

Tests Autoadaptativos Informatizados **(TADI)**

GUIÓN TRABAJO

TEMA 5: “TESTS ADAPTATIVOS INFORMATIZADOS **(TAI): ESTRUCTURA Y DESARROLLO. VENTAJAS E** **INCOVENIENTES”**

1. RESUMEN
2. Tests Informatizados (TI): Definición/ Origen. Características.
3. Tests Adaptativos Informatizados (TAI): Definición/ Origen. Características. ¿Cómo se crea y usa?
4. TAI vs TI (Tabla comparativa)
5. Tests Autoadaptativos Informatizados (TADI): Definición/ Origen. Características.
6. TAI vs TADI (Tabla comparativa)
7. Ventajas e Inconvenientes (TI, TAI, TADI)
8. CONCLUSIONES
9. Actividades
10. Bibliografía

ORIGEN Y DEFINICIÓN:

En las últimas décadas se ha producido un gran avance en el desarrollo de la informática. Esto ha influido directa e indirectamente sobre los diferentes ámbitos de nuestras vidas y campos de estudio e investigación; se ha producido una mejora en la realización de los mismos. Este avance de la informática, y el desarrollo de la teoría de respuesta a los ítems (TRI), ha generado la creación de nuevos procedimientos de medida diferentes a los típicos de papel y lápiz. Los primeros tests adaptativos en formato de lápiz y papel (Lord, 1971) y los test convencionales informatizados (Koson, Kitchen, Kochen y Stodolosky, 1970) han dado paso a los tests adaptativos informatizados (TAIs; Johnson y Mihal, 1973), entre los cuales, los denominados tests autoadaptados informatizados (TADIs; Rocklin y O'Donnell, 1987) constituyen una de sus variantes más recientes.

El desarrollo de la teoría de respuesta a los ítems (TRI) ha favorecido la aparición de los tests informatizados ya que una de las características principales de este modelo teórico es que la probabilidad de acierto ante las preguntas realizadas en los tests no va a depender del ítem, sino de la capacidad del sujeto; es decir, no sólo se considera la dificultad del ítem presentado sino que también se tiene en cuenta la capacidad del sujeto para lograr responder correctamente.

Por lo tanto, debido a la influencia del avance informático y la TRI surgieron en primer lugar los Tests Informatizados (TI). Éstos son los que utilizan el ordenador como medio de presentación de ítems, de entrada de las respuestas y de análisis e interpretación del rendimiento. Sin embargo, los Test Adaptativos Informatizados (TAI) consisten en un banco de ítems con propiedades conocidas (calibrados normalmente bajo algún modelo de la TRI) y un procedimiento para seleccionar los más apropiados para la habilidad de cada persona (un algoritmo informático que presenta ítems más difíciles o más fáciles según se acierten o fallen los anteriores).

Tal y como hemos visto previamente, con los TAIs, uno de los inconvenientes es que la ansiedad puede llegar a convertirse en un factor que condicione a los candidatos. Las personas difieren en sus reacciones a la dificultad del test, bien por la orientación motivacional hacia el éxito o el fracaso (Atkinson y Litwin, 1960), la ansiedad que experimentan ante una situación de evaluación (Rocklin y Thompson, 1985) o el concepto que tienen de si mismas (Vispoel, Rocklin y Wang, 1994). De modo que la uniformidad afectiva y motivacional generada por un TAI no siempre es una situación

óptima de evaluación y la incidencia en el rendimiento de los factores señalados puede afectar a la validez de las medidas (Rocklin, O'Donnell y Holst, 1995). Debido a esto, Rocklin y O'Donnell (1987) propusieron un procedimiento alternativo para intentar controlar la ansiedad de evaluación que pueden experimentar los sujetos cuando respondan a los TAIs, modificando el procedimiento de selección de los ítems.

La idea consistió en permitir a los evaluandos ajustar de forma dinámica el nivel de dificultad de los ítems que debían responder, hasta situarse en el nivel que les permitiera un rendimiento óptimo. Para ello, propusieron dividir el banco de ítems en varias categorías de dificultad (normalmente entre 5 y 8) y permitir al evaluando, antes de responder a cada ítem, situarse en la categoría que considere más adecuada para su nivel, pudiendo adaptar la situación de evaluación a su estado afectivo y motivacional; un test autoadaptado informatizado (TADI) selecciona en cada momento el ítem más informativo de la categoría elegida por el evaluando y proporciona feedback sobre el resultado obtenido (acierto o fallo). El resultado más destacado fue que con el TADI se lograban estimaciones de habilidad más elevadas, por lo que sugirieron que el sujeto al que se le administra un test autoadaptado tiene <<acceso a una variedad de información relevante (incluyendo estados afectivos y motivacionales) para la selección óptima de los ítems, mejorando en este sentido al algoritmo tradicional TAI>>. Por tanto, desde su punto de vista, los sujetos utilizan información adicional para mejorar su rendimiento en el test. Luego, la diferencia básica entre un TAI y un TADI reside en la forma de selección de ítems: el algoritmo informático vs. el propio sujeto; es decir, en el primer caso, el algoritmo se encarga de la selección de ítems a partir del rendimiento del sujeto en los ítems previos, y en el segundo, se permite al evaluando elegir el nivel de dificultad de cada ítem.

Un test autoadaptado informatizado (TADI), por ende, representa una aplicación alternativa de la TRI en contextos de evaluación informatizada.

FUNDAMENTOS:

En un TADI el banco de ítems se divide en varias categorías ordenadas de dificultad (normalmente entre 5 y 8), o estratos. Aunque la elección del número de

estratos depende del tamaño del banco de ítems, la longitud del test y de las limitaciones que pueda tener el software, se debería emplear el mayor número de estratos posible y el máximo de ítems por categoría, de este modo un mismo sujeto podría elegir todos los ítems de un mismo nivel. En todos los estudios sobre TADI se han utilizado ítems de elección múltiple, pero se admite cualquier formato que pueda corregir el software.

Según Olea et al. (1998), el éxito en la aplicación de un TADI va a depender en gran medida de las instrucciones presentadas al comienzo del test. De hecho en las instrucciones de los diferentes tests se suele hacer referencia a dos principios de la TRI. La primera es que cuando se puede informar a los sujetos acerca de que la valoración de su rendimiento en el test va a depender tanto de los niveles de dificultad que elijan como del número de ítems contestados correctamente, nos estamos refiriendo en términos no técnicos al método de estimación de máxima verosimilitud de la TRI., es decir, que deberemos darles las instrucciones sobre que se tendrá en cuenta la puntuación que obtengan dependiendo del nivel de dificultad, es decir, indicarles que la dificultad del test se tendrá en cuenta a la hora de obtener sus puntuaciones.

Por último, en segundo lugar, deberemos recomendar en las instrucciones a los sujetos que elijan los niveles superiores de dificultad a los que puedan responder adecuadamente. A primera vista puede ser contradictorio con respecto a lo que les comentamos en el punto anterior; sin embargo, es un intento sutil de evitar el problema de la estimación de errores típicos elevados para aquellos sujetos que seleccionen niveles de dificultad que no se ajustan razonablemente a sus niveles de habilidad.

ESTRUCTURA Y DESARROLLO:

Podemos establecer una secuencia prototípica en la construcción de un TADI; de hecho, se trata de la misma secuencia que se lleva a cabo con los TAI, la única diferencia estriba en el tipo de algoritmo utilizado, informático frente al del propio sujeto de los TADI.

✓ FASE 1: Planificación y Prospección:

En este momento se plantean los interrogantes de tipo económico, técnico y deontológico.

En primer lugar deberemos decidir presentar cuáles van a ser los objetivos de la evaluación y la viabilidad del TADI. Deberemos tener en cuenta si vamos a fabricar un

nuevo banco de ítems o utilizaremos alguno de los existentes junto a la validación de las puntuaciones, qué tipo de ítems utilizaremos (politómicos, dicotómicos...), cuánto tiempo será necesario para la construcción, evaluar los costes y beneficios para la creación del mismo...

En segundo lugar tendremos que comprobar la adecuación del contenido a las condiciones de la TRI. Fabricar un Banco de Ítems Calibrados (BIC) unidimensional supone satisfacer toda una serie de requerimientos de la TRI, como el ajuste de las respuestas de cada ítem a un modelo, la ausencia de anomalías producidas por otras características de los examinados, etc. Sin embargo, la elección de un contenido idóneo a la TRI facilitará este objetivo.

Por último, se deberá seleccionar el formato de los ítems, puesto que el tipo de ítem escogido afectará a la idoneidad del modelo TRI y sus prestaciones. Actualmente lo más habitual es efectuar el Diseño de Anclaje de Pruebas (DAP) con pruebas tipo test de lápiz y papel basadas en ítems de alternativa múltiple puntuados dicotómicamente. Esto supone que el formato y estructura de los reactivos se transformará a soporte informático, debiendo anticipar los posibles efectos de su traslado a la pantalla.

✓ **FASE 2: Producción del banco de ítems.**

A continuación lo que se debe hacer es elaborar las reglas de generación de los ítems. Existen diversas propuestas útiles para realizar esto dependiendo de cuál sea el objetivo del proyecto. Por ejemplo, si se trata de la medición de aptitudes será importante disponer de un modelo cognitivo de referencia.

Posteriormente habrá que determinar el tamaño del banco de ítems y del DAP. Es decir, un TADI será eficiente en la medida en que disponga de muchos ítems donde escoger, ya que se atenuarán los factores no controlados que pueden afectar a la estimación de la capacidad. Los DAP pueden variar desde estructuras muy simples hasta diseños con bloques encadenados y en red mucho más complejos.

También deben tomarse decisiones relativas a los examinados y a la administración del DAP, es decir, relativas a la cantidad de examinados, su motivación, condiciones de administración y seguimiento de la muestra. Con respecto a los DAP, las instrucciones de la prueba harán especial hincapié en no omitir, rectificar ni efectuar dobles marcas.

✓ **FASE 3: Calibración y Ensamblado del banco de ítems.**

En esta fase se obtendrán los datos y se llevarán a cabo los análisis previos por bloques

PROCESO DE APLICACIÓN:

1°) INICIO:

En primer lugar se deben dar las instrucciones para la realización del test, que, como hemos comentado en otras ocasiones, pueden aumentar o disminuir la probabilidad de éxito en la obtención de los resultados.

La sesión puede empezar de varias maneras, entre ellas nos encontramos con un ítem al azar, con uno centrado en la dificultad o bien ya desde este momento se deja a criterio del sujeto que está siendo examinado.

2°) CONTINUACIÓN:

El programa informa del acierto y fallo en cada respuesta siendo el *examinado* quien *elige el sentido* (más fácil – más difícil) del siguiente ítem (primer modo de presentación) o bien, el *examinado elige* tanto el *sentido* como la *distancia* en dificultad (segundo modo de presentación).

En el primer caso el programa combina elementos clásicos de la estrategia de máxima información y/o métodos bayesianos con la selección personal. El sujeto sólo decide la tendencia de la siguiente tarea pero no su diferencia en dificultad respecto al ítem que acaba de responder.

Por lo que el programa, tras cada presentación, debe reestimar el nuevo ítem que presentaría al sujeto de modo automático. Sin embargo, si la decisión del sujeto es contraria presenta uno de los ítems adyacentes al recién contestado. En la segunda opción, el examinado toma las decisiones de un modo completo ya que abarca la tendencia y distancia del nuevo ítem.

3°) FINAL:

Este momento tiene lugar cuando se alcanza determinado nivel de precisión o bien se agotan las nuevas elecciones. La puntuación queda fijada por la estimación de la pauta de aciertos y fallos según los procedimientos de la TRI.

VENTAJAS E INCOVENIENTES:

Los previsibles beneficios de los TADIs, fundamentalmente respecto al incremento en nivel de rasgo estimado y al descenso en ansiedad estado, se han intentado explicar de forma diferente. Así, Rocklin (1994) plantea la "hipótesis de autorregulación", según la cuál el evaluando que responde a un TADI ajusta su nivel de ansiedad al nivel que le resulta más adecuado para conseguir un rendimiento óptimo; Vispoel y Coffman (1994) plantean la "hipótesis de distracción": los evaluandos presentan más atención a la tarea en un TADI que en un TAI; Wise (1994) propuso la "hipótesis de control percibido", que mantiene que los efectos beneficiosos se deben a que el evaluando tiene mayor grado de control sobre una situación que le puede resultar estresante. Sin embargo, podemos decir que la efectividad de estos tipos de tests no se encuentran tanto en el nivel de ítems presentados como por la elevada carga motivante que los caracteriza.