

Tamizaje del desarrollo infantil: evaluación comparativa de tres instrumentos.

Francisco A. Leal Soto, Andrea A. Daudet Miranda, Paola M. Frugone Opazo, Pamela L. Silva Carvajal.

Artículo publicado en
Revista Iberoamericana de Diagnóstico y Evaluación Psicológica
Vol. 5, 1998.

Palabras clave
Desarrollo Infantil; Tamizaje (Screening); Evaluación de instrumentos

Resumen

Se evaluó la validez de tamizaje del *Denver Developmental Screening Test* (DDST, Frankenburg y Dodds, 1970), el Test de Desarrollo Psicomotor (TEPSI, Haeussler y Marchant, 1987) y el Inventario del Desarrollo Infantil (IDI, Ireton, 1992), en una muestra de 62 preescolares entre 3 años 0 meses y 4 años, 5 meses y 29 días, tomada de Jardines Infantiles públicos y privados de la ciudad de Arica, Chile. Como criterio se utilizó la Escala de Inteligencia de Stanford-Binet, en la revisión de Terman y Merrill, edición española (1975). Los resultados indican que ninguno de los instrumentos logra una validez de tamizaje óptima, aunque hay dos predictores que logran resultados aceptables, tanto desde el punto de vista de la sensibilidad como de la especificidad: la escala de Desarrollo Global del IDI y el Perfil del TEPSI. Los demás predictores fracasan en identificar los niños que deberían ser referidos por posible retraso en el desarrollo en porcentajes que hacen poco recomendable su utilización con fines de tamizaje. En particular, preocupa la escasa sensibilidad del DDST, ya que es utilizado masivamente en los jardines infantiles públicos con este fin, lo que, a la luz de estos resultados, debería ser revisado.

Screening of child development: a comparative assessment of three test

Abstract

Screening validity of the Denver Developmental Screening Test (DDST, Frankenburg & Dodds, 1970), the *Test de Desarrollo Psicomotor* (TEPSI, Haeussler & Marchant, 1987) and the Child Development Inventory (Ireton, 1992) was assessed in 62 preschool children ranging 3 years 0 month to 4 year, 5 month, 29 days, selected of public and private nursery of the city of Arica, Chile. The criterion was the Stanford-Binet Intelligence Scale, Terman-Merril revision, in the Spanish edition (1975). Results point that none of the three assessed tests obtains optimal levels of screening validity, but two of them obtains satisfactory results, even from sensitivity and specificity: the General Development scale of the IDI and the Profile of the TEPSI. Other predictors fails in recognize the children who would be referred by his/her possible developmental delay, in a number of cases that made its use not recommended with screening purpose. Small sensitivity of the DDST is specially alarming, by its massive use with this purpose in the public nursery, wich, as a result of this research, would seem convenient to review.

Tamizaje do desenvolvimento: avaliação comparativa de três instrumentos.

Resumo

Avaliou-se a validade de tamizaje do *Denver Development Screening Test* (DDST, Frankenburg y

Dodds, 1970), o *Test de Desarrollo Psicomotor* (TEPSI, Haussler y Marchant, 1987) e o *Inventario del Desarrollo Infantil* (IDI, Ireton, 1992) numa amostra de 62 crianças pre-escolares de 3 anos 0 mês a 4 anos, 5 meses e 29 dias, extraída das creches do governo e particulares da cidade de Arica, Chile. Utilizou-se como critério a Escala da Inteligência de Stanford-Binet, na revisão de Terman e Merrill, edição espanhola (1975). Os resultados mostram que nenhum dos instrumentos consegue a validade de tamizaje ótima, embora dois preditores tem alcançado resultados aceitáveis, tanto do ponto de vista da sensibilidade, como da especificidade: A escala do Desenvolvimento Global do IDI e o Perfil do TEPSI. Os demais preditores fracassam em identificar as crianças que deveriam ser referidos por possível atraso no desenvolvimento, em porcentagens que fazem pouco recomendável sua utilização com fins de tamizaje. Em particular preocupa a pouca sensibilidade do DDST, já que é utilizado massivamente nas creches públicas com este fim, o que, a luz destes resultados, deveria ser revisado.

El desarrollo infantil es preocupación de todos los gobiernos latinoamericanos, ya que se ha tomado clara conciencia del efecto de la intervención temprana en la mejoría de la calidad de vida de la población. Esta preocupación se ha acentuado y formalizado luego de la suscripción del Compromiso de Atención a la Infancia, y toma muy diversas formas en los distintos países, desde la atención en salud hasta la educación y la recreación. Una de tales formas es la evaluación sistemática y masiva del desarrollo infantil, en la forma de tamizaje (screening), con el objeto de detectar aquellos niños que requieren una atención preferencial, permitiendo focalizar los recursos hacia ellos y facilitando una intervención oportuna que pueda evitar o minimizar las consecuencias del retraso y/o las alteraciones en el desarrollo.

Hacerlo es un imperativo desde múltiples perspectivas: moral, ya que todos los niños tienen derecho al desarrollo pleno, por lo tanto la detección oportuna de sus alteraciones no es más que una extensión o condición de este derecho; social, ya que las condiciones de vida que afectan negativamente el desarrollo infantil se presentan con mayor fuerza y frecuencia en los sectores económicamente deprimidos, por lo que la detección oportuna del retraso permitiría intervenir a tiempo para corregir parcial o totalmente la injusticia generada a partir de tal desigualdad social; económicas, porque la intervención temprana implica un considerable ahorro en atención posterior a problemas de desarrollo, aprendizaje y fracaso escolar, entre otros; y científico-técnicas, ya que las intervenciones tempranas son de menor complejidad y mayor eficacia que las intervenciones posteriores (Lira, 1992).

La evaluación del desarrollo psicológico ha cobrado en Chile mayor importancia cada vez, ya que se ha cumplido la mayor parte de las metas referidas a la salud física de los niños: la mortalidad infantil ha sido controlada hasta límites bastante buenos, y la mayor parte de los niños es controlado en su desarrollo físico y su salud regularmente, disminuyéndose la desnutrición y la morbilidad (Lira, 1992-1993); en este contexto, las metas en salud infantil se amplían, y, al decir de M. I. Lira, "la meta en salud infantil no es, obviamente, que los niños no se mueran, se desnutran menos o se enfermen menos; la meta es el bienestar físico, psíquico y social del niño" (Lira, 1992-1993, pág. 16).

Concordantemente, uno de los objetivos del actual Gobierno de Chile es aumentar la cobertura en la educación preescolar, la que actualmente se brinda a través de varios sistemas públicos y privados, siendo el más importante la Junta Nacional de Jardines Infantiles (JUNJI), organismo estatal que tiene la mayor red de salas-cuna y jardines infantiles a lo largo del país, en los que atiende a una gran cantidad de niños entre 0 y 5 años. En todos estos jardines y salas-cuna se hace seguimiento del desarrollo de cada uno de los niños atendidos, con el objeto de identificar aquellos cuyo desarrollo no corresponde al esperado y programar intervenciones oportunas; con ese fin, se administra periódicamente una técnica de tamizaje, la que es aplicada por las propias

educadoras de párvulos a cada niño atendido en estos centros (Junta Nacional de Jardines Infantiles, 1996). Igual cosa se hace en los consultorios de salud, en que se controla periódicamente al niño desde el nacimiento hasta los cinco años (Ministerio de Salud, 1979).

El tamizaje es una técnica que permite identificar un posible problema o patología no evidente mediante la aplicación de exámenes u otros procedimientos que pueden ser aplicados rápidamente (Medina, 1987); el tamizaje no pretende ser un diagnóstico, sino una mera identificación de un posible problema, que debe ser evaluado en profundidad una vez que ha sido identificado como tal. El tamizaje es una forma de prevención, ya que su principal propósito es identificar la presencia potencial de alteraciones del desarrollo antes de lo que usualmente serían reconocidas sin la aplicación del tamizaje. Sin embargo, el tamizaje del desarrollo infantil no es sólo la detección de la eventual alteración; a juicio de Lessler (1972, cit. por Stangler, Huber y Routh, 1980), un buen tamizaje del desarrollo debe proveer información relevante acerca de las tareas de vida de los menores, que pueda ser utilizada para direccionar medidas de apoyo, y eventualmente conducir a evaluaciones y/o medidas preventivas adecuadas. De este modo, el tamizaje no termina con la detección de un eventual problema, sino que debe conducir a una intervención temprana adecuada, como una evaluación en profundidad, un apoyo directo a los padres, la inclusión en un programa de estimulación, u otra medida preventiva o correctiva.

El esfuerzo desplegado es grande, si se considera que se evalúa periódicamente a cada niño, para lo cual se utiliza, en el caso de la JUNJI, el Test de Desarrollo de Denver (Denver Developmental Screening Test, DDST, Frankenburg y Dodds, 1970), cuya aplicación toma alrededor de 15 minutos; y sus resultados incluyen la categorización del niño en una de tres categorías: normal, retraso o dudoso, además de su nivel de desarrollo en cuatro áreas: motricidad gruesa, motricidad fina, lenguaje y adaptación personal-social. Pese a la magnitud del esfuerzo, se carece de evaluaciones acerca de la bondad de este test como instrumento de tamizaje en Chile, y la información recogida en otros contextos es poco alentadora, presentándose dudas principalmente acerca de su capacidad para detectar a los niños que presentan riesgo o retraso en su desarrollo: estudios realizados en los Estados Unidos de Norteamérica indican que tendría una relativamente alta especificidad -capacidad para identificar a niños normales como tales- y una baja sensibilidad -capacidad para identificar a niños con riesgo o retraso como tales- al mismo tiempo que reporta índices altos de falsos negativos, que oscilan entre 50 y 94% (Glascoe y Borowitz, 1968; Moriarty, 1972, ambos en Carrasco, Cavour y Vergara, 1995).

En el caso de los consultorios de salud, se utiliza dos instrumentos: la Escala de Evaluación del Desarrollo Psicomotor (EEDP, Rodríguez, Arancibia y Undurraga, 1987), para los niños entre 0 y 2 años, y el Test de Desarrollo Psicomotor (TEPSI, Haeussler y Marchant, 1988), para los niños entre 2 y 5 años. Ambas son escalas desarrolladas en Chile y con normas chilenas, y evalúan el desarrollo en cuatro y tres áreas respectivamente: Coordinación, Social, Lenguaje y Motora, en el caso de la EEDP, y Coordinación, Lenguaje y Motricidad en el caso del TEPSI. En ninguno de estos casos, se dispone de evaluaciones de su validez de tamizaje, aunque en el caso de la EEDP, su uso como tal se encuentra tan validado que se está trabajando en técnicas de tamizaje aún más breves utilizándola como criterio (Lira, 1992-1993).

El presente trabajo tuvo como objetivo evaluar la validez de tamizaje del DDST y del TEPSI, así como de un instrumento de tamizaje alternativo a éstos, el Inventario del Desarrollo Infantil (Ireton, 1992), a fin de contribuir a una valoración del esfuerzo realizado y a una reorientación del mismo si es necesario.

Método

Participantes

Se incluyó a 62 preescolares de entre 3 años 0 meses y 4 años 5 meses 30 días, de los cuales 33 fueron niñas y 29 niños. 32 de ellos asistían a jardines infantiles dependientes de la JUNJI, lo que implica un nivel socioeconómico bajo y medio-bajo; y 30 asistían a jardines particulares, lo que implica un nivel socioeconómico medio-alto y alto. Para tomar la muestra, se definió inicialmente los jardines infantiles, y los preescolares fueron luego seleccionados al azar de entre aquellos que, asistiendo a los jardines previamente escogidos, cumplían con el requisito de edad preestablecido.

Instrumentos

Se utilizó tres instrumentos como predictores: el Denver Developmental Screening Test (DDST), de Frankenburg y Dodds (1970), en una versión adaptada en Chile (Himmel y cols., 1990; Himmel y cols., 1997); el Test de Desarrollo Psicomotor de 2 a 5 años (TEPSI, Haeussler y Marchant, 1980), instrumento desarrollado en Chile, que evalúa el desarrollo en las áreas de Motricidad, Coordinación y Lenguaje, entregando puntuaciones típicas por rango de edad para cada área y para el resultado global; y el Inventario del Desarrollo Infantil (IDI, Ireton, 1992), en una adaptación realizada en Chile (Leal, Daudet y Williams, 1994; Carrasco, Cavour y Vergara, 1994), aunque con las normas originales (Ireton, 1992). Este es un cuestionario de 300 preguntas, autoadministrado por los padres, que evalúa 8 áreas de desarrollo entre 1 y 5 años: Desarrollo social, Autoayuda, Motricidad gruesa, Motricidad fina, Lenguaje expresivo, Lenguaje comprensivo, Letras y Números, además de una escala resumida, la escala de Desarrollo Global, compuesta por los 70 ítems más discriminativos del inventario. En cada una de ellas, se obtiene una Edad de Desarrollo, a partir de las cuales puede obtenerse un Perfil del desarrollo del niño. Requiere que los padres tengan una escolaridad mínima de 8 años para responder en forma consistente (Donaire y Segovia, 1996).

Como criterio se utilizó la Escala de Inteligencia de Stanford-Binet, en la adaptación de Terman-Merrill, edición española (1975), que permite obtener un indicador de Edad Mental global, a partir del cual se obtiene un Coeficiente Intelectual; este instrumento es ampliamente usado en el país como instrumento de diagnóstico.

Procedimiento

Las evaluaciones fueron efectuadas por cuatro evaluadoras, estudiantes de último año de la carrera de Psicología, entrenadas en el uso de los instrumentos, entre Octubre y Diciembre de 1994. Cada niño fue evaluado con los 4 instrumentos en un período máximo de 3 semanas. En el caso de los jardines infantiles dependientes de la JUNJI, la aplicación del DDST se hace regularmente dos veces al año por el mismo personal del jardín, por lo que se tomó los resultados de esa aplicación, para evitar la repetición de la evaluación. El IDI fue entregado a los padres en una reunión, en la que se les informaba el objetivo del estudio, se les requería su colaboración y se les entregaba el cuestionario y las instrucciones para completarlo, solicitando su devolución en un plazo definido.

Una vez terminada la aplicación de los instrumentos, se procedió a su revisión, descartándose algunos protocolos por falta de correspondencia con los rangos de edad predefinidos o por incompletos, lo que significó que se descartaran 6 protocolos del IDI y 1 del DDST.

Se obtuvo los resultados en cada uno de los instrumentos; y a partir de dichos resultados, se clasificó a cada niño según los siguientes indicadores para cada test:

Terman-Merrill: Se clasificó en Normal o Anormal, según estuvieran por sobre o bajo el punto de corte de CI, el que se estableció en una desviación estándar bajo el promedio (CI=84). Se definió este punto

de corte, ya que generalmente se refiere a programas de estimulación a los niños que obtienen resultados por debajo de ese nivel en las pruebas de desarrollo.

DDST: Se clasificó en Referido y No referido, asumiendo como Referidos aquellos casos que fueron categorizados como Riesgo o Dudoso según el instrumento.

TEPSI: Se clasificó en Referido o No referido, según dos criterios:

Puntaje Escala Global (TEPSI-G): Se consideró Referidos los casos que obtuvieron un puntaje inferior al promedio en más de una desviación estándar ($T < 40$) en esta escala.

Perfil (TEPSI-P): Se consideró los resultados de las subescalas, considerándose Referidos aquellos casos que obtuvieron puntuación T inferior a 40 en una cualquiera de las tres subescalas.

IDI: Se clasificó en Referido o No referido, según 5 criterios:

Perfil, heterogéneos no referidos (IDI-HNR): Se consideró Referidos a los sujetos cuyos perfiles mostraban una cualquiera de las escalas en un nivel considerado de Riesgo o Retraso, es decir, con una Edad de Desarrollo bajo el 25% respecto a su edad cronológica. Los casos clasificados como Heterogéneos (es decir, que mostraban una diferencia de más de 50% de la edad cronológica entre algunas de sus escalas) no fueron incluidos en la clasificación Referidos.

Perfil, heterogéneos referidos (IDI-HR): Se consideró Referidos a los sujetos cuyos perfiles mostraban una cualquiera de las escalas en un nivel considerado de Riesgo o Retraso, es decir, con una Edad de Desarrollo bajo el 25% respecto a su edad cronológica. Los casos clasificados como Heterogéneos fueron incluidos en la clasificación Referidos.

Escala de Desarrollo Global, 25% (IDI-25): Se consideró Referidos a los sujetos que tuvieron una Edad de Desarrollo bajo el 25% de su edad cronológica, en la Escala de Desarrollo Global.

Escala de Desarrollo Global, 20% (IDI-20): Se consideró Referidos a los sujetos que tuvieron una Edad de Desarrollo bajo el 20% de su edad cronológica, en la Escala de Desarrollo Global.

Escala de Desarrollo Global, 15% (IDI-15): Se consideró Referidos a los sujetos que tuvieron una Edad de Desarrollo bajo el 15% de su edad cronológica, en la Escala de Desarrollo Global.

Cada uno de estos indicadores fue considerado un predictor, y se construyó para cada uno la matriz de cuatro casilleros para obtener los indicadores de validez de tamizaje en relación al criterio, el resultado en el Terman-Merrill. La matriz, tomada de Stangler, Huber y Routh (1980), se puede ver en la Tabla NE 1.

Los indicadores de validez de tamizaje calculados en base a esta matriz, según lo propuesto por Stangler, Huber y Routh (1980), son los siguientes:

1. Validez de Tamizaje. Indica el porcentaje de acuerdo entre el predictor y el criterio. Corresponde a $100(a+d)/(a+b+c+d)$.

2. Sensitividad. Es el porcentaje de casos anormales correctamente referidos por el predictor. Corresponde a $100(a)/(a+c)$.

3. Especificidad. Es el porcentaje de casos normales correctamente identificados por el predictor.

Corresponde a $100(d)/(b+d)$.

4. Tasa de sobrerreferencia en el número de referidos. Es el porcentaje de casos normales que fueron referidos por el predictor, en relación al total de casos referidos. Corresponde a $100(b)/(a+b)$.
5. Tasa de sobrerreferencia en la muestra total. Es el porcentaje de casos normales que fueron referidos por el predictor, en relación al total de casos tamizados. Corresponde a $100(b)/(a+b+c+d)$.
6. Validez de un test negativo. Es el porcentaje de casos normales correctamente no referidos por el predictor, en relación al total de casos no referidos. Corresponde a $100(d)/(c+d)$.
7. Tasa de subreferencia en el número de no referidos. Es el porcentaje de casos anormales que no fueron referidos por el predictor, en relación al total de casos referidos. Corresponde a $100(c)/(c+d)$.
8. Tasa de subreferencia en la muestra total. Es el porcentaje de casos anormales que no fueron referidos por el predictor, en relación al total de casos tamizados. Corresponde a $100(c)/(a+b+c+d)$.
9. Validez de un test positivo. Es el porcentaje de casos anormales correctamente referidos por el predictor, en relación al total de casos referidos. Corresponde a $100(a)/(a+b)$.
10. Tasa global de referencia. Es el porcentaje de casos referidos por el predictor, independientemente del acierto de la referencia en relación al criterio. Corresponde a $100(a+b)/(a+b+c+d)$.

Debe notarse que los indicadores 4 y 9, así como 6 y 7, son complementarios entre sí; se calculan y explicitan ambos por el diferente énfasis interpretativo que proporcionan.

Resultados

Los resultados se presentan en las Tablas NE 2 y NE 3. En la primera, se ha resumido las matrices de cada predictor en relación al criterio; en la segunda, se han resumido los indicadores obtenidos para cada predictor.

Discusión

Un test de tamizaje óptimo muestra una alta validez de tamizaje a partir de alta sensibilidad y alta especificidad, lo que se traduce en baja cantidad de sub y sobrerreferencias; es decir, al identificar correctamente tanto los casos anormales como los normales, es un test seguro al mismo tiempo que económico. Seguro, en el sentido de no dejar sin atención a casos efectivamente anormales, y económico, al no referir a diagnóstico innecesariamente a los casos efectivamente normales. Como se aprecia en los resultados, ninguno de los predictores utilizados satisface este criterio cabalmente, ya que, si bien varios predictores alcanzan una validez de tamizaje al menos satisfactoria (sobre 70%), en general lo hacen privilegiando o la sensibilidad o la especificidad, lo que se traduce en tasas altas de sub y sobrerreferencias.

En un extremo, se encuentra el caso del predictor TEPSE-G, que logra la más alta validez de tamizaje (82%): lo hace principalmente por una alta especificidad (92%), por lo que refiere incorrectamente a diagnóstico y/o tratamiento sólo al 6% del grupo total; sin embargo, su sensibilidad es apenas de 42%, que implica un 58% de falsos negativos en relación a los casos efectivamente anormales. Esto se traduce en un 11% de subreferencias sobre el total de tamizados, esto es, de cada 100 niños tamizados, 11 que deberían recibir diagnóstico y/o tratamiento dejan de recibirlo por no ser identificados. Es decir, es un predictor económico, pero poco seguro. Algo similar ocurre con los

predictores IDI-25, IDI-20 y DDST. Este último merece una observación especial, por su extraordinariamente baja sensibilidad, que se traduce en la mayor tasa de subreferencias (18% sobre el total de tamizados), que concuerda con los resultados obtenidos con el mismo instrumento en los Estados Unidos, en que se ha reportado sensibilidades que oscilan entre 50 y 6% (Glascoe y Borowitz, 1968; Moriarty, 1972, ambos en Carrasco, Cavour y Vergara, 1995). Esto resulta particularmente grave si se considera que es el instrumento usado regularmente por la JUNJI para su tamizaje; significa que casi un quinto de los niños tamizados no son referidos a diagnóstico y/o tratamiento, debiendo efectivamente recibirlo.

El extremo opuesto lo representa el predictor IDI-HR, que tiene una sensibilidad muy alta (90%), por lo que sólo el 2% de los tamizados deja de ser referido cuando lo necesita; sin embargo, la especificidad es sólo del 39%, por lo que el éxito en detectar los casos anormales se ve acompañado con una tasa de sobrerreferencias que equivale al 50% del total de los tamizados; es decir, es una alternativa segura, pero muy poco económica. Algo similar ocurre con el predictor IDI-HNR.

Entre ambos extremos, se sitúan los predictores TEPSI-P e IDI-15; siendo moderada su validez de tamizaje (76 y 71%, respectivamente), logran un mejor equilibrio entre la sensibilidad y la especificidad, produciendo por tanto menores tasas de sub y sobrerreferencias; en el caso de las subreferencias, se ubican alrededor del 5% sobre el total de tamizados, lo que coincide con el criterio sugerido por Stangler, Huber y Routh (1980); y en el caso de las sobrerreferencias, son del orden del 20% sobre el total de tamizados. Consecuentemente, la validez del resultado negativo es bastante alta (91 y 92%, respectivamente), pero la validez del resultado positivo es sólo aceptable (39 y 35% en cada caso). En definitiva, son predictores bastante seguros, aunque representan todavía un costo considerable en términos de diagnósticos y/o tratamientos innecesarios.

Dado que la prevalencia del trastorno es una de las variables que afecta significativamente la validez de un test de tamizaje -siendo ésta mejor en la medida que la prevalencia aumenta (Stangler, Huber y Routh, 1980)- es importante que los estudios de validez de tamizaje se realicen en muestras efectivamente representativas de la población en la que efectivamente se aplicará el tamizaje. En el presente caso, la muestra se obtuvo en condiciones naturales y sin exclusión de casos; la prevalencia observada en este estudio, 12%, es bastante esperable en una muestra tomada de una población normal, considerando que para el criterio se estableció un punto de corte situado una desviación estándar por debajo de la media, ya que la probabilidad asociada a dicho punto en una distribución normal ideal corresponde a 15,8%. Por esta razón, a pesar del tamaño relativamente pequeño de la muestra, estos resultados pueden ser considerados un buen indicador del comportamiento de estos predictores en condiciones de tamizaje masivo.

De acuerdo a Stangler, Huber y Routh (1980), los instrumentos de tamizaje deberían cumplir, además de las características psicométricas básicas de confiabilidad, validez y normas apropiadas, con los criterios de: aceptabilidad para los niños y sus padres, para los profesionales que utilizarán la información y para la comunidad en general; simplicidad en su aprendizaje, administración y puntuación, material requerido y capacitación del personal necesario para su administración; aplicabilidad, en el sentido de que las características y capacidades del instrumento sean consistentes con la prevalencia y características del trastorno a detectar, así como con las características de la población; y costo, en el sentido de que la aplicación del instrumento, incluyendo los materiales, la capacitación, el tiempo de los administradores y de los propios evaluados y sus familias, los costos de los falsos positivos y cualquier otro costo directo o indirecto de la aplicación del tamizaje, sean significativamente inferiores a los costos económicos o de pérdida de bienestar que significa la detección tardía o la no detección del problema tamizado. Estos criterios, a excepción del costo, son cumplidos en distinto grado por los predictores evaluados en este estudio. En general, son fácilmente aceptados por los padres y los niños, y tienen una buena recepción en la comunidad profesional

vinculada al desarrollo infantil; la excepción la constituye el IDI en su versión completa, a partir de la cual se obtienen los predictores IDI-HNR e IDI-HR, ya que por la gran cantidad de ítemes que incluye (recuérdese que está compuesto por 300 reactivos) puede ser tedioso o agotador para los padres, quienes pueden mostrarse algo reticentes a completarlo. Todos son simples tanto en su administración como en su interpretación, y requieren un mínimo de capacitación para su adecuado uso; en el caso del IDI, además de las competencias de los administradores del tamizaje, se requiere un mínimo de 8 años de escolaridad de la madre o el padre que completa el cuestionario, lo que en algunos casos podría ser una limitante, aunque en Chile la población general tiene una escolaridad media algo superior a 8 años, y, en todo caso, puede ser aplicado como entrevista cuando los padres no tienen la escolaridad requerida, como recomienda el autor del test (Ireton, 1992). Por otra parte, las características de construcción de todos los instrumentos, basados en fuertes teorías sobre el desarrollo infantil, así como su pertinencia en relación al problema-objetivo del tamizaje, es decir, el retraso en el desarrollo, los hacen igualmente aplicables.

El IDI, al requerir una participación activa de los padres, puede producir en ellos expectativas favorables que influyen en el desarrollo posterior de sus hijos, así como un mayor grado de compromiso con el desarrollo del niño, y, eventualmente, con un programa de tratamiento si fuera necesario, como han señalado varios autores (Brothy, Good y Nedler, 1975; Eisert y cols., 1980; Galey y cols., 1983; Ireton, Kampen y Lun, 1981); además, permite acceder a información acerca de conductas a las que los evaluadores no tienen acceso directo, como el comportamiento del niño en el hogar o en otros contextos distintos del consultorio o la institución educacional, ampliando y haciendo más estable el espectro de conductas evaluadas, razón por la cual la potencialidad del informe de los padres como evaluación de sus hijos ha sido destacada ampliamente (Anastasiow, 1981; Guerin y Gottfried, 1987; Ullman y Kausch, 1979). Estas características constituirían ventajas adicionales en relación a los demás instrumentos considerados, ventajas que podrían verse comprometidas por el exceso de reactivos y la consecuente eventual respuesta negativa de algunos padres; pero éstas se minimizan en los predictores que consideran la escala de Desarrollo Global, compuesta sólo por 70 ítemes, en los que, no obstante, se mantiene la característica de involucrar directamente a los padres en la observación de sus hijos, con el posterior mayor compromiso con su desarrollo, y el acceso a información de observaciones realizadas en distintos ámbitos de la vida del niño.

En relación al costo, los instrumentos difieren notablemente en sus costos de aplicación. El TEPSI toma aproximadamente 40 minutos, y debe ser aplicado íntegramente por un evaluador entrenado; el DDST requiere alrededor de 15 minutos, y no requiere la administración completa, ya que tiene criterios de línea base y techo o suspensión; y el IDI sólo requiere tiempo profesional para su corrección y puntuación, la que toma entre 5 y 10 minutos en el caso de aplicación completa, y no más de 3 minutos en el caso de aplicarse sólo la escala de Desarrollo Global.

Los predictores DDST, TEPSI-G, IDI-25 e IDI-20, por su alta especificidad y baja sensibilidad, resultan económicos en el corto plazo, ya que exigen una baja cantidad de diagnósticos y/o tratamientos; pero pueden resultar en un alto costo al largo plazo, por el efecto de diagnósticos tardíos y mayores costos en tratamientos y/o consecuencias de mayor gravedad para los casos no detectados.

Al contrario, los predictores IDI-HNR e IDI-HR, por su alta sensibilidad y baja especificidad, generan un alto costo asociado a la gran cantidad de diagnósticos y/o tratamientos innecesarios, aunque disminuyen el costo al largo plazo por la disminución de los efectos del retardo en el desarrollo y los tratamientos especializados a que puede dar lugar su detección tardía. A esto se suma el mayor esfuerzo exigido a los padres, al tener que completar el inventario en su forma íntegra.

Como ya se señaló, los predictores que equilibran de mejor manera sensibilidad y especificidad, TEPSI-P e IDI-15, son los que generan también el mejor equilibrio entre costos de corto

plazo asociados a las necesidades innecesarias de diagnóstico y/o tratamiento generadas por los falsos positivos, y de largo plazo, asociados a las consecuencias de un diagnóstico y/o tratamiento tardíos del retardo en el desarrollo generados por los falsos negativos.

Considerando estos criterios conjuntamente, los predictores IDI-15 y TEPSI-P, sin ser todo lo buenos que sería deseable, parecen ser los de mejor comportamiento, lo que apoyaría su selección como eventuales instrumentos de tamizaje del desarrollo infantil. Esto confirma la utilización del TEPSI realizada en los servicios de salud a través del programa de atención conocido como "control del niño sano", ya que el criterio de referencia es similar al utilizado en este estudio: un niño que resulta con cualquier área de desarrollo en el nivel de "riesgo" o "retraso", según el instrumento (lo que corresponde a una y dos desviaciones estándar bajo el promedio, respectivamente), es incorporado a un programa y reevaluado en un período breve. No obstante, por su ligeramente mejor sensibilidad y por las ventajas adicionales del IDI-15, en relación al menor tiempo y costo de administración, así como por requerir la participación activa de los padres, sería útil considerar seriamente la posibilidad de utilizar la escala de Desarrollo Global del IDI, para lo cual sería conveniente elaborar normas apropiadas (lo que ya se está haciendo, a través de un proyecto de investigación en curso); sin embargo, la utilización de las normas norteamericanas con el criterio del 15% por bajo la edad cronológica para referir, ha mostrado ser suficientemente útil como para sugerir desde ya su uso.

Desgraciadamente, los resultados del presente estudio confirman los hallazgos reportados en estudios anteriores realizados en los Estados Unidos de Norteamérica en relación al DDST, por lo que, a pesar de su simplicidad y aceptabilidad, su uso como instrumento de tamizaje resulta claramente desaconsejable; a la luz de estos resultados coincidentes, parecería que el gran esfuerzo desplegado por la JUNJI, evaluando sistemática y periódicamente a los niños que son atendidos en los establecimientos de su dependencia, podría no estar cumpliendo con su principal objetivo de identificar aquellos niños que requieren una atención especial para asegurar su desarrollo pleno.

Referencias

Anastasiow, J. (1981) "Early childhood education for the handicapped in the 1980's: Recommendations". Exceptional Children, 47 (4), 276-282.

Brothy, J., Good, T. y Nedler, S. (1975) Teaching in the preschool. New York: Harper & Row Publishers.

Carrasco, J., Cavour, P. y Vergara, M. (1995) Estudio de confiabilidad y validez de la traducción local del Inventario del Desarrollo Infantil de Harold Ireton (IDI). Arica, Chile: Universidad de Tarapacá, Seminario para optar al Título de Psicólogo y al grado de Licenciado en Psicología.

Donaire, D. y Segovia, M. (1996) Efecto del nivel de escolaridad del respondiente sobre las características psicométricas del IDI. Arica, Chile: Universidad de Tarapacá, Seminario para optar al Título de Psicólogo y al grado de Licenciado en Psicología.

Eysert, D., Spector, S., Shankaran, S., Faingenbaum, D. y Szego, E. (1980) "Mother's reports of their low birth weight infant's subsequent development on the Minnesota Child Development Inventory". Journal of Pediatric Psychology, 5 (4), 353-364.

Frankenburg, W. K. y Dodds, J. (1970) Denver Developmental Screening Test. Denver, CO: Ladoca Project and Publishing Foundation, Inc.

Galey, G., Macy, D., Schoen, J. y Solomon, G. (1983) "The Debt Project: Early intervention for

handicapped children and their parents". Exceptional Children, 49 (5), 447-448.

Guerin, D. y Gottfried, A. (1987) "Minnesota Child Development Inventories: Predictors of intelligence, achievement, and adaptability". Journal of Pediatric Psychology, 12 (4), 595-609.

Haeussler, I. M. y Marchant, T. (1988) TEPSI: Test de desarrollo psicomotor 2-5 años. Santiago de Chile: Ediciones Universidad Católica de Chile.

Himmel, E., Alvarez-Salamanca, E., Díaz, C. y Villalón, M. (1990) Manual Denver Developmental Screening Test. Santiago de Chile: Universidad Católica de Chile.

Himmel, E., Alvarez-Salamanca, E., Díaz, C. y Villalón, M. (1997) Adaptación y estandarización del Denver Developmental Screening Test (DDST), para la población preescolar chilena. En Scagliotti, B. y Pinto, A. (Eds.) (1997) Dificultades del aprendizaje. Avances en psicopedagogía. Santiago de Chile: Colección Monografías en Educación, Pontificia Universidad Católica de Chile, Facultad de Educación.

Ireton, H., Kampen, M. y Lun, K. (1981) "Minnesota Preschool Inventory identification of children at risk for kindergarten failure". Psychology in Schools, 18, 394-401.

Ireton, H. (1992) Child Development Inventory Manual. Minneapolis: Behavior Science Systems, Inc.

Junta Nacional de Jardines Infantiles, Depto. Técnico (1996) Líneas de Acción. Santiago de Chile.

Leal, F., Daudet, A. y Williams, E. E. (1994) Traducción, adaptación y evaluación del Child Development Inventory (CDI) en la ciudad de Arica. Arica, Chile: Universidad de Tarapacá, Informe Final proyecto 3739-93.

Lira, M. I. (1992) Construcción y evaluación de una técnica de tamizaje del retraso del desarrollo psicomotor (Primer año de vida). Santiago de Chile: Centro de estudios de desarrollo y estimulación psicosocial (CEDEP).

Lira, M. I. (1992-1993) "Construcción y evaluación de una técnica de tamizaje del retraso del desarrollo psicomotor (Primer año de vida)". Revista Chilena de Psicología, 13-14, 15-21.

Medina, E. (1987) Métodos epidemiológicos en Clínica y Salud Pública. Santiago de Chile: Escuela de Salud Pública, Universidad de Chile.

Ministerio de Salud de Chile (1979) Programa de salud del Niño y del Adolescente. Santiago de Chile.

Rodríguez, S., Arancibia, V. y Undurraga, C. (1987) Escala de evaluación del desarrollo psicomotor de 0-24 meses. Santiago de Chile: Galdoc.

Stangler, S., Huber, C. y Routh, D. (1980) Screening growth and development of Preschool Children: a guide for test selection. McGraw-Hill, New York.

Terman, L. y Merrill, M. (1975) Medida de la inteligencia: Método para el empleo de las pruebas de Stanford-Binet. Madrid: Espasa-Calpe, 9ª edición,

Ullman, D. G. y Kaush, D. F. (1979) "Early identification of developmental strengths and weaknesses in preschool children". Exceptional Children, 46 (1), 8-13.

Tabla N° 1
Matriz de cuatro casilleros para obtener los indicadores de Validez de Tamizaje

Predictor	Criterio	
	Anormal	Normal
Referido (Positivo)	a (correctamente referidos)	b (falsos positivos)
No referido (Negativo)	c (falsos negativos)	d (correctamente no referidos)

Tabla N° 2
Resumen de las matrices de cuatro casilleros para el cálculo de la Validez de Tamizaje para los ocho predictores

Predictor	Número de casos	a (correctamente referidos)	b (falsos positivos)	c (falsos negativos)	d (correctamente no referidos)
DDST	61	1	5	11	44
TEPSI-G	62	5	4	7	46
TEPSI-P	62	7	11	4	40
IDI-HNR	56	8	20	2	26
IDI-HR	56	9	28	1	18
IDI-25	56	1	2	9	44
IDI-20	56	3	5	7	41
IDI-15	56	7	13	3	33

Tabla N° 3

Indicadores de validez de tamizaje obtenidos por los predictores, siendo el criterio el Terman-Merril (Porcentajes)

Indicador	Predictor							
	DDST	TEPSI-G	TEPSI-P	IDI-HNR	IDI-HR	IDI-25	IDI-20	IDI-15
1. Validez de tamizaje	74	82	76	61	48	80	79	71
2. Sensitividad	08	42	64	80	90	10	30	70
3. Especificidad	90	92	78	57	39	96	89	72
4. Tasa de sobrerreferencias en el número de referidos	83	44	61	71	76	67	63	65
5. Tasa de sobrerreferencias en la muestra total	08	06	18	36	50	04	09	23
6. Validez de un test negativo	80	87	91	93	95	83	85	92
7. Tasa de subreferencias en el número de no referidos	20	13	09	07	05	17	15	08
8. Tasa de subreferencias en la muestra total	18	11	06	04	02	16	13	05
9. Validez de un test positivo	17	56	39	29	24	33	38	35
10. Tasa global de referencias	10	15	29	50	66	05	14	36