

CAPÍTULO 1

Las escalas de desarrollo

ESQUEMA

1. **¿Por qué necesitamos comprender el desarrollo?**
2. **¿Qué características deberían reunir las escalas de medida?**
3. **¿Responden las escalas disponibles a estas características?**
4. **¿Cómo podemos valorar técnicamente las escalas?**
5. **¿Cómo debería ser una escala de desarrollo?**

1. ¿POR QUÉ NECESITAMOS COMPRENDER EL DESARROLLO?

La psicología evolutiva ha buscado ofrecer, desde sus inicios, una teoría general sobre el proceso de desarrollo humano, tanto desde la perspectiva de los procesos implicados como de los estadios alcanzados, partiendo de la observación de casos individuales biográficos, como ya hicieran Tiedeman en el siglo XVIII o Preyer en el XIX (*El alma del niño*, 1882), considerado este último como el iniciador de la disciplina (Fernández y Gil, 1995). Preyer había recogido las tradiciones recapitulacionistas y embriologistas, mientras se sucedían las tesis lamarkistas y darwinistas, que trataban de explicar la evolución filogenética en una época en la que la psicología carecía de metodologías adecuadas a su objeto de estudio. Éstas se fueron desarrollando a lo largo del siglo XX, comenzando con el método psicogenético utilizado por Jean Piaget, quien dio un tratamiento empírico a la investigación epistemológica (Coll y Gillieron, 1995); y, desde los años 50, gracias al desarrollo de la computación, generadora de la llamada revolución cognitiva, que constituyó una alternativa viable al fracaso de los conductistas para dar cuenta de los procesos mentales de alto nivel. Como señala García Madruga (1991: 3):

Los años setenta y ochenta han visto la progresiva consolidación de los enfoques cognitivos del procesamiento de la información como la perspectiva teórica hegemónica dentro de la psicología evolutiva.

Neopiagetianos y sobre todo neoestructuralistas tratan de superar los problemas no resueltos aunando la capacidad explicativa de distintas disciplinas y métodos hasta desem-

bocar en la creación de una nueva ciencia cognitiva que, desde los años 80, se materializa en el *enfoque conexionista* y los modelos del *procesamiento distribuido en paralelo (PDP)*, capaz por su potencia simuladora de *los mecanismos que hay detrás del conocimiento humano* (Rumelhart, McClellan y otros, 1992:46) de atraer el consenso de la generalidad de los investigadores.

Si bien desde el principio la psicología evolutiva estaba impregnada de un aroma fisiológico, como se lee siguiendo a Damasio: «la frenología surgió como una curiosa mezcla de psicología temprana, neurociencia temprana y filosofía práctica» (Damasio, 2001:28), esto sirvió para distanciar las interpretaciones del desarrollo, visto por empiristas como un proceso determinístico, por racionalistas como una secuencia con grados de libertad, sobre la que el sujeto influye con su actividad, y otros insistieron en la necesidad de comprender al individuo dentro de su contexto histórico, pero todos acabaron encontrando los mismos problemas y, en las últimas décadas, aceptando una fructífera colaboración que puede acabar dando la razón a cada uno. Curiosamente, son aquellas aportaciones particulares que conservan el aroma fisiológico las que se han abierto camino colaborando en la constitución de la moderna ciencia cognitiva.

Desde hace siglos ha preocupado la comprensión de las causas del comportamiento humano, el pensamiento, el lenguaje y el aprendizaje de los niños. Se desarrollaron teorías sobre lo heredable y lo adquirido, sobre si el pensamiento necesita de las imágenes y sobre si el pensamiento es posible sin el lenguaje. Se intentó dar explicación a la evolución individual y a la de las especies. Se iniciaron estudios introspectivos y se construyeron laboratorios donde medir las distintas dimensiones del hombre. Todos partieron de la filosofía, y durante la segunda mitad del siglo XIX se produjo

la eclosión paradigmática que habría de escindir a nuestra joven ciencia de aquel tronco común. Nacieron en ese siglo figuras como Baldwin, Bolwby, Freud, Piaget, Vygotsky o Wallon, que hicieron las fecundas aportaciones conceptuales que han llevado a la constitución del principal cuerpo de conocimientos sobre la infancia del que disponemos en estos primeros años del siglo XXI.

Evidentemente ellos no contaron con las herramientas metodológicas que caracterizan hoy a la ciencia experimental, y sus respectivas miradas estaban condicionadas por las herencias filosóficas y los conocimientos científicos de la época. Pero contribuyeron entre todos a los cambios cuantitativos y cualitativos que, como señala García Madruga (1991), se han producido en las últimas décadas en la psicología evolutiva: nuevos temas de estudio como el desarrollo cognitivo y social previo a la adquisición del lenguaje; los notables progresos de las neurociencias, que han influido positivamente sobre las investigaciones de psicobiología evolutiva, y el aumento de la esperanza de vida que ha extendido el estudio psicológico a todo el ciclo vital. Dos nuevas perspectivas teóricas, el enfoque del procesamiento de la información y las teorías del paradigma dialéctico-contextual, han destacado asimismo la necesidad de abordar el estudio de los procesos evolutivos de forma inequívocamente científica, y en el segundo caso como un sistema integrado por múltiples variables. Esto también ha sido posible gracias a los avances metodológicos que utilizan nuevas tecnologías, como el vídeo o el ordenador, y a los nuevos procedimientos estadísticos y de diseño experimental.

El estudio de todos los fenómenos psicológicos se abordó desde una perspectiva evolutiva, ganando el enfoque del procesamiento de la información una posición hegemónica que

no ha perdido, pero que ha sufrido esperanzadoras metamorfosis al tener que enfrentarse a retos como la negación de su potencialidad para explicar el cambio cualitativo del desarrollo. Fodor (1986) lanzó este órdago a la psicología cognitiva y reclamó de nuevo un innatismo que negaba el propio desarrollo. Johnson-Laird (1990) explica este reto: los conceptos complejos no pueden construirse solo a partir de los más simples, o el todo no sería más que la simple suma de las partes. Algunos se resignan a una posición acumulativa (aditiva), pero otros, los llamados estructuralistas, buscan denodadamente mecanismos que aclaren cómo el desarrollo pasa por estadios cualitativa y gradualmente superiores.

Una línea de investigación procedente de Newell y Simon (1972) se centró en la construcción de programas que emulan el proceder de los expertos cuando se enfrentan a una tarea: son los sistemas de reglas de producción. Johnson-Laird (1990:164) aprecia que dichos sistemas son más parecidos a un lenguaje de programación que a teorías empíricas demostrables, pues proponen que las reglas que gobiernan el aprendizaje y la memoria son simbólicas:

...podría pensarse que es muy extraño que tales reglas, que tan fácilmente se parafrasean, estén bajo la superficie de nuestra vida mental. Si pudiésemos retirar la tapa de la mente y asomarnos dentro, ¿veríamos realmente un conjunto de principios dispuestos como en una constitución, capaces de cubrir cualquier eventualidad? Probablemente no.

Otra línea de trabajo más reciente ha planteado la revolucionaria idea de que la información puede estar representada de forma distribuida en muchas *unidades* de procesamiento, no requiriendo por lo tanto una memoria depositada en un lugar concreto; como veremos

más adelante, una arquitectura mental así constituida no necesita informarse con reglas explícitas, ya que es capaz de descubrir las. Se ha llamado *conexionistas* a los modelos que se suscriben a esta idea. Ellos consideran las unidades algo así como *neuronas idealizadas*; por eso se dice que sus redes distribuidas en paralelo son una metáfora del cerebro.

Sin embargo, y a pesar de las elegantes soluciones de las teorías conexionistas, la complejidad de los procesos estudiados ha puesto de manifiesto la necesidad de incluir en los diseños nuevas formas de representar y analizar el cambio, de explicar la variación individual. En palabras de Gutiérrez, Luque y García Madruga (2002:16): «lo que caracteriza los fenómenos estudiados no es tanto la regularidad y el orden de los procesos involucrados sino más bien todo lo contrario: la complejidad y el caos». Así, partiendo de la *teoría de los sistemas dinámicos* (Smith, 2003; Smith y Thelen, 2003; Spencer y Thelen, 2003; Thelen y Bates, 2003), se asumen múltiples modelos de cambio en el desarrollo, no necesariamente lineales, sino que admiten todo tipo de oscilaciones y fluctuaciones porque son el resultado *emergente* del funcionamiento interactivo y dinámico de los sistemas.

Entender cómo se produce el desarrollo nos permite adentrarnos en el control de las variables que lo facilitan o interfieren, de ahí la importancia de disponer de instrumentos válidos de medida que nos proporcionen indicadores del estado de los procesos y de los niveles alcanzados en cada momento, para decidir las estrategias de intervención tanto en la educación como en la clínica.

Las diversas concepciones del desarrollo han servido históricamente de soporte teórico a la construcción de escalas para su evaluación. Las actuales concepciones se interesan por el estudio microgenético de los fenómenos evo-

lutivos, pero será la perspectiva macrogenética la que permita al clínico acercarse al sujeto contemplándolo en su globalidad. Los tests de screening siguen este modelo y como recomiendan Glascoe y Saphiro (2002), todos los niños deberían ser evaluados con este tipo de herramienta en sus visitas pediátricas.

Las ventajas de estos instrumentos de cribado del desarrollo estriban en su facilidad y rapidez de aplicación, sin obligatoriedad de ser administrados por psicólogos especializados, sino por cualquier profesional que trabaje en estos ámbitos, contando incluso con la ayuda de los padres para resolver los cuestionarios; aunque si bien los datos que proporcionan no permiten realizar diagnósticos, sí son utilizados en atención primaria para después derivar el caso hacia el especialista que pueda caracterizar la alerta detectada utilizando instrumentos específicos (Gómez de Terreros y Valdés, 2005).

Nos proponemos revisar los principios que guiaron la construcción de las escalas más conocidas o utilizadas en el diagnóstico del desarrollo infantil para ver a qué necesidades dan respuesta desde su modelo y cuáles serían complementariamente las utilidades deseables que no satisfacen. Agradecemos su trabajo a todos aquellos estudiosos que, dirigidos por el profesor Francisco Secadas participaron en la confección de la Escala Observacional, de las Escalas de la Dra. McCarthy, del Inventario de Battelle, de las Escalas Bayley, a las Dras. Brunet y Lézine, y en general a cuantos nos han proporcionado modelos útiles a partir de los cuales podemos seguir evolucionando. Nuestras respuestas no van a superar sus esfuerzos, sino que se van a sumar a ellos, como es sabido que se produce el progreso en la ciencia.

2. ¿QUÉ CARACTERÍSTICAS DEBERÍAN REUNIR LAS ESCALAS DE MEDIDA?

Marchesi, Carretero y Palacios (1995) recogen en primer lugar la idea propuesta por Baltes (1973) sobre una *intervención enriquecedora*. Desde este punto de vista una escala ha de servir para garantizar el máximo nivel de desarrollo posible a cada individuo. La posibilidad de medir globalmente en cada momento el desarrollo alcanzado tendrá una repercusión preventiva, señalada por Caplan (1964), ya que desvela las desviaciones de funciones aisladas o del curso general del desarrollo. El hecho de disponer de un contraste válido y fiable estadísticamente, supone para el evaluador una guía de actuaciones enriquecedoras del desarrollo. El estimulador o el educador puede saber en cada momento cuál es la zona de desarrollo potencial –en el sentido de proximidad vygotskyana– para proponer el programa de condiciones educativas más adecuado para el despliegue de esas potencialidades, en los diferentes sistemas y /o contextos (Bronfenbrenner, 1980).

Esta utilidad enriquecedora es deseada y buscada permanentemente por los padres y educadores cuando se enfrentan al reto de la crianza y educación de los niños, y supone para los profesionales de la psicología una demanda a la que debemos responder. «La permanente posibilidad de optimización del desarrollo es, además de una realidad, un reto y una responsabilidad» (Marchesi, Carretero y Palacios, 1995:298).

En segundo lugar, y dado que los cambios evolutivos se producen dentro de la dialéctica entre la actividad del sujeto y las condiciones de su entorno ambiental, sería conveniente que las escalas evaluaran el comportamiento en los *múltiples sistemas* que interactúan entre sí, den-

tro de los que el individuo se desenvuelve: relaciones familiares, con extraños, conocidos, vivienda, parque, calle, barrio, colegio, etc. La escala debería recoger los hitos del desarrollo en el orden de consecución en que se producen, y debería ser el objeto de investigaciones a posteriori el interpretarlos y darles sentido dentro de una deseada teoría evolutiva general. Pero en principio, esos hitos, en tanto que hechos, deberían ser observables, para que resulten accesibles a las personas que puedan compartir con el niño la experiencia de cada uno de los entornos en que desarrolla su actividad.

Es por ello que, en tercer lugar, debemos aspirar a disponer de una escala que responda a una *metodología observacional* y no experimental, de manera que la cualificación de los usuarios podría no requerir el dominio de terminología y conceptos procedentes de diversas ciencias, ni la utilización de tecnologías sofisticadas ni el control de situaciones de prueba, que acaban restringiendo el acceso a los posibles usuarios. Sin duda, la deseable parsimonia del quehacer científico no debería producir menoscabo en la validez de la herramienta; y para evitarlo hay que introducir diversos controles, tales como la descripción pormenorizada de las conductas a evaluar, la fácil identificación de resultados extraños o la asistencia de una tecnología mecanizada de corrección e informe, bajo la supervisión de expertos. En definitiva, una escala debería gozar de la ventaja de servirse de situaciones y materiales que formen parte del entorno cotidiano del sujeto.

En cuarto lugar, una escala debe constituir una herramienta que proporcione una visión de unicidad y globalidad sobre el desarrollo. Esto se conseguiría utilizando una *única escala para todo el intervalo medido*; es decir, que incluya toda la etapa del desarrollo infantil. Una escala que recoja una descripción del desarro-

llo desde el nacimiento hasta la pubertad como una continuidad se enmarcaría dentro de la tradición organísmica (Bruner, Erikson, Freud, Piaget, Wallon, Werner), lo que supondría una *concepción estructural*, al incluir estadios o períodos con características diferenciales, que implican cambios cualitativos por transformación y no por simple suma.

Se deberá caracterizar cada uno de estos períodos a través de la interdependencia de funciones, que dan sentido a cada uno de ellos. El problema de la explicación de la transición de uno a otro no es el objetivo de ninguna escala, sino de la teoría sobre la que se sustenta el constructo medido. En cualquier caso, interpretaremos que el cambio cualitativo emerge de la interacción de las funciones, cuando éstas se alían cumplimentando el sentido que cada uno de los estadios representa para el desarrollo.

No eludimos que la división del desarrollo en estadios plantea diversos problemas teóricos que debemos reconocer y enfrentar. Uno de los más importantes es el de la dirección y determinación del desarrollo. Coll afirma que «los cambios estructurales que se producen son en cierto sentido unidireccionales, irreversibles y poseen una orientación hacia determinados estados finales» (Coll, 1979:65). Afortunadamente, ya Waddington (1957) nos ilustraba con la posibilidad de que el resultado final del desarrollo no sea unívoco sino que, al contrario, cabe la posibilidad de que pueda recorrer diferentes paisajes epigenéticos, desembocando en una multiplicidad de resultados. Entre los factores que generan esta diversidad están los de origen ambiental, que determinan las experiencias del sujeto. Marchesi, Carretero y Palacios (1995:311) revisaron los estudios que apuntan hacia tal variedad:

La diversidad sustituye a la universalidad hasta el punto de que los desfases entre las diferentes

tareas del estadio de las operaciones concretas, típicas de los sujetos occidentales, no suelen encontrarse en otras culturas, aunque se hallen otros.

De manera que los datos parecen apoyar una salida a esta controversia, en el sentido de que la maduración biológica condiciona en mayor medida las adquisiciones de las primeras etapas y el peso de las experiencias aumenta cada vez más, hasta el punto de que no se alcanzan los últimos estadios genéticamente posibles sin el apoyo específico del cultivo social y cultural.

En definitiva, parece posible describir el desarrollo del niño en estadios que le llevan en una dirección enriquecedora y más adaptativa.

La concreción de las diversas etapas debería responder al ya mencionado criterio de la observabilidad, de manera que la propia denominación de las etapas refleje la realidad del desarrollo psicomotor, guiado tanto por la maduración neuropsicológica como por las variables del contexto en el que se produce la adaptación: la capacidad inicial de movimiento, la capacidad de exploración del entorno o la progresiva capacidad de comunicación, son ejemplos que podrían representar hilos conductores para una escala de evaluación y medida del desarrollo.

En quinto lugar, recordemos de nuevo que si bien una escala de desarrollo no constituye una teoría, sí debe ser un ingenio al servicio del desarrollo, que debe adoptar una posición crítica al juzgarlo. Para ello debemos exigir a las escalas una *capacidad de predicción que sirva a la finalidad preventiva*. De cara a conseguirlo, las escalas deberían reflejar aquellas desviaciones de la norma que, provocadas por los diversos accidentes del paisaje epigenético, puedan condicionar el curso del desarrollo, para proceder en consecuencia a la implementación de

las decisiones y recursos que eviten la pérdida de niveles de desarrollo genéticamente previstos.

Nos enfrentamos al problema de tener que decidir qué alteraciones y qué desfases –respecto de la norma– deben ser considerados como una alerta que desencadene nuestra intervención. Siguiendo la lógica del modelo estructural ya propuesto, y dentro del que la secuencia del desarrollo natural se recorre en sentido unidireccional, «lo importante es la secuencia del desarrollo y no tanto la edad en que se producen las transiciones de un estadio al otro» (Marchesi, Carretero y Palacios, 1995:302). De manera que aunque los cambios suelen observarse a determinadas edades, la propia edad cronológica no puede ser interpretada como motor del cambio. Así, la edad se convertiría en un dato relativo, de modo que cualquier escala y las decisiones que de ella se derivaran deberían ser respetuosas con el tempo individual de cada niño.

La capacidad predictiva de los instrumentos de evaluación constituye una de sus más serias limitaciones, como señalan García-Tornel, García García, J. Reuter, Clow y L. Reuter, a propósito de la adaptación para España de la Escala de Desarrollo Infantil de Kent (EDIK), al reconocer que «la principal línea futura de investigación es conocer su valor predictivo haciendo un estudio a medio plazo con niños de riesgo». Estos mismos autores aceptan que el Cociente de Desarrollo (CD) «permite una detección precoz del ritmo de evolución de las diferentes áreas, pero en ningún caso de lo que será el futuro Cociente Intelectual». Y a la hora de tomar decisiones se mantienen en la posición secular de utilizar el CD «para determinar si el niño se desarrolla adecuadamente, si tiene riesgo de retraso en el desarrollo o si presenta un franco retraso» (García-Tornel y otros, 1996: 451).

Disponer de valores críticos en determinados índices sería quizá esencial para aventurar tendencias en el futuro curso del desarrollo; el cruce de variables asociadas a la edad cronológica y a la homogeneidad del perfil evolutivo podría resultar productivo en este sentido. Vamos a operativizar esta línea de trabajo más adelante aportando el concepto de *índice de armonía del desarrollo*.

Se adivina ya que, en sexto lugar, otra de las características que valoraremos como muy deseable en una escala será su *utilidad como herramienta investigadora*, al servicio de las teorías del desarrollo y de la investigación clínica y educativa.

Partiendo de las posibilidades que se abren al disponer de un *índice de armonía*, la escala que más adelante concretamos podría servir para correlacionar las distintas disarmonías que se caractericen –dependiendo de las funciones afectadas–, con sus resultados evolutivos parciales y finales, de manera que se puedan concluir predicciones sobre el curso de desarrollo que seguirá el sujeto, incluyendo la variable tipo de tratamiento, e incluso observar la importancia de la precocidad del diagnóstico.

- ¿Existen realmente períodos críticos para la adquisición de determinadas habilidades?
- ¿Cuáles serían los límites cronológicos que los definen?
- ¿Tendrían éstos relación con el tipo de disarmonía?
- ¿Se ampliarían estos límites, mejorando el pronóstico, si se tratara de otra determinada disarmonía?

Éstos y otros temas de estudio nos asaltarán probablemente en nuestro proyecto y en la continuidad que le deseamos.

En la **Tabla 1.1** enumeramos, a modo de resumen, el listado de las 6 características solicitadas a un instrumento de evaluación del desarrollo infantil, siguiendo el mencionado capítulo de Marchesi, Carretero y Palacios (1995).

3. ¿RESPONDEN LAS ESCALAS DISPONIBLES A ESTAS CARACTERÍSTICAS?

En la **Tabla 1.2** listamos las pruebas de valoración y seguimiento del desarrollo infantil, analizadas en Abellán (2003).

En la **Tabla 1.3** contrastamos cada una de las 11 herramientas de evaluación y diagnóstico del desarrollo infantil con cada uno de los seis aspectos mencionados en la Tabla 1.1.

La proporción de pruebas que cumplen con la condición, respecto del total de 11 pruebas revisadas se expresa a continuación:

1. *Que facilite una intervención enriquecedora:*
8/11: Desde una perspectiva cualitativa

observamos que la mayoría de las escalas ayudan a planificar la estimulación, como una consecuencia lógica de su diacronicidad; los hitos del desarrollo aparecen ordenados en la misma secuencia en que se presentan estadísticamente en el desarrollo natural.

2. *Que evalúe al niño dentro de múltiples sistemas:*

6/11: Son mayoría los autores que advierten la ventaja de evaluar al sujeto en los distintos ámbitos en los que desarrolla su actividad y a los que ha de adaptarse.

3. *Que utilice una metodología observacional:*

6/11: También, la mayoría confía más en lo que el niño ya viene demostrando que ha adquirido que en una evaluación basada en experiencias más o menos artificiales, a las que se somete al sujeto, y que llena de dudas en demasiadas ocasiones al experimentador, obligándole a recabar

Tabla 1.1. Resumen de las características solicitadas a las escalas de evaluación.

CARACTERÍSTICA	DEFINICIÓN
1 Intervención enriquecedora	Que sirva de guía para estimular y optimizar el desarrollo infantil
2 Múltiples sistemas	Que evalúe el comportamiento en los diversos entornos cotidianos
3 Metodología observacional	Que el evaluador verifique los hitos adquiridos en contexto natural
4 Concepción estructural	Que utilice una escala única a través de todos los niveles y etapas del desarrollo
5 Capacidad predictiva	Que demuestre capacidad de predicción para resultar útil a la prevención
6 Herramienta investigadora	Que haga aportaciones singulares que sirvan a la investigación del desarrollo

Tabla 1.2. Pruebas seleccionadas para el análisis.

Nº	NOMBRE Y REFERENCIA
1	Escala Observacional del Desarrollo (Secadas, 1992)
2	Escalas McCarthy de Aptitudes y Psicomotricidad para Niños (Cordero et al., 1996)
3	Inventario de Desarrollo Battelle (Fundació Catalana Síndrome de Down, 1989)
4	Escalas Bayley de Desarrollo Infantil (Bayley, 1977)
5	Guía Portage de Educación Preescolar (Bluma et al., 1995)
6	Escala de Desarrollo Psicomotor Primera Infancia, Brunet-Lézine (R) (Josse, 1997)
7	Escala de Desarrollo Infantil de Kent (García-Tornel et al., 1996)
8	Escala de Desarrollo Haizea-Llevant (Fernández-Álvarez, 1991)
9	Valoración Psicomotriz por el Método de Denver (Frankenburg y Dodds, 1967; Frankenburg et al., 1992; Salinas-Aracil, 2003)
10	Child Development Inventory (Ireton y Thwing, 1988)
11	Escala de Huzgiris y Hunt (Uzgiris y Hunt, 1975)

Tabla 1.3. Adaptación de las pruebas evaluadas a los criterios propuestos.

PRUEBAS	Intervención enriquecedora	Evalúa múltiples sistemas	Metodología observacional	Concepción estructural	Capacidad Predictiva	Herramienta investigadora
EOD	sí	sí	sí	sí	No	sí
MSCA	no	no	no	no	No	no
BATTELLE	sí	sí	sí	no	No	sí
BAYLEY	sí	no	no	no	No	sí
PORTAGE	sí	sí	sí	sí	No	no
BRUNET-L	sí	no	no	no	No	sí
EDIK	sí	sí	sí	no	No	no
HAIZEA-LI	no	sí	sí	no	No	no
DENVER	no	no	no	no	No	no
CDI	sí	sí	sí	no	No	sí
UZG. & H.	sí	no	no	sí	No	no

informaciones complementarias de personas cercanas al niño que pueden conocerle mejor.

4. *Que consista en una escala única dentro de una concepción estructural:*

3/11: Son pocos los autores que parten de una concepción del desarrollo basada en estadios, y los que lo hacen se limitan a describir la existencia de procesos más que a determinar las relaciones entre los mismos; en este punto parece que sólo la investigación dará con el tiempo algunas respuestas.

5. *Que demuestre capacidad de predicción para resultar útil en la prevención:*

0/11: Los autores o no abordan el problema por considerarlo demasiado alejado de sus pretensiones o plantean que sólo con arduas investigaciones longitudinales podría llegar a saberse el valor predictivo de sus obras. Una dificultad añadida es decidir si el valor predictivo quedaría demostrado hallando una alta correlación con el futuro Cociente Intelectual o con determinadas dificultades específicas de desarrollo (aprendizaje, enfermedad mental, etc.).

6. *Que sea válida como herramienta al servicio de la investigación:*

5/11: También una minoría ha conseguido desarrollar herramientas dotadas de suficiente complejidad como para servir a fines de investigación. Esta finalidad exige un diseño con un elevado número de ítems, áreas, procesos y un exhaustivo proceso de validación y fiabilización estadística, difícil de conseguir por la cantidad y complejidad de recursos que precisa.

4. ¿CÓMO PODEMOS VALORAR TÉCNICAMENTE LAS ESCALAS?

En Abellán (2003), cumplimentamos un cuestionario con 27 datos y una conclusión general, tomando como referencia el trabajo de Prieto y Muñiz (2000) sobre la evaluación de los tests utilizados en España. Lo hicimos siguiendo dos variables: el *grado de especialización* y el *grado de especificidad* de las distintas escalas de valoración y medida del desarrollo. También reflejamos en el análisis pormenorizado de las pruebas otros aspectos que podían ser interesantes para nuestro objetivo de hacer una nueva propuesta.

Entendíamos por *especialización* los aspectos relacionados con su utilización, como el rango de edades al que se aplican sin modificar la metodología, el número de factores o áreas del desarrollo que evalúan y el nivel de cualificación requerido a los usuarios de la prueba.

En cuanto a la *especificidad*, nos referíamos por tal a la rigidez disciplinaria, atendiendo a la fuente de la que proceden los ítems o reactivos del test, por ejemplo si se trata de hitos psicológicos, pedagógicos o datos biomédicos; así como la profesión de los usuarios.

En la **Tabla 1.4** se muestran estos aspectos, que sirvieron de guía para la elaboración del cuestionario de análisis de las pruebas habitualmente utilizadas en nuestro país.

Veamos un rápido resumen de esas informaciones (resumidas también en la **Tabla 1.5**), comenzando por los aspectos relacionados con la utilidad del test, lo que llamábamos *grado de especialización*:

I. *Rango de edades al que se aplica sin modificar la metodología:*

La prueba 1 de 0 a 17 años, cumplimentada por el observador de 0 a 9 años, y por el propio sujeto de 10 a 17 años.

Tabla 1.4. Aspectos evaluables de los tests utilizados en España.

Grado de Especialización	I. Rango de edades a que se aplica sin modificar la metodología
	II. Número de factores o áreas que evalúa
	III. Nivel de cualificación de los usuarios
Grado de Especificidad	IV. Fuentes de la prueba
	V. Usuarios a los que sirve
Otros aspectos	VI. Identificación de la prueba
	VII. Número de ítems y rango de edad
	VIII. Corrección, informe y tipo de soporte

La prueba 2 de 2 y medio hasta 8 años y medio.

La prueba 3 de 0 a 8 años.

La prueba 4 de 2 meses hasta 2 años y medio.

La prueba 5 de 0 a 6 años.

La prueba 6 de 2 a 30 meses.

La prueba 7 de 0 a 15 meses.

La prueba 8 de 0 a 6 años.

La prueba 9 de 1 mes a 6 años.

La prueba 10 de 15 meses a 6 años.

La prueba 11 de 0 a 24 meses.

Contamos con sólo una escala hasta 17 años, la prueba 1; con dos escalas hasta 8 años, las pruebas 2 y 3; con cuatro escalas hasta 6 años, las pruebas 5, 8, 9 y 10; y con cuatro *baby-test*, las pruebas 4, 6, 7 y 11. Por otra parte, algunas de las pruebas no están especializadas en los primeros momentos del desarrollo, como sucede con las pruebas 2, 4, 6, 9 y 10.

II. Número de factores o áreas que evalúan:

Resumimos la información en función de las áreas incluidas en el análisis:

Ninguna prueba: Inteligencia.

Prueba 2: Aptitudes.

Prueba 1: Habilidades y rendimiento académico.

Pruebas 1, 2, 3, 4, 5, 6, 7, 8, 9 y 10: Psicomotricidad.

Ninguna prueba: Alteraciones neuropsicológicas.

Pruebas 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 y 11: Desarrollo neurocognitivo.

Pruebas 1, 3, 4, 5, 6, 7, 8, 9 y 10: Desarrollo socioafectivo y adaptación.

Pruebas 1, 2, 3, 4, 5, 6, 7, 8, 9 y 10: Desarrollo lingüístico y oral.

Pruebas 1, 4, 8 y 10: Otros.

Observamos que el *desarrollo neurocognitivo* es valorado por todas las pruebas; la *psicomotricidad*, el *desarrollo socioafectivo-adaptativo* y el *lingüístico-oral* es valorado por la mayoría de las pruebas; sólo cuatro escalas se interesan por *otros aspectos*; y en un único caso por *aptitudes* o por *habilidades-rendimiento académico*. Al tratarse de escalas de desarrollo, no aparece ninguna que valore *inteligencia –ad hoc–* ni tampoco *alteraciones neuropsicológicas*.

III. Nivel de cualificación de los usuarios:

Resumimos la información para los distintos niveles:

Pruebas 1, y 5: Ninguna.

Ninguna prueba: Entrenamiento y acreditación específica.

Pruebas 1 y 5: Nivel A Tests de rendimiento y conocimiento.

Ninguna prueba: Nivel B Tests colectivos de aptitudes e inteligencia.

Pruebas 2, 3, 4, 6, 10 y 11: Nivel C Aplicación individual y complejidad.

Pruebas 7, 8 y 9: Otra: pediatras, maestros, etc.

Las pruebas 1 y 5 comparten poder ser utilizadas directamente por los padres o ser utilizadas por profesionales, ya que basta un mínimo nivel de cualificación. Pero por lo general, aunque los padres faciliten a los profesionales cuestionarios ya resueltos, la mayoría de las pruebas se adscriben al Nivel C (seis pruebas), dada su complejidad. Otras tres pruebas han sido diseñadas pensando específicamente en su manejo por parte de profesionales de la salud (incluidos los médicos pediatras) y/o de la educación.

En segundo lugar revisamos la información que compara las diversas escalas en cuanto a su rigidez disciplinaria o *grado de especificidad*:

IV. Fuente de la prueba:

La prueba 1 incluye ítems de manuales, escalas y estudios doctorales.

La prueba 2 incluye tareas consideradas como reflejo de la capacidad cognitiva y motórica, pero no se describe la fuente.

La prueba 3 incluye los ítems de distintas escalas y trabajos de investigación.

La prueba 4 incluye ítems de otras escalas.

La prueba 5 incluye los ítems de diversos tratados y escalas de evaluación, y se com-

plementó con ítems propuestos por los primeros usuarios.

La prueba 6 incluye ítems de otros test utilizados para observar directamente a bebés. En las pruebas 7, 8, y 9, los ítems proceden de la versión original en inglés.

Los ítems de la prueba 10 pertenecían a la anterior versión en inglés.

La prueba 11 utiliza clásicas tareas piagetianas.

Al parecer, la mayor parte de los ítems que se utilizan a la hora de construir una nueva escala de desarrollo proceden de tratados, manuales y otros tests elaborados durante las primeras décadas del siglo XX. Algunos autores complementan sus elementos con ítems procedentes de investigaciones contemporáneas en el momento de editar su propio test, y en unas pocas ocasiones se proponen ítems completamente originales. Podemos extraer dos conclusiones: una, que en general parece que los ítems son comunes al colectivo de los estudiosos y son utilizados por todos; y otra, que generalmente no se diferencia entre los ítems útiles para profesionales pertenecientes a una u otra disciplina.

V. Usuarios a los que sirve:

Resumimos la información acerca del área o áreas en que se aplica cada escala:

Pruebas 2, 3, 4, 5, 6, 8, 10, y 11: Psicología clínica.

Pruebas 2, 3, 5, 8, y 10: Psicología educativa

Ninguna prueba: Neuropsicología.

Ninguna prueba: Servicios sociales.

Pruebas 1, 3, 5, 7, 8, y 9: Otros: padres, educadores, investigadores, educación especial, pediatría...

De las once, ocho pruebas sirven a los psicólogos clínicos, lo que resulta lógico por la

importancia que la psicología evolutiva adquiere cuando lo que preocupa es el desarrollo de los niños. No es casualidad que el 68% de los psicólogos se consideran clínicos (Berduñas, 2002). El psicólogo clínico busca alteraciones del desarrollo como indicio de patología. De esas ocho, 5 pruebas sirven también para investigar las dificultades que los niños presentan para adaptarse al currículo educativo. Y hay seis pruebas que pueden ser utilizadas por médicos pediatras, padres y otros educadores. Al tratarse de escalas generales de desarrollo, no hemos revisado ninguna prueba con utilidad específica para neuropsicólogos y/o neuropediatras. Tampoco ninguna exclusiva para los trabajadores sociales.

Otros aspectos que hemos reflejado en el análisis pormenorizado de las pruebas pueden ser interesantes para nuestro objetivo final de una nueva propuesta. Veamos un rápido resumen de esas informaciones:

VI. Identificación de la prueba:

La prueba 1 está diseñada y baremada en España, siendo una obra original de autor español.

La prueba 2 procede de la versión original en inglés, y no ha sido baremada sistemáticamente para la población española.

Las pruebas 3 y 4 también proceden de su original en inglés, pero en la versión española se utilizan los baremos correspondientes a la población estadounidense.

La prueba 5 no utiliza baremos al tratarse de una guía de evaluación y estimulación, y ha sido traducida del inglés.

La prueba 6 fue traducida del francés, pero no baremada para España.

La prueba 7 no ha sido publicada en español, y se trata de una adaptación de la original en inglés, estandarizada con muestras catalanas.

La prueba 8 ha sido adaptada del inglés, estandarizada con muestras vasca y catalana, y editada para el País Vasco, pero no distribuida comercialmente.

La prueba 9 es una traducción incompleta de la original en inglés, y se ofrece en Internet.

La prueba 10 es una obra original en inglés que no ha sido traducida ni adaptada.

La prueba 11 es un documento interno de un equipo multiprofesional, que ha traducido parte de una obra original en inglés.

En resumen, nueve pruebas son originales en inglés, una en francés y otra en español. Por otra parte observamos que tres pruebas disponen de baremos elaborados con muestras españolas, pero de ellas sólo una es representativa de la población española, aunque no utilizó muestreo sistemático por cuotas.

VII. Número de ítems que incluye la prueba y rango de edad:

Prueba 1: 1657 ítems, de 0 a 17 años

Prueba 2: 134 ítems, de 2.5 a 8.5 años

Prueba 3: 341 ítems, de 0 a 8 años

Prueba 4: 244 ítems, de 2.5 años

Prueba 5: 578 ítems, de 0 a 6 años

Prueba 6: 150 ítems, de 0 a 2.5 años

Prueba 7: 252 ítems, de 0 a 15 meses

Prueba 8: 97 ítems, de 0 a 6 años

Prueba 9: 68 ítems, de 0 a 6 años

Prueba 10: 320 ítems, de 0 a 6 años

Prueba 11: 137 ítems, de 0 a 2 años

Se obtiene una media de 66 ítems para evaluar cada año de edad, pero entre unas y otras pruebas las diferencias son grandes, al igual que en los rangos de edad evaluados, ya que la desviación estándar es de 44 ítems aproximadamente.

Tabla 1.5. Resumen de los aspectos evaluados de las pruebas seleccionadas.

ASPECTO A EVALUAR	PRUEBAS	N	%
Rango de edades			
[0-17]	1	1	9%
[0-9]	2,3	2	18%
[0-6]	5,8,9,10	4	36%
[0-4]	4,6	2	18%
[0-2]	7,11	2	18%
Nº de factores que evalúa			
Aptitudes	2	1	9%
Habilidades y rendimiento académico	1	1	9%
Psicomotricidad	1,2,3,4,5,6,7,8,9,10	10	91%
Desarrollo neurocognitivo	1,2,3,4,5,6,7,8,9,10,11	11	100%
Desarrollo socioafectivo y adaptación	1,3,4,5,6,7,8,9,10	9	82%
Desarrollo lingüístico y oral	1,2,3,4,5,6,7,8,9,10	10	91%
Otros aspectos	1,4,8,10	4	36%
Nivel de cualificación de los usuarios			
Nivel A. Tests de rendimiento y conocimiento	1,5	2	18%
Nivel C. Aplicación individual y complejidad	2,3,4,6,10,11	6	54%
Salud y/o educación	7,8,9	3	27%
Fuentes de la prueba			
Tratados y manuales sobre desarrollo	1,5	2	18%
Otras escalas	1,3,4,5,6	5	45%
Estudios doctorales e investigaciones	1,3	2	18%
Propuestas de los usuarios	5	1	9%
Tareas clásicas piagetianas	11	1	9%
Obtenidos de la versión original en inglés	7,8,9,10	4	36%
No descritas	2	1	9%
Usuarios a los que sirve			
Psicología clínica	2,3,4,5,6,8,10,11	8	73%
Psicología educativa	2,3,5,8,10	5	45%
Educación y/o pediatría	1,3,5,7,8,9	6	54%

VIII. Corrección, informe y tipo de soporte:

La prueba 7 se puede corregir con lector óptico, pero en el resto de las pruebas no es posible una corrección automática o informatizada, y en ningún caso los editores ofrecen informes a partir de los resultados. La prueba 9 se ofrece en soporte informatizado a través de Internet, pero sólo resulta útil puntualmente, al no permitir contemplar toda la prueba, ni ofrecer informe automatizado.

5. ¿CÓMO DEBERÍA SER UNA ESCALA DE DESARROLLO?

Comenzamos este Capítulo 1 revisando el interés que el concepto de desarrollo ha tenido a largo de la historia de la ciencia. Definimos a continuación una serie de características que resultan exigibles a los artefactos de medida, a saber: Que sirvan de guía de estimulación evaluando el comportamiento mediante observación en contextos naturales, utilizando una escala única para todo el periodo medido y que incorpore valores como herramienta predictiva e investigadora.

Elaboramos después un cuestionario al que someter a las escalas disponibles analizando técnicamente su grado de especialización y de especificidad. Comprobamos que ninguna de las pruebas analizadas se adecuaba a los seis criterios exigidos desde dentro de nuestro planteamiento, lo que sumado a las limitaciones encontradas en el estudio pormenorizado, nos llevó a pensar en la necesidad de desarrollar una nueva herramienta de evaluación y medida del desarrollo infantil que intente superar las carencias que resumimos a continuación:

Respecto al grado de especialización:

- I. El rango de aplicabilidad suele centrarse en un período limitado de años,

obligando al examinador a cambiar de herramienta varias veces a lo largo de un seguimiento evolutivo. Uno de los principales inconvenientes que se encuentran en la práctica es que las herramientas homologables para dar continuidad al seguimiento no coinciden en su apreciación del desarrollo, en sus escalas de medida o en el tipo de unidades utilizadas, de manera que los informes se hacen confusos no sólo para los profesionales sino también para los padres, los cuales no suelen tener preparación técnica para comprender lo sucedido. Otro de los problemas habituales consiste en encontrar *zonas muertas* que las herramientas homologables dejan sin evaluar al no coincidir sus rangos de aplicabilidad. Lo deseable sería disponer de una medida continuada a lo largo de todo el desarrollo sin cambiar la metodología.

- II. El número de factores o áreas evaluadas intentando analizar el constructo *desarrollo general* suele limitarse a cuatro o cinco. Ocasionalmente, algunos autores incluyen otras áreas en relación a sus intereses particulares, pero por lo general no se justifica la decisión. Parece necesario integrar en el análisis todas aquellas funciones que permitan una visión global del desarrollo, atendiendo a sus vertientes orgánica y psicoeducativa. El número de funciones debería ser suficientemente elevado como para permitir diagnosticar el aspecto que necesita estimulación o rehabilitación, y suficientemente breve como para que las evaluaciones no resulten demasiado artificiales ni tediosas.
- III. La aplicación de las pruebas corre a cargo de profesionales especializados

en la mayoría de las ocasiones, de manera que su utilidad preventiva queda sistemáticamente limitada por no poder ser manejadas directamente por los padres y otros educadores. Hacer pruebas a los niños requiere la utilización de costosos recursos (temporales, económicos y profesionales), con lo que el número de niños realmente seguidos con garantías es siempre pequeño en relación a la población total. Como los padres se sienten incompetentes para acercarse a observar y comprender el desarrollo de sus hijos, muchos trastornos que deberían recibir una atención temprana no son detectados hasta que causan un problema mayor. Lo ideal sería una herramienta que pudiera ser utilizada por médicos, psicólogos, pedagogos, maestros, trabajadores sociales e incluso por los propios padres.

Respecto al grado de especificidad:

- IV. Nuestro análisis demuestra que los tests de desarrollo general no son realmente específicos para una u otra profesión, sino que, al contrario, parece que todos comparten un gran almacén de datos sobre lo que se considera normal en el desarrollo a una y otra edad. A pesar de la indicación de uso para diagnóstico pediátrico o psicológico, lo cierto es que estudian los mismos fenómenos, y la verdadera diferencia estriba en quién tomó la iniciativa de construir la herramienta.
- V. Resulta evidente que un curso alterado de desarrollo tendrá repercusiones sobre el desarrollo tónico-motor, la capacidad de comunicación y de relación con el entorno social, el desarrollo de la autonomía personal, el aprendi-

zaje reglado y no reglado, sobre el desarrollo intelectual en general, por lo que una prueba diagnóstica debería servir al mismo tiempo a los psicólogos clínicos y educativos, a los médicos neonatólogos, pediatras y neuropediatras, a todos los educadores y a gran variedad de estudiosos e investigadores. Los autores de las pruebas analizadas expresan por lo tanto una *limitación innecesaria*, ya que discriminan las pruebas en relación a la ciencia que profesan. El desarrollo es un objeto de estudio que compartimos entre todos los que de él entendemos. Evidentemente, cada profesional se ocupará después de su parcela y coordinará sus actuaciones dentro de un trabajo interdisciplinario que garantice una correcta toma de decisiones.

Otros aspectos analizados:

- VI. No se da el caso de una prueba diseñada enteramente en España, en lengua castellana y que haya sido baremada utilizando muestras representativas de toda la población española. Está claro que en nuestro país no se dedican esfuerzos suficientes a esta cuestión, y continuamos siendo subsidiarios de las iniciativas de nuestros colegas norteamericanos. Esta situación es claramente negativa desde varios puntos de vista, y este trabajo no es el lugar adecuado para recogerlos, aunque nos permitimos expresar que la carencia de instrumentos legítimos de evaluación transgrede principios éticos de nuestro código deontológico (véanse las normas desarrolladas por la *Comisión de Test* del Colegio Oficial de Psicólogos en <http://www.cop.es>).

- VII.** Existen *baby-tests*, y pruebas que evalúan hasta los 6 y 8 años, (aunque no todas comienzan desde el nacimiento), pero sólo la EOD alcanza hasta los 17, si bien en modalidad de autoencuesta a partir de los 10 años. Se aprecia cómo en general las pruebas no completan el desarrollo básico de la infancia, si consideramos ésta como el período que abarca desde el nacimiento hasta el inicio de la adolescencia, es decir hasta los 12 primeros años de la vida. Una prueba unificada que recogiera dicho período debería de contar con alrededor de 800 ítems, según la media calculada más arriba.
- VIII.** Las pruebas se comercializan casi exclusivamente impresas en papel, y los

editores no ofertan un servicio post-venta al usuario, ni como apoyo técnico para el entrenamiento ni como asesoría para la redacción de informes a partir de los resultados en dichas pruebas. Pensamos que la utilización de las nuevas tecnologías de la información puede revolucionar el formato y soporte de las pruebas, así como el tipo de relación que el editor mantiene con su cliente. Es posible que incluso algunos aspectos de la metodología de validación y fiabilización puedan beneficiarse de ello; hoy día estamos en disposición de hacerlo.

