



UNIVERSIDAD CÉSAR VALLEJO

**ESCUELA DE POSGRADO
PROGRAMA ACADÉMICO DE DOCTORADO EN PSICOLOGÍA**

**Propiedades Psicométricas del Test de Colores y Palabras (Stroop) en niños del
distrito de Víctor Larco**

**TESIS PARA OBTENER EL GRADO ACADÉMICO DE:
Doctora en Psicología**

AUTORA:

Mg. Lilia Marina Zegarra Pereda (ORCID: 0000-0001-7008-0657)

ASESORA:

Dra. Araujo Rojas, Dany Elizabeth (ORCID: 0000-0002-4685-3936)

LÍNEA DE INVESTIGACIÓN:

Psicometría

Trujillo – Perú

2020

DEDICATORIA

Esta tesis se la dedico con todo mi amor a mi madre, hermanos y sobrinos, ellos son el motor, la luz y el significado que le dan a mi vida y son los que me motivan a ser mejor persona cada día.

AGRADECIMIENTO

**Gracias al forjador de mi camino, a mi
Dios padre todo poderoso, gracias por
darme la madre, los hermanos y los
sobrinos que tengo y por permitirme disfrutar
de ellos siempre, sin ti mi Dios nada de esto
hubiera sido posible.**

Página del Jurado


DECLARATORIA DE AUTENTICIDAD

Yo, Lilia Marina Zegarra Pereda estudiante de la Escuela de Posgrado de la Universidad César Vallejo, sede Trujillo; declaro que la tesis titulada "Propiedades Psicométricas del Test de colores y palabras (Stroop) en niños del distrito de Victor Larco" presentada, en 63 folios para la obtención del grado académico de Doctorado en Psicología es de mi autoría.

Por lo tanto, declaro lo siguiente:

- He mencionado todas las fuentes empleadas en el presente trabajo de investigación identificando correctamente toda cita textual o de paráfrasis proveniente de otras fuentes, de acuerdo establecido por las normas de elaboración de trabajo académico.
- No he utilizado ninguna otra fuente distinta de aquellas expresadamente señaladas en este trabajo.
- Este trabajo de investigación no ha sido previamente presentado completa ni parcialmente para la obtención de otro grado académico o título profesional.
- Soy consciente de que mi trabajo puede ser revisado electrónicamente en búsqueda de plagio.
- De encontrar uso de material intelectual ajeno sin el debido reconocimiento de su fuente o autor, me someto a las sanciones que determinan el procedimiento disciplinario.

Trujillo, 04 de enero de 2020



Firma

Nombres y apellidos: Lilia Marina Zegarra Pereda.

DNI: 40922305

ÍNDICE

DEDICATORIA.....	ii
AGRADECIMIENTO.....	iii
PÁGINA DEL JURADO	IV
DECLARATORIA DE AUTENTICIDAD	V
ÍNDICE.....	vi
ÍNDICE DE TABLAS	vii
RESUMEN	viii
ABSTRACT	ix
I. INTRODUCCIÓN.....	1
II. MÉTODO.....	16
2.2. Variables.....	16
2.3. Población, muestra y muestreo	18
2.4. Técnicas e Instrumentos de recolección de datos.....	19
2.5. Procedimiento.	21
2.6. Método de análisis de datos.....	21
2.7. Aspectos éticos	22
III. RESULTADOS.....	24
IV. DISCUSIÓN.....	33
V. CONCLUSIONES.....	37
VI. RECOMENDACIONES.....	38
VII. PROPUESTA.....	39
VIII. REFERENCIAS.....	40
IX. ANEXOS.....	47

ÍNDICE DE TABLAS

Tabla 1	Operacionalización de variables.....	17
Tabla 2.	Estadísticos descriptivos y Correlación Interescalas del Test de Colores y Palabras.....	23
Tabla 3.	Análisis Factorial Exploratorio del Test de Colores y Palabras.....	24
Tabla 4.	Índice de consistencia interna mediante las puntuaciones factoriales.....	25
Tabla 5.	Índices de Estabilidad mediante correlación entre aplicaciones (interclases) e intraclase.....	26
Tabla 6.	Prueba de diferencias en el Test de Colores y Palabras por la variable sexo.....	27
Tabla 7.	Prueba de diferencias en el Test de Palabras y Colores según edad.....	28
Tabla 8.	Normas percentil y tipo T del Test de Colores y Palabras para niños y niñas de 7 y 8 años.....	29
Tabla 9.	Normas percentil y tipo T del Test de Colores y Palabras para niños y niñas de 9 y 10 años.....	30
Tabla 10.	Normas percentil y tipo T del Test de Colores y Palabras para niños y niñas de 11 a 12 años.....	31

RESUMEN

El proceso de evaluación de la atención en la infancia es necesario, frente a esto es indispensable contar con una herramienta adaptada a la localidad. Ante lo referido, la presente investigación, tuvo como objetivo principal determinar las propiedades psicométricas del Test de colores y palabras (Stroop) en niños del distrito de Víctor Larco, la muestra estuvo constituida por 575 niños de las instituciones educativas entre niños (319) y niñas (256). Los resultados, del análisis factorial exploratorio, evidencia la presencia de un solo factor, que explica el 70.84% de la varianza, con un KMO de .70 y un p-valor estadísticamente significativo ($p < .05$) en el Test de Bartlett, a la vez, las cargas factoriales oscilan de .625 a .891, con comunalidades de .391 a .793; la confiabilidad fue analizada a través del índice consistencia interna mediante las puntuaciones factoriales, empleándose el método de Mislevy y Bock (1990) presentando un alfa de .851, el método test-retest, muestra un coeficiente de correlación de Pearson (r) con valores entre .779 a .792 en los ítems, asimismo la correlación intraclase (CCI) cuenta con valores de .777 a .811. Las normas tipo percentil y puntuaciones tipificadas, están dadas por edades en tres grupos de 7 y 8 años, 9 y 10 años y de 11 a 12 años de edad, evidenciándose valores más altos en la media (M), mediana (Me) y moda (Mo) a más edad en los ítems de Palabras, Colores y Palabra-Color y de forma inversa con la escala de Interferencia.

Palabras clave: Validez, confiabilidad, atención, interferencia, test de colores Stroop, test-retest.

ABSTRACT

The process of evaluation of childhood care is important and necessary, in view of this, it is indispensable to have a tool adapted to the locality. Given the aforementioned, the present investigation, had as main objective to determine the psychometric properties of the Test of colors and words (Stroop) in children of the district of Victor Larco, the sample was constituted by 575 children of the educational institutions between men (319) and women (256). The results of the exploratory factor analysis show the presence of a single factor, which explains 70.84% of the variance, with a KMO of .70 and a statistically significant p-value ($p < .05$) in the Bartlett Test, at the same time, factor loads range from .625 to .891, with communalities from .391 to .793; Reliability was analyzed through the internal consistency index using factor scores, using the method of Mislevy and Bock (1990) presenting an alpha of .851, the test-retest method, shows a Pearson correlation coefficient (r) with values between .779 to .712 on the scales, the intraclass correlation (ICC) also has values from .777 to .811. The percentile-type norms and typified scores are given by ages in three groups of 7 and 8 years, 9 and 10 years and 11 to 12 years of age, showing higher values in the mean (M), median (Me) and fashion (Mo) at an older age in the Word, Color and Word-Color items and inversely with the Interference scale.

Keywords: Validity, reliability, attention, interference, Stroop color test, test-retest.

I. INTRODUCCIÓN.

La función ejecutiva es un constructo multifacético que se ha conceptualizado desde una variedad de contextos. En el nivel más global, la función ejecutiva está conformada por procesos psicológicos compuestos (Luria, 1966; Zelazo, Carter, Reznick y Frye, 1997). Las funciones ejecutivas incluyen la capacidad para planificar, dirigir y mantener la atención, la organización, el razonamiento abstracto, la resolución de problemas, el control motor, entre otras habilidades (Morgan y Lilienfeld, 2000).

En este sentido, la evaluación de la función ejecutiva, se puede encontrar en cualquier número de medidas en diversas literaturas por su naturaleza multifacética. Idealmente, una medida debería facilitar alguna indicación del acto ejecutivo inicial de inhibir la respuesta dominante (el tiempo de reacción, la tendencia a pesar de la retroalimentación externa, perseverar en una respuesta), así como otros procesos ejecutivos que se presentan, sin embargo, no hay indicadores inequívocos de función ejecutiva o disfunción (Morgan y Lilienfeld, 2000).

La evaluación neuropsicológica se basa en la idea, que las pruebas brindan una medida precisa y razonable sobre capacidades cognitivas de una persona. Sin embargo, otros autores sostienen que los resultados de las medidas de la función ejecutiva pueden verse confundidos por los efectos de la inteligencia (Chelune y Thompson, 1987; Riccio et al., 1994), la percepción y la memoria, así como el lenguaje internalizado (Denckla, 1996). Frente a esto, la complejidad del constructo y la naturaleza multifacética de las funciones ejecutivas dan como resultado serias dificultades conceptuales, al tratar de establecer la condición de los déficits y los rasgos del comportamiento (Morgan y Lilienfeld, 2000).

En tales circunstancias, las tasas de prevalencia de simulación en la aplicación de instrumentos neuropsicológicos han tenido un registro desde un 30 al 40% (Mittenberg, Patton, Canyock, & Condit, 2002) (Larrabee, 2010). En entornos militares, donde los contextos de ganancia secundaria (por ejemplo, relacionados con el estado de conexión del servicio y las calificaciones de

discapacidad) son particularmente comunes, las tasas estimadas de rendimiento inválido en pruebas neurocognitivas han estado entre el 40 y el 50% (Armistead-Jehle & Buican, 2012). La evaluación de la validez del rendimiento también es importante en la práctica clínica general, para garantizar que las confirmaciones de diagnóstico y la elaboración del tratamiento se basen en datos que sean razonablemente confiables y significativos, especialmente dado que el rendimiento puede ser inválido incluso cuando no hay incentivos externos (Schutte, Axelrod, & Montoya, 2015). Las guías de práctica profesional enfatizan la importancia de medir el sesgo de respuesta y la validez de desempeño durante la evaluación neuropsicológica (Bush et al., 2005; Heilbronner et al., 2009), y el uso de pruebas de validez de desempeño (PVT) dentro de la especialidad se ha convertido en una práctica clínica estándar (Sweet, Benson, Nelson, y Moberg, 2015).

Frente a esto, la evaluación neuropsicológica, es importante en cualquier contexto, para Posner y Rothbart (2014) resulta necesario, conocer el grado de atención o las dificultades que puedan presentar los estudiantes en relación a esta capacidad durante la etapa escolar, ya que, resulta fundamental y necesaria para el proceso de aprendizaje, sobre todo durante los primeros años de formación académica. Por lo que resulta oportuno medir esta variable neuropsicológica (atención) en escolares, ya que es aquí donde se observan diversas dificultades, según reportes de padres y maestros, los niños tienen muchos problemas para centrarse en su tarea, ignorar estímulos interferentes y dificultades en la organización, estas suelen presentarse desde primaria hasta secundaria y en la adultez sino se trata, generándose más problemas cuando los estudiantes pasan de aula en aula o cambian de maestros por curso, sumado a ello, deben cumplir con cada exigencia de dichos cursos, además de manejar materiales diversos, tareas y otras demandas propias de cada asignatura (Brown, 2016).

Reportes de investigaciones a nivel internacional, indican que las dificultades en las funciones ejecutivas pueden afectar a los niños regulares de diferentes formas, llámese regulares a aquellos niños sin ningún diagnóstico, es decir

pueden tener problemas en una, dos o en todas las áreas, como por ejemplo dificultades para cambiar su foco de atención de tareas, suprimir estímulos irrelevantes y seleccionar el estímulo correcto o importante, dificultades para priorizar tareas significativas, responder sin pensar, problemas en el seguimiento de instrucciones, entre otras señales de alerta de los niños en edad preescolar y escolar (Zelazo et al., 2013).

En el Perú, existe un alto índice en el diagnóstico en niños por ejemplo con déficit atencional, con una prevalencia, que oscila del 5 a 10% (lo cual es una elevada frecuencia las consultas a especialistas para atender estos casos) y en Latinoamérica, varía en el rango de 5 a 20%, (Velásquez y Ordóñez, 2015). Además, en la actualidad, es todo un desafío para el psicólogo clínico diagnosticar este trastorno, debido a que los indicadores conocidos son netamente clínicos, al poseer una alta comorbilidad con trastornos emocionales o conductuales, tales como retardo mental, trastorno específico del aprendizaje, trastornos de conducta, trastorno oposicionista desafiante, entre otros.

Como se puede apreciar la atención y la capacidad de ésta ante la interferencia está presente en la persona en especial en los niños-adolescentes, que están en pleno proceso de formación escolar, adquiriendo y fortaleciendo día a día nuevos aprendizajes y por otro lado están presentes aquellos con psicopatologías clínicas y neuropsicológicas, en quienes se busca conocer con que procesos cognitivos cuentan para maximizarlos y cuáles no, para tratar de recuperarlos, compensarlos o sustituirlos, por ende, se requiere de la probabilidad de contar con una herramienta que cuantifique el nivel del proceso atencional y el de la función ejecutiva.

En la actualidad, existen diversos instrumentos que miden la atención, entre ellos se encuentra el DiViSA-UAM y el TACI-UAM, ambos estandarizados, en niños, miden la atención en forma global, labilidad en la atención, la organización y el tipo de orden establecido, en una forma sencilla y atrayente para niños, sin embargo, para ser administrada se necesita de un programa informático y de una computadora, así como de familiaridad en el evaluado con el uso de computadoras para poder marcar correctamente usando el

mouse, limitando así su aplicación (Quiroga, Santacreu, Montoro, Martínez y Chun, 2011). Por otro lado, también está el d2, Test de Atención, de origen alemán y adaptado en España por Seisdedos (2012), con baremos percentilares de población general en grupos de 8 a 88 años de edad en varios grupos, la administración tiene una duración aproximada de 8 a 10 minutos. Asimismo, está el Toulouse-Piéron-Test Perceptivo y de atención, adaptada en España, tiene una duración de 10 minutos y se aplica a partir de los 17 años de edad, presenta baremos percentilares y eneatispos en adultos (Cordero et al., 2013). Existe también, Test del Trazo A-B, que busca medir la atención a nivel visual y el proceso en el cambio, de tarea, existen diversas adaptaciones, la más reciente y que presenta una adecuada validez y confiabilidad fue desarrollada por Puerta et al., (2018) sin embargo, está orientada para ser aplicada en universitarios, limitando así su aplicación.

Con todo lo expuesto anteriormente, resulta relevante trabajar con el instrumento de Colores y Palabras Stroop (Stroop, 1935) herramienta de evaluación útil y confiable (Lezak, Howieson y Loring, 2004). Además, se sabe, que han desarrollado varias versiones diferentes de pruebas Stroop (Comalli, Wapner y Werner, 1962; Trenerry, Crosson, DeBoe y Leber, 1989 en Arentsen, 2013) con variaciones en el color y el número de los elementos de prueba, el número de las subpruebas, y el procedimiento de administración. A pesar de estas variaciones, el paradigma básico de la prueba de Stroop ha seguido siendo el mismo: el rendimiento de un individuo en una tarea básica (por ejemplo, leer nombres de colores) y se compara con su desempeño en una tarea análoga en la que debe responderse una respuesta habitual inhibida para admitir uno poco usual (es decir, nombrar el color de la tinta en la que se imprimen las palabras de color con el nombre incongruente). El aumento en el tiempo que se tarda en realizar esta última tarea en comparación con la tarea básica se conoce como "el efecto de interferencia de Stroop" (Davidson, Zacks y Williams, 2003; Moering, Schinka, Mortimer, & Graves, 2004) y es considerada una medida general de la flexibilidad y el control cognitivo (Uttl & Graf, 1997) o el funcionamiento ejecutivo (Moering et al., 2004).

La prueba de Stroop se utiliza exclusivamente para evaluar la atención, la flexibilidad cognitiva y el efecto stroop (la inhibición de respuestas automáticas favoreciendo respuestas inusuales) (Stroop, 1935). Este instrumento ha sido validado en diferentes grupos de personas, entre ellas, pacientes con demencias, daño cerebral, psicopatología, drogadicción (Golden, 2007). Asimismo, la prueba ha sido útil para la identificación de alteraciones o trastornos psicológicos como: la esquizofrenia (Golden, 2007) y estrés postraumático (Bremmer et al., 2004, citado en Golden, 2007). Por otra parte, este instrumento tuvo un gran impacto para la detección de síndromes desatencionales y trastornos de aprendizaje (Lavado, Gómez, Sawada, Chomali y Alvarez, 2003).

Entre los estudios instrumentales realizados con el Stroop se encuentra a Arentsen et al., (2013), quienes investigaron la efectividad de la versión (Comalli de Stroop Test 1962) para detectar la invalidez en el rendimiento, en una muestra de pacientes con diagnóstico mixto. Esta versión de la tarea consta de tres ensayos (lectura de palabras, nombres de colores e interferencias), cada uno con 100 elementos; El rendimiento de la prueba se mide por tiempo, en segundos, para completar los 100 elementos. Los autores clasificaron a los examinados no creíbles como aquellos que cumplían con los criterios para una posible disfunción neurocognitiva no expresada (MND); aquellos clasificados como examinados creíbles fallaron menos de dos PVT, y no estaban en litigio ni solicitando discapacidad. Los resultados mostraron que los examinados no creíbles se desempeñaron significativamente peor que los examinados creíbles en todos los ensayos de la Prueba de Stroop. Aunque no se evaluó la precisión de la clasificación general de cada ensayo de Stroop, se informaron las estadísticas de sensibilidad y especificidad para un rango de puntajes brutos de corte. Con una especificidad del 90%, las pruebas de lectura de palabras (Word) y denominación de colores (Color) tuvieron una sensibilidad moderada (49–53%), mientras que la prueba de interferencia (Color – Word) tuvo una sensibilidad relativamente baja (29%). Los autores también examinaron el efecto de Stroop invertido (es decir, un mejor rendimiento de Word-Color que el que se predeciría con los rendimientos de Word y Color) para diferenciar a

los pacientes creíbles y los no creíbles, y encontraron que los ensayos individuales de Stroop fueron mucho más efectivos para detectar un rendimiento no válido.

Por su parte, Gutiérrez, Ramos y Vila (2018) en un estudio determinaron la validez y confiabilidad del instrumento Stroop. La muestra fue de 128 estudiantes, con edades comprendidas de 10 a 15 años (de nivel primario y secundario), pertenecientes a tres centros educativos y como característica sociodemográfica, del mismo nivel socioeconómico. En relación a la validez, el AFE evidencia la presencia de un solo factor, que explica el 64% de la varianza, así como una correlación estadísticamente significativa entre sus ítems ($p < .00$). La confiabilidad, se obtuvo mediante el alfa de Cronbach, siendo de .96 para la escala general; en relación a la validez, la prueba presenta una correlación significativa ($p < .00$) entre las respuestas obtenidas.

Asimismo, en Colombia, Rodríguez, Pulido y Pineda (2016) determinaron las propiedades psicométricas del Stroop, test de colores y palabras de Golden (2007) en una muestra colombiana no clínica. Los participantes fueron 1332 personas, 561 hombres y 711 mujeres entre 7 a 80 años de edad, divididos en tres grupos (niños, adultos y adultos mayores). Los resultados, en relación al análisis factorial exploratorio, evidencian la presencia de tres dimensiones, que explican el 86.84% de la varianza, el KMO evidencia valores que oscilan de .68 a .91. presenta también una validez convergente, realizado con el Trail Making Test (TMT), encontrándose una correlación negativa y significativa con valores entre -0.35 a -0.41. La confiabilidad fue obtenida a través del test-retest coeficiente de correlación intraclass CCI, presentando un alfa de .91 (en la primera lámina), de .87 en la segunda y en la tercera de .78. A la vez, se encontraron diferencias de acuerdo a la edad, estrato social y nivel de escolaridad.

Otro estudio realizado por Bezdicek, Lukavsky, Stepankova, Nikolai, Axelrod, Michalec, Růžička y Kopecek (2015) fue proporcionar datos normativos para adultos checos mayores y muy mayores en el Test de Stroop de Praga (PST) y probar su validez discriminativa en individuos con deterioro cognitivo leve (PD-MCI) de la enfermedad de Parkinson. La construcción del PST fue

modelada después de la prueba de Victoria Stroop. Se examinaron a 539 participantes de 60 a 96 años que cumplieran con estrictos criterios de inclusión. Después, se comparó las puntuaciones de PST para un grupo de 45 pacientes con PD-MCI con una muestra de adultos sanos (HAS) de 45 individuos de la misma edad y educación. Entre los resultados, en la muestra no clínica, se observaron influencias sólidas relacionadas con la edad y la educación en todas las puntuaciones de PST. No se observó ningún efecto de género. Para los casos clínicos, la condición de interferencia (PST – C) fue capaz de discriminar entre PD-MCI y HAS (todas las puntuaciones $p < .01$). El área bajo la curva (AUC) fue del 77% cuando se usó un corte de detección de ≤ 27 s, mostrando una sensibilidad del 82% y una especificidad del 53%. Un corte de diagnóstico más conservador de ≤ 33 s mostró una sensibilidad del 60% y una especificidad del 80%. Se concluye que la PST parece ser una herramienta útil para el diagnóstico de PD-MCI, especialmente en entornos de investigación en el Nivel II (Litvan et al., 2012) y para la atención / subtitulación de funciones ejecutivas y atención PD-MCI.

Por su parte, Maureira, Aravena, Gálvez y Flores (2014) realizaron un estudio que tuvo como principal objetivo conocer las propiedades psicométricas del test de interferencia de Stroop y de planificación de la Torre de Hanoi en 85 estudiantes de Educación Física de la USEK de Chile. Los valores alcanzados a partir de los resultados en las evidencias de confiabilidad mediante el test retest (a los 15 días) en ambos momentos, la confiabilidad fue de .884. El análisis de validez de constructo, se realizó a través de la prueba “t” para conocer, si existían diferencias significativas entre los cuartiles por edad, apreciándose un valor de 1,19 (d.e.=11,82) en el primer cuartil de edades inferiores y de -3,61 (d.e.=10,37) en las edades superiores, por lo que no existen diferencias significativas ($t=1,771$; $gl=72$; $p=0,081$) entre ambos resultados, a la vez, se aprecia un solo factor y una correlación altamente significativa ($p < .00$) entre los elementos que componen la prueba aspectos.

A la vez, Ramos (2009) realizó en Salamanca, la adaptación psicométrica del cuestionario Stroop. La muestra fue dividida en dos grupos, la primera conformada por 290 personas, con edades comprendidas de 25 a 80 años y

la segunda fue de 32 pacientes diagnosticados con Alzheimer. En relación a los resultados, el análisis factorial, evidencia una varianza explicada de 96.16, existe además una correlación estadísticamente significativa ($p < .00$) entre las palabras leídas y los colores nombrados; la confiabilidad fue obtenida con el alfa de Cronbach, siendo de .66 para la escala general, el factor 1 presenta un alfa de .87 y el dos de .52. En relación a los baremos, se hallaron diferencias en edad, el grado de instrucción y el nivel socioeconómico.

Por último, Conca e Ibarra (2004), realizaron una investigación, cuyo objetivo fue, estandarizar el Test de Colores y Palabras Stroop en niños y niñas entre los 8 a 12 años, escolarizados en escuelas urbanas de la Región Metropolitana, y pertenecientes a las distintas clases, la muestra estuvo conformada por 615 participantes. La confiabilidad, fue obtenida a través del método Entre calificadores, que se emplea para determinar el grado de comprensión de las instrucciones o ítems, comparados a partir de los resultados obtenidos en dos momentos; la validez fue obtenida a partir del análisis de la "t de student", encontrando diferencias estadísticamente significativas ($p < .00$) entre los grupos conformados por edades, datos que los investigadores aseguran guardan relación con los antecedentes empleados en su estudio. En los resultados se encontraron diferencias significativas en la edad y en el estrato socioeconómico, pero no en el género. Siendo estos procedimientos comparables con los de otros estudios psicométricos. Además, en la muestra de estandarización se analizaron las características psicométricas del test.

En relación a los aspectos teóricos, quién mejor explica el "efecto stroop" es la atención, la cual es definida como aquella capacidad que permite a la persona seleccionar y mantener un foco de atención un uno o varios estímulos presentes en el entorno (Puente, 2003) (López y García, 1999); se le considera también, como aquella habilidad cognitiva, que permite centrar la energía en un objeto, cualidad o atributo, contribuye al aprendizaje y permite la retención de información, a la vez, tiene una mayor retención de acuerdo al grado de maduración o edad de la persona (Boujon y Qualreau, 2004). A nivel psicológico, se asume como aquel proceso cognitivo, donde se le brinda una

mayor importancia a un elemento, considera como relevante para la persona (Mesulam, 2000).

Se considera, que la atención presenta tres elementos (De Vega, 2006): el primero de ellos, es la alerta, permite a la persona tener una mayor recepción de los estímulos, presenta dos tipos, la fásica (es el cambio repentino de atención) y el tónico (es aquella capacidad que permite a la persona poder tener continuar ejecutando cualquier actividad por un tiempo prolongado). El segundo es la atención sostenida, es el proceso, que involucra a la motivación y contribuye a que se mantenga la atención en un estímulo a pesar del cansancio o la fatiga mental. Y la atención selectiva, es la que permite a la persona seleccionar e integrar estímulos diversos.

En cuanto al paradigma Stroop hace referencia al control atencional, la cual se presenta en situaciones en las que se necesita de manera voluntaria controlar la acción poniéndose de manifiesto el funcionamiento de un sistema que centraliza la atención y dirige el esfuerzo en la ejecución de la acción o en el pensamiento. Siendo el control de las acciones o pensamientos y la selección de información, aspectos relacionado al sistema atencional (Posner y Rothbart, 1991).

Posner a su vez, expone tres sistemas neuropsicológicos, específicamente halladas en áreas del cerebro, las cuales se halla el proceso de atención selectiva y el control cognoscitivo, cuyo rol es primario para la elaboración de lo que vivenciamos de forma consciente (Posner y Raichle, 1994). De entre las tres, *la red atencional anterior (RAA)*, ubicada exactamente en el giro cingular anterior, es la que está más estrechamente implicada en el control ejecutivo (Posner y DiGirolamo, 1998). Así, cuando se necesita dominar determinadas respuestas habituales relacionadas con un estímulo, o en las que las respuestas presentan nuevas acciones secuenciadas de las que no se han aprendido del todo. Tal es el caso de las tareas tipo Stroop.

Existen dos principales teorías que fundamentan el motivo por la que este fenómeno (efecto stroop) se manifiesta:

1. La Teoría de la velocidad de procesamiento: Señala que la interferencia se da porque la lectura de palabras es más rápida que el denominar colores. Se han demostrado en investigaciones neuropsicológicas que el efecto Stroop es el resultado de interferir el proceso verbal. Activando al parecer los estímulos del Stroop es un proceso automático de respuesta verbal que dificulta con el nombramiento de los colores, anteriormente aprendido (McLeod, 1991).

2. La Teoría de la atención selectiva: Afirma que la interferencia se presenta porque el nombrar colores se necesita un mayor grado de atención que el de leer palabras. La interferencia se genera cuando se suprime el proceso automático que sería la lectura ante el nombramiento de colores, acción que requiere más control atencional por parte del individuo. Entendiéndose mejor, la palabra escrita en color genera una respuesta verbal automática que necesita de varias de las funciones neuropsicológicas que son requeridas para mencionar los colores. La lámina en donde se da la interferencia de Stroop mediría, exclusivamente, la capacidad del sujeto para separar los estímulos y ante dos situaciones conflictivas dar una respuesta correcta (McLeod, 1991).

La evidencia teórica del instrumento Stroop se ha realizado en población infantil (Golden, 2001). Comali, Wapner y Werner (1962, citado en Golden, 2001) sostuvieron que en la etapa infantil los niños de 7 años tienen el riesgo de desarrollar adecuadamente las habilidades de lectura e identificación de colores. Varios estudios han sugerido que los niños con TDAH demuestran deficiencias en la función ejecutiva (Kerns, McInerney y Wilde, 2001). Los niños y adultos con TDAH a menudo tienen dificultades para concentrarse, la atención y la inhibición de la respuesta. Además, se cree que la función cerebral está dañada hasta cierto punto en los niños que presentan TDAH. Las imágenes de resonancia magnética (IRM) revelaron que los niños con TDAH sus cerebros no muestran una normal asimetría frontal, encontrándose que el hemisferio derecho es más grande que el otro hemisferio. En cambio, los niños con TDAH mostraron un ancho frontal derecho más pequeño en comparación de los lóbulos frontales simétricos (Hynd, Semrud-Clikeman,

Lorys, Novey, y Eliopulos, 1990). La prueba de Stroop se ha utilizado para medir la inhibición cognitiva y la atención selectiva en individuos con TDAH.

La investigación ha indicado que los niños con discapacidades de aprendizaje (LD) muestran una función ejecutiva deficiente y un rendimiento más bajo en Stroop en comparación con los controles. Se obtuvieron tres estudios que incluyeron un grupo de niños con problemas de aprendizaje y controles. Se encontraron tamaños de efectos negativos en los tres estudios que indican que los niños con discapacidades de aprendizaje obtuvieron mejores resultados en todas las condiciones de la prueba de Stroop en comparación con los controles. Golden y Golden (2002) encontraron que al usar las tres escalas Stroop (Color, Word, Interference), podían discriminar a los participantes con discapacidades de aprendizaje de aquellos que no tenían discapacidades de aprendizaje con una precisión del 89%. Lazarus, Ludwig y Aberson (1984, citado por Dempster y Brainerd, 1995) informó que la medida de Color-Palabra del Stroop es una excelente medida de atención selectiva. Encontraron que los niños con LD obtuvieron puntajes significativamente más bajos que los niños sin LD en el puntaje Color Word de la prueba Stroop. Explicaron su hallazgo diciendo que “los niños que no eran LD a menudo dividían [agrupaban] los colores en la página de Color-Palabra de Stroop, mientras a menudo no lo hacían los niños con LD. Esto sugiere una deficiencia general en el procesamiento verbal y, específicamente, una velocidad de denominación lenta y falla en la fragmentación” (p. 85).

El aprendizaje está en relación directa con el bajo rendimiento escolar, lo que implica el no logro de los aprendizajes, es una vía final habitual de diversos trastornos, casusas y de mecanismos. Habitual es la presencia de variadas alteraciones, porque la disfunción cerebral en la niñez principalmente afecta a diversos procesos cognitivos (Shapiro, 2011).

Es por ello, para que exista un adecuado aprendizaje escolar, se requiere que los procesos cognitivos en el niño funcionen óptimamente, una disfunción en el sistema nervioso central genera alteraciones cognitivas, comportamentales y emocionales. Todos los procesos cognitivos son vitales, siendo la atención una de las bases y pilares del proceso de aprendizaje y es que sin atención

no habría memoria y sin memoria no habría gnosias, lenguaje, etcétera, en otras palabras, no habría medio u forma de codificar, asimilar y consolidar información y posteriormente evocarla.

Ante lo expresado no es un secreto que la principal preocupación de los padres con hijos en edad escolar o hijos adolescentes del nivel secundario sea el logro de aprendizajes, pero ante la realidad de que éstos (los aprendizajes) no son logrados o si los logran pero no se consiguen eficientemente, entonces se busca a menudo una explicación o se le atribuye a diversos factores, siendo uno de ellos o quizás las principales, la falta de atención, sensibilidad a la interferencia y un pobre control inhibitorio como principales causas por lo que los estudiantes no logran los aprendizajes esperados de sus menores hijos. Así mismo a diario podemos observar que padres llegan con sus hijos a las consultas con los profesionales en psicología, porque simplemente alguien les dijo eso o en ocasiones son los profesores lo que le atribuyen a la falta de atención y sensibilidad a la interferencia, como causas por la que sus hijos no aprenden eficientemente, pero realmente es ésta la causa, o son otros factores y es aquí donde esta Investigación toma fuerza he importancia ya que con la aplicación y validación de este instrumento se convertirá en una herramienta útil para detectar el nivel de Interferencia y a partir de ellos tomar las acciones correctivas idóneas y oportunas; lo que permitirá contribuir a la mejora o logro de los aprendizajes esperados de los niños y adolescentes objeto de estudio.

En relación a los aspectos psicométricos que se tomarán en este estudio se considerarán las evidencias psicométricas mediante los procedimientos de validez y confiabilidad. El concepto de validez a tenido una evolución en lo que es su conceptualización, Gregoory (2012) refiere que depende de la fidelidad con que mide lo que pretende medir. Por otro lado, Gregoory (2012) indicó que un test es válido cuando el constructo se correlaciona con algo. Ante ello, la evolución del concepto específico de la definición de la validez se refiere al grado en el que la evidencia empírica y la teoría sostienen la interpretación de las puntuaciones de los tests relacionada con un uso específico (APA, AERA & NCME, 1999). Este último concepto es el más

recomendado en el campo psicométrico para el desarrollo de estudios instrumentales.

Por otra parte, la confiabilidad ha sido un concepto muy discutido en el campo de la investigación; además, de la importancia que tiene para la precisión y consistencia de los instrumentos de medición. Asimismo, los estudios e investigaciones solicitan en su mayoría reportar los índices de confiabilidad de los instrumentos que se utilizan para la determinación del estudio. De esta manera, la confiabilidad se ha conceptualizado como el grado en que un instrumento genera resultados consistentes y coherentes (Hernández, Fernández y Baptista, 2014). Asimismo, el concepto de la confiabilidad que valora la consistencia y exactitud de la medida (Meneses et al, 2013).

El Test de Stroop es un instrumento neuropsicológico internacionalmente utilizado y más aún para fines de investigación. Se usa para evaluar la atención y el nivel de interferencia en personas normales, así como en el estudio y en la detección de disfunciones cerebrales que alteran la atención y la capacidad de control de interferencias e inhibición de respuestas automáticas, siendo estas últimas, capacidades de las funciones ejecutivas (Salgado – Pineda, Vendrell, Bargalló, Falcón, y Junqué, 2002, citado por Conca y Ibarra, 2004).

Finalmente se puede afirmar que en el Perú y la provincia de la Libertad no tiene un cimientto de pruebas psicológicas estandarizados para la evaluación y detección de estos problemas de aprendizaje que se suscitan en el campo educativo; asimismo, en el país existe carencia de test neuropsicológicos válidos y confiables, especialmente instrumentos que cumplan estas condiciones y que evalúen y midan la capacidad inhibitoria y la resistencia a la interferencia. En consecuencia, en la práctica clínica, tanto en adultos como infanto – juvenil, se observa escasez de normas peruanas para las distintas pruebas y baterías utilizadas, desmedrando la confiabilidad de las técnicas de apoyo a la evaluación neuropsicológica. Aporta a la confiabilidad de los diagnósticos contar con instrumentos y herramientas estandarizadas, ya que permite la creación de criterios de medición sistematizados, templando los parámetros diagnósticos. Por consiguiente, vendría bien adaptar un

instrumento así. Por otra parte, existe la necesidad de poder realizar investigaciones instrumentales para poder contar con instrumentos de medición contextualizados a la realidad peruana y así poder tener una certeza medición de esta variable. Asimismo, el estudio servirá como aporte empírico para la detección de la flexibilidad cognitiva el efecto Stroop en escolares de instituciones educativas públicas; ya que, mediante la validación del instrumento el profesional a cargo podrá contar con la herramienta útil y adaptada para la detección de estos indicadores. Por último, el estudio servirá como referente para la consecución de otras investigaciones relacionadas a la variable de estudio.

Por tal motivo el propósito de este estudio es determinar las propiedades psicométricas del inventario del test de colores y palabras (Stroop) en niños del distrito de Víctor Larco.

A la vez, se formuló la siguiente pregunta: ¿En qué medida la revisión de las propiedades psicométricas del test de colores y palabras (Stroop) en niños es significativo en el distrito de Víctor Larco?

La misma que se justifica:

Para el desempeño de aprendizaje, es hasta ahora de gran valor, el uso adecuado de la atención y su capacidad de control de interferencia del individuo en el ámbito educativo, ya que interviene de manera fundamental en el desarrollo cognitivo, así como en el desenvolvimiento comportamental del niño. Es por ello, que se evidencia la relevancia de adaptar instrumentos, a la realidad poblacional, que midan la capacidad atencional y su control de interferencia.

Por lo expresado anteriormente el presente estudio de investigación se justifica por ser de conveniencia, debido a la falta de test neuropsicológicos revisados en nuestro país y en nuestra localidad, servirá para contar con un test neuropsicológico que arrojen resultados más precisos y confiables.

Es de importancia metodológica, ya que permite tener instrumentos con validez y confiabilidad del test de colores y palabras (Stroop) para la población en la que se realizará la investigación; permitiendo a los profesionales de la

salud mental acercarnos a un diagnóstico contando con una herramienta psicométrica.

Otro aporte se realizaría en el ámbito de la investigación, ya que de una manera más rigurosa permitiría llevar a cabo investigaciones normadas, y obtener así en el área de la epidemiología y salud pública de la Salud Mental resultados más confiables.

Asimismo, si los resultados obtenidos de validez y confiabilidad son altos con puntajes entre el resultado del análisis de componentes principales de la rotación varimax, para la validez; y entre .77 y .88 para la confiabilidad, permitirá a que posteriores investigadores realicen estudios comparativos o programas de intervención.

Además, el estudio es de importancia práctica, porque ayudará y beneficiará a los profesionales en el campo de investigación que pretendan estudiar procesos cognitivos como la atención y funciones ejecutivas de esa población u otras y de esta manera poder contribuir a acercarnos a un diagnóstico más confiable, y a la labor psicoeducativa (labor preventiva – promocional). Asimismo, aportará empíricamente en hallazgos y resultados consistentes para un adecuado diagnóstico educativo. Asimismo, la investigación permitirá conocer datos estadísticos sobre cuál es el nivel de interferencia que poseen los niños del distrito de Víctor Larco con las que se trabajará.

Por tanto, se planteó el objetivo general:

Determinar las propiedades psicométricas del Test de colores y palabras (Stroop) en niños del distrito de Víctor Larco.

Además de los siguientes objetivos específicos:

- Determinar las evidencias de validez basadas en la estructura interna a través del análisis factorial exploratorio test de colores y palabras (Stroop) en niños del distrito de Víctor Larco.
- Determinar la confiabilidad por consistencia interna del test de colores y palabras (Stroop) en niños del distrito de Víctor Larco.

- Determinar la confiabilidad mediante el método de test retest del test de colores y palabras (Stroop) en niños del distrito de Víctor Larco.
- Elaborar las normas de tipo percentil del inventario del test de colores y palabras (Stroop) en niños del distrito de Víctor Larco.

II. MÉTODO.

2.1. Tipo y diseño de investigación.

El presente estudio corresponde al tipo tecnológico y de diseño instrumental, el cual según Montero y León (2007), pertenecen aquellas investigaciones dirigidos al desarrollo de pruebas y apartados, incluyendo el diseño o adaptación; así como, el estudio de las propiedades psicométricas de instrumentos de medición.

2.2. Variables

Control Atencional.

Definición Conceptual: Es la capacidad para clasificar diversos estímulos de su entorno y reaccionar selectivamente a esa información permitiendo al individuo anular por un determinado momento las fuentes de distracción (Golden, 2001).

Definición operacional: El control atencional será evaluada por el Test de Colores y Palabras de (Stroop) de Golden (2001).

Interferencia.

Definición Conceptual: Es el fenómeno resultante cuando una capacidad cognoscitiva relativamente controlada y voluntaria entra en conflicto con otro automático y no intencionado (Golden, 2001).

Definición operacional: La interferencia será evaluada por el Test de Colores y Palabras de (Stroop) de Golden (2001).

Validez:

Definición conceptual: Hace señalamiento a la capacidad de un instrumento de medición para cuantificar de forma significativa y adecuada el rasgo para cuya medición fue elaborado. De esta forma, un test de medida es válido en la manera en que las evidencias empíricas legitiman la interpretación de los resultados obtenidos por la prueba (Meneses, 2013).

Definición operacional: Se evaluará la validez de constructo mediante el análisis factorial exploratorio.

Confiabilidad:

Definición conceptual: Se refiere al grado en que un instrumento es consistente y coherente (Hernández, Fernández y Baptista, 2014).

Definición operacional: Se evaluará mediante el método test retest utilizando la consistencia interna.

2.2.1. Operacionalización de la variable.

Tabla 1.

Operacionalización de variables.

Variab les	Dimensiones	Indicadores	Escala de medición
Control Atencional	Interferencia	-	Intervalo
Validez de constructo	Exploratorio	Análisis Factorial	Cargas factoriales que superan el .30 Medida de Kaiser-Meyer-Olkin Test de Bartlett

			Varianza explicada
Confiabilidad	Consistencia Interna	Método de Mislevy & Bock	Bueno >.80
	Estabilidad	Test retest	
Normas	Puntuaciones T/ Percentil	Puntuaciones T/ Percentil	

2.3. Población, muestra y muestreo

Población.

La población estuvo constituida por 2 253 participantes, pertenecientes a 5 instituciones educativas del distrito de Víctor Larco; los cuales fueron I.E. Víctor Larco, I.E. José Antonio Encinas y I.E. Andrés Avelino Cáceres, I.E. Augusto Alva Ascurra y la I. E. Santa Edelmira.

Muestra.

De esta manera, la muestra estuvo constituida por 575 niños y niñas de las instituciones educativas entre niños (319) y niñas (256), para lo cual se empleó el método no probabilístico, debido a que los participantes de las instituciones educativas tuvieron las mismas posibilidades de poder pertenecer a la muestra (Hernández et al., 2014).

Muestreo.

El tipo de muestreo empleado fue por conveniencia, debido a la accesibilidad y a las facilidades, que los centros educativos brindaron durante la investigación (Hernández et al., 2014).

2.3.1. Criterios de selección:

Criterios de inclusión.

Asistir el día de la evaluación.

Niños que deseen participar de la evaluación voluntariamente.

Niños entre las edades de 7-12 años.

Criterios de exclusión.

Niños que abandonen la evaluación.

Niños que no logren comprender de forma correcta las indicaciones brindadas.

Niños que presenten déficit cognitivo.

Niños que presenten alteraciones perceptivo visuales (incluyendo el daltonismo), déficits auditivos o motores que impidieran o dificultaran ejecutar la prueba.

2.4. Técnicas e Instrumentos de recolección de datos.

2.4.1. Técnicas

Se llevó a cabo la encuesta que consiste en el método de investigación capaz de responder a problemas tanto en términos descriptivos como la de relación de variables, tras la recogida de información sistemática. (Cardona, Chiner & Lattur, 2006).

2.4.2. Instrumento.

Test de colores y palabras "Stroop"

Antecedentes de la Prueba

El test Stroop, surge a partir del aporte de la observación en relación, a las dificultades que presentaban las personas para identificar los nombres de colores. A partir de esto, los estudios de Cattell, evidenciaron que leer palabras no necesitaba de tanto tiempo como identificar el tono del color; en base a esto, Ligon consideraba que esta dificultad se debía a la poca práctica de la persona frente a este tipo de situaciones. Frente a esto, se crea la versión de Stroop, que contenía en un inicio, una hoja de trabajo, con cinco palabras escritas (rojo, verde, marrón, azul y purpura) en donde la palabra escrita era diferente a la tinta de color; otra hoja con rectángulos de

colores y la siguiente con las palabras de los colores, impresas en color negro.

Fue diseñada por Ridley Stroop, en el año 1935, con el objetivo de encontrar la mejor forma de medir “interferencia”. (Huguet, Galvaing, Monteil y Dumas, 1999; citados por Conca & Ibarra, 2004). Con su equipo de trabajo del George Peabody College, Stroop investigaba sobre la denominación de colores contra la lectura de palabras, generando la ocurrencia de crear un estímulo compuesto, en donde la palabra fuera distinta al color de la tinta en que impresa estaba la palabra (Ver protocolo), con el fin de apreciar la implicancia de las dimensiones del estímulo compuesto sobre la otra. Por lo tanto, se diseña el Test de Stroop, en el cual el individuo tiene la indicación de leer palabras en la primera lámina, luego en la segunda de denominar colores, y en la tercera de volver a denominar colores; en donde en ésta última lámina se observa la “interferencia” o “efecto Stroop” ya que el individuo debe nombrar el color del estímulo y no leerlo; interfiriendo así la lectura en la nominación (Golden, 2001). A partir de esto, otros investigadores, buscaron diseñar sus propias versiones para evaluar la atención e interferencia, realizando test con características propias.

a) Características Psicométricas.

La confiabilidad del Stroop, los investigadores en muestras de adultos utilizaron el método Test – Retest, con tiempos que oscilan entre un minuto y diez días entre ambas aplicaciones. Hensen (1965) obtuvo índices de 0.88, 0.79 y 0.71 para las tres puntuaciones directas. Golden (1975) halló puntuaciones de 0.89, 0.84 y 0.73 (N= 450) en la versión grupal y de 0.86, 0.82 y 0.73 (N=30) en aplicación individual (Golden, 2001).

En relación a la validez, existen también otras investigaciones que la han encontrado resultados significativos (Zajano, Hoyceanyls y Ouellette, 1981; Sichel y Chandler, 1969; en MacLeod, 1991). Además, se ha encontrado que la interferencia seguía apareciendo en igual medida aun de cambiar el estímulo de forma cualitativa (Mc Cown y

Arnoult, 1981; Regan, 1978 & Pretti, 1971; en MacLeod, 1991). Los autores que llevaron a cabo la estandarización hecha en España (2001) mencionan que deben tomarse con cuidado los puntajes hallados por individuos menos de 17 años y de más de 45 años, puesto que la muestra obtenida entre esas edades no fue del todo representativa; por lo tanto, no ha sido estadísticamente estudiada la confiabilidad y validez. Sin embargo, la experiencia clínica ha demostrado que el uso de esta prueba se ajusta a esas edades (Golden, 2001).

Para la baremación se consideró la muestra española y los resultados de la muestra española fueron superiores a la de 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, se pueda deber a como fue compuesta la muestra.

En cuanto a los datos de correcciones por edad en niños se podrá considerar cuando se disponga de muestras más numerosas. (Golden, 2007).

2.5. Procedimiento.

Para el desarrollo de la investigación, primero se tuvo que realizar las coordinaciones con las autoridades de las 5 instituciones educativas. Ante ello, se entregarán las cartas de presentación a los directores de cada institución educativa. En segunda instancia, se programaron las sesiones de aplicación en las instituciones educativas. Se informó a los participantes de los objetivos de investigación; así como, de la participación voluntaria de la investigación.

2.6. Método de análisis de datos

El análisis de datos inicio con la evaluación de supuestos de Multicolinealidad (Arias, 2008), la normalidad univariante mediante los índices de asimetría y curtosis g_1 y g_2 , sus valores estandarizados, z_1 y z_2 y el índice de asimetría y curtosis conjunta K_2 , identificándose una distribución normal en la escala de Colores y una distribución diferente a la normal en las escalas de Palabras y

Palabras-Color según el criterio de D'Agostino (1986) y Hair, Anderson, Tatham, & Black (2005), además el índice de normalidad multivariante de Mardia (1970), decidiéndose aplicar el Método de Máxima Verosimilitud en el AFE como lo sugiere Brown (2006).

Posteriormente se estimó el índice de confiabilidad por consistencia interna mediante las puntuaciones factoriales propuesto por Mislevy & Bock (1990), asimismo se estimó de forma puntual e intervalica los índices de confiabilidad por estabilidad mediante el método Test-ReTest utilizando el coeficiente de correlación de Pearson entre la primera y segunda aplicación del test y el coeficiente de correlación intraclase para las escalas del test. Finalmente se construyeron las normas específicas tipo percentil y puntuaciones típicas del test para cada rango de edades donde también se reportan estadísticos descriptivos para cada escala.

2.7. Aspectos éticos

En un inicio se informó a los adolescentes sobre los objetivos y la utilidad de la investigación; asimismo, se brindó la información correspondiente a los docentes y directores de las instituciones, en relación a los principios éticos a seguir antes, durante y después de la aplicación (Colegio de Psicólogos del Perú, 2017).

Asimismo, se tuvo en cuenta:

Consentimiento de participación en la investigación. Durante la evaluación, se brindó un consentimiento informado, documento que presentaba la información relacionada a la participación voluntaria, los datos de la investigadora y los criterios a seguir.

Confidencialidad. Toda la información recabada, ha sido usada por la investigadora, manteniendo siempre en reserva los datos de los participantes.

Veracidad. La información se obtuvo a través del instrumento mencionado, que es anónimo con la única finalidad de ser íntegros con la información

recogida. Además, se consideraron los principios éticos de: justicia, equidad, inclusión y beneficencia.

III. RESULTADOS

3.1. Evidencias de Validez de Estructura Interna.

Evaluación de supuestos univariados y multivariados.

En Tabla 2, se muestran los estadísticos descriptivos de las escalas del Test de Palabras y colores, útiles para verificar el cumplimiento de los supuestos de multicolinealidad mediante el rango de correlaciones entre ítems, hallándose valores entre .47 a .65, menores a .90 y evidenciando ausencia de redundancia en los contenidos medidos en los ítems (Arias, 2008), asimismo se observaron índices de asimetría de -.46 a .66 y sus valores estandarizados entre -4.47 a 6.42 e índices de curtosis de .33 a .91 y sus valores estandarizados entre 1.60 a 4.45, superando los valores G1 y G2 el rango de -1 a 1 y mayores a 1.96 en sus valores estandarizados Z1 y Z2, además de un índice de asimetría y curtosis conjunto univariado K^2 mayor a 5.99 en dos ítems, señalando el incumplimiento del supuesto de normalidad univariante como lo señalan D'Agostino (1986) y Hair, Anderson, Tatham, & Black (2005), finalmente el índice de curtosis multivariante mostró un valor de 4.37, inferior a 70, cumpliéndose el ajuste a una distribución normal multivariante (Mardia, 1970), decidiéndose aplicar el método de Máxima verosimilitud en el análisis Factorial Exploratorio como lo sugiere Brown (2006).

Tabla 2

Estadísticos descriptivos y Correlación Interescalas del Test de Colores y Palabras

Ítems	Estadísticos descriptivos univariados y multivariados							Correlaciones	
	M	DE	G ₁	Z ₁	G ₂	Z ₂	K ²	Palabras	Colores
Palabras	67.11	17.26	-.46	-4.47	.85	4.18	37.5		
Colores	46.99	12.04	.81	.81	.33	1.60	3.22	.65**	
Palabra - Color	24.02	7.92	.66	6.42	.91	4.45	61.0	.47**	.52**
							Normalidad Multivariante	4.37	

Nota: M: Media; DE: Desviación estándar, G₁: Asimetría Z₁: Asimetría estandarizada, G₂: Curtosis; Z₂: Curtosis estandarizada; K²: Curtosis y asimetría conjunta.

Unidimensionalidad.

Al aplicar el Análisis Factorial Exploratorio - AFE, las pruebas de adecuación muestral presentaron un KMO de .70 y un p-valor estadísticamente significativo ($p < .05$) en el Test de Bartlett, evidenciando la aplicabilidad del AFE, con el método de Máxima Verosimilitud, donde se evidenció un único factor con un autovalor (2.125) mayor a 1 y que explica un 70.84% de la varianza total contenida en los 3 ítems con cargas factoriales de .625 a .891 y comunalidades de .391 a .793 que evidencia la Unidimensionalidad.

Tabla 3

Análisis Factorial Exploratorio del Test de Colores y Palabras

Pruebas		
Medidas de adecuación		
Medida de Kaiser-Meyer-Olkin		
KMO		.70
Test de Bartlett		
χ^2		525.6
gl		3
p		<.001
Varianza Total Explicada		
Componentes	Autovalores	% Varianza explicada
1	2.125	70.84
2	.541	
3	.344	
Coefficientes estandarizados		
Ítems	Cargas	Comunalidades
Palabras	.625	.391
Colores	.891	.793
Palabra - Color	.742	.551

Nota: χ^2 : Estadístico Chi cuadrado; gl: Grados de libertad; p: Significancia

3.2. Evidencias de Confiabilidad

En la Tabla 4, se muestra que se estimó una varianza común de 1.735 de una varianza total de 3.00, obteniéndose un índice consistencia interna mediante las puntuaciones factoriales aplicando el método de Mislevy & Bock (1990) con un valor de .851, evidenciando una buena consistencia interna del Test de Colores y palabras.

Tabla 4

Índice de consistencia interna mediante las puntuaciones factoriales

Instrumento	Varianza común	Índice de Consistencia Interna
Test de colores y palabras	1.735	.851

En la Tabla 5, se muestran la estimación puntual e intervalica al 95% de los índices de correlación entre aplicaciones con el coeficiente de correlación de Pearson (r) con valores entre .779 a .812 en las escalas, asimismo la correlación intraclase (CCI) con valores entre .777 a .811 en los ítems, con una mínima diferencia (+/- .001) en los intervalos de confianza. Cabe destacar que los valores son similares por el tamaño de la muestra ($n=575$), dado que en muestras menores ($n<250$) la correlación entre aplicaciones tienda a sobreestimar la confiabilidad.

Tabla 5

Índices de Estabilidad mediante correlación entre aplicaciones (interclases) e intraclase.

Ítems	Correlación Interclase e IC95%			Correlación Intraclase e IC95%		
	r_{tr}	Li	Ls	CCI	Li	Ls
Palabras	.792	.727	.843	.792	.728	.843
Colores	.812	.752	.858	.811	.751	.857
Palabra - Color	.779	.710	.833	.777	.709	.831

Nota: r : Coeficiente de correlación de Pearson, CCI: Coeficiente de correlación intraclase; Li: Límite inferior; Ls: Límite superior.

3.3. Elaboración de Normas

En la Tabla 6, se evidencia una diferencia no significativa ($p \geq .05$) en los rangos promedio de la escala Interferencia y sus ítems Palabras, Colores, Palabra – Color entre varones y mujeres.

Tabla 6

Prueba de diferencias en el Test de Colores y Palabras por la variable sexo

Variables	Sexo		U de Mann-Whitney		
	Varones (n=319)	Mujeres (n=256)	U	Z	p
Palabras	281.4	296.3	38715.5	-1.07	.285
Colores	283.1	294.1	39269.5	-0.79	.430
Palabra - Color	290.6	284.7	39991.5	-0.42	.671
Interferencia	295.8	278.2	38328.0	-1.26	.206

Nota: U: Estadístico de Mann – Whitney, Z: Valor normal estándar, p: Significancia

Los resultados de la Tabla 7, se muestra evidencia estadísticamente significativa ($p < .05$) de una diferencia en las escalas de Palabras, Colores y Palabra-Color entre grupos de edades, con rangos promedio más altos a mayor edad en los tres ítems y rangos promedios menores en a grupos de mayor edad.

Tabla 7

Prueba de diferencias en el Test de Palabras y Colores según edad

Variables	Rangos promedio por Edad			Prueba de Kruskal Wallis		
	7 – 8 (n=140)	9 – 10 (n=213)	11 – 12 (n=222)	χ^2	gl	p
Palabras	142.43	290.70	377.21	171.65	2	<.001
Colores	165.90	292.81	360.38	118.03	2	<.001
Palabra - Color	206.29	273.45	353.49	70.14	2	<.001
Interferencia	328.83	268.93	280.54	11.71	2	.003

Nota: χ^2 : Estadístico Chi cuadrado, gl: Grados de libertad, p : Significancia

Con base en la evidencia de validez de constructo (Tabla 3) y confiabilidad (Tablas 4 y 5) del Test se procede a construir las normas tipo percentil y puntuaciones típicas tipo T para edades de 7 y 8 años, 9 y 10 años y de 11 a 12 años de edad en las Tablas 8, 9 y 10 respectivamente, evidenciándose valores más altos en la media (M), mediana (Me) y moda (Mo) a más edad en los ítems de Palabras, Colores y Palabra-Color y de forma inversa con la escala de Interferencia. Asimismo, los coeficientes de variación (CV) también muestran un comportamiento inverso a la edad, es decir, a más edad menos variabilidad.

Tabla 8

Normas percentil y tipo T del Test de Colores y Palabras para niños y niñas de 7 y 8 años

Pc	Palabras	Colores	Palabra - Color	Interferencia	T
99	87	60	52	314	73
95	80	54	34	150	66
90	70	51	30	110	63
85	67	50	27	60	60
80	66	46	26	44	58
75	62	44	25	25	57
70	60	43	23	12	55
65	59	42	22	0	54
60	58	40	21	-6	53
55	56	40	20	-12	51
50	55	38	19	-21	50
45	52	37	18	-28	49
40	49	37	18	-40	47
35	46	36	18	-51	46
30	43	34	17	-58	45
25	41	32	16	-69	43
20	40	31	15	-77	42
15	34	29	14	-91	40
10	27	28	12	-102	37
5	18	24	11	-115	34
1	9	4	2	-145	27
N	140	140	140	140	N
M	51.38	38.61	20.47	-9.94	M
Me	55	38	19	-21	Me
Mo	40	37	18	-70	Mo
DE	16.82	9.19	7.23	82.03	DE
g1	-.498	-.218	.994	1.226	g1
g2	.074	.760	2.323	2.102	g2
Mín.	9	4	2	-145	Mín.
Máx.	87	60	52	314	Máx.

Nota: M: Media, Me: Mediana, Mo: Moda, DE: Desviación Estándar, g1: Asimetría g2: Curtosis

Tabla 9

Normas percentil y tipo T del Test de Colores y Palabras para niños y niñas de 9 y 10 años

Pc	Palabras	Colores	Palabra - Color	Interferencia	T
99	109	77	53	232	73
95	89	63	40	83	66
90	85	60	33	37	63
85	80	59	30	18	60
80	80	57	28	3	58
75	78	55	27	-4	57
70	74	53	26	-13	55
65	73	51	25	-20	54
60	72	50	24	-25	53
55	70	49	23	-34	51
50	69	48	22	-41	50
45	67	46	21	-51	49
40	65	45	21	-57	47
35	63	44	20	-66	46
30	61	42	19	-74	45
25	60	41	19	-83	43
20	57	38	18	-101	42
15	54	35	17	-112	40
10	52	32	15	-130	37
5	44	25	12	-158	34
1	20	15	2	-207	27
N	213	213	213	213	N
M	67.90	46.85	23.23	-42.07	M
Me	69	48	22	-41	Me
Mo	80	49,00 ^a	20	-57	Mo
DE	13.75	11.35	7.57	68.86	DE
g1	-.199	-.374	.599	.480	g1
g2	.727	.148	1.080	1.268	g2
Mín.	20	15	2	-207	Mín.
Máx.	109	77	53	232	Máx.

Nota: M: Media, Me: Mediana, Mo: Moda, DE: Desviación Estándar, g1: Asimetría g2: Curtosis

Tabla 10

Normas percentil y tipo T del Test de Colores y Palabras para niños y niñas de 11 a 12 años

Pc	Palabras	Colores	Palabra - Color	Interferencia	T
99	125	84	57	180	73
95	99	73	40	79	66
90	96	69	38	36	63
85	90	62	34	22	60
80	88	60	33	15	58
75	85	59	31	3	57
70	83	57	30	-7	55
65	81	56	29	-15	54
60	80	55	29	-22	53
55	77	53	27	-32	51
50	75	52	26	-41	50
45	74	51	25	-49	49
40	71	50	25	-58	47
35	70	47	23	-64	46
30	68	46	22	-70	45
25	68	44	21	-81	43
20	66	43	20	-90	42
15	62	41	19	-100	40
10	60	39	19	-116	37
5	56	35	16	-145	34
1	43	24	14	-220	27
N	222	222	222	222	N
M	76.28	52.41	27.02	-37.85	M
Me	75	52	26	-41	Me
Mo	68	52	20	-82	Mo
DE	13.13	11.21	7.57	63.02	DE
g1	.355	.399	.768	.224	g1
g2	.111	.269	.862	.399	g2
Mín.	43	24	14	-220	Mín.
Máx.	125	84	57	180	Máx.

Nota: M: Media, Me: Mediana, Mo: Moda, DE: Desviación Estándar, g1: Asimetría g2: Curtosis

Interpretación: Los examinados que tienen un puntaje menos de 25 se ubican en un nivel bajo de control atencional, es decir son muy sensibles a la interferencia. Nivel medio bajo (30 a 45) indica que tiende a ser sensible a la interferencia. Nivel medio (50) que tiene un adecuado control atencional. Nivel medio alto (55 a 75) de control atencional significa que tiende a ser resistente a la Interferencia y un nivel alto o superior (80 a 99) que significa que posee resistencia a la interferencia.

IV. DISCUSIÓN.

La presente investigación, tuvo como objetivo determinar las propiedades psicométricas del Test de colores y palabras (Stroop) en niños del distrito de Trujillo. Los resultados son detallados a continuación:

Como primer análisis, se precedió a analizar, los estadísticos descriptivos de las escalas del Test de Palabras y colores, para poder determinar el cumplimiento o no, de los supuestos de multicolinealidad, a través de las correlaciones entre ítems, encontrándose valores que oscilaron de .47 a .65 (menores a .90), demostrándose así, una ausencia en la redundancia en los contenidos que miden las escalas (Arias, 2008), asimismo se observaron índices de asimetría de -.46 a .66 y sus valores estandarizados entre -4.47 a 6.42 e índices de curtosis de .33 a .91 y sus valores estandarizados entre 1.60 a 4.45, superando los valores G1 y G2 el rango de -1 a 1 y mayores a 1.96 en sus valores estandarizados Z1 y Z2, además de un índice de asimetría y curtosis conjunto univariado K2 mayor a 5.99 en dos ítems, que indican un incumplimiento del supuesto de normalidad univariante (D'Agostino, 1986), (Hair, Anderson, Tatham, & Black, 2005), finalmente el índice de curtosis multivariante presenta un valor de 4.37, inferior a 70, cumpliéndose el ajuste a una distribución normal multivariante (Mardia, 1970), decidiéndose aplicar el método de Máxima verosimilitud en el análisis Factorial Exploratorio como lo sugiere Brown (2006).

Frente a esto, se procedió a realizar el Análisis Factorial Exploratorio (AFE), presentando así, una adecuación muestral (KMO) de .70 y un p-valor estadísticamente significativo ($p < .05$) en el Test de Bartlett ($X^2=525.6$, $gl=3$, $p < .001$), conllevando a que se haga posible la aplicabilidad del AFE, con el método de Máxima Verosimilitud, donde se evidencia un único factor con un autovalor (2.125) mayor a 1, con una varianza explicada del 70.84% de la varianza total contenida en los 3 ítems con cargas factoriales de .625 a .891 y comunalidades de .391 a .793 que evidencia la Unidimensionalidad. Resultados similares, fueron encontrados por Gutiérrez et al., (2018) quienes determinaron de igual forma la presencia de un solo factor, que explica el 64% de la varianza, así como una correlación estadísticamente significativa entre

sus ítems ($p < .00$); de igual forma Maureira, Aravena, Gálvez y Flores (2014) determinó la presencia de un solo factor y una relación altamente significativa entre los elementos de la prueba. Por su parte Ramos (2009) determinó en su análisis la presencia de un solo factor, que explica el 96% de la varianza, así como una correlación estadísticamente significativa ($p < .00$) entre las palabras leídas y los colores nombrados.

Para el análisis de la confiabilidad, se estimó una varianza común de 1.735 de una varianza total de 3.00, obteniéndose un índice consistencia interna mediante las puntuaciones factoriales halladas, aplicándose el método de Mislevy & Bock (1990), presentando una fiabilidad de .851, evidenciando una buena consistencia interna del instrumento, considerado como adecuado, determinándose así una correlación entre los ítems de la prueba (Campo-Arias & Oviedo, 2008). Resultados similares fueron encontrados por Gutiérrez et al., (2018) quien empleó el alfa de Cronbach para determinar la confiabilidad, logrando encontrar una fiabilidad de .96; de igual Maureira et al., (2014) a través del alfa de Cronbach, determinó una confiabilidad de .88; por su parte, Ramos (2009) logró determinar una confiabilidad de .65.

Se realizó, además, el método test retest, como medida para estimar la confiabilidad, que en base a Reidi-Martínez (2015) permite conocer la temporalidad en los datos que el instrumento de evaluación brinda; se muestra una estimación puntual e interválica al 95% de los índices de correlación entre aplicaciones con el coeficiente de correlación de Pearson (r) con valores entre .779 a .812 en las escalas, asimismo la correlación intraclase (CCI) con valores entre .777 a .811 en las escalas, con una mínima diferencia ($\pm .001$) en los intervalos de confianza, frente a esto, es necesario, considerar que los valores son similares por el tamaño de la muestra ($n=575$), dado que en muestras menores ($n < 250$) la correlación entre aplicaciones tiende a sobreestimar la confiabilidad. Resultados similares, se pueden evidenciar en los estudios realizados por Rodríguez et al., (2016) quien determinó en el test-retest una fiabilidad de .91. Esto permite comprobar, que el test Stroop, posee adecuadas propiedades psicométricas, reflejadas en la temporalidad, conllevando a suponer que es capaz de medir

satisfactoriamente la atención en estudiantes, cuyas edades oscilen de siete a doce años.

Previo al objetivo de normas, se evidenció que no existe una diferencia significativa ($p \geq .05$) en los rangos promedio de la escala Interferencia y sus ítems Palabras, Colores, Palabra – Color entre varones y mujeres; sin embargo, se aprecia una diferencia estadísticamente significativa ($p < .05$) en las escalas entre grupos de edades, con rangos promedio más altos a mayor edad en los tres ítems y rangos promedios menores en a grupos de mayor edad.

Ante esto, se evidencian normas percentilares y tipo T para edades de 7 y 8 años, 9 y 10 años y de 11 a 12 años de edad en las Tablas 8, 9 y 10 respectivamente, evidenciándose valores más altos en la media (M), mediana (Me) y moda (Mo) a más edad en los ítems de Palabras, Colores y Palabra-Color y de forma inversa con la escala de Interferencia. Asimismo, los coeficientes de variación (CV) también muestran un comportamiento inverso a la edad, es decir, a más edad menos variabilidad. Estos resultados permiten determinar que, en base a las experiencias y el desarrollo físico, cognitivo y emocional, los evaluados con mayor edad, evidencian una mejor capacidad para prestar atención a los estímulos del ambiente, al respecto se debe tomar en cuenta, que la atención, va desarrollándose de acuerdo a la edad, estimulación y la enseñanza (Boujon y Qualreau, 2004); asimismo, conforme los niños van pasando etapas, presentan una mejor capacidad para retener, comprender y discriminar la información brindada, así como distinguirla de otros elementos (Ardila, Rosselli y Matute, 2010); al respecto, Rodríguez et al., (2016) refiere que la atención es un proceso que mejora con la edad y permite que se pueda tener un foco de atención mayor y discriminación de los estímulos que se encuentren presentes. En relación a la interferencia y variabilidad, la atención, representa un proceso de selección, de aquello considerado como relevante y se pueda así filtrar mejor los detalles y la persona evite prestar atención a diversos estímulos a la vez, que tiende a desarrollarse con la edad, el proceso de maduración y la estimulación

correcta, por ende, disminuye la sensación fatiga y permite que la información importante sea retenida pase a la memoria de largo plazo (Rosselló, 1998).

Finalmente, frente a los resultados obtenidos, se considera, que el Test de Colores y Palabras (Stroop) es una herramienta útil, válida y confiable, en la evaluación de la atención en niños del distrito de Víctor Larco, además, puede ser empleada en evaluaciones grupales e individuales, contribuyendo a que estudiantes o profesionales del campo de la psicología puedan emplearla en su trabajo para el diagnóstico oportuno o para futuras investigaciones.

V. CONCLUSIONES.

En base a los resultados se concluye que:

- Se obtuvieron las propiedades psicométricas, empleándose el Análisis Factorial Exploratorio - AFE, evidenciándose la presencia de un solo factor, con una varianza explicada del 70.84%, el KMO de .70 y un p-valor estadísticamente significativo ($p < .05$) en el Test de Bartlett, las cargas factoriales oscilan de .625 a .891 y comunalidades de .391 a .793.
- Se logró estimar la confiabilidad a través del índice consistencia interna mediante las puntuaciones factoriales, empleándose el método de Mislevy y Bock (1990) con un valor de .851.
- Se estimó, además, la confiabilidad a través del método test-retest, encontrándose una estimación puntual e intervalica al 95% de los índices de correlación entre aplicaciones con el coeficiente de correlación de Pearson (r) con valores entre .779 a .812 en las escalas, asimismo la correlación intraclase (CCI) con valores entre .777 a .811 en las escalas, con una mínima diferencia ($\pm .001$) en los intervalos de confianza.
- Se realizó como último objetivo, las normas tipo percentil y puntuaciones típicas tipo T para edades de 7 y 8 años, 9 y 10 años y de 11 a 12 años de edad en las Tablas 8, 9 y 10 respectivamente, evidenciándose valores más altos en la media (M), mediana (Me) y moda (Mo) a más edad en los ítems de Palabras, Colores y Palabra-Color y de forma inversa con la escala de Interferencia. Asimismo, los coeficientes de variación (CV) también muestran un comportamiento inverso a la edad, es decir, a más edad mayor variabilidad.

VI. RECOMENDACIONES.

En base a los resultados se recomienda:

- Emplear el Test de Colores para la evaluación y diagnóstico neuropsicológico en niños del distrito de Trujillo.
- Para la validez, se recomienda en futuras investigaciones realizar un análisis de validez convergente con el Test de Atención D2.
- Realizar la adaptación del instrumento en otros contextos del distrito y a nivel nacional, ya que representa una herramienta sencilla y útil, para medir la atención.

VII. PROPUESTA.

En nuestro país, la mayoría de instrumentos psicológicos y neuropsicológicos que son utilizados para evaluar y brindar diagnósticos no cuentan con la garantía de ser en su mayoría instrumentos válidos y confiables, utilizándose es más indiscriminadamente, sin considerar las consecuencias éticas y prácticas. Las incorrectas prácticas incluyen usar pruebas que en su mayoría son realizadas y normalizadas con el lenguaje, nivel socioeconómico y cultural de otros países, lo que constituye una constante preocupación de su aplicación en nuestra población. La prueba de J. R. Stroop se utiliza para evaluar atención, sensibilidad o resistencia a la interferencia, y funciones ejecutivas el cual vendría a ser la ausencia de acciones automáticas en favor de otras más poco usuales (Stroop, 1935, citado por Golden 2001).

Es, por tanto, que la presente investigación realizada, de tipo psicométrico, brinda como aporte principal, un instrumento que posee validez, confiabilidad y normas percentilares, para la evaluación de la atención y el control inhibitorio en estudiantes del nivel primario, aspecto importante y necesario durante el proceso de desarrollo (cognitivo, físico y emocional), a nivel personal, académico, social y familiar. En relación a los resultados evidenciados en la presente investigación, los profesionales, podrán hacer uso de la herramienta y a partir de esto, poder diagnosticar apoyándose en un instrumento válido y confiable y por ende detectar casos clínicos infantiles (TDAH, entre otros) y así derivar a los especialistas para esos casos, de igualmente se puede realizar evaluaciones grupales para obtener el nivel de control atencional en los niños y así proponer planes de intervención a nivel grupal e individual, para aquellos estudiantes con dificultades en la atención las cuales repercuten notablemente en su aprendizaje y por ende en su rendimiento escolar.

VIII. REFERENCIAS.

- Ardila, A.; Rosselli, M. y Matute, E. (2010). *Neuropsicología del desarrollo infantil*. México D.F.: Manual Moderno.
- Arentsen, T.; Boone, K.; Goldberg, H.; Cottingham, E.; Ziegler, E. y Zeller, M. (2013). Effectiveness of the Comalli Stroop Test as a Measure of Negative Response Bias. *Revista The Clinical Neuropsychologist* 27(6), 285.
- Arias, B. (2008). Desarrollo de un ejemplo de análisis factorial confirmatorio con LISREL, AMOS y SAS.
- Armistead-Jehle, P. & Buican, B. (2012). Evaluation context and Symptom Validity Test performances in a U.S. Military sample. *Archives of Clinical Neuropsychology*, 27 (8), 828-839. Recuperado de <https://academic.oup.com/acn/article/27/8/828/5194>
- Bezdicek, O.; Lukavsky, J.; Stepankova, H.; Nikolai, T.; Axelrod, B.; Michalec, J.; Růžička, E. & Kopecek (2015). The Prague Stroop Test: Normative standards in older Czech adults and discriminative validity for mild cognitive impairment in Parkinson's disease. *Journal of Clinical and experimental Neuropsychology*, 37(8):794-807. Recuperado de <https://www.ncbi.nlm.nih.gov/pubmed/26313510>
- Boujon, Ch. & Qualreau, Ch. (2004). *Atención, aprendizaje y rendimiento escolar*. Madrid: Narcea
- Brown, T. (2016). *¿Los niños superan las dificultades de las funciones ejecutivas cuando crecen?*. *Understood dificultades de aprendizaje y de atención*. Recuperado de <https://www.understood.org/es-mx/learning-attention-issues/child-learning-disabilities/executive-functioning-issues/do-kids-outgrow-executive-functioning-issues>
- Brown, T. (2006). *Confirmatory factor analysis for applied research*. New York: The Guilford Press.
- Buendía, L., Colás, P. y Hernández, F. (1998). *Métodos de Investigación en Psicopedagogía*. Madrid: McGraw-Hill.
- Bush, S., Ruff, R., Tröster, A., Barth, J., Koffler, S., Pliskin, N., Reynolds, C. & Cheryl H. Silver (2005). Symptom validity assessment: Practice issues and medical necessity NAN Policy & Planning Committee. *Archives of*

- Clinical Neuropsychology*, 20 (4), 419–426. Recuperado de <https://academic.oup.com/acn/article/20/4/419/2521>
- Colegio de Psicólogos del Perú. (2017). *Código de Ética y Deontología*. Perú. Recuperado de https://www.cpsp.pe/documentos/marco_legal/codigo_de_etica_y_deontologia.pdf
- Conca, B. e Ibarra, M. (2004). *Estandarización de la prueba de colores y palabras de Stroop en niños de 8 a 12 años para la región metropolitana, 2004* (Tesis de licenciatura). Universidad de Chile, Santiago, Chile. Recuperado de <http://repositorio.uchile.cl/handle/2250/113385>
- Cordero, A., Seisdedos, N., González, M. y De la Cruz, V. (2013). *Toulouse-Piéron revisado- Prueba perceptiva y de atención*. Madrid, España: TEA Ediciones.
- Chelune, G. & Thompson, L. (1987). Evaluation of the general sensitivity of the Wisconsin Card Sorting Test among younger and older children. *Developmental Neuropsychology*, 3, 81–89.
- D'Agostino, R. (1986). *Tests for the normal distribution*. New York: Marcel Dekker.
- Davidson, D. J., Zacks, R. T., & Williams, C. (2003). Stroop Interference, practice and aging. *Aging Neuropsychology and Cognition*, 10, 85-98. Recuperado de <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1761647/>
- De Vega, M. (2006). *La atención: Paradigma de atención dividida*. Madrid: Alianza.
- Dempster, F. & Brainerd, CH. (1995). *Interference and Inhibition in Cognition*. California: Academic Press, Inc,
- Denckla, M. (1996). *A theory and model of executive function: A neuropsychological approach*. Baltimore: Paul H. Brookes.
- Golden, C. (2001). *Stroop: Test de Colores y Palabras*. Madrid: TEA Ediciones.
- Golden, C. & Golden, Z. (2002). Patterns of performance on the Stroop Color and Word Test in children with learning, attentional, and psychiatric

- disabilities. *Psychology in the Schools*, 39(5), 489–495. Recuperado de https://nsuworks.nova.edu/cps_facarticles/585/
- Golden, C. (2007). *Stroop test de colores y palabras, manual* (5° Ed.). Madrid: TEA Ediciones.
- Gregoory, R. (2012), *Pruebas psicológicas. Historia principios y aplicaciones*. México: Pearson Educación, S.A.
- Gutiérrez, F.; Ramos, M. y Villa, O. (2018). Eficacia ejecutiva en tareas de interferencia tipo Stroop. Estudio de validación de una versión numérica y manual (CANUM). *Anales de Psicología*, 34(1), 186-194.
- Hair, J.; Anderson, R.; Tatham, R. y Black, W. (1999). *Análisis Multivariante*. (5ta Ed). Madrid: Prentice Hall.
- Heilbronner, R., et al. (2009). Neuropsychological Consequences of Boxing and Recommendations to Improve Safety: A National Academy of Neuropsychology Education Paper. *Archives of Clinical Neuropsychology* 24 (1) 11– 19.
- Hernández, R., Fernández, C. & Baptista, M. (2014). *Metodología de la Investigación*. México: McGRAW-HILL / Interamericana Editores, S.A. DE C.V. Recuperado de <http://observatorio.epacartagena.gov.co/wp-content/uploads/2017/08/metodologia-de-la-investigacion-sexta-edicion.compressed.pdf>
- Huguet, T., Galvaing, M., Monteil, J., & Dumas, F. (1999). Social presence effects in the Stroop task: further evidence for an attentional view of social facilitation. *Journal of Personality and Social Psychology*, 77(5), 1011-1025.
- Hynd, G.; Semrud-Clikeman, M.; Lorys, A.; Novey, E. & Eliopoulos, D. (1990). Brain morphology in developmental dyslexia and attention deficit disorder/hyperactivity. *Archives of Neurology*, 47(8), 919-926.
- Katz, M. (2006). *Multivariable analysis*. (2da Ed.). Cambridge: Cambridge University Press.
- Kerns, K., McInerney, R. & Wilde, N. (2001). Time reproduction, working memory, and behavioral inhibition in children with ADHD. *Child The Neuropsychologist*, 7(1), 21-31. Recuperado de <https://www.ncbi.nlm.nih.gov/pubmed/11815878>

- Larrabee, G. (2010). Detection of malingering using atypical performance patterns on standard neuropsychological tests. *The Clinical Neuropsychologist*, 17(3), 410-425.
- Lavado, P.; Gómez V.; Sawada M.; Chomalí M. & Álvarez M. (2003). Diagnósticos neurológicos en la atención primaria de salud en Santiago, Chile. *Revista Neurológica*, 36 (6), 518-522. Recuperado de <https://www.neurologia.com/articulo/2002265>
- Lezak, M.; Howieson, D. & Loring, D. (2004). *Neuropsychological assessment*. (4th ed.). New York: Oxford University Press.
- López, C. y García, J. (1997). *Problemas de atención en el niño*. Madrid: Pirámide.
- Luria, A. (1966). *Higher cortical functions in man*. NY: Basic Books.
- MacLeod, C. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109 (2) 163-203.
- Mardia, K. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika* 57, 519–530.
- Mauriera, F.; Aravena, C.; Gálvez, C. & Flores, E. (2014). Propiedades Psicométricas y datos normativos del test de Stroop y del test torre de Hanoi en estudiantes de educación física de Chile. *Revista GPU*, 10(3), 344-349.
- Mead, J.; Bruning, A.; Gill, M.; Steiner, A.; Acton, T y Vershon, A. (2002). Interactions of the Mcm1 MADS box protein with cofactors that regulate mating in yeast. *Mol Cell Biol*, 22(13), 4607-21. Recuperado de <https://www.ncbi.nlm.nih.gov/pubmed/12052870>
- Meneses, J.; Barrios, M.; Bonillo, A.; Cosculluela, A.; Lozano, L.; Turbany, J. y Valero, S. (2013). *Psicometría*. Barcelona: Editorial UOC Mesulam.
- Mislevy, R. & Bock, R. (1990). *BILOG 3 Item analysis and test scoring with binary logistic models*. Mooresville: Scientific Software.
- Mittenberg, W.; Patton, C.; Canyock, E. & Condit, D. (2002). Base rates of malingering and symptom exaggeration. *Journal of Clinical & Experimental Neuropsychology*, 24 (8), 809-828. Recuperado de <https://www.ncbi.nlm.nih.gov/pubmed/12650234>

- Moering, R. G., Schinka, J. A., Mortimer, J. A., & Graves, A. B. (2004). Normative data for elderly African Americans for the Stroop Color and Word Test. *Archives of Clinical Neuropsychology*, 19(1), 61-71. Recuperado de <https://www.sciencedirect.com/science/article/pii/S0887617702002196>
- Montero, I. & León, O. (2007). A guide for naming research studies in Psychology. *International Journal of Clinical and Health Psychology*, 7, 847-862. Recuperado de <http://www.redalyc.org/articulo.oa?id=33770318>
- Morgan, A. & Lilienfeld, S. (2000). A meta-analytic review of the relation between antisocial behavior and neuropsychological measures of executive function. *Clinical Psychology Review*, 20(1), 113–136.
- Posner, M. & DiGirolamo, G. (1998). Conflict, target detection and cognitive control. En: R. Parasuraman (Ed.) *The attentive brain* (pp. 401-423). Cambridge: MIT Press.
- Posner, M. & Raichle, M.E. (1994). *Images of mind*. New York: Scientific American Library.
- Posner, M. & Rothbart, M. (1991). Attentional mechanisms in conscious experience. *Foundations of neuropsychology. The neuropsychology of consciousness* (p. 91–111). Academic Press. Recuperado de: <https://doi.org/10.1016/B978-0-12-498045-7.50010-4>
- Posner, M. & Rothbart, M. (2014). Attention to learning of school subjects. *Trends in Neuroscience and Education*, 3(2), 14-17, doi: 10.1016/j.tine.2014.02.003
- Puente, A. (2003). *Cognición y aprendizaje. Fundamentos psicológicos*. Madrid: Pirámide.
- Puerta, I.; Dussán, C.; Montoya, D. y Landínez, D. (2018). Estandarización de pruebas neuropsicológicas para la evaluación de la atención en estudiantes universitarios. *CES Psicología*, 12(1), 17-31. DOI: <http://dx.doi.org/10.21615/cesp.12.1.2>
- Quiroga, A.; Santacreu, J.; Montoro, A.; Martínez, A. y Chun, P. (2011). Evaluación Informatizada de la Atención para Niños de 7 a 11 Años: El DiViSA-UAM y el TACI-UAM. *Clínica y Salud*, 22(1), 3-20.

- Ramos, S. (2009). *Adaptación del test de colores y palabras de Stroop en una muestra portuguesa. Influencia de la reserva cognitiva en la función ejecutiva de sujetos sanos y con enfermedad tipo alzheimer de inicio tardío*. (Tesis Doctoral). Universidad de Salamanca, Salamanca, España.
- Reidi-Martínez, L. (2013). Confiabilidad en la medición. *Investigación en Educación Médica*, (6), 107-111.
- Riccio, C.; Hall, J.; Morgan, A.; Hynd, G.; Gonzalez, J. & Marshall, R. (1994). Executive function and the Wisconsin Card Sorting Test: Relationship with behavioral ratings and cognitive ability. *Developmental Neuropsychology*, 10(3), 215–229. Recuperado de <https://psycnet.apa.org/record/1995-32658-001>
- Rodríguez, L.; Pulido, N. & Pineda, C. (2016). Propiedades psicométricas del Stroop, test de colores y palabras en población colombiana no patológica. *Revista Universitas Psychologica*, 15(2), 255-272. Recuperado de <http://dx.doi.org/10.11144/Javeriana.upsy15-2.ppst>
- Rosselli, M; Ardila, A; Santisi, M; Arecco, M.R; Salvatierra, J; Conde, A. & Lenis, B. (2002). Stroop effect in Spanish. *Journal of the International Neuropsychological Society*, 8, 819 - 827. Recuperado de https://www.researchgate.net/publication/11149631_Stroop_effect_in_Spanish-English_bilinguals
- Rosselló, J. (1998). *Psicología de la Atención Introducción al estudio de mecanismo atencional*. Madrid, España : Editorial Pirámide.
- Schutte, C., Axelrod, B. & Montoya, E. (2015). Making Sure Neuropsychological Data Are Meaningful: Use of Performance Validity Testing in Medico legal and Clinical Contexts. *Psychological Injury and Law* 8(2), 49. Recuperado de <https://psycnet.apa.org/record/2015-26989-001>
- Seisdedos, N. (2012). d2, *Test de Atención*. Madrid, España: TEA Ediciones.
- Shapiro, B. (2011). Bajo rendimiento escolar: una perspectiva desde el desarrollo del sistema nervioso. *Revista Médica Clínica*, 22(2), 218-225.

- Stroop, J. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643–662. Recuperado de <http://dx.doi.org/10.1037/h0054651>
- Sweet, J., Benson, L., Nelson, N. & Moberg, P. (2015). Professional Practice and ‘Salary Survey’: Professional Practices, Beliefs, and Incomes of U.S. Neuropsychologists. *The Clinical Neuropsychologist*, 29(8), 1069-1162. Recuperado de <https://www.tandfonline.com/doi/full/10.1080/13854046.2016.1140228>
- Uttl, B., & Graf, P. (1997). Color-Word Stroop test performance across the adult life span. *Journal of Clinical and Experimental Neuropsychology*, 19(2), 405-420. Recuperado de <https://www.ncbi.nlm.nih.gov/pubmed/9268815>
- Velásquez-Molina, E. & Ordoñez-Huamán, C. (2015). Trastorno por uso de sustancias y trastorno por déficit de atención e hiperactividad: frecuencias en pacientes varones hospitalizados en un servicio de adicciones. *Revista de Neuro-Psiquiatría*, 78(2), 73-79.
- Zelazo, P.; Anderson, J.; Richler, J.; Wallner, K.; Beaumont, J. y Weintraub, S. (2013). NIH Toolbox Cognition Battery (CB): measuring executive function and attention. *The Society for Research in Child Development, Inc.* 78(4),16-33. Recuperado de <https://www.ncbi.nlm.nih.gov/pubmed/23952200>
- Zelazo, P.; Carter, A.; Reznick, J. & Frye, D. (1997). Early development of executive function: A problem-solving framework. *Review of General Psychology*, 1(2), 198–226.

IX. ANEXOS.

Anexo 01 – Cuestionario

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

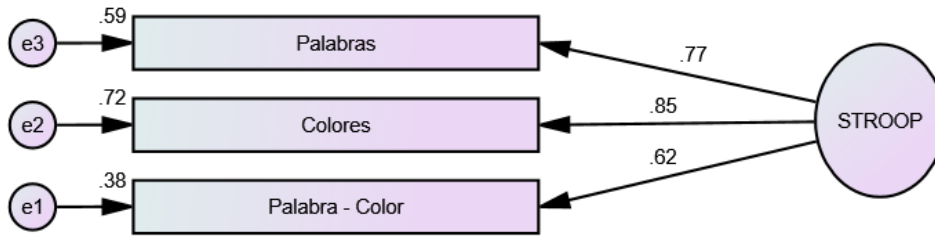


Copyright de la edición española © 1993 by TEA Ediciones, S.A., Madrid (España) - Traducido y adaptado con permiso del propietario original, Stoelting Company, Illinois (U.S.A.) - Edita: TEA Ediciones, S.A. - Prohibida la reproducción total o parcial. Todos los derechos reservados - Printed in Spain. Impreso en España por CIPSA; Orense, 68, 28020 Madrid - Depósito legal: M - 117 - 1994.

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

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

Anexo 02 – Identificación del modelo



Identificación del modelo

Variables	N	Identificación
Elementos conocidos		
Variables observadas	3	$(3*4)/2=6$
Parámetros diferentes a estimar	6	6
Cargas factoriales	3	
Correlaciones múltiples al cuadrado	3	
Grados de libertad		$(6-6)=0$
Identificación		Modelo identificado ($gl=0$)
Generalización		No procede