

Propiedades de medición del inventario de despistaje preescolar de Minneapolis (IDPM)

Angulo Ramos Marisol¹, Merino Soto César²

RESUMEN

Objetivo: evaluar exploratoriamente la validez y consistencia interna del Inventario de Despistaje Preescolar de Minneapolis (IDPM), y su relación entre con el TEPSI, en una muestra de niños peruanos. **Material y métodos:** participaron 52 niños, entre los 3 y 5 años de edad, provenientes de un programa no escolarizado de Educación Preescolar en Chorrillos (Lima). Se determinó la validez de los ítems (dificultad y discriminación del IDPM) y las correlaciones con los sub tests del TEPSI. **Resultados:** la consistencia interna fue 0.80; los ítems tuvieron moderada dificultad; las correlaciones entre el IDPM y TEPSI fueron estadísticamente significativas y entre moderadas y altas. **Conclusiones:** el IDPM es un instrumento con aceptables cualidades métricas y puede ser de utilidad para apoyar las evaluaciones de despistaje, y por lo tanto puede ser incorporado en los cursos de desarrollo del niño y prácticas profesionales.

Palabras clave: consistencia Interna, validez, evaluación de desarrollo, MPSI, TEPSI. Perú. (Fuente DeCs BIREME).

Measurement properties of Minneapolis Preschool Screening Inventory (MPSI)

Objective: Exploratory assess the validity and internal consistency of the instrument Minneapolis Preschool Screening Inventory and their relationship with TEPSI, in a sample of Peruvian children. **Material and Methods:** 52 children between 3 and 5 years old, from a program of Education Pre-school in Chorrillos (Lima). The study assesses the difficulty and discrimination of items, and the reliability and validity of MPSI with TEPSI. **Results:** The MPSI items show discriminative and medium difficulty; the internal consistency was 0.80; the MPSI-TEPSI correlations between moderate and high. **Conclusion.** In this initial assessment, the MPSI is an instrument with acceptable metric proprieties, and it can, therefore the MSPI can be incorporated into courses in child development and professional practices.

Key words: Internal consistency, validity, assessment development, MPSI, TEPSI. Peru. (Source DeCs BIREME).

¹ Licenciada en enfermería, docente en el Instituto Superior Tecnológico Privado Santa Rosa.

² Licenciado en psicología, docente-investigador el Instituto de Investigación de Psicología de la Universidad de San Martín de Porres; Lima, Perú.

INTRODUCCIÓN

El Ministerio de Salud, (MINSA - Perú), ha integrado para la población infantil de 2 a 5 años, el instrumento de evaluación, Test de desarrollo Psicomotor (TEPSI) (1), que viene siendo utilizada durante muchos años por las(os) enfermeras(os) en el área de crecimiento y desarrollo (CRED) de los centros de salud público; la finalidad de su uso es la detección temprana de posibles alteraciones en el desarrollo psicomotor infantil. Aunque es una herramienta útil y recomendada para hacer evaluaciones de despistaje temprano, parece existir reacciones negativas entre sus usuarios, ya que pueden poner en cuestionamiento su eficiencia en el contexto peruano de los centros de salud. Esta falta de eficiencia responde a las condiciones de la práctica asistencial de la enfermera en los centros estatales de salud, donde hay fuertes demandas de atención eficaz y eficiente a la población numerosa de familias que recurren a los servicios de salud infantil. Precisamente, el TEPSI es una herramienta que requiere más tiempo de lo justificado para una evaluación de despistaje, y por lo tanto no ayudaría a cumplir con lo estipulado por la norma del MINSA que exige que una evaluación de desarrollo debiera ocurrir en un tiempo de 30 a 45 minutos (2). Junto con las demandas de evaluación psicomotora, los enfermeros también desarrollan actividades alternas a la evaluación, por ejemplo el examen físico, vacunas, consejería, entre otros.

Por otro lado, los instrumentos utilizados en la edad preescolar han estado sujetos a constantes críticas respecto a sus características técnicas, que pueden poner los límites en la interpretación de sus resultados, y finalmente su utilidad en el terreno profesional y de investigación. Bajo los criterios cuantitativos y cualitativos desarrollados para evaluar sus características de adecuabilidad en preescolares, (3-5) la evaluación de instrumentos cognitivos preescolares ha arrojado resultados variables y aceptables. Estas han llegado a la conclusión que las baterías de pruebas preescolares de despistaje y diagnóstico individual, poseen una amplia variación de propiedades métricas, que van desde moderado hacia alto respecto a la consistencia interna y estabilidad test-retest. En estas revisiones también se observó que, aunque las evidencias de validez provienen de múltiples fuentes, los valores de validez también fueron relativamente inestables, ya que el tamaño de la muestra usada en los manuales de las pruebas revisadas fue pequeño en muchos casos (3). Además de las críticas publicadas, hay otros aspectos que deben moderar el entusiasmo por el uso acrítico de instrumentos en el nivel preescolar, por ejemplo, las normas antiguas de interpretación de puntajes, la disimilaridad de los contenidos

de las diferentes pruebas aunque la etiqueta nominal sea la misma, la practicidad de su estructura, etc. (4).

Dentro de los usos destinados de las pruebas en la edad preescolar, el despistaje o tamizaje podría ser considerado como el objetivo más frecuente en la práctica clínica. El despistaje es un proceso de evaluación para detectar a niños con pobres habilidades pre-académicas o problemas de desarrollo, y que contienen tareas consideradas como las más predictivas para el desarrollo general, o el aprendizaje de la lectura y/o las matemáticas. Mientras las pruebas de desarrollo están bien vinculadas con los programas que enfatizan las habilidades básicas, las pruebas de despistaje hacen un "escaneo" con materiales de rápida y sencilla aplicación, y con suficiente sensibilidad para detectar posibles problemas (4).

Actualmente, los instrumentos en uso para el despistaje del desarrollo psicomotor en el Perú son pocos, y es predominante el uso del TEPSI (Test de Desarrollo Psicomotor) (1) y el EEDP (Escala De Evaluación del Desarrollo Psicomotor) (5). Sobre estos instrumentos, apenas se han publicados estudios sobre sus propiedades de medición (es decir, validez y confiabilidad), y es preocupante que estos instrumentos formen parte de la batería estándar seleccionada por el MINSA y no sean respaldadas por evidencias científicas publicadas en muestras o revistas peruanas.

Una consideración para el uso de estos instrumentos en el sector público (por ejemplo el TEPSI), es que las condiciones laborales y administrativas del sector público de salud pueden limitar la aplicación de herramientas de evaluación del desarrollo debido a la carga de tareas administrativas. En consecuencia, se podría requerir un instrumento cuyo valor práctico se distinga por su capacidad para ser manejado eficientemente por el personal de salud, y que cumpla los requerimientos técnicos. Dado la carencia y/o las limitadas evidencias de medición de las pruebas de desarrollo psicomotor disponibles en Latinoamérica (particularmente en el Perú), y el reducido tiempo para aplicar una batería completa a un gran número de niños, se hace necesario no solo re-evaluar sus prioridades psicométricas, sino también crear o adaptar instrumentos clínicamente sensibles y culturalmente relevantes, y que cubran las recomendaciones de las mejores prácticas sobre su elegibilidad y características cualitativas y psicométricas en pruebas de despistaje de habilidades pre-académicas (4) y de las pruebas cognitivas en general (6, 7).

El objetivo de la presente investigación es evaluar

exploratoriamente las propiedades métricas del Inventario Preescolar de Despistaje de Minneapolis (Minneapolis Preschool Screening Inventory: MPSI) (8), el mismo que es una herramienta de evaluación de especial importancia en situaciones en que el tiempo, la urgencia de atención masiva y la cantidad de materiales, sean limitaciones importantes para la implementación de evaluaciones del desarrollo psicomotor. El MPSI está dirigido para niños de 3 a 6 años, aplicable para la detección de problemas de desarrollo psicomotor y del comportamiento social; sus áreas de evaluación incluyen: Motricidad Fina, Lenguaje, Cognición y Motricidad Gruesa. Es breve de administrar, barato y suficientemente fácil para aprender la administración y calificación (8) en profesionales (psicólogos, educadores, enfermeros) y para-profesionales (voluntarios y técnicos entrenados).

Nuestra investigación usará dos criterios: un criterio interno, en que se explorará las relaciones internas y las características estadísticas de los ítems; y un criterio externo basado en las relaciones con una prueba de despistaje del desarrollo psicomotor conocida en Latinoamérica, el Test de Desarrollo Psicomotor (TEPSI). Estos procedimientos analíticos de apoyarán en la Teoría Clásica de los Test (9), y que sustentará la obtención de la información métrica del MPSI.

MATERIAL Y MÉTODOS

Los participantes fueron 52 niños (52% varones), distribuidos en tres niveles de instrucción: 3 años (22,6%), 4 años (37,7%) y 5 años (37,7%). Todos estaban matriculados en uno de los Programas No Escolarizados de Educación Preescolar (PRONOEI), ubicados en el Distrito de Chorrillos al sur de Lima Metropolitana. El PRONOEI tiene la tarea de favorecer a los niños de familias económicamente pobres con aprendizajes pre-académicos mediante la estimulación de sus capacidades (10). Varios de estos programas están distribuidos en la zona urbana y semi-urbana del distrito, y cada programa contiene varios locales en que realizan las clases preescolares. Los locales preescolares de nuestros participantes se ubican en la zona de la población peri-urbana, que es frecuentemente de un nivel socioeconómico bajo o medio bajo. Cada aula cuenta con una Promotora que posee estudios técnicos en educación, y quien organiza y conduce las clases que varían entre 10 y 20 niños, aproximadamente. Los 4 locales desde los que provienen los niños fueron: N. (34,6%), I. (21,2%), BP (17,3%) y PM (26,9%); estos se en comunidades urbanas colindantes entre si y cada uno de estos contienen tres ambientes en los que distribuyen a los niños entre las edades.

La gran mayoría de los niños y sus familias vive en los alrededores de los PRONOEI; y han nacido en Lima, pero algunos de sus padres provienen de la Sierra del Perú, desde los que emigraron a Lima. El nivel educativo de las madres varía entre el nivel primaria y secundaria completa; y el rango de edades de los padres de los niños al momento de la evaluación esta entre los 26 a 30 años. La condición civil actual de la gran mayoría de las madres es la convivencia sin matrimonio oficializado; su ocupación principal se reparte entre las labores hogareñas y actividades económicas independientes. El ingreso económico se encuentra alrededor de los 500 soles mensuales. Generalmente, los niños y sus padres tienden a convivir con otros miembros de la familia materna o paterna, y por este motivo, los niños reciben actividades socializadoras de los padres y los demás familiares.

El Instrumento de Despistaje Preescolar de Minneapolis (MPSI) (8), es un instrumento que evalúa las habilidades del desarrollo psicomotor del niño relacionados con la escolaridad, creado en Estados Unidos. Las áreas principales de evaluación incluyen: Motricidad fina, Lenguaje, Cognitivo, Motricidad Gruesa. Está conformado por 50 ítems, que se califican usando un simple sistema de puntuación: «1» si es correcta y «0» si es incorrecta. Los ítems presentados tienen once subpruebas: Construcción, Copiado de Formas (7 ítems), Información (6 ítems), Emparejamiento (4 ítems), Sentencias, Salto y Balanceo (9 ítems), Nombrar Colores (5 ítems), Conteo (5 ítems), Preposiciones (4 ítems), Identificación de Partes del Cuerpo (5 ítems) y Repetición de Oraciones (5 ítems). La confiabilidad (consistencia interna) reportada en el manual ha demostrado ser alta. Originalmente, el instrumento experimentó cinco revisiones en etapas diversas, y fue revisada por psicólogos, educadores especiales y clínicos, profesores preescolares, de lenguaje demostrando su validez de contenido y su relevancia con criterios de rendimiento preescolar. El tiempo de administración aproximadamente de 12 a 15 minutos.

Se tradujeron las instrucciones de aplicación de la prueba, así como los ítems verbales del instrumento. Ambos elementos fueron sencillos de traducir, pues contienen expresiones directas sobre las acciones que deben hacer el examinador y el niño examinado. Sin embargo, se corroboró la exactitud de la traducción mediante el juicio de un psicólogo bilingüe, con experiencia en evaluaciones de niños. Hubo casi un completo acuerdo en su evaluación; y las discrepancias fueron resueltas discutiendo sobre las mejores opciones. La adaptación lingüística no enfatizó la traducción literal, sino la adecuación del contenido verbal a español hablado en Perú.

Se hicieron aplicaciones piloto a 10 niños de diferentes áreas urbanas y no urbanas, para obtener evidencias iniciales sobre lo apropiado de los ítems. Se observó que algunos ítems verbales (en la subescala preposiciones) y no verbales (en la subescala Partes del Cuerpo) no parecían adecuarse al nivel de dificultad que requería la prueba. Por lo tanto, se crearon ítem experimental para ser incluidos en el presente estudio. Para la subescala Partes del Cuerpo, se crearon los ítems Frente, Cintura y Hombro y en la subescala preposiciones, se crearon los ítems Lejos de la caja.

La recolección de datos se desarrolló en el semestre académico normal y en el horario regular de clases. Para la aplicación del instrumento se coordinó con una coordinadora de los PRONOEI de Chorrillos, definiendo e informado las fases del proyecto, los objetivos y las autorizaciones requeridas a los padres. Las evaluaciones se hicieron en un aula de ambiente cerrado, iluminado y alejado de ruidos extremos.

La administración de la prueba se hizo por un profesional de enfermería y psicología, previamente entrenados en la administración y calificación de la prueba.

El instrumento se aplicó individualmente, con un tiempo de duración máxima de 15 minutos. Para controlar la calidad de la información recolectada, se eliminaron aquellas resultados en que los niños que mostraron inconformidad y poca disponibilidad con la prueba.

Respecto a los análisis, se obtendrán evidencias de validez y consistencia interna de los ítems y de los puntajes del MPSI. Para evaluar la validez, primero se enfocará en la estructura interna del contenido, explorando las relaciones estadísticas entre los elementos y el puntaje que se derivan de ellas. Para ello, primero se hará un análisis de la dificultad y la discriminación de cada ítem; este último respecto a su subescala y al puntaje total del MPSI. También se estimará la homogeneidad de las respuestas mediante la correlación inter-ítem promedio.

El análisis de ítems se realizará con los estadísticos de dificultad y discriminación de los ítems. Algunos autores (11, 12) señalan que índices de discriminación aún de 0.20 son cuantitativamente aceptables en algunas circunstancias, pero las recomendaciones habituales indican que 0.30 es un nivel mínimamente aceptable (9, 13). Para propósitos descriptivos usaremos ambos criterios en una separación relevante para tomar decisiones sobre la validez de los ítems, tomando en cuenta la naturaleza exploratoria

del estudio. La dificultad del ítem se determinó por el porcentaje de examinados quienes responden correctamente al ítem. La evaluación de los niveles apropiados de dificultad usó el criterio de (11), en que se considera óptima una dificultad promedio de 0.50 y una distribución ideal de ítems entre 0.15 y 0.85, pero ya que las pruebas de despistaje se indican para detectar a niños con futuros problemas, entonces los ítems deberían tener un nivel de dificultad cercano a la tasa de selección deseada (11). Por lo tanto, como una prueba de despistaje generalmente tiene el propósito de ayudar a identificar el 10% más bajo de la población (14), los ítems deberían ser respondidos correctamente hasta el 90% de la muestra. Los ítems respondidos con una frecuencia mayor al 90% añaden poco valor discriminativo, tanto como los ítems incorrectamente respondidos por menos del 40% de la población (15). Fuera del rango de dificultad entre 0.40 y 0.90, los ítems podrían ser cuestionables para propósitos de despistaje (15). La discriminación de los ítems se evaluará por la correlación del ítem con su propia subescala y con el puntaje de la escala total (suma de todos los ítems): correlación ítem-test (Ritc). Se aplicará una corrección para controlar que estas correlaciones aumenten solo por el ítem se incluye en el puntaje (9).

El cálculo de la consistencia interna usará el coeficiente alfa de Cronbach (16, 17); su cálculo para las subpruebas usará todos sus ítems, pero para el puntaje de la prueba total, se evitará inflar la confiabilidad debido al número de ítems (8), y usamos los puntajes de las subáreas para el cálculo del coeficiente alfa. Adicionalmente, para la estimación de la consistencia interna del puntaje total, usaremos el coeficiente alfa estratificado (18). Este coeficiente incluye en su cálculo la variabilidad de las subpruebas, de tal modo que ajusta el cálculo con la varianza propia de cada elemento (prueba) independientemente (19).

El segundo enfoque será explorar las relaciones lineales con un criterio externo que converge en la misma finalidad de evaluación (16) del IDPM: el TEPSI. Esta correlación servirá para evaluar la covariación entre ambos instrumentos.

RESULTADOS

Dificultad de los ítems. La dificultad de los ítems (Tabla 1) se concentró dentro del rango 0.40 – 0.90 (80%), segregándose entre un 18.2% y 1.8% de ítems difíciles y fáciles, respectivamente. Los ítems más difíciles provinieron de la sección Partes del Cuerpo, y fueron considerados para su remoción. Los ítems experimentales adicionados a esta sección fueron moderadamente difíciles y

dentro del rango recomendado para ser incluidos en el instrumento.

Los ítems de Información fueron los más variables respecto a su frecuencia de respuesta, y en moderada magnitud, también ocurrió esto en Completamiento de Sentencias, Preposiciones y Repetición de Sentencias; estos

últimos comparten más relaciones conceptuales con el área del lenguaje. Por otro lado, los ítems menos variables fueron aquellos que comparten comunalidad con la habilidad visomotora, como Coordinación, Copiado, y Emparejamiento.

Discriminación de los ítems. En todos los análisis de dis-

Tabla 1. Índices de dificultad para la muestra total, de los ítems con su cluster

	N	%		N	%
Coordinación			Conteo		
Cubos			Conteo 2	38	73.1
Puente	47	90.4	Conteo 4	31	59.6
Torre 3	44	84.6	Conteo 3	29	55.8
Puente 2	43	82.7	Cuenta - 3	41	78.8
Torre 6	31	59.6	Cuenta - 7	35	67.3
Copiado			Preposiciones		
Copiar O	42	80.8	Arriba	18	34.6
Copiar +	45	86.5	Dentro	40	76.9
Copiar E	35	67.3	Debajo	27	51.9
Emparejamiento			Lado	20	38.5
Circulo	45	86.5	Partes del cuerpo		
T al revés	44	84.6	Rodilla	36	69.2
Letra M	44	84.6	Pulgar	12	23.1
V al revés	41	78.8	Menton	4	7.7
Información			Codo	15	28.8
Pies	38	73.1	Tobillo	11	21.2
Ojos	43	82.7	Frente (Prueba)	35	67.3
Ladra	26	50.0	Cintura (Prueba)	29	55.8
Ventana	15	28.8	Hombro (Prueba)	20	38.5
Grande	41	78.8	Repetición sentencias		
Oídos	28	53.8	Niña grande	36	69.2
Completamiento de sentencias			Siéntate piso	19	36.5
Niño - Niña	38	73.1	El gato a casa	30	57.7
Suben - Baja	27	51.9	Jugar en lluvia	35	67.3
Cuadrado - Redonda	30	57.7	Come su comida	29	55.8
Rápido - Lento	18	34.6			
Vuela - Nada	22	42.3			
Salto y balance					
Salta 1 pie	33	63.5			
Salta 2 pie	42	80.8			
Balance 1 pie	31	59.6			
Balance 2 pie	24	46.2			
Colores					
Verde	34	65.4			
Anaranjado	38	73.1			
Negro	38	73.1			
Amarillo	38	73.1			
Blanco	39	75.0			
Azul (Prueba)	37	71.2			
Rojo (Prueba)	42	80.8			

criminación, la correlación ítem-test corregido (Ritc) se elevaron cuando los ítems problemáticos se eliminaron del cálculo del puntaje del clúster. La siguiente descripción explica el procedimiento seguido. En los ítems del grupo Preposiciones, el ítem dentro correlacionó negativamente con su subescala (-0.17), y ello produjo un alfa de Cronbach de 0.18; pero su eliminación el coeficiente alfa a 0.42.

Respecto a los ítems del grupo Partes del cuerpo, la consistencia interna varió hasta su actual valor, pero su valor inicial fue 0.36, en sé que incluían todos los ítems originales ($r_{ii} = 0.09$). Se hizo una eliminación secuencial de los ítems iniciando con los ítems constantemente bajos en su Ritc. Esta eliminación secuencial se efectuó ya que la remoción de un ítem modificaba el Ritc de los demás ítem

Tabla 2. Índices de discriminación de los ítems con su cluster y con el puntaje total

	Rcluster	Rtotal		Rcluster	Rtotal
Coordinación			Conteo		
Cubos			Conteo 2	0,38	0,35
Puente	0,66	0,29	Conteo 4	0,41	0,49
Torre 3	0,66	0,40	Conteo 3	0,42	0,55
Puente 2	0,62	0,49	Cuenta - 3	0,35	0,28
Torre 6	0,44	0,66	Cuenta - 7	0,39	0,53
Copiado			Preposiciones		
Copiar O	0,31	0,25	Arriba	0,20	0,34
Copiar +	0,70	0,51	Debajo	0,23	0,15
Copiar E	0,57	0,57	Lado	0,38	0,24
Emparejamiento			Partes del cuerpo		
Circulo	0,02	0,28	Rodilla	0,63	0,31
T al revés	0,60	0,57	Codo	0,57	0,36
Letra M	0,69	0,38	Tobillo	0,26	0,16
V al revés	0,64	0,51	Frente(Prueba)	0,58	0,39
Información			Hombro(Prueba)	0,27	0,43
Pies	0,36	0,42	Repetición sentencia		
Ojos	0,32	0,39	Niña grande	0,27	0,39
Ladra	0,30	0,20	Siéntate piso	0,20	0,24
Ventana	0,49	0,54	El gato a casa	0,27	0,37
Grande	0,48	0,49	Jugar en lluvia	0,44	0,48
Oídos	0,45	0,62	Come su comida	0,26	0,39
Completamiento de sentencias					
Niño - Niña	0,28	0,29			
Suben - Baja	0,24	0,30			
Cuadrado - Redonda	0,46	0,58			
Rápido - Lento	0,66	0,56			
Vuela - Nada	0,44	0,36			
Salto y balance					
Salta 1 pie	0,45	0,62			
Salta 2 pie	0,44	0,41			
Balance 1 pie	0,51	0,43			
Balance 2 pie	0,24	0,33			
Colores					
Verde	0,60	0,52			
Anaranjado	0,79	0,51			
Negro	0,65	0,41			
Amarillo	0,69	0,48			
Blanco	0,77	0,57			
Azul (Prueba)	0,65	0,55			
Rojo (Prueba)	0,52	0,22			

restantes. De este modo, en la subescala Partes del Cuerpo la remoción de los ítems mentón ($R_{tc} = 0.02$) ni pulgar ($R_{itc} = -0.01$) produjeron un alfa de 0.64. Inmediatamente, la inclusión de los nuevos ítems experimentales (Frente, Cintura, Hombro) estabilizó estos estadísticos hacia una mayor discriminación. Sin embargo, El ítem Cintura constantemente permaneció en magnitud baja ($R_{itc} = 0.09$), así que finalmente fue removido. Los ítems que quedaron en esta subescala fueron, por lo tanto, Rodilla, Codo, Tobillo, Frente y Hombro.

En general, se puede observar que los índices de discriminación (Tabla 2), tendieron a ser ligeramente elevados entre los ítems con su propio agrupamiento que comparado con el puntaje total. Excepto un ítem con discriminación cero, el resto supera los límites recomendados para identificar ítems con adecuada discriminación. Los ítems alrededor de 0.25 pueden también cumplir con el criterio de discriminación por efecto del redondeo.

Consistencia interna.

La consistencia interna del puntaje total fue mayor a 0.80 y cercano a 0.90 por efectos del redondeo; estos niveles son muy apropiados y elevados para una prueba de

despistaje cuyas tareas son heterogéneas (Tabla 3). Tomando como criterios algunas de las recomendaciones de la literatura de medición (11, 19, 20), la consistencia del puntaje total del MPSI es aceptable para fines clínicos y serviría como ayuda para tomar de decisiones sobre el niño evaluado. La consistencia interna de las subescalas ha sido variable. El tamaño de los coeficientes alfa estuvo directamente relacionado con el número de ítems en cada subescala (8), observándose que los coeficientes tendieron a ser mayores cuando el número de ítems era mayor. Se observa también que los coeficientes alfa fueron influenciados también por la homogeneidad de los ítems (correlación inter-ítem.), que fue en general apropiada para habilidades de amplio espectro como en Repetición de Sentencias, Preposiciones, Información y Conteo. Los ítems de estas áreas señalan cierta heterogeneidad, pero aún son aceptables para asumir que internamente aportan a la medición de una habilidad (19, 21).

Casi todas las correlaciones entre el IDPM y el TEPSI fueron estadísticamente significativas en el nivel 0.001 (Tabla 4), lo que sugiere que las varianzas compartidas entre los componentes de las dos pruebas van más allá de

Tabla 3. Consistencia Interna y Homogeneidad de los clusters de ítems

	N ítems	Alfa	Rii
IDPM			
Coordinación total	7	0,81	0,41
Cubos	4	0,73	0,44
Copiado	3	0,58	0,32
Información	6	0,67	0,25
Emparejamiento	4	0,68	0,33
Completamiento de Sentencias	5	0,66	0,28
Salto y balance	4	0,62	0,30
Colores	7	0,88	0,51
Conteo	5	0,64	0,26
Preposiciones	3	0,42	0,20
Repetir sentencias	5	0,52	0,18
Partes del cuerpo	5	0,70	0,30
Total	50	0,86	0,44
TEPSI			
Coordinación	16	0,80	0,20
Lenguaje	22	0,77	0,15
Motricidad	12	0,68	0,14
Puntaje total	50	0,78	0,59

cero. Estas relaciones sistemáticas son esperadas, ya que las subpruebas del IDPM abordan conductas de amplio espectro tal como lo hacen las subescalas del TEPSI. El puntaje total del IDPM esta fuerte y linealmente asociado al área de Lenguaje y Coordinación, y más aún con el puntaje total del TEPSI. La varianza compartida es elevada entre ambos puntajes totales, y sugiere que la posición obtenida ordinal por el niño en el TEPSI puede ser replicada en el IDPM. Las correlaciones cercanas a cero con el género del niño indican que una sola norma de interpretación estadística podría construirse basados en las distribución de puntajes de la muestra total. Por otro lado, la covariación entre el puntaje total del IDPM con la edad es fuerte y moderadamente diferente que con el puntaje total del TEPSI, pero la diferencia entre estos coeficientes de validez no es estadísticamente diferente, $t(49) = -1.45, p > 0.05$.

El patrón de correlaciones con el TEPSI indica cierto traslape de las áreas de Coordinación ojo-mano y Lenguaje con los grupos de ítems del IDPM. Pero también se observan divergencias esperadas debido al contenido de los ítems del IDPM. Por ejemplo, Información, Completamiento de Sentencias, Colores, tienden a converger con el área de Lenguaje del TEPSI, pero la distancia de estas correlaciones no es tan diferente cuando se las compara con Coordinación del TEPSI. Este patrón tiende a ser más observable en las otras relaciones, ya que las diferencias de la validez con Coordinación y Lenguaje tiende a ser imperceptible.

En la Tabla 5 se muestran los estadísticos descriptivos básicos para generar normas de interpretación. Aunque no fue probada cuantitativamente, se observa que todos los puntajes avanzan en un patrón predecible desde los 3

Tabla 4. Correlaciones Pearson con el TEPSI y variables demográficas (Edad y Sexo)

	TEPSI				Edad	Sexo
	Coord.	Leng.	Motric.	Total		
Repetir Sentencias	0,566	0,544	0,413	0,601	0,499	-0,186
Coordinación	0,575	0,572	0,485	0,638	0,421	-0,020
Información	0,454	0,576	0,537	0,608	0,561	-0,112
Emparejamiento	0,312*	0,502	0,330*	0,461	0,407	-0,097
Completamiento de sentencias	0,580	0,667	0,359	0,656	0,404	-0,147
Salto y Balanceo	0,470	0,527	0,576	0,598	0,531	-0,107
Colores	0,672	0,592	0,291*	0,635	0,530	-0,081
Conteo	0,686	0,680	0,482	0,735	0,643	-0,178
Preposiciones	0,300*	0,422	0,423	0,440	0,411	0,029
Partes de Cuerpo(3)	0,419	0,415	0,241 ns	0,435	0,390	-0,157
Total	0,760	0,810	0,590**	0,860	0,700	-0,155

* $p < 0,05$

** $p < 0,01$

a : Todas las correlaciones sin asteriscos, son significativas en $p < 0,001$ o menos

ns : $p > 0,05$,

a los 5 años. Sin embargo, las diferencias entre el grupo de 4 y 5 años son mayores en el IDPM que entre los mismos grupos en el puntaje total del TEPSI. Aparentemente, el IDPM puede ser más sensible a los cambios en las diferentes edades, pues contiene aprendizaje que son fuertemente influenciados por los programas preescolares.

DISCUSIÓN

El presente estudio inició la búsqueda de evidencias para validar una herramienta para la evaluación del desarrollo psicomotor potencialmente aplicable en el Perú. Se evaluaron aspectos relacionados con el funcionamiento esta-

Tabla 5: Estadísticos descriptivos para las pruebas IDPM y TEPSI.

Edad	IDPM		TEPSI							
			Coordinación		Lenguaje		Motricidad		Total	
	M	DE	M	DE	M	DE	M	DE	M	DE
3	20,25	7,32	7,42	2,843	9,92	3,872	6,33	1,826	23,67	6,243
4	32,40	8,58	11,40	2,703	16,60	2,854	8,05	1,761	36,05	5,853
5	39,85	6,05	13,35	1,496	18,10	2,049	9,65	1,843	41,10	3,447
Total	32,46	10,44	11,23	3,240	15,63	4,275	8,27	2,188	35,13	8,416

dístico de los ítems y de los puntajes derivados de estos ítems. Por ejemplo, la homogeneidad de los ítems mostró que las subescalas están relacionados en áreas que reflejan una aparente unidimensionalidad en cada una de ellas, especialmente para los ítems de los grupos de Coordinación, Colores y el puntaje Total. Los ítems de otras subescalas, como repetir sentencias, tuvo un comportamiento estadístico heterogéneo debido a su baja correlación inter-ítem. En general, siguiendo algunas pautas interpretativas (16, 21), los ítems cubren un amplio rango de homogeneidad y apuntan a áreas de contenido de moderado espectro.

La capacidad discriminativa de cada ítem también se midió con un índice clásico, y casi la totalidad de los ítems demostraron que pueden diferenciar entre niños de elevado desempeño frente a los de bajo desempeño. Ya que el puntaje total es la suma de todos los ítems, los ítems mostraron menor discriminación con el puntaje total que con sus respectivas subescalas, pero aun así estas diferencias no fueron severas como para determinar que los ítems no apoyan en la evaluación de una solamedida del desempeño psicomotor, es decir, el puntaje Total. Además de los índices de discriminación, la consistencia interna también aportó evidencia favorable de la validez interna del IDPM. La magnitud alcanzada por el coeficiente de consistencia interna indica que el contenido de los ítems ha permanecido unificado y que la varianza de error es moderadamente pequeña.

Para estimar el coeficiente alfa del puntaje total, se utilizaron las agrupaciones de ítems en lugar de los ítems individuales, debido que el número de ítems inflaría la magnitud del coeficiente Cronbach(18). Las diferencias en la consistencia interna de los puntajes totales entre el TEPSI y el IDPM estuvieron influenciadas por el número de ítems. Esta característica infla el coeficiente alfa aun cuando las correlaciones entre los ítems son bajas Cronbach(18); sin embargo, examinando los componentes del puntaje total

del IDPM, la consistencia interna de cada uno es aceptable considerando que están compuestos por un número pequeño de ítems y que, en general, su homogeneidad tiende a ser moderadamente alta. Por lo tanto, a nivel de los ítems y a nivel de los agrupamientos de ítems, el IDPM ha mostrado ser internamente homogénea ya que sus partes componentes están interrelacionadas justificando la expresión del desempeño del niño en un solo puntaje. Aunque la unidimensionalidad no se ha evaluado directamente por métodos más apropiados y refinados como el análisis factorial (16, 22), la consistencia interna puede darnos un acercamiento a esta cualidad psicométrica.

Los resultados del presente estudio pueden ser evaluados como favorables para el IDPM, y el alcance de este instrumento. El IDPM puede ser potencialmente incorporado en los cursos de desarrollo del niño, en la investigación científica y en la práctica profesional, especialmente cuando las demandas de rapidez, eficiencia y bajo costo son factores decisivos para implementar programas de intervención preventiva. Su administración requiere menos tiempo de evaluación y entrenamiento, y las tareas son bien aceptadas por el niño.

A diferencia del TEPSI, el IDPM da únicamente un puntaje total que sirve para tomar decisiones sobre el niño en términos de derivación a evaluación. Este aspecto ha sido intencionalmente planificado por los autores del IDPM, ya que el IDPM no tiene por objetivo hacer un diagnóstico o identificar problemas establecidos en los niños. Y esto es congruente con la finalidad de los instrumentos de despistaje (4, 23, 24), Aunque el TEPSI se diseñó para esta misma finalidad, la interpretación de los puntajes permite establecer etiquetas diagnósticas que influenciarán al examinador poco entrenado. Las siguientes acciones sobre el niño podrían ser influenciadas por esta indebida identificación, ya que el examinador puede considerar que no es necesaria una segunda evaluación a profundidad pues el niño ha sido identificado como «retraso». Esto evita

que el niño sea evaluado por otros instrumentos más finos como lenguaje, etc.

Otra ventaja del IDPM es que ofrece materiales adicionales para efectuar un seguimiento del desempeño del niño basado en la observación de la profesora. Estos materiales no formaron de esta investigación preliminar, pero su exploración forma parte de futuros objetivos a esta investigación.

Durante la evaluación, se pueden implementar otros criterios descriptivos de la conducta de examen del niño, ya que las conductas durante el proceso evaluativo influyen en la magnitud del puntaje (25). Como herramienta única de detección de problemas de desarrollo, el IDPM parece ser eficiente, pero se requieren más datos que repliquen los presentes resultados. Una combinación con otras medidas de despistaje basado en otras fuentes de información como los padres (12, 26), la observación de juego espontáneo o estructurado (27) y profesores, puede incrementar la eficacia diagnóstica.

Los datos de validez original del IDPM son favorables sobre su predictividad del rendimiento escolar en los primeros años de escolaridad, y más aún si los ítems están intencionalmente vinculados con el futuro rendimiento escolar, como ocurre con la más reciente versión revisada del IDPM (8). Instrumentos de despistaje preescolar diseñados para ser sensibles a problemas pre-académicos están involucrados en las evaluaciones de despistaje de problemas en el desarrollo (28), y hay evidencias satisfactorias sobre su poder predictivo aún en un rango de 5 a 7 años después de la evaluación (9). Pero esta aparente ventaja debe ser validada empíricamente cuando se intenta adaptar un instrumento de un contexto cultural a otro.

Aún en instrumentos de despistaje diseñados dentro de un fuerte programa de validación y construcción de sus componentes, hay problemas de exactitud diagnóstica que varían desde severa (por ejemplo el Denver-II (29) a moderada inexactitud (por ejemplo, el inventario Battelle (29), que no provienen del error del muestreo o la confiabilidad de sus puntajes, sino que requieren modificación en sus métodos de calificación o de su estructura interna (29, 30). En el estudio, las adaptaciones realizadas resultaron en un mejoramiento de las propiedades internas de los ítems y, consecuentemente, del puntaje total. En contraste, los ítems poco discriminativos no contribuyen a la validez interna ni a la confiabilidad, y esta última propiedad métrica pondrá un límite a los coeficientes de validez obtenidos de la aplicación de los puntajes.

REFERENCIAS BIBLIOGRÁFICAS

1. Haeussler I, Marchant T. Test de Desarrollo Psicomotor. 1a ed. Santiago: Universidad Católica; 1985.
2. MINSA. Norma Técnica de Salud para el Control del crecimiento y desarrollo de la Niña y Niño Menor de Cinco Años, Resolución Ministerial N° 990-2010/MINSA; 2010.
3. Emmons MR, Alfonso VC. Critical review of the technical characteristics of current preschool screening batteries. *Journal of Psychoeducational Assessment*. 2005;23(2):111-27.
4. Hasbrouck JE, Tindal GA, Marston D. *Preschool Assessment Classroom-based Assessment: Testing for Teachers*. Columbus OHIO: Merrill; 1990. 273-91 p.
5. Rodríguez S, Arancibia V, Undurraga C. *Escala de Evaluación del Desarrollo Psicomotor de 0 a 24 meses*. Santiago de Chile: Editorial Galdoc; 1985.
6. Flanagan DP, Nutall EV, Romero I, Kalesnik J. *Assessment of cognitive functioning in preschoolers. Assessing and screening preschoolers*. 2a ed. Allyn & Bacon: New York; 1999. 186 – 217 p.
7. Alfonso VC, Flanagan DP. *Best practices in the use of the Stanford-Binet Intelligence Scales. Fifth Edition (SB5) with preschoolers*. Bracken BA, Nagle R (Eds.). 4a ed. Mahwah NJ: Erlbaum: *Psychoeducational assessment of preschool children*; 2006. P. 267-95.
8. Lichtenstein R. *Minneapolis Preschool Screening Inventory*. Minneapolis, MN. Minneapolis Public Schools. 1982;38(8):12-5.
9. Nunnally JC, Bernstein IJ. *Teoría psicométrica*. México, D. F: McGraw-Hill; 1995.
10. Merino C, Díaz M, Zapata L, Benites L. *School psychology in Peru*. In: Jimerson SR, Oakland T, Farell PT (Eds). *The Handbook International of School Psychology*. Oakland: Sage; 2006. p. 299-307.
11. Anastasi A, Urbina S. *Psychological testing*. 7a ed. Upper Saddle River, NJ: Prentice-Hall; 1997.
12. Glascoe FP, Byrne KE. The usefulness of the Battelle Developmental Inventory Screening Test. *Clinic Pediatrics*. 1993;32(5):273-80.
13. Stagnitti K, Unsworth C, Rodger S. Development of an assessment to identify play behaviours that discriminate between the play of typical preschoolers and preschoolers with pre-academic problems. *Canadian Journal of Occupational Therapy*. 2000;67(5):291-303.
14. Meisels SJ, Marsden DB, Wiske MS, Henderson LW. *Early Screening Inventory – Revised. Examiner’s manual*. Ann Arbor. Michigan: Rebus, Inc; 1997.

15. Parush S, Yochman A, Jessel AS, Shapiro M, Mazor-Karsenty T. Construct validity of the Miller assessment for preschoolers and the pediatric examination of educational readiness for children. *Physical Occupational Therapy in Pediatrics*. 2002;22(2):7-27.
16. Clark LA, Watson D. Constructing validity: Basic issues in scale development. *Psychological Assessment* 1995;7(3):309-19.
17. Cronbach LJ. Coefficient alpha and the internal structure of test. *Psychometrika*. 1951;16:297-334.
18. Cronbach LJ, Schoneman P, Mckie D. Alpha coefficient for stratified-parallel test. *Educational and Psychological Measurement*. 1995;25:291- 312.
19. Charter RA. A breakdown of reliability coefficients by test type and reliability method, and clinical implications of low reliability. *The Journal of General Psychology*. 2003;130(3):290-304.
20. Chicchetti DV, Sparrow SA. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. *American Journal of Mental Deficiency*. 1981;86(2):309 - 19.
21. Briggs S, Cheek JM. The role of factor analysis in the development and evaluation of personality scales. *Journal of Personality*. 1986;54:106-48.
22. Hattie JA. Methodology review: Assessing unidimensionality of tests and items. *Applied Psychological Measurement*. 1985;9:139-64.
23. Perrin J EG, Kaminer RK, Lacamera R, Nackashi JA, Poncher JR, Randall V, Wachtel RC, Ziring PR. Screening infants and young children for developmental disabilities. *American Academy of Pediatrics*. 1994;93(5):863-5.
24. Mathews J, Shames G, Wiig E. The professions of speech-language pathology and audiology. *Human communication disorders*. 2a ed. Columbus, OH: Charles E. Merrill Publishing; 1986. P. 3-26.
25. Glutting J, Oakland T. GATSB: Guide to the assessment of test session behavior. Orlando: The Psychological Corporation; 1993.
26. Wake M, Garner B, Gallagher S. Does parents' evaluation of developmental status at school entry predict language, achievement, and quality of life 2 years later. *Ambulatory Pediatrics*. 2005;5(3):143-9.
27. Simmos JO. Fluharty Preschool Speech and Language Test: Analysis of construct validity. *Journal of Speech and Hearing Disorders*. 1988;5:168-74.
28. Parush S, Winokur M, Goldstand S, Miller LJ. Prediction of school performance using the Miller Assessment for Preschoolers (MAP): A validity study. *American Journal of Occupational Therapy*. 2002;6(5):547-55.
29. Glascoe FP, Byrne KE, Ashford LG, Johnson KL, Chang B, Strickland B. Accuracy of the Denver-II in developmental screening. *Pediatrics*. 1992;59(6):1221-5.
30. Glascoe FP, Sandler H. . Value of parents' estimates of children's developmental ages. *Journal of Pediatr*. 1995;127(5):831-5.

Correspondencia

Marisol Angulo Ramos
Calle Enrique Palacios 430, Chorrillos, Lima 9
Correo electrónico: noa_c22@yahoo.es

Forma de citar este artículo: Angulo-Ramos M, Merino-Soto C. Propiedades de medición del Inventario de Despistaje Preescolar de Minneapolis (IDPM). *Rev enferm Herediana*. 2012;5(2):114-124.