

MANUAL TÉCNICO Y GUÍA DE USO **SMART**

Este manual ha sido elaborado por los psicólogos/as del departamento de Investigación y Desarrollo de Central Test.

El test **SMART** 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

1. Introducción	3
2. Descripción y uso	5
2.1. Aplicación	5
2.2. Estructura del cuestionario	6
2.3. Administración, idiomas y adaptaciones	6
2.4. Seguridad	6
2.5. Administración supervisada	7
2.6. Consideraciones éticas y técnicas.	7
2.6.1. Marco legal y ético	8
2.6.2. Privacidad y seguridad de los datos	8
2.6.3. Características del Sistema informático y disponibilidad del servidor	8
2.6.4. Informes	9
2.7. Suministro y disponibilidad	9
3. Fundamento teórico y desarrollo de SMART	10
3.1. Historia y teoría de la medición de la inteligencia	10
3.2. Evaluación de la habilidad cognitiva en el trabajo	12
3.3. Historia de los tests adaptativos informatizados (TAIs)	12
3.4. Banco de ítems calibrados	13
3.5. Teoría de Respuesta al ítem, TRI	13
3.5.1. Parámetros del ítem	13
3.5.1. a. Dificultad y discriminación del ítem	14
3.5.2. Ventajas del modelo TRI y estudios de validación	15
4. Módulos de SMART	16
4.1. SMART – LOGICAL , razonamiento lógico	16
4.1.1. Desarrollo	16
4.1.2. Baremos	17
4.1.3. Fiabilidad	20
4.1.3. a. Consistencia interna	20

4.1.4. Parámetros de los ítems	20
4.2. SMART – NUMERICAL , razonamiento numérico	22
4.2.1. Desarrollo	23
4.2.2. Baremos	24
4.2.3. Fiabilidad	26
4.2.3. a. Consistencia interna	26
4.2.4. Parámetros de los ítems	26
4.3. SMART – VERBAL , razonamiento verbal	28
4.3.1. Desarrollo	28
4.3.2. Baremos	29
4.3.3. Fiabilidad	31
4.3.3. a. Consistencia interna	32
4.3.4. Parámetros de los ítems	32
5. Puntuación	35
5.1. Puntuación y percentiles	35
5.2. Puntuación combinada	35
Referencias	37

1. Introducción

SMART es un test adaptativo informatizado (TAI) para evaluar la *inteligencia* de las personas. Está compuesto por tres módulos que evalúan la habilidad de *razonamiento lógico* (**SMART – LOGICAL**), *razonamiento verbal* (**SMART – VERBAL**) y *razonamiento numérico* (**SMART – NUMERICAL**).

SMART se realizó partiendo del *Test de Razonamiento* de Central Test. En su desarrollo se revisó el marco teórico y las diferentes conceptualizaciones de la inteligencia a lo largo de la historia de la psicología y las teorías vigentes en la actualidad.

La inteligencia es un concepto complejo, sobre el que se han realizado múltiples investigaciones. Una teoría importante es la de Charles Spearman, que consideraba que existía un factor subyacente general de habilidad cognitiva, conocido como el factor *g*. Este factor se podía medir y utilizarse para realizar comparaciones entre las personas. En la teoría de bifactorial de la inteligencia (1927), Spearman proponía la existencia de una capacidad cognitiva *general* (*g*) y una serie de habilidades cognitivas *específicas* (*s*). Las habilidades cognitivas específicas son independientes entre sí, pero todas y cada una de ellas se solapan con la habilidad cognitiva general (*g*).

En la década de 1940, Raymond Cattell indica la existencia de dos tipos de inteligencia diferentes, la inteligencia *fluida* y la *cristalizada*. La inteligencia *fluida* serían la capacidad de razonamiento básica, como la habilidad de establecer relaciones y solucionar problemas. La inteligencia *cristalizada* sería la habilidad cognitiva desarrollada en función del aprendizaje y la experiencia, incluyendo los conocimientos adquiridos. Utilizamos la inteligencia *fluida* para solucionar problemas abstractos, con los que no se está familiarizado/a, la inteligencia *cristalizada* permite resolver problemas concretos, que se basan en la experiencia y los conocimientos adquiridos (Cattell, 1963).

SMART evalúa las habilidades cognitivas específicas (*s*), según la teoría de Spearman. La puntuación obtenida en cada uno de los módulos se combina para obtener una puntuación general que se correspondería con el factor *g* de Spearman. Además, siguiendo la distinción propuesta por Cattell, **SMART** evalúa la inteligencia *fluida* en el módulo de razonamiento lógico, **SMART – LOGICAL** y la *cristalizada* en los módulos verbal y numérico; **SMART – VERBAL** y **SMART – NUMERICAL**.

El formato de **SMART** está especialmente adaptado para evaluar conceptos como la inteligencia. Hay diferentes tipos de tests para medir niveles de habilidad, por ejemplo, en el caso de la inteligencia, los *test tradicionales* consisten en una serie de preguntas fijas, con una sola respuesta correcta, para todas las personas. En función del número de preguntas contestadas de forma correcta, en un tiempo determinado, se obtiene una puntuación. Estos tests tienen problemas de seguridad, ya que las personas pueden compartir las preguntas con otras, que pueden entrenarse en ellas, invalidando los resultados por el efecto del aprendizaje. **SMART** soluciona este problema con su formato de *test adaptativo informatizado*. Este formato implica que cada persona responda a un conjunto diferente de preguntas. Los TAIs como **SMART**, evalúan la habilidad de las personas de forma precisa, ya que se ajustan a la habilidad que muestra la persona al contestar cada pregunta concreta. Cuando se responde

de forma correcta a una pregunta, la siguiente tendrá un nivel de dificultad mayor, mientras que, si la respuesta es incorrecta, se presentará una pregunta con una dificultad menor.

Este formato presenta grandes ventajas, pero su desarrollo y administración tiene grandes dificultades técnicas y psicométricas. Sin embargo, merece la pena hacerles frente, ya que su utilización tiene grandes ventajas. Por ejemplo, **SMART** se adapta a la persona, la experiencia de cada persona es única y la evaluación del nivel cognitivo se hace con mucha precisión. La calidad de la experiencia de usuario/a mejora, el test es más breve porque no se pierde el tiempo en solucionar preguntas demasiado fáciles y puede reducir la exposición a preguntas frustrantes. Este aumento en la calidad de los resultados y de la experiencia.

Desde el punto de vista técnico, la mayor ventaja de los tests adaptativos informatizados, como **SMART** es su alta seguridad y el mantenimiento de la calidad a lo largo del tiempo. Cada ítem se calibra de forma específica, lo que permite aumentar el banco de ítems de forma continua, manteniéndose la calidad y seguridad a lo largo del tiempo.

En conclusión, las ventajas de utilizar **SMART** son:

- Los ítems son precisos, coherentes en el factor evaluado, claros, equitativos y sin sesgos culturales.
- Seguridad, diferentes configuraciones de ítems para cada persona
- Menor tiempo, se contestan menos ítems que en tests convencionales.
- Amplio banco de ítems para cada nivel de dificultad
- Mayor precisión, uso de la teoría de respuesta al ítem (TRI)
- Alta calidad de la experiencia de usuario/a con informes personalizables.

En este manual y guía de uso se proporciona información útil sobre la creación y desarrollo de **SMART**, su marco teórico, características psicométricas y orientación sobre su uso y aplicación.

2. Descripción y uso

SMART es un test para evaluar la habilidad cognitiva o “inteligencia” de las personas, orientado a las necesidades específicas de las organizaciones y empresas. Los módulos de **SMART** permiten evaluar de forma objetiva las habilidades cognitivas de las personas en razonamiento *lógico, verbal y numérico*. Su uso combinado permite la obtención de una estimación de la inteligencia general. Además, sus características técnicas permiten la adición de nuevas competencias cognitivas, en el futuro, lo que permite adaptarse a las necesidades de evaluación de las competencias cognitivas a largo plazo.

Existen diferentes formas de definir la inteligencia, aunque una de las más prácticas es la definición de uno de los principales teóricos sobre la inteligencia, David Wechsler, en 1944, “la inteligencia es la capacidad total o combinada de una persona de actuar en función de objetivos, pensar de forma racional y adaptarse al ambiente”.

2.1. Aplicación

SMART es un test específico para empresas y organizaciones en la evaluación del personal o candidaturas, siendo también apto para su uso en centros educativos superiores, con fines de orientación laboral o selección.

Se ha diseñado para su uso por profesionales de recursos humanos, técnicos/as de selección y personal responsable de la evaluación del personal, con el propósito de apoyarles en sus procesos de evaluación en selección y contratación, movilidad interna y promociones, orientación laboral o programas de desarrollo del personal. Este test se puede utilizar en cualquier momento de la carrera profesional de una persona y es adecuado para todos los niveles de cualificación. En la siguiente tabla se muestran algunas de las aplicaciones de **SMART**.

Tabla 1. Ejemplos de uso de **SMART**.

Selección y contratación	Movilidad interna	Centros educativos
Comparación del perfil cognitivo de la persona con el del puesto.	Comparación del perfil cognitivo de la persona con el nuevo puesto	Comparación del perfil cognitivo del estudiante con el necesario para el módulo /tipo de estudio.
La adecuación óptima permite alta productividad del personal y la empresa.	Identificar los mejores ajustes a los proyectos o puesto-persona	Orientación profesional

SMART se puede utilizar en combinación con otros tests para obtener un perfil más completo de las personas. En la siguiente tabla se presentan algunos ejemplos de combinación.

Tabla 2. Ejemplo de combinación con otros tests de Central Test.

Test complementario	Objetivo	Ventaja
Perfil Pro 2©	Desarrollo profesional	Comparación del perfil de <i>personalidad + cognitivo</i> con el del puesto/ profesión.

2.2. Estructura del cuestionario

SMART, es un test adaptativo que se administra *online*. Cada módulo está compuesto por un banco de ítems de diferentes niveles de dificultad, para cada módulo existe un número de preguntas determinado, con límite de tiempo por pregunta. Cada persona responde a una selección de ítems particular, ya que las preguntas se seleccionan en función de si la respuesta a la anterior es correcta o no. El objetivo es presentar ítems que se ajusten progresivamente mejor al nivel cognitivo de la persona para obtener una estimación precisa de su nivel.

2.3. Administración, idiomas y adaptaciones

SMART se realiza *online* y la plataforma de administración es compatible con los sistemas operativos *Windows*, *MAC-OS* y *Linux*. Al igual que el resto de nuestros tests, su administración se realiza de forma remota y sin supervisión. En el caso de que se necesite supervisar la administración, es posible organizarlo en función de las necesidades. Debido a las características de este tipo de tests no es posible su administración en papel. Está disponible en español, francés e inglés.

SMART puede ser utilizado, aunque existan algunas necesidades perceptivas especiales. Por ejemplo, es posible ajustar la apariencia en la pantalla. Por ejemplo, si existen problemas visuales o se necesita un tamaño de fuente mayor, se puede aumentar presionando "CTRL" y "+" simultáneamente. Las instrucciones son accesibles en todo momento. Para otro tipo de ajustes, puede consultar con el departamento de atención al cliente de Central Test.

2.4. Seguridad

Los test de tipo adaptativo, como **SMART**, tienen una alta seguridad, debido a varias de sus características. En primer lugar, el banco de ítems para cada nivel de dificultad es amplio y es improbable que dos personas de un mismo nivel cognitivo contesten un test compuesto por las mismas preguntas, esto previene la posibilidad de compartir, aprender las respuestas o copiar las respuestas de otra persona. En segundo lugar, para evitar el uso sistemático de unos ítems particulares y garantizar que todos los ítems de un mismo nivel de dificultad se usen igualmente, se ha organizado su selección al azar. Estas características hacen posible que cada pregunta se utiliza solo con un grupo reducido de personas, garantizando la seguridad y la calidad del test.

Además, el test mantiene la calidad a lo largo del tiempo, se realiza el seguimiento de los ítems, comparando las dificultades de los ítems a medida que se van utilizando, garantizando una evaluación equitativa para todas las personas (Chae et al., 2000). **SMART** permite la inclusión de nuevos ítems a lo largo del tiempo, ya que sus características se evalúan de forma independiente, permitiendo la renovación y actualización del test de forma sistemática.

2.5. Administración supervisada

La administración de **SMART** es *online* y sin supervisión. *Central Test* cuenta con un sistema de supervisión *online*, puede contactar con el servicio de atención al cliente si desea más información al respecto. En el caso de realizar una administración supervisada presencial, en una localización determinada, es importante que se respeten condiciones semejantes para todos/as. Se puede tener en cuenta que:

- Las instrucciones de realización del test se presentan, antes de comenzar el test, en la pantalla. No obstante, los administradores/as pueden asegurarse de que se han entendido de forma correcta.
- Las instrucciones del test son iguales para todos/as y durante la administración supervisada es importante respetar condiciones similares. Instrucciones diferentes pueden hacer que el proceso deje de ser imparcial y equitativo para todos/as y existe el peligro de que algunas personas o grupo/s tengan ventaja frente al resto.

El enlace para realizar el test es único para cada persona y se recibe mediante correo electrónico, junto con otro enlace a la página web con la cuenta de usuario/a y la información de acceso, que es única para cada persona y está protegida mediante contraseña. En casos excepcionales, en que no se incluye el correo electrónico al crear al usuario/a, se puede generar en la cuenta cliente directamente (esta opción no se recomienda). Se pueden determinar fechas determinadas, durante las que el enlace será válido, para evitar que el proceso de evaluación se dilate innecesariamente en el tiempo. Además, para aumentar la seguridad, se puede configurar la evaluación para que se complete desde la dirección IP en que se abre por primera vez. Por último, el tiempo de realización se registra con el fin de hacer el seguimiento de los tests y permite detectar anomalías al responder el test o posibles fallos del sistema.

2.6. Consideraciones éticas y técnicas

A continuación, se presenta el marco legislativo y las actuaciones de *Central Test* para garantizar un uso adecuado de **SMART**. El uso de test psicométricos ha de ser equitativo, justo y objetivo para cualquier persona, independientemente de su grupo social o características personales. La legislación y buenas prácticas sobre protección y privacidad de la información generada garantizan el buen uso de la información durante la administración y después de la realización.

2.6.1. Marco legal y ético

Central Test es un editor de test psicométricos *online*. Su sede está en París, Francia. Su marco legal es el vigente en Francia y la Unión Europea. *Central Test* está sujeto a la ley y convenciones legales de la ley de trabajo francesa y la ley del 6 de enero de 1978 respecto de las libertades sobre los archivos e información tecnológica. Además, desde mayo de 2018, *Central Test* cumple con la RGPD, Reglamento General de Protección de Datos elaborada por el Consejo y Parlamento Europeos.

Los datos personales y los resultados de los test no pueden compartirse con terceros sin el consentimiento explícito de la persona. Las personas tienen derecho de acceso, rectificación y eliminación de sus datos y de acceso a sus resultados. La persona que administra el test debe permitirlo siempre que se solicite. Para ejercer estos derechos, la persona puede contactar con la organización que administró el test o a *Central Test* directamente.

2.6.2. Privacidad y seguridad de los datos

Central Test garantiza la seguridad y la privacidad de los datos generados durante la administración de los test *online*. Los códigos de acceso están cifrados en la base de datos y nadie tiene acceso a ellos. Las contraseñas se generan de forma automática y aleatoria por ordenador; ninguna empresa u organización tiene acceso. El sistema cuenta con la instalación de un módulo de reescritura URL para generar un enlace y un parámetro de acceso seguro para el usuario/a.

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 ©. La certificación prueba nuestro compromiso con la seguridad y la protección de las evaluaciones de *Central Test* respecto a ataques para suplantar la identidad. También se registra el tiempo de realización del test y se hace un seguimiento de la dirección IP.

2.6.3. 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 datos especializado cuyas características son:

- Conexiones privadas
- Ancho de banda de 100 Mbps (por servidor)
- Fuentes de alimentación adicional
- 4 horas de tiempo de reiniciación del equipo
- 2 horas de tiempo de intervención en caso de incidentes
- Supervisión y asistencia continuada 24h/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; Safari 7 o superior.

2.6.4. Informes

Los informes se generan *online* inmediatamente, o en pocos minutos, una vez que se ha completado el test y se han enviado las respuestas. El informe de **SMART** es claro, accesible y se puede personalizar según las necesidades. El informe predefinido de *Central Test* proporciona información útil para la interpretación y sobre el concepto y su importancia en el trabajo; la puntuación obtenida en percentiles, ya que permite compararla de forma sencilla con los resultados de las demás personas; se incluyen también comentarios de interpretación de la puntuación y consejos para el desarrollo personal. Se puede descargar en .pdf y es posible configurar versiones diferentes en función del uso, por ejemplo, uno para la empresa y otro para los candidatos/as.

2.7. Suministro y disponibilidad

Las condiciones de suministro y disponibilidad pueden variar según el país y es necesario tener la formación legal necesaria para aplicar e interpretar tests psicométricos según la legislación vigente en el país. Desde el servicio de consultoría se ofertan formaciones en uno o varios tests, que garantizan la adquisición de los conocimientos necesarios para la interpretación adecuada de los perfiles obtenidos.

SMART y formaciones asociadas pueden solicitarse por internet, en su cuenta cliente de *Central Test*, o a nuestros socios comerciales o personal de ventas, en cualquier parte del mundo.

También puede obtener información sobre nuestros precios, productos relacionados o servicios de consultoría poniéndose en contacto con el servicio de atención al cliente de *Central Test*; atencion.cliente@centraltest.com.

3. Fundamento teórico y desarrollo de SMART

En la primera parte de este capítulo se explican los modelos teóricos del concepto de inteligencia y su uso en el desarrollo de los tests **SMART**. En la segunda parte se introducen los modelos psicométricos y los parámetros analizados para la selección de los ítems.

3.1. Historia y teoría de la medición de la inteligencia

La inteligencia es la capacidad de adaptarse y aprender de la experiencia, permite solucionar problemas, entender conceptos abstractos y las consecuencias de lo que sucede. La investigación sobre el concepto y la evaluación de la inteligencia ha sido extensa a lo largo de la historia de la psicología. El punto de partida fue el estudio científico de las diferencias humanas y la inteligencia, que comenzó con Galton en 1869. Este investigador utilizó pruebas y medidas de carácter heterogéneo, conocidas como el método biométrico, con el objetivo de evaluar las habilidades motoras en una muestra muy amplia de personas. Su trabajo fue pionero en el desarrollo de los test de aptitudes y la psicometría.

Aunque la definición del concepto de inteligencia se continúa debatiendo y estudiando, existe consenso en considerarla como la habilidad para procesar y entender información nueva. A continuación, se revisan diferentes teorías que se han tenido en cuenta en el desarrollo de **SMART**.

Teoría bifactorial de la inteligencia de Charles Spearman

El psicólogo Charles Spearman definió la inteligencia general como la “habilidad para inferir relaciones y establecer correlaciones”. Introdujo la teoría bifactorial en 1927, en base a sus investigaciones sobre la naturaleza del constructo de inteligencia, utilizando el análisis estadístico de las puntuaciones en tests de IQ. Su teoría propone que todas las habilidades cognitivas diferentes tienen un factor común subyacente, conocido como “factor *g*”. Por otro lado, las habilidades cognitivas específicas son independientes entre sí y se denominan “factores *s*” (específicos).

En sus estudios, Spearman (1904), exploró las relaciones estadísticas entre las puntuaciones de diferentes tests de inteligencia y encontró que había una correlación positiva entre cualquier par de funciones mentales, atribuyó estas correlaciones a un factor diferente y común que denominó *g*, ya que los factores *s* presentaban correlaciones bajas entre ellos. En base a estos datos, Spearman postuló que el factor *g* estaba en la base de todas las habilidades cognitivas y permitía predecir el rendimiento de las personas, siendo la razón principal por la que se encuentran relaciones positivas entre las puntuaciones de los diferentes tests de inteligencia. La técnica estadística que Spearman utilizó fue el análisis factorial. Esta técnica derivó de los resultados obtenidos en su investigación sobre las habilidades cognitivas, en que, si una persona obtenía buenas puntuaciones en un test, solía obtener también buenas puntuaciones en otros tests y lo contrario; las personas con bajas puntuaciones en un test solían obtener puntuaciones bajas también en los demás tests, deduciendo el factor *g* y los factores específicos *s*.

El factor *g* también se conoce como la *habilidad mental general*. El concepto de inteligencia general o “factor G” está ampliamente aceptado en la actualidad. Los tests de inteligencia que evalúen mejor este factor general correlacionarán en mayor medida con los demás y los tests de habilidad cognitiva que implican razonamiento abstracto son los que correlacionarán en mayor medida con el factor *g* o capacidad de inteligencia general, que en la actualidad se considera que es la esencia de la inteligencia y viene determinada de forma genética mayormente (Hally, 2012). Esta asunción es aceptada de forma general, aunque aún no está probada.

Teoría de la inteligencia fluida y cristalizada de Cattell

El psicólogo Raymond Cattell propuso en 1940, una teoría sobre la inteligencia, con dos componentes diferenciados, la inteligencia fluida y la cristalizada. La “inteligencia fluida” sería la habilidad de establecer relaciones abstractas y resolver problemas generales, no se vería afectada por los factores ambientales como la calidad de la educación o el enriquecimiento del entorno de crianza. En contraste, la “inteligencia cristalizada” haría referencia a los conocimientos adquiridos y sí se vería influida por factores ambientales (Cattell, 1963).

En su libro sobre la estructura de la inteligencia (1987), *“Intelligence: Its Structure, Growth and Action”* (Inteligencia: Su estructura, desarrollo y actuación), Cattell definió la habilidad de razonamiento general como inteligencia fluida, por su flexibilidad para poder utilizarse en cualquier tipo de problema de razonamiento (Vinney, 2020). La inteligencia fluida sería el potencial de la persona independientemente de los factores ambientales o la calidad de la educación. La inteligencia cristalizada haría referencia a las habilidades adquiridas mediante aprendizaje y experiencia, se trataría del aprendizaje en diferentes ámbitos relativamente independientes unos de otros (Vinney, 2020).

La inteligencia fluida se suele evaluar mediante tests de razonamiento lógico o abstracto. La inteligencia cristalizada se evalúa mediante tests de habilidad específicos, como pueden ser los de razonamiento verbal o numérico. Aunque la inteligencia fluida y la cristalizada hacen referencia a dos tipos de inteligencia diferentes, están estrechamente relacionados. La inteligencia fluida se usa cuando aprendemos, así, cuando nos exponemos a información nueva, es la inteligencia fluida la que nos permite analizar y entender las relaciones y estructura lógica. Una vez hemos realizado el aprendizaje, este pasa a nuestra memoria a largo plazo y se “almacena como conocimiento cristalizado”. Por ejemplo, cuando aprendemos una habilidad nueva como cocinar, es la inteligencia fluida la que nos ayuda a establecer relaciones entre los ingredientes, cantidades, texturas para preparar comidas que nos gusten; pero cuando seguimos una receta concreta, es la inteligencia cristalizada la que en base a la experiencia y conocimientos que tenemos nos permite realizarla correctamente, realizar cambios o detectar errores. En un examen de matemáticas, las fórmulas y los significados de los símbolos serían cuestiones relacionadas con la inteligencia cristalizada, mientras que la estrategia para solucionar un problema que no hemos realizado previamente implicaría la inteligencia fluida (Vinney, 2020).

3.2. Evaluación de la habilidad cognitiva en el trabajo

La evaluación de la inteligencia es importante en el trabajo ya que la investigación ha demostrado que la habilidad cognitiva predice el rendimiento en el trabajo (Salgado et al (2003); Schmidt & Hunter (2004)). Es útil y relevante evaluar la habilidad cognitiva identificada en el análisis del puesto como necesaria para un adecuado desempeño. Los tests de competencia cognitiva como **SMART** permiten comparar el perfil cognitivo de la persona con el del puesto y determinar su adecuación.

En instituciones educativas permite comparar el perfil cognitivo de la persona con el de los estudios que se desea realizar, o en orientación laboral permite identificar si existe ajuste entre el perfil cognitivo y el de la profesión. La naturaleza objetiva de estas evaluaciones, permite la comparación de forma sencilla y objetiva con las personas de su grupo normativo.

3.3. Historia de los tests adaptativos informatizados (TAIs)

Los TAIs o tests adaptativos informatizados se desarrollaron principalmente a partir de las décadas de los 70 y 80 del siglo veinte, su creación fue posible gracias a la potencia creciente de los ordenadores para procesar datos, que posibilitaba análisis psicométricos más complejos, con mayor número de datos (Weiss, 1983).

La ventaja principal de los TAIs es que pueden tener en cuenta un abanico muy amplio de niveles de dificultad a la hora de evaluar competencias, permiten adaptar el nivel de dificultad eligiendo un ítem dentro del conjunto de ítems del mismo nivel y así definir el nivel de dificultad que es apropiado para una persona determinada (Meijer, 1999). En el ejército de EEUU compararon la versión adaptativa informatizada de la batería de aptitudes del ejército (ASVAB) con la versión en papel y lápiz para seleccionar a los reclutas (Gade, Johnson & Pliske, 1984), encontrando que la calidad de la versión informatizada era igual o mejor y su aplicación era más efectiva.

En el desarrollo de los TAIs fue importante contar con el software adecuado, este se caracteriza por estar compuesto de módulos que se ocupan de diferentes aspectos del test de forma independiente; MicroCAT se introdujo en 1984 por la Assessment Systems Corporation permitiendo el desarrollo de estos test y la administración de los bancos de ítems (Rocklin, Roos, Wise, & Yoes, 1996).

En la actualidad, los TAIs han acreditado sus ventajas para reducir costes, tiempo, mejorar la experiencia de uso y proporcionar evaluaciones más precisas, por lo que es lógico pensar que su uso se mantendrá o aumentará en el futuro. Thompson y Weiss (2011) establecieron los 5 componentes clave de los tests adaptativos informatizados, que se citan a continuación:

- banco de ítems calibrados por niveles
- algoritmo para el inicio del test
- algoritmo para la continuación del test
- algoritmo de cálculo de la puntuación
- criterio de finalización.

3.4. Banco de ítems calibrados

Los tests adaptativos informatizados requieren un banco de ítems calibrados suficientemente amplio. Los parámetros característicos de cada ítem se almacenan para su uso cuando sea necesario. La mayor garantía de la fiabilidad y seguridad del test es que el banco de ítems será tan amplio como sea posible (Flaughner, 2000).

En **SMART** los ítems se probaron y compararon para determinar sus parámetros. Debido a la amplitud del banco de ítems, para poder ponerlos a prueba, se crearon varias listas para cada módulo con ítems diferentes. Se utilizaron una serie de ítems "anchor ítems" como referencia para realizar las comparaciones, estos ítems aparecían al inicio de todas las listas. Se calculó la dificultad de los ítems para cada una de las listas y los "anchor ítems" se utilizaron para derivar constantes para cada lista y realizar las comparaciones (Kolen & Brennan, 1995)

3.5. Teoría de Respuesta al ítem, TRI

La teoría de respuesta al ítem (TRI) se utilizó inicialmente para calcular la puntuación de los tests de habilidad. Su utilidad y desarrollo se ha ampliado en las últimas décadas y actualmente se considera que es una mejora en relación a la Teoría clásica de los tests (TCT), tanto en la evaluación de la personalidad, como de la habilidad (Crocker & Algina, 1986). La teoría está compuesta por diferentes modelos estadísticos para explicar las relaciones entre los constructos teóricos no observables y la forma en que se manifiestan las características observadas. En psicometría, los modelos TRI buscan establecer una relación entre cada ítem específico, las respuestas de las personas y el constructo o variable que se quiere medir.

Hay 4 supuestos básicos para la evaluación de la habilidad en TRI, que, si se cumplen, permiten concluir que las diferencias observadas entre las respuestas correctas se deben a diferencias en el nivel en la variable medida:

- La *probabilidad de la respuesta correcta* aumenta a medida que el nivel en la característica aumenta.
- La *unidimensionalidad*, asume que se evalúa una única característica latente, siendo esta característica la que se muestra en las respuestas observadas para cada ítem.
- La *independencia*, asume que la respuesta a un ítem no se ve influenciada por sus respuestas en los demás.
- La *invarianza*, permite estimar los parámetros del ítem a partir de cualquier posición de la curva de respuesta al ítem y a partir de cualquier grupo de personas que hayan contestado al ítem.

3.5.1. Parámetros del ítem

Los parámetros del ítem hacen referencia a las características del ítem y su posición en un continuo que evalúa la variable a medir. Permiten hacer predicciones sobre las respuestas de las personas, asumiendo que cada respuesta en un ítem refleja el nivel latente en la variable (θ). El nivel de habilidad, θ , se puede considerar que es la probabilidad de la persona de proporcionar una respuesta correcta en un ítem determinado. Cuanto mayor sea θ , más

probable será que la persona de una respuesta correcta. Según el modelo de dos parámetros de la TRI, la probabilidad se calcula según la fórmula siguiente:

$$P(\theta) = \frac{1}{1+e^{-L}} = \frac{1}{1+e^{-a(\theta-b)}}$$

Siendo:

b el parámetro de dificultad

a el parámetro de discriminación

$L = a(\theta - b)$ la desviación logística (logit)

θ nivel de habilidad.

La representación gráfica sería la curva características del ítem (CCI). En estadística aplicada, el intervalo de θ va de -3 a +3.

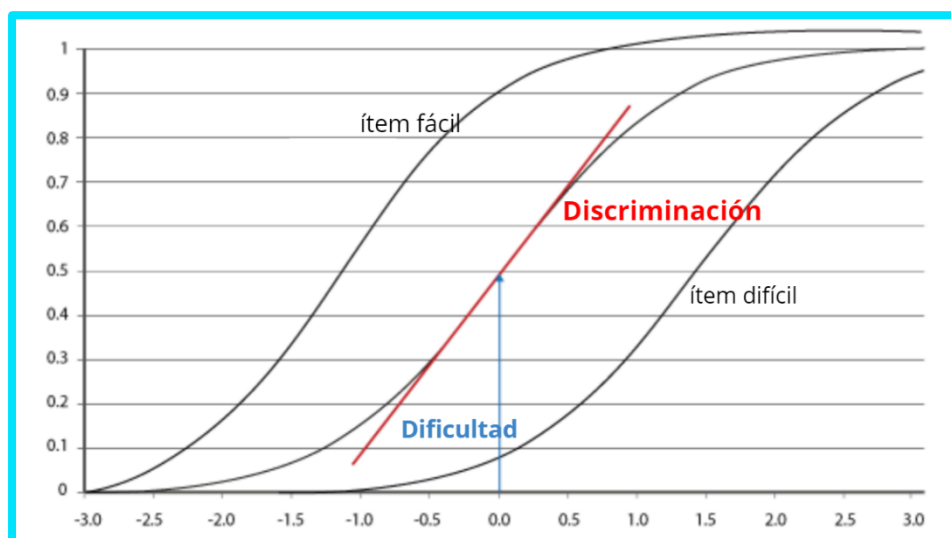


Figura 1. CCI para un ítem fácil, medio y difícil; índices de discriminación y dificultad.

3.5.1. a. Dificultad y discriminación del ítem

Según el modelo de los dos parámetros de la TRI, la curva característica se construye en función de la *dificultad* y la *discriminación* del ítem. La dificultad del ítem indica el funcionamiento del ítem en una escala de habilidad, que se calcula en base a una probabilidad media (el 50% de las personas seleccionan la opción correcta). En la curva característica del ítem, CCI, que se muestra en la figura anterior, la dificultad de los ítems es diferente en función de su posición más a la izquierda o más a la derecha. Los ítems fáciles estarán a la izquierda y los difíciles a la derecha.

La discriminación del ítem viene determinada por la probabilidad de seleccionar la respuesta correcta en función del nivel de habilidad (Baker, 2001). La discriminación es muy importante

ya que permite comparar personas con niveles semejantes en la variable medida. Es necesario tener ítems con mayores niveles de discriminación para determinar la posición de cada persona en el continuo de la variable medida. La discriminación del ítem viene determinada por la pendiente de la curva, cuanto más inclinada sea la pendiente de la curva, mayor será la discriminación del ítem.

3.5.2. Ventajas del modelo TRI y estudios de validación

La TRI presenta ventajas para mejorar la precisión y fiabilidad de los tests psicométricos. La característica principal que diferencia la TRI de otras teorías es que se pueden calcular la fiabilidad y el error de medida para cada ítem en particular, además de para el test en su conjunto. Esto proporciona mayor precisión al seleccionar los ítems y al construir los tests y reducir el número de ítems necesarios en el test que realizará cada persona (Thomas, 2011). La TRI permite utilizar información sobre el ítem "*función de información del ítem*" y del test "*función de información del test*" a la hora de tomar decisiones.

Una ventaja importante de la TRI es que permite la construcción de tests adaptativos informatizados, como **SMART**. Los tests adaptativos contruidos utilizando la TRI estiman la habilidad de la persona en función del nivel de dificultad de los ítems presentados y las respuestas correctas. Se parte de una primera pregunta seleccionada según un criterio predeterminado (en nuestro caso al azar) y en función de los criterios predefinidos para las respuestas que se dan, se van presentando unos ítems u otros. Los criterios de selección de ítems tienen el objetivo de identificar con precisión el nivel de habilidad de la persona. Este formato permite gran precisión con un número pequeño de ítems, mejorando la fiabilidad y reduciendo el tiempo de realización (Vispoel et al., 1999, Thomas, 2011). Otra ventaja de la TRI es el supuesto de invarianza de los parámetros del ítem; los parámetros del ítem no dependen de la habilidad de las personas o de los ítems que se les presentan (Bortolotti et al., 2012). La invarianza implica que los parámetros funcionan de forma semejante en cualquier población, permitiendo una medida uniforme en los diferentes grupos y la recombinación en diferentes subtests de los ítems (Magno, 2009). Por esta razón, se considera que la TRI permite inferir como se responderá a los ítems de forma más precisa que otras teorías, como la teoría clásica de los tests.

Existen estudios adicionales que ponen de relieve los beneficios de la TRI a la hora de identificar el nivel de un ítem y sus parámetros individuales. Permite una mayor precisión que otras técnicas y proporciona más información sobre las relaciones entre los ítems y las variables que miden (Ryan & Brockmann, 2009). Así mismo, mejora la precisión para identificar el funcionamiento diferencial del ítem, ya que diferencias reales en la habilidad de diferentes grupos pueden diferenciarse de prejuicios y los prejuicios en sí mismos, pueden analizarse en mayor profundidad (Thomas, 2011).

4. Módulos de SMART

SMART está compuesto por tres módulos que evalúan diferentes tipos de competencia cognitiva; lógica, numérica y verbal. En este capítulo se revisará el marco teórico subyacente a cada uno, su desarrollo específico, los baremos y la fiabilidad de cada uno de los módulos. Al final del capítulo, en el apartado 4.4, se proporciona el peso que cada uno de los módulos de **SMART** tiene en el factor general de inteligencia (*g*).

4.1. SMART – LOGICAL, razonamiento lógico

La capacidad de establecer relaciones entre diferentes elementos es el razonamiento lógico y se señala en la investigación como la más relacionada con la habilidad mental general. En las organizaciones hay muchas tareas que necesitan de un cierto nivel de competencia lógica para poder realizarlas con éxito. Algunos campos donde la capacidad de razonamiento lógico es muy importante son en legislación, tecnología, medicina, ciencias de la información...

Numerosas investigaciones muestran que la inteligencia y el rendimiento laboral están altamente relacionados (Schmidt & Hunter, 2004). La mejor forma de evaluar la inteligencia general es con tests de razonamiento abstracto, que consisten en mostrar una serie de figuras o matrices, donde se ha de identificar la que mejor encaja con el patrón. La razón principal porque estos test se utilizan de forma general y son tan populares es porque están "libre de cultura" (Cattell, 1971) o, al menos, no hay factores culturales afectando los resultados, como ocurre en otros tests de habilidades específicas, por ejemplo, numérica o verbal.

Niveles de inteligencia fluida relacionados con una mejor:

- Identificación de patrones y tendencias en información nueva
- Claridad de pensamiento y análisis adecuado de conceptos complejos
- Capacidad para encontrar soluciones lógicas a los problemas.
- Aprendizaje y comprensión de nuevas habilidades y conocimientos

Niveles de inteligencia fluida menores están relacionados con dificultades para:

- Llegar a conclusiones lógicas y/o correctas
- Tomar decisiones basadas en evidencias o en argumentación racional.
- Entender conceptos complejos y/o abstractos
- Identificar problemas y entender sus causas.

4.1.1. Desarrollo

El equipo de Investigación y Desarrollo de *Central Test* siguió las directrices de Thompson y Weiss (2011) para el desarrollo de los ítems para *razonamiento lógico* de **SMART** mencionadas en el apartado 3.3.

El primer paso consistió en crear un número suficientemente grande de ítems para evaluar el razonamiento lógico. Cada pregunta o ítem está compuesta por matrices de 3 x 3, cada

cuadrado de la matriz, excepto el del lado inferior derecho, contiene figuras situadas de una forma determinada. La distribución de las figuras se da en función de una serie de reglas. La tarea de la persona consiste en identificar cual, entre varios patrones, completa de forma correcta la matriz. La siguiente figura es un ejemplo de este tipo de ítem.

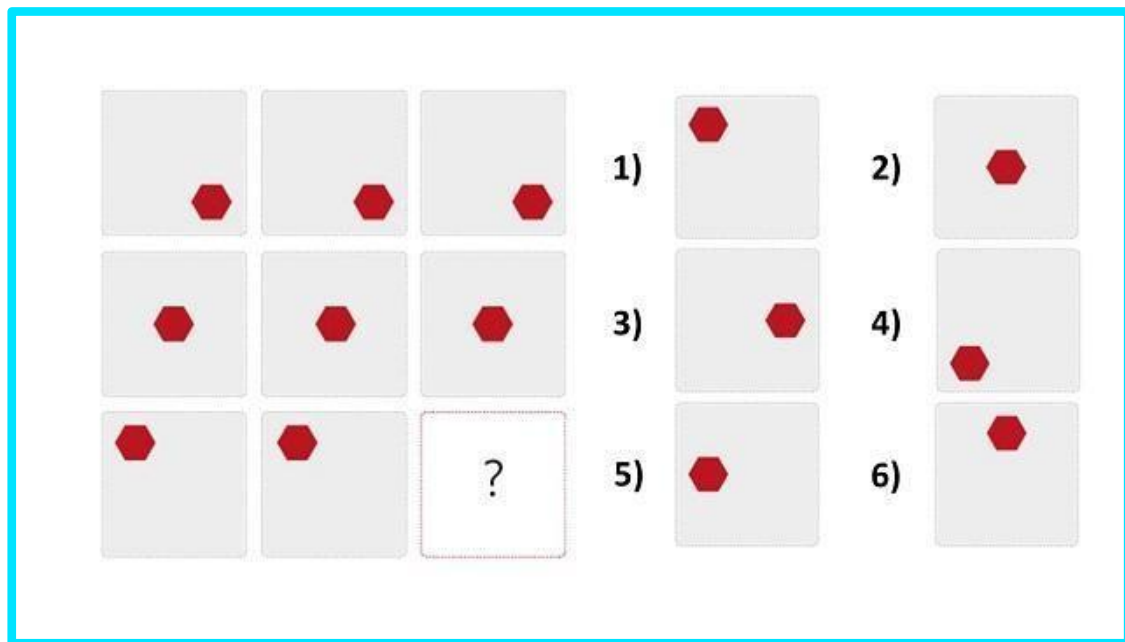


Figura 2. Ejemplo de ítem de SMART – LOGICAL.

Una vez creado el banco de ítems, se procedió a realizar la calibración de los ítems para obtener los parámetros de cada ítem que se utilizan en el test. Como se contaba con una gran cantidad de ítems no era posible probarlos en una sola vez. Se eligieron unos ítems para tener una referencia, “anchor ítems”. La selección de los ítems de referencia se realizó mediante un pretest previo para determinar qué ítems tenían una dificultad media ya que la investigación indica que este nivel de dificultad es el más apropiado cuando se realizan estudios con tests relacionados (Sinharay, Haberman, Holland, and Lewis, 2012). Se procedió a elaborar 10 subtests con todos los ítems; cada subtest estaba compuesto por 13 o 15 ítems, combinando los ítems generales del banco con los de referencia. Se administraron a un total de, N = 2137 participantes en el *Testing Center del departamento de I+D de Central Test*, con instrucciones en inglés y francés, y se analizó el funcionamiento diferencial de los ítems, se encontró que no existían diferencias en función del idioma de las instrucciones. En base a estos resultados y con la consideración de que los ítems no dependían del lenguaje, los parámetros de la TRI (teoría de respuesta al ítem) se calcularon para todos los ítems conjuntamente, independientemente del lenguaje.

4.1.2. 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 del grupo de población para el que se ha desarrollado la prueba, las personas que comparten las mismas características configuran el grupo representativo de una persona y se

llama «grupo normativo». Se pueden definir grupos representativos en función de diferentes características.

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 clasificarlas. Los parámetros estadísticos permiten a los investigadores/as que desarrollan tests y a los psicólogos/as que los utilizan hacer comparaciones entre las personas y así poder tomar decisiones basadas en información objetiva. Los baremos permiten dar significado a las puntuaciones de los test, estimar el rendimiento y comparar el nivel en un factor, característica, o rasgo en relación con el resto de la población.

En tests desarrollados específicamente para organizaciones y empresas los baremos permiten a los responsables de recursos humanos, técnicos, personal directivo, formadores/as o coordinadores/as de equipos de trabajo:

- Comprobar el ajuste persona - puesto y/o realizar las intervenciones pertinentes
- Evaluar las características de la persona en función del objetivo de la evaluación
- Identificar el ajuste potencial con el grupo de trabajo o la cultura de la organización.
- Identificar la respuesta probable ante el estrés o cambios en el ambiente de trabajo.

En la elaboración de los baremos para **SMART – LOGICAL**, se contó con una muestra de N=644 profesionales, que realizaron la versión final del test, ya como test adaptativo informatizado (TAI), en noviembre y diciembre de 2020. El test se administró a personas de diferentes partes del mundo, con instrucciones en inglés, francés y español, los ítems son semejantes para todas las personas, ya que no contienen información verbal. Los datos se analizaron para descartar que el test se hubiera realizado de forma descuidada o aleatoria. Las personas participantes provenían de la plataforma de estudios estadísticos de *Central Test, Testing Centre* y en la plataforma para estudios de mercado *CINT*. Las personas que participaron mediante el *Testing Centre* obtenían el informe, que podían utilizar con fines de desarrollo personal, mientras que los participantes de *CINT* obtuvieron una pequeña compensación económica gestionada por CINT, por el tiempo invertido.

La puntuación directa del test **SMART – LOGICAL** están entre el intervalo -3 y +3. En las tablas siguientes se proporciona la media de esta puntuación directa. Esta puntuación no se muestra en el informe por sus dificultades de interpretación, se proporciona en este manual para que se puedan comparar con los parámetros que se muestran en el apartado 4.1.4. La media para comparar con la puntuación de las personas es la que aparece en la columna "Media puntuación".

Tabla 3. SMART – LOGICAL, baremos para género.

Género	n	Media puntuación directa	Media puntuación	Desviación típica
Femenino	337	-0.69	4.7	1.01
Masculino	327	-0.76	5.4	1.02

Tabla 4. SMART – LOGICAL, baremos por edad.

Edad	n	Media puntuación directa	Media puntuación	Desviación típica
18-25	178	-0.61	5.7	1.04
26-33	170	-0.63	5.7	1.00
34-41	137	-0.79	4.9	0.97
42-49	79	-0.78	4.6	0.97
50+	100	-0.97	3.9	1.06

Tabla 5. SMART – LOGICAL, baremos por nivel de estudios.

Nivel de estudios	n	Media puntuación directa	Media puntuación	Desviación típica
Doctorado/ Master	50	-0.55	6.3	1.01
Licenciatura/ Diplomatura/ Grado	133	-0.62	5.1	0.87
Formación profesional	53	-0.67	5.3	1.01
Bachiller	43	-0.91	4.5	1.17

4.1.3. Fiabilidad

La Fiabilidad indica en qué medida el test está libre de error y si 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, interés, actitudes o habilidades de la persona. Un problema importante a tener en cuenta durante el desarrollo de 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 validez de las conclusiones en función de los resultados depende de su consistencia y precisión. Una alta fiabilidad en un test permite confiar que las decisiones se toman en función de información objetiva.

4.1.3. a. Consistencia interna

La fiabilidad es el cálculo de la precisión de la medición que hace el test. Habitualmente, se define como la relación entre la varianza de la puntuación verdadera y la varianza de la puntuación observada en el test. Por ejemplo, hasta qué punto las puntuaciones diferentes obtenidas por las personas en el test que refleja diferencias genuinas entre ellos/as. El valor de la fiabilidad se sitúa entre 0 y 1, donde 0 indica que las puntuaciones en el test no reflejan en ningún modo las puntuaciones reales y 1 indica que las puntuaciones del test reflejan completamente las puntuaciones reales.

La consistencia interna indica en qué medida los ítems de una prueba se relacionan entre ellos de forma coherente y miden el mismo factor. La consistencia interna se calcula con el Alpha de Cronbach (α), proporciona el valor de la media de todas las posibles correlaciones entre los ítems sin depender de un indicador único de fiabilidad que podría tener cantidades de error elevadas. El alpha de Cronbach va de 0 a 1, con un valor mínimo recomendado de 0.7, donde 1 indicaría una fiabilidad perfecta (Nunnally, 1978). Los valores por encima de 0.9 se consideran excelentes, valores entre 0.7 y 0.9 son buenos, entre 0.6 y 0.7 son aceptables y por debajo no se consideran aceptables.

La fiabilidad de **SMART – LOGICAL** se calculó para la muestra descrita en los apartados anteriores. Se obtuvo una fiabilidad de 0.8, lo que indica una buena consistencia interna en este módulo.

4.1.4. Parámetros de los ítems

Los parámetros de los ítems muestran las características de cada ítem en la evaluación del factor. En el caso de **SMART – LOGICAL**, se han calculado los parámetros para una escala de habilidad que va de -3 a +3. Los ítems de este módulo tienen una configuración muy similar, variando simplemente el número de elementos y características a tener en cuenta. El parámetro que más información proporciona es la dificultad de los ítems que se proporciona en la tabla a continuación. Los valores menores indican que la dificultad del ítem es baja y los valores altos que la dificultad del ítem es alta.

Tabla 6. SMART – LOGICAL, dificultad de los ítems.

Ítem	Dificultad del ítem	Ítem	Dificultad del ítem	Ítem	Dificultad del ítem
1	-1.19	35	0.519510854	68	0.326047841
2	-0.75	36	1.028810313	69	-0.857268049
3	1.04	37	1.419273231	70	-0.812881657
4	1.96	38	1.113693556	71	-0.049435725
5	0.85	39	-1.551640279	72	0.190250789
6	-3.65	40	0.366721016	73	-0.129331229
7	-0.03	41	1.478691502	74	0.39442819
8	-1.09	42	-0.22512786	75	0.021582502
9	1.13	43	-0.321358501	76	1.379806079
10	0.81	44	1.26207295	77	0.65186926
11	-1.5	45	-0.959980025	78	1.770406324
12	0.35	46	-0.942483545	79	-3.558688655
13	1.04	47	0.973381028	80	-0.987533607
14	1.04	48	-0.426337381	81	0.10564456
16	-0.501700281	49	0.579710225	82	-0.016791395
17	-0.234173829	50	-0.870144022	83	0.770296885
18	-0.121985316	51	-1.185987525	84	0.980187093
19	2.121784928	52	0.744167217	85	0.735315183
20	0.99126992	53	-0.062988402	86	0.875241989
21	-0.208284172	54	-0.856106533	87	0.131880836
22	0.741003239	55	0.659942283	88	0.770296885
23	0.775522781	56	1.102123188	89	-2.754087094
24	0.174006803	57	0.603792327	90	0.135341465

25	-1.571650387	58	0.807335918	91	-0.269883272
26	-2.25877077	59	-0.625138854	92	-1.106760446
27	-0.540969812	60	1.411519718	93	0.135341465
28	0.573279458	61	-1.71061073	94	0.276289199
29	2.272509595	62	0.348428705	95	0.478901567
30	-0.039557641	63	-0.815376193	96	1.606483444
31	1.761812013	64	1.411519718	97	1.113166373
32	1.186116557	65	1.288424969	98	0.619849302
33	-2.035474765	66	0.986283313	99	1.452916318
34	-1.016875847	67	2.351515981		

4.2. SMART – NUMERICAL, razonamiento numérico

En muchos trabajos es necesaria una buena capacidad de razonamiento numérico, siendo imprescindible tener una buena competencia para razonar con información numérica en algunos de ellos.

SMART - NUMERICAL evalúa la habilidad de razonamiento numérico, la competencia para razonar y realizar operaciones con números. En este módulo se ha considerado la habilidad de razonamiento numérico como la capacidad de establecer la secuencia adecuada y la estrategia para resolver problemas, en vez de evaluar simplemente cálculo mental.

Niveles altos de habilidad numérica permiten:

- Sacar conclusiones adecuadas de bases de datos numéricos.
- Facilidad y seguridad al utilizar números y gráficos.
- Facilidad para identificar errores en el razonamiento matemático.
- Entender e interpretar adecuadamente las conclusiones basadas en información numérica.

Niveles bajos de habilidad numérica hacen más difícil:

1. Calcular e interpretar bases de datos
2. Realizar correctamente tareas que impliquen transformar o analizar datos.
3. Entender las consecuencias derivadas del razonamiento con datos numéricos.

4.2.1. Desarrollo

El equipo de Investigación y Desarrollo de *Central Test* siguió las directrices de Thompson y Weiss (2011) para el desarrollo de los ítems para *razonamiento numérico* de **SMART** mencionadas en el apartado 3.3.

El primer paso consistió en crear un banco de ítems suficientemente amplio para evaluar el razonamiento numérico para cada idioma. Cada pregunta consiste en un problema matemático con cuatro alternativas de respuesta, de las cuales solo una es correcta. Los problemas implican razonar con números, conocimientos sobre las relaciones entre los números, conversiones y operaciones matemáticas. En el ejemplo a continuación, se puede ver la estructura de los ítems.

1	Fernando tiene \$1200 cuando va al casino. Si gasta la mitad de su dinero cada hora. ¿Cuánto dinero le queda después de 4 horas?
a)	\$ 150
b)	\$ 50
c)	\$ 100
d)	\$ 75

Figura 3. Ejemplo de ítem de **SMART – NUMERICAL**.

Una vez creado el banco de ítems, se procedió a realizar la calibración de los ítems para obtener los parámetros de cada ítem que se utilizan en el test. Como se contaba con una gran cantidad de ítems no era posible probarlos en una sola vez. Se eligieron unos ítems para tener una referencia, “anchor ítems”. La selección de los ítems de referencia se realizó mediante un pretest previo para determinar qué ítems tenían una dificultad media, ya que la investigación indica que este nivel de dificultad es el más apropiado cuando se realizan estudios con tests relacionados (Sinharay, Haberman, Holland, and Lewis, 2012).

En total, se crearon 12 tests diferentes con los ítems de razonamiento numérico de referencia y generales, cada test contenía 12 ítems y se probó con al menos 300 personas. En la calibración de los ítems de **SMART – NUMERICAL** participaron más de 3000 personas. La administración de los tests tuvo lugar desde *julio de 2019 a octubre de 2020*. Las personas participantes provenían de la plataforma de estudios estadísticos de *Central Test*, *Testing Centre* y en la plataforma para estudios de mercado *CINT*. Las personas que participaron mediante el *Testing Centre* obtenían sus resultados, por lo que podían utilizar el informe con fines de desarrollo personal, mientras que los participantes de *CINT* obtuvieron una pequeña compensación económica gestionada por *CINT*, por el tiempo invertido. Con estos datos se analizaron las propiedades de los ítems y si alguno presentaba funcionamiento diferencial FDI.

4.2.2. 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 del grupo de población para el que se ha desarrollado la prueba, las personas que comparten las mismas características configuran el grupo representativo de una persona y se llama «grupo normativo». Se pueden definir grupos representativos en función de diferentes características.

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 clasificarlas. Los parámetros estadísticos permiten a los investigadores/as que desarrollan tests y a los psicólogos/as que los utilizan hacer comparaciones entre las personas y así poder tomar decisiones basadas en información objetiva. Los baremos permiten dar significado a las puntuaciones de los test, estimar el rendimiento y comparar el nivel en un factor, característica, o rasgo en relación con el resto de la población.

En tests desarrollados específicamente para organizaciones y empresas los baremos permiten a los responsables de recursos humanos, técnicos, personal directivo, formadores/as o coordinadores/as de equipos de trabajo:

- Comprobar el ajuste persona - puesto y/o realizar las intervenciones pertinentes
- Evaluar las características de la persona en función del objetivo de la evaluación
- Identificar el ajuste potencial con el grupo de trabajo o la cultura de la organización.
- Identificar la respuesta probable ante el estrés o cambios en el ambiente de trabajo.

En la elaboración de los baremos para **SMART – NUMERICAL**, se contó con una muestra de N=322 profesionales cuya lengua principal es el español, que realizaron la versión final del test, ya como test adaptativo informatizado (TAI), en noviembre y diciembre de 2020. Los datos se analizaron para descartar que el test se hubiera realizado de forma descuidada o aleatoria. Las personas participantes provenían de la plataforma de estudios estadísticos de *Central Test, Testing Centre* y en la plataforma para estudios de mercado *CINT*. Las personas que participaron mediante el *Testing Centre* obtenían el informe, que podían utilizar con fines de desarrollo personal, mientras que los participantes de *CINT* obtuvieron una pequeña compensación económica gestionada por CINT, por el tiempo invertido.

La puntuación directa del test **SMART – NUMERICAL** están entre el intervalo -3 y +3. En las tablas siguientes se proporciona la media de esta puntuación directa. Esta puntuación no se muestra en el informe por sus dificultades de interpretación, se proporciona en este manual para que se puedan comparar con los parámetros que se muestran en el apartado 4.2.4. La media para comparar con la puntuación de las personas es la que aparece en la columna “Media puntuación”.

Tabla 7. SMART – NUMERICAL, baremos por género.

Género	n	Media puntuación directa	Media puntuación	Desviación típica
Femenino	169	-0.44	5.0	2.65
Masculino	153	-0.43	5.0	2.61

Tabla 8. SMART – NUMERICAL, baremos por edad.

Edad	n	Media puntuación directa	Media puntuación	Desviación típica
18-25	94	0.15	5.5	2.60
26-33	70	-0.64	3.9	2.58
34-41	50	-0.69	4.3	2.47
42-49	30	-0.10	4.9	2.60
50+	39	0.25	5.9	2.39

Tabla 9. SMART – NUMERICAL, por nivel de estudios.

Nivel de estudios	n	Media puntuación directa	Media puntuación	Desviación típica
Doctorado/ Master	17	0.17	5.8	2.71
Licenciatura o Diplomatura/ Grado	57	0.15	5.8	2.67
Formación profesional	30	0.15	5.9	2.59
Bachiller	40	-0.24	5.0	2.49

4.2.3. 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 cuestión más importante a tener en cuenta en el desarrollo de 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 permite confiar que las decisiones se toman en función de información objetiva.

4.2.3. a. Consistencia interna

La fiabilidad es el cálculo de la precisión de la medición que hace el test. Habitualmente, se define como la relación entre la varianza de la puntuación verdadera y la varianza de la puntuación observada en el test. Por ejemplo, hasta qué punto las puntuaciones diferentes obtenidas por las personas en el test que refleja diferencias genuinas entre ellos/as. El valor de la fiabilidad se sitúa entre 0 y 1, donde 0 indica que las puntuaciones en el test no reflejan en ningún modo las puntuaciones reales y 1 indica que las puntuaciones del test reflejan completamente las puntuaciones reales.

La consistencia interna indica en qué medida los ítems de una prueba se relacionan entre ellos de forma coherente y miden el mismo factor. La consistencia interna se calcula con el Alpha de Cronbach (α) para cada factor y/o dimensión, proporciona el valor de la media de todas las posibles correlaciones entre los ítems sin depender de un indicador único de fiabilidad que podría tener cantidades de error elevadas. El alpha de Cronbach va de 0 a 1, con un valor mínimo recomendado de 0.7, donde 1 indicaría una fiabilidad perfecta (Nunnally, 1978). Los valores por encima de 0.9 se consideran excelentes, valores entre 0.7 y 0.9 son buenos, entre 0.6 y 0.7 son aceptables y por debajo no se consideran aceptables.

La fiabilidad de **SMART – NUMERICAL** se calculó para la muestra descrita en los apartados anteriores. Se obtuvo una fiabilidad de 0.91, lo que indica una excelente consistencia interna en este módulo.

4.2.4. Parámetros de los ítems

Los parámetros de los ítems muestran las características de cada ítem en la evaluación del factor. En el caso de **SMART – NUMERICAL**, se han calculado los parámetros para una escala de habilidad que va de -3 a +3. Los ítems de este módulo son complejos ya que el éxito al contestarlos depende de diferentes tipos de conocimientos sobre la misma temática, es importante tener en cuenta tanto la discriminación como la dificultad de cada ítem que se presentan en la tabla a continuación. Los valores menores indican que la dificultad del ítem es baja y los valores altos que la dificultad del ítem es alta. La discriminación indica en qué medida el ítem diferencia entre nivel en la habilidad de las personas.

Tabla 10. SMART – NUMERICAL, parámetros de los ítems.

Ítem	Discriminación del ítem	Dificultad del ítem	Ítem	Discriminación del ítem	Dificultad del ítem
1	1.004779892	-0.812672804	39	0.869996123	-0.788301129
2	0.548959723	0.227851212	40	0.798383847	-0.321956873
3	1.087215762	-1.196688319	41	1.006840691	-1.50930649
4	2.099082437	-2.148704361	42	1.38584914	0.065578199
5	0.753447533	-0.487412037	43	1.558747412	-0.085750703
6	0.453279023	-0.587770998	44	0.627468411	-1.183273747
7	1.124090509	0.187716735	45	2.352806742	-2.082431289
8	2.148427545	0.945276779	47	1.421643887	0.24181745
9	1.33168022	0.901055333	49	0.650652923	-0.936304313
10	0.742661279	0.005816289	51	0.437369543	-0.227357659
11	0.537310712	-0.908441193	52	0.036749615	-1.386718526
12	0.48891785	0.021478671	53	0.915919367	-0.541245492
13	1.164390745	-0.431659846	55	0.740094619	1.177028456
14	1.040602416	0.071119825	58	1.450727121	0.858877511
15	1.749875521	0.97417084	59	1.538742856	-1.147650031
16	1.627113799	0.369287565	60	0.518081289	-1.33457542
17	0.985835012	-0.911268806	61	1.43753855	-0.679448472
18	1.597878218	2.521190983	63	0.560278966	-0.833694784
19	1.400964809	0.531933359	66	0.523729164	-0.727789469
20	0.812879071	-0.021355712	69	0.568844427	-0.401752175
21	1.956546249	0.129085717	71	0.25673564	-1.532792592
22	0.665773391	-1.642713164	72	0.855936302	-0.314601222
23	2.359044958	0.867063415	76	0.89845636	-0.284407092
24	0.822592253	-0.342209005	77	0.545768752	-0.652886141
25	0.572436356	0.313792943	78	0.596180972	-0.691633728
26	0.841350337	-0.177029034	79	0.54234389	-0.311740266
27	0.454420695	-0.483183278	80	1.011432718	-0.133664295

29	0.313771527	-1.207797291	81	0.801108894	0.088774395
30	1.102159254	0.375609573	82	0.851281697	-0.222310614
32	1.088039039	-0.482129594	84	0.604392106	-0.521382447
33	0.634868301	0.133676408	85	0.981072939	-1.112416396
34	2.22965507	0.71349082	86	0.400205749	-0.668242556
36	0.767049553	0.104174055	88	0.72692394	-1.767151217
37	1.618402622	0.74486841	89	0.94682956	-0.179646295
38	0.852187119	0.00267569			

4.3. SMART – VERBAL, razonamiento verbal

SMART - VERBAL evalúa la habilidad cognitiva de razonamiento verbal, que consiste en entender y utilizar correctamente las relaciones entre las palabras y la información presentada de forma verbal. **SMART – VERBAL** evalúa el razonamiento verbal mediante analogías, que requieren que se entiendan las relaciones lógicas entre conceptos y significados, para establecer la semejanza o diferencia entre los conceptos.

Niveles elevados de razonamiento verbal se asocian con:

- Alta competencia para comunicarse y mayor facilidad para presentar la información de forma convincente o persuasiva.
- Mayor facilidad para utilizar argumentaciones complejas y aprender nuevos conceptos.
- Mayor facilidad para identificar los errores de razonamiento verbal o en las argumentaciones.
- Facilidad para entender e interpretar las consecuencias de los argumentos que se presentan de forma verbal.

Niveles bajos de razonamiento verbal pueden suponer una mayor dificultad para:

1. Comunicarse de forma eficaz, o para convencer y persuadir a los demás.
2. Realizar tareas que impliquen entender argumentos complicados o aprender nuevos conceptos.
3. Trabajar de forma precisa y eficaz con información verbal.
4. Entender las consecuencias de la información presentada de forma verbal.

4.3.1. Desarrollo

Durante el desarrollo de los ítems de **SMART – VERBAL**, el equipo de I+D siguió las directrices de Thompson & Weiss (2011), explicadas en el apartado 3.3.

El primer paso consiste en crear una gran cantidad de ítems de razonamiento verbal. El ejemplo de la figura, a continuación, muestra el formato. Se evalúa la habilidad cognitiva verbal mediante analogías que implican poner en relación los elementos verbales e identificar las semejanzas o diferencias entre ellos.

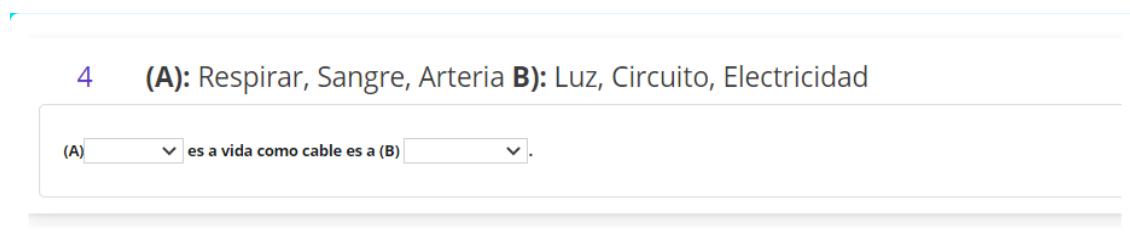


*Figura 4. Ejemplo de ítem de **SMART – VERBAL**.*

Una vez creado el banco de ítems, se procedió a realizar la calibración de los ítems para obtener los parámetros de cada ítem que se utilizan en el test. Como se contaba con una gran cantidad de ítems no era posible probarlos en una sola vez. Se eligieron unos ítems para tener una referencia, "anchor ítems". La selección de los ítems de referencia se realizó mediante un pretest previo para determinar qué ítems tenían una dificultad media ya que la investigación indica que este nivel de dificultad es el más apropiado cuando se realizan estudios con tests relacionados (Sinharay, Haberman, Holland, and Lewis, 2012).

En total, se crearon 10 tests diferentes con los ítems de razonamiento numérico de referencia y generales, cada uno contenía 15 ítems y se probó con al menos 300 personas. En la calibración de los ítems de **SMART – VERBAL** participaron más de 3000 personas, y la administración de los tests tuvo lugar desde *julio de 2019 a octubre de 2020*. Las personas participantes provenían de la plataforma de estudios estadísticos de *Central Test, Testing Centre* y en la plataforma para estudios de mercado *CINT*. Las personas que participaron mediante el *Testing Centre* obtenían sus resultados, por lo que podían utilizar el informe con fines de desarrollo personal, mientras que los participantes de *CINT* obtuvieron una pequeña compensación económica gestionada por CINT, por el tiempo invertido. Con estos datos se analizaron las propiedades de los ítems y si alguno presentaba funcionamiento diferencial FDI.

4.3.2. 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 del grupo de población para el que se ha desarrollado la prueba, las personas que comparten las mismas características configuran el grupo representativo de una persona y se llama «grupo normativo». Se pueden definir grupos representativos en función de diferentes características.

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 clasificarlas. Los parámetros estadísticos permiten a los investigadores/as que desarrollan tests y a los psicólogos/as que los utilizan hacer comparaciones entre las personas y así poder tomar decisiones basadas en información objetiva. Los baremos permiten dar significado a las

puntuaciones de los test, estimar el rendimiento y comparar el nivel en un factor, característica, o rasgo en relación con el resto de la población.

En tests desarrollados específicamente para organizaciones y empresas los baremos permiten a los responsables de recursos humanos, técnicos, personal directivo, formadores/as o coordinadores/as de equipos de trabajo:

- Comprobar el ajuste persona - puesto y/o realizar las intervenciones pertinentes
- Evaluar las características de la persona en función del objetivo de la evaluación
- Identificar el ajuste potencial con el grupo de trabajo o la cultura de la organización.
- Identificar la respuesta probable ante el estrés o cambios en el ambiente de trabajo.

En la elaboración de los baremos para **SMART – VERBAL**, se contó con una muestra de N=366 profesionales cuya lengua principal es el español, que realizaron la versión final del test, ya como test adaptativo informatizado (TAI), en noviembre y diciembre de 2020. Los datos se analizaron para descartar que el test se hubiera realizado de forma descuidada o aleatoria. Las personas participantes provenían de la plataforma de estudios estadísticos de *Central Test, Testing Centre* y en la plataforma para estudios de mercado *CINT*. Las personas que participaron mediante el *Testing Centre* obtenían el informe, que podían utilizar con fines de desarrollo personal, mientras que los participantes de *CINT* obtuvieron una pequeña compensación económica gestionada por *CINT*, por el tiempo invertido.

La puntuación directa del test **SMART – VERBAL** están entre el intervalo -3 y +3. En las tablas siguientes se proporciona la media de esta puntuación directa. Esta puntuación no se muestra en el informe por sus dificultades de interpretación, se proporciona en este manual para que se puedan comparar con los parámetros que se muestran en el apartado 4.3.4. La media para comparar con la puntuación de las personas es la que aparece en la columna "Media puntuación".

Tabla 11. SMART – VERBAL, baremos por género.

Género	n	Media puntuación directa	Media puntuación	Desviación típica
Femenino	192	-0.44	5.1	1.25
Masculino	174	-0.36	5.0	1.25

Tabla 12. **SMART – VERBAL**, baremos por edad.

Edad	n	Media puntuación directa	Media puntuación	Desviación típica
18-25	89	-0.24	5.1	1.19
26-33	100	-0.24	5.3	1.07
34-41	65	-0.24	5.4	1.10
42-49	38	-0.32	4.8	0.97
50+	32	-0.48	4.3	1.37

Tabla 13. **SMART – VERBAL**, baremos por el nivel de estudios.

Nivel de estudios	n	Media puntuación directa	Media puntuación	Desviación típica
Doctorado/ Master	26	0.05	6.6	0.94
Licenciatura / Diplomatura/ Grado	87	-0.06	5.6	1.14
Formación profesional	32	-0.54	4.7	1.18
Bachiller	49	-0.23	4.3	1.02

4.3.3. Fiabilidad

La Fiabilidad indica en qué medida el test está libre de error y si 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 cuestión más importante a tener en cuenta en el desarrollo de 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 permite confiar que las decisiones se toman en función de información objetiva.

4.3.3. a. Consistencia interna

La fiabilidad es el cálculo de la precisión de la medición que hace el test. Habitualmente, se define como la relación entre la varianza de la puntuación verdadera y la varianza de la puntuación observada en el test. Por ejemplo, hasta qué punto las puntuaciones diferentes obtenidas por las personas en el test que refleja diferencias genuinas entre ellos/as. El valor de la fiabilidad se sitúa entre 0 y 1, donde 0 indica que las puntuaciones en el test no reflejan en ningún modo las puntuaciones reales y 1 indica que las puntuaciones del test reflejan completamente las puntuaciones reales.

La consistencia interna indica en qué medida los ítems de una prueba se relacionan entre ellos de forma coherente y miden el mismo factor. La consistencia interna se calcula con el Alpha de Cronbach (α) para cada factor y/o dimensión, proporciona el valor de la media de todas las posibles correlaciones entre los ítems sin depender de un indicador único de fiabilidad que podría tener cantidades de error elevadas. El alpha de Cronbach va de 0 a 1, con un valor mínimo recomendado de 0.7, donde 1 indicaría una fiabilidad perfecta (Nunnally, 1978). Los valores por encima de 0.9 se consideran excelentes, valores entre 0.7 y 0.9 son buenos, entre 0.6 y 0.7 son aceptables y por debajo no se consideran aceptables.

La fiabilidad de **SMART – VERBAL** se calculó para la muestra descrita en los apartados anteriores, cuyo idioma era el español, entre 18 y 65 años. Se obtuvo una fiabilidad de 0.78, lo que indica una buena consistencia interna en este módulo.

4.3.4. Parámetros de los ítems

Los parámetros de los ítems muestran las características de cada ítem en la evaluación del factor. En el caso de **SMART – VERBAL**, se han calculado los parámetros para una escala de habilidad que va de -3 a +3. Contestar con éxito los ítems de este módulo requiere entender las relaciones lógicas entre los conceptos, pero también depende de los conocimientos, por ello, es importante tener en cuenta tanto la dificultad como la discriminación de cada ítem para poder entender sus características. Los valores para cada ítem se indican a continuación. Valores menores indican que la dificultad del ítem es baja y los valores altos que la dificultad del ítem es alta. La discriminación indica en qué medida el ítem diferencia el nivel en la habilidad de las personas.

Tabla 14. **SMART – VERBAL**, parámetros de los ítems.

Ítem	Discriminación del ítem	Dificultad del ítem	Ítem	Discriminación del ítem	Dificultad del ítem
1	0.9984459059	0.6253175639	54	0.3608929934	0.1542563602
2	1.222905322	0.589528174	56	1.048444527	1.111835179
3	1.451304899	0.4317329595	57	0.8426607723	-0.2895166129

4	1.057947374	0.6349013215	58	0.8350417366	-0.4333315416
5	1.220332446	1.366858913	60	0.6117325617	-3.05754589
7	1.362836966	0.1699263547	61	1.10030713	-1.020320439
8	1.244054198	-0.6560076992	62	1.429378948	-0.5445099824
9	1.06245074	-1.502095731	63	1.079242921	2.180156971
10	0.90412936	0.5102536363	64	0.7225559527	3.413606212
11	1.158872909	0.5688042767	65	1.048444527	1.111835179
12	0.8513845157	-1.626386957	66	1.239128604	0.8685695073
13	0.874350066	-0.9924042223	67	0.9850476034	-1.809800772
14	1.813588601	1.492510752	68	0.5286739356	-1.590572757
16	0.6438830845	-3.125290183	70	0.2172134549	-2.611963689
17	1.205979431	-2.933652031	71	0.9985905435	2.074489117
18	0.9192465325	-1.153892489	72	0.4576776323	1.940482454
20	0.2645358898	-0.3449732507	73	0.2393732767	2.770278474
21	0.6591252926	0.471871365	74	0.4831476902	-0.1755989281
22	0.3643536586	1.096499106	75	1.239128604	0.8685695073
23	0.266952469	-3.428886583	78	0.7378654397	-3.655330394
24	0.444492811	2.32535302	79	0.6645078431	0.3906852862
27	1.350960625	-0.1183153853	80	1.002510018	-0.4411044104
28	1.044313979	1.709516073	81	0.737947638	-2.907103578
29	1.06085645	-0.5415917329	82	0.3155597061	-2.758284846
30	0.9015562276	0.3002296656	87	0.3066154929	-0.5781682163
31	0.2431018418	-1.312991656	88	0.7556980464	0.03691063132
32	0.6433372836	0.13762667	92	0.3663874939	-1.201336476
35	1.289319389	-1.663960485	93	0.446660991	2.421763961
36	0.6946552855	-0.8319401985	96	1.922375083	2.57125567

37	0.6907497382	2.664153061	97	0.6304095821	-3.264919927
38	1.1116376	1.892533615	98	1.219584077	-0.4998510284
39	0.4313459089	2.751039759	99	1.736931384	0.1330447586
40	1.010902234	1.932820414	100	1.020996732	2.333274577
42	0.702993946	0.5211247837	101	0.6721374915	2.587880014
47	0.5934595505	-1.387732557	102	1.200657041	0.3160543529
48	0.3414803478	-2.407756748	103	0.9077155558	4.613061308
49	0.288161468	5.506940575	105	1.922375083	2.57125567
50	0.4772274191	0.9243053515	106	0.9984459059	0.6253175639
51	0.3480914317	2.750418005			

5. Puntuación

En los tests de tipo adaptativo la puntuación bruta proporciona muy poca información en sí misma, porque cada persona responde preguntas diferentes, por esta razón en **SMART** se ha optado por facilitar la comparación de las puntuaciones obtenidas por las personas.

5.1. Puntuación y percentiles

SMART presenta sus puntuaciones en los informes en percentiles para permitir la interpretación sencilla de lo que significa la puntuación de la persona y poder compararla con su grupo de referencia.

Se obtiene una puntuación de 0 a 10, que es el percentil y muestra el nivel de habilidad de la persona en comparación con su grupo de referencia, lo que permite dotar de significado a los resultados. Además, en el informe se presenta un gráfico con el percentil en una escala de 0 a 100. Por ejemplo, una puntuación de 7.4 se corresponde con un percentil de 74, que significa que esta persona tiene un nivel de habilidad superior al 74% de las personas de su grupo de referencia. El texto de interpretación de las puntuaciones en cada módulo **SMART** están en función de esta puntuación u explican lo que implica el nivel de habilidad obtenido. Los textos de interpretación están organizados en 4 niveles. Por ejemplo, percentiles bajos estarán acompañados de textos que indican que la persona puede tener dificultades para realizar las tareas de razonamiento evaluadas en el módulo correspondientes, mientras que percentiles altos tendrán un texto de interpretación que indicará que la persona puede realizar estas tareas de razonamiento con facilidad.

5.2. Puntuación combinada

Si se contestan los tres módulos **SMART** se obtiene una puntuación combinada que refleja la inteligencia general de la persona y se muestra en la página del candidato/a en la cuenta cliente conocido como "Talent Insights Report".

El cálculo de la puntuación combinada de **SMART** tiene en cuenta el peso (o importancia) de cada uno de los tipos de razonamiento en el factor general (g). El peso del factor en g indica la correlación del módulo con el factor general, se puede consultar la revisión del marco teórico sobre el factor g en el *apartado 3.1* de este manual. El peso en el factor g se puede calcular mediante un análisis factorial de una muestra de participantes o consultando la literatura científica, que indica cual el peso del factor de cada tipo de razonamiento.

La correlación más alta es para los tests de razonamiento abstracto, ya que están libres de cultura (Colom et al., 2004). Los tests que evalúan características cognitivas específicas, donde el papel de la cultura y de la experiencia individual es más importante, tendrán una carga en g menor.

Tabla 15. Peso en G de diferentes tipos de razonamiento.

Tipo de test cognitivo	Peso en G
Razonamiento numérico	0.44
Razonamiento verbal	0.56
Razonamiento lógico	0.71

Referencias

- Baker, F. (2001). *The Basics of Item Response Theory (2nd ed.)*. College Park, MD: ERIC Clearinghouse on Assessment and Evaluation.
- Bortolotti, S., Tezza, R., de Andrade, D., Tezza, R., & de Sousa Junior, A. (2012). Relevance and advantages of using the item response theory. *Qual Quant* 47, 2341–2360.
- Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, 54(1), 1-22. <http://dx.doi.org/10.1037/h0046743>
- Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Boston: Houghton Mifflin.
- Cattell, R. B. (1987). *Intelligence: Its Structure, Growth, and Action*. Elsevier Science Publishers.
- Chae, Sunhee, Kang, Unson, Jeon, Eunhwa, & Linacre, J. M. (2000). Development of Computerized Middle School Achievement Test. *MESA Memorandum No. 69*. Seoul, South Korea: Komesa Press.
- Colom, R., Rebollo, I., Palacios, A., Juan-Espinosa, M., & Kyllonen, P. (2004). Working memory is (almost) perfectly predicted by *g*. *Intelligence*, 32(1), 277-296. <https://doi.org/10.1016/j.intell.2003.12.002>
- Crocker, L., and Algina, J. (1986). *Introduction to Classical and Modern Test Theory*. Orlando, FL: Harcourt.
- Flaughner, R. (2000). Item Pools. In Wainer, H. (Ed.) *Computerized adaptive testing: A Primer*. Mahwah, NJ: Erlbaum.
- Galton, F. (1869). *Hereditary Genius*. London: MacMillan.
- Gade, P. A., Johnson, R. M., & Pliske, R. M. (1984). Cross-Validation of the Computerized Adaptive Screening Test (CAST). *Research Report 1372*. U. S. Army Research Institute for the Behavioral and Social Sciences.
- Hally, T. (2012). *The g-factor*. https://www.researchgate.net/publication/275659645_The_g-Factor
- Magno, C. (2009). Demonstrating the difference between Classical Test Theory and Item Response Theory using derived test data. *International Journal of Educational and Psychological Assessment*, 1(1), 1-11.
- Meijer, R. R. (1999). Computerized Adaptive Testing: Overview and Introduction. *Applied Psychological Measurement*. 23 (3), 187-194.
- Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw-Hill.

- Rocklin, T. R., Roos, L. L., Wise, S. L., & Yoes, M. E. (1996). Conducting Self-Adapted Testing Using MicroCAT. *Educational and Psychological Measurement*, 56 (5), 821-827.
- Ryan, J., & Brockmann, F. (2009). *A Practitioner's Introduction to Equating with Primers on Classical Test Theory and Item Response Theory*. Retrieved from <https://eric.ed.gov/?id=ED544690>
- Salgado, J.F., Anderson, N., Moscoso, S., Bertua, C., de Fruyt, F. & Rolland, J.P. (2003) A Meta-analytic Study of General Mental Ability Validity for Different Occupations in the European community. *Journal of Applied Psychology*, 88 (1068-1081)
- Schmidt, F. L., & Hunter, J. (2004). General Mental Ability in the World of Work: Occupational Attainment and Job Performance. *Journal of Personality and Social Psychology*, 86 (1), 162-173.
- Sinharay, S., Haberman, S., Holland, P., & Lewis, C. (2012). A Note on the Choice of an Anchor Test in Equating. *ETS research report series*.
- Spearman, C. (1904). 'General Intelligence': Objectively Determined and Measured. *American Journal of Psychology*, 15, 201-292.
- Spearman, C. (1927). *The abilities of man*. London: Macmillan.
- Thomas, M. (2011). The Value of Item Response Theory in Clinical Assessment: A Review. *Assessment*, 18(3), 291-307.
- Thompson, N. A., & Weiss, D. J. (2011). A framework for the development of computerized adaptive tests. *Practical Assessment, Research & Evaluation*. 16 (1)
- Vinney, C. (2020). *Fluid Versus Crystallized Intelligence: What's the Difference?* <https://www.thoughtco.com/fluid-crystallized-intelligence-4172807>
- Vispoel, W., Rocklin, T., Wang, T., & Bleiler, T. (1999). Can Examinees Use a Review Option to Obtain Positively Biased Ability Estimates on a Computerized Adaptive Test? *Journal of Educational Measurement*, 36(2), 141-157.
- Wechsler, D. (1944). *The measurement of adult intelligence (3rd ed.)*. Baltimore: Williams & Wilkins.
- Weiss, D. J. (1983). *New Horizons in Testing: Latent Trait Test Theory and Computerized Adaptive Testing*.

INFORMACIÓN DE CONTACTO

Servicio de atención al cliente: **atención.cliente@centraltest.com**

Servicios de formación y consultoría : **consulting@centraltest.com**



www.centraltest.es