarrollo

En el desarrollo de cada uno de sus tests *Central Test* invierte en una extensa investigación tanto de la teoría como de los resultados en empresa del constructo a evaluar. En el desarrollo de **VOCATION** se investigó la relación entre los intereses y el comportamiento, el ajuste laboral y se identificó el soporte teórico que permitiera el desarrollo de un test útil, práctico y fácil de utilizar.

2.1. Marco teórico

VOCATION está basado en la teoría de intereses profesionales de John Holland (1959).

El Modelo de John Holland. La Tipología de Holland, también conocida como el *Modelo RIASEC*, fue desarrollada por Holland (1959) y se basa en la idea de que el interés vocacional de una persona o la elección de carrera vienen determinados por los “tipos de personalidad” subyacentes. El modelo es conocido por la sencillez de su enfoque, su estructura simple y su fácil accesibilidad para evaluar los intereses vocacionales. Aunque cada ser humano es tremendamente complejo, Holland (1977) cree que cada persona tiende a encajar en uno de los 6 tipos de personalidades básicas o una combinación de ellas, que él llama tipos - realista, investigador, artístico, social, emprendedor y convencional. Holland (1997) estipula que la elección de carrera constituye una expresión de su personalidad y que el Modelo RIASEC puede utilizarse también para identificar cuál es el ambiente profesional adecuado (Gottfredson, 1999).

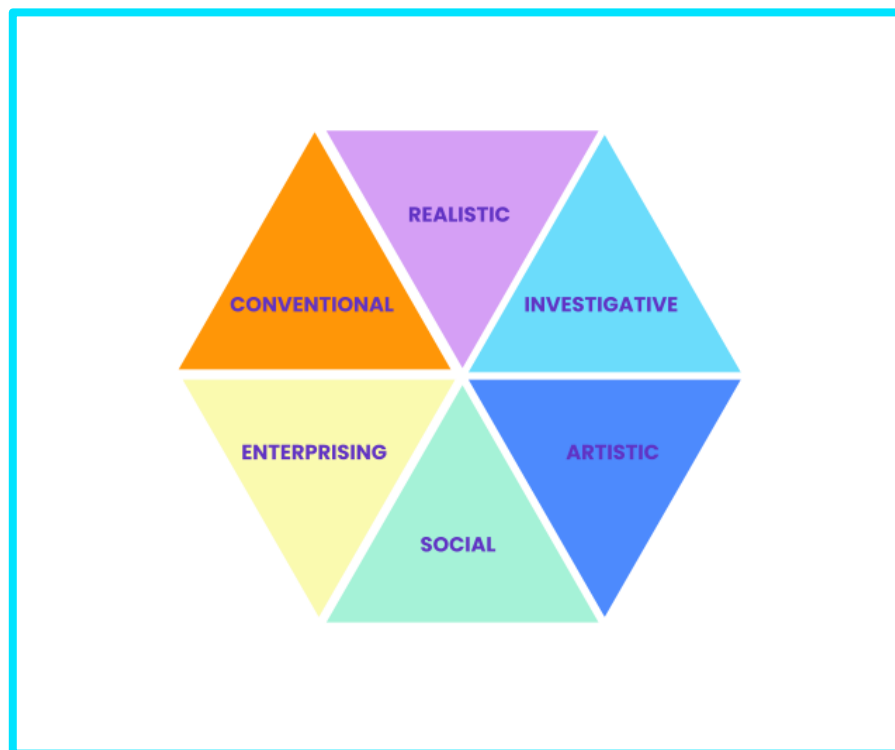


Figura 2. *Modelo RIASEC de Holland (1977).*

Holland (1959) desarrolló su teoría durante el transcurso de su carrera, en 1950 se alistó en el ejército para ejercer de psicólogo en prácticas. Nauta (2010) declaró que muchas de las ideas que Holland desarrolló en su teoría fueron fruto de sus prácticas de graduación y de las dificultades a las que se enfrentó cuando ejercía de orientador profesional. También comenta cómo Holland se dio cuenta cuando servía en el ejército, que las personas encajaban en tipos de personalidad, según su vocación profesional. Holland observó que las contrataciones del ejército se basaban en los resultados de las entrevistas y fue entonces cuando se le ocurrió la idea de clasificar a las personas en tipos psicológicos (Nauta, 2010). Holland, avalado por su experiencia en el campo de la orientación profesional, empezó a darse cuenta de las limitaciones de test vocacionales disponibles en aquellos tiempos. La mayor parte de ellos consistía en pruebas largas, generalmente tipo Likert, y no existía ninguna lista de ocupaciones que pudiera encajar con los resultados de las evaluaciones. Esto llevó a Holland (1985) a diseñar un inventario de intereses vocacionales. Holland continuó elaborando su teoría durante los años siguientes. Publicó la SDS (búsqueda autodirigida) en 1970 y ha sido revisada cuatro veces desde entonces. Así mismo, publicó el buscador de carreras en 1977; consiste en un sistema de clasificación ocupacional. Aunque se jubiló en 1980, siguió desarrollando su teoría. En 1997, la revisó y añadió un aspecto de identidad medioambiental (Staff, 2008).

Nauta (2010) explica cómo Holland basándose en seis tipos de personalidad, desarrolló el Modelo RIASEC. Las siglas son un acrónimo de *Realista, Investigador, Artístico, Social, Emprendedor y Convencional*. Cada tipo de personalidad está definido por una combinación de intereses, actividades preferidas, creencias, habilidades, valores y característica. Según Nauta (2010), la mayor aportación de Holland es darse cuenta del hecho de que la mayoría de la gente busca un ambiente de trabajo que le permita desarrollar sus competencias y aptitudes, asumiendo el papel en los que se sienten más cómodos/as. La teoría de Holland ha sido ampliamente aceptada y su aplicación ha cristalizado en un cambio radical de paradigma en la medición de las decisiones profesionales, la satisfacción y la estabilidad profesional (Bullock, Andrews and Braud, 2010).

Relaciones entre las dimensiones del RIASEC y los rasgos de personalidad. Se han llevado a cabo muchos estudios sobre la relación entre las dimensiones RIASEC y los rasgos de personalidad. Larson, Rottinghaus y Borgen (2002) realizaron un meta-análisis de las relaciones entre las dimensiones RIASEC de Holland y los factores de personalidad de los Big Five (B.F.); compuestos por apertura, responsabilidad, extroversión, amabilidad, neuroticismo (Costa & McCrae, 1992). En el meta-análisis se encontraron correlaciones significativas para la dimensión *Artística e Investigadora* (RIASEC) con “*Apertura*” (B.F.), *Emprendedor y Social* (RIASEC) con “*Extraversión*” (B.F.), así como *Social* (RIASEC) con “*Amabilidad*” (B.F.). Barrick et al (2003) también encontraron las correlaciones anteriormente citadas y adicionalmente correlación entre *Convencional* (RIASEC) con “*Responsabilidad*” (B.F.). Sin embargo, parece que no existe correlación entre ningún factor del modelo RIASEC y “*Neuroticismo*” (B.F.) (Schinka, Dye, & Curtiss, 1997).

Estas correlaciones sugieren que se pueden inferir los intereses ocupacionales a partir de los rasgos de personalidad y viceversa; los rasgos de personalidad a partir de los

intereses ocupacionales. Sin embargo, la investigación apunta a que los intereses ocupacionales son más fiables a la hora de predecir la elección de un puesto que la personalidad (Fruyt & Mervielde, 1999), por lo tanto, pueden ser predictores adecuados del éxito que se tendrá en el desempeño de un determinado puesto de trabajo (Iddekinge, Putka, Roth, & Lanivich, 2011).

División de los factores RIASEC. La investigación ha demostrado que las dimensiones del modelo RIASEC de Holland (1958) son muy amplias. Los seis factores, Realista, Investigador, Artístico, Social, Emprendedor y Convencional son extensos, por lo que, considerando la complejidad del sector profesional actual, deberían subdividirse para identificar un rango más amplio de diferencias individuales. Dingemanse, Van Amstel, De Fruyt, & Wille (2007) estudiaron con más atención la idea de que deberían existir más divisiones *dentro* de los tipos RIASEC en lugar de *entre* ellos. Crearon el *Career Insight Questionnaire (CIQ)*, (Cuestionario de exploración profesional) en el que reconocieron 15 “elementos básicos que construyen los tipos de Holland”, a partir de los seis factores originales del modelo *RIASEC* (Dingemanse, Van Amstel, De Fruyt & Wille, 2015, p.237). Por ejemplo, Realista se dividió en “Artesano”, “Actividades al aire libre” y “Mecánica/Construcción”, e Investigador en “Ciencia” y “Teoría”. Así, Dingemanse et al (2015) compararon sus factores con los rasgos de personalidad de los Big Five de Costa y McCrae, y descubrieron que las correlaciones entre los rasgos de personalidad de los Big Five y las divisiones de los factores de RIASEC no se mantenían. Así, encontraron que existía una correlación negativa entre “Amabilidad” (B.F.) y el componente científico de Investigador (RIASEC), lo cual implica que, si una persona obtiene una puntuación alta en amabilidad, solo podrá tener puntuación alta en el componente curiosidad intelectual de Investigador (RIASEC).

Partiendo de los resultados de esta investigación, en la elaboración del test **VOCATION** se decidió dividir cada uno de los factores RIASEC en dos, con el objetivo de lograr una representación mucho más precisa del rango de los intereses ocupacionales de cada persona. Se explican en detalle en el apartado Definiciones del factor.

2.2. Desarrollo del Test

El diseño de **VOCATION** ha pasado por varias etapas. Su precursor, el *Test de intereses vocacionales-R* se elaboró en 2006. Durante la revisión para desarrollar **VOCATION** los ítems y factores del cuestionario anterior se actualizaron, el objetivo era que los nuevos ítems reflejaran las tendencias del mercado laboral actual, evitar distorsiones en las respuestas y la deseabilidad social. Esto permitiría a *Central Test* proporcionar una herramienta mejorada, que facilitase la toma de decisiones a los profesionales en sus procesos de evaluación y así mismo orientar de forma adecuada a las personas evaluadas respecto a la trayectoria profesional. Por último, en la revisión de **VOCATION** se cuidó que tanto la estructura de los factores como de los ítems se ajustara a altos niveles de fiabilidad y validez.

2.2.1. Factores y elaboración de ítems

Las fases de desarrollo de **VOCATION** fueron:

- Se recogió información mediante entrevistas a clientes y las respuestas a los cuestionarios en talleres de aprendizaje. El objetivo era asegurarse, que tanto el cuestionario como los factores medidos respondían a las necesidades del mercado actual.
- Un panel de expertos compuesto por profesionales de la psicometría e investigadores revisó los ítems para comprobar la validez teórica.
- El cuestionario se probó seis veces en una muestra de población amplia y representativa. Se analizaron los resultados y se retiraron y/o modificaron los ítems basándose en sus cualidades psicométricas para lograr una fiabilidad y validez alta.

Todos los cambios se realizaron con la idea de garantizar que **VOCATION** midiera de forma precisa y rigurosa las opciones profesionales y los intereses ocupacionales del mercado actual. La siguiente tabla muestran dos ejemplos de cómo se modificaron los ítems durante este proceso.

Tabla 3. Ejemplos de cambios en los ítems durante la revisión.

Factor	Ítem del test anterior	VOCATION
Artístico, A1	Ayudar escribiendo el guión	Expresar su creatividad mediante la redacción de un guión
Emprendedor, E2	En la oficina, a cargo de los planes empresariales.	Impulso de nuevos proyectos para atraer visitantes

2.2.2. Pruebas preliminares

Las pruebas preliminares se llevaron a cabo con una serie de profesionales de ambos géneros y procedentes de diferentes lugares en el mundo, con el objetivo de comprobar que las preguntas fueran relevantes, claras, precisas, concisas, sin sesgos o rasgos culturales y que representasen adecuadamente el comportamiento laboral a evaluar. Después de realizar seis pruebas preliminares en inglés, el cuestionario final se probó con 309 profesionales de habla inglesa de Reino Unido. A continuación, se realizaron pretests con personas de habla inglesa de diferentes continentes, para comprobar que el cuestionario era útil para medir los intereses vocacionales, a nivel internacional.

Una vez recogidos los datos de las pruebas preliminares, se calculó la consistencia interna comprobando la sensibilidad de cada ítem en relación a su factor. Se

identificaron los factores cuya fiabilidad era inferior a 0,55. Se comprobó la independencia de los factores entre sí y se revisaron, modificaron o cambiaron los ítems ambiguos, probando así el test hasta obtener los niveles de calidad necesarios.

Una vez finalizada la versión inglesa se tradujeron y adaptaron culturalmente los ítems al español. El test se probó cinco veces con población de habla hispana y los ítems se seleccionaron de igual forma que con la versión inglesa. Una vez obtenidos los niveles de calidad necesaria se puso en fase piloto para realizar una última comprobación con los clientes de *Central Test*.

2.2.3. Definiciones de factor

Los factores RIASEC son amplios y es posible tener una mayor precisión si se identifican elementos que los componen, en **VOCATION** se han dividido para mejorar su funcionalidad, siguiendo las investigaciones citadas anteriormente. En consecuencia, en **VOCATION** se utilizan los seis factores de RIASEC, pero cada uno de ellos se descompone en dos subfactores.

El informe para el usuario/a del test contiene una breve descripción para cada factor de los rasgos de la personalidad, valores, aptitudes potenciales, estilo de aprendizaje y entorno que podrían atribuirse a la persona si sus respuestas son coherentes con interés en el factor. Los tipos de personalidad se elaboraron utilizando la investigación de Holland (1985) y otros estudios que relacionan los rasgos de personalidad con el modelo RIASEC, discutidos anteriormente. Los *valores* asociados con cada factor reflejan aquellos valores de mayor importancia para la persona. Las *aptitudes potenciales* son las habilidades que la persona puede adquirir con mayor facilidad. El *estilo de aprendizaje* indica que método de aprendizaje es más eficaz para lograr que la persona comprenda y asimile nueva información. El *entorno* consiste en el estilo de entorno profesional que puede encajar con la persona, basándose en sus propios intereses.

Presentamos a continuación la definición de cada factor y de cada sub factor, así como la información adicional relacionada con los rasgos de personalidad, valores, aptitudes potenciales, estilo de aprendizaje y entorno.

Realista: Las personas con intereses de tipo "realista" tienen preferencia por actividades de tipo físico o manual. Orientados hacia la consecución de resultados, les atraen actividades que requieren destreza y resistencia física, por ello disfrutan de actividades realizadas en el exterior o que impliquen el uso de sus manos, como puede ser montar o reparar objetos. Este factor se subdivide en R1 y R2.

(R1) Actividades físicas y en el exterior: Preferencia por realizar actividades físicas y realizar el trabajo fuera de entornos cerrados o de oficina.

(R2) Intereses técnicos y manuales: Preferencia por tareas que impliquen construir objetos, trabajar con las manos y/o usan herramientas o maquinaria.

Personalidad: Tenaz, auténtico, natural y práctico.

Valores: Sentido común, pragmatismo y esfuerzo.

Aptitudes potenciales: Destreza, facilidad para las actividades técnicas, desenvoltura en actividades mecánicas y resistencia física.

Estilo de aprendizaje: Enfoque objetivo y pragmático, aprendizaje mediante el uso de ejemplos.

Entorno: Mayor comodidad en entornos que les permitan obtener resultados concretos, usar equipamiento técnico o trabajo en el campo. Tienden a estar incómodos en áreas cerradas como una oficina.

Investigador/a: El interés investigador/a incluiría a aquellas personas que disfrutan explorando y desarrollando teorías. Les gusta investigar, realizar estudios o solucionar problemas para poder entender mejor un área de conocimiento o el entorno que les rodea. Este factor se subdivide en I1 e I2.

(I1) Curiosidad y aprendizaje intelectual: Preferencia por actividades que requieren pensamiento crítico o exista la posibilidad de aumentar el conocimiento.

(I2) Ciencia y tecnología: Deseo de investigar y explorar temas científicos y relacionados con nuevas tecnologías.

Personalidad: Gran curiosidad por distintos temas, búsqueda de aprendizaje y deseo de entender su entorno. Suelen ser precisos, con inclinación intelectual y mentalidad científica.

Valores: Curiosidad intelectual, pensamiento crítico y lógico.

Aptitudes potenciales: Investigación, razonamiento lógico, capacidad de análisis y aprendizaje.

Estilo de aprendizaje: Enfoque teórico e investigación sistemática.

Entorno: Mayor comodidad en entornos orientados a la reflexión y que favorezcan la aplicación de conocimientos sobre la productividad.

Artístico/a: Las personas con interés en actividades "artísticas" tendrán propensión a expresarse de manera creativa mediante la música, el dibujo u otras formas de expresión artística. Estas personas tienden a tener la necesidad de expresar sus ideas y sentimientos de forma artística. Preferirán actividades que les permitan ser intuitivos e imaginativos y pueden sentir rechazo hacia proyectos convencionales que no permitan improvisación. Este factor se subdivide en A1 y A2 como se muestra a continuación:

(A1) Expresión y sentido estético: Interés en actividades que requieran sensibilidad artística y la expresión de las emociones.

(A2) Creatividad y diseño: Preferencia por crear, imaginar e innovar.

Personalidad: Intuición, creatividad, idealismo e independencia. Sensibilidad hacia el arte y la estética.

Valores: Apreciación por la belleza, la originalidad, la imaginación y la libertad.

Aptitudes potenciales: Originalidad, imaginación y la habilidad de apreciar diferentes perspectivas.

Estilo de aprendizaje: Metodologías que favorezcan la creatividad, intuición y la comunicación mediante el uso del arte, la palabra, el discurso, así como la visualización o conceptualización de ideas.

Social: Las personas con interés por las actividades "sociales" prefieren estar rodeados de gente, comunicándose o ayudando en la medida de lo posible. Les gusta conocer a gente nueva, así como cooperar o sentir que son miembros de un grupo. Asimismo, les gusta llegar a los demás, escuchan con atención y son sensibles a las necesidades de los demás. Este factor se subdivide en S1 y S2 como se describe a continuación.

(S1) Dedicación a los demás: Deseo de ayudar, escuchar, dar información, educar, aconsejar y ayudar a la comunidad.

(S2) Relaciones personales: Relacionarse y comunicarse con los demás. En su relación con clientes o compañeros de trabajo gusta de dar explicaciones o consejo.

Personalidad: Paciente, comprensivo y atento; amable, afectuoso y cordial.

Valores: Altruismo, cooperación y generosidad.

Aptitudes potenciales: Empatía, escucha y habilidades de comunicación.

Estilo de aprendizaje: Aprendizaje mediante la relación con los demás y el trabajo en equipo.

Entorno: Preferencia por entornos que faciliten el contacto, la colaboración y la comunicación con otras personas.

Emprendedor/a: Las personas con interés en actividades "emprendedoras" suelen disfrutar de actividades que impliquen algún riesgo y la influencia de otras personas. Pueden disfrutar guiando, coordinando o persuadiendo a otros de sus ideas o realizando ventas. Disfrutarán tomando la iniciativa y comunicándose con los demás. Este factor se subdivide en E1 y E2 como se explica a continuación.

(E1) Emprendedor: Inclínación hacia el desarrollo de nuevas actividades, materiales, ideas o proyectos, asumir riesgos y manejar situaciones difíciles.

(E2) Liderazgo: Deseo de animar, motivar y coordinar o guiar a otras personas o a equipos.

Personalidad: Abiertos/as, dinámicos/as, asertivos/as y emprendedores/as.

Valores: Asunción de riesgos, búsqueda de prestigio y competitividad.

Aptitudes potenciales: Toma de iniciativas, liderazgo y motivación de los demás.

Estilo de aprendizaje: Aprendizaje mediante la acción y la participación.

Entorno: Cómodo en entornos competitivos, donde pueden asumir riesgos, ponerse a prueba y avanzar rápidamente.

Convencional: El término “convencional” corresponde a un trabajo organizado y meticuloso, la aplicación de normas y la planificación. A las personas de tipo “convencional” les gusta acatar reglas precisas, ser organizadas, eficaces y respetar la autoridad. No se desaniman con las tareas rutinarias. Este factor se divide en C1 y C2, tal como se indica a continuación:

(C1) Metódico/a: Hace referencia a la preferencia por organizar y planificar tareas, así como el deseo de controlar la calidad y mantener estructura en el ámbito laboral.

(C2) Interés por los datos y los números: Preferencia para trabajar y razonar con números. Tienen facilidad, por ejemplo, en trabajos de contabilidad e informática.

Personalidad: Las personas de tipo “convencional” tienden a ser meticulosas, perfeccionistas, conformistas y algo introvertidas.

Valores: Exactitud, estabilidad, eficacia.

Aptitudes potenciales: Análisis de datos, meticulosidad, capacidad para manejar números.

Estilo de aprendizaje: Aprenden siguiendo las normas, instrucciones y procedimientos establecidos.

Entorno: Cómodos en ambientes estructurados y jerarquizados donde las normas y los procedimientos están claramente definidos.

2.2.4. Perfiles combinados

Hasta la fecha, la mayoría de las herramientas RIASEC determinan perfiles de intereses ocupacionales identificando dos o tres intereses principales a partir de los 6 factores RIASEC y combinándolos entre sí para identificar qué ámbito o actividad profesional puede interesar al individuo. Holland señaló la importancia de utilizar combinaciones de factores, y así lo explica en el manual de su prueba, que es auto aplicada SDS (1970). Holland utilizaba una combinación de tres factores, en la cual el usuario/a debía tener en cuenta todas las variantes posibles de la combinación de estos tres factores. En consecuencia, el perfil de los intereses ocupacionales no debe inferirse partiendo únicamente de su grado de interés, por un único factor de RIASEC.

La combinación de factores RIASEC para elaborar el perfil de intereses ocupacionales de la persona viene avalada por diferentes estudios que han demostrado la superposición de los seis factores RIASEC con la personalidad. Por ejemplo, tanto Barrick et al (2003), como Larson, Rottinghaus y Borgen (2002) descubrieron que existía una estrecha relación entre extraversión y los factores emprendedor y social del modelo de Holland. Estos estudios muestran que la personalidad es compleja y es normal que las personas tengan diferentes intereses profesionales, por lo tanto, es importante que los perfiles incluyan más de un factor de RIASEC.

En consecuencia, los informes de **VOCATION** proporcionan el perfil de intereses ocupacionales donde se combinan dos intereses en que se obtiene mayor puntuación y que se consideran relevantes para la persona en igual medida, junto con el tercer interés en que se obtiene mayor puntuación y que se considera de menor relevancia para la persona. Como complemento, se presenta una lista de posibles profesiones. La utilización de una combinación de los tres factores/ intereses principales permite definir de forma más precisa y concreta las ocupaciones que pueden encajar mejor con la persona. Además, tener en cuenta un tercer interés elevado, pero con un nivel de relevancia ligeramente inferior permite proporcionar una visión más completa del perfil ocupacional, manteniendo la flexibilidad de que la persona se centre solo en los dos intereses de mayor relevancia, dejando de lado el tercero, en el caso de que no se sienta identificación con ese tercer factor/interés.

2.2.5. Formato de respuesta ipsativa

Los ítems en **VOCATION** se combinan en parejas y se presentan en un formato de elección forzada. Los ítems ipsativos o de elección forzada requieren que las personas que responden el test realicen juicios comparativos del tipo “más parecido a mí” o “menos parecido a mí”. Este formato permite superar muchas de las limitaciones asociadas con los ítems de la escala Likert (Brown & Maydeu-Olivares, 2011a; Chan, 2003; Saville & Willson, 1991; ten Holt, van Duijn, & Boomsma, 2010).

A continuación, enumeramos las ventajas de utilizar los ítems de elección forzada:

- Evitar sesgos de respuesta como el sesgo de afirmación (estar de acuerdo con todo)

- Limitar la posibilidad de distorsión de las respuestas y la deseabilidad social
- Comparar las puntuaciones obtenidas por una misma persona en los intereses de una forma más precisa.

Los ítems de elección forzada presentan una serie de ventajas respecto a otro tipo de escalas, pero también presentan limitaciones respecto al tipo de información que proporcionan. Los datos obtenidos son “datos ipsativos” lo cual significa que la suma de las puntuaciones se mantiene constante. La Teoría clásica de los test se consideraba inadecuada para interpretar y validar las respuestas con datos ipsativos (Cornwell & Dunlap, 1994; Loo, 1999; Meade, 2004), por las siguientes razones:

- Dificultan la comparación entre personas ya que las puntuaciones son relativas, no absolutas, por ejemplo, el candidato no obtiene un nueve en “Realista” sino que su mayor interés es “Realista”.
- Las puntuaciones tienen una correlación lineal que distorsiona la validez de constructo de la escala.
- La falta de varianza distorsiona las correlaciones con criterios externos.
- No se ajusta a los supuestos necesarios para las estimaciones tradicionales de fiabilidad.

Estas limitaciones no presentan un problema real en un test como **VOCATION**, debido al objeto de la evaluación, que son los intereses profesionales. Habitualmente, el motivo para realizar esta evaluación es proporcionar orientación profesional o para comprobar que puestos se ajustan mejor con el perfil de la persona.

En cualquiera caso, en la actualidad la comprensión de las pruebas ipsativas ha mejorado y se ha avanzado en la investigación, por lo que actualmente es posible superar algunas de las limitaciones, como se explica en detalle en el apartado siguiente.

2.2.6. Teoría de respuesta a los ítems de Thurstone (TRI) y las respuestas ipsativas

Los test psicométricos utilizan diferentes tipos de formatos de respuesta para obtener la información sobre la persona. Hay formatos en que la persona ha de valorar hasta qué punto está de acuerdo o no con un enunciado, como la escala Likert, este formato es susceptible a los sesgos de respuesta, como puede ser la aquiescencia, la deseabilidad social o respuestas extremas (Wetzel, Böhnke & Brown, 2016). Para evitar estos sesgos se puede utilizar un formato de respuesta alternativo, el formato ipsativo o de respuesta obligatoria, este formato es menos susceptible a los sesgos de respuesta (Brown & Bartram, 2013; Hontangas et al., 2015; Paulhus, 1991; Saville & Willson, 1991). Los test con formato de respuesta obligatoria están especialmente indicados en situaciones en que las personas tienen una gran motivación para intentar sesgar las respuestas, como puede ser el contexto de la selección de personal (Christiansen, Burns, & Montgomery, 2005; Stewart, Darnold, Zimmerman, Parks, & Dustin, 2010). Sin embargo, el formato de respuesta obligatoria tiene la desventaja de que proporciona una puntuación que representa la preferencia por una característica u otra más que el nivel de la persona en la característica, esto es debido a que la puntuación en una característica está ligada a la

puntuación en las demás. Por lo tanto, las puntuaciones de las personas en este tipo de tests no se pueden comparar con las de otras personas (Cattell, 1944; Hicks, 1970) y en el caso de realizarlas hay distorsiones y la fiabilidad y validez se reducen.

Para solucionar este tipo de problemas, se introdujo un método de calcular las puntuaciones para los formatos de respuesta obligatoria: La teoría de respuesta al ítem de Thurstone (TRI) (Brown & Maydeu-Olivares, 2011). Al aplicar este modelo se obtienen puntuaciones absolutas para cada factor o característica, lo que permiten la comparación entre las personas que realizan el test. Existen otros modelos de cálculo para solucionar este problema, pero los métodos basados en las técnicas de la teoría de Thurstone son las mejor establecidas y proporcionan una solución más completa (Brown, 2016). Este modelo de cálculo permite diseñar tests con dos o más alternativas de respuesta (ítems) en cada pregunta.

La teoría de respuesta al ítem de Thurstone: En los test de respuesta forzada, las preguntas están compuestas por dos o más ítems entre los que se debe elegir en función de las instrucciones del test. En algunos casos solo es necesario elegir entre dos ítems, como en el caso de **VOCATION**, en otras ocasiones, hay que hacer una ordenación de más de dos ítems según el orden de preferencia. En cualquier caso, lo que se produce es una ordenación y la lógica que se aplica en el cálculo siempre es la misma (Brown & Maydeu-Olivares, 2011).

La lógica subyacente es que las personas están comparando cada ítem con cada uno de los demás, en juicios individuales (Maydeu-Olivares & Böckenholt, 2005; Thurstone, 1931). Se sigue la siguiente fórmula en las comparaciones de dos ítems en un grupo de n ítems: $\tilde{n} = n(n - 1) / 2$. Si y_{jk} es la respuesta dada por una persona ante la pregunta de si prefiere el ítem j más que el ítem k en un par de ítems. Si se elige el ítem j , tendremos que $y_{jk} = 1$. En el caso de que sea k el ítem elegido, tendremos que $y_{jk} = 0$. Para aplicar un modelo a estas respuestas, la teoría de respuesta al ítem de Thurstone se apoya en la "Ley del juicio comparativo" (Thurstone, 1927). Esta ley implica que cada ítem j se refiere a una utilidad, que podemos denominar t_j . Esta t_j es la variable de valor psicológico o índice de deseabilidad del ítem (Maydeu-Olivares & Böckenholt, 2005). Se prefiere el ítem j en vez del ítem k si la utilidad de j es mayor que la de k , por ejemplo: si $t_j \geq t_k$. Se asume que cada utilidad latente de t_j es una función lineal de la característica medida latente η_j . Hay que tener en cuenta que se hace la asunción de que cada ítem mide una sola característica, esta asunción también se hace en la validez del test para calcular el ajuste del modelo.

Se han comparado los modelos de TRI con los resultados de otro tipo de escalas de valoración y otros tipos de formatos de respuesta forzada. Las puntuaciones obtenidas en formatos de respuesta forzada basados en modelos TRI correlacionan con los resultados de otras escalas de valoración en gran medida (una media de 0.76 en Guenole et al., 2018; y de 0.79 en Lee, Lee, & Stark, 2018). Sin embargo, sí parece que la forma en que los ítems cargan en el factor varía en las dos condiciones (respuesta forzada y escalas de valoración), por lo que se obtienen diferentes medias en la variable (Guenole et al., 2018).

Las puntuaciones obtenidas siguiendo el modelo TRI tienen igual o mejor validez convergente y discriminante que otros procedimientos de puntuación para los ítems de elección forzada (Anguiano-Carrasco et al., 2015; Lee et al., 2018). Si se comparan con ítems en que se debe hacer una valoración (Anguiano-Carrasco et al., 2015; Brown & Maydeu-Olivares, 2013; Lee et al., 2018) encontraron que la puntuación siguiendo TRI tiene valores estimados ligeramente menores de validación. Aunque, la escala de validación utilizada en ambos estudios (Anguiano-Carrasco et al., 2015; Lee et al., 2018) tenía un formato de valoración, en vez de ipsativo y esto puede haber hecho que hubiera un efecto del método con la versión utilizada para la validación. Además, ninguno de estos estudios se realizó en situaciones en que distorsionar las respuestas supusieran una ventaja para los participantes. En este caso, es problemático comparar la validación de los valores estimados de la respuesta forzada si se comparan con puntuaciones obtenidas con un formato diferente (por valoración), cuando el criterio es una puntuación de respuesta forzada. Brown et al. (2017) modelaron respuestas a ítems de valoración no solo de forma convencional, sino en una segunda condición en que se incluía un factor de distorsión adicional. Los ítems de valoración analizados con este método obtuvieron validez convergente similar comparados con los ítems de respuesta forzada basados en TRI. En conclusión, no existen investigaciones sobre la validación en que los ítems de respuesta forzada basados en TRI tengan unos valores estimados de validez más altos que las de las escalas de ítems que se valoran en una escala.

Equivalencia en la preferencia de los ítems: Lógicamente, para que las preguntas con ítems de respuestas forzada funcionen es necesario que todos los ítems de la pregunta sean igual de deseables. Por esta razón, Brown y colaboradores (2017, p. 142) establecieron que: “para evitar distorsiones en las respuestas es importante que la deseabilidad de la característica/s medidas, en cada conjunto de ítems y para cada pregunta, sea similar”. De esta forma, los tests de respuesta forzada solo evitarán la distorsión si se construyen con cuidado de que no exista una respuesta correcta.

Tradicionalmente se utilizan las valoraciones del nivel de deseabilidad de los pretest, para diseñar conjuntos de ítems para cada pregunta, con un mismo nivel de deseabilidad (Christiansen et al., 2005). De esta forma, las combinaciones de ítems en test de respuesta forzada deberían tener en cuenta las valoraciones de deseabilidad:

- a) para la situación en que van a administrarse
- b) realizadas por personas de la misma población que con la que se utilizará el test

Otra característica de los ítems a tener en cuenta es si son positivos o negativos hacia la característica medida y su combinación en el test. Los primeros estudios de simulación muestran que la medición de la característica medida es más precisa si el test está compuesto por conjuntos de ítems que evalúan la característica en preguntas solo del mismo tipo, o positivas hacia el factor o solo negativas en el factor (Brown & Maydeu-Olivares, 2011). Se desaconseja mezclar ambos tipos de ítems en el cuestionario y utilizar ítems positivos hacia la característica por las siguientes razones:

- Ítems negativos pueden aumentar el esfuerzo cognitivo de los participantes
- La inclusión de ítems negativos puede aumentar la varianza metodológica de forma considerable (Dueber, Love, Toland, & Turner, 2018)
- La inclusión de estos ítems negativos hace a los tests de respuesta forzada más vulnerables a las distorsiones de respuesta
- Los test que incluyen tanto ítems positivos como negativos pueden distorsionarse con más facilidad y tendrán menor precisión que las simulaciones.

Desde una perspectiva práctica, la aplicabilidad del modelo depende en gran medida en su habilidad para estimar de forma correcta las puntuaciones de la persona en el factor con ítems solo positivos o solo negativos para el factor. Esta propiedad se ha estudiado con un único estudio de simulación hasta el momento (Brown & Maydeu-Olivares, 2011). Lógicamente, esta primera investigación no fue suficiente para considerar todas las condiciones relevantes en la construcción de un test. Otra limitación de esta primera investigación de Brown y Maydeu-Olivares (2011) fue que su modelo de evaluación solo tenía en cuenta la comparación de los estimadores de la fiabilidad. Además, las implementaciones hasta el momento de los estimadores de los modelos TRI lo hace solo mediante la estadística de frecuencias, incluso si la estadística bayesiana tiene más ventajas, por ejemplo, la posibilidad de incorporar información previa y estimar la distribución conjunta completa de todos los parámetros del modelo (Gelman et al., 2013). Entre otras cosas, esto último permite obtener distribuciones específicas individuales de las puntuaciones en la característica sin tener que hacer ninguna simplificación de las asunciones.

En conclusión, para **VOCATION**, la TRI de Thurstone se utilizó para revisar, reescribir y validar los ítems que miden los 6 factores RIASEC y los 12 subfactores, que se presentan en 60 preguntas de dos opciones de respuesta, y asegurándose que cada pregunta estaba compuesta por ítems de igual nivel de deseabilidad y positivos para el factor.

3. Puntuación

La puntuación bruta es el número total de respuestas elegidas por una persona para cada factor medido. En **VOCATION**, las puntuaciones brutas han de ser transformadas a una escala que permita su interpretación y tener un criterio de comparación entre las personas. Las puntuaciones brutas pueden también baremarse. Los baremos permiten la comparación de la persona con el grupo de referencia.

3.1. Relación entre puntuaciones brutas y baremadas

Para realizar la comparación de las puntuaciones obtenidas por una persona en un test, se necesita un grupo normativo adecuado, en que la puntuación de esta persona cobra sentido. El grupo normativo óptimo es aquel que representa mejor las características de la persona. En la evaluación de muchas variables psicológicas, como los rasgos de personalidad, intereses, motivaciones, valores, actitudes y percepciones, se considera que las puntuaciones de las personas se distribuirán de una forma determinada, conocida como distribución normal, también conocida como campana de Gauss, debido a su forma. Las variables psicológicas se dan en cualquier población, pero pueden aparecer en diferentes niveles. Esto implica que una característica determinada se distribuirá en función de la curva normal para cada población concreta, en la que tendremos una media, alrededor de la cual estarán la mayor parte de las puntuaciones, así como una serie de puntuaciones situadas por encima y por debajo de esta media que se considerarán puntuaciones extremas cuanto más se alejen de la media (Kaplan & Saccuzzo, 2008).

Los grupos normativo siguen la distribución normal y posibilitan la interpretación de la puntuación de la persona. Para que la distribución normal usada en la comparación de las puntuaciones sea relevante, la muestra de personas utilizada debe reflejar las características de la población que se prevea someter a test. Si se desarrolla un test para ingenieros/as, la población normativa debe estar compuesta por ingenieros/as y si el grupo normativo es una población de enfermeros/as es probable que no sea adecuado para seleccionar farmacéuticos/as. Los grupos normativos puede tener muy diferentes tamaños, desde cien a cien mil. Sin embargo, cuanto mayor sea el tamaño de la muestra del grupo de referencia, mayor será la aproximación a la distribución normal de la población real total.

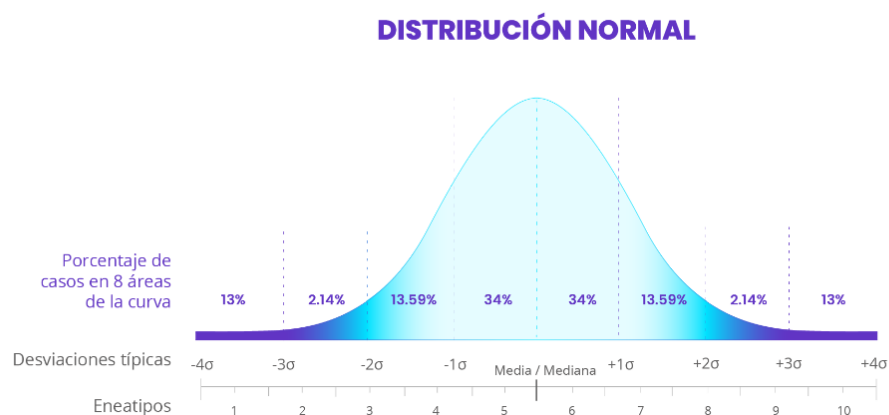


Figura 4. Curva de distribución normal.

El proceso de baremación consiste en convertir las puntuaciones brutas en puntuaciones típicas. Este proceso proporciona una escala de medición que facilita la interpretación de las puntuaciones de los test. En **VOCATION**, los estadísticos aplicados a las puntuaciones fueron: 1) Media, 2) Mediana, 3) Desviación típica, 4) Puntuación típica, 5) transformación a eneatis y 6) Error típico de medida. La ilustración 4 describe varios tipos de puntuaciones derivadas utilizados para la validación de **VOCATION**.

Tabla 4. Estadísticos y transformaciones para la baremación de VOCATION.

Estadístico	Símbolo	Definición y características
Media*	μ	La media se calcula dividiendo la suma de todas las puntuaciones entre el número de participantes. Su fórmula es: $\mu = (\sum x) / N,$ $\sum X$ es el sumatorio de todas las puntuaciones. N es el total de observaciones
Mediana*	$\tilde{x}$	La mediana se obtiene ordenando todas las puntuaciones desde la más baja hasta la más alta y seleccionando la puntuación que se encuentra justo a la mitad.

Desviación típica	σ	Mide la variación entre las puntuaciones de los participantes. Si es pequeña, indica que las puntuaciones están próximas a la media y hay poca variación; si es elevada indica que hay grandes diferencias entre las puntuaciones. La fórmula es: $\sigma = \sqrt{[\sum (X_i - \mu)^2 / N]}$ μ es la media X_i es cada uno de los elementos de la muestra N es el total de observaciones
Puntuación eneatiípica	Eneatipo	Puntuación estandarizada con puntuaciones que varían de 1 a 10. La media es 5.5 y la desviación típica es 2. La puntuación obtenida se redondea el número entero más cercano La fórmula de las puntuaciones eneatiípicas es: $\text{Eneatipo} = (\chi - \mu / \sigma)^2 + 5.5$ X es la puntuación de la persona μ es la media σ es la desviación típica de la muestra $2+5.5$ sería la regla de conversión a eneatiipos
Error típico de medida	SE	Refleja la fiabilidad de un test, indica la dispersión de los errores de medida asociados con las puntuaciones del test obtenidas en la muestra. Identifica el nivel de error mediante la estimación de la puntuación verdadera a partir de la observada. A menor error, más precisa es la evaluación. Su fórmula es: $SE = \sigma \sqrt{1 - \text{Fiabilidad}}$

*La media y la mediana son medidas de la tendencia central de un grupo de puntuaciones.

3.2. Conversión de las puntuaciones directas a puntuaciones típicas

Para cada uno de los ítems de **VOCATION**, la escala de puntuaciones máxima y mínima en cada factor es de 0 a 10. La conversión de las puntuaciones directas se realiza en un proceso de cuatro pasos:

Paso 1: Las puntuaciones ipsativas se calculan para cada escala sumando todas las respuestas de elección forzada de cada factor. En cada respuesta, marcamos con 1 la respuesta elegida y 0 la rechazada.

Paso 2: Las puntuaciones se transforman a una escala de 0 a 10 con esta fórmula:

$$puntuación = \left(\frac{puntuación\ bruta - (min)}{(max) - (min)} \right) \times 10$$

donde:

max es puntuación máxima posible de cada factor

min es puntuación mínima posible de cada factor

Paso 3: Para facilitar las comparaciones, las puntuaciones de cada escala se bareman, se calcula la media y la desviación típica del grupo normativo. A continuación, se obtiene la puntuación típica para cada persona mediante la siguiente fórmula:

$$z = (puntuación - (\mu/x))$$

Paso 4: Las puntuaciones típicas se convierten en eneatis del 1 al 10 utilizando la fórmula:

$$Eneatis = z * 2 + 5.5$$

3.3. Error típico de medida

Las puntuaciones de cualquier test no son perfectas; siempre existe una cierta distancia entre la puntuación obtenida y la verdadera. De ahí que sea necesario estimar la cantidad de error de cada puntuación. La cantidad de error asociada con una puntuación recibe el nombre de error típico de medida (SE). SE nos ayuda a entender cuál es la verdadera puntuación de un candidato/a y en qué medida es diferente de la puntuación obtenida (Harvill, 1991). El procedimiento ideal para calcular SE es pasar el test en múltiples ocasiones a una misma persona durante un periodo de tiempo largo. La variación entre las puntuaciones obtenidas en cada ocasión sería el error típico de medida. Sin embargo, este procedimiento ideal de cálculo de SE es inviable y es necesario utilizar una fórmula para calcularlo.

El error típico de medida o SE proporciona una estimación que se suma o resta a la puntuación obtenida, estableciendo el límite por encima y por debajo de la puntuación obtenida. El conjunto de puntuaciones comprendidas entre estos límites se conoce como intervalo de confianza, que indica los márgenes entre los que se encuentra la puntuación verdadera.

El error típico de medida y la fiabilidad tienen una relación inversa, si el SE aumenta la fiabilidad disminuye y viceversa. Por ejemplo, para una serie de puntuaciones

determinadas, el 68% de las veces, se obtendrá una puntuación cuyo intervalo de confianza estará entre +1.0 y -1.0 de la puntuación verdadera. En el 95% de los casos, la puntuación observada variará entre +1.96 y -1.96 y en el 99% de las veces, la puntuación verdadera estará comprendida entre +2.58 y -2.58. El intervalo de confianza del 95% se calcula utilizando la fórmula:

$$\text{Intervalo de confianza} = \text{puntuación} \pm 1.96 \times SE$$

Por lo tanto, en **VOCATION**, un factor/interés en el que se obtenga una puntuación de 5 y tenga un error típico de medida (SE) de 0,2 tiene un intervalo de confianza entre 4.608 y 5.392. Esto significa que existe una probabilidad del 95% de que la puntuación verdadera esté comprendida en este intervalo

3.4. Características del grupo normativo adecuado

El grupo de comparación o grupo normativo da significado a la puntuación de una persona. La interpretación de una misma puntuación en un test puede ser diferente en función de si fue obtenida por una persona con estudios superiores, un administrador/a o una persona en un cargo directivo.

La interpretación sólo será pertinente si se compara a la persona con su grupo de referencia. Los parámetros para definir qué población es adecuada para una persona varían, aunque son habituales el país de procedencia, la edad, el género, el nivel de educación, el número de años de experiencia profesional, el estatus socio profesional y/o el grupo profesional o industrial más apropiado para el candidato/a. Existe la posibilidad de generar los parámetros de la población a posteriori, tomando como referencia el grupo de personas que se ha evaluado para compararlos entre sí.

3.5. Generación de parámetros para una población concreta

Es posible generar parámetros estadísticos para una población determinada si se cuenta con una muestra de por lo menos 50 personas. La generación de los parámetros permite que las comparaciones entre personas y su grupo de referencia reflejen adecuadamente lo que se quiere evaluar. La generación de los parámetros está indicada cuando no se dispone de población de referencia; las funciones a desempeñar son nuevas; la organización tiene una función empresarial poco habitual/ basada en el cliente y en casos en que el número de candidatos/as por puesto es mucho mayor o menor de lo esperado.

La generación de parámetros estadísticos se hace calculando la media y la desviación típica para cada factor y se dan puntuaciones a los resultados siguiendo los pasos explicados anteriormente. Esto permite a la organización comparar a las personas con un grupo normativo específico y relevante.

3.6. Puntuación en los informes

Los informes de **VOCATION** reflejan el proceso de puntuación y cada persona obtiene una puntuación de 0 a 10. El informe ofrece una descripción detallada de cada uno de los aspectos de los seis factores RIASEC del test. Se escoge aquel nivel que se corresponde con la puntuación del candidato/a de entre seis diferentes posibilidades y en él se detallan sus preferencias y las actividades u ocupaciones por las que sentirá mayor inclinación. Las puntuaciones elevadas se describirán con mayor intensidad y las puntuaciones bajas se describirán con menor intensidad. Por ejemplo, un candidato/a con una puntuación baja en un factor determinado puede describirse como “no parece que muestre mucho interés”. Por el contrario, una puntuación elevada, indicará “un marcado interés” en el factor relevante. Las puntuaciones deben interpretarse en términos relativos, el orden de los perfiles escogidos es más importante que la puntuación absoluta.

La combinación de perfiles de los informes de **VOCATION** muestra el proceso de puntuación. Los perfiles representan la combinación de los tres factores representados por las puntuaciones más altas. Se calculan considerando que las puntuaciones de los primeros dos intereses serán las más elevadas y de igual importancia, mientras que la puntuación obtenida en el tercer interés tendrá una relevancia ligeramente inferior.

4. Descripción de la muestra

La muestra está compuesta por N=564 personas de ambos géneros, todos países de habla hispana, edades entre 16 y 70 años y diferentes niveles profesionales. Los datos se recogieron entre diciembre de 2019 y noviembre de 2020 y provienen del centro de investigación Testing Centre de *Central Test*, clientes y la plataforma de estudios de mercado Amazon Turk. Las personas pertenecientes a la plataforma Amazon Turk recibieron una pequeña retribución por su participación.

Para garantizar la calidad de los datos se eliminaron las respuestas poco fiables utilizando criterios como el tiempo de realización del test o evidencias de que se contestó al azar. Para las personas de Amazon Turk se comprobó su nivel de español mediante preguntas de control ya que la muestra contenía personas que viven en países cuya lengua principal no es el español.

Tabla 5. País de origen.

País de origen*	(%)	N
Argentina	3.19	18
Colombia	29.43	166
Costa Rica	2.48	14
Ecuador	3.19	18
España	15.78	89
Estados Unidos	21.45	121
Honduras	6.91	39
México	5.50	31
Perú	1.77	10
Venezuela	2.66	15

*Otros países con participación menor a 10: Belice, Bolivia, Brasil, Chile, República Dominicana, El Salvador, Francia, Alemania, Guatemala, Nicaragua, Panamá, Puerto Rico, Uruguay

Tabla 6. Género.

Género	(%)	N
Femenino	53.19	300
Masculino	46.81	264

Tabla 7. Edad.

Grupos de edad	(%)	N
-25	46.45	262
26-33	18.62	105
34-41	15.60	88
42-49	9.93	56
50-56	4.96	28
+57	4.43	25

Tabla 8. Nivel de estudios.

Nivel de estudios	N
Doctorado / Master	66
Licenciatura / Grado	104
Diplomatura	63
Bachiller	36
Formación profesional	32
Estudiando	201
Sin datos	62

5. Baremos

Los baremos nos permiten comparar la puntuación en un test de una persona con la de las demás personas que han realizado la prueba. El grupo de personas que realiza el test son una muestra representativa del grupo de población para el que se ha desarrollado la prueba y se llama “grupo normativo”. Cuanto más se parezcan las características de la muestra con las de la población diana, más se podrá confiar en el valor de la distribución obtenida en función de sus respuestas como guía para clasificar las respuestas. Los parámetros estadísticos permiten a las personas que desarrollan test y los psicólogos/as hacer comparaciones entre las personas y así poder tomar decisiones basadas en información objetiva. Los baremos nos permiten dotar de significado a las puntuaciones de los test y valorar el desempeño, el nivel en un factor, interés, rasgo en relación con el resto de la población.

Las siguientes tablas muestran la media, la desviación típica, las puntuaciones máximas y mínimas para la población de habla hispana en los factores de **VOCATION**, para la muestra total y está organizada en función de la edad y el género.

Tabla 9. Baremos por factor para toda la muestra.

Factor	Media	Mediana	Desviac Típica	Punt mínima	Punt máxima	Intervalo	SE
R	3.95	3.5	2.27	0	10	10	0.10
I	5.15	5	2.36	0	10	10	0.10
A	5.20	5.5	2.23	0	10	10	0.09
S	4.92	5	2.42	0	10	10	0.10
E	5.47	5.5	2.21	0.5	10	9.5	0.09
C	5.30	5.5	2.27	0	10	10	0.10

R=Realista, I=Investigador, A=Artístico, S=Social, E=Emprendedor, C=Convencional

Tabla 10. Baremos por subfactor para toda la muestra.

Sub- factor	Media	Mediana	Desviac Típica	Punt mínima	Punt máxima	Intervalo	SE
R1	4.17	4	2.67	0	10	10	0.11
R2	3.73	3	2.45	0	10	10	0.10

I1	5.93	6	2.23	0	10	10	0.09
I2	5.75	6	2.59	0	10	10	0.11
A1	4.91	5	2.75	0	10	10	0.12
A2	5.26	5	2.59	0	10	10	0.11
S1	4.93	5	2.81	0	10	10	0.12
S2	4.92	5	2.45	0	10	10	0.10
E1	5.68	6	2.21	0	10	10	0.10
E2	5.64	6	2.59	0	10	10	0.11
C1	4.96	5	2.27	0	10	10	0.10
C2	4.11	4	2.71	0	10	10	0.11

R1= Actividades físicas y al exterior, R2= Intereses técnicos y manuales, I1=curiosidad y aprendizaje intelectual, I2= Ciencia y tecnología, A1=Sentido y expresión estética, A2= Creatividad y diseño, S1=Dedicación a los demás, S2= relaciones sociales, E1= emprendedor, E2= Liderazgo, C1= Metódico, C2= Interés por los datos y los números.

Tabla 11. Baremos por factor y género de la muestra.

Factor	Media	Mediana	Desviac Típica	Punt mínima	Punt máxima	Intervalo	SE
Mujeres							
R	3.58	3.5	2.09	0	9	9	0.12
I	4.98	5	2.24	0	10	10	0.13
A	5.46	5.5	2.24	0.5	10	9.5	0.13
S	5.47	5.5	2.29	0.5	10	9.5	0.13
E	5.46	5.5	2.13	0.5	10	9.5	0.12
C	5.04	5	2.22	0	10	10	0.13
Hombres							
R	4.38	4	2.40	0	10	10	0.15
I	5.34	5.5	2.48	0	10	10	0.15
A	4.89	5	2.17	0	10	10	0.13
S	4.31	4	2.42	0	10	10	0.15

E	5.48	5.5	2.30	0.5	10	9.5	0.14
C	5.60	5.5	2.31	0.5	10	9.5	0.14

R=Realista, I=Investigador, A=Artístico, S=Social, E=Emprendedor, C=Convencional

Tabla 12. Baremos por subfactor y género de la muestra.

Sub-factor	Media	Mediana	Desviac Típica	Punt mínima	Punt máxima	Intervalo	SE
Mujeres							
R1	3.83	3	2.51	0	10	10	0.15
R2	3.32	3	2.22	0	9	9	0.13
I1	5.69	6	2.21	0	10	10	0.13
I2	5.11	5	2.62	0	10	10	0.15
A1	5.39	5	2.70	0	10	10	0.16
A2	5.51	6	2.56	0	10	10	0.15
S1	5.64	6	2.67	0	10	10	0.15
S2	5.29	5	2.38	0	10	10	0.14
E1	5.73	6	2.10	1	10	9	0.12
E2	5.46	6	2.56	0	10	10	0.15
C1	5.23	5	2.28	0	10	10	0.13
C2	3.78	4	2.70	0	10	10	0.16

Hombres

R1	4.56	4	2.79	0	10	10	0.17
R2	4.19	4	2.61	0	10	10	0.16
I1	6.21	6	2.21	0	10	10	0.14
I2	6.48	7	2.34	0	10	10	0.14
A1	4.37	4	2.72	0	10	10	0.17
A2	4.97	5	2.61	0	10	10	0.16
S1	4.12	4	2.75	0	10	10	0.17
S2	4.49	4.5	2.48	0	10	10	0.15

E1	5.62	6	2.32	0	10	10	0.14
E2	5.85	6	2.60	0	10	10	0.16
C1	4.64	5	2.24	0	10	10	0.14
C2	4.48	4	2.67	0	10	10	0.16

R1= Actividades físicas y al exterior, R2= Intereses técnicos y manuales, I1=curiosidad y aprendizaje intelectual, I2= Ciencia y tecnología, A1=Sentido y expresión estética, A2= Creatividad y diseño, S1=Dedicación a los demás, S2= relaciones sociales, E1= emprendedor, E2= Liderazgo, C1= Metódico, C2= Interés por los datos y los números.

Tabla 13. Baremos por factor y grupos de edad.

Factor	Media	Mediana	Desviac Típica	Punt mínima	Punt máxima	Intervalo	SE
-25							
R	3.99	3.5	2.20	0	10	10	0.14
I	5.23	5	2.34	0	10	10	0.14
A	5.28	5.5	2.29	0	10	10	0.14
S	5.16	5.5	2.25	0	9.5	9.5	0.14
E	5.59	5.5	2.06	1	9.5	8.5	0.13
C	4.74	4.5	2.10	0	10	10	0.13
26-33							
R	4.24	4	2.28	0	9.5	9.5	0.22
I	5.53	5.5	2.52	1	10	9	0.25
A	4.83	5	2.11	0.5	9.5	9	0.21
S	4.20	4	2.41	0	10	10	0.23
E	5.05	5	2.08	0.5	10	9.5	0.20
C	6.14	6	2.06	1.5	10	8.5	0.20
34-41							
R	4.16	4	2.29	0.5	10	9.5	0.24
I	4.57	4.5	2.5	0	9.5	9.5	0.27
A	5.32	5.5	2.31	0.5	9.5	9	0.25

S	4.57	4	2.70	0	10	10	0.29
E	5.43	5.5	2.28	0.5	10	9.5	0.24
C	5.95	6.5	2.36	0.5	10	9.5	0.25
42-49							
R	3.92	3.50	2.36	0	9.5	9.5	0.32
I	5.05	5.25	2.16	1	8.5	7.5	0.29
A	5.13	5	2.13	1	10	9	0.28
S	5.34	6	2.36	0	9.5	9.5	0.32
E	5.13	5	2.55	0.5	9.5	9	0.34
C	5.42	5.5	2.41	1	10	9	0.32
50-56							
R	3.04	2	2.43	0	10	10	0.46
I	4.84	4.75	1.93	1.5	8.5	7	0.36
A	5.14	5.25	2.02	1	8.5	7.5	0.38
S	5.7	6	2.89	1	10	9	0.55
E	6.07	5.75	2.58	1	10	9	0.49
C	5.21	5.25	2.41	1	10	9	0.46
+57							
R	2.68	2	2.12	0	8.5	8.5	0.42
I	5.34	5.5	2.04	1	8	7	0.41
A	5.62	5.5	2.24	1.5	10	8.5	0.45
S	4.96	5.5	2.15	1	8	7	0.43
E	6.18	6	2.49	0.5	10	9.5	0.5
C	5.22	5	2.63	0.5	1	9.5	0.53

R=Realista, I=Investigador, A=Artístico, S=Social, E=Emprendedor, C=Convencional

Tabla 14. Baremos por subfactor y grupos de edad.

Sub-factor	Media	Mediana	Desviac Típica	Punt mínima	Punt máxima	Intervalo	SE
-25							
R1	4.25	4	2.65	0	10	0	0.16
R2	3.73	3	2.31	0	10	10	0.14
I1	5.68	6	2.18	0	10	10	0.13
I2	5.31	5	2.63	0	10	10	0.16
A1	5.24	5	2.79	0	10	10	0.17
A2	5.31	5.5	2.68	0	10	10	0.17
S1	5.12	5.	2.72	0	10	10	0.17
S2	5.19	5	2.23	0	10	10	0.14
E1	5.66	6	2.12	1	10	9	0.13
E2	5.57	6	2.37	0	10	10	0.15
C1	5.16	5	2.17	0	10	10	0.13
C2	3.74	3	2.81	0	10	10	0.17
26-33							
R1	4.66	5	2.69	0	10	10	0.26
R2	3.83	4	2.46	0	9	9	0.24
I1	6.14	6	2.32	1	10	10	0.23
I2	6.32	6	2.37	2	10	8	0.23
A1	5.24	5	2.85	0	10	10	0.28
A2	5.52	6	2.66	0	10	10	0.26
S1	4.28	4	2.68	0	10	10	0.26
S2	4.13	4	2.52	0	10	10	0.25
E1	5.24	5	2.15	0	9	9	0.21
E2	5.10	6	2.61	0	10	10	0.26
C1	4.90	5	2.21	0	10	10	0.22
C2	4.61	4	2.40	0	10	10	0.23

34-41

R1	4.17	4	2.47	0	10	10	0.26
R2	4.15	4	2.70	0	10	10	0.29
I1	6.15	6	2.36	0	10	10	0.25
I2	6.08	6	2.64	1	10	9	0.28
A1	4.18	4	2.81	0	10	10	0.3
A2	4.84	5	2.62	0	10	10	0.28
S1	4.53	4	2.99	0	10	10	0.32
S2	4.60	5	2.7	0	10	10	0.29
E1	5.48	5.5	2.24	0	9	9	0.24
E2	5.70	6	2.8	0	10	10	0.3
C1	5.23	5	2.39	0	10	10	0.25
C2	4.88	5	2.82	0	10	10	0.30

42-49

R1	4.07	4	2.8	0	10	10	0.37
R2	3.77	3	2.53	0	10	10	0.34
I1	6.2	6	2.15	1	10	9	0.29
I2	5.59	6	2.68	0	10	10	0.36
A1	4.59	4.5	2.51	1	9	8	0.33
A2	5.20	5	2.45	1	10	9	0.33
S1	5.45	6	2.7	0	10	10	0.36
S2	5.23	6	2.56	0	10	10	0.34
E1	5.91	6	2.31	1	10	9	0.31
E2	6.07	6	2.60	1	10	9	0.35
C1	4.29	4	2.6	0	9	9	0.35
C2	3.64	4	2.38	0	8	8	0.32

50-56

R1	3.18	3	2.31	0	10	10	0.44
R2	2.89	2	2.82	0	10	10	0.53

I1	6.14	6	1.88	3	10	7	0.36
I2	6.39	6	1.91	2	10	8	0.36
A1	3.64	3.5	1.91	1	8	7	0.36
A2	4.64	4	2	2	9	7	0.38
S1	5.64	5.5	3.4	0	10	10	0.64
S2	5.75	6	2.8	0	10	10	0.53
E1	6.71	7	2.05	2	10	8	0.55
E2	6.57	7	2.89	0	10	10	0.55
C1	4.36	4	2.3	1	10	9	0.43
C2	4.07	4	2.76	0	10	10	0.52

+57

R1	2.68	2	2.88	0	10	10	0.58
R2	2.68	2	1.82	0	8	8	0.36
I1	6.12	6	2.24	2	9	7	0.45
I2	6.36	7	2.58	1	10	9	0.52
A1	4.72	4	2.25	1	9	8	0.45
A2	5.76	6	1.98	0	9	9	0.4
S1	5.12	5	2.79	1	10	10	0.56
S2	4.8	5	2.14	1	8	7	0.43
E1	6.76	7	2.62	2	10	8	0.52
E2	6.48	7	3.02	1	10	10	0.6
C1	4.28	4	2.11	1	9	8	0.42
C2	4.24	5	2.42	0	9	9	0.48

R1= Actividades físicas y al exterior, R2= Intereses técnicos y manuales, I1=curiosidad y aprendizaje intelectual, I2= Ciencia y tecnología, A1=Sentido y expresión estética, A2= Creatividad y diseño, S1=Dedicación a los demás, S2= relaciones sociales, E1= emprendedor, E2= Liderazgo, C1= Metódico, C2= Interés por los datos y los números.

Tabla 15. Baremos por factor y nivel educativo.

Factor	Media	Desv. Típica	Mediana	Punt mínima	Punt máxima	Intervalo	SE
<i>Doctorado / Master</i>							
R	3.17	2.05	3	0	9.5	9.5	0.25
I	5.13	2.09	5.5	1	9	8	0.26
A	5.42	2.17	5.5	1.5	10	8.5	0.27
S	4.92	2.58	5	0	10	10	0.32
E	5.62	2.49	5.5	0.5	10	9.5	0.31
C	5.73	2.25	6	0.5	9.5	9	0.28
<i>Licenciatura / Grado</i>							
R	3.74	2.34	3.5	0	10	10	0.23
I	4.96	2.62	4.5	0	10	10	0.26
A	5.51	2.09	5.5	0.5	10	9.5	0.2
S	4.68	2.6	4	0	10	10	0.25
E	5.03	2.3	5	0.5	10	9.5	0.23
C	6.08	2.32	6	0.5	10	9.5	0.2
<i>Diplomatura</i>							
R	3.93	2.54	3.5	0	10	10	0.32
I	5.11	2.19	5	1	10	9	0.28
A	4.79	1.95	5	0.5	9	8.5	0.25
S	5.02	2.32	5	0.5	9	8.5	0.29
E	5.45	2.38	5	0.5	10	9.5	0.3
C	5.7	2.46	6	1	10	9	0.31
<i>Bachiller</i>							
R	4.94	2.27	5	1	9	8	0.38
I	4.46	2.68	4.5	0	9	9	0.45
A	4.74	2.06	4.5	0.5	9.5	9	0.34
S	5.21	2.58	5.5	0	9.5	9.5	0.43

E	4.94	2.04	4.75	2	9	7	0.34
C	5.69	2.18	5.5	1	10	9	0.36

Formación Profesional

R	3.92	2.43	3.5	0	10	10	0.43
I	5.5	2.59	5.5	1	10	9	0.46
A	5.38	2.04	5.5	0.5	9	8.5	0.36
S	4.44	2.76	3.75	0.5	10	9.5	0.49
E	4.97	2.24	5	1	10	9	0.4
C	5.8	1.98	5.75	2.5	9	6.5	0.35

Estudiando

R	3.99	2.17	4	0	10	10	0.15
I	5.24	2.25	5	0.5	10	9.5	0.16
A	5.26	2.44	5.5	0	10	10	0.17
S	5.15	2.24	5	0	9.5	9.5	0.16
E	5.7	1.99	5.5	1	10	9	0.14
C	4.65	2.04	4.5	0	10	10	0.14

R=Realista, I=Investigador, A=Artístico, S=Social, E=Emprendedor, C=Convencional

Tabla 16. Baremos por subfactor y nivel educativo.

Sub-factor	Media	Desv. Típica	Mediana	Punt mínima	Punt máxima	Intervalo	SE
<i>Doctorado / Master</i>							
R1	3.24	2.31	3	0	10	10	0.28
R2	3.09	2.4	3	0	10	10	0.3
I1	6.39	2.13	7	1	10	9	0.26
I2	6.17	2.32	6	0	10	10	0.29
A1	4.32	2.43	4	0	10	10	0.3
A2	5.23	2.08	5	1	10	9	0.26

S1	4.91	2.99	4	0	10	10	0.37
S2	4.94	2.61	5	0	10	10	0.32
E1	6.26	2.23	6	1	10	9	0.28
E2	6.61	2.28	7	0	10	10	0.33
C1	4.52	2.25	4	1	10	9	0.28
C2	4.32	2.59	4	0	10	10	0.32

Licenciatura / Grado

R1	3.94	2.64	3	0	10	10	0.26
R2	3.54	2.57	3	0	10	10	0.25
I1	6.33	2.17	7	0	10	10	0.21
I2	6.25	2.61	6	1	10	9	0.26
A1	4.79	2.98	5	0	10	10	0.29
A2	5.36	2.75	6	0	10	10	0.27
S1	4.75	2.92	5	0	10	10	0.29
S2	4.62	2.6	5	0	10	10	0.25
E1	5.77	2.25	6	0	10	10	0.22
E2	5.069	2.52	6	0	10	10	0.25
C1	4.71	2.36	4	1	10	9	0.23
C2	4.26	2.63	4.5	0	10	10	0.26

Diplomatura

R1	4.02	2.97	3	0	10	10	0.37
R2	3.84	2.62	4	0	10	10	0.33
I1	5.94	2.33	6	1	10	9	0.29
I2	6.08	2.66	7	1	10	9	0.34
A1	4.81	2.44	5	0	10	10	0.31
A2	5.16	2.4	5	0	10	10	0.3
S1	5.38	2.7	6	0	10	10	0.34
S2	4.67	2.42	5	1	10	9	0.31
E1	5.38	2.21	5	1	10	9	0.28

E2	5	2.79	5	0	10	10	0.35
C1	5.11	2.56	5	0	10	10	0.32
C2	4.62	2.7	5	0	10	10	0.34
<i>Bachiller</i>							
R1	5.36	2.38	5.5	0	9	9	0.4
R2	4.53	2.76	5	0	9	9	0.46
I1	5.83	2.61	6	0	10	10	0.44
I2	5.31	2.67	5	0	10	10	0.45
A1	4.86	2.97	4	0	9	9	0.49
A2	5.03	2.59	5	0	9	9	0.43
S1	5.22	2.68	5	0	9	9	0.45
S2	5.19	2.7	6	0	10	10	0.45
E1	4.86	1.99	5	0	9	9	0.33
E2	4.72	2.57	4.5	0	10	10	0.43
C1	4.92	2.38	5	1	10	9	0.4
C2	4.14	2.53	4	0	9	9	0.42
<i>Formación profesional</i>							
R1	4.03	2.82	4	0	10	10	0.5
R2	3.81	2.44	3	0	10	10	0.43
I1	5.97	2.13	7	2	10	8	0.38
I2	6.25	2.34	6	3	70	7	0.41
A1	5.38	2.78	5	1	10	9	0.49
A2	5.25	2.85	5.5	0	10	10	0.5
S1	4.56	3.11	4	0	10	10	0.55
S2	4.31	2.79	4	0	10	10	0.49
E1	5.5	2.29	6	0	10	10	0.4
E2	5.72	2.53	6	0	10	10	0.45
C1	5	1.93	5	2	8	6	0.34
C2	4.22	2.67	4.5	0	9	9	0.47

Estudiando

R1	4.23	2.68	4	0	10	10	0.19
R2	3.74	2.24	3	0	10	10	0.16
I1	5.62	2.2	6	0	10	10	0.16
I2	5.27	2.55	5	0	10	10	0.18
A1	5.16	2.79	5	0	10	10	0.2
A2	5.29	2.66	5	0	10	10	0.19
S1	5.02	2.76	5	0	10	10	0.19
S2	5.27	2.19	5	0	9	9	0.15
E1	5.78	2.12	6	1	10	9	0.15
E2	5.65	2.44	6	0	10	10	0.17
C1	5.15	2.14	5	0	10	10	0.15
C2	3.78	2.76	3	0	10	10	0.19

R1= Actividades físicas y al exterior, R2= Intereses técnicos y manuales, I1=curiosidad y aprendizaje intelectual, I2= Ciencia y tecnología, A1=Sentido y expresión estética, A2= Creatividad y diseño, S1=Dedicación a los demás, S2= relaciones sociales, E1= emprendedor, E2= Liderazgo, C1= Metódico, C2= Interés por los datos y los números.

6. Fiabilidad

La Fiabilidad indica en qué medida el test está libre de error y sus ítems miden el mismo constructo. Para que un test sea fiable se deben obtener los mismos resultados de forma consistente a lo largo del tiempo. La fiabilidad de un test es importante porque nos permite tener confianza en que las puntuaciones en el test realmente representan el comportamiento habitual, el interés, las actitudes o habilidades de la persona. La primera preocupación cuando se desarrolla un test es en qué medida los errores de medida arbitrarios pueden influir en las respuestas al test. Si un test carece de fiabilidad tampoco será válido, puesto que la precisión de sus resultados depende de la consistencia de los resultados. Una alta fiabilidad en un test previene los errores que se pueden dar en función de los sesgos característicos de los procesos de selección.

6.1. Consistencia interna

La consistencia interna es la medida en que los ítems de una escala miden la misma dimensión o factor. La técnica utilizada en **VOCATION** para evaluar la consistencia interna es el *coeficiente omega*, ya que permite evaluar hasta qué punto convergen entre sí los ítems de una escala. Su definición matemática sería la cantidad de variación de la puntuación en la escala de todos los factores subyacentes. Existen otras formas de evaluar la consistencia interna, por ejemplo, el *alpha de Cronbach*, es una técnica muy utilizada, pero no es el coeficiente más adecuado para analizar la fiabilidad en los test ipsativos. Alpha asume que todos los ítems de la escala son igual de importantes y este supuesto puede hacer que se distorsione la fiabilidad aumentando o disminuyendo el número de ítems de la escala, independientemente de su relevancia real.

En **VOCATION** se utiliza el *coeficiente omega* como alternativa ya que tiene en consideración tanto las cargas factoriales, como el error de medida (Dunn, Baguley, & Brunsden, 2014; Raykov, 1997; Schmitt, 1996). De forma semejante a otros coeficientes para analizar la fiabilidad, la consistencia interna varía entre 0 y 1, y es mayor cuanto más se acerca su valor a 1. Una simplificación de cómo interpretar el coeficiente es que un valor de .70 o superior es lo recomendable (Nunnally, 1978), los valores superiores a 0.9 serían excelentes; entre 0,7 a 0,9 buenos; entre 0,6 y 0,7 aceptables; entre 0,5 a 0,6 bajos y menores de 0,5 inaceptables. El coeficiente omega debería ser $\geq 0,7$.

La fórmula para el coeficiente omega es:

$$\omega = \frac{\left[\sum_{i=1}^i \lambda \right]^2}{\left[\sum_{i=1}^i \lambda \right]^2 + \left[\sum_{i=1}^i 1 - \lambda_i^2 \right]}$$

La fiabilidad de **VOCATION** se calculó con la muestra descrita con anterioridad de 564 personas. La siguiente tabla describe la fiabilidad de **VOCATION**.

Tabla 17. Fiabilidad factores.

Factor	Fiabilidad
R	0.95
I	0.94
A	0.95
S	0.94
E	0.94
C	0.93

R=Realista, I=Investigador, A=Artístico, S=Social, E=Emprendedor, C=Convencional

Todos los factores tienen una fiabilidad excelente, ya que en todos es superior a 0.9, siendo el factor convencional, con 0.93 el que tiene una fiabilidad menor. Además, la media de la consistencia interna de los factores de **VOCATION** es de 0.94 lo que indica que la escala tiene una alta consistencia interna y por lo tanto es fiable.

Tabla 18. Fiabilidad subfactores.

Subfactor	Fiabilidad
R1	0.97
R2	0.96
I1	0.96
I2	0.7
A1	0.95
A2	0.93
S1	0.98
S2	0.97
E1	0.97
E2	0.96
C1	0.74
C2	0.93

R1= Actividades físicas y al exterior, R2= Intereses técnicos y manuales, I1=curiosidad y aprendizaje intelectual, I2= Ciencia y tecnología, A1=Sentido y expresión estética, A2= Creatividad y diseño, S1=Dedicación a los demás, S2= relaciones sociales, E1= emprendedor, E2= Liderazgo, C1=

Metódico, C2= Interés por los datos y los números.

Se comprobó también la fiabilidad de cada uno de los subfactores. Todos tienen una fiabilidad por encima de 0.7. Siendo la más elevada S1 con 0.98 y la menor I2 con 0.7. La media de la fiabilidad de los subfactores es de 0.91. Esto muestra que la fiabilidad es buena para I2 y C1 y excelente para el resto de los subfactores.

6.2. Fiabilidad Test-Retest

La fiabilidad test-retest implica la administración del test dos veces a la misma muestra de personas en condiciones similares y comparar los resultados obtenidos por cada individuo, de esta forma se estima si las puntuaciones del test son estables a lo largo del tiempo. Cuanta mayor semejanza se encuentre entre las puntuaciones en ambos casos mayor será la fiabilidad test -retest. El coeficiente de correlación de Pearson nos proporciona el valor de la correlación entre las dos puntuaciones que se encuentra entre 0 y 1. Cuanto más se acerca la correlación a 1, mayor es la fiabilidad test-retest.

La fórmula del coeficiente de correlación de Pearson es:

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

donde $[\bar{X}]$ e $[\bar{Y}]$ son las medias de cada una de las veces que se respondió el test y X e Y son cada uno de los valores de cada una de las dos variables.

Para **VOCATION**, la fiabilidad test-retest se realizó con la versión inglesa en Reino Unido, con una muestra de N=88, el intervalo entre la administración de la prueba la primera y la segunda vez fue de 3 a 6 semanas. Los participantes tenían entre 21 y 72 años de edad. La fiabilidad test-retest se evaluó para las puntuaciones de las personas en todos los factores y se utilizó el coeficiente de Pearson para calcular las correlaciones ente las puntuaciones obtenidas la primera y la segunda vez.

Tabla 19. Fiabilidad Test-retest por factor.

Factor	r
R	0.84
I	0.76
A	0.84
S	0.80

E	0.74
C	0.87

R=Realista, I=Investigador, A=Artístico, S=Social, E=Emprendedor, C=Convencional

Las correlaciones test-retest para **VOCATION** muestran excelentes resultados, en que la menor correlación es de 0,74 y la media 0,81. Por lo tanto, se puede concluir que **VOCATION** tiene una alta fiabilidad en las puntuaciones a lo largo del tiempo y una persona que realice la prueba en dos momentos puntuales obtendrá de forma consistente puntuaciones semejantes.

7. Validez

La validez de un test indica si la prueba mide lo que dice medir, si tiene un buen diseño y si refleja de forma precisa el comportamiento de la persona. Para que un test sea válido es necesario que exista correspondencia entre cómo se interpretan las puntuaciones del test y la teoría en la que se basa el test. Por lo tanto, la fiabilidad del test estará en función de la consistencia de los resultados a lo largo del tiempo y la validez del test estará en función del nivel de precisión de los resultados y la posibilidad de hacer predicciones con confianza. Es importante tener en cuenta que es posible que un test tenga alta fiabilidad, se puede obtener la respuesta equivocada de forma consistente, y no sea válido. Sin embargo, la fiabilidad si es un prerequisite para que un test tenga validez, es imposible que un test sea válido si no es fiable primero, ya que la precisión de los resultados necesita que exista consistencia en las respuestas. Establecer la validez del test permite a los/las profesionales tomar decisiones adecuadas, evitar errores caros o actuaciones con falta de ética o incurrir en riesgos legales.

El origen etimológico de validez proviene del latín “validus”, que significa “fuerte”. Siguiendo las recomendaciones de organizaciones internacionales para la utilización de tests en el ámbito de la psicología y educación (Standards; APA, AERA, & NCME, 2014), la validación es el proceso mediante el cual se proporciona la evidencia científica para la correcta interpretación de los resultados. Las evidencias aportadas en este proceso proceden del; contenido del test; procesos implicado en la conducta de respuesta al test; estructura interna; relación con otras variables y las consecuencias del proceso de evaluación. **VOCATION** se ha validado mediante un proceso de varias etapas para garantizar la precisión, representatividad, consistencia, claridad, equitatividad e independencia de cultura de los ítems y los factores.

El concepto de validez es amplio y existen diferentes tipos que se discuten en los apartados siguientes. Un concepto de validez que cada vez se tiene más en cuenta en la validez aparente, se refiere a si la prueba parece medir lo que mide desde el punto de vista de las personas que lo contestan (Nevo, 1985); es importante porque incide en la calidad de la experiencia de la persona evaluada; a mayor validez aparente del test más probable será obtener la cooperación y el compromiso de las personas que contestan. La validez aparente y la actitud hacia el test influye en la motivación, concentración e incluso en la ansiedad con la que los candidatos/as realizan la prueba (McCarthy & Goffin, 2003). Durante el estudio de **VOCATION** se preguntó a las personas de la muestra si parecía medir preferencias laborales o vocacionales, para asegurarse de que las preguntas sobre situaciones fueran claras y no llevasen a error. En el siguiente apartado se aporta información en detalle de diferentes aspectos de la validez que se consideraron para **VOCATION**.

7.1. Validez concurrente

La validez concurrente mide si un método de medición coincide con los resultados de otro método similar que ha sido probado previamente. Para establecer la validez

concurrente de **VOCATION** se necesitaba que los componentes del test correlacionaran de forma significativa con los componentes de otros cuestionarios bien establecidos. La validez concurrente requiere que las pruebas que se comparan se realicen en un mismo momento temporal, de esta forma se comprueba que la correlación es pura y no existen cambios debidos al paso del tiempo. En las tablas 16 a la 18 se muestran los resultados en detalle para la validez concurrente de **VOCATION**.

7.1.1. La escala de intereses vocacionales Oregón de IPIP

Se realizó un estudio de la validez concurrente de la versión inglesa de **VOCATION**. Para realizar este estudio se utilizó la escala de intereses vocacionales Oregón de IPIP. El IPIP (International Personality Item Pool) es una colección de más de 3000 ítems de personalidad de uso público, a partir de los cuales se han construido sobre 250 escalas de personalidad reconocidas a nivel internacional. El proyecto comenzó en 1996 y está disponible en más de 25 idiomas. El IPIP cuestionario para esta validación se extrajo de la página web IPIP (Goldberg, 1999).

Se eligió la escala de intereses vocacionales Oregón (ORVIS- Oregon Vocational Interest Scale) como un test adecuado para investigar la validez concurrente de **VOCATION** ya que ambas evalúan los mismos constructos teóricos (Pozzebon, 2011). A continuación, se muestra la correspondencia entre los factores de **VOCATION** y los de ORVIS.

Tabla 20. Correspondencia entre los factores de VOCATION y ORVIS.

VOCATION	ORVIS
R	Production
I	Analysis
A	Creativity
S	Altruism
E	Leadership
C	Organization

R=Realista, I=Investigador, A=Artístico, S=Social, E=Emprendedor, C=Convencional

Se pidió a los participantes: “sin tener en cuenta el salario, indique en qué medida está interesado/a en las siguientes actividades de la escala”. La escala de respuesta era una escala tipo Likert de 5 niveles que iba desde en completo desacuerdo (1) a completamente de acuerdo (5).

Tabla 21. Correlaciones entre los factores de **VOCATION** y **ORVIS** (IPIP).

	Production	Analysis	Creativity	Altruism	Leadership	Organization
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. * $p < .05$, ** $p < .01$, *** $p < .001$.

R=Realista, I=Investigador, A=Artístico, S=Social, E=Emprendedor, C=Convencional

Las correlaciones entre los factores y subfactores de **VOCATION** y **ORVIS** son altas, para los factores varían entre 0.50 de Realista y Production a 0.78 de Emprendedor y Leadership, por lo tanto, se puede concluir que la validez concurrente es elevada y **VOCATION** se ajusta al modelo de intereses ocupacionales. Para un análisis en mayor profundidad también se analizaron las correlaciones entre los subfactores de **VOCATION** y **ORVIS** que se muestran en la tabla a continuación. Los resultados siguen mostrando coherencia en su mayor parte, aunque en el caso de E1 también obtuvo una correlación alta con creatividad, esto puede ser debido a que este subfactor también tiene en cuenta la proactividad necesaria para emprender que tiene relación con la creatividad y saber ver las oportunidades.

Tabla 22. Correlaciones entre los sub- factores de **VOCATION** y **ORVIS** (IPIP).

	Production	Analysis	Creativity	Altruism	Leadership	Organization
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. * $p < .05$, ** $p < .01$, *** $p < .001$.

7.2. Validez de constructo

La validez de constructo hace referencia al grado en que un test mide lo que dice medir; en qué medida las inferencias que se hacen en función de las respuestas dadas son apropiadas. La validez de constructo es básica para la validez en su conjunto, tanto la aparente como la real, de la prueba y es realmente más relevante si el concepto es muy subjetivo. Si el constructo no está definido de forma clara y minuciosa no es posible establecer relaciones entre el constructo y los comportamientos (Cronbach & Meehl, 1955). En ocasiones no existe una unidad para medir los constructos aceptada universalmente; incluso conceptos muy utilizados como la inteligencia están abiertos a especulación y debate. Por lo tanto, la validez de constructo de un test puede verse comprometida si el constructo: se define de forma demasiado amplia o restringida; se confunde con uno de sus componentes; se confunde con constructos relacionados que enmascaran el efecto del constructo principal, el que se quiere medir; se desarrolla sin la suficiente investigación. Por ello, si el objetivo es desarrollar una escala que mida las preferencias laborales y vocacionales, se deben identificar las teorías que definen los intereses profesionales, identificar qué componentes están incluidos y determinar con precisión los atributos positivos y negativos asociados para cada dimensión.

En el ámbito laboral, la validez de constructo es de particular relevancia porque proporciona legitimidad a la evaluación del comportamiento del personal y permite una mayor aceptación; permite entender las implicaciones para la organización de los resultados; permite a las organizaciones y las personas diseñar intervenciones para alcanzar los objetivos de rendimientos previstos. La validez de constructo responde a la cuestión de si **VOCATION** realiza una definición y clasificación de las preferencias laborales y vocacionales de forma integrada en un marco teórico y que se pueda utilizar de forma práctica. Para **VOCATION**, la validez de constructo se estableció utilizando dos enfoques: el análisis factorial, prestando especial atención a las correlaciones entre los factores y el análisis factorial confirmatorio, como se muestra en las secciones siguientes.

7.2.1. Análisis factorial y correlaciones entre los factores

Todos los factores de **VOCATION** miden preferencias laborales o vocacionales, aunque cada uno de ellos mide una preferencia diferente. Por lo tanto, si una persona obtiene una puntuación elevada en un factor, eso no significa necesariamente que obtendrá una puntuación elevada en todos los factores. El coeficiente de correlación es la forma numérica de cuantificar la relación entre dos variables y va de -1 a 1. Cuanto más se acerca el valor a 1 mayor es la relación. El coeficiente de correlación no implica causalidad; simplemente señala la relación entre dos variables, pero es importante tener en cuenta que puede haber terceras variables explicativas. Lo habitual es utilizar el coeficiente de Pearson (1985) en el análisis de las correlaciones entre los factores. En la tabla 22 se muestran las correlaciones entre los diferentes factores.

Tabla 23. Correlaciones entre los factores.

	R	I	A	S	E
I	-0.17				
A	-0.27	-0.32			
S	-0.26	-0.28	0.02		
E	-0.18	-0.23	-0.22	-0.12	
C	-0.11	-0.04	-0.19	-0.43	-0.21

R=Realista, I=Investigador, A=Artístico, S=Social, E=Emprendedor, C=Convencional

Las correlaciones en la tabla muestran las relaciones entre varios factores de **VOCATION**. La mayor parte de las correlaciones son negativas, lo que indica independencia entre los factores que evalúan cada interés profesional. La única correlación positiva, aunque muy baja es entre S y A con 0.02. La mayor independencia entre los factores ocurre entre C y S con -0.43.

7.2.2. Análisis factorial confirmatorio y exploratorio

El análisis factorial confirmatorio y exploratorio, CFA y EFA respectivamente, se utilizan para analizar la cualidad de los ítems individuales y la teoría subyacente a los constructos teóricos. Tanto CFA como EFA son análisis factoriales. El análisis factorial consiste en realizar una matriz de correlaciones, donde se agrupan los ítems que tiene correlaciones significativas, esto permite comprobar si los ítems pertenecen al factor que se desea medir y solo a ese factor. El análisis factorial tiene como objetivo agrupar el conjunto de ítems en una serie de variables comunes; que se obtendrán eliminando los ítems no relevantes y agrupando los que tienen alta correlación en el factor, de esta forma se identifican las variables comunes.

El objetivo de EFA es identificar y poner de relevancia la importancia de cada dimensión (porcentaje de varianza), mientras que CFA es un enfoque en el que se plantea una hipótesis, se especifica en número de factores que se espera encontrar y se clasifican según los valores de las relaciones esperada, esto sirve para determinar en qué grado los indicadores observados reflejan el constructo subyacente (Schreiber, Nora, Stage, Barlow & King, 2006). De esta forma, CFA proporciona una estimación numérica del grado en que la estructura de los factores propuesta refleja de forma adecuada y precisa como se manifiesta el constructo en la vida real. Por lo tanto, CFA explora el grado en que la medición de un constructo coincide con la concepción teórica de ese constructo.

El análisis factorial confirmatorio (CFA) con indicadores para los factores se llevó a cabo para comprobar la convergencia entre los resultados con el modelo teórico de la prueba. CFA es un tipo de análisis factorial usado para comprobar que el constructo representa lo que se intenta medir. Existen dos enfoques básicos para comprobar la estructura factorial de la escala. Un enfoque estrictamente confirmatorio es aquel en el que el modelo se prueba utilizando los test de bondad de ajuste para elaborar el modelo de ecuaciones estructurales (SEM) para comprobar que el modelo propuesto tras realizar la investigación es consistente con los datos obtenidos. El otro enfoque, permite poner a prueba dos o más modelos para identificar cuál se ajusta mejor. Para comprobar la validez de **VOCATION** se utilizó un enfoque confirmatorio estricto ya que sólo existe un modelo para **VOCATION**.

CFA se utiliza para comprobar la adecuación de los datos al modelo, por lo tanto, examinar los índices de bondad de ajuste es fundamental. Existen diferentes pruebas estadísticas para determinar el ajuste del modelo a los datos en CFA y SEM. Cuanto mayor sea el ajuste entre los índices y los requisitos estadísticos, mejor será el modelo. El Chi cuadrado (X^2) está considerado como el índice principal de la bondad de ajuste. Sin embargo, debido a los problemas asociados con la dependencia del X^2 , como la sensibilidad al tamaño de la muestra o su inadecuación si no se respetan los supuestos de la distribución normal. En estos casos es recomendable utilizar medidas de ajuste relativas. Chi cuadrado es una medida de ajuste absoluta, proporciona una indicación básica de hasta qué punto la teoría propuesta encaja con los datos y no depende de la comparación con el modelo de base (Jöreskog & Sörbom, 1993).

En la validación de **VOCATION** se escogieron como medidas adicionales del ajuste, la Media de la raíz cuadrada del error de aproximación (RMSEA, Steiger y Lind, 1980); el Índice comparativo de ajuste (ICA, Bentler, 1990); y el Índice de Tucker-Lewis (TLI); porque es menos probable que estos índices se vean afectados por el tamaño de la muestra (Fan, Thompson & Wang, 1999). Además, utilizar diferentes índices es necesario para reflejar las diferentes formas del ajuste del modelo (Hooper, Coughlan & Mullen, 2008). Al igual que X^2 , RMSEA es una medida de ajuste absoluta que evalúa la cantidad de error asociado con el modelo propuesto. Cuanto menor sea RMSEA, más robusto será el modelo; índices de 0,5 o menores se consideran un buen ajuste (Hu & Bentler, 1999). Esto se debe a que RMSEA tiene en cuenta los aspectos residuales del modelo y evalúa el grado en que el modelo propuesto no se ajusta. Tanto el CFI como el TLI son índices del ajuste incrementales o relativos que evalúan la mejora conseguida de un modelo en

comparación con la independencia teniendo en cuenta las variables que no correlacionan. (Hooper et al., 2008; Holtzman & Vezzu, 2011). Mientras que niveles elevados en CFI y TLI (por encima de 0,9 o superiores) indican un ajuste excelente, se ha de tener en cuenta que CFI o TLI de 0,90 o superiores no se corresponden necesariamente con RMSEA de 0,50 o menores (Hu & Bentler, 1999).

Se llevaron a cabo dos análisis CFA en el modelo de 6 factores y en el modelo de 12 subfactores. Las tablas 23 y 24 muestran los índices de bondad de ajuste de los resultados del modelo.

Teoría de respuesta a los ítems de Thurstone: La teoría de respuesta a los ítems (TRI) es un conjunto de modelos matemáticos, que describen en términos probabilísticos, las relaciones entre la respuesta de una persona en un cuestionario y su nivel en la variable medida (Hays & Fayes 2005, p.55). En los test ipsativos, la persona que responde debe hacer juicios comparativos entre dos o más ítems para dar una respuesta. Sin embargo, en test normativos, cada juicio se mide en términos absolutos para cada atributo evaluado, en vez de comparar preferencias entre los diferentes atributos. Las puntuaciones en los test clásicos no muestran este proceso psicológico de comparación, sino que ignoran el hecho de que las personas muestran una preferencia y no un juicio absoluto. La ley de Thurstone de juicios comparativos (Thurstone, 1927; 1931) proporciona resultados basados en los juicios comparativos de las personas. Para poder analizar los datos utilizando el modelo de respuesta a los ítems de Thurstone, los datos brutos deben transformarse en datos binarios siguiendo el procedimiento descrito por Brown y Maydeu-Olivares (2011) para los modelos de respuesta de comparación por pares y posicionamientos en una escala. En **VOCATION**, las 60 preguntas consisten en grupos de 2 ítems que constituyen las opciones de respuesta. Las respuestas observadas en los dos enunciados a comparar se codificaron utilizando la fórmula que se muestra a condición donde i representa al par $\{a, b\}$;

$$y_i = \begin{cases} 1 & \text{if item } a \text{ is preferred over item } b \\ 0 & \text{if item } b \text{ is preferred over item } a \end{cases}$$

El modelo de respuesta a los ítems de Thurstone se ajustó para resultados binarios de comparaciones entre pares de enunciados con una estructura de 6 y 12 factores latentes. Se generaron un total de 60 resultados binarios para 60 bloques de dos ítems. Se utilizó la Mplus version 7.1 (Muthén & Muthén, 2014) para estimar los parámetros de los ítems para correlaciones tetracóricas usando los algoritmos de mínimos cuadrados sin ponderar.

Tabla 24. Análisis factorial confirmatorio utilizando TRI para el modelo de 6 factores.

N	X ²	df	X ² /df	p	RMSEA	P ≤ .05	CFI	TLI
564	2670	1635	1.633	0	0.034	1	0.874	0.863

Tabla 25. Análisis factorial confirmatorio utilizando TRI para el modelo de 12 subfactores.

N	X ²	df	X ² /df	p	RMSEA	P ≤ .05	CFI	TLI
564	2043	1584	1.289	0	0.023	1	0.944	0.937

Las tablas 24 y 25 muestran el resultado del ajuste del modelo. Debido al tamaño de la muestra, se utilizaron los índices más robustos, RMSEA y CFI/TLI para evaluar el ajuste del modelo. Los resultados de estos índices muestran que ambos modelos tienen un buen ajuste y que el modelo de 12 factores tiene mejor ajuste que el modelo de 6 factores. Esto confirma la validez del modelo de **VOCATION**. El hecho de que el modelo de 12 factores ajuste mejor que el modelo de 6 proporciona más confianza en la idea de que cada factor del modelo RIASEC se puede dividir en dos subfactores.

7.2.3. Validez consecuencial

La validez consecuencial surge de la preocupación por la ética en relación a las consecuencias en la sociedad de las evaluaciones. Puede definirse como un tipo de validez relacionada con las consecuencias intencionadas o no, positivas o negativas en función de la interpretación y uso de la prueba (Messick, 1995). Messick (1980, p.1012) explica que “los test no deberían evaluarse solo en términos de sus propiedades métricas, sino que, las herramientas de evaluación deberían evaluarse en términos de sus consecuencias sociales potenciales”. La validez consecuencial es importante dentro de la validez general y debe integrarse en el conjunto más que considerarlo como un concepto aislado (Hubley & Zumbo, 2011). La validez consecuencial es especialmente importante en **VOCATION** debido al alto coste asociado a la contratación de un candidato y la legislación de los gobiernos sobre empleo, diversidad e inclusión de grupos demográficos específicos. Información más en detalle sobre la validez consecuencial se dan en el capítulo 8.

8. VOCATION y la igualdad de oportunidades

Las organizaciones tienen cada vez más presión para utilizar métodos y herramientas imparciales, objetivos y precisos en la evaluación para la selección, contratación, desarrollo, promoción y evaluación del rendimiento de los empleados. En la actualidad, el conjunto de los trabajadores/as tiene mayor acceso a todo tipo de información y los candidatos suelen estar mejor informados que en el pasado, además la legislación local e internacional se está volviendo progresivamente más estricta, ambas situaciones repercuten en mayor presión para las empresas. La igualdad de oportunidades debe tenerse en cuenta siempre en cualquier evaluación en el trabajo a día de hoy, y se considera que una prueba de evaluación es válida solamente si es apropiada para su uso por diferentes grupos. **VOCATION** es una prueba imparcial y precisa de las preferencias laborales y vocacionales por lo que los responsables de RRHH y psicólogos/as pueden tener confianza en que en su aplicación se está respetando la diversidad.

Cuando se realiza una evaluación, los instrumentos utilizados no deben tener un impacto que perjudique a ningún grupo evaluado. El impacto negativo ocurre si hay una diferencia significativa en la cantidad de personas seleccionadas de un grupo respecto a otro. Esto tiene gran importancia, principalmente por el perjuicio hacia el grupo discriminado, además también puede tener consecuencias económicas y legales como elegir a la persona equivocada para un trabajo o tener que afrontar penalizaciones o juicios por prácticas discriminatorias. El impacto negativo es una práctica discriminatoria si en que la no elección de una persona concreta se debe a su origen étnico, género, grupo de edad o cualquier otra característica demográfica que le haga pertenecer a un grupo potencialmente vulnerable. Existe legislación relativa al respeto al derecho a evaluaciones imparciales, objetivas, sin prácticas que impliquen discriminación en cada uno de los países de habla hispana para proteger los derechos de las personas que participan en cualquier proceso de evaluación en el ámbito laboral.

Por lo tanto, es importante conocer si las diferencias que se dan en los resultados entre los grupos son significativas o no, para poder explorar a que son debidas.

8.1. T-tests y ANOVAS

Para explorar las diferencias entre los grupos se llevó a cabo un t-test y ANOVA por género y edad, cuyos resultados se muestran a continuación.

Diferencias entre los grupos de edad: Se dividió a los participantes en 6 grupos en función de su edad: menores de 25 (n=262), 26 a 33 (n=105), 34 a 41 (n=88), 42 a 49 (n=56), 50 a 56 (n=28) y por encima de 57 (n=25).

Tabla 26. ANOVA según factor y grupo de edad.

	- 25	26- 33	34 -41	42 -49	50 -56	+ 57	F	df	p-value
R	3.99	4.24	4.16	3.92	3.04	2.68	3.032	5.00	0.01*
I	5.23	5.53	4.57	5.05	4.84	5.34	1.844	5.00	0.102
A	5.28	4.83	5.32	5.13	5.14	5.62	4.422	5.00	0.487
S	5.16	4.2	4.57	5.34	5.7	4.96	3.708	5.00	0.002*
E	5.59	5.05	5.43	5.13	6.07	6.18	2.143	5.00	0.059
C	4.74	6.14	5.95	5.42	5.21	5.22	7.913	5.00	0*

R=Realista, I=Investigador, A=Artístico, S=Social, E=Emprendedor, C=Convencional

* indica que la diferencia es significativa

Según se puede observar en la tabla, las diferencias en las puntuaciones de R, S y C según el grupo de edad si presentan diferencias significativas. Para explorar estas diferencias en mayor profundidad se realizó un t-test a posteriori comparando cada par de grupos de edad entre sí en cada uno de los factores.

Tabla 27. T-test a posteriori para los factores para cada par de grupos de edad.

Edad	-25	26 - 33	34 - 41	42 - 49	50- 56
R					
26 - 33	0.329				
34 - 41	0.539	0.797			
42 - 49	0.835	0.386	0.534		
50 - 56	0.033*	0.012*	0.02*	0.09	
+57	0.005*	0.001*	0.003*	0.022*	0.566
I					
26 - 33	0.262				
34 - 41	0.022*	0.004*			
42 - 49	0.612	0.218	0.227		
50 - 56	0.404	0.165	0.595	0.693	
+57	0.821	0.711	0.148	0.612	0.439
A					

26 - 33	0.082				
34 - 41	0.864	0.125			
42 - 49	0.664	0.408	0.618		
50 - 56	0.763	0.508	0.708	0.986	
+57	0.462	0.111	0.558	0.365	0.437
<i>S</i>					
26 - 33	0*				
34 - 41	0.045*	0.293			
42 - 49	0.607	0.004*	0.06		
50 - 56	0.258	0.003*	0.03*	0.519	
+57	0.692	0.156	0.47	0.51	0.264
<i>E</i>					
26 - 33	0.033*				
34 - 41	0.542	0.234			
42 - 49	0.158	0.813	0.437		
50 - 56	0.273	0.029*	0.177	0.066	
+57	0.202	0.021*	0.131	0.049*	0.858
<i>C</i>					
26 - 33	0*				
34 - 41	0*	0.553			
42 - 49	0.038*	0.05	0.161		
50 - 56	0.285	0.05	0.126	0.688	
+57	0.304	0.062	0.146	0.707	0.992

R=Realista, I=Investigador, A=Artístico, S=Social, E=Emprendedor, C=Convencional

* indica que la diferencia es significativa

El ANOVA para los factores RIASEC para los grupos de edad encontró diferencias significativas en R, S y C. En el análisis posterior en que se utilizó la comparación por pares con la corrección de Bonferroni se exploraron en profundidad estas diferencias.

Para R, se encontró que las diferencias eran significativas para el grupo + de 57 con los grupos de personas menores a 50 y para el grupo de 50 a 56 con los grupos menores de 42. Esto indica que las personas eligen los ítems del factor R menos a medida que tienen mayor edad. Indicando una disminución del interés en actividades laborales que impliquen actividad física o destreza manual según se aumenta en edad.

Para S, se encontraron diferencias significativas entre los grupos de -25 con 26 a 33 y 34 a 41 en que las personas de menor edad tienen mayor interés en S, también hay diferencias entre las personas de 26 a 33 con los grupos de 42 a 49 y 50 a 56, en que las personas de más edad tienen más interés en S que las personas de 26 a 33. Por último, las personas de 50 a 56 tienen más interés que las personas de 34 a 41 en S de forma significativa. Es interesante ver que las personas del grupo de edad más joven y las personas de mayor edad tengan un interés mayor en S. Aunque puede haber diferentes razones puede ser que las personas jóvenes estén en una fase de exploración que los lleve a querer conectar con los demás y las personas mayores ya no tienen tantas cargas familiares o presión por su desarrollo profesional y de ahí que su foco de interés cambie hacia intereses más sociales, en línea con lo que propone la teoría Life Span de Donald Super (1996).

Para C se encontraron diferencias significativas entre -25 y los grupos de edades intermedias, de 26 a 33, 34 a 41 y 42 a 49. Los grupos de edad intermedios parecen tener un interés significativamente mayor en C. Aunque puede haber diferentes explicaciones, una razón puede ser la menor exposición al mundo laboral por parte de las personas menores de 25, este desconocimiento de lo que en realidad supone el interés C puede influir en que se elijan los ítems relacionados con este factor en el test.

Para un mayor conocimiento de las diferencias se llevó el análisis ANOVA para los grupos de edad en los subfactores. Los resultados globales se muestran a continuación.

Tabla 28. ANOVA según grupo de edad para los subfactores.

	- 25	26- 33	34 -41	42 -49	50 -56	+ 57	F	df	p-value
R1	4.25	4.66	4.17	4.07	3.18	2.68	3.157	5.00	0.008*
R2	3.73	3.83	4.15	3.77	2.89	2.68	2.151	5.00	0.058
I1	5.68	6.14	6.15	6.2	6.14	6.12	1.255	5.00	0.282
I2	5.31	6.32	6.08	5.59	6.39	6.36	3.562	5.00	0.003*
A1	5.24	5.24	4.18	4.59	3.64	4.72	3.753	5.00	0.002
A2	5.31	5.52	4.84	5.2	4.64	5.76	1.258	5.00	0.281
S1	5.12	4.28	4.53	5.45	5.64	5.12	2.524	5.00	0.028*
S2	5.19	4.13	4.6	5.23	5.75	4.8	4.048	5.00	0.001*

E1	5.66	5.24	5.48	5.91	6.71	6.76	3.629	5.00	0.003*
E2	5.57	5.1	5.7	6.07	5.57	6.48	2.585	5.00	0.025*
C1	5.16	4.9	5.23	4.29	4.36	4.28	2.541	5.00	0.027*
C2	3.74	4.61	4.88	3.64	4.07	4.24	3.511	5.00	0.003*

R1= Actividades físicas y al exterior, R2= Intereses técnicos y manuales, I1=curiosidad y aprendizaje intelectual, I2= Ciencia y tecnología, A1=Sentido y expresión estética, A2= Creatividad y diseño, S1=Dedicación a los demás, S2= relaciones sociales, E1= emprendedor, E2= Liderazgo, C1= Metódico, C2= Interés por los datos y los números.

* indica que la diferencia es significativa

En el análisis para los subfactores se encontraron más diferencias significativas entre los grupos de edad, solo los subfactores de A (A1 y A2), R2 y I1 no mostraron diferencias por edad. Para analizar entre que grupos se daban las diferencias se llevó un análisis a posterior entre los pares de grupos de edad para los subfactores.

Tabla 29. T-test a posterior para los subfactores por grupos de edad.

Edad	-25	26 - 33	34 - 41	42 - 49	50 - 56
R1					
26 - 33	0.184				
34 - 41	0.802	0.203			
42 - 49	0.642	0.180	0.826		
50 - 56	0.041*	0.008*	0.084	0.144	
+57	0.004*	0*	0.013*	0.029*	0.493
R2					
26 - 33	0.713				
34 - 41	0.159	0.364			
42 - 49	0.905	0.88	0.361		
50 - 56	0.086	0.071	0.018*	0.121	
+57	0.041*	0.034*	0.008*	0.064	0.751
I1					
26 - 33	0.074				
34 - 41	0.09	0.988			

42 - 49	0.117	0.884	0.898		
50 - 56	0.299	1	0.992	0.917	
+57	0.348	0.963	0.956	0.886	0.97
<i>I2</i>					
26 - 33	0*				
34 - 41	0.015*	0.508			
42 - 49	0.463	0.083	0.262		
50 - 56	0.034*	0.899	0.572	0.175	
+57	0.02*	0.949	0.628	0.21	0.962
<i>A1</i>					
26 - 33	0.984				
34 - 41	0.001*	0.007*			
42 - 49	0.102	0.15	0.381		
50 - 56	0.003*	0.006*	0.361	0.133	
+57	0.357	0.392	0.383	0.841	0.15
<i>A2</i>					
26 - 33	0.417				
34 - 41	0.143	0.058			
42 - 49	0.768	0.407	0.422		
50 - 56	0.196	0.099	0.725	0.356	
+57	0.406	0.719	0.118	0.366	0.118
<i>S1</i>					
26 - 33	0.009*				
34 - 41	0.088	0.523			
42 - 49	0.431	0.012*	0.057		
50 - 56	0.349	0.022*	0.068	0.761	
+57	0.997	0.175	0.355	0.627	0.497
<i>S2</i>					
26 - 33	0*				

34 - 41	0.047*	0.18			
42 - 49	0.916	0.006*	0.128		
50 - 56	0.249	0.001*	0.029*	0.356	
+57	0.436	0.216	0.718	0.458	0.154
<i>E1</i>					
26 - 33	0.094				
34 - 41	0.496	0.448			
42 - 49	0.435	0.0629	0.245		
50 - 56	0.015*	0.001*	0.009*	0.112	
+57	0.016*	0.001*	0.009*	0.106	0.939
<i>E2</i>					
26 - 33	0.11				
34 - 41	0.667	0.101			
42 - 49	0.184	0.021*	0.403		
50 - 56	0.02*	0.007*	0.12	0.4	
+57	0.09	0.015*	0.183	0.508	0.897
<i>C1</i>					
26 - 33	0.303				
34 - 41	0.82	0.309			
42 - 49	0.008*	0.103	0.015*		
50 - 56	0.073	0.263	0.076	0.891	
+57	0.621	0.221	0.064	0.991	0.901
<i>C2</i>					
26 - 33	0.005*				
34 - 41	0*	0.493			
42 - 49	0.804	0.029*	0.007*		
50 - 56	0.534	0.345	0.167	0.49	
+57	0.373	0.535	0.296	0.354	0.819

R1= Actividades físicas y al exterior, R2= Intereses técnicos y manuales, I1= curiosidad y aprendizaje intelectual, I2= Ciencia y tecnología, A1= Sentido y expresión estética, A2= Creatividad y diseño, S1= Dedicación a los demás, S2= relaciones sociales, E1= emprendedor, E2= Liderazgo, C1= Metódico, C2= Interés por los datos y los números.

* indica que la diferencia es significativa

En R1 se encontraron diferencias significativas entre los mayores de 57 con todos los grupos de edad hasta 49 años. El grupo de edad de 50 a 56 también mostró diferencias significativas con los dos primeros grupos más jóvenes, de -25 y 26 a 33. Una vez más estas diferencias indican que las personas tienen menos interés por actividades físicas y al exterior a medida que aumenta la edad. Aunque también hay que destacar que tanto R1 como R2 son poco populares, con medias por debajo de 5.5 en todos los grupos de edad, esto puede ser debido a que los trabajos relacionados con este interés tienen asociado de forma tradicional un estatus menor a nivel social.

Para R2 se encontraron diferencias significativas entre los grupos de 50 a 56 con 34 a 41 y entre los grupos de + 57 con los grupos de edad menores a 41 años. Estas diferencias vuelven a mostrar la misma tendencia observada en R y en R1, siendo los ítems correspondientes a los intereses técnicos y manuales escogidos en menor medida según se avanza en edad. Esto puede ser debido a diferentes razones, como su menor estatus o el mayor deseo de dedicarse a actividades en que se deban gastar menores niveles de energía a medida que se avanza en edad.

En I2, ciencia y tecnología nos encontramos diferencias significativas entre el grupo de edad menor con el resto de grupos de edad menos el intermedio 42 a 49. De todos los grupos, incluido 42 a 49, los menores de 25 son los que eligen en menor medida los ítems relacionados con el subfactor (este parece ocurrir también con el grupo de 42 a 49, pero en menor medida). Se podrían lanzar varias hipótesis sobre el porqué de esta situación, pudiendo ocurrir que las personas más jóvenes no tienen una experiencia tan directa de la aplicabilidad de la ciencia y tecnología en el trabajo, como personas de mayor edad.

Para A no se habían encontrado diferencias significativas según la edad, pero en el análisis entre pares, se encontró que en A1, sentido y expresión estéticas si que había diferencias significativas entre algunos grupos. La tendencia observada es a elegir los ítems de este subfactor en menor medida, según se avanza en edad. Esto puede indicar una mayor preferencia por expresar las emociones de forma artística en el trabajo en personas más jóvenes. Las diferencias significativas se encontraron en el grupo 34 a 41, con menor interés que los grupos menores de 25 y de 26 a 33; y también entre el grupo 50 a 56, con un interés significativamente menor que los grupos de menos de 25 o 26 a 33.

En S1, dedicación a los demás y S2, relaciones sociales, muestran una situación semejante a la observada y analizada en el factor global S, siendo el interés mayor en los menores de 25 y los mayores de 42. Solo se atenúa la tendencia un poco en el grupo de mayor edad, en que vuelve a decrecer. Las diferencias significativas para S1 se dan entre los menores de 25 con el grupo entre 26 y 33, en que el grupo de menores de 25 tiene un

mayore interés en S1, y entre el grupo 26 a 33 con los grupos 42 a 49 y 50 a 56, en que los grupos de mayor edad muestran un interés significativamente mayor en el factor. Las diferencias significativas en S2 se dan entre los menores de 25 con un interés mayor en S2 que los grupos de 26 a 33 y 34 a 41; el grupo de 26 a 33 que tiene un interés significativamente menor que los grupos 42 a 49 y 50 a 56. Y el grupo 34 a 41 que muestra un interés que muestra un interés significativamente menor con el grupo 50 a 56.

E1, emprendedor y E2, liderazgo, siguen tendencias similares. Más marcada si cabe en E1. Esta tendencia es a tener una mayor probabilidad de identificarse con los ítems de estos subfactores a medida que se avanza en edad. Además, se puede decir que son un factor y subfactores muy populares, ya que la mayor parte de las medias en los grupos de edad están por encima de 5.5. Una posible explicación es el mayor estatus que las profesiones asociadas con el factor y subfactores tienen, el mayor deseo de independencia en el trabajo e incluso de asumir el liderazgo puede venir por una mayor confianza de las personas en sí mismas a medida que se aumenta en edad, debido a la experiencia que se acumula. Las diferencias son significativas para E1 en los grupos de edad entre 50 a 56 y más de 57 con los grupos de menores de 25 hasta 41 años de edad, siendo mayor para los mayores de 50. En E2, se dan diferencias significativas entre 50 a 56, mayor interés, con los grupos menores de 25 y de 26 a 33; también entre los grupos de 42 a 49 y más de 57, mayor interés, que el grupo de edad entre 26 y 33.

El factor C, junto con los subfactores C1 y C2 parece ser poco popular, con todas las medias en los grupos de edad por debajo de 5.5. En C1, metódico, parece que hay un interés inicial, que va decreciendo según se aumenta en edad, para C2, interés por los datos y los números, los niveles de interés son bajos en todos los grupos de edad, siendo los más bajos en la gente más joven, por debajo de 25 y en el grupo 42 a 49. El hecho de que para el grupo 42 a 49 el interés en C1 y C2 sea menor que otros grupos de edad y de forma significativa con otros grupos puede ser debido a muchas razones, se puede reflexionar sobre el momento vital de muchas de estas personas como hace la teoría ya mencionada de Super (1996), la experiencia acumulada puede hacer que no se dependa tanto de la organización y planificación o la necesidad de depender de datos. En cualquier caso, se necesitaría analizar la información en mayor profundidad para poder proporcionar explicaciones claras al respecto. Los grupos entre los que se encontraron diferencias significativas para el subfactor C1 fueron, el grupo 42 a 49 mostró un interés significativamente menor que el grupo de menores de 25 y el de 34 a 41. En el subfactor C2, se encontraron diferencias significativas entre el grupo 42 a 49, interés menor, que los grupos 26 a 33 y 34 a 41; así como el grupo de menores de 25 con un interés significativamente menor que los grupos 26 a 33 y 34 a 41.

Género: Se analizaron las diferencias entre las puntuaciones de los dos géneros mediante un t-test para los factores de **VOCATION**.

Table 30. T-test para género en los factores.

Factor	Femenino	Masculino	T.	df	p
R	3.57	4.37	- 4.2408	562	0*
I	4.98	5.34	- 1.8351	562	0.067
A	5.46	4.89	3.0723	562	0.002*
S	5.46	4.3	5.8468	562	0*
E	5.46	5.47	0.0645	562	0.948
C	5.04	5.59	- 2.9112	562	0.003*

R=Realista, I=Investigador, A=Artístico, S=Social, E=Emprendedor, C=Convencional

* indica que la diferencia es significativa

Los t-test encontraron diferencias significativas entre hombres y mujeres en los factores R, A S y C. Los hombres tienen un interés mayor por R y C, mientras que las mujeres tienen más interés por A y S. Esto parece ser congruente con las expectativas y estereotipos sociales sobre las profesiones correspondientes para cada género.

Tabla 31. T-test para género en los subfactores.

Subfactores	Femenino	Masculino	T.	df	p
R1	3.83	4.56	-3.29	562	0.001*
R2	3.32	4.18	-4.27	562	0*
I1	5.69	6.21	-2.796	562	0.005*
I2	5.10	6.47	-6.506	562	0*
A1	5.38	4.36	4.459	562	0*
A2	5.51	4.96	2.4804	562	0.013*
S1	5.64	4.12	6.657	562	0*
S2	5.29	4.49	3.915	562	0*
E1	5.73	5.61	0.6221	562	0.534
E2	5.46	5.84	-1.784	562	0.074

C1	5.23	4.64	3.093	562	0.002*
C2	3.78	4.48	-3.088	562	0.002*

R1= Actividades físicas y al exterior, R2= Intereses técnicos y manuales, I1=curiosidad y aprendizaje intelectual, I2= Ciencia y tecnología, A1=Sentido y expresión estética, A2= Creatividad y diseño, S1=Dedicación a los demás, S2= relaciones sociales, E1= emprendedor, E2= Liderazgo, C1= Metódico, C2= Interés por los datos y los números.

* indica que la diferencia es significativa

También se encontraron diferencias significativas entre hombres y mujeres en los subfactores. Los hombres siguen teniendo mayor preferencia que las mujeres en R, pero en el caso de C, solo seguían teniendo mayor preferencia en C2, mientras que C1 fue preferido por las mujeres de forma significativa. Los hombres también tienen preferencia de forma significativa por I. Por el contrario, las mujeres mostraron preferencia por ambos A y S, además del ya mencionado C1. Para ambos subfactores de E no se encontraron diferencias significativas entre hombres y mujeres. Los datos están en línea con las expectativas y estereotipos generales para ambos géneros. Es posible que a medida que la sociedad vaya cambiando y se normalice en mayor medida la homogeneidad de trabajos y tareas entre los géneros, también cambien los datos sobre los intereses profesionales.

9. Información y consideraciones adicionales

VOCATION se ha diseñado para su uso en entornos organizacionales. Es importante hacer un uso adecuado de los baremos, puntuaciones e interpretaciones para guiar las decisiones en RRHH de cara a la selección y contratación, desarrollo profesional, coaching y evaluación del rendimiento. La interpretación exacta de los test psicométricos ayuda a los evaluadores y a los candidatos a: a) contextualizar la evaluación, b) comparar a los candidatos de forma eficaz respecto a las habilidades requeridas; c) identificar con precisión las áreas de desarrollo, d) maximizar el ajuste entre las necesidades del puesto y los intereses del candidato y e) promover que la organización sea un lugar integro y justo donde trabajar. En este capítulo se proporciona una guía para que los profesionales de RRHH entiendan e interpreten con éxito las puntuaciones del candidato en **VOCATION** de cara a alcanzar los objetivos de la organización.

El software usado en la elaboración de los informes de **VOCATION** proporciona gráficos y comentarios personalizados de los 6 factores y 12 subfactores relacionados con la puntuación del candidato. El informe debe ser interpretado solamente por profesionales que hayan realizado el curso de formación acreditado por *Central Test* o tengan la capacitación legal para realizar evaluaciones psicométricas y que estén familiarizados con este manual. La capacitación y asimilación del proceso de evaluación son esenciales ya que las puntuaciones se interpretan en función de los datos relacionados con la fiabilidad, la validez y el error típico de medida, así como los ajustes que se hayan podido realizar durante la evaluación. Las puntuaciones deben ser presentadas de forma apropiada (claridad, relevancia y precisión) y a las personas pertinentes.

9.1. Interpretación de las puntuaciones

Todas las evaluaciones de *Central Test* han sido diseñadas con para ser utilizadas en el área de RRHH. Por lo tanto, no se pueden sacar conclusiones clínicas de los resultados, aunque estos sean extremos. En los informes existen seis diferentes niveles posibles en total que se determinan por la puntuación de un candidato. La tabla 30 muestra un ejemplo de cómo se describiría un determinado nivel de uno de los factores. Las puntuaciones extremas se definen como los valores más bajos o más altos posibles de un factor. En el caso de **VOCATION**, la puntuación mínima se fija en 0 y la máxima en 10. Una consecuencia del formato ipsativo de la evaluación es la imposibilidad de obtener puntuaciones extremas en todos los factores.

Cada factor está compuesto de dos subfactores. Las puntuaciones altas en un factor determinado indican que existe una fuerte probabilidad de que la persona evaluada tenga interés en ese factor. Las puntuaciones medias (por ejemplo, de 4 a 7) indican similitud con el grupo normativo y pueden indicar menor grado de preferencia individual. Sin embargo, todos los comentarios deberían ser considerados como hipótesis que necesitan completarse con otras fuentes de información, como las entrevistas, otras evaluaciones, el contenido del CV y las referencias personales.

Tabla 32. Tipo de hipótesis en función de la puntuación.

Nivel 1	Nivel 1	Nivel 1	Nivel 1	Nivel 1	Nivel 6
0 a 2	2.1 a 3.5	3.6 a 5.0	5.1 a 6.4	6.5 a 7.8	7.9 a 10
No demuestra ningún interés	Muy bajo interés	Interés bajo, solo en ocasiones	Ligero interés	Interés	Fuerte interés

9.2. Necesidades organizacionales: Elección de carrera, contratación y desarrollo

VOCATION constituye una herramienta eficaz en la evaluación del personal, ya que, por un lado, identifica los intereses profesionales actuales y por el otro, permite seleccionar a los candidatos/as para a un puesto basándose en las áreas ocupacionales que encajan con sus intereses vocacionales, así mismo el candidato/a al puesto tiene la posibilidad de elegir con éxito una trayectoria profesional adecuada a sus intereses.

Decidir cuál es la trayectoria profesional adecuada es una tarea compleja debido a la cantidad abrumadora de opciones profesionales en la actualidad. **VOCATION** permite identificar el trabajo más adecuado para un candidato/a gracias a sus intereses de ocio y tiempo libre. Optar por una carrera en que pueda desarrollar sus intereses permite a la persona conseguir una mayor satisfacción profesional.

Esto pone de manifiesto la importancia que tiene la evaluación de los intereses profesionales en la selección de personal. Conocer las preferencias de los candidatos durante el proceso de selección permite elegir a aquel candidato con más posibilidades de sentirse satisfecho con su trabajo y que quiera, por consiguiente, permanecer en la empresa. Con la ayuda de **VOCATION**, la organización será capaz de inferir lo apropiado de un trabajo para un candidato/a basándose en una evaluación objetiva de sus intereses. Si los intereses no se corresponden con los del puesto de trabajo, puede que exista el riesgo de una falta de ajuste entre ambos.

La selección de personal es un área organizacional importante que puede consumir mucho tiempo y suponer gastos elevados. Además, los errores en la toma de decisiones respecto a la contratación tienen repercusiones adicionales en la productividad, capacidad y sostenibilidad de la organización. Así, alguien que trabaja en la administración financiera y jurídica puede sentirse más cómodo si tiene una preferencia por las tareas que implican respetar normas estrictas, sin embargo, una persona que trabaja en marketing encajará mejor si tiene preferencia por tareas creativas. Por esto, se recomienda que antes de administrar **VOCATION**, se cuente con una descripción detallada del puesto con los requisitos personales basados en un análisis del puesto.

VOCATION ayuda a los profesionales de la selección de personal mediante la generación de los perfiles del puesto en la cuenta cliente donde se pueden especificar cuáles son las puntuaciones deseadas para cada factor. De esta forma se puede comparar con mayor facilidad el perfil buscado con el perfil del candidato.

Además, **VOCATION** puede combinarse con otras evaluaciones de desempeño que son un elemento fundamental en la evolución y éxito de la carrera profesional de la persona. La evaluación del desempeño debe hacerse de forma justa y proporcionada a expectativas realistas de la actuación de la persona, permite hablar sobre los logros y expectativas futuras. **VOCATION** ayuda en la evaluación de desempeño porque puede poner en relación la eficiencia con el interés mostrado en el perfil. Si hay un bajo ajuste entre el puesto y los intereses puede ser normal que no se de un rendimiento muy alto y se puede explorar la posibilidad de un traslado a un lugar o tarea en que se encaje mejor.

VOCATION es eficaz en la identificación de áreas de desarrollo, pero no debería utilizarse para realizar un seguimiento directo de la evolución del rendimiento del empleado. Recomendamos, por lo tanto, recurrir a evaluaciones del desempeño objetivas u otro tipo de evaluaciones en combinación con **VOCATION**.

Referencias

- Anguiano-Carrasco, C., MacCann, C., Geiger, M., Seybert, J. M., & Roberts, R. D. (2015). Development of a forced-choice measure of typical-performance emotional intelligence. *Journal of Psychoeducational Assessment*, 33, 83–97.
<http://doi:10.1177/0734282914550387>
- 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.
<http://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.
- 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>
- Dueber, D. M., Love, A. M., Toland, M. D., & Turner, T. A. (2018). Comparison of single response format and forced-choice format instruments using Thurstonian item response theory. *Educational and Psychological Measurement*.
<http://doi:10.1177/0013164417752782>
- 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>
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian Data Analysis, Third Edition*. Boca Raton: Chapman and Hall/CRC.
- 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.
- Guenole, N., Brown, A., & Cooper, A. J. (2018). Forced-choice assessment of work-related maladaptive personality traits: Preliminary evidence from an application of thurstonian item response modeling. *Assessment*, 25, 513-526.
<http://doi:10.1177/1073191116641181>
- 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.
- Lee, P., Lee, S., & Stark, S. (2018). Examining validity evidence for multidimensional forced choice measures with different scoring approaches. *Personality and Individual Differences*, 123, 229-235.
<http://doi:10.1016/j.paid.2017.11.031>
- Maydeu-Olivares, A., & Böckenholt, U. (2005). Structural equation modeling of paired comparison and ranking data. *Psychological Methods*, 10, 285-304.
<http://doi:10.1037/1082-989X.10.3.285>
- 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)*.
- Super, D. E. (1996). A life-span, life-space approach to career development. In D. Brown, L. Brooks & Associates (eds), *Career Choice and Development*, 2nd ed. San Francisco: Jossey-Bass, pp.1978-261
- 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.

INFORMACIÓN DE CONTACTO

Servicio al cliente: atencion.cliente@centraltest.com

Servicios de formación y consultoría: consulting@centraltest.com



www.centraltest.com