

Tema 5: Introducción a los tests informatizados (TIs).

Licenciatura de Psicología:
*Desarrollos actuales de la medición:
Aplicaciones en evaluación psicológica.*
José Antonio Pérez Gil
Dpto. de Psicología Experimental.
Universidad de Sevilla.

Agradecimientos: a M^a José Díaz Domínguez y Ángela Baña Moreno.

Tema 5

Unidimensionalidad: Definición y Evaluación.

1. Introducción.
- 2.-Procedimiento general de construcción de una prueba o test.
- 3.- La construcción del test desde las diferentes teorías.
4. Estudio sobre los test que se ofertan
5. Bibliografía

1. INTRODUCCIÓN

Con la denominación de test informatizados hacemos referencia a test que, existiendo en formato de papel y lápiz, se trasladan a soporte informático, siendo presentados los mismos ítems y con la misma secuencia de presentación a todos los sujetos. Por ello se consideran como “primera generación” de test informatizados existiendo además los test informatizados de procesos cognitivos que, no siendo convencionales tienen de común con estos que se presentan completamente, es decir mediante procedimientos no adaptativos.

No obstante se hace necesario distinguir entre test informatizados y test adaptativos informatizados, para ello haremos uso de lo que James Dean Brown nos dice en su artículo “Computers in Language Testing: Present research and some future directions”. Según él, los TAIS son un subtipo de los tests informatizados, lo que los hacen diferentes son tres características:

1. En los TAIS, los ítems son seleccionados y se presentan al candidato según el nivel.
2. El test se acaba justo cuando se ha especificado el nivel del alumno.
3. Son relativamente cortos tanto en el número de preguntas como en el tiempo que se necesita para completarlos.

En cuanto a los informatizados pretenden optimizar diferentes aspectos del proceso de aplicación-corrección-interpretación de un test. Así existen ciertos grados de automatización en el proceso, por ejemplo algunos exámenes universitarios y test comercializados se caracterizan por presentar un nivel mínimo de automatización ya que la presentación de los ítems es en formato de papel y lápiz, un medio de respuesta mediante hoja electrónica y la corrección mediante lectora electrónica. El siguiente nivel consiste en la presentación de los ítems en formato de papel y lápiz, mientras que tanto las respuestas como el proceso de datos hasta llegar a estimar el nivel del sujeto se realizan mediante ordenador. Y por último, el nivel máximo de automatización incluye además la presentación de los ítems e instrucciones por pantalla o por otro tipo de periféricos como altavoces, video-disco...

En resumen, *un test informatizado es el que utiliza el ordenador como medio de presentación de ítems, de entrada de las respuestas y de análisis e interpretación del rendimiento.*

Evolución de los test informatizados

Con los primeros test informatizados lo que se pretendía era solucionar los problemas relacionados con la corrección e interpretación de las pruebas psicológicas convencionales existentes ya en papel y lápiz.

Así, centrándonos en la década de los 30, podemos decir que tuvo lugar la primera experiencia de corrección automatizada mediante ordenador.

En la década de los 40, se aplican los desarrollos analógicos a la puntuación automatizadas del Strong Vocational Interest Blank(SVIB) y del Minnesota Multiphasic Personality Inventory (MMPI).

En los 50, comienzan las primeras experiencias en la aplicación de hojas de respuesta electrónica y lectoras ópticas, que sustituyen a los sistemas basados en tarjetas.

En los años 60 se desarrollaron sistemas automatizados de interpretación para tests como el 16PF, el California Psychological Inventory, el Rorschach o el WAIS.

En los 70 comienzan las primeras experiencias sobre la aplicación informatizada de test. Así el MMPI es de los primeros que puede presentarse en pantalla y responderse mediante el teclado. También se diseñaron administraciones informatizadas y otros test de personalidad como el Eysenck Personality Inventory, el 16PF.

Gracias a la tecnología digital se pudo proceder a la informatización de otros tipos de tests como los de inteligencia general y aptitudes, que requiere la presentación de ítems más complejos.

En los años 80 se produce un incremento exponencial de versiones informatizadas de tests convencionales de diversos contenidos, aplicados en contextos escolares, organizacionales o clínicos.

En el contexto de los Recursos Humanos, se desarrolló la batería CAM(un conjunto de tests informatizados que miden diversas habilidades relacionadas con el procesamiento de la información).

En cuanto a los contextos clínicos se han diseñado versiones informatizadas de entrevistas estructuradas y también se han realizado ensayos con los que se denominan entrevistas adaptativas informatizadas, las que permiten establecer diversos procedimientos de ramificación e itinerarios diversos dependiendo de las respuestas que van dando los sujetos. En concreto, en el ámbito de la psicoterapia se han desarrollado tests informatizados para identificar pensamientos irracionales o programas integrados para evaluar y superar progresivamente problemas relacionados con la toma de decisiones en diversas situaciones.

Ventajas de los tests informatizados

Los tests informatizados se caracterizan por la aportación de una serie de ventajas, las relacionadas con las condiciones de aplicación y con el procesamiento de las respuestas y su interpretación.

Con respecto a las primeras, destacamos las siguientes:

- Requieren menos tiempo. Según el estudio realizado por Bounderson, Intuye y Olsen (1989), se necesita más tiempo cuando el procedimiento de respuesta se hace mediante hoja electrónica que cuando se emiten mediante el teclado, así mismo, las

diferencias en el tiempo invertido en la realización del test de papel y lápiz e informatizado son mayores a medida que los sujetos tienen menor edad.

- Reducen la posibilidad de copia. La aplicación informatizada minimiza la posibilidad de los sujetos conozcan el contenido de los ítems o las respuestas correctas.

- Facilitan que las condiciones sean similares para todos los sujetos. El ordenador permite dar instrucciones similares y controlar el tiempo de exposición de cada ítem.

- Reducen determinados costes.

En cuanto a las relacionadas con el procesamiento de las respuestas y su interpretación, señalar que el ordenador reduce el tiempo dedicado a la corrección y baremación de las respuestas a los tests, así como la posibilidad de que se cometan errores en dicho proceso. El sistema informatizado también nos permite hacer baremaciones que nos da información sobre el nivel de la persona para posteriormente ponderar de forma diferente las diversas opciones de respuesta a los ítems.

Por otra parte, los procedimientos de elaboración de informes se fundamentan en un conjunto de reglas programadas en un algoritmo y en una base de datos de palabras y frases de las cuales se construye el informe final. Según Moreland(1991) existen tres tipos de informes:

Descriptivos. Caracterizados por la interpretación de cada subescala sin referirse a las restantes que constituyen el test. Estas interpretaciones añaden cierto grado al nombre de la escala o indican la posición del sujeto en la subescala según varias categorías ordenadas. Los informes descriptivos más elaborados añaden información estadística sobre el rendimiento del sujeto en relación al grupo normativo.

Condicionados. Proporcionan cierta interpretación en la medida que se cumplen ciertas condiciones en diferentes subescalas. Esta indicado como un primer acercamiento al problema.

Consultivos. Son los más cercanos al tipo de informe que proporcionaría un experto que utiliza frecuentemente el test, utiliza para ello el lenguaje técnico apropiado y formas de redacción semejantes a las que se realizan en los informes clínicos, siendo mucho más elaborados que los informes condicionados.

Además de las ventajas citadas con anterioridad es importante resaltar que la tecnología informática nos permite el diseño de nuevos instrumentos dentro de lo que se hace llamar “la nueva generación de test”. En estos test el uso del ordenador es fundamental ya que permite un control adecuado del tiempo de presentación de los ítems y de las latencias de respuesta.

Un ejemplo de estos últimos es la batería WICAT que incluye tareas de reconocimiento de figuras en un fondo enmascarado, secuencias de símbolos que se

incrementan progresivamente para evaluar la amplitud de la memoria etc. La batería CAM mide diversas variables procesuales en tres dominios estímulares diferentes.

Aspectos controvertidos.

Una duda que se plantea a la hora de automatizar un test de papel y lápiz es la de si ambos van a proporcionar puntuaciones equivalentes, ya que el hecho de cambiar el formato de presentación del test no tiene porque implicar un cambio en el baremo de referencia.

Una de las primeras revisiones sobre la equivalencia entre test convencionales y su versión informatizada realizadas por Mazzeo y Harvey (1988) dieron lugar a los siguientes resultados: 1) en la mayoría de estudios se obtuvieron puntuaciones medias diferentes, aunque pequeñas y de poca significación práctica, 2) entre los trabajos que encontraron diferencias, la mayoría obtuvieron medias más altas en las versiones originales no informatizadas, 3) los test de rapidez se respondían en menos tiempo mediante el teclado, 4) en determinados test informatizados de personalidad se produjo un número más elevado de elecciones de la opción “no lo sé”. Estos resultados no son concluyentes ya que entre las versiones informatizadas se da mucha heterogeneidad en cuanto al nivel de automatización, además se produjeron deficiencias técnicas importantes que podrían explicar el menor rendimiento de los sujetos en estas pruebas, así como que en escasos trabajos se permitió entrenamiento con el procedimiento de respuesta con la revisión de las respuestas proporcionadas.

Otras revisiones confirman que en la mayoría de los estudios de equivalencia sobre tests de rendimiento se obtienen puntuaciones medias superiores en las versiones no informatizadas. Otra revisión la realizan Rolls y Feltham (1993) en la que se concluye que las diferencias encontradas en test de personalidad son pequeñas o nulas, no aconsejándose una nueva baremación para las versiones informatizadas, cosa que no ocurre con respecto a los tests de aptitudes y neuropsicológicos.

Por otra parte, gracias a un trabajo reciente se concluye que los parámetros de dificultad de los ítems resultan equivalentes si en la versión informatizada se proporciona al sujeto las mismas condiciones de flexibilidad que en la versión no informatizada.

Por lo que a la validez de las interpretaciones informatizadas se refiere, se han realizado bastantes estudios rigurosos sobre la utilidad de los informes generados mediante ordenador. Moreland (1991) diferencia dos líneas de trabajo fundamentales:

- a. Los estudios comparativos con criterios externos, que consiste en comparar las valoraciones que se realizan en el informe automatizado con otra fuente como expertos clínicos con gran experiencia en la aplicación de las versiones no automatizadas.
- b. Estudios sobre la satisfacción de los clientes. Realizados para comprobar el grado en que manifestaban satisfacción los profesionales de la psicología y de la psiquiatría en relación a la claridad, adecuación de la información o utilidad.

Por otra parte, Endres, Guastello y Rieke (1992) para mejorar la calidad de las interpretaciones automáticas propusieron la fiabilidad metainterpretativa, que sirve para comprobar el grado de precisión con que un sistema de interpretación automática refleja en texto las puntuaciones del test.

En cuanto a las propiedades psicométricas de las medidas procesuales debemos destacar el trabajo de Roznowski (1993), el cual argumenta que las medidas procesuales reflejarán diferencias individuales estables si las tareas incluyen niveles heterogéneos de complejidad.

El verdadero problema de la validación de test informatizado de procesos cognitivos requiere nuevas estrategias que tienen que ver con la comprobación empírica de que el test detecta las previsiones derivadas del modelo teórico, con la planificación de estudios predictivos para demostrar la utilidad de las medidas en situaciones reales donde intervienen los procesos que se pretende medir.

2.-PROCEDIMIENTO GENERAL DE CONSTRUCCIÓN DE UNA PRUEBA O TEST.

Según Murphy y Davidshofer, (1994) la mayoría de los instrumentos de medida consisten en el desempeño de alguna tarea, la observación de la conducta o el autoinforme.

Por otra parte, todos los instrumentos tienen en común los siguientes aspectos; se trata de una técnica sistemática o estandarizada en la que la tarea, las instrucciones, las respuestas, el procedimiento de aplicación, la corrección, y la interpretación están pautadas y son equivalentes para todos los sujetos a los que se les aplica.

En general, los pasos habituales a seguir para la elaboración de un instrumento de medida (test, cuestionarios o cualquier tipo de escala) según Croker y Algina (1986), son los siguientes:

- Delimitación del objetivo para el que van utilizarse las puntuaciones.
- Definición del constructo que se pretenda medir con el instrumento.
- Descripción general de los componentes del constructo.
- Diseño del instrumento.
- Redacción de los ítems.
- Análisis de la calidad de los ítems.
- Estudio de la fiabilidad del instrumento.
- Estudio de la validez del instrumento.
- Elaboración de las normas de aplicación, interpretación, baremación de las puntuaciones.

De este modo a la hora de diseñar un instrumento de medida es preciso hacer una delimitación formal y teórica-conceptual.

En la **delimitación formal** se incluye a su vez el número y tipo de ítems y número de alternativas de respuesta. El número de ítems a priori es difícil de fijar, se deben construir como mínimo el doble de los que se usarán para la versión final y se debe ajustar el número de ítems al tiempo disponible para realizar el test.

Por lo que al tipo de ítems se refiere, existen dos tipos de respuestas; las abiertas, en las que el sujeto debe reproducir la respuesta y resultan difíciles de cuantificar aunque son muy adecuadas para evaluar conceptos complejos. Estas a su vez pueden ser de *completar, sustituir o corregir* (por ejemplo, la.....es condición necesaria pero no suficiente para que exista validez), de *respuestas cortas* (por ejemplo, ¿cómo afecta el número de ítems al coeficiente de fiabilidad de un test?), y de *ensayo* (por ejemplo, diferencia entre la TCT y TRI).

Las de respuesta cerradas en las que el sujeto selecciona las respuestas de entre dos o más respuestas propuestas, además son fáciles de cuantificar, precisas y fiables por lo que son adecuadas para evaluar una gran diversidad de conceptos. Estos tipos de respuestas son características de los test de ejecución máxima y los test de ejecución típica. Los primeros contienen preguntas de *verdadero y falso*, de *elección múltiple o de relacionar*, y los de ejecución típica (como cuestionarios de personalidad o aptitudes, escalas tipo Likert), que estarían formadas por alternativas del tipo *grado de acuerdo* (muy en desacuerdo-muy de acuerdo), de *evaluación* (nada importante-muy importante) y de *frecuencia* (siempre-nunca).

Así mismo, para cuantificar las respuestas con ítems politómicos o test de ejecución típica, el sujeto ponderará o graduará sus respuestas en escalas de 3 a 9 posibles valores aunque la mayoría de los instrumentos usan 5 puntos. La puntuación total del test se obtendrá al sumar los valores que se hayan asignado a cada ítem. En test con ítems dicotómicos por lo general a las respuestas correctas se le asigna el valor de 1 y a las erróneas 0, la puntuación total del test se obtendría de la suma de las respuestas correctas.

En la **delimitación teórica-conceptual** debemos identificar cual es el objetivo del instrumento por lo que es preciso presentar el objetivo desde un punto de vista comportamental que obliga a declarar explícitamente el objetivo para el que van a utilizarse las puntuaciones. Esto tiene la ventaja de que convierte al objetivo comprensible y ayuda a la definición del constructo. En lo que a ésta última se refiere está compuesta de una definición operacional y una definición estructural. La primera consiste en definir los comportamientos que se pueden considerar como indicadores del constructo y se identifica con el modelo de medida del constructo. Con respecto a la definición estructural se trata de una teoría consistente y sustantiva que relaciona el constructo medido con otros constructos empíricos de su entorno, esto es lo que dentro de la técnica de modelos estructurales se denomina parte estructural de un modelo o teoría.

Por último, realizar una descripción general de los componentes del constructo.

Otro aspecto a tener en cuenta en la construcción de un instrumento es la redacción de los ítems, para esto es necesario desarrollar el ítem partiendo de las ideas relevantes con el fin de elaborar ítems que midan objetivos importantes con el máximo

poder discriminativo. También es necesario presentar el ítem con claridad y precisión de tal forma que se evite el introducir elementos no funcionales demasiado específicos o que puedan sugerir las respuestas, así como eliminar el uso de frases que puedan estar estereotipadas. Según Muñiz existen una serie de errores frecuentes que han de evitarse como son:

- Las alternativas de respuestas que no resultan igualmente plausibles
- Los enunciados que aparecen en forma negativa.
- La diferencia en la longitud de las alternativas
- Reiteración innecesaria en el texto.
- Cifras desordenadas en las alternativas de respuestas.
- Abuso de la frase “todas /ninguna de las anteriores” en las alternativas de respuestas.

El siguiente paso para construir un test sería la evaluación del instrumento. Esta fase es crucial ya que todas las medidas obtenidas al aplicar una prueba contienen error y por tanto se hace necesario la evaluación de la calidad de los instrumentos de medida. La calidad de los ítems, la fiabilidad y la validez son las características fundamentales de un buen instrumento de medida y son propiedades exigibles a un buen test y por lo tanto, criterios a tener en cuenta para evaluarlos con criterios de calidad.

Esto implica un análisis de los ítems del test y un estudio de la fiabilidad y validez del mismo.

Nos centraremos primero en el análisis de los ítems, para ello debemos hacer un examen de la calidad individual de cada uno de los ítems que componen un instrumento de medida, lo que implica la obtención de información descriptiva y estadística.

La información descriptiva se refiere a la calidad técnica de sus elementos y en general supone la obtención de evidencias de validez de los mismos. Su objetivo sería establecer juicios razonables sobre el grado en que las evidencias aportadas permiten emprender acciones basadas en el modelo de medida, en este caso, a través de los ítems o indicadores. Lo importante por tanto no es plantear si el sistema de ítems o indicadores es válido en sí mismo sino que su validez esta en relación con los objetivos a medir.

De acuerdo con ello, es necesario asignar a los ítems o indicadores distintas evidencias de validez, como pueden ser la validez de contenido, de constructo y de criterio. La información estadística supone la estimación de parámetros para cada ítem (Dentro de la TCT, se estima mediante el índice de dificultad, discriminación y validez), el análisis de los distractores o alternativas incorrectas de respuestas y el examen del posible funcionamiento diferencial de los ítems en grupos de interés.

En cuanto a la fiabilidad se refiere, se trata de un criterio de calidad relacionado con la precisión de las medidas obtenidas con un test y que proporciona información acerca del grado en el que las puntuaciones empíricas son reflejo de las verdaderas y del grado de ajuste entre puntuaciones empíricas y verdaderas. En general, una medida es fiable si es precisa, estable y consistente.

Así, los procedimientos que se utiliza para constatar la fiabilidad se basan en el modelo lineal de regresión y utilizan como índice el coeficiente de correlación de Pearson, atendiendo a la estabilidad temporal de las puntuaciones obtenidas con el test, ya sea éste test- retest o formas paralelas, y la consistencia interna del test (alpha o dos mitades).

Otro aspecto importante es la *validez del test*, entendiendo ésta como el criterio de calidad relacionado con el grado en que un instrumento mide aquello que dice medir.

Lo último que se lleva a cabo en la construcción de un instrumento de medida es la evaluación de los sujetos, que consiste en la asignación, interpretación y equiparación de las puntuaciones. Para ello, las puntuaciones directas se transforma en estandarizadas y con éstas últimas se confeccionarán *baremos*. Existen dos tipos de baremos, los referidos al criterio y los referidos a la norma. Los primeros dan información sobre el grado en que el sujeto posee la característica o dominio medido. Estos baremos se ven posibilitados gracias a las mediciones que se realizan desde la TRI (teoría de respuesta al ítem). Los referidos a la norma, van a informar de lo posición que ocupa el sujeto, en el rasgo medido, respecto a su grupo de iguales. Las mediciones de este tipo de baremos se realizan desde la TCT (teoría clásica de los test).

No obstante, existen diferencias entre ambos tipos de baremos en cuanto a finalidad de la evaluación (los de criterio, estiman el rendimiento o conducta del sujeto en los objetivos que mide el instrumento y los de norma, pone de manifiesto las diferencias interindividuales en el rasgo que mide), la construcción del instrumento(en los de criterio hay que hacer una clara especificación del dominio de contenido y del uso pretendido por el instrumento, y en los de la norma los ítems se construye derivados de alguna teoría de rasgo o de constructo), la selección de los ítems (en los de criterio se seleccionan en función de los objetivos y uso predeterminado del instrumento y en los de la norma se seleccionan para poner de relieve las diferencias interindividuales, maximizar la varianza del test, y elegir los ítems de dificultad), el significado de las puntuaciones (en los de criterio la puntuación representa el estimador muestral de la conducta del sujeto en el dominio y en los de de la norma ésta es un indicador de la puntuación verdadera en el rasgo latente que se mide) y por último, la interpretación de las puntuaciones (en los de criterio, la puntuación tiene significado en términos absolutos y en los de la norma, la puntuación tiene significado sólo con relación al grupo normativo de referencia).

3. LA CONSTRUCCION DEL TEST DESDE LAS DIFERENTES TEORIAS

Para plantear las diferentes teorías que existen en la construcción del test, es preciso definir en primer lugar qué es una teoría de test. Esta es una teoría que proporciona modelos para las puntuaciones de los test, es decir, modeliza matrices de datos que tiene la respuesta que una muestra o grupo de sujetos han dado a cada uno de los elementos de un test. Así, una teoría del test sirve para saber hasta qué punto una medida obtenida en un momento determinado proporciona una estimación adecuada del nivel real en que posee el sujeto la característica psicológica que supuestamente se está evaluando. Por consiguiente, una teoría del test sirve para dar cuenta del error de medida inherente a toda medición

psicológica: estimación del error y proporcionar una estimación del rasgo o característica evaluada: estimación de la característica de interés.

Las principales teorías de tests que han surgido son la Teoría clásica de los test que ha pretendido desarrollar un modelo estadístico que contemplara los errores de medida. Además, esta teoría asume que la puntuación empírica que obtiene un sujeto cuando se le administra un test (X) es función del nivel real o verdadero en que el sujeto posee la característica o rasgo que está evaluando dicho test (V) y el error de medida que siempre se introduce en cualquier proceso de medida (E). Por tanto, la relación entre el comportamiento observable (X) y el nivel del sujeto en la variable no observable (V) es una relación lineal.

Entre sus supuestos encontramos que el nivel real del sujeto en la característica de interés es la media de los valores que se obtendría de forma empírica en caso de administrar el mismo test al sujeto en idénticas condiciones de medida un número infinito de veces, la independencia de las puntuaciones verdaderas y los errores de medida y la independencia de los errores de medida cometidos con distintas forma del test.

No obstante, esta teoría clásica de los test presenta algunas limitaciones como es la no invarianza de las mediciones respecto de los instrumentos utilizados. Otra gran limitación que presenta se refiere al hecho que en este enfoque las propiedades de los instrumentos de medida dependen en exceso de los objetos medidos, es decir, de los sujetos a los que se aplica el test. Por ejemplo, el índice de dificultad de los ítems está en función de la competencia de las personas a las que se la pasa el test; el índice de discriminación de los ítems y el coeficiente de fiabilidad dependen en gran medida de la variabilidad de la muestra.

A estas limitaciones anteriores se suma la dificultad real que existe para construir test paralelos, concepto que es clave en la TCT y que afecta directamente al cálculo del coeficiente de fiabilidad y longitud de los tests.

Como alternativa a las limitaciones anteriores surge un nuevo enfoque denominado Teoría de Respuesta al ítem, la cual va a permitir una construcción y análisis mucho más potente de los tests y cuyo objetivo principal es “conseguir medidas invariantes respecto de los sujetos medidos y de los instrumentos utilizados”.(Muñiz,1997). La clave del éxito de esta teoría va a consistir en la unificación de dos conceptos, el de separación de parámetros e invarianzas de los mismos. Para conseguir su objetivo la TRI ha desarrollado un conjunto de modelos matemáticos, los cuales asumen que la probabilidad de que una persona emite una determinada respuesta ante un ítem puede ser descrita en función de la posición de la persona en el rasgo y de una o más característica de ítem.

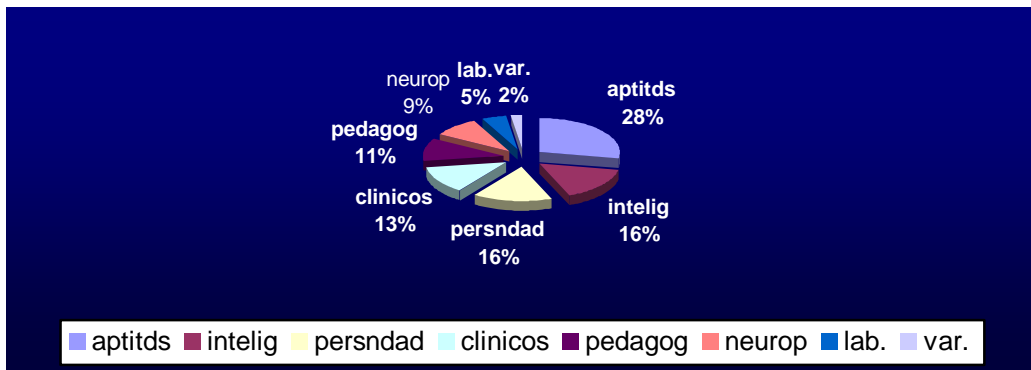
Por tanto, sus principales supuestos se van a referir a la naturaleza del rasgo que se pretende medir y a las relaciones que se esperan entre las respuestas de los ítems.

4. ESTUDIO SOBRE LOS TEST QUE SE OFERTAN

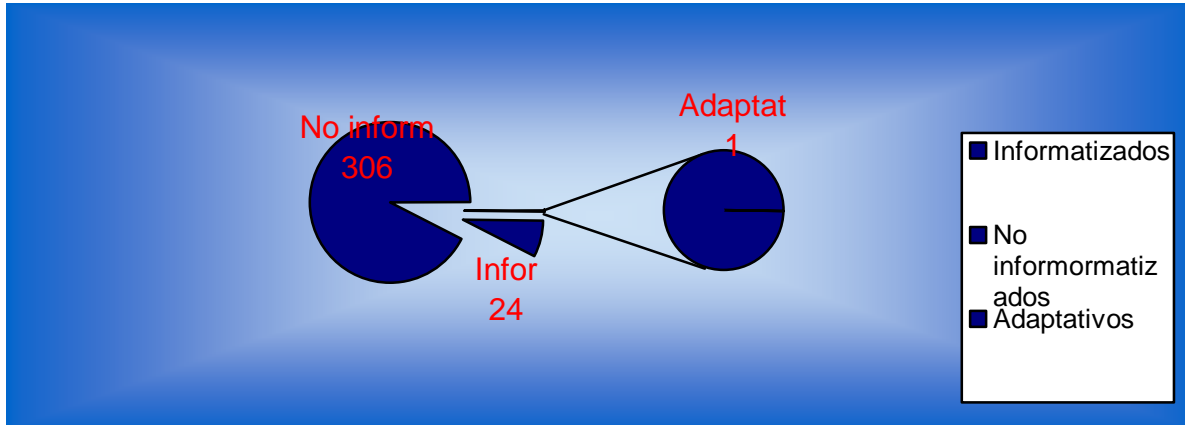
El siguiente estudio se ha llevado a cabo con el fin de poner en conocimiento de los futuros profesionales los posibles instrumentos que, hoy por hoy, se pueden usar según estemos interesados en evaluar las distintas áreas y utilizando los diferentes procedimientos de aplicación, ya sea los tradicionales de papel y lápiz o los informatizados.

Para ello, nos hemos apoyado en dos catálogos, el Psymtéc y en el catálogo de TEA ediciones 2004. No obstante durante esta revisión hemos tenido en cuenta que además de los test que aparecen en los mismos, existen multitud de test que no se ofertan debido a que se encuentran en determinadas instituciones que impiden el uso de los mismos por parte de los profesionales de manera personal y que, por tanto, habría que acudir a las mismas en caso de que se necesitara tal instrumento.

Los resultados que hemos obtenido se presentan en las siguientes gráficas:



Como es posible apreciar en la gráfica, en el análisis realizado se han encontrado un gran número de test que miden aptitudes, ya sean estas específicas (por ejemplo Toulouse-Piéron para medir atención y percepción) o generales (como el test de aptitudes básicas para la informática), en concreto son 92 test los que se han hallado. Otra de las áreas más estudiadas es la de la inteligencia, en esta se han encontrado 53 instrumentos (por ejemplo el Test de inteligencia no verbal). A esta área le sigue la de personalidad en la que se incluyen 52 instrumentos como por ejemplo el 16 PF-5; en la dedicada a cuestionarios clínicos situamos 43 tests o escalas como el MMPI-2. Por último las áreas que cuentan con un menor número de tests son la pedagógica, neuropsicológica y clima laboral, en la primera de ellas situamos un total de 35 (evaluación de los problemas en la edad preescolar), en la segunda 31 (CLIFTON, procedimiento de evaluación de ancianos) y en la última 17 (orientación profesional). Además de estas áreas hemos añadido una última que hemos llamado varios debido que en esta se han incluido instrumentos que no se han podido incluir en ninguna otra como son la entrevista, idiomas, test sobre psicología del deporte etc. hasta llegar a un total de 7 (ver Anexo).



Podemos decir que si tomásemos como muestra representativa los 330 instrumentos analizados el 7.27%(24) serían informatizados y sólo uno de estos últimos sería adaptativo informatizado (TRASI, test informatizado para la evaluación del razonamiento secuencial y la inducción).

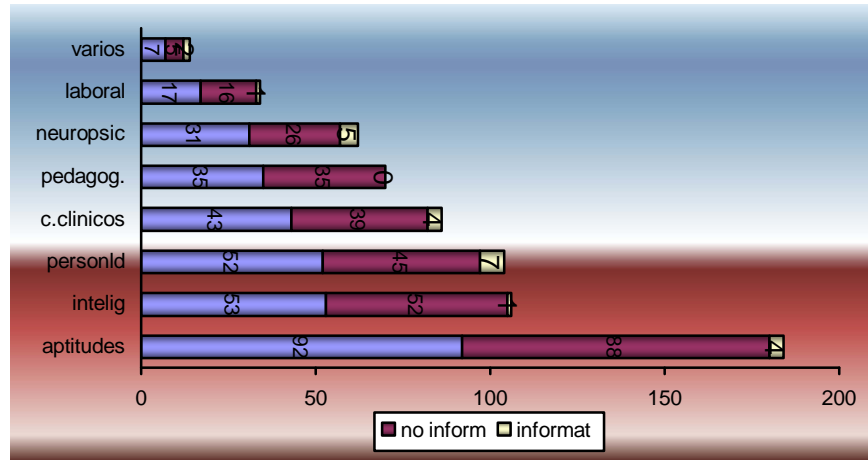
Dado que durante el estudio solo hemos podido encontrar un test informatizado adaptativo se hace necesario centrar la atención en éste.

Los autores de este test son Rubio y Santacreu, siendo el primero que se ha comercializado en Europa; su aplicación puede ser tanto individual como colectiva a sujetos de 18 años en adelante y con un nivel intelectual superior, cuenta con un banco de 98 ítems con una solución informática que decide cuales son los ítems más adecuados para el sujeto que está siendo evaluado, teniendo siempre presente el rendimiento que está obteniendo el mismo.

El principal objetivo de este instrumento es evaluar el razonamiento general secuencial y la inducción como componentes de la habilidad intelectual general.

Debido a que es un instrumento adaptativo permite que se le aplique a cada sujeto los elementos más apropiados para su nivel de aptitud, así como obtener una estimación mejor de la misma; además, el TRASI permite establecer la precisión que se desee de la medida en función de la cantidad de elementos aplicados al sujeto, también controla la tasa de exposición de elementos para evitar que estos se difundan entre los posibles evaluandos.

El test registra los datos personales del evaluado, presenta las instrucciones, selecciona y administra los elementos más adecuados, estima la habilidad del sujeto y genera un informe con la puntuación típica. Cuenta con una media de 50 y una desviación tipo de 10, los baremos se han conseguido con más de 7.500 universitarios españoles de diferentes titulaciones.



Con el análisis de esta gráfica lo que se pretende es poner de manifiesto los test que son informatizados y los que no, dentro de todas las áreas analizadas en la gráfica anterior.

Como es posible apreciar en la gráfica los test que miden personalidad son los que cuentan con un mayor número de instrumentos informatizados (CTI, inventario de pensamiento constructivo) seguido de los neuropsicológicos y de clima laboral, que cuentan con 5 instrumentos informatizados cada uno de ellos (neurología clínica; IPP, inventario de intereses y preferencias laborales). En cuanto a los cuestionarios clínicos y de aptitudes, estos cuentan con 4 informatizados (CAQ, cuestionario de análisis clínico y BAMI, test de aptitudes mentales informatizado), con 2 informatizados se encuentra la categoría de varios y con uno las áreas de inteligencia (el cual es además adaptativo) y de clima laboral, no pudiendo encontrar ninguno informatizado en el área de pedagógicos.

Por ello añadimos que es necesario para posteriores estudios un análisis más minucioso de este tipo de instrumentos para confirmar que, como es de sospechar, existen más instrumentos informatizados aunque no se encuentran al alcance de todos los profesionales.

5. Bibliografía

Olea, J. y Hontanga, O. (1999). Test informatizados de primera generación. En J. Olea, V. Ponsoda y G. Prieto, G. (Eds.), **Tests Informatizados. Fundamentos y aplicaciones.** Cap. 5, p.p. 111-125. Madrid: Pirámide.

A N E X O

PSYMTEC Y TEA EDICIONES 2004

1. APTITUDES

1.1. Aptitudes específicas

Tests no informatizados

- 1.- Escalas de valoración de memoria (MAS)
- 2.- Evaluación de memoria.
- 3.- Escala de evaluación de aprendizaje.
- 4.- Test de modelos visuales.
- 5.- Test conductual de memoria de Rivermead para niños.
- 6.- Test de la torre de Hanoi.
- 7.- Laberintos.
- 8.- Test de vocabulario en imágenes peabody.
- 9.- Escala de desarrollo del lenguaje de Reynell.
- 10.- Prueba del desarrollo inicial del lenguaje.
- 11.- Procedimiento de evaluación del lenguaje, 3ª edición.
- 12.- Escala de desarrollo del lenguaje temprano.
- 13.- Test de Token. Valoración del lenguaje descriptivo.
- 14.- Test de vocabulario expresivo y receptivo.
- 15.- Test de vocabulario expresivo.
- 16.- Test de vocabulario receptivo.
- 17.- Batería de valoración del lenguaje.
- 18.- Test de vocabulario para sordos.
- 19.- Inventario del desarrollo del lenguaje revisado.
- 20.- Test de habilidad fonológica.
- 21.- Desarrollo del habla y del lenguaje.
- 22.- Programa para el desarrollo y mejora de la capacidad articulatoria.
- 23.- Programa Hamlet.
- 24.- Crucigrama de sinfonos.
- 25.- Test de discriminación auditiva.
- 26.- EDAF. Evaluación de la discriminación auditiva y fonológica.
- 27.- MASS. Prueba de valoración de la percepción auditiva.
- 28.- Evaluación de la percepción fonológica.
- 29.- Test de visión de stycar.
- 30.- Escala de inteligencia de Haptic (ciegos).
- 31.- Test de aptitudes viso-perceptivas (no motor). TVPS.
- 32.- Test de percepción horizontal vertical.
- 33.- Tests motores de Ozerrestki.
- 34.- Test de punteado.
- 35.- Test de destreza manual de Minnesota.
- 36.- Test de relación espacial de Minnesota.

- 37.- Destreza digital de O'connor.
- 38.- Destreza manual con pinzas de O'connor.
- 39.- Test de destreza manual de Purdue.
- 40.- Test de coordinación visomotora.
- 41.- Test de aptitud manipulativa de Roeder.
- 42.- Lafayette Pegboard.
- 43.- Test de coordinación ambos brazos.
- 44.- Test de herramientas manuales.
- 45.- Dibujo en espejo automático.
- 46.- Clasificación de tarjetas.
- 47.- BENNETT. Test de destreza en el manejo de herramientas.
- 48.- CL. Cuadrado de letras.
- 49.- COE. Comprensión de órdenes escritas.
- 50.- CRAWFORD. Test de destreza con pequeños objetos.
- 51.- CREA. Inteligencia creativa.
- 52.- D2. Test de atención.
- 53.- IC. Instrucciones complejas.
- 54.- Macquarrie. Test de aptitudes mecánicas.
- 55.- MO ½. Método y orden.
- 56.- MVR. Memoria visual de rostros.
- 57.- RM. Rapidez motora.
- 58.- SIT. Situación-1.
- 59.- STROMBERG. Test de destreza de.
- 60.- TP. Toulouse-Pierón.
- 61.- TSAV. Test semántico de aptitud verbal.
- 62.- Visomotora. Coordinación.
- 63.- Caras. Percepción de diferencias.
- 64.- FI. Formas idénticas.
- 65.- MAI. Memoria auditiva inmediata.
- 66.- Batería de exploración visual de Mcdowell.
- 67.- Laminas pseudoisocromaticas.
- 68.- MY. Test de memoria.
- 69.- PIC. Prueba de imaginación creativa.
- 70.- SEASHORE. Aptitudes musicales.
- 71.- TO-1. Palabras, test de observación.
- 72.- BAIRES.- Test de aptitud verbal.

Tests informatizados

- 1. 73.- Test de la torre de Hanoi computerizado.
- 2. 74.- Visionlab.
- 3. 75.- Test de golpeo digital computerizado.

1.2. Aptitudes generales

Tests no informatizados

1. 76.- ABG. Aptitudes básicas generales.
2. 77.- ABI. Aptitudes básicas para la informática.
3. 78.- AMPE-F. Aptitudes mentales primarias equivalentes.
4. 79.- BAC. Batería para la actividad comercial.
5. 80.- BC. Batería de conductores.
6. 81.- BO. Batería de operarios.
7. 82.- BPA. Batería de pruebas de admisión.
8. 83.- BS.- Batería de subalternos.
9. 84.- BTA-R. Batería de tareas administrativas revisada.
10. 85.- DAT-5. Test de aptitudes diferenciales.
11. 86.- GCT. Test de aptitudes administrativas.
12. 87.- GMA. Evaluación de grado medio y alto.
13. 88.- PMA. Aptitudes mentales primarias.
14. 89.- APT. Test de pronóstico estructurado.
15. 90.- IGF. Inteligencia general factorial.
16. 91.- TEA. Test de aptitudes escolares.

Test informatizado

1. 92.- BAMI. Bateria de aptitudes mentales informatizadas.

2.- PRUEBAS DE INTELIGENCIA

Tests no informatizados

- 1.- Escala de desarrollo psicomotor de la primera infancia. O. Brunet-I. Lezine Revisión 1997.
- 2.- Escala manipulativa internacional de Leiter-R.
- 3.- Escala de inteligencia de adultos.
- 4.- Test de inteligencia breve no verbal.
- 5.- Escala de desarrollo de Kaufman.
- 6.- Evaluación del pensamiento cognitivo de Kaufman.
- 7.- Brazelton. Escala para la evaluación del comportamiento neonatal.
- 8.- Escala de desarrollo motriz de Lincoln-Oseretsky.
- 9.- Escala de inteligencia Stanford- Binet 4ª Edición.
- 10.- Recopilación Stanford- Binet (Forma L-M).
- 11.- Recopilación Stanford- Binet (4ª Edición).
- 12.- Escala de potencial intelectual.
- 13.- Escala de inteligencia de Merrill Palmer.
- 14.- Escala extendida de Merri Palmer.
- 15.- Test de inteligencia general.
- 16.- Test de inteligencia no verbal.
- 17.- Inteligencia práctica. Primera serie (Adaptación española).

- 18.- Test de formación de conceptos.
- 19.- Test del dibujo de la figura humana de Goodenough.
- 20.- Proyecto de inteligencia "Harvard".
- 21.- Proyecto de inteligencia "Harvard". Educación primaria.
- 22.- Test de inteligencia verbal.
- 23.- COGAT. Thorndike, Hagen.
- 24.- BETA.- BETA III. Instrumento no verbal de inteligencia.
- 25.- Cambios. Test de flexibilidad cognitiva.
- 26.- D-48. Test de dominós.
- 27.- D-70. Test de dominós.
- 28.- IG-2. Inteligencia general, nivel 2.
- 29.- Monedas. Aptitud de tipo superior.
- 30.- OS. Otis sencillo. Test de inteligencia general.
- 31.- Raven. Matrices progresivas.
- 32.- RP-30. Resolución de problemas.
- 33.- TIG. Test de inteligencia general. Serie de dominós.
- 34.- TISD.- Test de interpretación selectiva de datos.
- 35.- Alexander. Escala.
- 36.- Battelle. Inventario de desarrollo.
- 37.- BSID. Escala Bayley de desarrollo infantil.
- 38.- EOD. Escala observacional del desarrollo.
- 39.- K-ABC: Batería de evaluación de Kaufman para niños.
- 40.- K- BIT. Test breve de inteligencia de Kaufman.
- 41.- MSCA. Escalas Mccarthy de aptitudes y psicomotricidad para niños.
- 42.- TOMAL. Test de memoria y aprendizaje.
- 43.- TONI-2. Test de inteligencia no verbal.
- 44.- WAIS-III. Escala de inteligencia de Wechsler para adultos.
- 45.- Claves. Para la evaluación con el Wais-III.
- 46.- WISC-R. Escala de inteligencia de Wechsler para niños-revisada.
- 47.- WPPSI.- Escala de inteligencia para preescolar y primaria.
- 48.- EPA-2. Evaluación del potencial de aprendizaje.
- 49.- Factor G. Test de escala 1.
- 50.- Factor G. Test de escala 2 y 3.
- 51.- Primaria. Test de aptitudes cognoscitivas, I y II.
- 52.- Naipes G, niveles elemental, medio y superior.

Test informatizado adaptativo

1. TRASI. Test informatizado para la evaluación de razonamiento secuencial y la inducción.

3. PERSONALIDAD

Tests no informatizados

- 1.- Inventario y valoración de la personalidad.
- 2.- Índice de autoestima.
- 3.- Escala de depresión y ansiedad en la adolescencia.
- 4.- Test de relaciones objetales. Phillipson.
- 5.- Escala de depresión de Reynolds (niños).
- 6.- Escala de depresión de Reynolds (adultos).
- 7.- Temas, cuenta una historia.
- 8.- Terapia cognitiva de la depresión.
- 9.- Test de Rorschach.
- 10.- Test del dibujo de la familia.
- 11.- PY-EAS-D. Escala de estrategias de afrontamiento del stress para directivos.
- 12.- Test de los colores.
- 13.- Cuestionario de la salud general.
- 14.- Cuestionario de trastornos de personalidad.
- 15.- Cuestionario de rasgos de personalidad.
- 16.- AFA.- Autoconcepto forma A.
- 17.- AF-5. Autoconcepto forma 5.
- 18.- ASPA. Cuestionario de aserción en la pareja.
- 19.- CACIA. Cuestionario de autocontrol infantil y adolescencia.
- 20.- CEP. Cuestionario de personalidad.
- 21.- BFQ.- Cuestionario Big Five.
- 22.- CPS.- Cuestionario de personalidad situacional.
- 23.- EPQ-R. Cuestionario de personalidad de Eysenck.
- 24.- MIPS. Inventario Millon estilos personalidad.
- 25.- NEO PI-R. Inventario personalidad neo-revisado.
- 26.- PPG-IPG. Perfil e inventario de personalidad.
- 27.- CPQ. Cuestionario de personalidad de niños.
- 28.- EHS. Escala de habilidades sociales.
- 29.- ESFA. Escala de satisfacción familiar por adjetivos.
- 30.- ESPA. Escala de socialización parental en adolescentes.
- 31.- EPI. Cuestionario de personalidad.
- 32.- ESPQ. Cuestionario factorial de personalidad.
- 33.- IAC. Inventario de adaptación de conducta.
- 34.- Test del árbol.
- 35.- Fábulas de Düss.
- 36.- Familia.
- 37.- Figura humana.
- 38.- FTT. Test de los cuentos de hadas.
- 39.- H-T-P. (Casa, árbol, persona)
- 40.- Pata negra.
- 41.- TAT, CAT-A Y CAT-H. Test de apercepción temática.

- 42.- Szondi test.
- 43.- Z test.
- 44.- L.E.A. 12
- 45.- Cuestionario de rasgos de personalidad normal.

Tests informatizados

- 1. 46.- CTI. Inventario de pensamiento constructivo.
- 2. 47.- IPV. Inventario de personalidad para vendedores.
- 3. 48.- TPT. Test de personalidad de TEA.
- 4. 49.- 16 PF-5. Cuestionario factorial de personalidad 5ª edición.
- 5. 50.- HSPQ. Cuestionario de personalidad para adolescentes.
- 6. 51.- 16 PF. Cuestionario factorial de personalidad.
- 7. 52.- 16 PF-APQ. Cuestionario de personalidad para adolescentes.

4. CUESTIONARIOS CLINICOS

Tests no informatizados

- 1.- Valoración de discapacidad cognitiva severa.
- 2.- Evaluación del estado mental en ancianos.
- 3.- LOTCA Geriátrico.
- 4.- Escalas de evaluación de niños y adolescentes superdotados.
- 5.- PY-ABSTRACT. Test de screening.
- 6.- Evaluación de la conducta adaptativa.
- 7.- Valoración de discapacidad profunda.
- 8.- SEEP. Sistema de evaluación para personas plurideficientes.
- 9.- Batería de discapacidad severa.
- 10.- Instrumento de evaluación del comportamiento para niños autistas y con problemas.
- 11.- Test de desórdenes de atención, autismo e hiperactividad.
- 12.- Escala de evaluación de autistas de Gilliam.
- 13.- Test de la torre de Londres.
- 14.- Evaluación del procesamiento lingüístico en la afasia.
- 15.- Caja de juegos diagnóstico terapéutico.
- 16.- Diagnostico de problemas de aprendizaje.
- 17.- Corrección del trazo, lápiz eléctrico.
- 18.- Escala d deficientes mentales adultos.
- 19.- Prueba acacia.
- 20.- A-D. Cuestionario de conductas antisociales-delictivas.
- 21.- BASC.Sistema de evaluación de la conducta de niños y adolescentes.
- 22.- CAS. Cuestionario de ansiedad infantil.
- 23.- CDS. Escala de presión para niños.
- 24.- CET-DE. Cuestionario estructural tetradimensional para la depresión.
- 25.- CMAS-R. Escala de ansiedad manifiesta niños.

- 26.- EAE. Escala de apreciación de estrés.
- 27.- EBP. Escala de bienestar psicológico.
- 28.- EDI-2. Inventario de trastornos de la conducta alimentaria.
- 29.- EDAH. Evaluación del trastorno por déficit de atención con hiperactividad.
- 30.- ISRA. Inventario de situaciones y respuestas de ansiedad.
- 31.- MMPI-A. Inventario multifásico de personalidad de Minnesota para adolescentes.
- 32.- MCMI-II. Inventario multiaxial clínico de Millon.
- 33.- MACI. Inventario clínico de Millon para adolescentes.
- 34.- PNP. Cuestionario.
- 35.- SCL-90-r. Cuestionario de 90 sistemas.
- 36.- STAI. Cuestionario de ansiedad estado/rasgo.
- 37.- STAIC. Cuestionario de ansiedad estado/rasgo en niños.
- 38.- STAXI-2. Inventario de expresión de ira estado-rasgo.
- 39.- TAMAI. Test autoevaluativo multifactorial de adaptación infantil.

Tests informatizados

- 1.- Evaluación psicológica computerizada Viena Test System.
- 2.- CAQ. Cuestionario análisis clínico.
- 3.- JAS. Inventario de actividad de Jenkins forma C.
- 4.- MMPI-2. Inventario multifásico de personalidad de Minnesota-2.

5. PEDAGOGICOS

Tests no informatizados

- 1.- BEL. Batería de evaluación de la lectura.
- 2.- Estudio de las habilidades tempranas,(SECA).
- 3.- Batería R-habilidades cognitivas y de conocimiento.
- 4.- Evaluación de los problemas en la edad preescolar.
- 5.- Evaluación de las habilidades lectora escolar.
- 6.- Test de mejora educacional.
- 7.- Test de análisis de lectoescritura.
- 8.-ACRA. Estrategias de aprendizaje.
- 9.- AEI. Aptitudes en educación infantil.
- 10.- A-RE-HA. Análisis del retraso del habla.
- 11.- BAPAE. Batería de aptitudes para el aprendizaje escolar.
- 12.- BLOC. Batería del lenguaje objetivo y criterial.
- 13.- BOEHM. Test de conceptos básicos.
- 14.- CHTE. Cuestionario de hábitos y técnicas de estudio.
- 15.- CLT-CLOZE. Dos pruebas de comprensión lectora.

- 16.- CMP. Prova control de la maduración del nen a preescolar.
- 17.- COLE. Test de comprensión lectora.
- 18.- Comprensión lectora.
- 19.- CO+PD, Test de comprensió verbal I programa de desenvolupament.
- 20.- DIE. Diagnóstico integral del estudio.
- 21.- ECL-1 y 2. Evaluación de la comprensión lectora.
- 22.- EDIL. Exploración de las dificultades individuales de lectura.
- 23.- EPP. Escala de evaluación de psicomotricidad en preescolar.
- 24.- EVOCA. Estimación del vocabulario.
- 25.- FROSTIG. Desarrollo de la percepción visual.
- 26.- IHE. Inventario de hábitos de estudio.
- 27.- ITPA. Test Illinois de aptitudes psicolingüísticas.
- 28.- Lectura.
- 29.- MFF-20. Test de emparejamiento de figuras conocidas.
- 30.- RA+PD. Test de raonament verbal I programa de desenvolupament.
- 31.- PROESC. Batería de evaluación de los procesos de escritura.
- 32.- PROLEC. Batería de evaluación de los procesos lectores de los niños de educación primaria.
- 33.- PROLEC-SE. Evaluación de los procesos lectores en alumnos de tercer ciclo de primaria y secundaria.
- 34.-TAEL. Test for assessment of english language.
- 35.- TALE. test de análisis lectoescritura.

6. NEUROPSICOLOGIA

Tests no informatizados

- 1.- Batería neuropsicológica Luria Nebraska.
- 2.- Test y materiales de la Batería neuropsicológica Halstead-Reitan.
- 3.- Sistema de evaluación neuropsicológica Halstead Russell.
- 4.-Test Barcelona "Programa integrado de exploración neuropsicológica".
- 5.- Cognistat Examen Neurocognitivo.
- 6.- EL diagnóstico neuropsicológico de Luria.
- 7.- Test de inatención visual.
- 8.- Test gestáltico visomotor de Bender.
- 9.- Test neuropsicológico breve de Sterling.
- 10.- Tests seleccionados del laboratorio de neuropsicología de Benton.
- 11.- Examen de afasia.
- 12.- Test de reconocimiento escalado.
- 13.- Test para el examen de la afasia.
- 14.- Test de retención visual de Benton.
- 15.- BOSTON. Test de vocabulario de evaluación de la afasia y de los trastornos relacionados.
- 16.- CAMDEX-R. Prueba de exploración Cambridge revisada para la valoración de los trastornos mentales en la vejez.

- 17.- CLIGTON. Escala de madurez mental de Columbia.
- 18.- CMMS. Escala de madurez mental de Columbia.
- 19.- CSAT. Tarea de atención sostenida en la infancia.
- 20.- CUMANIN. Cuestionario de madurez neuropsicológica infantil.
- 21.- EDAF. Evaluación de la discriminación auditiva y fonológica.
- 22.-HARRIS. Test de dominancia lateral.
- 23.- HPL. Test de homogeneidad y preferencia lateral.
- 24.- Luria-DNA. Diagnóstico neuropsicológico de adultos.

Tests informatizados

- 1.- Fundamentos neuroanatómicos.
- 2.- Neurología clínica.
- 3.- Enfermedades neurológicas.
- 4.- CANTAB. Diagnóstico neurológico computerizado.
- 5.- REHACOM. Rehabilitación neurológica computerizado.

7. CLIMA LABORAL

Tests no informatizados

- 1.- Sistema Harrington-O Shea para la selección de carreras y profesiones.
- 2.- Elección de carrera.
- 3.- Orientación profesional.
- 4.- Batería psicotécnica selección profesional EOS.
- 5.- CLA. Clima laboral.
- 6.- EAE. Escalas de apreciación del estrés.
- 7.- MPS. Escala de motivaciones psicosociales.
- 8.- NEG0. Habilidades en la negociación.
- 9.- SIV. Cuestionario de valores interpersonales.
- 10.- SPV. Cuestionario de valores personales.
- 11.- ACS. Escala de afrontamiento adolescente.
- 12.- AECS. Actitudes y estrategias cognitivas sociales.
- 13.- BAS. Batería de socialización.
- 14.- CIPSD. Cuestionario de intereses profesionales.
- 15.- KUDER-C. Registro.
- 16.- SMAT. Test de motivación en adolescentes.

Test informatizado

- 1.- IPP. Inventario de intereses y preferencias profesionales.

8. *VARIOS*

Tests no informatizados

- 1.-Entrevista.
- 2.- Persona y profesión.
- 3.- TAEL. Prueba de inglés.
- 4.- TFI. Prueba de francés.
- 5.-PY-BTPD-S. Batería de test psicológicos para deportistas Salamanca.

Test informatizado

- 1.- TOEIC. Test of English for international communication.
- 2.-TRC. Análisis observacional codificado.