

**S  
T  
R  
O  
O  
P**

**TEST  
DE  
COLORES  
Y  
PALABRAS**

CHARLES J. GOLDEN, PH. D.

**Manual**



**Ediciones**



# **STROOP**

## **Test de Colores y Palabras**

Charles J. Golden, PH.D.

**Traducción y adaptación:**

*Departamento I+D de TEA Ediciones*

## **MANUAL**

(3º edición)

**TEA**

PUBLICACIONES DE PSICOLOGÍA APLICADA

Serie menor núm. 226

TEA Ediciones, S. A.

MADRID 2001

## ÍNDICE

PRÓLOGO .....	
1. INTRODUCCIÓN .....	
1.1. Ficha técnica .....	
1.2. Descripción general .....	
2. APLICACION Y CORRECCION .....	
Aplicación individual .....	
Aplicación en grupo .....	
Repetición de la aplicación .....	
Puntuación .....	
Fiabilidad .....	
Datos ESTADÍSTICOS .....	
Puntuación de interferencia .....	
Influencia de la edad en los niños .....	
Retraso mental .....	
Influencia del sexo .....	
3. DIAGNOSTICO DE DISFUNCIONES CEREBRALES .....	
Puntuaciones bajas en palabras .....	
Todas las puntuaciones son bajas .....	
P y C normales; PC bajo .....	
Todas las puntuaciones son normales .....	
Observaciones conductuales .....	
P normal; C y PC bajo .....	
El Stroop en baterías de tests .....	
Ejemplos de casos .....	
Cómo distinguir desórdenes orgánicos de los de tipo psiquiátrico .....	
Estrés .....	
Drogas .....	
4. RESULTADOS DEL STROOP Y ESTRES .....	
Cambios en el Sistema Nervioso Autónomo .....	
Niveles hormonales .....	
EEG (Electroencefalograma) .....	
5. STROOP, PERSONALIDAD Y COGNICIÓN .....	
Creatividad .....	
Cognición .....	
Automatización .....	
6. EL STROOP Y LA PSICOPATOLOGIA .....	

Estudios con puntuación "serial" .....	
7. RESULTADOS EN MUESTRAS ESPAÑOLAS.....	
7.1. Muestra .....	
7.2. Resultados .....	
7.3. BAREMOS .....	
8. REFERENCIAS .....	
ANEXO A .....	
ANEXO B .....	

## PRÓLOGO

Las investigaciones realizadas con el Test Stroop de Colores y Palabras han demostrado que examina procesos psicológicos básicos útiles en el estudio de la neuropsicología humana y de los procesos cognitivos. Las dimensiones básicas evaluadas por el Stroop que se asocian con la flexibilidad cognitiva, la resistencia a la interferencia procedente de estímulos externos, la creatividad, la psicopatología y la complejidad cognitiva juegan un papel evidente en muchos procesos cognitivos interrelacionados que determinan la habilidad individual para afrontar el estrés cognitivo y procesar informaciones complejas. El Stroop puede ser de utilidad en estudios referidos a campos tan importantes como la neuropsicología, la neurofisiología, la personalidad y los procesos cognitivos en psicopatología. Ofrece a los investigadores de dichas disciplinas una medida básica y fiable de procesos subyacentes importantes para el estudio de los procesos cognitivos.

Además, el Stroop ha demostrado en las investigaciones y en la práctica clínicas que se trata de un test clínico eficaz, tanto para la evaluación de disfunciones cerebrales como para la evaluación de psicopatologías en general. Puede ser utilizado como un test de evaluación o como parte de una batería más general. Su administración rápida (5 minutos), su validez y fiabilidad, su facilidad de aplicación hacen del Stroop un instrumento especialmente atractivo.

La presente publicación del Stroop pretende cumplir varios objetivos. En primer lugar, ponerlo a disposición de los profesionales clínicos e investigadores con una presentación homogénea y normalizada, lo que hasta ahora no se había realizado. Existían diversas formas y versiones con formatos, instrucciones y tiempos diferentes, cada una de ellas con ventajas e inconvenientes. La forma que presentamos pretende recoger las principales ventajas de las versiones existentes para servir tanto al campo clínico como al de la investigación y aprovechar la amplia literatura existente sobre la estructura del Stroop (ver capítulos 1 y 2).

El segundo objetivo es conseguir que los psicólogos sean conscientes de los usos potenciales del Stroop en el campo clínico (capítulos 3 y 6) y en el campo experimental (capítulos 4, 5 y 6). Nuestra experiencia demuestra que el Stroop es una valiosísima aportación a la evaluación neuropsicológica de pacientes clínicos, así como una herramienta versátil para la investigación experimental. Estamos convencidos de que el uso del Stroop servirá también para identificar nuevas áreas en las que puede ser muy útil. La bibliografía que se incluye sobre aplicaciones experimentales pretende, a la vez, presentar el estado actual de las mismas y estimular la realización de investigaciones más extensas con este instrumento.

Deseo animar a los usuarios del Stroop a compartir con nosotros cualquier experiencia adicional a fin de que podamos mantener este manual lo más actualizado posible. Creemos que con las investigaciones realizadas hasta la fecha sólo hemos tocado algunos de los posibles usos del Stroop y que futuras investigaciones abrirán nuevos e importantes campos para su utilización.

Charles J. Golden, Ph.D.  
*Nebraska Psychiatric Institute*  
*University of Nebraska Medical Center*  
1978

## 1. INTRODUCCION

### 1.1 FICHA TECNICA

**Nombre original:** Stroop Color and Word Test.

**Nombre de la adaptación española:** STROOP. Test de Colores y Palabras.

**Autor:** Charles J. Golden, Ph.D.

**Adaptación:** Departamento de I+D de TEA Ediciones, S.A.

**Administración:** Preferentemente individual.

**Duración:** 5 minutos.

**Aplicación:** De 7 a 80 años.

**Finalidad:** Detección de problemas neurológicos y cerebrales y medida de la interferencia.

**Baremación:** Baremos de población española general.

**Material:** Manual y un juego de tres páginas de palabras y colores.

### 1.2. DESCRIPCIÓN GENERAL

El Test Stroop de Colores y Palabras se desarrolló a partir de investigaciones de los primeros psicólogos experimentales que observaron que la identificación de colores era siempre más lenta en adultos que sabían leer que la lectura de los nombres de colores.

El primer informe publicado sobre este fenómeno se debe a Catell (1886), quien estimó que las palabras se podían leer e identificar en 1/4 de segundo mientras que la identificación de un tono de color requería un tiempo dos veces más largo. Cattell atribuyó esta diferencia a la idea de que ver y nombrar una palabra era una asociación automática mientras que nombrar un tono de color era el fruto de un esfuerzo consciente para elegir y decir el nombre del mismo. Pensó que la lectura de palabras era automática como resultado de una dilatada práctica en la tarea de lectura. Sin embargo, Brown (1915) demostró que, incluso con una práctica intensiva, el nombramiento de colores nunca era tan rápido como la lectura de palabras.

Ligon (1932) expuso la teoría de que el nombrar colores era más lento por causa de un factor especial de naturaleza orgánica. La identificación de palabras, por su parte, dependía más de la práctica que de un factor orgánico especial. Por tanto, la lectura de palabras sería susceptible de grandes mejoras con la práctica mientras que la velocidad de nombrar colores estaba siempre limitada por un factor orgánico innato.

Stroop (1935) sugirió que la diferencia entre ambas tareas era debida a que los colores estaban asociados a una variedad de respuestas conductuales mientras que las palabras sólo estaban asociadas a un tipo de respuesta conductual, la lectura. Con el fin de facilitar el estudio de las relaciones entre colores y palabras, Stroop diseñó el test que ha llegado a ser conocido como Test Stroop de Colores y Palabras.

La primera versión del Stroop contenía una hoja con las palabras "rojo", "verde", "marrón", "azul" y "púrpura" impresas en diez columnas y diez filas. Todas las palabras estaban impresas en tinta de color pero siempre en colores diferentes a los indicados por la palabra (por ejemplo, la palabra "rojo" podía estar impresa en tinta azul pero nunca en tinta roja). Una segunda página del test estaba formada por pequeños rectángulos de colores. La última página estaba formada por las palabras de los colores citados impresas en tinta negra.

El test de Stroop llamó la atención, sobre todo, por la conducta de los sujetos en la página en la que no coincidía el nombre de cada color con el color de la tinta usada para imprimirlo. Se comprobó que cuando se pedía al sujeto que leyese las palabras, lo hacía tan rápidamente como cuando las palabras estaban escritas en tinta negra pero que cuando se pedía al sujeto que nombrase el color de la tinta con que estaba escrita la palabra el tiempo aumentaba casi un 50% en relación con el tiempo necesario en la página con rectángulos coloreados. Esta fuerte disminución en la velocidad de identificación de los colores se conoce como "efecto de interferencia color-palabra".

Los estudios de Stroop fueron seguidos de cientos de experiencias publicadas sobre este test. Esta prueba ha atraído una atención tan grande por su fiabilidad en la identificación de diferencias individuales (Jensen & Rohwer, 1966) y por su naturaleza en cierto modo paradójica (Dyer, 1973; Bonis, 1968). Se han realizado numerosas experiencias sobre el uso del Stroop en investigación de procesos cognitivos y de la personalidad, en psicopatología y en el diagnóstico y comprensión de disfunciones cerebrales orgánicas.

Dado el interés que tienen los efectos de la interferencia en el Stroop, se han hecho diversos intentos para desarrollar otros tests que midiesen la misma dimensión. Por ejemplo, White (1969) diseñó un test parecido con las palabras norte, sur, este y oeste, situadas en las esquinas de un cuadrado, de forma que fuesen incoherentes la posición en el cuadrado y el significado de la palabra. Shor (1970) presentó series de flechas dirigidas hacia arriba, abajo, izquierda y derecha con las palabras arriba, abajo, izquierda y derecha escritas sobre las flechas pero sin coincidir con la dirección indicada por las mismas. Estos estudios demostraron que los tiempos de reconocimiento se incrementaban en un 10 ó 20% lo que resulta un cambio mínimo comparado con los efectos detectados en el Stroop. Además, ciertos estudios realizados en nuestro laboratorio (inéditos) no encontraron correlación entre la interferencia del Stroop y la interferencia medida por estos tests, lo que indica que hay un factor especial de relación entre el citar colores y el leer palabras que no se halla presente en esos tests pseudo-Stroop, factor al que se deben los resultados únicos obtenidos con el Stroop (Dyer, 1972).

Algunos estudios neuropsicológicos han demostrado que la interferencia del Stroop se produce como consecuencia de interferir el proceso verbal y no en la fase de respuesta ni en la confusión del sujeto (Wheeler, 1977). Parece que los estímulos del Stroop activan un proceso automático de respuesta verbal que interfiere con el nombramiento de los colores aprendido conscientemente. El sujeto realiza la tarea o bien ejecutando las dos respuestas secuencialmente (lectura de la palabra seguida del nombramiento del color) o bien suprimiendo mediante control voluntario la respuesta automática de lectura de palabras (Golden, 1976).

La originalidad del Stroop radica en el hecho de que la palabra coloreada suscita una respuesta verbal automática que requiere muchas de las mismas funciones neuropsicológicas que son necesarias para nombrar los colores. Además, la velocidad de



ambas reacciones (leer palabras y nombrar colores) es tal que la respuesta de leer palabras ocupa los canales neuropsicológicos que, al mismo tiempo, la respuesta de nombrar colores necesita para poder ser procesada. En las versiones de las flechas y direcciones, por ejemplo, la velocidad del proceso es mucho menor, lo que permite que la respuesta de leer la palabra se dé antes de que el sujeto esté listo para atribuir a la flecha una palabra de dirección. Además, la palabra y la flecha no son tan inseparables una de otra como lo son los colores y las palabras.

Todo esto indica que la página de interferencia del Stroop mide básicamente la capacidad del individuo para separar los estímulos de nombrar colores y palabras. Ciertas personas son capaces de hacerlo y pueden suprimir la respuesta de lectura y concentrarse en la tarea de nombrar los colores; otras personas no son capaces de suprimir el nombramiento de la palabra y han de procesar tanto la palabra como el color antes de responder; en otros sujetos, por fin, las respuestas de palabra y color están íntimamente confundidas debido a niveles altos de interferencia.

Los estímulos del Stroop afectan, por tanto, a niveles básicos, a la capacidad del sujeto para clasificar información de su entorno y reaccionar selectivamente a esa información. La detección de esa habilidad básica hace que el test Stroop sea útil en la investigación de una amplia serie de procesos psicológicos fundamentales tanto en personas normales como en discapacitados.

## **2. APLICACIÓN Y CORRECCIÓN**

Para el desarrollo de una versión normalizada del test Stroop fue necesario analizar las formas disponibles, que tenían algunas características diferentes:

### **a) Número de colores**

La versión original del Stroop constaba de cinco colores pero en versiones posteriores se usaron tres, cuatro o cinco colores (Thurstone & Mellinger, 1953; Broverman, 1960). Golden (1975) comparó los resultados usando versiones con tres, cuatro y cinco colores, sin encontrar diferencias significativas entre las tres versiones. Dado que las tres versiones eran igualmente fiables para producir el efecto de interferencia, se escogió la versión más simple (tres colores) a la hora de construir la versión normalizada.

### **b) Forma de presentación de los colores en la 2da página**

Los estímulos usados para presentar los colores en la 2da página han sido tan variados como rectángulos, círculos, cruces esvásticas y sílabas sin sentido (Stroop, 1935; Jensen, 1965; Golden, 1974). A la hora de hacer la elección es de gran importancia garantizar que los colores de los estímulos usados en la página de los colores coinciden con los colores incluidos en la página de interferencia color-palabra. En esta última página se presentan palabras (p.e.: ROJO) y no rectángulos coloreados, lo que hace que el color parezca más claro de lo que en realidad es por efecto de los espacios en blanco. Por ese motivo, se decidió utilizar un estímulo similar (la palabra) en la página de colores. También era necesario que el estímulo no tuviese ningún valor semántico y, en consecuencia, se pensó que la presentación más adecuada sería formar grupos de "X", más concretamente grupos de cuatro "X" para mantener un número de signos muy similar al de los nombres de los colores (AZUL, ROJO, VERDE).

### **c) Presentación de los elementos**

La mayor parte de las versiones del Stroop consistían en matrices de 10 por 10 o de 5 por 20 elementos que, a su vez, podían leerse por filas o por columnas. Estos detalles tal vez sean secundarios en el caso de sujetos normales pero sujetos con deficiencias (p. ej.: pacientes con daño cerebral) pueden desorientarse en la página cuando tienen que leer por filas (Silverstein, 1965). Con el fin de reducir al mínimo este problema, en la versión normalizada se adoptó el formato de 5 columnas de 20 elementos, indicando a los sujetos que trabajen por columnas. Esto reduce el número y la dificultad de los cambios de orientación espacial y tiene también la ventaja de permitir que el test se imprima en hojas de papel de tamaño normalizado A4 (210 x 297 mm.).

### **d) Forma de puntuación**

Se han empleado dos métodos principales de puntuación: tiempo que se tarda en completar 100 elementos y número de elementos realizados en cierto lapso de tiempo. Aunque varios estudios demostraron que, en personas normales, los resultados de ambos métodos son iguales, se ha preferido el segundo con un tiempo de 45 segundos por varias razones:

- El test tiene un tiempo límite. En ciertos sujetos discapacitados el tiempo de completar 100 elementos puede subir a 10 ó 15 minutos, sin que ello añada información útil.
- La versión de 45 segundos puede usarse tanto para aplicación individual como en grupo (ver más abajo).
- La realización de 100 elementos puede producir una gran frustración en niños pequeños y en personas discapacitadas y puede llevar a que rehúsen continuar o a producir otros problemas similares.

### **Descripción de la versión normalizada**

La versión normalizada (Golden, 1975) consta de tres páginas, conteniendo cada una 100 elementos distribuidos en cinco columnas de 20 elementos.

El contenido de cada página es el que se detalla a continuación:

La primera página está formada por las palabras "ROJO", "VERDE" y "AZUL" ordenadas al azar e impresas en tinta negra en una hoja de tamaño A4. No se permite que la misma palabra aparezca dos veces seguidas en la misma columna.

La segunda página consiste en 100 elementos iguales ("XXXX") impresos en tinta azul, verde o roja. El mismo color no aparece dos veces seguidas en la misma columna. Los colores no siguen el mismo orden de las palabras de la primera página.

La tercera página consiste en las palabras de la primera página impresas en los colores de la segunda, mezclados ítems por ítem. El primer ítem es el color del ítem 1 de la 1ª página impreso en la tinta del color del ítem 1 de la 2ª página. No coincide en ningún caso el color de la tinta con el significado de la palabra.

## **APLICACIÓN INDIVIDUAL**

Las páginas se presentan al sujeto en el orden siguiente:

- **Primera:** la que contiene nombres de colores (ROJO, VERDE, AZUL) impresos en

tinta negra.

- **Segunda:** la que está formada por filas de X impresas en colores distintos (ROJO, VERDE, AZUL).
- **Tercera:** la que contiene nombres de colores (ROJO, VERDE, AZUL) que se presentan impresos en un color distinto al que corresponde a la palabra escrita.

Las hojas se colocarán directamente frente al sujeto en una superficie plana. Si lo desea, éste las puede girar hasta un ángulo máximo de 45°, a la derecha o a la izquierda. No se permitirá que este giro supere el ángulo indicado ni que el sujeto las levante o las separe de la superficie en que estén colocadas. La tercera página debe estar situada en la misma posición que la segunda: si el sujeto mantiene la segunda página en la misma posición que tenía cuando se le entregó, deberá mantener también la tercera en esa misma posición, sin girarla en ningún sentido. No se permite tapar las hojas de ninguna forma.

A continuación se leerán las siguientes instrucciones:

*"Esta prueba trata de evaluar la velocidad con que Usted puede leer las palabras escritas en esta página. Cuando yo se lo indique, deberá empezar a leer en voz alta las columnas de palabras, de arriba a abajo, comenzando por la primera (SEÑALAR LA PRIMERA COLUMNA DE LA IZQUIERDA) hasta llegar al final de la misma (MOSTRAR CON LA MANO, MOVIENDOLA DE ARRIBA A ABAJO EN LA PRIMERA COLUMNA); después continuará leyendo, por orden, las siguientes columnas sin detenerse (MOSTRAR CON LA MANO LA SEGUNDA COLUMNA, LA TERCERA, ETC.)*

*Si termina de leer todas las columnas antes de yo le indique que se ha terminado el tiempo concedido, volverá a la primera columna (SEÑALAR) y continuará leyendo hasta que dé la señal de terminar.*

*Recuerde que no debe interrumpir la lectura hasta que yo diga "¡Basta!" y que debe leer en voz alta tan rápidamente como le sea posible. Si se equivoca en una palabra, yo diré "No" y Ud. corregirá el error volviendo a leer la palabra correctamente y continuará leyendo las siguientes sin detenerse. ¿Quiere hacer alguna pregunta sobre la forma de realizar esta prueba?"*

Estas instrucciones se pueden repetir tantas veces como sea necesario, hasta que el sujeto comprenda claramente qué es lo que tiene que hacer. Después se continúa diciendo:

*"¿Está preparado?...Entonces ¡Comience!"*

En este momento se pone el cronómetro en marcha. Cuando hayan transcurrido 45 segundos se dice:

*"¡Basta! Rodee con un círculo la última palabra que ha leído. Si ha terminado toda la página y ha vuelto a empezar ponga un 1 dentro del círculo. Ahora pase a la página siguiente."*

Las instrucciones para la SEGUNDA PÁGINA son iguales que las de la primera excepto el comienzo que es el siguiente:

*"En esta parte de la prueba se trata de saber con cuánta rapidez puede nombrar los colores de cada uno de los grupos de X que aparecen en la página"*

Si el sujeto no presenta ningún tipo de alteración, se le propondrán brevemente las siguientes instrucciones:

*"Este ejercicio se realiza de forma similar al de la página anterior. Comience en la primera columna, y nombre los colores de los grupos de X que hay en ella, de arriba a abajo, sin saltar ninguno; luego continúe la misma tarea en las restantes columnas. Recuerde que debe nombrar los colores tan rápidamente como le sea posible"*

Si el sujeto tiene alguna perturbación o dificultad se le repetirán las instrucciones completas. Como en la primera página. Se concede un tiempo de 45 segundos.

Para comenzar la TERCERA PARTE de la prueba se seguirán las instrucciones incluidas a continuación:

*"Esta página es parecida a la utilizada en el ejercicio anterior. En ella debe decir el color de la tinta con que está escrita cada palabra, sin tener en cuenta el significado de esa palabra. Por ejemplo (SE SEÑALA LA PRIMERA PALABRA DE LA COLUMNA), ¿qué diría Ud. en esta palabra?"*

Si la respuesta del sujeto es correcta se continúa leyendo las instrucciones; si es incorrecta, se dice:

*"No, esa es la palabra que está escrita. Lo que Ud. tiene que decir es el color de la tinta con que se ha escrito. Ahora (**SEÑALAR EL MISMO ELEMENTO**), ¿qué diría al mirar esta palabra?"*

Si contesta correctamente, se dice:

*"De acuerdo, eso es correcto".*

Se continúa señalando la segunda palabra y diciendo:

*"¿Qué diría Ud. en esta palabra?"*

Si contesta correctamente se prosigue la aplicación diciendo:

*"Bien, ahora continuará haciendo esto mismo en toda la página. Comenzará en la parte de arriba de la primera columna (SEÑALAR) y llegará hasta la base de la misma; luego continuará de la misma manera en las columnas restantes. Debe trabajar tan rápidamente como le sea posible. Recuerde que si se equivoca tiene que corregir su error y continuar sin detenerse. ¿Quiere hacer alguna pregunta?"*

Si la respuesta es incorrecta se repiten de nuevo estas instrucciones. Esto se hará tantas veces como sea necesario, hasta que el sujeto comprenda lo que tiene que hacer o se tenga la certeza absoluta de que le va a resultar imposible realizar la tarea.

Como en los casos anteriores, las instrucciones se repetirán tantas veces como sea necesario. Cuando el sujeto no tenga dudas sobre lo que debe hacer, se inicia la aplicación diciendo:

*"¡Puede comenzar!"*

Se pone en marcha el cronómetro y cuando hayan transcurrido 45 segundos, se dice:

*"¡Basta! Rodee con un círculo la última palabra que ha dicho."*

En este momento se recoge la prueba y se anotan los datos de identificación del sujeto (nombre y apellidos, edad y sexo).

## **APLICACIÓN EN GRUPO**

La aplicación en grupo es similar a la que se realiza individualmente pero las palabras "en voz alta" deben sustituirse por "para sí mismo". Las instrucciones deben darse con detalle para cada página con el fin de asegurarse de que todos los sujetos las comprenden bien. Cuando se termine el tiempo previsto para cada página (el mismo que en la aplicación individual) se pedirá a los sujetos que rodeen la última palabra o grupo de figuras a que hayan llegado.

La aplicación en grupo sólo puede hacerse cuando se trata de sujetos que no van a necesitar ayuda individual, por ejemplo, estudiantes mayores, adultos, etc. La aplicación en grupo no debe usarse con poblaciones disminuidas o con personas con problemas cerebrales, esquizofrenia, etc. Todas las aplicaciones que se recogen en los baremos españoles se han realizado en forma individual por lo que en caso de aplicación colectiva dichos haremos deben usarse con prudencia.

## **REPETICIÓN DE LA APLICACIÓN**

En ciertas ocasiones es conveniente detener la aplicación del test y volver a comenzarla. Si en una página el sujeto duda y no da la primera respuesta en tres segundos, el examinador debe decir:

*"¡Alto! ¿Ha comprendido bien las instrucciones?"*

Deben entonces repetirse éstas, si se cree necesario, y se vuelve a empezar la aplicación de la página.

Si el sujeto no trabaja con la rapidez con la que parece posible, el examinador dirá:

*"¡Alto! Recuerde que debe leer las palabras (o decir los colores) lo más rápidamente que pueda. Vamos a empezar otra vez esta página".*

En cualquier página el sujeto puede señalar con el dedo los elementos. Sin embargo, en algunos casos, en la tercera página puede tapar toda la palabra salvo una sola letra, lo que interrumpe el procesamiento verbal de la palabra. En ese caso, el examinador dirá:

*"¡Alto! Puede señalar con el dedo las palabras si quiere, pero no puede taparlas. Vamos a empezar de nuevo y no tape la palabra".*

Igualmente, el sujeto debe ser interrumpido si gira la página más de 45 grados o si la levanta de la mesa. La interrupción debe realizarse inmediatamente al detectar el fallo.

También se interrumpirá al sujeto si hace una tarea diferente de la indicada (normalmente leer las palabras en la tercera página) y se le repetirán las instrucciones.

Como norma práctica para facilitar el seguimiento de la prueba, se aconseja que el examinador tenga delante de sí un ejemplar de la segunda página lo que permite comprobar si el sujeto se equivoca o no tanto en dicha página como en la tercera, ya que los colores de las palabras escritas en ésta siguen el mismo orden que en la página dos.

Conviene también repetir el test a los sujetos que presenten resultados correspondientes a modelos fuertemente patológicos. Si el sujeto tiene problemas relevantes, obtendrá resultados similares en la segunda aplicación mientras que los sujetos normales obtendrán resultados más normales la segunda vez. En general, se debe repetir el test si:

- la puntuación de nombrar colores (2da página) es igual o mayor que la de leer palabras (1ra página); la puntuación obtenida en la tercera página es igual o mayor que la obtenida en cualquiera de las otras dos;
- la puntuación de color-palabra (3a página) es inferior a 20;
- se dan otras circunstancias que el examinador considere sospechosas.

No hay inconveniente en aplicar el test dos veces como norma habitual. Los individuos normales mostrarán una mejora en la segunda aplicación mientras que los sujetos con problemas orgánicos mostrarán poca o ninguna variación en las puntuaciones. En estudios experimentales tanto la primera como la segunda aplicaciones produjeron los mismos resultados al correlacionar sus puntuaciones con las obtenidas en pruebas de personalidad o cognitivas. Sin embargo, en dichos estudios conviene usar siempre la misma puntuación, ya sea la primera o la segunda.

## **PUNTUACIÓN**

En el test Stroop se obtienen tres puntuaciones principales:

- **P** es el número de palabras leídas en la primera página
- **C** es el número de elementos realizados en la página de los colores (2da página).
- **PC** es el número de elementos realizados en la tercera página

Los errores no se cuentan pero producen una puntuación total algo menor ya que se hace que el sujeto repita el elemento.

Diversos autores han propuesto otras puntuaciones secundarias para medir la interferencia u otras habilidades. Muchas de esas puntuaciones pretendían obtener una medición pura de la interferencia a partir de la página tres, independiente de la velocidad del sujeto en la lectura de palabras y en la identificación de colores. Con el fin de saber cuál de esas puntuaciones aportaba una información única y original se realizó un análisis factorial cuyos resultados se presentan en la tabla 1. Como puede verse, el factor 2 es el mejor representado por las puntuaciones, diferenciales PC-PC estimado y PC-C, que son medidas de interferencia "pura" corregidas por factores de velocidad. En general se recomienda el uso de la primera puntuación porque es más estable en poblaciones con disfunciones. La

fórmula para calcular el valor de PC estimado se presenta en el anexo A. Con el fin de simplificar la terminología la puntuación PC estimada se expresará también como PC'.

**TABLA 1.** Análisis factorial de las puntuaciones del Stroop

<b>Variable</b>	<b>Factor 1</b>	<b>Factor 2</b>	<b>Factor 3</b>
PC	0,663	0,723	0,187
C	0,938	-0,190	0,283
P	0,737	-0,044	-0,669
P/C	0,076	-0,128	0,984
C+P	0,943	-0,122	-0,307
PC/C	-0,587	0,781	-0,203
PC-PC'	0,003	0,997	-0,038
PC-C	0,108	0,980	0,080

### **FIABILIDAD**

La fiabilidad del Stroop se ha mostrado muy consistente en las diversas versiones existentes. En todos los casos, los investigadores han usado el método test-retest con tiempos comprendidos entre un minuto y 10 días entre las dos aplicaciones. Jensen (1965) obtuvo índices de 0,88, 0,79 y 0,71 para las tres puntuaciones directas. Golden (1975) obtuvo valores de 0,89, 0,84 y 0,73 (N=450) en la versión colectiva y de 0,86, 0,82 y 0,73 (N=30) en aplicación individual. La fiabilidad que se obtuvo con sujetos sometidos a las dos formas (N=60) fue de 0,85, 0,81 y 0,69. En las mismas muestras indicadas, la fiabilidad del factor de interferencia (PC - PC') es igual a la de la tercera página (0,7).

### **DATOS ESTADÍSTICOS**

La tabla 2 recoge los datos estadísticos básicos obtenidos en diversos estudios americanos, mientras que los correspondientes a las aplicaciones realizadas en España se incluyen en el capítulo 7. Puede comprobarse que los resultados derivados de las diversas muestras son muy parecidos cuando se comparan grupos de edades similares. Con el fin de comparar con mayor facilidad las puntuaciones directas conviene convertirlas en puntuaciones típicas T (con media 50 y desviación típica 10) usando la tabla incluida en el anexo B.

Las puntuaciones directas de los sujetos de más de 45 años o de menos de 17 deben ser corregidas mediante los datos del anexo B antes de usar las tablas de puntuaciones T.

Hay que destacar que las correcciones por edad para los grupos superiores (más de 65 años) e inferiores (menos de 17 años) deben tomarse con reservas hasta que se disponga de muestras más amplias. No obstante, la experiencia clínica demuestra que estas puntuaciones son adecuadas en la práctica.

En todo tipo de análisis, para considerar significativa una diferencia en puntuaciones, ésta debe ser de al menos 10 puntos T. Los límites considerados normales se encuentran entre 35 y 65 puntos T en cualquiera de las puntuaciones.

**TABLA 2. Puntuaciones medias obtenidas en el Stroop**

Elementos contestados en 45 segundos			
MUESTRA	PALABRA	COLOR	COLOR-PALABRA
STROOP (1935) (N=100)	110*	71*	41*
Jensen (1965) (N=436)	118*	78*	45*
Jensen (1965) (N=436)	121*	81*	51*
Adultos jóvenes (15-45)	108,2	80,4	44,7
Adultos mayores (46-64)	99,6	75,8	38,8

\*Estas puntuaciones fueron estimadas matemáticamente partiendo del tiempo tardado en resolver 100 elementos.

### PUNTUACION DE INTERFERENCIA

Dado que la puntuación de interferencia proviene de las obtenidas en las páginas de lectura y de colores, no necesita ser corregida en función de la edad. Teóricamente, la puntuación media es cero y la desviación típica es 10. Los sujetos con puntuación superior a cero tienen una alta resistencia a la interferencia. Aunque esta puntuación no es muy significativa a efectos de diagnosticar disfunciones cerebrales orgánicas (capítulo 3), sí lo es como instrumento de investigación sobre creatividad y estilos cognitivos (capítulos 4 y 5). Las puntuaciones de interferencia se calculan siempre sobre puntuaciones directas de P, C y PC ya corregidas por la edad.

Teóricamente, esta puntuación corregida debe medir una dimensión pura de flexibilidad cognitiva, como se indica en los capítulos 4 a 6. Por tanto, es de gran utilidad para buscar individuos con alto o bajo potencial de flexibilidad cognitiva, adaptación al estrés cognitivo y creatividad. A diferencia de otros tests, el Stroop está poco afectado por el rendimiento del individuo. El rendimiento real depende mucho de factores tales como la oportunidad, capacidades físicas, apoyo de otros, cociente intelectual, etc.; factores que no son evaluados por el Stroop. En forma similar, cuando la puntuación de interferencia se relaciona con la habilidad de un sujeto para soportar el estrés hay que pensar que la conducta real de la persona es una combinación de este factor con la cantidad de estrés experimentada en su vida. Así, aunque una persona con alta resistencia a la interferencia puede ser más estable bajo el estrés, la cantidad de tensión o estrés que puede soportar es finita.



## INFLUENCIA DE LA EDAD EN LOS NIÑOS

El primer gran estudio sobre las velocidades de lectura de palabras e identificación de colores en niños fue realizado por Ligon (1932), que aplicó a 635 alumnos de escuelas públicas páginas similares a las P y C del Stroop. Comprobó que en los dos casos la velocidad iba en aumento con la edad y que la diferencia entre las dos páginas era casi una constante.

Stroop (1935) propuso, en relación con los datos de Ligon, que se constase el número de respuestas dadas en 100 segundos más que el tiempo tardado en dar 100 respuestas y comprobó que la puntuación pasaba de 22,6 en los más pequeños a 73,6 en los chicos mayores. Los estudios más ambiciosos en relación con la edad se realizaron por Comalli, Wapner y Werner (1962). Usaron la versión de tres colores descrita más arriba (Broverman, 1960) y examinaron a 235 sujetos con edades entre 7 y 80 años. Los sujetos se agruparon en 11 grupos de edades: 7, 8, 9, 10, 11, 12, 13, 17-19, 25-34, 35-44 y 65-80. El grupo más numeroso era el de los 11 años (29) y el más reducido el de 25-34 (14). Los resultados para las páginas P y C fueron similares a los de Ligon (1932). En la página PC los niños de 7 años fueron los más lentos (20 elementos en 45 segundos) mientras que los del grupo 17-19 años tuvieron resultados similares a los de adultos jóvenes de la tabla 2.

Los autores han interpretado los resultados en términos de la teoría evolutiva comparativa que basa su análisis en los conceptos de diferenciación jerárquica e integración, indicando que el Stroop depende de la diferenciación en la lectura de palabras y en el nombramiento de colores. Los niños, dicen, están menos diferenciados en su desarrollo.

Rand, Wapner, Werner y MacFarland (1963) intentaron un análisis más fino sobre cuatro grupos de edad (6, 9, 12 y 16 años). Además de las puntuaciones básicas, obtuvieron con el Stroop siete categorías adicionales. Primero analizaron el número de errores, comprobando que éstos decrecían con la edad en las tres páginas. Asimismo, los errores aumentaban al pasar de P a C y de C a PC. En segundo lugar contaron las respuestas inadecuadas en PC entendiendo por tal las respuestas que no se referían ni al color de la tinta ni a la palabra escrita, por ejemplo, decir azul cuando la palabra es "verde" y está escrita en rojo. Este tipo de respuesta también decrecía con la edad aunque los resultados eran similares para los grupos de 9 y 12 años. Los de 6 años daban el valor de inadecuaciones más alto y los de 16 el más bajo (sólo 1,0 de media).

Una tercera categoría se refería a respuestas mal articuladas, por ejemplo, decir "verda" en vez de verde. En las páginas de colores esta función resultó más alta para los grupos de 6 y 16 años y más baja para los de 9 y 12. La repetición de una palabra de color entre ítems se consideró como "error de inserción de palabras de color", lo que tendía a decrecer con la edad, al igual que las omisiones de algún elemento. Estas omisiones, en el grupo de los 16 años, supusieron una cantidad insignificante (0,03 de media). Las últimas dos categorías fueron "frases lingüísticas insertadas" ("Quiero decir", "Esto es muy difícil") y frases insertadas no lingüísticas ("Oh", "Ah"). Las primeras decrecían con la edad mientras que las segundas aumentaban. Los autores supusieron que las frases no lingüísticas servían a los grupos mayores como sustitutos de las frases lingüísticas de la infancia con las que el niño se da instrucciones a sí mismo ("Lee el color, no la palabra"). Rand, et al. (1963) expusieron la teoría de que los sujetos mayores dicen estas frases internamente y las frases no lingüísticas son un efecto lateral de este proceso.

Peretti (1969) encontró los mismos efectos de la edad que Comalli y aventuró la hipótesis de que los alumnos de primaria responden peor a la página PC porque la fuerza de los hábitos de leer palabras y de nombrar colores es casi igual.

Se han realizado dos intentos de desarrollar una versión del Stroop para niños de preescolar. El primero planteaba el conflicto de forma y color (Butollo, Bauer y Riedle, 1971). Los autores usaron como medida el número de respuestas correctas y omisiones, encontrando más interferencia en los niños de tres años que en los de seis, un resultado esperable a tenor de las teorías evolutivas de Rand y otros (1963).

Cramer (1967) desarrolló una versión del Stroop para preescolares usando dibujos como estímulos (manzana, árbol, sol, agua) y encontró un efecto de interferencia cuando se pedía al niño que suprimiese la forma y dijese el color. El autor concluyó que la forma es el concepto dominante en la edad preescolar.

Aunque no hay datos que permitan establecer baremos definitivos para niños, en el anexo B presentamos las correcciones por edad basadas en los estudios citados y en nuestras propias investigaciones. Se aconseja, sin embargo, el empleo de estos datos con la necesaria prudencia. A pesar de que los datos son todavía experimentales parece que algunas de las relaciones entre las páginas del Stroop están establecidas con bastante claridad. Por ejemplo, a cualquier edad por encima de los 7 años, en un niño que sabe leer, la puntuación directa PC debe ser inferior a la puntuación directa de color y ésta inferior a la puntuación directa de palabras. Las diferencias entre las tres puntuaciones deberían ser crecientes y acercarse a las normas de adultos hacia los 15 ó 17 años.

## **RETRASO MENTAL**

Diversos autores han estudiado los resultados del Stroop en sujetos retrasados. Silverstein y Franklin (1965) compararon dos grupos de chicos retrasados de unos 17 años de edad con edades mentales de 6-7 y de 10-11 años, midiendo la inteligencia con el test de Vocabulario de Peabody y modificando el Stroop de forma que se orientase con una regleta la línea por la que iba el sujeto para evitar que se perdiese y no supiese por dónde iba. No se encontraron diferencias significativas entre los dos grupos lo que, según los autores, pudo deberse a una inadecuada estimación de la edad mental o al cambio efectuado en la aplicación del Stroop en relación con el sistema normal.

Das (1969) comparó un grupo de sujetos retrasados de entre 10 y 15 años con un grupo de chicos normales de la misma edad aplicando el Stroop y también, a los chicos retrasados, el test de Matrices Progresivas de Rayen. Los chicos con retraso eran más veloces en el nombramiento de colores que en la lectura de palabras lo que les situaba en un grado de desarrollo correspondiente al primer curso o menos (Rand et al., 1963; Comalli et al., 1962). Das (1969) analizó sus datos en términos de la puntuación derivada (C-P)/C que puede considerarse como un factor de dificultad por el color. Los retrasados fueron los más bajos en esta puntuación, un nuevo signo de haber alcanzado sólo un nivel bajo de desarrollo. Das (1969) encontró, también, correlaciones significativas entre las puntuaciones del Raven y las de P (-0,24), C (-0,29) y el factor de dificultad por el color (-0,21): Asimismo encontró correlaciones significativas del CI con P (-0,45) y C (-0,23).

En otro estudio, Das (1970) trabajó con chicos de edad mental comprendida entre 7 y 12 años, comprobando que al aumentar la edad mental eran mayores los efectos de la interferencia, un efecto que ya había sido observado en niños normales de 6 a 8 años (Comalli et al., 1962). El autor encontró una correlación significativa de la edad mental con

las puntuaciones P y C pero no con la de PC. La interferencia se calculó mediante la fórmula  $PC/C$  y se comprobó que correlacionaba positivamente con la edad mental (0,3). A diferencia de lo observado por Das (1969), no se encontró correlación significativa con el CI. Das (1970) indicó que ello podía deberse a que el rango de los CI de la muestra era más restringido. Das (1970) confirmó después las conclusiones de su primer estudio comprobando que las puntuaciones de los retrasados eran inferiores a las de los sujetos normales de igual edad mental en todas las pruebas del Stroop.

Wolitzky, Hofer y Shapiro (1972) sustituyeron las palabras en colores por números a fin de que no fuese necesario saber leer y bastase reconocer los números del 1 al 9. Los sujetos con retraso mental mostraron una mayor interferencia en esta tarea que los normales.

Leisman (1971) obtuvo un resultado similar con niños espastohemipléjicos. Los chicos tenían CI normal de acuerdo con las Matrices Progresivas pero su puntuación de interferencia en el Stroop se mantenía constante después de los 8 años, mientras que en los sujetos normales decrecía constantemente.

Leisman (1971) interpretó estos datos en función de una menor diferenciación e integración de los niños hemipléjicos (Comalli et al., 1962), característica de la edad de 6 a 8 años.

## **INFLUENCIA DEL SEXO**

Entre los investigadores está extendida la opinión de que las mujeres tienen mayor capacidad para nombrar los colores (Brown, 1915; Ligon, 1932; Stroop, 1935; Jensen, 1965). Los datos de Ligon (1932) mostraban diferencias significativas a favor de las mujeres en todas las edades de 6 a 15 años. Stroop (1935) encontró diferencias significativas en la página C en estudiantes de bachillerato. Jensen (1965), con una muestra de 436 personas no tituladas encontró una correlación biserial de 0,23 entre el sexo y la puntuación de nombrar colores (significativa al nivel de confianza de 0,01).

Ligon (1932) y Stroop (1935) no encontraron diferencias significativas entre hombres y mujeres en la página de palabras. Tampoco los estudios de Jensen (1965) y Jensen y Rohwer (1966) aportan diferencias significativas en este aspecto.

Las cosas no son, sin embargo, tan claras cuando se trata de la página colores-palabras. Stroop encontró una ligera diferencia a favor de los hombres pero no significativa estadísticamente. Gardner, Hozman, Klein, Linton y Spence (1959) no encontraron una diferencia significativa en los tiempos de esta página pero detectaron diversos modelos de correlación entre varios tests cognitivos, entre los cuales estaba el Stroop. Desgraciadamente el valor de estos datos se ve afectado por la alta correlación existente entre sexo y edad (0,37) y porque la muestra sólo contenía 30 hombres y 30 mujeres.

Las únicas diferencias significativas entre sexos en la página de color-palabras fueron encontradas por Peretti (1971) que en una muestra de 50 varones y 50 mujeres obtuvo una puntuación media de 48,2 en las mujeres y de 55,1 en los varones, lo que es una diferencia significativa al 0,01. En un estudio anterior, Peretti (1969) encontró diferencias significativas en tres grupos de edad correspondientes a niveles escolares diferentes. En el conjunto las mujeres presentaban resultados mejores que los hombres en PC. Aunque las diferencias no eran significativas, Peretti argumentó que los datos apoyaban la idea de que las mujeres son mejores que los hombres.

Golden (1974) comparó un grupo de 117 estudiantes mujeres con otro de 102 estudiantes varones, empleando la forma colectiva del Stroop. No encontró diferencias significativas en la página de palabras (razón crítica de Student  $t=1,33$ ; g.l.=217) ni en la página de colores ( $t=1,86$ ; g.l.=217). Sin embargo, encontró diferencias significativas a favor de las mujeres en la tercera página ( $t=1,97$ ; g.l.=217). Tampoco aparecía diferencia en la puntuación de interferencia ( $t=0,41$ ). En consecuencia, aunque parece haber una cierta diferencia a favor de las mujeres en el nombramiento de colores, esta ventaja desaparece cuando se tiene en cuenta la interferencia. En todo caso, las diferencias entre grupos son pequeñas y sin trascendencia en el trabajo clínico o de investigación, incluso cuando son estadísticamente significativas.

### 3. DIAGNÓSTICO DE DISFUNCIONES CEREBRALES

El Test Stroop de Colores y Palabras se usa como instrumento para la detección de disfunciones cerebrales solo e incluido en baterías más amplias. El Stroop tiene varias ventajas como instrumento de evaluación: se puede administrar en un tiempo corto, sólo exige un nivel educativo elemental, no está sujeto a influencias culturales y puede traducirse con facilidad a diferentes idiomas. Debe, sin embargo, como cualquier otro test, ser usado con prudencia y conociendo sus puntos fuertes y sus puntos débiles.

Diversos estudios han mostrado que la capacidad de lectura simple y la de nombrar colores pueden verse alteradas en pacientes con daño cerebral. En general, las alteraciones en la lectura de palabras se relacionan con daños en el hemisferio izquierdo, mientras que las alteraciones en el nombramiento de colores pueden deberse a lesiones en el hemisferio derecho (si la dificultad se presenta en la capacidad de reconocer y clasificar los estímulos de color) o en el hemisferio izquierdo (si la incapacidad consiste en no poder atribuir un nombre al estímulo de color) (Golden, 1978).

Perret (1974) aplicó una versión modificada del Stroop a 118 sujetos con lesiones cerebrales unilaterales y limitadas y comprobó que los efectos de la interferencia aparecían más acusadamente en sujetos con daños en el hemisferio izquierdo, sobre todo en la zona anterior o frontal. Golden (1976) usó el Stroop con 141 sujetos con daño cerebral y normales. Como era de esperar, encontró que la puntuación de nombramiento de colores era el mejor indicador de la existencia de lesiones, sobre todo si existía daño en los dos hemisferios cerebrales. Esta única puntuación era suficiente para discriminar los pacientes normales y psiquiátricos de los pacientes con lesiones cerebrales en el 87% de los casos. En dicho estudio se diagnosticó correctamente al 84,6% de los pacientes con lesiones y al 88,9% de los sujetos de control. Los pacientes psiquiátricos fueron clasificados correctamente en el 83% de los casos. El peor resultado se obtuvo con el grupo de sujetos con lesiones en el hemisferio derecho, diagnosticando correctamente el 74% de los casos. Los puntos de corte más adecuados para suponer la existencia de lesiones cerebrales fueron de 75 puntos en la página de palabras, 58 en la de colores y 25 en la de palabras-colores. Sin embargo, para el diagnóstico es más útil el análisis de ciertos modelos o patrones de resultados que establecer puntos de corte. Estos patrones se describen y comentan a continuación.

## **PUNTUACIONES BAJAS EN PALABRAS**

Los sujetos con puntuaciones muy bajas en esta parte (menos de 50), resultados normales en colores y superiores en palabras-colores, presentan un daño cerebral que produce dislexia pura, es decir, incapacidad para leer a pesar de no observarse otras carencias. El Stroop es una herramienta muy eficaz para diagnosticar dislexia, tanto en niños como en adultos. La puntuación P es baja a causa de la incapacidad para leer aunque el nombramiento de colores es normal, como es lógico en una dislexia pura que no afecta a otras funciones. La puntuación clave para diagnosticar la dislexia pura es la PC: si el sujeto es disléxico, habrá poca o ninguna interferencia en la página tercera y la puntuación PC se aproximará a la puntuación C. En los sujetos en que la lentitud en la lectura se deba a problemas psiquiátricos o al rechazo a leer, se producirá una interferencia normal o alta. Esta técnica es, por tanto, de utilidad en el diagnóstico de niños discapacitados.

## **TODAS LAS PUNTUACIONES SON BAJAS**

Este patrón de puntuaciones suele asociarse con lesiones en el hemisferio izquierdo o de tipo difuso. A menudo se producen también problemas de pronunciación, que sugieren la existencia de problemas motores en el habla o con perseveración, que es la repetición de la misma respuesta. La perseveración es con frecuencia indicio de la existencia de lesiones cerebrales en el lóbulo frontal del hemisferio izquierdo. Además, si hay poca interferencia suelen ser lesiones del hemisferio izquierdo, mientras que una interferencia normal sugiere la existencia de un problema más difuso.

## **P NORMAL; C Y PC BAJO**

Este patrón suele asociarse con la idea de lesiones cerebrales en el hemisferio derecho que causan incapacidad para clasificar las claves de color. En algunos casos raros (Golden, 1977) este patrón resulta de la incapacidad para citar verbalmente los colores.

## **P Y C NORMALES; PC BAJO**

Este patrón suele asociarse con lesiones prefrontales aisladas, sobre todo en el lado izquierdo o bilaterales. Es un patrón sensible a formas tempranas de atrofia cerebral que afectan a zonas prefrontales y también a atrofas posteriores producidas por causas tales como el abuso crónico de drogas. Por tanto, es útil para detectar la posible existencia de demencia presenil y otras disfunciones que, de otra forma, presentan manifestaciones escasas.

## **TODAS LAS PUNTUACIONES SON NORMALES**

Este resultado puede darse no sólo en sujetos normales sino también en algunos que tienen daños cerebrales, sobre todo si éstos se han recuperado o si son reducidos y situados en el hemisferio derecho. Por ello un Stroop normal no descarta totalmente la existencia de daño cerebral, pero sí reduce mucho la posibilidad de que exista. La obtención de resultados normales descarta normalmente la existencia de un desorden

amplio y progresivo, que suele producir interferencias significativas en la conducta. De esta forma, si el sujeto presenta algún desorden de conducta importante, deberá pensarse más en la existencia de un problema psiquiátrico o de otro tipo que en la presencia de una disfunción cerebral.

### **OBSERVACIONES CONDUCTUALES**

Como se ha dicho, es importante observar cómo realiza el sujeto el test y no sólo la puntuación obtenida. Por ejemplo, si un sujeto tiene dificultades para seguir el orden correcto de las columnas, aún intentándolo, puede ser indicio de una grave carencia espacial que suele acompañar a disfunciones cerebrales del hemisferio derecho; la perseveración puede indicar la existencia de daños en el hemisferio izquierdo; puntuaciones bajas debidas a la incapacidad de contener la frustración, mostradas negándose a seguir la tarea o mediante reacciones de ira, suelen ser indicativas de la existencia de desórdenes psiquiátricos o de lesiones cerebrales agudas. Generalmente, estas lesiones son suficientemente serias como para producir otros síntomas referentes a problemas motores, sensoriales o de lateralidad.

Hay que recordar que, aunque el Stroop puede usarse con individuos de muy escasa formación, para no cometer errores de diagnóstico en desórdenes del hemisferio izquierdo, es necesario asegurarse que el sujeto adulto no es analfabeto.

### **COMO DISTINGUIR DESÓRDENES ORGÁNICOS DE LOS DE TIPO PSIQUIÁTRICO**

Aunque los pacientes psiquiátricos suelen mostrar dificultades al realizar el Stroop en comparación con los sujetos normales, suele ser posible distinguir los dos grupos. Golden (1976) comprobó que los sujetos psiquiátricos obtenían valores relativamente normales en las tres puntuaciones, con medias de un 10% más bajas que los sujetos normales y, en general, con puntuaciones T superiores a 40. Si las puntuaciones son inferiores a este valor hay que atender a los patrones descritos más arriba. En pacientes psiquiátricos es frecuente encontrar un valor bajo en P, un valor alto o normal en C y un valor bajo en PC, lo cual es muy raro en los pacientes de origen orgánico. Golden (1976) piensa que la puntuación C por sí sola es capaz de discriminar entre los grupos de pacientes orgánicos y psiquiátricos con una fiabilidad del 83%. En el caso de individuos con síntomas de esquizofrenia debe aplicarse la prueba dos veces y debe usarse la segunda puntuación para el diagnóstico.

El uso de ciertos medicamentos puede afectar a los resultados del Stroop, al igual que ocurre con otros tests neuropsicológicos. Por ello, se debe aplicar la prueba cuando el sujeto esté menos afectado por los efectos psicomotores agudos que producen algunas medicinas psicoactivas. Es fácil confundir los efectos de un exceso de medicación con los derivados de daños cerebrales ya que ciertas sustancias afectan a las funciones del cerebro, aunque sus efectos sean temporales.

### **EL STROOP EN BATERÍAS DE TESTS**

El Stroop, por su breve tiempo de aplicación, puede añadirse con facilidad a otras baterías de pruebas. Además, su sensibilidad a los desórdenes derivados del lóbulo frontal, lo hacen especialmente interesante porque muchos otros tests neuropsicológicos no tienen en cuenta esta zona. (Golden, 1978). También su capacidad para discriminar entre pacientes psiquiátricos y orgánicos es de gran valor en las baterías de tests.

## **EJEMPLOS DE CASOS**

### **Paciente nº 1**

El paciente tiene un CI de 135 pero ha tenido problemas en el trabajo en los últimos seis meses sin razón aparente. Los resultados en otros tests aplicados al sujeto son normales pero la fuerte disparidad entre las puntuaciones típicas obtenidas en el Stroop (P=80; C=75; PC=22) hace pensar en la existencia de un daño en el lóbulo frontal. El paciente fue enviado a un neurólogo quien detectó un tumor en el área prefrontal izquierda.

### **Paciente nº 2**

Se trata de un niño de 12 años que se había considerado normal y sin síntomas de dislexia hasta que tuvo un accidente de coche. Las puntuaciones típicas del Stroop fueron para cada prueba de 23, 54 y 46 respectivamente. Este modeló no coincidía con el de dislexia pura (la puntuación PC debería haber sido más alta) por lo que se pensó más en la existencia de un problema emocional que uno de aprendizaje, lo que fue confirmado en terapia.

### **Paciente nº 3**

Las puntuaciones típicas obtenidas, muy bajas (P=22; C=20; PC=12), hicieron pensar en la existencia de un daño cerebral difuso, lo que fue confirmado mediante scanner. El neurólogo diagnosticó al paciente enfermedad de Alzheimer.

### **Paciente nº 4**

El Stroop se aplicó con posterioridad a una intervención quirúrgica de un tumor en el lóbulo frontal-temporal-parietal derecho y obtuvo las siguientes puntuaciones: P=43; C=21; PC=14.

### **Paciente nº 5**

Un chico de 14 años, con dislexia pura, presentaba puntuaciones T muy típicas de esta disfunción: P=8; C=45; PC=75.

### **Paciente nº 6**

Se pretendía discriminar entre un posible episodio esquizofrénico y un daño cerebral orgánico. Sus puntuaciones T fueron en cada página de 40, 43 y 38 respectivamente, lo que indujo a pensar en una reacción esquizofrénica. El posterior diagnóstico del neurólogo y el psiquiatra confirmó esta hipótesis.

### **Paciente nº 7**

Se examinó al sujeto para conocer los efectos residuales de un traumatismo recibido en la cabeza un año antes, obteniendo puntuaciones normales en torno al 50 en todas las variables. Otras pruebas dieron resultados similares con la excepción de ciertos problemas motores. No se hallaron problemas cognitivos de ningún tipo. Este caso demuestra que el

Stroop es poco sensible a daños subcorticales que no afectan al hemisferio cerebral o a los procesos básicos de atención.

#### **4. RESULTADOS DEL STROOP Y ESTRÉS**

La relación entre los resultados obtenidos en el Stroop y ciertos factores asociados con el estrés ha sido un tema de gran interés. Esta prueba se utiliza tanto como test para provocar estrés como para medir el nivel de estrés. Los datos disponibles muestran claramente la utilidad del Stroop para investigar las características psicológicas y fisiológicas del estrés.

#### **ESTRÉS**

Agnew y Agnew (1963) estudiaron los efectos del estrés sobre los resultados del Stroop. Se aplicó el test de Stroop y los laberintos de Porteus a 50 universitarios en condiciones normales y en condiciones de estrés. El estrés se provocaba diciendo a los sujetos que recibirían una descarga cada vez que tuviesen resultados inferiores al nivel "medio". Además, se les decía que los resultados estaban relacionados con la inteligencia y que, como universitarios, debían hacerlo bien. La ausencia de estrés se producía diciendo a los sujetos que sólo se quería conocer el test, sus puntuaciones y sus instrucciones y que la tarea podía resultar aburrida.

Los autores usaron el valor (PC + C) para medir los resultados del Stroop, suponiendo que la tensión afectaría a las páginas de colores y colores-palabras. En la primera aplicación se encontraron diferencias significativas entre los sujetos sometidos a alta presión y los que no lo estaban. Sin embargo, una segunda aplicación, habiendo invertido la situación para cada sujeto, no mostró diferencias entre ambos grupos, lo que llevó a los autores a suponer que el estrés es una variable significativa sólo cuando el test es nuevo para el sujeto.

Hardinson y Purcell (1959) usaron el Stroop para clasificar los sujetos en dos grupos, flexibles y rígidos. Después fueron clasificados en dependientes o independientes con la Escala de Preferencias de Personalidad de Edwards. De los 130 sujetos disponibles originalmente se incluyeron 15 en cada grupo, manteniendo un total de 60 sujetos.

A ciertos sujetos se les provocó el estrés mediante refuerzos verbales, amenazas, etc. En el conjunto no se encontraron diferencias en rendimiento debidas al estrés pero analizando los grupos menores se comprobó que los sujetos del grupo flexibles e independientes lo hicieron mejor que los del grupo dependientes y rígidos. La puntuación empleada en el Stroop fue PC menos PC'.

Teece y DiMartino (1965) comprobaron que el estrés aumentaba los resultados cuando se mantenían constantes los estímulos auditivos y visuales. Sin embargo, el rendimiento decrecía cuando el estímulo auditivo era diferente, efecto análogo al uso de la página PC. Hochman (1967) encontró que el rendimiento en PC decrecía cuando se usaba una versión del Stroop sobre taquistoscopio y había fuerte presión.

Houston y Jones (1967) evaluaron los resultados del Stroop con tres condiciones: silencio, ruido que había que ignorar y ruido al que había que atender. Se usaron sólo las páginas 2a y 3a con 42 sujetos. Las puntuaciones en C fueron muy similares en las tres situaciones y las puntuaciones PC se incrementaron ligeramente cuando había ruido al que no se debía atender. En otro estudio posterior, Houston (1969) comprobó que las puntuaciones PC eran más altas en condiciones de ruido que de silencio. Consideró que estos resultados sugerían



que la necesidad de ignorar estímulos adicionales (ruido) podía favorecer el rendimiento del sujeto en una tarea que también exigía ignorar palabras.

## **DROGAS**

Hollingsworth (1923) indicó que el alcohol dificultaba la tarea de nombrar colores duraderos que las dosis pequeñas y que la cuantía de la interferencia producida por que las dosis mayores producían efectos más una dosis de alcohol era de tipo muy individual. El mismo autor (1912) encontró que la cafeína no producía efectos sobre el nombramiento de colores, independientemente de las dosis administradas.

Estos primeros estudios de Hollingsworth despertaron mucho interés por estudiar los efectos de diversos tipos de drogas sobre el rendimiento de tests como el Stroop.

Callaway y colaboradores usaron el Stroop para estudiar un fenómeno que llamaron "atención concentrada" (Callaway y Dembo, 1958) que, según los autores, se debía a drogas estimulantes y permite al sujeto concentrar su atención e ignorar estímulos irrelevantes. De esta forma el sujeto rinde mejor en pruebas como la página PC del Stroop ya que puede ignorar el estímulo que supone la palabra. Asimismo, la interferencia disminuye porque en la página C no hay estímulos distractivos que eliminar.

Según esta teoría, las drogas depresoras deberían producir un efecto contrario: peores resultados en PC y en interferencia. Este efecto se produce porque las drogas amplían el campo de atención y hacen más difícil que el sujeto suprima los estímulos de distracción.

Callaway (1959) estudió este tema aplicando a tres grupos independientes de universitarios amobarbital (un depresor), metanfetamina (estimulante) y un placebo. Se examinó a los sujetos con el Stroop antes de tomar el producto y una hora después. El grupo del estimulante obtuvo menor interferencia que el grupo del placebo mientras que en el grupo del depresor aumentaba la interferencia.

En un estudio posterior Callaway y Stone (1969) revisaron sus datos y mostraron que los estimulantes mejoraban todas las puntuaciones del Stroop mientras que los depresores las empeoraban.

Quarton y Talland (1962) encontraron diferencias significativas en las tres puntuaciones básicas del Stroop entre dos grupos a los que se había suministrado un depresor (pentobarbital) y un estimulante (metanfetamina). El cambio era igual en las tres páginas.

Callaway y su equipo (Callaway y Band, 1958; Callaway y Dembo, 1958) estudiaron los efectos de la atropina en el Stroop. Los resultados se complicaban porque la atropina puede ser estimulante o depresor según la dosis administrada y porque puede producir efectos visuales secundarios. Los resultados fueron análogos a los de estudios anteriores.

Ostfield y Aruguete (1962) comprobaron que la hysocina disminuía los resultados por lo que se pensó que se trataba de un depresor. Esto se confirmó al comprobar que este producto también produce una depresión de la respuesta en el electroencefalograma.

Wapner y Krus (1960) comprobaron que los sujetos normales obtenían mejores resultados en el Stroop que los esquizofrénicos, tanto habiendo tomado un placebo como una droga (LSD). También se vio que el LSD empeoraba significativamente los resultados en cualquiera de los grupos, lo que se interpretó considerando que esta droga produce un resultado más inmaduro desde el punto de vista evolutivo (ver capítulo anterior).

Lehman y Knight (1961) encontraron mejores resultados con la ingestión de Anfetamina-d si el sujeto era adicto pero no si era normal.

Baskowitz, Korchin y Oken (1960) no encontraron efectos sobre los resultados del Stroop al administrar adrenalina. Por otra parte, Miller, Drew y Kiplinger (1972) tampoco vieron resultados significativos administrando marihuana a los sujetos.

## **CAMBIOS EN EL SISTEMA NERVIOSO AUTÓNOMO**

Oken, Grinker, Heath, Herz, Korchin, Sabshi y Schwartz (1962) investigaron la reacción al estrés de dos grupos con puntuaciones significativamente diferentes en el Stroop. El estrés se produjo administrando a los sujetos una cápsula y diciéndoles que se estaba haciendo un estudio sobre drogas y que la cápsula, entre otros efectos, podía producir problemas cardíacos. Durante el tiempo que los sujetos estaban preocupados se hacían diversas mediciones como tasa cardíaca, tasa respiratoria, resistencia de la piel, GSR, presión sanguínea y motilidad estomacal. Sólo se produjo una diferencia significativa entre los dos grupos en la medición de la tasa respiratoria.

Frankenhaeuser, Froeber, Hagdahl, Rissler, Bjorkvall y Wolff (1967) midieron las respuestas fisiológicas del sujeto en situación de estrés provocada con una versión modificada de la página palabras-colores del Stroop. Los elementos del Stroop (nombres de colores en un color incongruente) se presentaban de uno en uno a un ritmo rápido y variable. Al mismo tiempo, el sujeto oía los nombres de los colores por medio de auriculares. Por último, se decía a los sujetos que la mayor parte de los estudiantes lo hacían muy bien después de un poco de práctica. La medición realizada a las 15 mujeres que formaban el grupo fue la conductancia de la piel y se observó que ésta decrecía a medida que la tarea se hacía más fácil y menos estresante, de lo que puede deducirse que la tarea de la tercera página produce estrés en los momentos iniciales de su aplicación.

Sin embargo, la mayor parte de los estudios realizados no han demostrado la existencia de cambios en el sistema nervioso autónomo durante la aplicación del Stroop. Houston (1969) utilizó dos versiones de la prueba con colores más o menos fácilmente distinguibles y no encontró diferencias en el pulso de los sujetos. Lo mismo ocurrió en situaciones de silencio y ruido. Trackray y Jones (1971) tampoco encontraron alteraciones de la respiración y la tasa cardíaca al aplicar PC, por lo que supusieron que el Stroop sólo afecta a procesos de la atención que tienen poca relación con el estado

general. Daniel (1971) sugirió que estos resultados podían en parte deberse a las grandes diferencias existentes entre sujetos.

## **NIVELES HORMONALES**

Frankenhaeuser et al. (1967) descubrieron que al aplicar versiones fuertemente estresantes del Stroop se producía una disminución de las secreciones de adrenalina a medida que la situación de estrés iba siendo menor, mientras que la noradrenalina sólo presentaba fluctuaciones aleatorias, lo que coincide con la idea de que la tarea del Stroop produce estrés. En un segundo estudio, Frankenhaeuser, Mellis, Rissler, Bjorkvall y Parker (1968) dividieron a los sujetos en dos grupos atendiendo tanto a las secreciones de adrenalina como de noradrenalina en momentos de inactividad y de estrés producido por el Stroop, encontrando que los sujetos que lo hacían mejor en situación de estrés también tendían a tener tasas de secreción hormonal altas, particularmente importantes en el caso de la noradrenalina.

Los trabajos de Broverman y su equipo (1972) se centraron en las relaciones fisiológicas con la capacidad de "automatización", que correlaciona positivamente con los resultados obtenidos en el Stroop.

Los varones fuertemente automatizadores tienden a producir niveles más altos de secreción de las hormonas masculinas (Klaiber, Broverman y Kobayashi, 1967) lo que puede indicar un mayor desarrollo de las características sexuales secundarias en este tipo de sujetos. En segundo lugar, Broverman, Klaiber, Kobayashi y Vogel (1968) informaron que la capacidad de automatización depende de la función adrenérgica central. Las tareas que definen la automatización (p.ej.: nombramiento de colores) parecen mejorar en presencia de las anfetaminas, sustancias que estimulan la actividad del sistema nervioso adrenérgico. La cloropromacina, por el contrario, bloquea estos procesos y reduce los resultados de las capacidades automatizadas.

Broverman et al. (1972) consideran que las hormonas sexuales inhiben la actividad de la monamina oxidasa (MAO). Una actividad excesiva de la MAO perjudica al sistema nervioso central adrenérgico y es, por tanto, contraria a la automatización. Como consecuencia de ello, las hormonas sexuales, la testosterona y los estrógenos pueden incrementar la automatización inhibiendo la acción de la MAO.

## **EEG (ELECTROENCEFALOGRAMA)**

Vogel, Breverman, Klaiber y Kun (1967) supusieron que si los sujetos muy automatizadores son de tipo dominante adrenérgico ello debería poderse ver en el electroencefalograma y que los sujetos fuertemente automatizados deberían mostrar menos respuestas de tensión. Citan trabajos en los que las "respuestas de pulsión" son menos frecuentes cuando se ingiere droga, porque éstas inducen estados adrenérgicos. Los resultados del estudio confirmaron esta hipótesis.

## **5. STROOP, PERSONALIDAD Y COGNICIÓN**

Debido a las grandes diferencias individuales encontradas al aplicar el Stroop, algunos autores han investigado las relaciones de la prueba con varias dimensiones de la personalidad y el área cognitiva. El primer estudio de este tipo se hizo por Thurstone y Mellinger (1953), intentando relacionar los resultados del Stroop con un estilo de personalidad caracterizado por un alto grado de control y de tolerancia a la frustración. Pero el estudio no permitió llegar a conclusiones válidas porque empleó una medida del Stroop que mezclaba otros factores, como la velocidad de lectura.

Algunos autores han intentado relacionar el Stroop con las dimensiones de Eysenck, extraversión (E) y neuroticismo (N). Jensen (1965) aplicó el Stroop y el Inventario de Personalidad de Maudsley a 436 sujetos, encontrando correlaciones significativas al nivel 0,05 de E con C (-0,13), con P (-0,10) y con PC (-0,14) y de N con C (0,11) y con PC (0,10). Una vez corregidas para tener en cuenta la fiabilidad de las puntuaciones de interferencia, de N y de E, se encontró una correlación entre interferencia y E de -0,12 (significativa al 5%) y de 0,06 entre interferencia y N (no significativa). En todo caso, se aprecia que las correlaciones son muy bajas.

Callaway (1959), con una muestra de sólo 28 sujetos, encontró una correlación de 0,43 entre interferencia y extraversión. Alperson (1968) no encontró correlación significativa entre el Stroop y las escalas E y N.

Daniel y Skondackove (1970) tampoco encontraron correlación al usar la versión estándar del Stroop pero sí cuando usaron una versión con mayor interferencia procedente de una cinta que recitaba nombres de colores. En este caso no había correlación con N pero sí de E con PC (0,42) y con interferencia (0,52).

Stein y Lager (1966) aplicaron el Stroop y el ACL (Lista de Adjetivos) a 81 mujeres, sin que se apreciaran diferencias significativas en el conjunto. Después tomaron los 10 sujetos con interferencia más alta y los 10 sujetos con interferencia más baja y los compararon con los baremos del test, comprobando que el grupo de interferencia alta se desviaba de las normas generales en 13 de las 24 escalas del ACL mientras que el grupo de baja interferencia no presentaba diferencias. Los autores no compararon los dos grupos pero los datos del estudio muestran que diferían en nueve de las 24 escalas. Los autores interpretan los resultados identificando el grupo de baja interferencia como mujeres de personalidad media normal, mientras que el de alta interferencia es presentado como "un grupo de mujeres que actúan con menor eficacia y tienen una capacidad menor para adaptarse y resolver problemas internos e interpersonales. El grupo de alta interferencia (GAI) está compuesto por personas con ansiedad y que tienden a verse como personas con mayores problemas que los demás. Los conflictos que sienten se centran en torno a impulsos agresivos y a inclinaciones sexuales controladas".

## **CREATIVIDAD**

Gamble y Kellner (1968) supusieron que las personas creativas harían mejor el Stroop que las no creativas. Como medida de la creatividad usaron el test de Asociaciones Remotas de Mednicks, lo que resultó ser una elección incorrecta porque este test tiene un gran peso de factores como la capacidad verbal y el cociente intelectual.

Golden (1975) examinó a 450 estudiantes con la versión colectiva del Stroop y calculó la correlación entre la puntuación de interferencia del Stroop y varias medidas de la creatividad de tipo verbal (Test Improvement de Rowenton), no verbal (Test Matchstick

de Guilford) y calificaciones dadas por los profesores. Se encontraron diferencias significativas en las tres puntuaciones entre los estudiantes con resultados altos, medios y bajos en la puntuación de interferencia. Las calificaciones de los profesores presentaron una correlación de 0,42 con la puntuación de interferencia (N=212). Golden opinó que el Stroop se relaciona con la creatividad porque exige que el sujeto conciba con rapidez y precisión nuevas estrategias cognitivas para tratar estímulos sencillos. Indicó también que el Stroop frente a otros tests de creatividad tiene la ventaja de no requerir elevada capacidad verbal, ni destreza manual, ni un determinado nivel cultural o de experiencia.

En un estudio posterior, Golden (1976) comprobó que los sujetos con buen resultado en la medida de interferencia del Stroop se veían más seguros de sí, más maduros, más abiertos a la investigación y más dispuestos a rechazar ciertos comportamientos ligados con psicopatologías. En conjunto, estos rasgos coinciden con los descritos por Dauw (1966) para el pensamiento sano y creativo.

## COGNICIÓN

Klein (1954) definió el control cognitivo como el mecanismo que actúa "para resolver exigencias de adaptación inmediata". Examinó a 100 estudiantes con el Stroop para formar grupos de 20 sujetos que fuesen altos o bajos respectivamente en interferencia. Cada grupo fue después dividido en subgrupos que estuviesen deseosos o satisfechos. Aplicando una tarea de percepción para estimar tamaños, Klein comprobó que los sujetos con alta interferencia (GAI) tendían a sobreestimar el tamaño de los objetos mientras que los sujetos con baja interferencia (GBI) tendían a infravalorar el tamaño.

Las personas GAI mostraban una menor variabilidad, lo que, según Klein, parece indicar "una forma de control más estrecha o supresiva que refleja, quizás, esfuerzos concretos para mantener los juicios en línea con lo que indican las fuentes de información externas". Klein supone que se trata de una tendencia a "empaquetar" cada estímulo y reducir solapamientos dentro de los límites del campo cognitivo. Más exactamente, toleran mal la ambigüedad y desean que las cosas estén establecidas con claridad. Como consecuencia de ello, sus resultados en el Stroop se ven perturbados por el solapamiento de estímulos y la ambigüedad provocada.

Las personas GBI, por el contrario, muestran mayor flexibilidad y pueden guiar sus juicios en un contexto más amplio, incluyendo sus propios sentimientos. Tienen una mayor tolerancia a la ambigüedad. Klein (1954) no aventura ninguna hipótesis sobre por qué tienden a sobreestimar el tamaño del estímulo.

Klein (1954) encontró también que los sujetos GAI (con alta interferencia) y deseosos tendían a ver en un taquistoscopio un campo periférico menor que los sujetos deseosos o insatisfechos y con baja interferencia (GBI).

Esta diferencia no se producía en el caso de los sujetos satisfechos. Los sujetos GAI de este estudio también hacían identificaciones más correctas de los estímulos.

En un estudio realizado con el test de Rorschach, Podell y Phillips (1959) hicieron un análisis factorial con una amplia batería de tests que incluía el Stroop. Emplearon 20 tests

cognitivos que permitían obtener 36 puntuaciones diferentes en dos muestras reducidas de varones, una de 32 sujetos y otra de 37, lo que nos obliga a considerar los resultados con cierta cautela ya que lo reducido de la muestra hace que los resultados factoriales sean poco fiables. Los autores encontraron que el Stroop saturaba significativamente en dos factores: fluidez verbal (puntuaciones C y PC) y labilidad (PC). La labilidad se definía como la capacidad para dar respuestas independientemente de las respuestas anteriormente dadas.

Loomis y Moskowitz (1958) detectaron que los sujetos GBI son más capaces de integrar los solapamientos y los estímulos ambiguos, lo que coincide con las experiencias de Klein (1954). Hochman (1971) comprobó que los sujetos GBI son menos susceptibles a los estímulos competitivos.

Golden (1975) dividió una muestra de 219 estudiantes en dos grupos, respecto al valor PC' (según la fórmula del anexo A). Cada grupo se sometió a 20 mediciones cognitivas y el autor dedujo que tres tipos de tests se relacionaban con los resultados de la puntuación PC del Stroop:

1. Tareas sensibles a perturbaciones externas (principalmente tareas automáticas).
2. Tareas que contienen estímulos irrelevantes o distractivos.
3. Tareas en las que las propias respuestas del sujeto actúan como elementos que inhiben los resultados posteriores.

A partir de estos criterios, el autor infirió que el grupo que obtenía mejor resultado que el previsible (alta resistencia a la interferencia) tendría mejores resultados en 12 de las 20 tareas incluidas en el estudio. Los resultados demostraron un rendimiento mejor en 11 de los 12 tests previstos, mientras que en el duodécimo el rendimiento era también algo mejor pero no significativamente. En los restantes 8 tests ninguno arrojó una "t" (de Student) superior a 0,97 al comparar los dos grupos. El autor indicó que la habilidad básica medida por el Stroop es la habilidad para seleccionar de forma flexible información relevante del entorno y, por tanto, para adaptarse a nuevas circunstancias y para tener en general una actuación más efectiva. Asimismo, dedujo que estos individuos deberían tener menor tendencia a desarrollar conflictos psicopatológicos. Por último, Golden indicó que el Stroop podría ser una buena medida de las mejoras obtenidas en psicoterapia en aquellos casos en que se desea que el sujeto evolucione desde conductas rígidas y defensivas hacia otras más abiertas, flexibles y en contacto con las exigencias de la realidad.

## **AUTOMATIZACIÓN**

El concepto de automatización es esencial en los estudios realizados por Broverman y sus colaboradores durante 16 años. El enfoque inicial se dirigió a estudiar ciertas dimensiones cognitivas en relación con las tres puntuaciones básicas del Stroop, pero después fue evolucionando hacia investigaciones más amplias referidas a aspectos psicológicos, cognitivos y fisiológicos de la automatización poco relacionados con el Stroop (Broverman, Clarkson, Klaiber y Vogel, 1972).

Lazarus, Baker, Broverman y Mayer (1957) usaron en principio el Stroop para diferenciar dos tipos cognitivos, verbal y no verbal, usando el ratio o cociente P/C para medir esta

dimensión. Si el valor obtenido era superior a 1,4 el sujeto era definido como verbal y como no verbal en caso contrario. Los sujetos verbales eran más capaces de reproducir literalmente los párrafos en una tarea de reproducción de textos. Los autores atribuyeron este hecho a que los sujetos verbales son más maduros (de acuerdo con el trabajo de Comalli et al., 1962) y más habituados a usar símbolos abstractos (verbales).

En otro estudio, Broverman y Lazarus (1958) llamaron "conceptual dominante" (CD) al grupo verbal y "perceptivo motor dominante" (PMD) al grupo no verbal. Dedujeron que cuanto más fuerte fuese un subsistema cognitivo (como leer) menos ligado estaría a distracciones. Los sujetos conceptuales están dominados por su sistema verbal y, por tanto, deben ser menos susceptibles de distracción. Formularon también la hipótesis que cuando dos subsistemas entran en conflicto el más dominante de los dos prevalece en perjuicio del sistema menos dominante. Se comprobó que, en condiciones de distracción, los sujetos conceptuales eran más capaces de mantener su conducta en tareas numéricas de suma.

Broverman (1960) consideró necesario añadir una segunda dimensión cognitiva: la capacidad de automatización. Afirmó que la distinción conceptual-perceptivo se limita sólo a tareas que son nuevas, difíciles o que exigen concentración, mientras que cuando se trata de tareas bien conocidas la automatización se convierte en lo más importante, tanto si la tarea es originalmente de tipo conceptual como si es de tipo perceptivo motor.

Broverman (1960) dijo que teóricamente el resultado en una tarea nueva o difícil depende mucho de la transferencia de conductas previamente aprendidas. La cantidad de transferencia posible depende, a su vez, de la experiencia existente con tareas similares. La experiencia a lo largo de la vida de diferenciación conceptual o perceptiva y motora determina que una persona esté dominada por unos u otros tipos de aspectos.

Sin embargo, después de un entrenamiento específico, el resultado en la tarea depende más de dicho entrenamiento que del aprendizaje pasado en general. Si el nivel de entrenamiento es elevado, el aprendizaje tiende a disminuir la superioridad innata de algunas personas. Si se supone que esta superioridad es constante en diferentes conductas, las tareas más aprendidas (automatizadas) se acercan a dicho límite y reflejan las diferencias de los sujetos en cuanto a aprendizaje. Esta diferencia es la que subyace en los resultados diferentes de sujetos muy o poco automatizadores.

Broverman (1962) se sintió insatisfecho con esta definición inicial y volvió a definir la automatización subyacente como la tasa con la que los sujetos eran capaces de beneficiarse de la práctica. Demostró que la curva de aprendizaje de los sujetos fuertemente automatizadores era más acusada de forma que después de varios intentos se producía una diferencia significativa entre los dos grupos. Se supone, sin embargo, que el sujeto débilmente automatizador puede alcanzar un rendimiento similar si puede realizar un número adicional de ensayos.

Broverman indica también que la fatiga puede explicar algunas de estas diferencias. Si un sujeto se cansa rápidamente al repetir la tarea, obtendrá menos beneficios de la práctica que un sujeto que no se fatigue tan pronto. Así, la hipótesis de que los sujetos automatizadores se cansan menos rápidamente que los poco automatizadores, permitirá predecir que los sujetos que se fatigan en seguida obtendrán peores rendimientos en tareas que exigen una repetición rápida y sostenida de un número de respuestas (p.ej.: nombrar colores).

Broverman (1960) confirmó estas hipótesis comprobando que la capacidad de automatización puede calcularse mediante la interferencia del Stroop ( $PC - PC'$ ). También demostró que en los sujetos automatizadores se dan menos distracciones y que son capaces de ejecutar mejor tareas repetitivas (como afilar lápices).

Broverman (1962) también introdujo cambios en la forma de calcular el índice de automatización de un sujeto, a partir de datos obtenidos de una amplia batería de tests cognitivos. Todas las puntuaciones, incluidas las del Stroop, se convirtieron en puntuaciones típicas y se les restó la puntuación típica media del sujeto, lo que daba una medida diferencial para cada tarea del nivel relativo al que el sujeto hacía cada tarea en relación con su nivel medio. Estos datos se sometieron a un análisis factorial de conglomerados (método acumulativo de Tryon) en ocho muestras que iban desde varones adultos y parejas casadas hasta varones y mujeres delincuentes y pacientes esquizofrénicos. Estos análisis definieron factores de automatización muy similares cuya saturación mejor era la puntuación diferencial típica de PC, con un valor que oscilaba entre 0,31 y 0,72.

En una muestra posterior de 206 gemelos, el peso de PC fue de 0,42 mientras que el de C era 0,59 y el de P era de 0,60. En la muestra más grande usada por Broverman (1962) el peso de PC fue de 0,50 mientras que el de C fue de 0,68, lo que sugiere que C es más adecuado para predecir la automatización que PC. Un estudio posterior (Broverman, Broverman y Klaiber, 1966) con 50 sujetos varones produjo un resultado equívoco en este punto: la tarea de nombrar colores tenía una saturación de 0,70 mientras que la de PC era 0,66 en el factor de automatización.

Broverman (1962) y Broverman (1964) comprobaron que los sujetos conceptual dominante (CD) tienen mayor nivel cultural y mejores profesiones que los PMD y que los sujetos automatizadores tienen niveles profesionales más altos que los poco automatizadores pero no mayor nivel cultural. Se comprobó también que esto no se debía a diferencias socioeconómicas: los automatizadores habían tenido de niños un nivel socioeconómico inferior al que habían tenido los poco automatizadores.

Broverman (1962) comparó la automatización en 20 parejas casadas, encontrando que en 10 casos era mejor el resultado del marido y en los otros 10 era mejor el de la mujer. También se aplicó a las parejas un cuestionario que incluía dos preguntas dirigidas a evaluar el éxito de su matrimonio: "En general, ¿está Ud. satisfecho de su matrimonio?" y "¿Qué nivel de satisfacción cree que tiene su pareja en el matrimonio?". Las parejas mostraron un mayor nivel de satisfacción cuando el marido era el más fuertemente automatizador y viceversa. Se obtuvieron los mismos resultados usando tanto la medida de automatización de Broverman más antigua (1962) como la más moderna (1964). Sería interesante investigar si estos resultados están relacionados con las actitudes sobre el modelo de matrimonio, por ejemplo, sobre quién debe jugar un papel dominante en la familia. Broverman (1962) indica que los fuertemente automatizadores son más dominantes y asertivos, lo que nos daría una imagen tradicional del matrimonio: el matrimonio es más feliz si el marido asume su papel "natural" y es menos feliz cuando la mujer es la dominante o cuando ésta siente que el marido no asume sus responsabilidades maritales. Si esta relación existe, sería también interesante determinar si se mantiene en un matrimonio en que ninguna de las dos partes es dominante.

Broverman (1962) informa también que los más automatizadores asumen la mayor parte de las responsabilidades cotidianas del matrimonio, de donde deduce que la capacidad



de automatización ayuda a manejar la casa y la familia. Sugiere, asimismo, que es tal vez este mismo factor el que explica el éxito laboral de los automatizadores.

En una muestra de 32 chicas delincuentes con una media de edad de 14 años, Broverman (1962) descubrió que las más automatizadoras escogían amigos más jóvenes mientras que las poco automatizadoras preferían amigos mayores, lo que puede indicar respectivamente un deseo o necesidad de dominar o de ser dominado. Las chicas más automatizadoras empezaban a tener novio más tarde, lo que Broverman interpretó como un deseo de evitar asumir el tradicional papel femenino pasivo. En el caso de delincuentes varones, los fuertemente automatizadores tienden a rebelarse o a adoptar el papel de "malos" en la escuela dos años antes que los demás. Entre los esquizofrénicos crónicos, los más automatizadores tenían tendencia a ser más agresivos.

En conjunto, estas observaciones hacen suponer que "la automatización... se presenta en diferentes formas y contextos como un impulso interno hacia la dominancia interpersonal, la independencia personal, la evitación de los papeles sociales de sumisión, el inconformismo y la rebeldía a la autoridad" (Broverman, 1962).

Los automatizadores tienden también a ser menos indulgentes consigo mismos. Broverman (1962) comprobó que las chicas menos automatizadoras empezaban a beber alcohol un año antes y tenían su primera relación sexual 10 meses antes, mientras que los chicos poco automatizadores se masturbaban con mayor frecuencia y se casaban dos años antes.

Broverman, Broverman y Klaiber (1966) comprobaron que la fatiga es un factor significativo de diferenciación entre los sujetos más o menos automatizadores cuando el número de ensayos es muy elevado, lo que confirmó las hipótesis previas del propio Broverman (1962).

Broverman (1962) intentó encontrar razones que justificasen el mayor desarrollo de los fuertemente automatizadores y observó que los padres de éstos solían tener ocupaciones más bajas que las de los padres de los menos automatizadores. Como explicación de este hecho sugirió que entre los valores de las clases medias y bajas figura la necesidad de luchar por salir adelante mientras que las clases altas forman niños más conformistas y menos motivados. Reconoce, sin embargo, que la motivación no puede explicar todas las diferencias psicológicas citadas más arriba entre automatizadores y no automatizadores.

También se observó (Broverman, 1962) que entre los automatizadores había muchos primogénitos y que las chicas delincuentes poco automatizadoras solían tener madres muy jóvenes o muy mayores, lo que, sin embargo, no se comprobó en el caso de los chicos delincuentes.

Broverman (1962) indicó que en parejas casadas los más automatizadores solían despertarse más temprano que los poco automatizadores, lo que indica que la actividad del sistema nervioso central puede estar relacionada con la automatización. Esta idea viene también apoyada por los estudios sobre las hormonas del sistema nervioso central (Klaiber, Broverman y Kobayashi, 1967).

La automatización está relacionada con ciertas disfunciones del aprendizaje (Broverman et al., 1962). Así, los problemas de lenguaje y lectura son más frecuentes en los chicos

que en las chicas, al igual que ocurre con la baja automatización. Los niños con dificultades para leer con fluidez en voz alta tienen resultados significativamente peores que los otros niños en tareas de automatización. Estos lectores con dificultades no presentaban, sin embargo, dificultades para realizar las pruebas de Comprensión, Vocabulario o Semejanzas del WISC. Broverman (1964) comprobó que tanto los que tienen dificultades para leer como los poco automatizadores tienen peores resultados en tareas de tipo perceptivo. Mathewson obtuvo resultados similares en su tesis doctoral (Broverman et al., 1972).

Broverman et al. (1972) encontraron también relación entre hiperactividad y automatización. Las características de conducta son, según los autores, las asociadas a la baja automatización: "se distraen fácilmente, se cansan en seguida de una tarea... y tienen cambios de atención fluctuantes y, en consecuencia, dificultades para concentrarse y para obtener buenos resultados en forma estable" (Broverman, Clarkson, Klaiber y Vogel, 1972). Los poco automatizadores, de acuerdo con lo indicado más arriba, son deficientes en adrenergia central y, si esta teoría es cierta, como afirma Broverman, puede ser beneficioso tratarles con estimulantes adrenérgicos.

La misma idea se recoge en un estudio sobre chicos preadolescentes (Broverman et al., 1972) en el que los autores comprobaron que los chicos hiperactivos ejecutaban el test como los poco automatizadores, en contraposición a los resultados obtenidos con muestras de chicos normales. El mismo esquema no se producía en niños más pequeños (6 a 8 años), que tenían resultados peores en todas las tareas mientras que los poco automatizadores tenían resultados relativamente mejores en tareas perceptivas. Según los autores esto pone en tela de juicio las teorías fisiológicas ya que hay un cambio crítico en la capacidad de reestructuración hacia los 9-11 años. A partir de investigación con animales, los autores han propuesto una teoría para responder a estos hechos.

En la fase I los niños son relativamente adrenérgico dominantes porque los procesos colinérgicos que actúan como inhibidores adrenérgicos todavía no están en funcionamiento.

En la fase II aumenta la actividad adrenérgica establemente pero los procesos colinérgicos se hacen operativos y alcanzan una tasa muy elevada. Como los más automatizadores tienden a madurar antes, es fácil que a los 6-8 años encontremos a los automatizadores en la fase II mientras que los poco automatizadores aún estarían en la fase I y, por ello, los primeros superarían a los segundos tanto en las funciones adrenérgicas (fuerte automatización) como en las colinérgicas (automatización débil).

En la fase III las actividades colinérgicas pasan a dominar a las adrenérgicas. En la fase IV se restablece la superioridad de los procesos adrenérgicos. La velocidad y duración de cada fase depende de las cantidades de testosterona por medio de la acción de MAO.

## **6. EL STROOP Y LA PSICOPATOLOGÍA**

Otro campo de investigación ha sido el estudio de las relaciones entre la propensión a la interferencia y la psicopatología. Wapner y Krus (1960) encontraron una reducción

significativa en la rapidez de los esquizofrénicos en las tres páginas del Stroop. Weiss y Sherman (1962) encontraron una correlación de 0,40 entre la Escala de Ansiedad Manifiesta y la interferencia del Stroop en esquizofrénicos crónicos pero no en esquizofrénicos agudos. Asimismo, obtuvieron una correlación de 0,67 en un grupo de paranoicos de largo plazo mientras que los paranoicos de corto plazo presentaban una correlación de -0,41.

Peixotto y Rowe (1969) examinaron con el Stroop a esquizofrénicos, psiconeuróticos y sujetos normales, presentando los elementos por medio de taquistoscopia y añadiendo una cuarta página elaborada por Klein (1964) que consistía en palabras asociadas a color de manera remota, impresas en colores incongruentes. Los autores obtuvieron diferencias significativas entre los grupos tanto en la página PC como en la adicional. El análisis se basó en la técnica de la covarianza múltiple con correcciones por edad e inteligencia. No se hicieron correcciones sobre los niveles generales de rendimiento en P y C.

Golden (1976) halló diferencias significativas en las tres puntuaciones básicas del Stroop entre un grupo de sujetos normales y otro de pacientes psiquiátricos siendo, sin embargo, mayor la diferencia en PC que en las páginas de palabras y de colores. La media de interferencia en los pacientes normales fue 0,0 (lo esperable) mientras que en los pacientes psiquiátricos fue de -6,1.

Golden (1975) clasificó a los sujetos en tres grupos según su perfil en el MMPI:

- A. Puntuación alta en la escala 2 (D) o en la escala 8 (Sc).
- B. Puntuación alta en las escalas 1 (Hs), 3 (Hy), 7 (Pt) y 9 (Ma).
- C. Puntuaciones altas en cualesquiera de las otras escalas.

Todos los sujetos eran estudiantes sin antecedentes psiquiátricos. Los resultados mostraron diferencias significativas entre los tres grupos en la puntuación de interferencia. Los sujetos con puntuaciones altas en Sc o en D (grupo A) tenían peores puntuaciones, mientras que los sujetos del grupo C, sin problemas psicopatológicos, mostraban más resistencia a la interferencia en el Stroop. El autor indicó que los sujetos con mayor tendencia a la interferencia tienen también mayor tendencia a seguir conductas estereotipadas características de los perfiles psicopatológicos. Estas conductas estereotipadas son ligeras en ambientes libres de estrés pero pueden causar problemas neuróticos, e incluso psicóticos, serios en condiciones de estrés. Así, los sujetos con mejores resultados en el Stroop son también los más resistentes al estrés como demuestran sus perfiles más normales en el MMPI.

### **ESTUDIOS CON PUNTUACIÓN "SERIAL"**

Smith y Nyman hicieron el proyecto de investigación más ambicioso sobre los resultados del Stroop con pacientes psiquiátricos, pese a lo cual no se ha prestado suficiente atención a este trabajo. Jensen y Rohwer (1966) piensan que ello se debe a que es un estudio difícil de leer e interpretar y a que los autores usaron un método de corrección del Stroop llamado puntuación "serial".

En esta técnica, los autores usaron sólo la página PC y la presentaron a los sujetos cinco veces tomando tiempos para cada grupo de 20 elementos. Es decir, se tomaban los tiempos al final de los elementos 20°, 40°, 60°, 80° y 100°. Después de un minuto de descanso se volvía a aplicar la prueba hasta contar con un total de 25 mediciones (Smith y Klein, 1953).

En el primer estudio con esta técnica aplicaron el test a 48 estudiantes de Harvard y Radcliffe obteniendo las puntuaciones individuales de acuerdo con la técnica serial. El sistema requería que el investigador calculase la dispersión de las cinco medidas de cada presentación de la página, así como la variabilidad general atribuida por regresión a cada puntuación. La variabilidad residual era, por tanto, la diferencia entre esas dos puntuaciones. Los autores usaron la raíz cuadrada de la variabilidad residual para identificar la variable lineal V, mientras que la variabilidad debida a la regresión fue llamada R.

A partir de estas dos medidas identificaron tres tipos de sujetos: Los estabilizadores (S) se caracterizaban por obtener resultados constantes de forma que la variabilidad era cero o próxima a cero; en los acumulativos (C) toda la variabilidad podía atribuirse a R mientras que nada o casi nada era debido a V; los disociados (D), por el contrario, tenían puntuaciones que oscilan arriba y abajo de forma que la mayor parte de la variabilidad no era lineal, por lo que podían tener alta V y baja R (Smith y Klein, 1953; Smith y Nyman, 1962).

Para clasificar a los sujetos se calcularon la V y la R de cada presentación de forma que cada sujeto obtenía en total cinco puntuaciones V y cinco R. Las medianas de V y R se calcularon en todos los sujetos y todos los ensayos. Si un intento de un sujeto tenía un valor R inferior a la mediana y un valor de V superior a la mediana, era clasificado como S; si R era alto y V bajo, era clasificado como C; si ambos valores eran altos, era clasificado como CD, es decir, una combinación de C y D.

Cada sujeto fue clasificado cinco veces y su adscripción definitiva a un grupo se hizo en función de su tendencia dominante. Por ejemplo si había resultado C, C, C, D, S era considerado C; si había sido S, S, D, C, S era considerado S, y si había resultado C, D, CD, CD, y C se consideraría C, ya que CD contaba como C y como D. Se descartó al 40 % de los sujetos cuyo patrón típico no coincidía con su patrón en las dos primeras páginas. De los 48 sujetos iniciales 9 fueron clasificados como C, 12 como D y 8 como S.

Smith y Klein (1953) desarrollaron predicciones sobre estos tipos a partir de sus resultados característicos en el Stroop, considerando que algunos de los aspectos del Stroop eran más importantes a estos efectos. En primer lugar, la combinación de colores y palabras varía de forma impredecible por lo que es poco lo que el sujeto puede aprender del elemento anterior. Es más, añadieron, el mejor resultado se obtiene olvidando al máximo el elemento anterior y no dejando que interfiera en los siguientes.

Los acumulativos trabajan como si la tarea fuese cada vez más difícil: "Cada nueva exigencia de aislar la respuesta desarrolla un proceso como de bola de nieve de dificultad cada vez mayor". Dedujeron que en tareas que requieren un modelo previo ese modelo permanece en elementos posteriores y perturba el resultado si el primer modelo ya no es aplicable. En segundo lugar, cuando se produce el cambio ocurre de forma exagerada como si el sujeto quisiese recuperar el tiempo perdido. Por último, si hay claramente procedimientos alternativos, el tipo C usará los modelos que han sido más reforzados en el pasado.

Los disociados pueden adolecer, según los autores, de falta de la atención requerida para ejecutar la tarea del Stroop, lo que causará una considerable interferencia y lentitud en la terminación de la página. Los D son capaces de cambiar de modelo, pero sólo si se les da un respiro suficiente para hacerlo. Su capacidad para hacer un cambio con esfuerzo es de corta duración, por lo que pueden hacerlo bien un rato y luego volver a cometer faltas.

Los tipos E (estabilizadores) encuentran el mejor modo de responder y lo mantienen, pues son capaces de mantener la atención y de aislar cada elemento de forma constante durante muchos ensayos. Estos sujetos son capaces de cambiar de modelo cuando lo necesitan sin pérdida de tiempo. Como consecuencia, Smith y Klein (1953) indicaron que eran capaces de descartar los modelos inadecuados, sobre todo cuando así lo indicaban las instrucciones, y de elegir entre las diversas opciones, no por la experiencia anterior, sino en función de lo que es más adecuado a la situación, siendo también capaces de volver a un modelo anterior si las condiciones lo aconsejan.

Con el fin de contrastar estas teorías aplicaron varios tests, el primero de los cuales consistía en estimar el tamaño de unos cuadrados que se presentaban a los sujetos. Los cuadrados se presentaban en series siendo el tamaño medio de cada serie ligeramente más grande que el de la precedente. Se comprobó que los sujetos veían un cambio exagerado entre la primera y la segunda series y después se estabilizaban en valores en torno a los tres cuartos del crecimiento real de tamaño. En los sujetos C la estimación inicial es bastante ajustada y se hace cada vez más precisa al aumentar el tamaño de los cuadrados. Hacia la serie quinta o sexta toman conciencia de su error y reaccionan con exceso pasando a sobreestimar el tamaño de los cuadrados. Se trata de las reacciones adecuadas y esperadas en cada grupo. También como era de esperar, los disociados tienden a hacer estimaciones que oscilan hacia arriba y hacia abajo.

Otra prueba usada consistía en caras camufladas, pero resultó demasiado fácil para los estudiantes, como para considerarse válida. Los sujetos D obtuvieron los mejores resultados pero también los peores. Los autores supusieron que ello era debido a que los D son potencialmente más autistas que los C o los S y se basan menos en datos objetivos.

Otros tests aplicados por los autores confirmaron sus teorías. En general la conducta de cada grupo en las tareas de tipo cognitivo era similar a su comportamiento en el Stroop, por lo que los autores concluyeron que "los estilos inferidos a partir de modelos seriales de conducta cognitiva permiten hacer predicciones sobre el rendimiento a obtener en diversas tareas" (Smith y Klein, 1953). Sin embargo, es preciso tomar estos resultados con prudencia porque las muestras eran reducidas y por el sesgo debido al alto número de sujetos descartados.

En estudios posteriores se aplicó un sistema de corrección diferente (Smith y Nyman, 1962) y se añadieron algunos tipos secundarios a los tipos primarios descritos en el primer estudio (Nyman y Smith, 1960). Estos tipos secundarios derivaron de un análisis más detallado de los cinco valores de V y R obtenidos en cada sujeto. Estas puntuaciones se usaron para calcular la variabilidad de la regresión en cada grupo de valores. En este sistema se llama  $R_r$  a la regresión de los cinco valores de R y  $R_v$  a la regresión de los cinco valores de V. Las variabilidades residuales se identificaron de forma análoga,  $V_r$  y  $V_v$ . Un sujeto con  $R_r$  y  $R_v$  bajas es calificado como  $Srv$ ; si  $R_r$  es alto, el sujeto se considera como  $Cr$ ; si  $V_v$  y  $R_v$  son altos el sujeto es  $CD_v$ , etc. En Smith y Nyman (1962) puede verse la lista completa de combinaciones.

Se hicieron varios estudios con este nuevo sistema de corrección (Smith y Nyman, 1959; Nyman y Smith, 1960; Nyman y Smith, 1959; Nillson y Smith, 1962; Nyman y Smith,

1961). Smith y Nyman (1959) examinaron a 78 sujetos de una clínica psiquiátrica sueca comparando los resultados con los síntomas, la conducta, la edad, la inteligencia y el nivel educacional de los mismos. Con los "baremos" de este grupo estudiaron a 51 pacientes ortopédicos, 48 psicóticos y 40 chicas delincuentes. Nyman y Smith (1959) hicieron un estudio más minucioso con 37 chicas delincuentes. Nyman y Smith (1960) hicieron una investigación con 48 pacientes psiquiátricos mientras que Nyman y Smith (1961) estudiaron el problema de las aberraciones neuróticas en 26 pacientes. Nillson y Smith (1962) estudiaron a un grupo de 114 pacientes.

De estos estudios se dedujo que los S primarios eran más frecuentes entre los sujetos normales que entre los pacientes psiquiátricos (Smith y Nyman, 1959). También aparecían con más frecuencia entre los verdaderos psicópatas que entre los delincuentes, lo que llevó a pensar que los delincuentes no son, como a veces se les ha llamado, psicópatas (Nyman y Smith, 1960). También comprobaron (Nyman y Smith, 1960) que el modelo S era más frecuente entre sujetos que habían mejorado con la terapia, lo que indicó que el Stroop puede servir como indicativo del éxito de una terapia.

Los acumulativos primarios (C) aparecían con mayor frecuencia en los sujetos caracterizados por su ansiedad manifiesta, así como en los que presentaban lo que los autores llamaron "defensas aversivas" (Smith y Nyman, 1959). Eran, en cambio, poco frecuentes entre los pacientes depresivos (Smith y Nyman, 1959). También encontraron con mayor frecuencia este tipo en grupos caracterizados por fobias u obsesiones que también tenían un nivel alto de ansiedad manifiesta (Nyman y Smith, 1961). Esto no es de extrañar pues, como habíamos visto (Smith y Klein, 1953), el rendimiento de los C decrece con el tiempo lo que puede deberse a un aumento de la ansiedad.

Los disociados primarios (D) aparecieron con mayor frecuencia en sujetos evasivos e infantiles (Smith y Nyman, 1959), lo que tampoco es de extrañar dada la descripción de sus resultados en el Stroop (Smith y Klein, 1953). Smith y Nyman (1959) comprobaron que los D eran incluso más similares cuando se eliminaba del grupo a los sujetos de elevada ansiedad y también (Nyman y Smith, 1961) que los sujetos D no suelen ser obsesivos o compulsivos.

Los CD primarios, no descritos por Smith y Klein (1953), resultaron relacionados con síntomas de retraso, rigidez, depresión, ansiedad y conflictos graves (Smith y Nyman, 1959). Aparecieron con mayor frecuencia en grupos de psicóticos y más raramente en sujetos normales (Smith y Nyman, 1959). También se vio que este tipo aparecía con mayor frecuencia en casos en que había fracasado la terapia (Nyman y Smith, 1960).

Smith y Nyman (1962) estudiaron también las características secundarias. Un factor de regresión estable (Sr) indica que el rendimiento es igual en todas las páginas y aparece con mayor frecuencia en sujetos normales (Smith y Nillson, 1959). También es característica de los sujetos normales una variación residual estable en cada página (Smith y Nyman, 1959).

Los Cr (acumulativos con crecimiento constante de los valores de R) y CDr (valores de R fluctuantes) se encontraron en grupos de pacientes no depresivos, rígidos y orgánicos (Smith y Nyman, 1959). Un valor V secundaria alto se encontró más en psicóticos y menos en neuróticos, siendo raro en sujetos normales (Smith y Nyman, 1959). Valores altos de V secundaria y bajos de C aparecieron en sujetos con retraso motor en contraposición a los sujetos con V secundaria baja y C alto (Nillson y Smith, 1962). Los

sujetos caracterizados por valores oscilantes de V secundaria y R tendían a ser más tensos que los sujetos con modelos menos fluctuantes (Nillson y Smith, 1962).

Cuando el valor alto de V secundaria y R se combinaba con un tipo primario de C o CD era indicativo de que los sujetos responderían favorablemente a la terapia (Nyman y Smith, 1960). Entre los psicópatas aparecía con frecuencia un valor creciente de V secundaria pero sin CD (Nyman y Smith, 1959).

Los autores buscaron otras puntuaciones adicionales (Smith y Nyman, 1962). En primer lugar, calcularon la diferencia entre los 20 primeros ensayos en la primera presentación de la página y los 20 primeros ensayos en la segunda presentación, dividido por la ganancia de tiempo entre los 20 primeros ensayos en la segunda presentación y el final del test (Xo). Una segunda medida fue la diferencia entre los primeros y los segundos veinte ensayos en la primera presentación de la página, dividido por la mejora de puntuación entre los segundos 20 elementos de la primera presentación de la página y el final del test (Yo). Smith y Nyman (1959) y Nyman y Smith (1961) encontraron valores más altos de Xo e Yo en sujetos obsesivos compulsivos. También encontraron diferencias significativas en estas medidas por razón del sexo (Smith y Nyman, 1962).

Smith y Borg (1964) plantearon el problema del retest usando la técnica de corrección serial. Al contrario de los sistemas normales de puntuación del Stroop que no cambian los resultados en el retest (Jensen, 1965), Smith y Borg (1962) comprobaron que la repetición del test con la técnica serial no produce los mismos resultados. Esto parece deberse en parte a la complejidad de los métodos de corrección usados y hace que el lector deba ser prudente al interpretar los resultados de estas experiencias. Smith y Borg (1962) no consiguieron desarrollar una versión acromática del Stroop que produjese resultados similares.

Las técnicas empleadas por Smith y sus colaboradores presentan varios problemas. Así, Jensen y Rohwer (1966) criticaron la falta de fiabilidad del sistema de corrección serial así como la falta de validación cruzada. Además, la mayor parte de estos estudios se realizaron en Suecia sin que se conozcan otros estudios que puedan validarlos. Sjoberg (1969) criticó el hecho de que hubiesen sido eliminados muchos sujetos al hacer el estudio, 40% en el estudio inicial (Smith y Klein, 1953) y 20% en los posteriores (Smith y Nyman, 1962). Sjoberg (1969) también criticó el sistema de corrección por razones estadísticas: el sistema de Smith no tiene en cuenta la varianza no lineal, exige simplificaciones y no minimiza las condiciones de error.

## 7. RESULTADOS EN MUESTRAS ESPAÑOLAS.

### 7.1. MUESTRA

La versión normalizada del test Stroop se ha aplicada a un total de 254 sujetos con edades comprendidas entre los 7 y los 80 años de edad.

La distribución de sujetos en los grupos de edad considerados en el Anexo B ha sido la siguiente:

Niños (7 a 15 años):33

Adultos jóvenes (16-44 años): 160

Adultos de 45 a 64 años:49

Adultos mayores (65-80 años): 12

La distribución por sexos es la siguiente:

Varones: 119

Mujeres: 135

En todos los casos se ha aplicado el Stroop de forma individual.

No se obtuvieron los datos referentes al nivel cultural y socioeconómico de los sujetos pero una gran parte de los sujetos adultos examinados (en torno al 80%) tenían estudios superiores o nivel similar.

### 7.2. RESULTADOS

	P		C		PC		INT.	
	Med.	Dt.	Med.	Dt.	Med.	Dt.	Med.	DA.
<b>Población general *</b>	118	18	77	14	49	10	2,70	8,53
<b>Varones *</b>	118	18	76	14	49	11	3,63	8,54
<b>Mujeres *</b>	118	18	78	13	49	09	1,88	8,13
<b>Niños de 7 a 15 años</b>	79	19	53	12	31	08	1,89	5,75
<b>Adultos jóvenes</b>	119	20	79	14	50	11	2,71	9,05
<b>Adultos de 45 a 64 años</b>	113	14	71	13	43	09	1,73	7,25
<b>Adultos mayores</b>	96	20	54	13	34	09	8,67	7,04

\* Los datos de la población general y los referidos a varones y mujeres incluyen los resultados de los sujetos más jóvenes y mayores corregidos por la edad como se indica en el anexo B.

La puntuación de interferencia se ha calculado como se indica en el anexo A, tomando siempre las puntuaciones corregidas por la edad.

A la vista de los resultados presentados podemos hacer los siguientes comentarios:

- Los resultados de la muestra española son superiores a los obtenidos en poblaciones americanas en las cuatro puntuaciones del Stroop.



- La media de las puntuaciones de interferencia que teóricamente debe ser cero resulta ser de 2,7 en la muestra española lo que, tal vez, puede deberse a la composición de la propia muestra.
- Las diferencias entre varones y mujeres son prácticamente inexistentes en P, C y PC. Por el contrario, los varones resultan ser más resistentes a la interferencia (3,63) que las mujeres (1,88).
- Los datos de corrección por edad en el caso de los niños podrán ser objeto de ulteriores consideraciones cuando se disponga de muestras más numerosas.
- En el caso de los adultos de 45 a 64 años, la corrección por edad indicada en el anexo B resulta correcta, pues si sumamos las cantidades 8, 4 y 5 a las medias de las puntuaciones P, C y PC obtenidas en el grupo se acercan mucho a las medias de la población general:

$$P : 113 + 8 = 121 \text{ (Media general= 118)}$$

$$C : 71 + 4 = 75 \quad \text{(Media general= 77)}$$

$$PC : 43 + 5 = 48 \quad \text{(Media general= 49)}$$

- Por el contrario, en los adultos mayo-res la corrección por edad indicada en el anexo B resulta insuficiente para las puntuaciones P y C, si bien también en este caso sería deseable contar con muestras más amplias para llegar a conclusiones definitivas:

$$P : 96 + 14 = 110 \text{ (Media general= 118)}$$

$$C : 54 + 11 = 65 \quad \text{(Media general= 77)}$$

$$PC : 34 + 15 = 49 \quad \text{(Media general= 49)}$$

### 7.3. BAREMOS

PUNTUACIONES DIRECTAS				Puntuación típica T
P	C	PC	INT.	
179	121	83	29,9	80
175	118	81	28,1	78
171	115	79	26,2	76
167	113	76	24,4	74
163	110	74	22,6	72
159	107	72	20,8	70
155	104	70	19,0	68
151	101	68	17,2	66
147	99	65	15,4	64

	143	96	63	13,6	62
	139	93	61	11,8	60
	135	90	59	10,0	58
	131	87	57	8,1	56
	127	85	54	6,3	54
	123	82	52	4,5	52
	119	79	50	2,7	50
	115	76	48	0,9	48
	111	73	46	-0,9	46
	107	71	43	-2,7	44
	103	68	41	-4,5	42
	99	65	39	-6,3	40
	95	62	37	-8,2	38
	91	59	35	-10,0	36
	87	57	32	-11,8	34
	83	54	30	-13,6	32
	79	51	28	-15,4	30
	75	48	26	-17,2	28
	71	45	24	-19,0	26
	67	43	21	-20,8	24
	63	40	19	-22,6	22
	59	37	17	-24,4	20
Media	119.00	79.00	50.00	2,71	50.00
D.t.	20.00	14.00	11.00	9,05	10.00

## 8. REFERENCIAS

Broverman, D.M., Clarkson, F.E., Klaiber, E.L., Vogel, W.: Learning disabilities: Multidisciplinary approaches to identification, diagnosis and remedial education. New York: MacMillan, 1972.

Callaway, E. and Stone, G.: Drugs and Behavior. New York: Willey, 1960.

Golden, C.J.: Diagnosis and rehabilitation in clinical neuropsychology. Springfield, Illinois: Charles C. Thomas, 1978.

Golden, C.J.: Stroop Color and Word Test. A manual for clinical and experimental uses. Wood Dale, Illinois: Stoelting Co., 1978.

Lehman, H.E. and Knight, D.A.: Systems Extrapramidal et Neuroleptics. Montreal: Editions Psychiatriques, 1961.

Thurstone, L.L.: A factorial Study of Perception. Chicago: University of Chicago Press, 1944.

**Nota:** El manual original en inglés relaciona un total de 113 artículos de publicaciones profesionales referentes al Stroop y a las investigaciones realizadas con este instrumento. El lector puede remitirse a dicho manual si le interesa conocer con mayor detalle algunas de las informaciones incluidas en este Manual.

## ANEXO A

### CÁLCULO DE LA PUNTUACIÓN DE INTERFERENCIA

Para determinar la puntuación de interferencia pura se debe calcular la diferencia entre la PC y la PC' (PC estimada). Cuanto mayor es la puntuación resultante, menos susceptible a interferencia es el sujeto, por lo que la variable podría definirse como "resistencia a la interferencia".

La fórmula para calcular la PC estimada (PC') es la siguiente:

$$PC' = \frac{C \times P}{C + P}$$

y la interferencia es igual a:

$$PC - PC' = PC - \frac{C \times P}{C + P}$$

La PC estimada (PC') debe obtenerse con esta fórmula en todos los casos.

### ORIGEN DE LA FÓRMULA

La fórmula deriva de suponer que la forma más sencilla de realizar la página PC es leer primero la palabra y después nombrar el color. Así, el tiempo necesario para resolver un elemento en la tercera página sería la suma del tiempo de leer una palabra más el tiempo de nombrar un color.

El tiempo de leer una palabra es 45 segundos dividido por la puntuación P, el número de respuestas dadas en la página 1. El tiempo de nombrar un color es, a su vez, 45/C. Así, el tiempo para un elemento de PC sería:

$$\frac{45}{P} + \frac{45}{C}$$

o lo que es lo mismo:

$$\frac{45(P+C)}{P \times C}$$

El número de ítems realizados en 45 segundos será:

$$\frac{45}{\frac{45(P+C)}{P \times C}}$$

lo que, simplificado, da:

$$\frac{1}{\frac{(P+C)}{P \times C}}$$

que es igual a:

$$\frac{(P+C)}{P \times C}$$

---

## ANEXO B

### CÁLCULO DE LAS PUNTUACIONES T

#### 1. Adultos jóvenes

La puntuación T para los adultos jóvenes (entre 16 y 44 años) se obtiene a partir de las puntuaciones directas consultando la tabla I. La puntuación de interferencia se calcula como se ha indicado en el Anexo A (PC-PC').

Tabla I. Puntuaciones T para los datos del Stroop

Puntuaciones directas (edad corregida)				
Puntuación T	Palabra	Color	Color-Palabra	Interferencia
80	168	125	75	30
78	164	122	73	28
76	160	119	71	26
74	156	116	69	24
72	152	113	67	22
70	148	110	65	20
68	144	107	63	18
66	140	104	61	16
64	136	101	59	14
62	132	98	57	12
60	128	95	55	10
58	124	92	53	8
56	120	89	51	6
54	116	86	49	4
52	112	83	47	2
50	108	80	45	0
48	104	77	43	-2
46	100	74	41	-4
44	96	71	39	-6
42	92	68	37	-8
40	88	65	35	-10
38	84	62	33	-12
36	80	59	31	-14
34	76	56	29	-16
32	72	53	27	-18
30	68	50	25	-20
28	64	47	23	-22
26	60	44	21	-24
24	56	41	19	-26

22	52	38	17	-28
20	48	35	15	-30

---

## 2. Adultos de 45 a 64 años

Para los adultos con edades comprendidas entre 45 y 64 años la puntuación directa debe incrementarse como se indica a continuación para obtener la puntuación corregida:

P corregida por la edad =  $P + 8$

C corregida por la edad =  $C + 4$

PC corregida por la edad =  $PC + 5$

Una vez obtenidas las edades corregidas puede usarse la tabla I para determinar la puntuación T. La interferencia debe calcularse a partir de las puntuaciones ya corregidas por la edad.

## 3. Adultos mayores

Las correcciones a realizar por causa de la edad en el caso de adultos de 65 a 80 años son las siguientes:

P corregida por la edad =  $P + 14$

C corregida por la edad =  $C + 11$

PC corregida por la edad =  $PC + 15$

Como en el caso anterior, una vez obtenidas las edades corregidas puede usarse la tabla I para determinar la puntuación T y la interferencia debe calcularse a partir de las puntuaciones ya corregidas por la edad.

## 4. Niños

Las puntuaciones de los niños deben también corregirse por la edad antes de usar la tabla I y de calcular las puntuaciones de interferencia. Las correcciones a realizar figuran en la tabla II.

Aunque estos datos son todavía experimentales y deben tomarse con prudencia, se observa que las correcciones a realizar varían notablemente en función de la edad, por lo que hay que tener en cuenta la edad exacta del sujeto. Esto es particularmente cierto en el caso de niños retrasados o adelantados, para los cuales la corrección por la edad cronológica podría ser incorrecta. En los casos en que la corrección por edad sea incoherente con los resultados de otros tests o los resultados escolares puede ser útil aplicar las correcciones correspondientes al grupo de edades inferior o superior.

Tabla II. Correcciones por edad para niños

Datos Experimentales			
Edad	Palabra	Color	Color-Palabra
7	52	40	26
8	46	36	24
9	41	29	20
10	34	24	16
11	26	16	11
12	15	10	7
13	10	7	5
14	5	0	2
15	3	0	0
16	0	0	0





# STROOP

## Test de Colores y Palabras

Nombre: \_\_\_\_\_

Edad: \_\_\_\_\_ Sexo: \_\_\_\_\_ Fecha: \_\_\_\_\_

### PARA USO DEL PROFESIONAL

	PD	PT
P		
C		
PC		
$\frac{P \times C}{P + C} = PC'$		
$PC - PC' = \text{INTERF.}$		

**NO ABRA EL CUADERNILLO  
HASTA QUE SE LE INDIQUE**



ROJO	AZUL	VERDE	ROJO	AZUL
VERDE	VERDE	ROJO	AZUL	VERDE
AZUL	ROJO	AZUL	VERDE	ROJO
VERDE	AZUL	ROJO	ROJO	AZUL
ROJO	ROJO	VERDE	AZUL	VERDE
AZUL	VERDE	AZUL	VERDE	ROJO
ROJO	AZUL	VERDE	AZUL	VERDE
AZUL	VERDE	ROJO	VERDE	ROJO
VERDE	ROJO	AZUL	ROJO	AZUL
AZUL	VERDE	VERDE	AZUL	VERDE
VERDE	ROJO	AZUL	ROJO	ROJO
ROJO	AZUL	ROJO	VERDE	AZUL
VERDE	ROJO	AZUL	ROJO	VERDE
AZUL	AZUL	ROJO	VERDE	ROJO
ROJO	VERDE	VERDE	AZUL	AZUL
AZUL	AZUL	ROJO	VERDE	ROJO
ROJO	VERDE	AZUL	ROJO	VERDE
VERDE	ROJO	VERDE	AZUL	AZUL
ROJO	AZUL	ROJO	VERDE	ROJO
VERDE	ROJO	VERDE	AZUL	VERDE

[illegible]

ROJO	AZUL	VERDE	ROJO	AZUL
VERDE	VERDE	ROJO	AZUL	VERDE
AZUL	ROJO	AZUL	VERDE	ROJO
VERDE	AZUL	ROJO	ROJO	AZUL
ROJO	ROJO	VERDE	AZUL	VERDE
AZUL	VERDE	AZUL	VERDE	ROJO
ROJO	AZUL	VERDE	AZUL	VERDE
AZUL	VERDE	ROJO	VERDE	ROJO
VERDE	ROJO	AZUL	ROJO	AZUL
AZUL	VERDE	VERDE	AZUL	VERDE
VERDE	ROJO	AZUL	ROJO	ROJO
ROJO	AZUL	ROJO	VERDE	AZUL
VERDE	ROJO	AZUL	ROJO	VERDE
AZUL	AZUL	ROJO	VERDE	ROJO
ROJO	VERDE	VERDE	AZUL	AZUL
AZUL	AZUL	ROJO	VERDE	ROJO
ROJO	VERDE	AZUL	ROJO	VERDE
VERDE	ROJO	VERDE	AZUL	AZUL
ROJO	AZUL	ROJO	VERDE	ROJO
VERDE	ROJO	VERDE	AZUL	VERDE