



**UNIVERSIDAD AUTÓNOMA METROPOLITANA**  
**UNIDAD XOCHIMILCO**

---

**DIVISIÓN DE CIENCIAS BIOLÓGICAS Y DE LA SALUD**

**VALORACIÓN DE LA SENSIBILIDAD Y ESPECIFICIDAD DE LA  
PRUEBA DE TAMIZAJE DEL DESARROLLO DENVER II.**

**PARA OBTENER EL GRADO DE  
MAESTRA EN REHABILITACIÓN NEUROLÓGICA**

**PRESENTA**

**LIC. MARÍA DEL ROSARIO OLIVERA MORENO**

**COMITÉ TUTORAL**

**DIRECTOR: DR. ROLANDO RIVERA GONZALEZ**

**ASESORA: M.R.N. IVONE MARTÍNEZ VÁZQUEZ**

**ASESORA: M.R.N. FABIOLA SOTO VILLASEÑOR**

México, D.F

Febrero 2010

## **Agradecimientos.**

*Mi más profunda gratitud al pueblo mexicano que hizo posible la realización de este trabajo de investigación.*

*Mi reconocimiento y eterna gratitud a mi Director de tesis y asesoras quienes hicieron posible con sus recomendaciones académicas, dedicación y paciencia la culminación de mi trabajo de investigación.*

*De manera especial, agradezco a las familias de la bella comunidad de Tlaltizapan Morelos por la calidez con la que me recibieron y a mis queridas compañeras Flavia, María Luisa, Magdalena, Jovita, Lupita y a todo el personal del CER-INP por su compromiso con las familias de la comunidad y por la amistad y cariño que me brindaron.*

*De igual forma mi agradecimiento a mis maestros y maestras del Laboratorio de seguimiento del Neurodesarrollo y de la UAM-Xochimilco.*

*A mis compañeras de la Generación XXVII, muchas gracias por su apoyo y amistad.*

## INDICE

<b>INTRODUCCIÓN.....</b>	<b>1</b>
<b>I.- LA EVALUACIÓN DEL DESARROLLO INFANTIL.....</b>	<b>4</b>
1.1.- <i>Los comienzos del estudio del niño.</i>	4
1.2.- <i>Generalidades de los métodos e instrumentos para la evaluación del desarrollo infantil.</i>	11
<b>II.- VALIDACIÓN DE LAS PRUEBAS PSICOLÓGICAS.....</b>	<b>17</b>
2.1.- <i>Validez de los instrumentos de medición</i>	17
2.2.- <i>Tipos de validez</i>	17
2.3.- <i>Indicadores estadísticos de validez</i>	19
<b>III.- PRUEBA DE TAMIZAJE DEL DESARROLLO DENVER II (DDST-II).....</b>	<b>22</b>
3.1.- <i>Antecedentes Históricos</i>	22
3.2.- <i>Descripción de la prueba</i>	24
3.3.- <i>Estudio de Validación</i>	27
3.4.- <i>Estudios realizados en México</i>	29
<b>IV. ESCALA DE DESARROLLO INFANTIL DE GESELL.....</b>	<b>34</b>
4.1.- <i>Principios básicos</i>	34
4.2.- <i>Descripción de la prueba</i>	35
<b>V.- METODOLOGÍA.....</b>	<b>42</b>
5.1.- <i>Justificación</i>	42
5.2.- <i>Objetivo General</i>	43
5.3.- <i>Objetivos Específicos</i>	43
5.4.- <i>Tipo de estudio</i>	43
<b>VI.- MATERIALES Y MÉTODOS.....</b>	<b>43</b>
6.1.- <i>Población de estudio</i>	43
6.2.- <i>Criterios de Inclusión</i>	44
6.3.- <i>Criterios de Exclusión</i>	44
6.4.- <i>Ubicación del estudio</i>	44
6.5.- <i>Muestra</i>	44
6.6.- <i>Variables</i>	44
6.7.- <i>Instrumentos de evaluación del desarrollo</i>	45
6.8.- <i>Procedimiento</i>	47
6.9.- <i>Análisis estadístico</i>	47

<b>6.10.- Ética</b>	<b>48</b>
<b>VII.- RESULTADOS</b> .....	<b>49</b>
<b>VIII. DISCUSIÓN</b> .....	<b>69</b>
<b>IX. CONCLUSIONES</b> .....	<b>74</b>
<b>X. REFERENCIAS BIBLIOGRAFICAS</b> .....	<b>75</b>
<b>ANEXOS</b> .....	<b>80</b>

## INTRODUCCIÓN

Las diversas organizaciones pediátricas americanas y británicas en los últimos años han señalado y publicado la importancia del monitoreo del desarrollo durante el cuidado preventivo de la salud del niño dentro del marco del proceso de vigilancia del mismo. La importancia de la vigilancia del desarrollo está en distinguir niños con alto y bajo riesgo de presentar trastornos durante su desarrollo y mayor probabilidad de futuras incapacidades, lo cual permitirá una intervención temprana y un monitoreo del desarrollo de forma periódica, este proceso de vigilancia incluye el empleo de pruebas de tamiz que se caracterizan por ser simples, confiables, válidas (sensibles y específicas) rápidas en su aplicación y de fácil interpretación (1).

Las pruebas de tamizaje o filtro son recomendadas para el seguimiento del desarrollo del niño sano, cuya importancia estriba en poder diferenciar o identificar correctamente a los niños con desviaciones en su desarrollo de aquellos que no presentan ningún tipo de alteración de manera confiable, es decir, que sus indicadores de validez se ubiquen en los rangos esperados de probabilidad de acierto estadísticamente confiables. Estos indicadores se relacionan con la capacidad de los instrumentos para detectar la mínima diferencia (entre niños sanos y enfermos) considerada clínicamente significativa (2).

Las recomendaciones de organismos internacionales sobre la vigilancia del desarrollo del niño recomiendan el uso de pruebas de tamizaje con mayor sensibilidad para detectar alteraciones en las diferentes áreas del desarrollo infantil y que reflejen la evolución y cambios en las conductas consideradas críticas del desarrollo tomando en cuenta la diversidad individual y los contextos medio-ambientales a los que están expuestos los infantes, para lo cual se requiere del control de algunas variables que influyen en el desarrollo tales como: género, condición de riesgo al momento del nacimiento, escolaridad materna, características socio-económicas etc., ya que estas condiciones repercuten en mayor o menor grado en las diferentes áreas del desarrollo.

En las pruebas de tamizaje se consideran las expresiones conductuales observables durante la evolución del desarrollo como indicadores válidos para identificar las desviaciones del mismo, indicadores que deben ser considerados en un contexto poblacional específico, por lo que, las pruebas de tamizaje requieren de estudios de validación en los que se realicen ajustes en términos de secuencia y temporalidad en la adquisición de las conductas esperadas (2).

Tomando en cuenta la necesidad de que las pruebas de tamizaje deben ser validadas observando las características poblacionales del grupo al que se pretende valorar y siendo el "Denver II Developmental Screening Test" uno de los instrumentos más utilizados mundialmente para la detección de niños en riesgo de presentar alteraciones del desarrollo y no habiéndose estandarizado en población infantil mexicana, se consideró necesario la valoración de la validez de la prueba por criterio externo (sensibilidad y especificidad), empleando como criterio estándar a la "Escala de desarrollo infantil de Gesell". Se procedió a contrastar los coeficientes generales de desarrollo de ambas pruebas así como los coeficientes de desarrollo por áreas, para determinar los indicadores de sensibilidad y especificidad de la prueba de tamizaje del desarrollo Denver II y establecer la validez de la prueba según grupos de edad.

El presente estudio "Valoración de la sensibilidad y especificidad de la prueba de tamizaje del desarrollo Denver II" se realizó como parte del proyecto "Perfiles de desarrollo de niños sanos empleando indicadores de 4 pruebas tamíz y una prueba diagnóstica de desarrollo" que se inició en Junio del 2006 en conjunto con el Laboratorio de Seguimiento del Neurodesarrollo del Instituto Nacional de Pediatría de la ciudad de México.

Dicho proyecto se realiza en el Centro de Estudios Rurales del Instituto Nacional de Pediatría en Tlaltizapan Morelos, centro de atención a la infancia en donde se vienen desarrollando investigaciones sobre el crecimiento y desarrollo infantil normal y en situaciones de desventaja por desnutrición, aplicando una serie de instrumentos para la evaluación de factores biológicos, psicológicos y sociales que lo impactan.

El objetivo principal del proyecto que viene implementándose en el centro es el determinar la edad y secuencia de adquisición de cada una de las conductas sensoriomotoras, adaptativas sociales y cognitivas de neonatos y lactantes, incluidas en cuatro instrumentos filtros (Denver II, Vanedela, CAT/CLAMS y Norma oficial Mexicana) para la vigilancia del desarrollo, empleando los procedimientos propios de cada una de las pruebas para su aplicación y establecer la prueba tamíz que con sus respectivos procedimientos mejor ajusta a las características de niños sanos que habitan en la comunidad de Tlaltizapan.

Así, de este proyecto se desprende el interés del presente estudio para valorar la validez por criterio externo de la Prueba de tamizaje del desarrollo Denver II, en una población de niños de cero a veinticuatro meses de edad residentes en la comunidad de Tlaltizapan. Los resultados obtenidos en el estudio nos demuestra las ventajas y desventajas que el instrumento tiene para detectar a la población a riesgo de alteraciones en el desarrollo, nos plantea la necesidad de buscar la forma de mejorar la especificidad de la prueba y el establecimiento de los indicadores de validez en diferentes grupos de edad.

## **I.- LA EVALUACIÓN DEL DESARROLLO INFANTIL.**

Para apreciar las características y formas en que evoluciona a lo largo del desarrollo de la conducta del niño y para determinar si este desarrollo obedece a lo esperado como normalidad, es necesario establecer patrones o normas de comportamiento promedio en cada una de las etapas en que se ha dividido el desarrollo psicológico del niño para su estudio e investigación. La graduación de los patrones de maduración sirve como escala reguladora del curso "normal" promedio de la evolución conductual. Como en todos los aspectos de la medición se debe seguir la metodología establecida para obtener mediciones objetivas válidas y confiables, en el caso del desarrollo del niño, es necesario investigar en grandes grupos de niños de todas las edades para obtener la media del comportamiento en cada uno de los niveles de desarrollo, tomando en cuenta la variabilidad individual. A través de las expresiones conductuales observables se puede determinar la edad que un individuo manifiesta a partir de la comparación con la conducta promedio establecida (3).

El desarrollo del niño se observa desde que los seres humanos existen, sin embargo, su estudio científico formal es relativamente nuevo, las concepciones acerca del niño y los métodos de investigación para su evaluación fueron cambiando con el tiempo, a través de los años fue la consolidación del método científico el que le dio estructura propia al estudio del desarrollo infantil, este estudio científico al cumplir con los criterios establecidos permitieron observar, registrar, comprobar hipótesis, y propagar las conclusiones de los diversos descubrimientos, de tal manera que se logre explicar en forma satisfactoria y predecir la conducta infantil.

### **1.1.- Los comienzos del estudio del niño.**

Las primeras observaciones sobre la conducta infantil en occidente datan del siglo XIV en Francia cuando comienzan a registrarse acontecimientos familiares en los denominados "Libros de Raison". Hasta entonces era una época en la que había un escaso aprecio a la infancia. Durante la época medieval y del renacimiento las representaciones pictóricas de los niños eran escasas y cuando aparecían en algunas pinturas eran adultos en pequeño

tanto por los rasgos como por las proporciones de las partes del cuerpo. Eran épocas en las que los adultos no veían a los niños cualitativamente distintos de ellos, o como si no tuvieran necesidades especiales, los niños eran considerados como algo divertido semejante a un animal y llegado el momento pasaban a formar parte de la sociedad adulta con las mismas obligaciones que estos.

Fue desde el siglo XVII hasta finales del siglo XIX en que algunas tendencias filosóficas importantes de la época prepararon el camino con sus fundamentos teóricos para el estudio del desarrollo infantil. La cuestión principal para entonces en epistemología consistía en contrastar la razón contra el sentido de percepción como medio para adquirir el conocimiento, es así como frente a esta postura René Descartes introduce el concepto de la dualidad en el hombre, la existencia del alma y el cuerpo, siendo que, posteriormente el alma y el cuerpo se estudiarían a través de dos métodos diferentes. Por un lado la captación del alma con sus ideas innatas, y por el otro, las investigaciones empíricas y experimentales acerca de la mecánica de los cuerpos (4). Estos planteamientos de Descartes sirvieron para dar origen a la escuela Empirista a fines del siglo XVII fundada por el filósofo John Locke quien describió la teoría de la "Tabla rasa" por la cual concebía la mente del niño como una hoja en blanco susceptible de ser escrita por el ambiente en que se encontrara, para los empiristas la experiencia es la única fuente válida del conocimiento en el contacto de los sentidos con las cosas (4), como contraposición a esta concepción surgió en Alemania la corriente filosófica denominada el Racionalismo, la cual planteaba que la causa principal del conocimiento reside en el pensamiento, en la razón misma y que en el hombre existen ideas innatas y por lo tanto la mente no era un papel en blanco; uno de los seguidores de esta corriente fue el filósofo francés Jean Jacques Rousseau quien afirmaba que los niños estaban dotados de un sentido moral innato. Para entonces la discusión de filósofos y en especial de los médicos se centraba en el papel relativo de la herencia y el ambiente en la persona (4).

Mientras los filósofos debatían sobre cuestiones como la naturaleza esencial de los seres humanos, un estudio documentado de un niño que había crecido en

aislamiento podría proporcionar evidencia del impacto relativo de la "naturaleza" y la "crianza", fue en 1799 en la provincia de Aveyron en Francia donde se encontró a un niño salvaje que fue llamado Víctor y enviado a una escuela para niños sordomudos quedando a cargo del médico Jean Itard quien ensayó nuevos métodos para entrenar al niño a discriminar sus respuestas emocionales así como a la instrucción en lenguaje, pensamiento, comportamiento social y moral convirtiéndose así en pionero de la educación especial y desarrollando un nuevo método de estudio como fueron las observaciones sobre casos de niños excepcionales (5).

En 1870 la Sociedad Pedagógica de Berlín publica el resultado de una encuesta sobre los contenidos de la mente infantil al entrar a la escuela, era el primer documento que se publicó y que se realizó con un número amplio de sujetos y fue en 1883 que Stanley Hall, considerado pionero del movimiento de estudio del niño, quien publicaría en los EE. UU. varios estudios que son considerados los primeros trabajos estadísticos de la época (5).

En 1882 el médico-fisiólogo Preyer brindó uno de sus mayores aportes en el estudio psicológico infantil con su obra "El alma del niño", en el cual describía el desarrollo físico y mental del niño, acuñando por primera vez en su obra el término de Psicogénesis (5).

A través de la historia acerca de las concepciones de la infancia y de los estudios realizados encontramos que existe una relación entre estas posturas teóricas y los instrumentos empleados en el estudio del desarrollo infantil, lo cual se pudo apreciar en el surgimiento de la psicometría que aportó una nueva metodología de estudio en el desarrollo infantil. Las primeras aplicaciones psicométricas tuvieron lugar en íntima conexión con los comienzos de la psicología experimental y sirvieron para la investigación de ciertos aspectos de la psicofisiología humana, posteriormente se extendieron al campo de la psicología diferencial, sirviendo de soporte al desarrollo de los "Tests mentales", pasando a ser estos instrumentos en sí mismos denominados "métodos psicométricos" (6).

El estudio de diferencias cuantitativas en diversas áreas sensoriales había sido iniciado por psicólogos y fisiólogos mucho antes del siglo XIX, de forma rudimentaria. Sin embargo, fue Fechner quien continuó con las investigaciones psicofisiológicas de Weber acerca de las relaciones entre estímulos y respuestas sensoriales, adoptando métodos experimentales para establecer la medida de la intensidad de las sensaciones. De esta forma la Psicología experimental se centra cada vez más en la ideación de métodos e instrumentos que permitan obtener medidas del comportamiento humano.

En 1879, Wundt inaugura en la Universidad de Leipzig el primer laboratorio experimental y realiza trabajos para establecer como se interrelacionan las sensaciones con las percepciones y las palabras con las ideas estableciendo un puente entre la psicofisiología y la psicología introspectiva. Uno de sus continuadores fue J. Mc. Cattell, quien estudio los mecanismos de la atención y la asociación de palabras y acuñó el término "Test mental" que utilizó por primera vez en 1890 (6)

Así pues, los primeros psicólogos experimentales fueron los pioneros de la psicometría. Sin embargo, la aparición de los tests mentales en los albores del siglo XX influiría de tal forma en el desarrollo de la psicometría, que incluso hoy en día es a estos métodos, y a los procedimientos estadísticos en que se fundan a los que denominamos "métodos psicométricos".

El interés sobre los fenómenos sensoriales se reflejaba en la naturaleza de los primeros tests psicológicos pero también pusieron de manifiesto la necesidad de un control riguroso de las condiciones en que se hacían las observaciones, de esta forma, se demostró la importancia de hacer observaciones sobre todos los sujetos bajo condiciones tipificadas, esta tipificación de procedimientos llegó a ser una de las características especiales de los tests psicológicos (6).

Francis Galton, biólogo inglés, desempeñó una gran influencia en el desarrollo del movimiento de pruebas. El campo psicométrico floreció gracias a que propuso por primera vez dos conceptos estadísticos importantes: regresión lineal y correlación, que permitieron estudiar la inteligencia y la relación entre los puntajes de pruebas entre padres e hijos. En 1884 estableció un laboratorio

en Londres abierto al público en el que se proporcionaban medidas de las capacidades físicas y mentales (6).

Con base en los trabajos de Galton, Karl Pearson desarrolló diferentes métodos estadísticos psicométricos, como por ejemplo el Coeficiente de Correlación de Pearson, mediante la cual se puede medir el grado de concordancia o discrepancia entre dos variables.

James Mckeen Cattell en 1890 se dedicó en América a la difusión de los tests mentales y en el establecimiento de laboratorios de psicología experimental. Cattell compartía la idea con Galton de que era factible obtener una medida de las funciones intelectuales mediante tests de discriminación sensorial y tiempo de reacción, estos intentos de Cattell se vio fomentado también por el hecho de que las funciones simples pueden medirse con precisión y exactitud, mientras que el desarrollo de medidas objetivas para funciones más complejas aparecía en aquel momento como una tarea casi inaccesible (7).

En Alemania algunos Investigadores también contribuyeron en el desarrollo de los tests como por ejemplo: Munsterberg, Ebbinghaus y Carl Wernicke, este último desarrolló investigaciones sobre localización cerebral y planteó un conjunto de preguntas destinadas a detectar deficiencias mentales.

Casi al mismo tiempo, en Francia, Alfred Binet, Victor Henri y Theodore Simon desarrollaban métodos para estudiar diversas funciones mentales. Encontraron la clave para medir la inteligencia al orientar su trabajo a procesos mentales superiores en lugar de las funciones sensoriales simples. Su trabajo culminó en 1905, con la Escala Binet-Simon, su aportación era una nueva combinación de las pruebas ya existentes, escogidas y ordenadas por su nivel de dificultad. La escala de 1905 se considera como la primera prueba de inteligencia moderna, en la que se demostró algún interés por los datos normativos, sirvió para diagnosticar en forma objetiva los grados de retraso mental en niños, y se convirtió en el prototipo de escalas subsiguientes para evaluar la habilidad mental (6).

En 1916 Izard y Simon crean una escala integrada por un conjunto de pruebas para niños de 3 a 24 meses en la que se fijan diferentes etapas evolutivas determinadas por el desarrollo perceptivo, indicándose el desarrollo intelectual del infante, esta fue una de las primeras pruebas que estableció etapas evolutivas en la infancia.

Con la introducción de las escalas de Binet-Simon, el movimiento de pruebas empezó a florecer en E.U.A. Lewis Terman estandarizó las escalas Binet-Simon en 1916 y con Maud Merrill las revisó en 1937 y 1960. Sin embargo en 1917, Robert Yerkes señaló ciertas deficiencias en el formato de las escalas de edades y por consiguiente desarrolló su propia escala de reactivos. Según Yerkes, las pruebas para la escala de edades se seleccionan con base en las proporciones de éxitos y fracasos de determinados grupos de diferentes edades, este procedimiento asume que importantes formas de conducta aparecen en los diferentes puntos del desarrollo (6).

En E.U.A. Wechsler-Bellevue estudió las pruebas estandarizadas disponibles a finales de la década de 1930, y seleccionó 11 sub-escalas diferentes para formarla. La norma que Wechsler adoptó en su búsqueda de sub-escalas fue la naturaleza global de la inteligencia. Concebía la inteligencia como una parte del todo de la personalidad misma. Las escalas Wechsler fueron construidas para considerar los factores que contribuían a la inteligencia global efectiva del individuo. Sin embargo, no se hicieron intentos para diseñar una serie de sub-escalas para medir habilidades primarias, ni ordenar las sub-escalas en una jerarquía de importancia relativa. El C.I. global obtenido de esta escala representaba un índice de habilidad mental general (6).

Los estudios estadísticos sobre la naturaleza de la inteligencia habían explorado las interrelaciones entre las puntuaciones obtenidas por numerosas personas en una gran variedad de tests. El psicólogo inglés Charles Spearman inició dichas investigaciones durante la primera década del presente siglo, en 1927 propuso la Teoría de dos factores para explicar los tipos de correlaciones que observó en las pruebas colectivas de inteligencia.

Sin embargo, el estudio del niño continuaba siendo un campo menor y alejado aun del análisis de la experimentación, seguía aun basado en la observación y apoyado en el empleo de pruebas estandarizadas. En los EE.UU. Arnold Gesell desarrolla una escala para la valoración del diagnóstico del desarrollo infantil, basado en una visión maduracionista en la que el ambiente tiene una influencia secundaria; siguiendo sus criterios Brunet-Lezine publicó la escala de desarrollo de la primera infancia. En 1933 Bayley construye las escalas de desarrollo infantil de amplia utilización en el campo clínico y en la investigación.

En esa misma época en Europa Jean Piaget, con una concepción diferente a la maduracionista, respondiendo al modelo genetista-cognitivo plantea la evaluación infantil en términos de procesos, valiéndose del método clínico-experimental de exploración, las pruebas más representativas basadas en la teoría piagetana son las de Casati y Lezine publicadas en Francia en 1968 y las de Uzgiris-Hunt publicadas en 1975 en los EE.UU (7).

En estos últimos tiempos el uso de la computadora en psicometría ha tenido un amplio desarrollo aplicándose en las distintas etapas de la evaluación y la mayoría de los tests psicológicos cuentan con versiones computarizadas, la Asociación de Psicólogos Americanos cuenta con una amplia variedad de programas para administrar y/o interpretar tests por computadora. Se han diseñado también una variedad de paquetes estadísticos que incluyen rutinas y menús para resolver problemas de investigación en este dominio (8).

La psicometría proporcionó las bases científicas a las ciencias del comportamiento humano planteando teorías acerca de cómo se miden las diferencias individuales y cómo se explica y fundamenta dicha medición. En los últimos tiempos esta disciplina con sus nuevos métodos e instrumentos ha permitido medir de manera objetiva, válida y confiable los cambios que se producen a lo largo del desarrollo de la conducta del niño, estableciendo patrones o normas de comportamiento promedio en cada una de las etapas del desarrollo del niño para su estudio e investigación (3).

## **1.2.- Generalidades de los métodos e instrumentos para la evaluación del desarrollo infantil.**

Como consecuencia de los cambios en la concepción infantil, el estudio sistemático del desarrollo del niño comenzó en épocas muy recientes, debido a preocupaciones diversas, como son: el de tipo práctico relacionado con la educación, o el de tipo médico interesados en la salud y bienestar del niño o el de tipo científico como los estudios estadísticos que se iniciaron a finales del siglo XIX y que hoy son muy frecuentes, lo cual ha permitido realizar evaluaciones sistemáticas y objetivas, que fundamentan y justifican que se pueda hablar de una Evaluación Infantil dotada de una entidad propia y con un tratamiento específico enfocado al estudio del niño.

### **1.2.1 Razones para la evaluación infantil.-**

Los motivos por los cuales los investigadores han centrado su interés en el estudio de la infancia, se puede resumir en 3 razones:

- **El cambio acelerado.-** Los primeros años de vida del niño es un periodo en el que ocurren muchos cambios en el desarrollo que en cualquier otro, es aquí donde la psicología evolutiva o del desarrollo ha permitido explicar estos cambios acelerados, recordemos que la psicología evolutiva se ocupa de los cambios de comportamiento que se presentan en el transcurso de la vida así como de las diferencias y semejanzas interpersonales respecto a dichos cambios (9).
- **Influencia a largo plazo.-** Otra razón para el estudio infantil es que los acontecimientos y experiencias de los primeros años influyen en el posterior desarrollo del individuo (10).
- **Procesos adultos complejos.-** Los investigadores que intentan comprender las conductas complejas de los adultos encuentran útil examinar como son esas conductas durante los periodos de desarrollo y es en el niño en crecimiento, en su desarrollo de habilidades y capacidades en que los investigadores han aprovechado este hecho

como ayuda para comprender los mecanismos de los diferentes aspectos funcionales de la conducta adulta (10).

### **1.2.2 Pautas para la evaluación infantil.-**

Cada niño tiene una variedad de capacidades que pueden evaluarse por diferentes medios ya sean cuantitativos como cualitativos, el objetivo es evaluar las capacidades y las limitaciones del infante. Las siguientes pautas constituyen un fundamento importante para la evaluación infantil:

- a) La evaluación del niño ha de contemplar una variedad de instrumentos y de informantes que permitan obtener la información necesaria sobre el desempeño del niño, ya que las conductas de este en los primeros años de vida es muy cambiante debido a las condiciones biológicas y ambientales en las que se encuentra.
- b) Adecuación psicométrica de las pruebas a aplicarse.- Las pruebas con referencia a la norma son indispensables en la evaluación clínica y psico-educativa. Estas pruebas se estandarizan en un grupo definido al cual se le denomina Grupo de Norma y se colocan en una escala de manera que cada puntuación individual refleje un rango dentro del grupo. Estas pruebas se han desarrollado para evaluar áreas que incluyen: inteligencia, habilidades en lectura y aritmética, habilidades visomotoras, motricidad fina y gruesa y conductas adaptativas. Se recomienda que se empleen aquellas pruebas que satisfagan las normas estadísticas aceptables (confiabilidad y validez). Las pruebas con referencia a la norma proporcionan información acerca del funcionamiento de un niño en las áreas cubiertas por la prueba, brindan un índice para evaluar los cambios del desarrollo y los efectos terapéuticos (6).
- c) Normas evolutivas de comparación.- Cualquier medición que se hace de un niño tiene que ir referida a un determinado nivel de edad, ya que la variación de la medida en función de la variable edad es esencial.

- d) Atención al entorno del niño.- Para entender y evaluar la conducta de un niño es necesario saber cuales son las características de su entorno. La plasticidad del niño le hace especialmente sensible a la influencia de este, y por otra parte el aprendizaje social es una de las vías más potentes para la adaptación al medio, sobretodo en los primeros estadios de su vida.

### **1.2.3 Categorías para el análisis de los métodos empleados en la evaluación infantil.**

Cuando se pretende sistematizar las posibles formas de evaluación infantil surgen una serie de interrogantes, en muchas ocasiones la evaluación de distintos aspectos o dimensiones se lleva a cabo con diferentes metas, marcos teóricos y utilizando distintos instrumentos de medición, lo cual le da complejidad a la evaluación del desarrollo infantil (38).

Para realizar una evaluación infantil debemos tener en cuenta las siguientes categorías:

- a) La evaluación puede utilizar dos tipos de métodos:
- Métodos objetivos como la observación, la entrevista, los cuestionarios y los tests.
  - Métodos subjetivos como escalas de apreciación, lista de adjetivos etc.
- b) Confiabilidad de la medición que la prueba o pruebas que se vayan a emplear sean capaces de proporcionar evidencia de que las puntuaciones obtenidas son consistentes a lo largo del tiempo y los examinadores, es decir, que son pruebas confiables.
- c) Por la característica a evaluar pueden ser:
- Características físicas en donde se distinguen los temas genéticos, sensoriales, neurológicos y endocrinológicos.
  - Características Psicológicas; que incluye las dimensiones personal, intelectual y social (38).

- d) En cuanto al ambiente una evaluación puede ser familiar, escolar o contextual.
- e) De acuerdo al propósito u objetivo de la evaluación, sobresalen las categorías educativo, clínico y de investigación.
- f) Por su extensión se pueden evaluar procesos o funciones aisladas o varios aspectos del desarrollo.
- g) Por su aproximación diagnóstica se pueden emplear dos tipos de pruebas las pruebas de tamizaje o screening y las pruebas diagnósticas (38).

El interés por la evaluación psicológica a los niños en los últimos años deriva de la importancia que tiene para el diagnóstico precoz de las alteraciones en el desarrollo, tomando en cuenta que las diferentes respuestas de los niños ante los mismos estímulos provienen de dos fuentes de variación: las competencias del recién nacido y las variaciones ambientales en interacción constante, con el fin de proveer de intervenciones preventivas y/o terapéuticas, lo cual ha sido posible con la participación de diversas ciencias como la psicología evolutiva, psicopatología, etología entre otras, que han dado rigor empírico a la evaluación del desarrollo en la primera etapa de la vida humana. Dos aspectos han sido determinantes para la evaluación en este campo:

- La toma de conciencia del valor que tienen estos primeros años para el desarrollo del niño, así se plantean diferentes teorías explicativas con respecto al desarrollo infantil.
- El surgimiento de inventarios de conducta, los que derivaron en escalas evolutivas, surgiendo posteriormente los registros de conductas y las escalas de desarrollo.

Una forma de clasificación de los instrumentos a utilizarse en la evaluación infantil es aquel que responde a criterios y técnicas psicométricas como las Pruebas Diagnósticas que establecen un criterio general de normalidad y permiten obtener Cocientes de desarrollo (CD), miden la cantidad de

comportamientos que se adquieren a una edad determinada, considerando dos elementos para ello: las "conductas blanco" (comportamientos considerados críticos del desarrollo en el área que se intenta destacar y cuya constancia, bajo diversas condiciones socioculturales tiende a permanecer) y los patrones de referencia poblacionales con las que se cotejan constantemente. Un ejemplo de este tipo de instrumentos son las Escalas de desarrollo que toman en cuenta el criterio estadístico de frecuencia de una conducta a una determinada edad estableciendo como normativo y esperable lo más usual, es decir, la adquisición, dominio y autonomía por parte de los sujetos, de conductas acaecidas a determinada edad; entre las que destacan la Escala de desarrollo de Gesell, las Escalas Bayley de desarrollo infantil entre otras (7).

En esta categoría se encuentran también las Pruebas de tamiz o de filtro empleadas para estimar el desarrollo e identificar a los niños en riesgo y que requerirán otros procedimientos más intensivos de diagnóstico, estos instrumentos se califican con tres posibilidades: normal, dudoso y anormal o alterado. Para conformar los reactivos empleados como tamiz se parte de identificar aquellas conductas blanco con mayor sensibilidad y especificidad, buscando indicadores confiables de riesgo para determinadas alteraciones del desarrollo. Las Pruebas de Tamizaje o Screening, son pruebas psicológicas que permiten estimar el desarrollo del niño y tienen por objetivo identificar a los que están en riesgo y que requerirán otros procedimientos más intensivos de diagnóstico. Entre estas pruebas podemos mencionar la prueba de tamizaje del desarrollo Denver II, CAT/ CLAMS entre otros (11)

Existen varias pruebas filtro para la detección y seguimiento de las alteraciones del desarrollo como la Escala de desarrollo Denver II, la Escala observacional del desarrollo entre otros, las características de estas pruebas tamiz dependen del objetivo para las que se diseñan:

- a) Aplicación masiva
- b) Vigilancia del niño sano
- c) Detección de riesgos determinados
- d) Evaluación de programas de fomento a la salud o cambios en las condiciones económicas, de crianza, hábitos etc.

La selección de una prueba tamíz en el contexto del seguimiento del niño sano y para la detección de riesgos, incluye ciertos criterios específicos, como:

- a) Considerar el contexto de seguimiento de la prueba, esto es que refleje la evolución y cambios en el desarrollo de las conductas blanco.
- b) Que sean útiles como elementos orientadores del manejo del desarrollo en aspectos de promoción de la salud.
- c) Su capacidad para orientar la necesidad del apoyo diagnóstico y terapéutico en condiciones de alteración.
- d) Que tengan buena sensibilidad y especificidad (validez de la prueba) con respecto a la detección de riesgos para el desarrollo.
- e) Que tengan buena capacidad predictiva (valor positivo y negativo) para secuelas permanentes.

## **II.- VALIDACIÓN DE LAS PRUEBAS PSICOLÓGICAS.**

### **2.1.- Validez de los instrumentos de medición**

La validez de los instrumentos de medición tiene que ver con lo que miden y qué tan bien lo hacen, nos indican qué se puede inferir a partir de sus resultados. El rasgo medido por determinada prueba sólo puede definirse mediante el examen de las fuentes objetivas de la información y las operaciones empíricas utilizadas para establecer su validez. La validez no puede expresarse en términos generales, no es posible decir que es "alta" o "baja", sino que hay que referirse al uso particular para el que se planea utilizar el instrumento (12).

### **2.2.- Tipos de validez**

Se pueden utilizar diferentes procedimientos para averiguar la validez de un test, la elección de uno u otro procedimiento depende de los fines para los que se destina el test y de sus características intrínsecas.

Podemos mencionar los tipos de validación más importantes, como son:

**2.2.1. Validez de contenido.-** Supone la evaluación y examen sistemáticos del contenido de la prueba para determinar si los reactivos son o no representativos del campo que se pretende medir.

**2.2.2. Validez de constructo.-** Es el grado en que una prueba mide un constructo o rasgo teórico. La validez de constructo requiere de la acumulación gradual de diversas fuentes de información; cada constructo se deriva de las interrelaciones establecidas entre medidas conductuales y se forma para organizar y dar cuenta de las concordancias observadas en la respuesta. Cualquier dato que explique la naturaleza del dato considerado y sobre las condiciones que afectan su desarrollo y manifestaciones constituye una evidencia apropiada para su validación, como los cambios en el desarrollo específicamente en la diferenciación por edad, al contrastar la edad cronológica con los resultados de una prueba estas nos permitirán observar si existe un incremento progresivo conforme aumenta la edad (12).

**2.2.3. Validez por criterio.-** Este procedimiento de validación indica la efectividad de la prueba para predecir el desempeño del individuo en actividades específicas. La medida de criterio contra la que se validan los resultados del instrumento puede obtenerse aproximadamente al mismo tiempo que los resultados de la prueba o después de un intervalo establecido. Los estándares de evaluación utilizan estas relaciones temporales entre el criterio y la prueba como base para diferenciar entre la validación concurrente y la predictiva.

- Concurrente (o de diagnóstico); se refiere a, si los resultados de la prueba están o no relacionados con alguna medida de criterio disponible en la actualidad. Este tipo de validación es adecuada para las pruebas que se emplean para diagnosticar el estado actual más que para predecir los resultados futuros (12).
- Predictiva (o de pronóstico); se refiere a la correlación entre puntajes de prueba o rendimiento conforme a un criterio, con un intervalo de tiempo entre la administración de la prueba y el rendimiento conforme al criterio (6).

La validación con base a criterios externos, como es el caso de este estudio, se determina a través del acuerdo entre el instrumento criterio o estándar (Escala de desarrollo infantil de Gesell) y el instrumento a evaluar (Prueba de tamizaje del desarrollo Denver II), esto permite conocer de manera adecuada la capacidad predictiva del instrumento. Mediante este método se calcula la sensibilidad, especificidad y los valores predictivos positivos y negativos de las pruebas de tamizaje.

Las relaciones que se obtienen entre ambos instrumentos se establecen a partir de la forma en que se presentan los resultados según la tabla de contingencia que muestra el número de evaluaciones que coinciden con la calificación de la prueba a evaluar o de tamizaje, que para el presente estudio es la prueba de tamizaje del desarrollo Denver II con relación a la prueba estándar Escala de desarrollo infantil de Gesell.

Prueba estándar (Escala de Gesell)				
Prueba a evaluar (Denver II)		+	-	Total de casos
		(alterado)	(Sin alteración)	
	+	Verdadero Positivo (VP)	Falso Positivo (FP)	(VP+FP)
	-	Falso Negativo (FN)	Verdadero Negativo (VN)	(FN+VN)
Total de casos	(VP+FN)	(FP+VN)	(VP+FP+VN+FN)	

VP=verdadero positivo, VN=verdadero negativo, FP=falso positivo, FN=falso negativo

Cuadro 1. Tabla de contingencia para la comparación de los resultados de la prueba a evaluar (Denver II) con el criterio estándar (escala de desarrollo de Gesell)

### 2.3.- Indicadores estadísticos de validez

En 1947, Yerushalmy introduce los términos de sensibilidad y especificidad como indicadores estadísticos que evalúan el grado de eficacia de una prueba diagnóstica, estos indicadores miden la discriminación diagnóstica de una prueba en relación a un criterio de referencia, cuyos resultados se consideran la verdad y esperar resultados similares cuando son aplicados en diferentes países, grupos o ámbitos. Los indicadores estadísticos que evalúan el grado de eficacia o validez de un instrumento son:

**2.3.1. Sensibilidad de la prueba,** se define como la capacidad del instrumento para identificar correctamente aquellos individuos que están a riesgo de presentar alteraciones en el desarrollo. Se determina a través de la proporción de positivos verdaderos sobre el total de casos positivos determinada por el instrumento criterio o estándar.

$$S = (VP) / (VP+FN)$$

**2.3.2. Especificidad de la prueba**, es la capacidad que tiene la prueba para identificar correctamente aquellos individuos que no presentan riesgo de presentar alteraciones en su desarrollo, es decir se clasifica al individuo como normal. Se determina a través de la proporción de negativos verdaderos sobre el total de casos negativos determinada por el instrumento criterio o estándar.

$$E = (VN) / (VN+FP)$$

**2.3.3. Valor predictivo positivo**, indica la probabilidad de que el individuo detectado con riesgo por el instrumento realmente lo presente, se expresa mediante la relación entre los positivos verdaderos y el total de positivos detectados por la prueba a evaluar.

$$VPP = (VP) / (VP+FP)$$

**2.3.4. Valor predictivo negativo**, indica la probabilidad de que un individuo detectado sin riesgo, realmente lo sea, se expresa por la relación entre los negativos verdaderos y el total de detectados por la prueba a evaluar (13).

$$VPN = (VN) / (VN+FN)$$

**2.3.5. Prevalencia**, se estima a partir del total de sujetos diagnosticados como positivos (alterados) por la prueba Estándar Respecto al total de sujetos evaluados.

$$P = (VP+FN) / (VP+FP+VN+FN)$$

**2.3.6. Índice de validez o proporción correcta de aciertos**, se define como la proporción de individuos clasificados correctamente. En términos de la tabla 2x2 básica, el índice de validez responde a la siguiente fórmula:

$$IV = (VP+VN) / (VP+FP+VN+FN)$$

Feinstein demuestra cómo este índice depende, no solamente de la sensibilidad y la especificidad, sino también de la prevalencia de la enfermedad.

Como ejemplos de algunos estudios realizados en México en los que se evaluó la validez de pruebas de tamizaje empleando un criterio externo podemos mencionar el estudio en el que se utilizó como estándar de oro la Escala de desarrollo infantil de Gesell a un punto de corte 90 (PC 90). Se aplicaron las pruebas (Vanedela-. Gesell) a una muestra de 184 niños de uno a dieciocho meses del INP y del Centro de Investigación Materno Infantil del grupo de Estudios al nacimiento, residentes en comunidades urbanas del nivel socioeconómico medio-bajo del Distrito Federal. Los resultados del Vanedela fueron comparados con los del estándar determinándose para la prueba una alta sensibilidad en cada una de las edades de corte de la prueba siendo adecuada para detectar alteraciones en distintas áreas del desarrollo, especificidad variable para algunas edades y regular para otras, un valor predictivo positivo y negativo alto en algunas edades y regular para otras edades (14).

En la investigación realizada en Tlaltizapan Morelos se trabajo con una muestra de 687 valoraciones con una población de niños de 0-12 meses y al aplicar la escala de desarrollo infantil de Gesell como estándar de oro se obtuvieron un alto número de retrasos en el desarrollo (67%) y por cortes de edad (50-93%), a la edad de 3 meses se presenta el mayor número de retrasos en el desarrollo. Cuando se compararon ambos resultados de las dos pruebas (CAT/CLAMS-Gesell) se encontró una sensibilidad de la prueba CAT/CLAMS global de 38% y una especificidad de 93%. El valor predictivo positivo de la prueba global fue de 92%, el valor predictivo negativo de la prueba global fue de 42% (15).

### **III.- PRUEBA DE TAMIZAJE DEL DESARROLLO DENVER II (DDST-II).**

La prueba de tamizaje del desarrollo Denver II, es el instrumento más utilizado para examinar los progresos en el desarrollo de los niños. Desde el diseño y publicación de la primera versión en 1967 la prueba ha sido utilizada en varios países del mundo, lo que indujo a que la prueba fuera revisada, surgiendo la versión DDST-II, que es la que actualmente se utiliza (16).

#### **3.1.- Antecedentes Históricos**

La Prueba de tamizaje del Desarrollo de Denver en su primera versión fue publicada en 1967 para ayudar a los cuidadores de la salud a detectar los problemas potenciales del desarrollo en niños. Desde su publicación original, el Denver ha sido adaptado para su uso y estandarizado en varios países, por ejemplo en Argentina se realizó el estudio cuyo objetivo fue la conformación del instrumento denominado Prueba Nacional de Pesquisa (PRUNAPE) en base a la prueba de tamizaje Denver II, se trabajó con una muestra nacional de 3573 niños sanos del país. Las percentilas seleccionados de la edad de cumplimiento de las conductas y sus respectivos intervalos de confianza fueron calculados por regresión logística, utilizando la edad o alguna transformación de la edad como variable independiente y el cumplimiento o no cumplimiento de la conducta como variable dependiente y fue sometida a un proceso de validación, en el cual se estimó una sensibilidad y especificidad del 80% y 93% respectivamente; un valor predictivo positivo y negativo del 94% y 78% (17).

En la investigación para la estandarización de la prueba de tamizaje del desarrollo Denver (DDST) en Arabia y África, se valoró a 936 niños de Oriente Medio y Norte de África. La muestra se dividió en tres grupos de edad desde el nacimiento a los seis años. En consecuencia, una nueva forma de detección DDST fue diseñado y presentado para la población estudiada (18).

En Trinidad y Tobago se estandarizó la prueba de tamizaje Denver II con una muestra de 969 niños sanos de 3 semanas de nacidos hasta los 6 años, los resultados indicaron que en el área de lenguaje las conductas tienden adquirirse más tardíamente en comparación con la muestra de estandarización

del Denver II y en el área motor grueso las conductas tienden adquirirse antes en comparación con la muestra del Denver II (24).

Como resultado de su amplia difusión y utilización, la prueba fue revisada por su autor el Dr. William Frankenburg en 1988 en un nuevo estudio en el estado de Colorado, EUA. y se culminó con la creación del Denver II (16).

Los motivos por los cuales se procedió a su revisión fueron expuestos por el autor y consistían en:

- La necesidad de agregar más ítems de lenguaje;
- La adaptación de la normas poblacionales de 1967 por las de 1988;
- Las características específicas de los ítems de la prueba (como la dificultad en la administración y/o puntaje) y;
- La adecuación de la prueba en varios subgrupos (como grupos étnicos, sexos, niveles de educación maternal, y lugar de residencia cuando diferencias clínicamente significativas existían en las normas de los subgrupos).

Otro motivo fue la falta de sensibilidad de la prueba en predecir un estado posterior, como la ejecución en la escuela. Así como la forma en que era administrada y/o interpretada.

Las razones anteriores fueron fundamentales en la decisión de revisar la prueba, re-estandarizarla, modificar su interpretación y enfatizar el entrenamiento de los evaluadores y la periódica evaluación efectiva en la administración de la prueba (16).

A pesar de estas aparentes mejoras, el Denver II fue publicado sin información sobre su validez o precisión en la identificación de niños con o sin problemas, se refiere en el manual de entrenamiento que se establecieron los valores normativos para las conductas de la nueva versión.

En el proceso de revisión y re-estandarización del Denver I se hicieron varias modificaciones en los 105 ítems originales. Estas incluyeron la omisión de ítems debido a su limitado valor clínico o a la dificultad en su administración o

calificación. El Denver II, difiere de su predecesor en la adición de 20 nuevos ítems, la mayoría de los cuales fueron integrados en el sector de lenguaje expresivo y habilidades de articulación. La selección final de los 125 ítems de la prueba se basó en los siguientes criterios: facilidad de administración y puntaje, presentación de los ítems tanto para el niño como para el evaluador, confiabilidad inter-evaluadores, puntajes mínimos de "rechazo" y de "no oportunidad", diferencias mínimas en subgrupos y una progresión fluida paso a paso de edades en la cual el 90% de los niños puedan realizar las tareas propuestas.

Se utilizó una cuota de muestra que contenía mayor número de niños pequeños para asegurar igual precisión al determinar las edades en las cuales un 25%, 50%, 75% y 90% de los niños pudieran ejecutar cada uno de los ítems potenciales. Los niños fueron agrupados de acuerdo a los siguientes subgrupos demográficos: educación de la madre, etnicidad, género y lugar de residencia.

La muestra de la ciudad de Denver incluyó 1,039 niños y la muestra de los condados fuera de Denver 1,057 haciendo un total de 2,096 niños, los cuales aparentemente fueron valorados una sola vez (16).

### **3.2.- Descripción de la prueba**

El objetivo de la prueba es:

- el tamizaje de niños desde las dos semanas de nacimiento hasta los seis años de edad para posibles problemas de desarrollo.
- confirmación de sospecha en el desarrollo con una medición objetiva.
- monitoreo de niños con un alto riesgo de problemas de desarrollo.

El Denver II no es una prueba de inteligencia (IQ), no es una prueba que predice la habilidad futura de adaptación o la habilidad intelectual. No está diseñada para generar niveles de diagnóstico como problemas de aprendizaje, trastornos de lenguaje o alteraciones emocionales y nunca sustituirá una evaluación de diagnóstico o un examen físico. Además, la prueba está

diseñada para comparar la ejecución de un niño en una variedad de tareas con la ejecución de otros niños de la misma edad.

El Denver II consta de 125 ítems, los cuales están agrupados en cuatro secciones para evaluar las siguientes áreas o funciones:

- Área Personal-social (25 ítems) evalúa los hábitos de crianza que se expresan en las actitudes del niño para relacionarse con las personas de su entorno, por ejemplo: "dice adiós", "imita actividades caseras", "se viste sin ayuda", "se lava dientes con ayuda".
- Área Motor fino-adaptativo (29 ítems) valora la madurez neuromotriz del niño en sus habilidades de coordinación ojo-mano para la aproximación, prensión y manipulación de objetos y realizar dibujos; evalúa su capacidad de utilizar sus experiencias previas en la resolución de problemas de la vida diaria, por ejemplo: "toma dos cubos", "forma torre de cubos", "garabatea".
- Área de Lenguaje (39 ítems) comprende toda forma de comunicación visible y audible como la imitación y comprensión de lo que expresan otras personas, ejemplo: "imita sonidos", "parlotea", "señala dibujos", "dice papá, mamá".
- Área Motor grueso (32 ítems) evalúa las capacidades motrices del niño, estos ítems comprenden las reacciones posturales, equilibrio y movimientos corporales globales, por ejemplo: "sentado sin ayuda", "se para solo", "camina bien", "corre", "salta en un pie".

El material a utilizarse, es muy simple y de fácil adquisición:

Bola de lana roja, botones o pasas, sonaja, lápiz rojo, botella pequeña, campana, pelota de tenis, papel blanco, 8 bloques o cubos de 1 pulgada por lado, muñeca pequeña y biberón y taza de plástico (19).

La administración de la prueba puede durar entre 20 y 25 minutos, ya que es de rápida y fácil aplicación.

### 3.2.1 Calificación e Interpretación de la prueba.

Para iniciar la calificación de la prueba se calcula la edad exacta del niño y se marca sobre la hoja de evaluación, se evalúan todos aquellos ítems que son intersectados por la línea de edad y las que están ligeramente atrás de la línea de edad y los más cercanos hacia la derecha para establecer si el niño presenta algunas conductas adelantadas para su edad (tres ítems por área).

La calificación se determina dependiendo si la respuesta del niño cae dentro o fuera del rango normal esperado de éxitos de cada ítem para la edad, las normas en las que se basa la prueba indican cuando el 25%, 50%, 75% y el 90% de niños pasan cada ítem.

La calificación de cada ítem se realiza de acuerdo a las siguientes claves:

- P (pasó) si el niño realiza la tarea
- F (falló) si el niño no realiza una tarea que la hace el 90% de la población de referencia.
- NO (no oportunidad) si el niño no realiza la conducta por motivos ajenos a él.
- R (rechazó) si el niño por alguna situación no colaboró para la evaluación, automáticamente se convierte en un NO y debe evaluarse en la siguiente visita.

Una vez calificados los ítems estos se interpretan de la siguiente manera:

- Normal, cuando el niño pasa el ítem a la edad requerida.
- Adelantado, cuando el niño pasa la conducta que está a la derecha de la línea de edad.
- Precaución, falla la conducta cuando la línea de edad atraviesa entre las percentilas 75 y 90.
- Retraso, cuando el niño falla la conducta que está a la izquierda de la línea de edad.

Para la interpretación Global de la prueba se deberán tomar en cuenta los siguientes criterios:

- Normal, cuando pasa todos los reactivos o tiene como mínimo una precaución.
- Sospechoso, cuando presenta un retraso o dos precauciones (19).

Se han planteado otras formas de calificación, por ejemplo Glascoe (1992) propone los siguientes criterios para la interpretación general de la prueba:

- Normal, cuando el niño presenta las conductas requeridas para su edad o máximo presenta una precaución.
- Sospechoso, cuando presenta un retraso o dos o más precauciones.
- Alterado, cuando presenta dos o más retrasos.

En la investigación realizada en el Centro de estudios rurales de Tlaltizapan se implementó una nueva forma de calificación por área y de forma general (20).

Calificación por áreas:

- Normal, si no presenta retraso o no más de una precaución en el área.
- Sospechoso, cuando se presentan dos o más precauciones en el área.
- Alterado, cuando presenta un retraso o más en el área.

Calificación general de la prueba:

- Normal, ningún área con alteración y hasta 2 áreas sospechosas.
- Sospechoso, cuando presenta un área con alteración o más de 2 áreas sospechosas.
- Alterado, cuando dos o más áreas presentan alteración (20).

### **3.3.- Estudio de Validación**

Debido a que la prueba Denver II, fue publicada sin tener evidencia de su validez o precisión en la identificación de niños con o sin problemas, el Dr. **Frances Page Glascoe** y sus colaboradores realizaron un estudio en el estado de Denver en los E.U.A. en 1992 para evaluar la precisión del Denver II comparando sus estándares con los de otras pruebas de tamizaje. Estas incluyeron sensibilidad (es adecuado tener aproximadamente 80%), especificidad (es preferible 90%) y valor de predicción positivo (es adecuado

70%). La población representativa estuvo conformada por 104 niños entre los 3 y los 72 meses de edad, que asistían a guarderías diurnas. Para determinar la presencia de problemas del desarrollo se les administró a los niños de manera individual pruebas para medir la inteligencia, el lenguaje, sus logros y su comportamiento adaptativo, un segundo examinador que desconocía los resultados de la batería de diagnóstico, administró el Denver II. Los problemas de desarrollo fueron encontrados en el 17% de los niños y el 83% fueron diagnosticados como normales. Los resultados del Denver II fueron: 38% de los niños recibieron puntajes normales, 26% tuvieron puntajes anormales, 33% tuvieron puntajes dudosos y 3% fueron considerados no verificables ya sea por rechazados o por otros comportamientos oposicionales.

Al analizar los datos se agruparon los puntajes de dudosos/no verificable con los resultados de normales, encontrándose una sensibilidad del 56%, una especificidad del 80% el valor predictivo positivo del 37%. Cuando se agrupó los dudosos/no verificables con los resultados anormales se obtuvo una sensibilidad del 83%, una especificidad del 43% y un valor predictivo positivo del 23%.

Dado que el Denver II identificó como anormal a muchos niños que no presentaban deficiencias en su desarrollo, se trató de ubicar con precisión las posibles fuentes de error.

Se agrupó a los niños por edades de 0-23 meses, de 24-47 y de 48-72 meses. La sensibilidad en todos los grupos etarios fue relativamente alta, con un 73% en el grupo de 25- 47 meses, un 100% en los grupos de 0-24 meses y de 48-72 meses de edad. La especificidad se mantuvo levemente baja: 30% para el grupo de 48-72 meses, 44% para los de 0-23 meses y 53% para los de 24-47 meses. El valor predictivo positivo fue del 18% al 31%. En consecuencia, el Denver II estableció un alto porcentaje de falsos positivos en todos los grupos etarios.

Estos resultados mostraron que los niños con problemas de desarrollo incluyendo a los que presentaban problemas leves, fueron detectados correctamente por el Denver II, es decir que sus niveles de sensibilidad se

acercan a los estándares de las pruebas de tamizaje, sin embargo, casi la mitad de los niños sin problemas de desarrollo también recibieron puntajes de sospechosos en el Denver II.

Debido al alto porcentaje de falsos positivos en el Denver II se buscó las fuentes del error en la prueba, se pudo observar que los niños que tuvieron calificaciones de retrasos o precauciones en las áreas de motor fino-adaptativo, personal social o motor grueso fueron a menudo los sobre referidos. Solo la presencia o ausencia de retrasos o precauciones en el área de lenguaje discriminó niños con problemas de niños sin problemas,

Las recomendaciones que se desprenden de este estudio son: la eliminación o reemplazo de ítems que fallan al discriminar el desempeño en pruebas diagnósticas; la revisión del criterio de puntajes, quizás dándole más peso al desempeño en el dominio del lenguaje y clarificando el criterio de referencia en relación al desempeño en todos los dominios. El protocolo de la prueba debería ser revisado para reflejar estos cambios también en el manual del evaluador (21).

#### **3.4.- Estudios realizados en México**

A continuación se describen algunos de los trabajos realizados en México empleando la prueba de tamizaje del desarrollo Denver en su primera y segunda versión, por ejemplo:

A) En Yucatán se llevó a cabo una investigación en la que se conformaron tres grupos integrados uno por familias de la clase trabajadora, otra por familias de clase media y un tercero por familias que provenían de la zona rural que hablaban Maya, cada familia acudían a diferentes tipos de atención médica, los trabajadores acudían a las clínicas del instituto de protección al niño, los de la clase media acudían a pediatras privados y los de la zona rural acudían al Instituto Mexicano de Seguridad Social.

La muestra estuvo conformada por 288 niños (96 en cada grupo) todos diagnosticados como niños sanos, cada niño fue evaluado con la Escala motora de Bayley y el DDST.

Los resultados fueron que de acuerdo a la prueba de Denver todos los niños resultaron normales a excepción de uno, se reportó una gran concordancia entre el DDST y la Escala motor de Bayley; no existieron diferencias en los niños en cuanto al desarrollo motor, así como entre los tres grupos.

Cuando las puntuaciones de los niños de Yucatán fueron comparadas con normas de los niños de los EUA fueron significativamente avanzadas para las diferentes edades solo para los doce meses presentaban retraso.

Los niños de Yucatán presentaron una mejor habilidad en funciones como coordinación ojo-mano, restablecer y alcanzar objetos, transferidos de una mano a otra y el uso de pinza fina a más temprana edad que lo observado en la muestra de los E.U.A.

La prueba Denver (DDST) fue incapaz de detectar niños con más de una DS. por debajo del promedio en la Escala motriz de Bayley (22).

B) En el estudio realizado en el Centro de Estudios Rurales de Tlaltizapán Morelos, se aplicó la prueba de tamizaje del desarrollo Denver II utilizando tres criterios para su interpretación y se identificó los componentes de la prueba que más se relacionan con la detección de alteraciones, comparando las diferencias en las edades de adquisición de referencia con los de la población de estudio (20).

La muestra estuvo conformada por 796 evaluaciones realizadas a 94 niños del poblado de Tlaltizapan-Morelos. Los niños fueron evaluados mensualmente con la Prueba de Denver II. Para la calificación total de la prueba se aplicaron tres criterios:

- a) El estipulado por el manual, en el que basta un retraso o dos precauciones para que la prueba se considere como sospechosa y normal cuando se cumplen todos los reactivos y no hay más de una precaución.
- b) El propuesto por Glascoe en 1992 se considera normal sino hay retrasos y máximo una precaución, sospechoso un solo retraso o dos o más precauciones y anormal con dos o más retrasos.

- c) El tercer criterio propuesto por el equipo investigador al cual denominaron "Propuesta", que consistió en calificar cada una de las áreas para luego hacer la calificación global de la prueba.

Los resultados obtenidos cuando se analizó la calificación final por área, a partir del número de reactivos con retrasos o precauciones, el área motora presentó mayor número de alteraciones (46%), lenguaje (31.3%), adaptativa y personal social con 13.3% y 14.1% respectivamente. Cuando se analizó el número de áreas alteradas con el criterio "propuesta" el 45.2% presentaron un área alterada, 28.6% las que no presentaron ninguna área alterada, el 26.1% de las valoraciones mostraron dos o más áreas alteradas.

Tomando en cuenta las alteraciones globales y utilizando el criterio de calificación del Manual, el 12.18% de las evaluaciones calificaron como normales y el resto como sospechosas. De acuerdo al criterio de Glascoe, el porcentaje de evaluaciones normales se mantuvo respecto al criterio anterior, pero el 51.4% quedó como sospechoso y un 36.4% como Anormales. El criterio Propuesta dio como normales al 28.64%, sospechosas al 45.22% y 26.13% de valoraciones alteradas.

Cuando se analizó el total de valoraciones realizadas a cada niño se encontró que de acuerdo al manual no calificó ni a un solo niño sin alteración, solo 2 casos presentaron menos de la mitad de sus valoraciones alteradas y el 53% de los niños resultaban con todas sus valoraciones alteradas. Con el criterio de Glascoe, el 6% de niños obtuvo sus valoraciones alteradas, se incrementó discretamente a 55% con más de la mitad de las valoraciones alteradas. El criterio Propuesta arrojó un 22% de niños con la mitad o todas sus valoraciones alteradas, 23.4% de los niños sin ninguna alteración y 54% con menos de la mitad de las evaluaciones alteradas.

Se observó que las secuencias del desarrollo en las percentilas 90 hasta los doce meses edad en que se hizo el corte en el área adaptativa es igual a la secuencia observada por los autores, aunque algunas de ellas se logren 2 meses después a lo esperado, las áreas lenguaje y motora presentan cambios en la secuencia de 6 y 7 reactivos respectivamente. Se observó también que

las diferencias de edades en que se logró la P90 mostro siempre retrasos en la población de estudio en todas las conductoras motoras exploradas, mientras que en las áreas adaptativas, lenguaje y personal social, las secuencias consecutivas de retrasos se interrumpen a los 8, 10 y 13 meses de edad respectivamente (20).

C) En el Centro de Investigación de Tlaltizapan Morelos se realizó un estudio para lo cual se conformaron 6 grupos de edad de RN-3 meses, 4-6 meses, 7-9 meses, 10-12 meses, 13-18 meses y 19-24 meses. En el estudio se tomaron en cuenta la interpretación de Glascoe para Denver II, los resultados obtenidos fueron que en todos los cortes de edad se encontró un alto porcentaje de casos sospechosos, detectando que los niños de la comunidad de Tlaltizapan no logran conductas que entre el 75-90% de niños de la población de referencia adquieren a la edad esperada, los resultados obtenidos en este estudio muestran similitudes con los resultados obtenidos en la investigación de Trinidad y Tobago (24) en cuanto a los retrasos de adquisición de las conductas, tanto en el área motora como de lenguaje, lo cual podría explicarse como resultado de las diferencias en las prácticas de crianza de los niños y del contexto social y cultural donde residen.

En los cortes de edad se encontró que del grupo de RN-3 meses fueron diagnosticados como normales el 22% y con riesgo de retraso (sospechosos y alterados) el 78%; en el grupo de 4-6 meses fueron diagnosticados como normales el 12% y con riesgo de retraso 88%; en el grupo de 7-9 meses fueron diagnosticados como normales el 33% y con riesgo de retraso 67%; en el grupo de 10-12 meses fueron diagnosticados como normales el 31% y con riesgo de retraso de 69%; en el grupo de 13-18 meses fueron diagnosticados como normales el 29% y con riesgo de retraso 71%; en el grupo de 19-24 meses fueron diagnosticados como normales el 60% y con riesgo de retraso el 40% (23).

### **3.5.- Estudios realizados en otros países.**

A) En Brasil en una muestra de 1363 niños nacidos a término se les aplicó la prueba Denver II, obteniéndose que el 34% de los niños presentaron sospecha

de atraso en el desarrollo, hubieron diferencias en las 4 áreas evaluadas, en el área motora 15% de sospecha de atraso, personal social 5.5%, motricidad fina 1.6% y en el área de lenguaje 1%. En el análisis multivariado se verificó que los niños que tenían mayor riesgo de sospecha de atraso en su desarrollo fueron los que nacieron con bajo peso, edad gestacional menor a 37 semanas, los que tenían más de 3 hermanos y los que habían recibido leche materna por menos de 3 meses o no habían sido amamantados; los niños que presentaron un índice de peso/edad los 6 meses igual o menor a menos de 2 DS. Tuvieron un riesgo 10 veces mayor de sospecha de atraso en el desarrollo. En la población estudiada el grupo más desfavorecido acumuló los factores que determinan mayor posibilidad de atraso en el desarrollo de los niños (25).

B) Se llevó a cabo en Venezuela una investigación con una población de 15 lactantes sanos de 9 meses de edad, sin antecedentes obstétricos y perinatales patológicos a los cuales se les aplicó la prueba de Denver II y la prueba de CAT/CLAMS hasta los 12 meses de edad. Una vez tabulados los datos los resultados fueron que 11 pacientes fueron clasificados por ambas pruebas como normales en su neurodesarrollo, 2 pacientes fueron diagnosticados con retardo por ambas pruebas y 2 pacientes clasificados por la prueba CAT/CLAMS con retardo en el neurodesarrollo no coincidió con la prueba Denver II. Se concluyó que ambas pruebas son útiles en la pesquisa de trastornos de neurodesarrollo, aplicándose la prueba de CAT/CLAMS en un tiempo significativamente menor que la prueba Denver II (26).

C) En el estudio realizado en Massachusetts Boston (27), se comparó una muestra de 86 niños provenientes de familias sin hogares vs. 134 niños provenientes de familias con hogares, los cuales estaban dirigidos solo por mujeres (madres de familia), en ambos grupos las madres fueron entrevistadas y se aplicó tests estandarizados a las madres y a los niños (Denver II, Inventario de depresión para niños, Escala de ansiedad manifiesta para niños, Listado de conductas de Simons y Listado de conductas para niños). Los resultados del Denver II indicaron que el 54% de los niños de pre-escolar sin hogares presentaron mayor retraso en el desarrollo en comparación con el 16% de los niños con hogares.

#### **IV. ESCALA DE DESARROLLO INFANTIL DE GESELL.**

##### **4.1.- Principios básicos**

Comportamiento o conducta.- Son las reacciones del niño, sean estas reflejas, voluntarias, espontáneas o aprendidas (parpadear, asir un objeto). Así como el cuerpo del niño crece, su conducta evoluciona, a medida que el SNC se modifica bajo la acción del crecimiento, la conducta se diferencia y cambia.

Desarrollo.- Es un proceso continuo y se puede examinar a través del examen de los modos de conducta.

Modos de conducta.- Son respuestas del sistema neuromotor frente a una situación específica. Resultan del proceso del desarrollo integral en ordenada progresión de etapas sucesivas, estos productos conductuales son la consecuencia de la continua interacción entre la dotación genética y el ambiente.

Nivel de desarrollo.- La conducta en el niño aumenta y asume formas características por medio de un proceso continuo que se inicia con la concepción y procede mediante ordenada sucesión, etapa por etapa, representando cada una de ellas un "nivel de madurez. Los modos de conducta son expresiones sintomáticas de cierto nivel de desarrollo (28).

Normas de conducta.- Los modos de conducta presentan en cada nivel variaciones individuales, pero el estudio de gran número de niños ha demostrado que los modos de conducta se adhieren estrechamente a un promedio central y que su orden de aparición (la sucesión del desarrollo) es significativamente uniforme. Esos estudios controlados del curso normal promedio del crecimiento de la conducta ha permitido establecer criterios de madurez "normas tipificadas de conductas". Estos modos de conductas tipificadas se emplean en el examen evolutivo para identificar y valorar, por comparación la conducta que un cierto sujeto exhibe. Gesell estudió por intervalos regulares la totalidad de las conductas significativas en el proceso del desarrollo desde las 4 semanas a los 42 meses (con pausas de un mes entre 4 semanas y 56 semanas; pausas de tres meses entre los 15 meses y los

24 meses y pausas de un semestre entre los 2 años y los 5 años) siendo un total de 21 etapas (niveles de madurez).

Edades claves.- Del total de niveles de madurez se seleccionaron algunos pocos, los que del curso longitudinal del desarrollo en los 3 primeros años mejor servían como muestras y puntos de referencia a los fines diagnósticos.

Se establecieron 8 edades claves que son: 4, 16, 28 y 40 semanas; 12, 18, 24 y 36 meses, ellas representan periodos integrativos y de modificaciones principales en los centros de organización.

Escala de la conducta.- La serie graduada de las normas de conducta que corresponde a cada una de las 8 edades claves, entre las 4 semanas y los 3 años constituye una vara de medida, una "escala de los niveles de desarrollo" que (a la manera de los baremos en psicometría) permite juzgar la edad en términos de conducta y la conducta en términos de edad (28).

Diagnóstico del desarrollo.- Es la traducción de los valores de conducta en valores de edad. Se logra mediante la observación discriminada de las formas de conducta de un sujeto y de su estimación por comparación con las normas tipificadas, el diagnóstico del desarrollo es la valoración de los modos de conducta observados en un sujeto, dado en términos de modos de conductas normalizados; el resultado de una confrontación entre las observaciones hechas y las normas estandarizadas. Dado que los modos de conductas constituyen síntomas, un diagnóstico del desarrollo es en esencia la apreciación de la madurez del S.N. con la ayuda de normas de conductas (28).

#### **4.2.- Descripción de la prueba**

Naturaleza de examen evolutivo.- La escala de Gesell es un instrumento ideado para estimar la madurez e integridad del sistema nervioso del niño apoyada en las etapas de la conducta, información histórica y datos actuales del niño. Nos permite la diferenciación y la exploración detallada de respuestas normales y anormales.

Para el examen de la conducta Gesell ideó un repertorio de pruebas funcionales que se aplican en posición supina, sentado, con sostén, sobre la

mesa, prueba de dibujo, pruebas verbales y en postura libre; que estimulan rendimiento motriz en el niño sometándolo así a una observación adecuada y oportuna del conjunto total-transversal y longitudinal de la conducta en cada uno de los 5 campos y en las 8 edades claves.

El examen evolutivo hace una revisión completa del comportamiento y la eficiencia de la conducta con vistas a investigar el sistema nervioso en su calidad de generador del mecanismo de la conducta.

Finalidad del examen evolutivo.- Responde a 2 objetivos:

1.- Describir precozmente defectos y desviaciones del desarrollo que de otra manera escaparían totalmente o demorarían en el diagnóstico. Sirve para revelar lesiones, defectos y retardos en la organización del SN. En el examen evolutivo la conducta infantil puede mostrarse normal, exhibir un retardo simple o desviaciones que señalen defectos en la estructura neural o disminuciones en el funcionamiento neuromotriz.

2.- Facilitar un mejor consejo orientador en aspectos prácticos del tratamiento, del cuidado infantil y de las relaciones paterno-filiales, pues brinda un mayor conocimiento de las características significativas de la personalidad del niño a tratar.

El diagnóstico evolutivo comprende el examen de 5 campos de conducta, representativos de los aspectos del crecimiento:

Área Adaptativa.- Evalúa la organización de los estímulos, percepción de relaciones, descomposición de totalidades en sus partes componentes y la reintegración de éstas. Coordinación de movimientos oculares y manuales para alcanzar y manipular objetos, habilidad motriz en la solución de problemas prácticos. La conducta adaptativa es la precursora de la inteligencia, que utiliza la experiencia previa para la solución de nuevos problemas.

Área motriz gruesa.- Valora la coordinación motora de los movimientos del cuerpo, el equilibrio de la cabeza, sentarse, pararse, gatear, caminar etc.

Área motriz fina.- Evalúa el uso de manos y dedos para la aproximación, asir un objeto, colocarlo sobre otro, manipulación del objeto.

Área de lenguaje.- Evalúa toda forma de comunicación visible y audible, sean gestos, movimientos posturales, vocalizaciones, palabras, frases u oraciones. La conducta del lenguaje incluye la imitación y comprensión de lo que expresan otras personas.

Área personal-social.- Comprende las reacciones personales del niño ante el medio sociocultural en el que vive, por ejemplo control de la micción y defecación, capacidad para alimentarse, higiene y habilidades para el juego, colaboración y reacción adecuadas frente a la enseñanza y a las convenciones sociales.

El material empleado en la prueba es el siguiente:

Sonaja, aro colgante, 3 aros de colores, taza, cubos, bolita, botella, campana, tablero con 3 figuras geométricas, pelota chica, caja de pruebas con bloque cuadrado, libro de imágenes, papel y lápiz, laminas con dibujos, pelota grande, formas coloreadas, figuras de rombo y cruz, dibujo de hombre incompleto (28).

#### **4.2.1 Administración de la escala.-**

La prueba consta de 24 grupos de edad conformados de la siguiente manera: 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, 56 semanas y 15, 18, 21, 24, 30, 36, 42, 48, 54, 60 meses, cada grupo de edad explora las cinco áreas de conductas. Para iniciar la valoración se debe seleccionar el grupo de pruebas más próximo a la edad cronológica y establecer el nivel más alto de habilidad del niño para obtener su máximo rendimiento.

Se debe realizar un interrogatorio a la madre o acompañante del niño sobre los 5 campos de conducta de manera progresiva, es decir, avanzando desde la zona evolutiva en la que se ubica el niño hacia la inmediata superior hasta establecer el límite máximo de habilidades del niño.

Para la presentación de los materiales se recomienda iniciar la prueba con la presentación de los objetos manipulables como los cubos y dejar las pruebas posturales al final de la evaluación con miras a reducir su efecto alterante (28).

#### **4.2.2 Calificación e interpretación de la escala.-**

El protocolo de examen contiene un cuadro de modos de conductas características, ordenadas en campos y niveles correspondientes a la edad del niño y dos columnas más, en la primera columna denominada nivel de desarrollo (N) se registran los diferentes signos que representarán el nivel de madurez del niño de acuerdo a la presencia o ausencia de la conducta:

El signo más (+) cuando el niño presenta claramente la conducta examinada.

El signo menos (-) cuando el niño no presenta la conducta examinada.

Los signos (+-) cuando el niño presenta la conducta examinada de forma incipiente, no se ha integrado totalmente.

La letra (N) cuando el niño no muestra la conducta esperada pero en su lugar presentó un modo más maduro.

En la segunda columna de observación (O) se anotan características observadas:

A (anormal) cuando un modo de conducta está presente o es incipiente pero su expresión es anormal.

Inc. (incapacidad) cuando una conducta no puede cumplirse por causa de un impedimento sensorial.

R (rechazo) cuando el niño rechaza una situación determinada.

X indica que no se dispone de información de un campo en particular por omisión al momento de la evaluación o por falta de datos históricos.

En el caso de algunos reactivos también pueden acreditarse estos signos en base a la información que nos brinde la madre.

Una vez que se ha calificado el cuadro de conductas, se procederá a calcular la Edad de desarrollo del niño en cada uno de los 5 campos de desarrollo, para luego calcular el Coeficiente de desarrollo el cual nos dará un índice aproximado del grado y ritmo del desarrollo que nos permita predecir la proporción de desarrollo normal que obtendrá el niño en el futuro.

El Coeficiente de Desarrollo (CD) nos proporcionará una estimación objetiva del estado evolutivo del niño y se refleja por la relación existente entre la Edad de madurez obtenida o Edad de Desarrollo (ED) y la Edad Cronológica (EC) real expresada en valor porcentual:

$$\text{CD} = \frac{\text{ED}}{\text{EC}} \times 100 = \frac{\text{Edad de madurez obtenida}}{\text{Edad cronológica}} \times 100$$

Las fluctuaciones en los C.D general ó específico (áreas) denotan la existencia de factores intrínsecos o extraños que el examinador deberá interpretar. Las fluctuaciones dentro de los 10 a 15 puntos pueden ser leves y deberse a la variabilidad normal de la conducta (28).

En México el procedimiento de evaluación de la escala de Gesell se tradujo al español y se condensó el campo motor grueso y fino en uno, estableciéndose así solo cuatro áreas de conducta a diferencia de la versión original de la escala, el orden de las áreas se presentan de la siguiente forma: área motora, área adaptativa, área de lenguaje y área personal-social. La escala presenta desde 1 mes de edad hasta los 42 meses un total de 385 ítems, pero para efectos de este estudio se consideró hasta los 36 meses haciendo un total de 374 ítems distribuidos en las 4 áreas de conducta de la siguiente manera: área motora 111 ítems, área adaptativa 113 ítems, área de lenguaje 70 ítems y en el área personal-social 80 ítems en total.

El uso de la escala se extendió en México, a partir de los años 50 utilizada principalmente por psicólogos; pero fue el Dr. Joaquín Cravioto quien empleó la escala en el campo de la investigación en su estudio sobre el impacto de la desnutrición infantil sobre la inteligencia, realizando algunas adaptaciones al instrumento como el introducir nuevos ítems en algunas áreas y edades y estableciendo una tabla de calificación en días ponderados, con el fin de que cada adquisición de conducta fuera computada para obtener el coeficiente general de desarrollo y el coeficiente de desarrollo en cada campo de

conducta. En la actualidad estas modificaciones en la escala se continúan utilizando en el Laboratorio de Seguimiento del Neurodesarrollo del Instituto Nacional de Pediatría ya que permite determinar el máximo logro alcanzado por el infante y conocer las dificultades a las que se enfrenta para adquirirlas, lo cual orienta al programa de intervención llamado cuidado integral (29).

La forma de calificación que se tomo en cuenta para la presente investigación fue la establecida por el Dr. Cravioto, en la cual cada ítem en cada edad adquiere un valor individual de acuerdo al número de ítems contenidos en la edad clave, tomando en cuenta que cada nuevo ítem positivo va adquiriendo un valor superior al anterior. La calificación total es la suma de los valores relativos expresados en días equivalentes a la edad de la muestra usada en la estandarización. En la calificación se toma en cuenta la potencialidad que la presencia de un nuevo ítem tiene en el establecimiento de posibles combinaciones con otros ítems presentes. El sistema consiste en dar una calificación mayor a aquel niño que presente un número mayor de ítems positivos, se da a cada uno de los ítems un valor en días para lo cual se suman los numerales. A la suma de los numerales se le da un valor arbitrario en días correspondiente a 30 días este se divide entre la suma de los numerales dando el valor de un ítem en días equivalentes (30).

En la evaluación se consideran la edad cronológica del niño, su control postural, su conducta espontánea y un interrogatorio previo al cuidador sobre los cuatro campos de conducta, se determina el grupo de pruebas que se aplicara en la evaluación a partir de que el infante realice todas las conductas (edad base) se sigue a los grupos superiores hasta encontrar el límite máximo y se termina la evaluación cuando el niño no realiza las conductas de una edad clave. Una vez que se califican los ítems con la clave positivo o negativo (presencia, ausencia) se establece la edad de línea base y se busca el equivalente de días en tabla y se cuenta los ítems positivos que hay en cada columna ubicando su equivalencia en días, este procedimiento se realiza área por área. Para establecer el CD por área se divide el total de cada área entre la edad del niño en días y se multiplica por cien. Para obtener el CGD se suman los porcentajes de las áreas y se divide entre cuatro.

La interpretación del nivel de desarrollo tomando en cuenta el CGD es el siguiente:

CD	Nivel de desarrollo
90 +	Normal
89-76	Retardo leve
75-65	Retardo moderado
50-26	Retardo severo

La Escala de Desarrollo infantil elaborada por Arnold Gesell y C. Amatruda es una prueba mundialmente difundida cuyos fundamentos teóricos se comprueban o se validan al aplicarla al campo apropiado, es un instrumento que abarca un amplio rango de edad desde el nacimiento hasta los cinco años. Durante más de veinte años sus autores estudiaron a una gran cantidad de niños normales y los observaron desde su nacimiento, examinándolos cada mes y después a intervalos más amplios. En esas condiciones, evaluaron veinticinco niveles consecuentes de edad, hasta llegar a los cinco años. Siendo, por lo tanto ésta una de las pruebas con mayor aplicación en el mundo, por sus registros sistematizados, por su rigor en el estudio longitudinal de casos y por el amplio número de sujetos estudiados con y sin alteraciones en su desarrollo y pudiendo establecerse a partir de estos registros evolutivos, edades promedio o medidas de crecimiento, es que se emplea esta escala como el estándar de oro para el presente trabajo de investigación (31)

## **V.- METODOLOGÍA.**

### **5.1.- Justificación**

En los últimos años el interés por el estudio del desarrollo infantil se a incrementado en las instituciones de salud pública con el objetivo de detectar tempranamente las alteraciones que se presenten y poder atender al niño antes de que se constituyan en problemas mayores de difícil manejo.

La Organización Mundial de la Salud señala que 5 de cada 100 niños en el mundo padecen alteraciones en el desarrollo, siendo un problema mayor en los países no industrializados, de tal forma que es recomendación internacional la vigilancia del niño sano y la detección temprana de las alteraciones del desarrollo. En México la Norma Oficial Mexicana 031-SSA2-1999 para la atención de la salud del niño, establece la obligatoriedad de dar seguimiento al desarrollo del niño bajo un esquema de tamiz.

Partiendo de la idea de que el tamizaje tiene 2 objetivos principales:

- Vigilar el desarrollo de los niños en un proceso continuo de seguimiento, con objeto de promover su salud y estimular el desarrollo en condiciones culturalmente sensibles e inclusivas del contexto familiar.
- Detectar desviaciones del desarrollo resultados de riesgos médicos o sociales y potencial expresión de secuelas que conducen a discapacidad.

Se hace necesario el empleo de un instrumento confiable para detectar riesgo de alteraciones en el neurodesarrollo basados en indicadores que permitan identificar apropiadamente el daño neurológico infantil y otras causas de alteraciones en el desarrollo, así como, llevar a cabo de forma confiable el seguimiento del desarrollo en niños normales sin saturar la capacidad de los recursos de diagnóstico con los que cuenta el sector salud, complementándose con la creación y puesta en marcha de programas de intervención tan pronto sea detectada la presencia de signos de desviación propiciando una respuesta oportuna y precoz, evitando el establecimiento de la secuela y patrones funcionales patológicos de difícil y costoso manejo para las familias . Por lo que

el uso de instrumentos debe proveer de evidencia que valide su capacidad de detección de los problemas de una forma adecuada.

### **5.2.- Planteamiento del problema**

Siendo la prueba Denver II un instrumento de tamiz muy utilizado en la población infantil a nivel mundial, aun se desconoce la sensibilidad y especificidad del instrumento en población infantil mexicana, lo cual nos llevó a dar respuesta con el presente trabajo de investigación a la siguiente pregunta: ¿Cuál es la validez por criterio externo de la prueba de tamizaje del desarrollo de Denver II, respecto a la valoración de la escala de desarrollo infantil de Gesell aplicado en lactantes de 0-2 años?

### **5.3.- Objetivo General**

Evaluar la validez por criterio externo de la Prueba de tamizaje del desarrollo de Denver II, con respecto a la Escala de desarrollo infantil de Gesell en niños de 0 a 2 años residentes en un poblado del Estado de Morelos.

### **5.4.- Objetivos Específicos**

- Evaluar la sensibilidad, especificidad y valores predictivos positivo y negativo de la prueba global respecto al CGD del estándar.
- Evaluar la validez según áreas del Denver II respecto al Coeficiente General y por áreas del estándar.
- Evaluar la validez según grupos de edad.

### **5.5.- Tipo de estudio**

Observacional, descriptivo, transversal, prospectivo.

## **VI.- MATERIALES Y MÉTODOS.**

### **6.1.- Población de estudio**

Lactantes de 0 a 24 meses de edad incluidos en el programa de vigilancia del desarrollo, pertenecientes al Protocolo de Investigación "Perfiles de desarrollo de niños sanos empleando indicadores de 4 pruebas tamiz".

### **6.2.- Criterios de Inclusión**

Lactantes de ambos géneros

Tener menos de 24 meses de edad en su primera valoración

Cubran criterio de encontrarse sanos en la fecha de valoración

Aceptación por parte de los padres o tutores de entrar al estudio, previa información de las actividades que se realizarán a su hijo/a.

### **6.3.- Criterios de Exclusión**

Diagnóstico de morbilidad o condición neurológica evidente expresada por el niño.

### **6.4.- Ubicación del estudio**

El estudio se realizó en las instalaciones del Centro Rural de Estudios de Tlaltizapan Morelos, del Instituto Nacional de Pediatría.

### **6.5.- Muestra**

Se consideró una muestra de 104 niños con 1,921 valoraciones, residentes del municipio de Tlaltizapan a los cuales se les dio seguimiento longitudinal en su desarrollo aplicándoles 4 pruebas psicológicas, de las cuales una es la Prueba de Tamizaje del desarrollo Denver II y se les aplicó la Escala de desarrollo infantil de Gesell como estándar de oro para la contrastación con la prueba Denver II.

Se conformaron 5 grupos de edad los cuales fueron: 0-6 meses, 7-12 meses, 13-18 meses, 19-24 meses, 25-30 meses, para evaluar la validez de la prueba en cada corte de edad.

### **6.6.- Variables**

Variables Generales:

Género

Edad

Grado de estudio del padre

Grado de estudio de la madre

Puesto en el trabajo del padre

Puesto en el trabajo de la madre

Ingreso familiar  
Porcentaje del ingreso destinado a alimentación  
Grado de hacinamiento

**Variables Independientes:**

Resultados en la calificación del test de Denver II  
Resultado global general y por áreas.  
Resultados en la aplicación de la prueba de Gesell  
Resultado global general y por áreas.

**Variables Dependientes:**

Resultado general en la aplicación de la prueba de Gesell  
Resultado por áreas de la prueba de Gesell

**6.7.- Instrumentos de evaluación del desarrollo**

Se emplearon dos instrumentos para evaluar el desarrollo de los niños, la Escala de desarrollo infantil de Arnold Gesell que evalúa cuatro áreas de desarrollo que son la motora, adaptativa, lenguaje y personal social. La calificación que se implementó fue el criterio establecido por el Dr. Joaquín Cravioto, en el que a los reactivos considerados como positivos se les establece su equivalencia en días de desarrollo con lo cual se establece la edad en días de desarrollo para cada área, calculándose un coeficiente general de desarrollo y por áreas según su relación con la edad cronológica.

Los criterios de calificación tomados en cuenta para la interpretación de la prueba fueron:

Normal: coeficiente general de desarrollo 90% o más; coeficiente por áreas de desarrollo se considera a partir de 76% o más.

Alterado: coeficiente general de desarrollo menor a 90%; coeficiente por áreas menor a 76%.

Se aplicó la Prueba de tamizaje del desarrollo Denver II creada por William Frankenburg, sus 125 reactivos se agrupan por áreas de desarrollo: motor grueso, motor fino-adaptativo, lenguaje y personal-social (19).

Para la interpretación general de la prueba tanto para la población total como para los grupos de edad se utilizó el criterio de Glascoe:

- Normal, cuando el niño presenta las conductas requeridas para su edad o máximo presenta una precaución.
- Sospechoso, cuando presenta un retraso o dos o más precauciones.
- Alterado, cuando presenta dos o más retrasos (21).

Para la comparación de los resultados globales con la prueba estándar, se definió para la prueba Denver II primero como negativos o sin alteración a los normales, mientras que los sospechosos y alterados en el criterio de Glascoe como positivos o alterados; luego se consideraron como negativos o sin alteración a los normales y sospechosos y como positivos solo a los alterados en el criterio de Glascoe quedando los resultados agrupados de la siguiente manera:

- a) normales vs. sospechosos y alterados que corresponde al criterio del manual (N/SyA).
- b) normales y sospechosos vs. alterados que corresponden al criterio de Glascoe (NyS/A).

Para la comparación de los resultados por áreas para la población total y por grupos de edad de la prueba Denver II se utilizó el criterio de Rivera:

Calificación por área:

- Normal, cuando se cumplen todos los ítems o presenta máximo una precaución en el área, que para el análisis del presente estudio se consideran negativos o sin alteración.
- Sospechoso, cuando se presentan dos o más precauciones en el área, que también fueron considerados como positivos o sin alteración.
- Alterado, cuando presenta un retraso o más en el área, considerados como positivos o alterados (20).

	Criterio del manual		Criterio Glascoe	
	Denominación	Tratamiento en el análisis de la validez	Denominación	Tratamiento en el análisis de la validez
<b>Ningún retraso o no más de una precaución</b>	NORMAL	Negativo o sin alteración	NORMAL	Negativo o sin alteración
<b>Un retraso o más de una precaución</b>	SOSPECHOSO	Positivo o alterado	SOSPECHOSO	Negativo o sin alteración
<b>Dos o más retrasos</b>	SOSPECHOSO	Positivo o alterado	ALTERADO	Positivo o alterado

**Cuadro 2. Clasificación de los resultados en la prueba Denver II para el análisis de validez en el presente estudio.**

### **6.8.- Procedimiento**

Se explicó a los padres de la comunidad los objetivos de la investigación y se les invitó a participar, los padres que se integraron al estudio se les tomó sus datos demográficos (nombre, edad, nivel de estudios, ocupación, sueldo etc.). A cada niño se le abrió un expediente previa evaluación pediátrica para establecer las condiciones de salud del lactante si estas eran las requeridas para el estudio se procedía aplicar la Prueba de tamizaje de desarrollo Denver II y la Escala de desarrollo infantil de Gesell.

A la población de niños se les valoró mensualmente por evaluadores estandarizados, quienes inmediatamente procedían a la calificación y registro de los resultados de las pruebas en los formatos respectivos, quedando archivados en los expedientes clínicos de cada niño.

Los resultados obtenidos de ambos instrumentos de estudio eran registrados, posteriormente, en la base de datos con los cuales se realizaron los análisis estadísticos correspondientes.

### **6.9.- Análisis estadístico**

Se realizó un análisis de frecuencia de acuerdos y discrepancias entre la Prueba de Tamizaje de Denver II y la Escala de desarrollo infantil de Gesell, empleando la Tabla de Contingencia de 2X2 para el cálculo de la sensibilidad, especificidad, valores predictivos positivos y negativos, se utilizó el Programa "Análisis Epidemiológico de datos Tabulados" (EPIDAT).

#### **6.10.- Ética**

Se contó con el consentimiento informado y suscrito de los padres al inicio de la investigación.

Se explicó a los padres los beneficios de la valoración periódica a sus niños, posibilitando detectar desviaciones en el desarrollo.

Se brindaba a los padres recomendaciones para el manejo adecuado de los niños que presentaban alguna alteración en su desarrollo.

En casos de franca alteración que señalaba algún tipo de riesgo para secuela en el desarrollo los casos eran canalizados para el manejo del neurodesarrollo y programa de intervención temprana por personal adscrito al Laboratorio de seguimiento del neurodesarrollo del Instituto Nacional de Pediatría y a su vez de ser necesario este los remitía a otra instancia de los servicios de salud disponibles en la comunidad.

## VII.- RESULTADOS

Se estudiaron un total de 104 niños desde recién nacidos hasta los 30 meses de edad, obteniéndose un total de 1,921 valoraciones sobre el cual se realizaron los análisis estadísticos correspondientes.

### 7.1. Características de la población.

**Cuadro 2. Características socio-demográficas de los padres.**

<b>Edad del padre</b>		<b>Edad de la madre</b>	
Promedio	28.1 años	Promedio	24.97 años
DS.	10 años	DS.	6.14 años
<b>Escolaridad del padre</b>		<b>Escolaridad de la madre</b>	
	%		%
1. Media superior completa	20.7	1. Media superior completa	24.7
2. Media básica completa	57.7	2. Media básica completa	47.7
3. Primaria completa	14.1	3. Primaria completa	20.3
4. Primaria incompleta	5.7	4. Primaria incompleta	5.3
5. Analfabeto	1.7	5. Analfabeto	1.7
<b>Grado cursado de escolaridad del padre</b>		<b>Grado cursado de escolaridad de la madre</b>	
Promedio	9.5 años	Promedio	9.49 años
DS.	3.25 años	DS.	3.09 años
<b>Posición en el trabajo del padre</b>		<b>Posición en el trabajo de la madre</b>	
	%		%
1. Funcionario o profesionista	2.7	1. Funcionario o profesionista	0.9
2. Empleado a nivel técnico (comerciante establecido o estudiante)	17.4	2. Empleado a nivel técnico (comerciante establecido o estudiante)	3.6
3. Trabajador manual asalariado (obrero, intendencia o limpieza)	23.8	3. Trabajador manual asalariado (obrero, intendencia o limpieza)	1.8
4. Trabajo manual independiente u hogar (mecánico, carpintero)	43.1	4. Trabajo manual independiente u hogar (mecánico, carpintero)	91.8
5. Subempleado o desempleado (vendedor Ambulante)	11.9	5. Subempleado o desempleado (vendedor Ambulante)	1.8

En cuanto a los datos demográficos tenemos que el rango de edad del padre es de 17 a 50 años, el nivel de escolaridad alcanzada mayormente observado es la Media básica completa (57.7%). En el caso de la madre el rango de edad es de 15 a 39 años, el nivel de escolaridad alcanzado más observado igualmente es la Media básica (47.7%). La principal actividad económica del padre es el Trabajo manual independiente (43.1%). En el caso de la madre la ocupación predominantemente es el hogar (91.8%).

**Cuadro 3. Características socio-económicas de las familias.**

<b>Número de salarios mínimos (ingreso familiar)</b>	
Promedio	4.44 salarios
DS	1.24
<b>Salario</b>	
	%
1. <2,000	7.4
2. <3,400	47.2
3. 3,400-6000 al mes	27.7
4. >6,000	17.5
<b>Salario neto</b>	
Promedio	4,168 pesos
DS	2,868
<b>Ingreso familiar destinado a la alimentación</b>	
	%
1. <30 %	2.7
2. 30-50%	49.5
3. >51%	47.7
<b>Número de miembros de la familia</b>	
	%
1. 3 o 4	69.2
2. 5	13.1
3. 6	8.7
4. 7	6.1
5. 8 o más	2.6
<b>Número de miembros de la familia</b>	
Promedio	4.12
DS.	1.44
<b>Número de personas por dormitorio</b>	
	%
1. 2 o menos	19.2
2. 3 -4	69.2
3. 5	7.8
4. 6	2.6
5. 7 o más	0.8
<b>Número de personas por dormitorio</b>	
	%
1. 1 persona	2.6
2. 2 personas	16.6
3. 3 personas	42.1
4. 4 personas	28.0
5. 5 personas	7.8
6. 6 personas	1.7
7. 7 personas	0.8
<b>Número de dormitorios</b>	
	%
1. 2 o menos	74.1
2. 3 o 4	25.0
3. 5 o más	0.8
<b>Número de dormitorios</b>	
Promedio	1.92
DS.	1.03
<b>Número de personas por cama</b>	
	%
1. 2 o menos	55.2
2. 3 o 5	31.5
3. >5 o suelo	13.1
<b>Número de camas</b>	
	%
1. 2 o menos	49.5
2. 3 o 4	29.2
3. 5	18.5
4. 6 o más	1.7

En lo referente a las condiciones de los hogares, las familias están compuestas predominantemente por 3 o 4 personas (69.2%) el número de dormitorios por vivienda es de 2 o menos en el 74.1% y el número de personas por dormitorio predominantemente es de 3 o 4 personas (69.2%); el número de camas por hogar es de 2 o menos en el 49.5%, seguido de 3 o 4 camas (29.2%), el número de personas por cama es de 2 o menos en el 55.2% seguido de 3 o 5 (31.5%).

El ingreso económico familiar en promedio es de 4.4 a 5 salarios mínimos siendo el salario neto promedio de 4,168 pesos al mes. El 41.5% de las familias destinan entre el 31-50% de sus ingresos para la alimentación familiar y el 47.7% destina más del 51% de sus ingresos para la alimentación familiar.

El promedio del gasto neto para la alimentación familiar es de 2,144 pesos.

\*Tipo de cambio 11 pesos = 1 dólar US; 1 salario mínimo = 1,950.

## **7.2. Indicadores de validez en la población general.**

### **7.2.1. Criterio de calificación del manual.**

A continuación se reportan los resultados obtenidos de la comparación de la prueba de tamizaje Denver II con la Escala de desarrollo infantil de Gesell la cual se consideró como estándar para la valoración de la validez por criterio externo del instrumento tamíz.

1.- Al contrastar los resultados globales de la escala de desarrollo infantil de Gesell (estándar) punto de corte 90(PC90) y los resultados globales de la prueba Denver II y que corresponden según los criterios del manual (N/SyA). Del total de 1921 valoraciones, la prueba estándar reportó 1121 valoraciones como alteradas, de las cuales la prueba Denver II reportó 1019 como alteradas y 102 como sin alteración, obteniéndose una sensibilidad de 90.9%, con un I.C. de 89.1- 92.6.

La prueba estándar reportó 800 valoraciones sin alteración, de las cuales la prueba Denver II reportó 313 valoraciones como sin alteración y 487 valoraciones como alteradas, obteniéndose una especificidad de 39.1%, con un I.C. de 35.6-42.5.

La prueba Denver II reportó 1506 valoraciones como alteradas de las cuales 1019 valoraciones fueron confirmadas por la prueba estándar, obteniéndose un valor predictivo positivo de 67.6%.

La prueba Denver II reportó 415 valoraciones como sin alteración de las cuales 313 valoraciones fueron confirmadas por la prueba estándar y 102 fueron reportadas como alteradas, obteniéndose un valor predictivo negativo de 75.4%. El índice de validez es de 69.3% y la prevalencia del riesgo para alteración en el desarrollo es de 58.3%.

**Tabla 1. Indicadores de validez de la prueba Denver II global (criterio manual) respecto al criterio estándar de Gesell (PC 90)**

Prueba de Gesell (estándar)				
Prueba Denver II	Alterado	Alterado	Sin alteración	Total
	Sin alteración	102	313	415
	Total	1121	800	1921

		IC (95%)
Sensibilidad	90.9%	89.1-92.6
Especificidad	39.1%	35.6-42.5
V. predictivo positivo	67.6%	65.2-70.0
V. predictivo negativo	75.4%	71.1-79.6
Índice validez	69.3%	67.2-71.4
Prevalencia	58.3%	56.1-60.5

### 7.2.2. Criterio de calificación de Glascoe.

2.- Al contrastar los resultados globales de la Escala de desarrollo infantil de Gesell (estándar) PC90 y la prueba Denver II y que corresponden al criterio de calificación de Glascoe (NyS/A). De un total de 1921 valoraciones, la prueba estándar reportó 1121 valoraciones alteradas, de las cuales la prueba Denver II reportó 895 valoraciones como alteradas y 226 valoraciones como sin alteración, obteniéndose una sensibilidad de 79.8%, con un I.C de 77.5-82.3.

La prueba estándar reportó 800 valoraciones sin alteración, la prueba Denver II reportó 462 valoraciones como sin alteración y 338 valoraciones como alteradas, obteniéndose una especificidad de 57.7%, con un I.C. de 54.2-61.2

La prueba Denver II reportó 1233 valoraciones como alteradas de las cuales 895 valoraciones fueron confirmadas por la prueba estándar, obteniéndose un valor predictivo positivo de 72.5%.

La prueba Denver II reportó 688 valoraciones como sin alteración de las cuales 462 valoraciones fueron confirmadas por la prueba estándar y 226 fueron reportadas como alteradas, obteniéndose un valor predictivo negativo de 67.1%. El índice de validez es de 70.6% y la prevalencia del riesgo para alteración en el desarrollo es de 58.3%.

**Tabla 2. Indicadores de validez de la prueba Denver II global (según Glascoe) respecto al criterio estándar de Gesell (PC 90)**

Prueba de Gesell (estándar)				
Prueba Denver II	Alterado	Alterado	Sin alteración	Total
		895	338	
	Sin alteración	226	462	688
	Total	1121	800	1921

		IC (95%)
Sensibilidad	79.8%	77.4-82.2
Especificidad	57.7%	54.2-61.2
V. predictivo positivo	72.5%	70.6-75.1
V. predictivo negativo	67.1%	63.5-70.7
Índice validez	70.6%	68.5-72.7
Prevalencia	58.3%	56.1-60.5

### 7.3.- Indicadores de validez por Grupos de edad.

#### 7.3.1.- Criterio de calificación de Glascoe.

1.- Al contrastar los resultados globales de la Escala de desarrollo infantil de Gesell (estándar) PC90 y los resultados globales de la prueba Denver II (NyS/A). para el grupo de edad de 0-6 meses y con un total de 356 valoraciones, la prueba estándar reportó 267 valoraciones alteradas, de las cuales la prueba Denver II reportó 224 valoraciones como alteradas y 43 valoraciones como sin alteración obteniéndose una sensibilidad de 83.9%, con un I.C. de 79.3-88.4.

La prueba estándar reportó 89 valoraciones sin alteración, la prueba Denver II reportó 37 valoraciones como sin alteración y 52 valoraciones como alteradas, obteniéndose una especificidad de 41.5%, con un I.C. de 30.7-52.3

La prueba Denver II reportó 276 valoraciones como alteradas de las cuales 224 valoraciones fueron confirmadas por la prueba estándar como alteradas, obteniéndose un valor predictivo positivo de 81.1%.

La prueba Denver II reportó 80 valoraciones como sin alteración de las cuales 37 valoraciones fueron confirmadas por la prueba estándar y 43 fueron reportadas como alteradas, obteniéndose un valor predictivo negativo de 46.2%. El índice de validez es de 73.3% y la prevalencia del riesgo para alteración en el desarrollo es de 75%.

**Tabla 3. Indicadores de validez de la prueba Denver II (criterio Glascoe) respecto al criterio estándar de Gesell (PC 90) para el grupo de 0-6 meses de edad.**

Prueba de Gesell (estándar)				
Prueba Denver II	Alterado	Prueba de Gesell (estándar)		Total
		Alterado	Sin alteración	
	Alterado	224	52	276
	Sin alteración	43	37	80
	<b>Total</b>	<b>267</b>	<b>89</b>	<b>356</b>
IC (95%)				
<b>Sensibilidad</b>		83.9%		79.3-88.4
<b>Especificidad</b>		41.5%		30.7-52.3
<b>V. predictivo positivo</b>		81.1%		76.3-85.9
<b>V. predictivo negativo</b>		46.2%		34.7-57.8
<b>Índice de validez</b>		73.3%		68.5-78.0
<b>Prevalencia</b>		75%		70.3-79.6

2.- Al contrastar los resultados globales de la Escala de desarrollo infantil de Gesell (estándar) PC90 y los resultados globales de la prueba Denver II (NyS/A). para el grupo de edad de 7-12 meses y con un total de 473 valoraciones, la prueba estándar reportó 311 valoraciones alteradas de las cuales la prueba Denver II reportó 263 valoraciones como alteradas y 48 valoraciones como sin alteración obteniéndose una sensibilidad de 84.5%, con un I.C. de 80.3-88.7

La prueba estándar reportó 162 valoraciones sin alteración, la prueba Denver II reportó 73 valoraciones como sin alteración y 89 valoraciones como alteradas, obteniéndose una especificidad de 45%, con un I.C. de 37.0-53.0

La prueba Denver II reportó 352 valoraciones como alteradas de las cuales 263 valoraciones fueron confirmadas por la prueba estándar como alteradas, obteniéndose un valor predictivo positivo de 74.7%.

La prueba Denver II reportó 121 valoraciones como sin alteración de las cuales 73 valoraciones fueron confirmadas por la prueba estándar y 48 fueron reportadas como alteradas, obteniéndose un valor predictivo negativo de 60.3%. El índice de validez es de 71% y la prevalencia del riesgo para alteración en el desarrollo es de 65.7%.

**Tabla 4. Indicadores de validez de la prueba Denver II (criterio Glascoe) respecto al criterio estándar de Gesell (PC 90) para el grupo de 7-12 meses de edad.**

Prueba de Gesell (estándar)				
Prueba Denver II		Prueba de Gesell (estándar)		Total
		Alterado	Sin alteración	
	Alterado	263	89	352
	Sin alteración	48	73	121
	<b>Total</b>	<b>311</b>	<b>162</b>	<b>473</b>
				<b>IC (95%)</b>
<b>Sensibilidad</b>		<b>84.5%</b>		<b>80.3-88.7</b>
<b>Especificidad</b>		<b>45.0%</b>		<b>37.0-53.0</b>
<b>V. predictivo positivo</b>		<b>74.7%</b>		<b>70.0-79.4</b>
<b>V. predictivo negativo</b>		<b>60.3%</b>		<b>51.2-69.4</b>
<b>Índice validez</b>		<b>71.0%</b>		<b>66.8-75.2</b>
<b>Prevalencia</b>		<b>65.7%</b>		<b>61.3-70.1</b>

3.- Al contrastar los resultados globales de la Escala de desarrollo infantil de Gesell (estándar) PC90 y los resultados globales de la prueba Denver II (NyS/A). para el grupo de edad de 13-18 meses y con un total de 362 valoraciones, la prueba estándar reportó 201 valoraciones alteradas de las cuales la prueba Denver II reportó 164 valoraciones como alteradas y 37 valoraciones como sin alteración, obteniéndose una sensibilidad de 81.5%, con un I.C. de 75.9-87.2

La prueba estándar reportó 161 valoraciones sin alteración, la prueba Denver II reportó 88 valoraciones como sin alteración, y 73 valoraciones como alteradas, obteniéndose una especificidad de 54.6%, con un I.C. de 46.6-62.6

La prueba Denver II reportó 237 valoraciones como alteradas de las cuales 164 valoraciones fueron confirmadas por la prueba estándar, obteniéndose un valor predictivo positivo de 69.2%.

La prueba Denver II reportó 125 valoraciones como sin alteración de las cuales 88 valoraciones fueron confirmadas por la prueba estándar y 37 fueron reportadas como alteradas, obteniéndose un valor predictivo negativo de 70.4%. El índice de validez es de 69.6% y la prevalencia del riesgo para alteración en el desarrollo es de 55.5%.

**Tabla 5. Indicadores de validez de la prueba Denver II (criterio Glascoe) respecto al criterio estándar de Gesell (PC 90) para el grupo de 13-18 meses de edad.**

Prueba de Gesell (estándar)				
		Alterado	Sin alteración	Total
Prueba Denver II	Alterado	164	73	237
	Sin alteración	37	88	125
Total		201	161	362

		IC (95%)
Sensibilidad	81.5%	75.9-87.2
Especificidad	54.6%	46.6-62.6
V. predictivo positivo	69.2%	63.1-75.2
V. predictivo negativo	70.4%	62.0-78.8
Índice validez	69.6%	64.7-74.4
Prevalencia	55.5%	50.2-60.7

4.- Al comparar los resultados globales de la Escala de desarrollo infantil de Gesell (estándar) PC90 y los resultados globales de la prueba Denver II (NyS/A). para el grupo de edad de 19-24 meses y con un total de 292 valoraciones, la prueba estándar reportó 155 valoraciones alteradas de las cuales la prueba Denver II reportó 112 valoraciones como alteradas y 43 valoraciones como sin alteración, obteniéndose una sensibilidad de 72.6%, con un I.C de 64.8-79.6

La prueba estándar reportó 137 valoraciones sin alteración, la prueba Denver II reportó 86 valoraciones como sin alteración, y 51 valoraciones como alteradas, obteniéndose una especificidad de 62.7%, con un I.C. de 55.5-72.7

La prueba Denver II reportó 163 valoraciones como alteradas de las cuales 112 valoraciones fueron confirmadas por la prueba estándar como alteradas, obteniéndose un valor predictivo positivo de 68.7%.

La prueba Denver II reportó 129 valoraciones como sin alteración, de las cuales 86 valoraciones fueron confirmadas por la prueba estándar y 43 fueron reportadas como alteradas, obteniéndose un valor predictivo negativo de 66.6%. El índice de validez es de 67.8% y la prevalencia del riesgo para alteración en el desarrollo es de 53.0%.

**Tabla 6. Indicadores de validez de la prueba Denver II (criterio Glascoe) respecto al criterio estándar de Gesell (PC 90) para el grupo de 19-24 meses de edad.**

Prueba de Gesell (estándar)				
Prueba Denver II	Alterado	Alterado Sin alteración		Total
		112	51	
	Sin alteración	43	86	129
	<b>Total</b>	<b>155</b>	<b>137</b>	<b>292</b>
				IC (95%)
<b>Sensibilidad</b>			<b>72.6%</b>	<b>64.8-79.6</b>
<b>Especificidad</b>			<b>62.7%</b>	<b>54.3-71.2</b>
<b>V. predictivo positivo</b>			<b>68.7%</b>	<b>61.2-76.1</b>
<b>V. predictivo negativo</b>			<b>66.6%</b>	<b>58.1-75.1</b>
<b>Índice validez</b>			<b>67.8%</b>	<b>62.2-73.3</b>
<b>Prevalencia</b>			<b>53.0%</b>	<b>47.1-58.9</b>

5.- Al comparar los resultados globales de la Escala de desarrollo infantil de Gesell (estándar) PC90 y los resultados globales de la prueba Denver II (NyS/A). para el grupo de edad de 25-30 meses y con un total de 261 valoraciones, la prueba estándar reportó 130 valoraciones alteradas de las cuales la prueba Denver II reportó 101 valoraciones como alteradas y 29 valoraciones como sin alteración, obteniéndose una sensibilidad de 77.6%, con un I.C. de 70.1-85.2.

La prueba estándar reportó 131 valoraciones sin alteración, la prueba Denver II reportó 84 valoraciones como sin alteración, y 47 valoraciones como alteradas, obteniéndose una especificidad de 64.1%, un I.C de 55.5-72.7.

La prueba Denver II reportó 148 valoraciones como alteradas de las cuales 101 valoraciones fueron confirmadas por la prueba estándar como alteradas, obteniéndose un valor predictivo positivo de 68.4%.

La prueba Denver II reportó 113 valoraciones como sin alteración de las cuales 84 valoraciones fueron confirmadas por la prueba estándar y 29 fueron reportadas como alteradas, obteniéndose un valor predictivo negativo de 74.3%. El índice de validez es de 70.8% y la prevalencia del riesgo para alteración en el desarrollo es de 49.8%.

**Tabla 7. Indicadores de validez de la prueba Denver II (criterio Glascoe) respecto al criterio estándar de Gesell (PC 90) para el grupo de 25-30 meses de edad.**

Prueba de Gesell (estándar)				
Prueba Denver II	Alterado Sin alteración	Alterado Sin alteración		Total
		101	47	
		29	84	113
	Total	130	131	261

		IC (95%)
Sensibilidad	77.6%	70.1-85.2
Especificidad	64.1%	55.5-72.7
V. predictivo positivo	68.4%	60.4-76.0
V. predictivo negativo	74.3%	65.8-82.8
Índice validez	70.8%	65.1-76.5
Prevalencia	49.8%	43.5-56.0

### 7.3.2.- Criterio de calificación según Manual.

1.- Al comparar los resultados globales de la Escala de desarrollo infantil de Gesell (estándar) PC90 y los resultados globales de la prueba Denver II (N/SyA), para el grupo de edad de 0-6 meses y con un total de 356 valoraciones, la prueba estándar reportó 267 valoraciones alteradas de las cuales la prueba Denver II reportó 250 valoraciones como alteradas y 17 valoraciones como sin alteración, obteniéndose una sensibilidad de 93.6%, con un I.C. de 90.5-96.7.

La prueba estándar reportó 89 valoraciones sin alteración, la prueba Denver II reportó 15 valoraciones como sin alteración, y 74 valoraciones como alteradas, obteniéndose una especificidad de 16.8%, un I.C de 8.5-25.1

La prueba Denver II reportó 324 valoraciones como alteradas de las cuales 250 valoraciones fueron confirmadas por la prueba estándar como alteradas, obteniéndose un valor predictivo positivo de 77.1%.

La prueba Denver II reportó 32 valoraciones como sin alteración, de las cuales 15 valoraciones fueron confirmadas por la prueba estándar, obteniéndose un valor predictivo negativo de 46.8%. El índice de validez es de 74.4% y la prevalencia de alteración en el desarrollo es de 75%.

**Tabla 8. Indicadores de validez de la prueba Denver II (según manual) respecto al criterio estándar de Gesell (PC 90) para el grupo de 0-6 meses de edad.**

Prueba de Gesell (estándar)				
Prueba Denver II	Alterado	Sin alteración		Total
		250	74	
	Sin alteración	17	15	32
	<b>Total</b>	<b>267</b>	<b>89</b>	<b>356</b>
				<b>IC (95%)</b>
<b>Sensibilidad</b>	<b>93.6%</b>		<b>90.5-96.7</b>	
<b>Especificidad</b>	<b>16.8%</b>		<b>8.5-25.1</b>	
<b>V. predictivo positivo</b>	<b>77.1%</b>		<b>72.4-81.9</b>	
<b>V. predictivo negativo</b>	<b>46.8 %</b>		<b>28.0-65.7</b>	
<b>Índice validez</b>	<b>74.4%</b>		<b>69.7-79.1</b>	
<b>Prevalencia</b>	<b>75%</b>		<b>70.3-79.6</b>	

2.- Al comparar los resultados globales de la Escala de desarrollo infantil de Gesell (estándar) PC90 y los resultados globales de la prueba Denver II (N/SyA), para el grupo de edad de 7-12 meses y con un total de 473 valoraciones, la prueba estándar reportó 311 valoraciones alteradas de las cuales la prueba Denver II reportó 289 valoraciones como alteradas y 22 valoraciones como sin alteración, obteniéndose una sensibilidad de 92.9%, con un I.C. de 89.9-95.9

La prueba estándar reportó 162 valoraciones sin alteración, la prueba Denver II reportó 41 valoraciones como sin alteración, y 121 valoraciones como alteradas, obteniéndose una especificidad de 25.3%, un I.C de 18.3-32.3

La prueba Denver II reportó 410 valoraciones como alteradas de las cuales 289 valoraciones fueron confirmadas por la prueba estándar como alteradas, obteniéndose un valor predictivo positivo de 70.4%.

La prueba Denver II reportó 63 valoraciones como sin alteración, de las cuales 41 valoraciones fueron confirmadas por la prueba estándar, obteniéndose un valor predictivo negativo de 65%. El índice de validez es de 69.7% y la prevalencia de alteración en el desarrollo es de 65.7%.

**Tabla 9. Indicadores de validez de la prueba Denver II (según manual) respecto al criterio estándar de Gesell (PC 90) para el grupo de 7-12 meses de edad.**

Prueba de Gesell (estándar)				
Prueba Denver II	Alterado Sin alteración	Alterado	Sin alteración	Total
		289	121	410
	22	41	63	
	<b>Total</b>	<b>311</b>	<b>162</b>	<b>473</b>
				<b>IC (95%)</b>
Sensibilidad	92.9%		89.9-95.9	
Especificidad	25.3%		18.3-32.3	
V. predictivo positivo	70.4%		65.9-75.0	
V. predictivo negativo	65.0%		52.5-77.6	
Índice validez	69.7%		65.5-74.0	
Prevalencia	65.7%		61.3-70.1	

3.- Al comparar los resultados globales de la Escala de desarrollo infantil de Gesell (estándar) PC90 y los resultados globales de la prueba Denver II (N/SyA), para el grupo de edad de 13-18 meses y con un total de 362 valoraciones, la prueba estándar reportó 201 valoraciones alteradas de las cuales la prueba Denver II reportó 187 valoraciones como alteradas y 14 valoraciones como sin alteración, obteniéndose una sensibilidad de 93%, con un I.C. de 89.2-96.8

La prueba estándar reportó 161 valoraciones sin alteración, la prueba Denver II reportó 60 valoraciones como sin alteración, y 101 valoraciones como alteradas, obteniéndose una especificidad de 37.2%, un I.C de 29.4-45.0

La prueba Denver II reportó 288 valoraciones como alteradas de las cuales 187 valoraciones fueron confirmadas por la prueba estándar como alteradas, obteniéndose un valor predictivo positivo de 64.9%.

La prueba Denver II reportó 74 valoraciones como sin alteración, de las cuales 60 valoraciones fueron confirmadas por la prueba estándar, obteniéndose un valor predictivo negativo de 81%. El índice de validez es de 68.2% y la prevalencia de alteración en el desarrollo es de 55.5%.

**Tabla 10. Indicadores de validez de la prueba Denver II (según manual) respecto al criterio estándar de Gesell (PC 90) para el grupo de 13-18 meses de edad.**

Prueba de Gesell (estándar)				
		Alterado	Sin alteración	Total
Prueba Denver II	Alterado	187	101	288
	Sin alteración	14	60	74
Total		201	161	362
				IC (95%)
Sensibilidad		93.3%		89.2-96.8
Especificidad		37.2%		29.4-45.0
V. predictivo positivo		64.9%		59.2-70.6
V. predictivo negativo		81.0%		71.4-90.6
Índice validez		68.2%		63.3-73.1
Prevalencia		55.5%		50.2-60.7

4.- Al comparar los resultados globales de la Escala de desarrollo infantil de Gesell (estándar) PC90 y los resultados globales de la prueba Denver II (N/SyA), para el grupo de edad de 19-24 meses y con un total de 292 valoraciones, la prueba estándar reportó 155 valoraciones alteradas de las cuales la prueba Denver II reportó 130 valoraciones como alteradas y 25 valoraciones como sin alteración, obteniéndose una sensibilidad de 83.8%, con un I.C. de 77.7-89.9

La prueba estándar reportó 137 valoraciones sin alteración, la prueba Denver II reportó 65 valoraciones como sin alteración, y 72 valoraciones como alteradas, obteniéndose una especificidad de 47.4%, un I.C de 38.7-56.1

La prueba Denver II reportó 202 valoraciones como alteradas de las cuales 130 valoraciones fueron confirmadas por la prueba estándar como alteradas, obteniéndose un valor predictivo positivo de 64.3%.

La prueba Denver II reportó 90 valoraciones como sin alteración, de las cuales 65 valoraciones fueron confirmadas por la prueba estándar, obteniéndose un valor predictivo negativo de 72.2%. El índice de validez es de 66.7% y la prevalencia de alteración en el desarrollo es de 53%.

**Tabla 11. Indicadores de validez de la prueba Denver II (según manual) respecto al criterio estándar de Gesell (PC 90) para el grupo de 19-24 meses de edad.**

Prueba de Gesell (estándar)				
		Alterado	Sin alteración	Total
Prueba Denver II	Alterado	130	72	202
	Sin alteración	25	65	90
Total		155	137	292
				IC (95%)
Sensibilidad		83.8%		77.7-89.9
Especificidad		47.4%		38.7-56.1
V. predictivo positivo		64.3%		57.5-71.2
V. predictivo negativo		72.2 %		62.4-82.0
Índice validez		66.7%		61.2-72.3
Prevalencia		53.0%		47.1-58.9

5.- Al comparar los resultados globales de la Escala de desarrollo infantil de Gesell (estándar) PC90 y los resultados globales de la prueba Denver II (N/SyA), para el grupo de edad de 25-30 meses y con un total de 261 valoraciones, la prueba estándar reportó 130 valoraciones alteradas de las cuales la prueba Denver II reportó 117 valoraciones como alteradas y 13 valoraciones como sin alteración, obteniéndose una sensibilidad de 90%, con un I.C. de 84.4-95.5.

La prueba estándar reportó 131 valoraciones sin alteración, la prueba Denver II reportó 63 valoraciones como sin alteración, y 68 valoraciones como alteradas, obteniéndose una especificidad de 48%, un I.C de 39.1-57.

La prueba Denver II reportó 185 valoraciones como alteradas de las cuales 117 valoraciones fueron confirmadas por la prueba estándar como alteradas, obteniéndose un valor predictivo positivo de 63.2%.

La prueba Denver II reportó 76 valoraciones como sin alteración, de las cuales 63 valoraciones fueron confirmadas por la prueba estándar, obteniéndose un valor predictivo negativo de 82.8%. El índice de validez es de 68.9% y la prevalencia de alteración en el desarrollo es de 49.8%.

**Tabla 12. Indicadores de validez de la prueba Denver II (según manual) respecto al criterio estándar de Gesell (PC 90) para el grupo de 25-30 meses de edad.**

Prueba de Gesell (estándar)				
		Alterado	Sin alteración	Total
Prueba Denver II	Alterado	117	68	185
	Sin alteración	13	63	76
Total		130	131	261
IC (95%)				
Sensibilidad		90.0%		84.4-95.5
Especificidad		48.0%		39.1-57.0
V. predictivo positivo		63.2%		56.0-70.4
V. predictivo negativo		82.8 %		73.7-92.0
Índice validez		68.9%		63.1-74.7
Prevalencia		49.8%		43.5-56.0

**7.4.- Indicadores de validez de la prueba Denver II según área del desarrollo evaluada. Criterio Rivera (NyS/A).**

**Área Motora.-**

Se confrontó el área motora de la prueba Denver II con el área motora de la escala de desarrollo infantil de Gesell de la población total de estudio y se encontró una sensibilidad de 63.58% y una especificidad de 81.66% con un valor predictivo positivo de 79.56% y un valor predictivo negativo de 66.64%.

Se calculó también por rangos de edad, encontrándose que en el rango de edad de 19-24 meses hay una mayor sensibilidad de 74.14% y una especificidad de 67%, la más baja sensibilidad se encontró en el rango de edad de 7-12 meses con un 56.38% y una especificidad de 87.43%.

**Tabla 1. Indicadores de Validez por criterio externo del área motora de la prueba de Denver II (criterio Rivera), según prueba de Gesell (PC=90)**

Área Motora	V. Positivo	F. Positivo	V. Negativo	F. Negativo	Sensibilidad	Especificidad	V.P. Positivo	V.P. Negativo	I. Validez	Prevalencia
Población Total	646	166	739	370	63.58	81.66	79.56	66.64	72.10	52.89
0-6 meses	206	25	37	88	70.07	59.68	89.18	29.60	68.26	82.58
7-12 meses	168	22	153	130	56.38	87.43	88.42	54.06	67.86	63.00
13-18 meses	79	49	203	31	71.82	80.56	61.72	86.75	77.90	30.39
19-24 meses	86	58	118	30	74.14	67.05	59.72	79.73	69.86	39.73
25-30 meses	90	9	109	53	62.94	92.37	90.91	67.28	76.25	54.79

### Área Adaptativa.-

Se confrontó el área adaptativa de la prueba Denver II con el área adaptativa de la escala de desarrollo infantil de Gesell de la población total de estudio y se encontró una sensibilidad de 14.29% y una especificidad de 97% con un valor predictivo positivo de 83.83% y un valor predictivo negativo de 51.42%.

Se calculó también por rangos de edad, encontrándose que en el rango de edad de 7-12 meses hay una mayor sensibilidad de 29.47% y una especificidad de 86.55%, la más baja sensibilidad se encontró en el rango de edad de 25-30 meses con un 2.48% y una especificidad de 100%.

**Tabla 2. Indicadores de Validez por criterio externo del área adaptativa de la prueba de Denver II (criterio Rivera), según prueba de Gesell (PC=90)**

Área Adaptativa	V. Positivo	F. Positivo	V. Negativo	F. Negativo	Sensibilidad	Especificidad	V.P. Positivo	V.P. Negativo	I. Validez	Prevalencia
Población Total	140	27	889	840	14.29	97.05	83.83	51.42	54.27	51.69
0-6 meses	25	1	47	258	8.83	97.92	96.15	15.41	21.75	85.50
7-12 meses	89	23	148	213	29.47	86.55	79.46	41.00	50.11	63.85
13-18 meses	16	3	233	110	12.70	98.73	84.21	67.93	68.78	34.81
19-24 meses	5	0	218	69	6.76	100	100	75.96	76.37	25.34
25-30 meses	3	0	140	118	2.48	100	100	54.26	54.79	46.36

### Área de Lenguaje.-

Se confrontó el área del lenguaje de la prueba Denver II con el área de lenguaje de la escala de desarrollo infantil de Gesell de la población total de estudio y se encontró una sensibilidad de 35.74% y una especificidad de 82.62% con un valor predictivo positivo de 75.41% y un valor predictivo negativo de 46.29%.

Se calculó también por rangos de edad, encontrándose que en el rango de edad de 0-6 meses hay una mayor sensibilidad de 62.5% y una especificidad de 67.27%, la más baja sensibilidad se encontró en el rango de edad de 19-24 meses con un 14% y una especificidad de 100%.

**Tabla 3. Indicadores por criterio externo del área de lenguaje de la prueba de Denver II (criterio Rivera), según prueba de Gesell (PC=90)**

Área de Lenguaje	V. Positivo	F. Positivo	V. Negativo	F. Negativo	Sensibilidad	Especificidad	V.P. Positivo	V.P. Negativo	I. Validez	Prevalencia
Población Total	411	134	637	739	35.74	82.62	75.41	46.29	54.55	59.86
0-6 meses	85	72	148	51	62.50	67.27	54.14	74.37	65.45	38.20
7-12 meses	178	60	103	132	57.42	63.19	74.79	43.83	59.41	65.54
13-18 meses	73	1	64	224	24.58	98.46	98.65	22.22	37.85	82.04
19-24 meses	26	0	107	159	14.05	100	100	40.23	45.55	63.36
25-30 meses	38	0	111	112	25.33	100	100	49.78	57.09	57.47

### Área Personal social.-

Se confrontó el área personal-social de la prueba Denver II con el área personal social de la escala de desarrollo infantil de Gesell de la población total de estudio y se encontró una sensibilidad de 34.97% y una especificidad de 85.37% con un valor predictivo positivo de 80.75% y un valor predictivo negativo de 42.77%.

Se calculó también por rangos de edad, encontrándose que en el rango de edad de 13-18 meses hay una mayor sensibilidad de 55% y una especificidad de 77.87%, la más baja sensibilidad se encontró en el rango de edad de 0-6 meses con un 12.98% y una especificidad de 97.87%.

**Tabla 4. Indicadores por criterio externo del área personal social de la prueba de Denver II (criterio Rivera), según prueba de Gesell (PC=90)**

	V. Positivo	F. Positivo	V. Negativo	F. Negativo	Sensibilidad	Especificidad	V.P. Positivo	V.P. Negativo	I. Validez	Prevalencia
<b>Área Personal-social</b>										
<b>Población Total</b>	<b>428</b>	<b>102</b>	<b>595</b>	<b>796</b>	<b>34.97</b>	<b>85.37</b>	<b>80.75</b>	<b>42.77</b>	<b>53.25</b>	<b>63.72</b>
<b>0-6 meses</b>	<b>34</b>	<b>2</b>	<b>92</b>	<b>228</b>	<b>12.98</b>	<b>97.87</b>	<b>94.44</b>	<b>28.75</b>	<b>35.39</b>	<b>73.60</b>
<b>7-12 meses</b>	<b>92</b>	<b>30</b>	<b>189</b>	<b>162</b>	<b>36.22</b>	<b>86.30</b>	<b>75.41</b>	<b>53.85</b>	<b>59.41</b>	<b>53.70</b>
<b>13-18 meses</b>	<b>132</b>	<b>27</b>	<b>95</b>	<b>108</b>	<b>55</b>	<b>77.87</b>	<b>83.02</b>	<b>46.80</b>	<b>62.71</b>	<b>66.30</b>
<b>19-24 mese</b>	<b>65</b>	<b>1</b>	<b>58</b>	<b>168</b>	<b>27.90</b>	<b>98.31</b>	<b>98.48</b>	<b>25.66</b>	<b>59</b>	<b>53.26</b>
<b>25-30 meses</b>	<b>58</b>	<b>26</b>	<b>96</b>	<b>81</b>	<b>41.73</b>	<b>78.69</b>	<b>69.05</b>	<b>54.24</b>	<b>59</b>	<b>53.26</b>

## VIII. DISCUSIÓN

Como en el presente estudio se establece la validez por criterio externo de la "Prueba de tamizaje del desarrollo Denver II" se decidió confrontar sus resultados con la "Escala de desarrollo infantil de Gesell", a la manera de "Estándar de Oro", también utilizada en otras investigaciones (14, 15, 30) ya que permite caracterizar eficazmente el desarrollo de los niños e identificar aquellos que se encuentran a riesgo para alteraciones del desarrollo.

Siendo la prueba de Gesell-Amatruda una de las pruebas con mayor consistencia interna, por sus registros sistematizados, por su rigor en el estudio longitudinal de casos, por el amplio número de sujetos estudiados con alteraciones en su desarrollo así como con aquellos que presentaban desarrollo normal y pudiendo establecerse a partir de estos registros evolutivos, edades de desarrollo y coeficientes del desarrollo, es que se emplea esta escala como el estándar de oro para el presente trabajo de investigación.

Para determinar el punto de corte en el presente estudio se decidió por el del coeficiente del desarrollo ( $\geq 90\%$  normal y  $\leq 90\%$  alterado) con base a algunos trabajos realizados en México que utilizaron la escala, como el de la evaluación de la validez del instrumento (Vanedela) Valoración neuroconductual del desarrollo del lactante (14); en el estudio sobre la evaluación del desarrollo motor de acuerdo al nivel socioeconómico (32); en la investigación de niños con desnutrición y sus modificaciones en la evolución de su conducta (33); en el estudio sobre el neurodesarrollo en los dos primeros años de vida de niños con hipotiroidismo congénito (34) entre otros.

Glascoe (21) en su estudio para evaluar la validez del Denver II comparando sus estándares con las de otras pruebas que variaban de acuerdo a la edad del niño, tales como Escalas de desarrollo infantil Bayley, Batería evaluativa de Kaufman o la Escala de Inteligencia Stanford-Binet, Escala de comportamiento adaptativo de Vineland y la Prueba de tamizaje de lenguaje y habla preescolar Fluharty, utilizó como referencia estadística para establecer como adecuados en una prueba de tamizaje contar con una sensibilidad del 80%, especificidad

del 90% y un valor predictivo positivo de 70%, referencias que se toman como punto de comparación para la interpretación de los resultados del presente estudio realizados en la comunidad de Tlaltizapan, según los cuales observamos que la prueba Denver II sobre-detecta a niños con riesgo para alteraciones del desarrollo en contraste con la prueba estándar, al realizar los análisis estadísticos correspondientes la prueba se encuentra dentro de los estándares de validez esperados para las pruebas de tamizaje expresados en un alto nivel de sensibilidad del 80% para los resultados globales, considerando a los sospechosos aun como sin alteración, según los criterios propuestos por Glascoe antes descritos; mientras que la sensibilidad sería del 90.9% si se consideran los sospechosos como alterados que finalmente es el criterio del manual y que no diferencia unos de otros.

Para los grupos de edad, la sensibilidad se encuentra entre el 72.6% y el 84.5% considerando el criterio de Glascoe sospechosos como sin alteración y al utilizar el criterio del manual sospechosos como alterados la sensibilidad aumenta a valores entre 83.8% y 93.6%.

Sin embargo como es de esperarse, al utilizar un criterio de más severidad para interpretar los resultados de la aplicación afectaría la especificidad la cual de por si ya se muestra baja con el criterio sospechosos como sin alteración que es del 57.7% en la población total y baja al 39.1% si se consideran como alterados. A lo largo de los cortes de edad la especificidad se ubica entre 41.5% y 64.1% utilizando el criterio de Glascoe de sospechosos como sin alteración, siendo la más baja a la edad de corte de 0-6 meses y la más alta de 25-30 meses y del 16.8% al 48% utilizando el criterio de sospechosos como alterados, siendo la más baja a la edad de corte de 0-6 meses y la más alta de 25-30 meses. El valor predictivo positivo para los resultados globales es de 72.5% al considerar el criterio de sospechosos como sin alteración y de 67.6% al considerar el criterio de sospechosos como alterados; para los cortes de edad este valor predictivo positivo varia entre 69.2% y 81.2%, siendo la más baja para el grupo de edad de 13-18 meses y la más alta para el grupo de edad de 0-6 meses, considerando el criterio de sospechosos como sin alteración y al considerar el criterio de sospechosos como alterados el valor

predictivo positivo varia entre 63.2% y 77.1%, siendo la más baja para el corte de edad de 25-30 meses y la más alta para el corte de edad de 0-6 meses. De lo cual se desprende que hay una mayor aproximación a los resultados del estándar si se utilizan los criterios de Glascoe respecto a los propuestos inicialmente en el manual.

En cuanto a los resultados obtenidos en la muestra estudiada, los índices de prevalencia son de un 58.3% lo cual nos indica un alto porcentaje en cuanto a la frecuencia de alteraciones del desarrollo en la población infantil de Tlaltizapan. Para los grupos de edad el índice de prevalencia es alto en todos los grupos con porcentajes entre el 50% y 75%; utilizando los resultados de la aplicación del Gesell en población del distrito federal la prevalencia también se ha reportado como alta.

Además en el estudio no se excluyeron pacientes que presentaran alto riesgo, ya que en un estudio de validación por criterio externo, ello no es necesario, por lo mismo debe tomarse con cuidado el asumir que la prevalencia encontrada represente de alguna manera a la población general de la comunidad.

Con respecto a las variaciones que se observan en el estudio de los valores de sensibilidad, especificidad, valor predictivo positivo y valor predictivo negativo en los diferentes cortes de edad pueden ser explicados por la diferencia en el número de ítems que conforman y evalúan cada área de las pruebas por ejemplo: a los 11 meses Denver II evalúa las cuatro áreas y Gesell solo evalúa tres áreas, a los 16 meses Denver II evalúa 38 ítems y Gesell no evalúa esta edad.

En estudios realizados en otros países aplicando la prueba Denver II como en la investigación realizada en Sao Paulo Brasil (35) con una muestra no-clínica de 120 niños para detectar riesgo para problemas del desarrollo en los cuatro primeros años de vida se detectó alteraciones del desarrollo en un 33%; en el estudio realizado en la ciudad de Pelotas Brasil (25) en una muestra de 1,363

niños de 12 meses de edad de una cohorte de niños nacidos en hospitales con el objetivo de verificar la prevalencia de sospecha de atraso en el desarrollo se detectaron 34% de alteraciones en el desarrollo; en USA en la ciudad de New York en un estudio longitudinal de 5 años (36) a 92 niños de pre-escolar se detectó un 46% de alteraciones en el desarrollo que se expresaron años después en problemas de aprendizaje. En México, un estudio realizado en el hospital La Raza (37) en una muestra de 30 niños prematuros con antecedentes de hipoxia perinatal se detectó un 33% de alteraciones en el desarrollo; en estudios realizados en la comunidad de Tlaltizapan implementando el criterio de calificación de Glascoe en la aplicación del Denver II en un estudio de niños de bajo riesgo (23) para conocer la frecuencia de riesgo de alteraciones en el desarrollo se detectó por grupos de edad un alto porcentaje de alteraciones en el desarrollo de entre el 40% y el 88%. En otra investigación (20) realizada en la misma comunidad de Tlaltizapan con un total de 796 valoraciones realizadas a lactantes y empleando el criterio del manual se detectó un 88% de alteraciones en el desarrollo.

Los resultados que reportan estas investigaciones apoyan los resultados de la prevalencia del riesgo para alteraciones del desarrollo en población infantil, de manera especial las realizadas en la comunidad de Morelos que son similares a las del presente estudio en el que se encontró un alto índice de prevalencia de riesgo para el desarrollo.

En el análisis de las áreas de conducta utilizando el criterio de Rivera, de acuerdo a los resultados del estudio para la población total el área adaptativa es la que más baja sensibilidad presenta con un 14.2% y especificidad de 97%, seguida por las áreas de lenguaje y personal social con un 36% y 35% de sensibilidad y especificidad de 83% y 85.3% respectivamente, el área motora presenta una sensibilidad del 64% y especificidad del 82%, variaciones que pueden explicarse por los ítems que conforman cada área y cuyo nivel de exigencia no se cumplen en la edad establecida, como se reportó en el estudio realizado en la comunidad de Tlaltizapan (20).

Al analizar los resultados por áreas para los diferentes grupos de edad encontramos que el área motora es la que mejor detecta niños con riesgo para alteraciones del desarrollo específicamente para las edades de 0 a 6 meses, 13 a 18 meses y de 19 a 24 meses ya que presentan una sensibilidad entre 70% y 74%. Las áreas de lenguaje, personal social presentan una baja y variada sensibilidad en todos los grupos de edad, siendo el área adaptativa la que más baja sensibilidad presenta de todas las áreas en los diferentes grupos de edad.

## IX. CONCLUSIONES

1. La Prueba de tamizaje Denver II presenta adecuada sensibilidad para el total de la población, así como para cada uno de los grupos de edad. La sensibilidad es mayor si se utiliza el criterio del manual respecto al criterio alterno propuesto por Glascoe.
2. La especificidad es baja y variable tanto para la población total como para los diferentes grupos de edad. Su uso implica un sobrediagnóstico de casi la mitad de los niños que no presentan problemas como si los tuvieran, situación que se corrige parcialmente mejorando la especificidad si se utiliza el criterio de interpretación propuesto por Glascoe.
3. Por rangos de edad el criterio de Glascoe (NyS/A) provee una mejor relación entre la sensibilidad y especificidad, por lo mismo es más recomendable utilizar el criterio propuesto por Glascoe.
4. Usando el Criterio de Glascoe el valor predictivo positivo es alto de 73% para la población total y para los grupos de edad varía entre 69% y 81%; el valor predictivo negativo para la población total es de 67%, así como para algunos grupos de edad, solo para los grupos de 13 a 18 meses y de 25 a 30 meses son de 70% y 74%.
5. En el análisis por áreas, la sensibilidad de la prueba es baja en todas las áreas del desarrollo para la población total.
6. Se recomienda el uso de la prueba en edades de 0 a 6 meses, 13 a 18 meses y de 19 a 24 meses, específicamente del área motora ya que presentan una sensibilidad entre el 70% y 74%.

## **X. REFERENCIAS BIBLIOGRAFICAS**

1. Dworkin P. British and American recommendations for developmental monitoring: The role of surveillance. *Pediatrics*. 1989; 84(6):1000-10.
2. American Academy of Pediatrics. Screening infants and Young children for developmental disabilities. *Pediatrics*. 1994;93(5):863-5
3. Morales L. *Psicometría aplicada*. México: Ed. Trillas; 1980.
4. Papalia D. *Desarrollo humano*. Colombia: Ed. Mc Graw Hill;1992.
5. Delval J. *Lecturas de psicología del niño: Las teorías, los métodos y el desarrollo temprano*. Madrid: Ed. Alianza; 1978.
6. Sattler J. *Evaluación infantil*. México: Ed. El manual moderno 3ª ed; 1998.
7. Buela-Casal G, Sierra J. *Manual de evaluación psicológica: Técnicas y aplicaciones*. Madrid: Editores Siglo veintiuno; 1997.
8. Tornimberí S, Pérez E, Fernández A. *Introducción a los tests psicológicos*. Córdoba: Ed. Brujas; 2004.
9. Baltes B, Reese W, Nessel R. *Métodos de investigación en psicología evolutiva: Enfoque del ciclo vital*. Madrid: Edics. Morato; 1981.
10. Vasta R, Haith M, Miller S. *Psicología infantil*. Barcelona: Ed. ARIL S.A; 1996.
11. Cohen J, Swerdlik M. *Pruebas y evaluación psicológica: Introducción a las pruebas y la medición*. México: Ed. Mac Graw Hill; 2000
12. Anastasi A, Urbina S. *Tests psicológicos*. México: Ed. Prentice Hall; 1998.
13. Greenberg R, Daniels S, Flanders D, Boring J. *Epidemiología médica*. México: Ed. Manual moderno: 2002.

14. Martínez I. Evaluación de la sensibilidad y especificidad del instrumento de detección: Valoración Neuroconductual del desarrollo del lactante. (Vanedela) {tesis para Maestría}. México: UAM-X. DAS. CBS.; 2001.
15. Cazarez A. Evaluación de la validez por criterio externo de las escalas Clinical Adaptative Test & Clinical Linguistic Auditory Milestone Scale CAT/CLAMS en niños de 0-12 meses en Tlaltizapan Morelos {tesis para Maestría}. México: UAM-X. DAS. CBS.; 2008.
16. Frankenburg WK, Dodds J, Archer P, Shapiro H, Bresnick b. The Denver II: A mayor developmental screening test. *Pediatrics* 1992; 89(1): 91-7.
17. Lejarraga H, Kelmansky D, Pascucci MC, Salamanco G. Prueba Nacional de Pesquisa PRUNAPE. Buenos Aires: Ediciones Fundación Hospital Garrahan; 2005.
18. Al-Naquib N, Frankenburg WK, Mirza H, Yazdi AW, al-Noori S. The standardization of the Denver Developmental Screening Test on Arab children from the Middle East and north Africa. *J Med Liban.* 1999 Mar-Apr; 47(2):95-106.
19. Frankenburg WK, Dodds J. Denver II Screening Manual. Denver Co.: Denver Developmental Materials. Inc; 1990.
20. Rivera R, Sanchez C, Corral I, Figueroa M, Martínez I, Oliveros ML, Mendoza F, Ortiz M, Rogel J, Casares A. Experiencias del uso del test de Denver II para evaluar el perfil del desarrollo psicomotor en lactantes de una comunidad del estado de Morelos. *Memoria de la CV reunión de la Asociación de Investigación Pediátrica* 2007; 105: 199-212
21. Glascoe P, Byrne E, Ashford G, Johnson L, Chang B, Strickland B. The accuracy of the Denver II in developmental screening. *Pediatrics* 1992; 89 (6):1221-5.
22. Solomons G, Solomons HC. Motor development in Yucatecan infants. *Develop. Med. Child Neurol* 1975; 17: 41-46.

23. Oliveros ML. Detección de riesgo de retrasos del desarrollo en niños de 1 a 24 meses en Tlaltizapan Morelos {tesis para licenciatura}. México; 2009.
24. Ramcharan J, Ali Z, Adams J, Simeon D. Standardization of the Denver Development Screening Test (DDSTII) for Trinidadian children. Caribbean Health Research Council Documents.2005. <http://www.chrccaribbean.org/files/GRANT%20REPORTS%202005/J.%20Ramcharan.pdf>
25. Halpern R, Giugliani J, Victoria G, Barros C, Horta L. Factores de riesgo para sospecha de atraso en el desarrollo neuropsicomotor a los 12 meses de vida. Arch. Pediatr 2002; 73 (1).
26. Sutil AP, Tovar RR. Trastornos del neurodesarrollo: Prueba de Denver vs. Cat/Clams. Universidad Central de Venezuela 2003.
27. Bassuk EL, Rosenberg L. Psychosocial characteristics of homeless children and children with homes. Pediatrics. 1990 Mar; 85(3):257-61.
28. Gesell A, Amatruda C. Diagnóstico del desarrollo normal y anormal del niño. México: Ed. Paidós; 1992.
29. Martínez I, Alvarado R. Examen evolutivo de la conducta de Arnold Gesell y Catherine Amatruda. En: Confederación Nacional de Pediatría de México, Editorial Revistas Profesionales. Neurodesarrollo y estimulación temprana en pediatría. México: 2008.p. 78-92.
30. Ayala S. Relación entre estimulación en el hogar y conducta adaptativa a los 3 años de edad en niños de un medio rural {tesis para licenciatura}. México: UNAM.; 1971.
31. Carmen S, Alvarado G, Romero G, Granados D, Ballesteros B, Mandujano M. et al. Laboratorio de seguimiento del neurodesarrollo (Guía para la organización y procedimientos). México: editor Universidad Autónoma Metropolitana-Xochimilco; 2004.
32. Ontiveros E, Cravioto J, Sánchez C, Barragán-Mejía G. Evaluación del desarrollo motor en función de género, estimulación disponible en el hogar y nivel socioeconómico en niños de 0 a 3 años de edad del

- área rural. En: Fundación derechos de la infancia, editores. La desnutrición infantil en México. México: 2003. p.503-516.
33. Robles B, Ramos-Galván R, Cravioto J. Valoración de la conducta del niño con desnutrición avanzada y de sus modificaciones durante la recuperación. En: Fundación derechos de la infancia, editores. La desnutrición infantil en México. México: 2003. p.435-453.
  34. Sánchez C. Neurodesarrollo en los 2 primeros años de vida de niños con hipotiroidismo congénito detectados por tamiz metabólico, tratamiento hormonal reconstitutivo y estimulación sistematizada {tesis doctoral}. México: UAM-X. ; 2004.
  35. María-Menguel S, Martins B. Factores de riesgo para problemas de desarrollo infantil. Rev. Latinoamericana de Enfermagen 2007; 15.
  36. Diamond E. Effectiveness of the Denver II in identifying children at risk for learning problems. Center for child study, Skidmore college, Saratoga Springs, NY 1990; 83(3).
  37. Salazar A, Ramirez E, Gonzales R, Alva E. Modificaciones de la escala de Denver en la evaluación de las condiciones del neurodesarrollo, en niños atendidos con hipoxia neonatal en una unidad de terapia intensiva. Revista Mexicana de Neurociencias 2006; 7 (1):88-89.
  38. Del Barrio V. Evaluación del niño. En: TEA ediciones. La evaluación psicológica en el año 2000. Madrid: 1997. p.214-217.

#### **Bibliografía consultada.**

39. Buena Lara J. Prueba de tamizaje del desarrollo de Denver II. <http://Jaime bueno.files.wordpress.com/2007/04/denver-teoria.doc>
40. Caon G, Kittel Ries L. Rastreo de la conducta interactiva en la primera niñez. Universidad de Oriente. Santiago de Cuba 2003; 101: 325-329
41. Fernández P, Pértegas D. Pruebas diagnosticas: sensibilidad y especificidad. Universidad de Epidemiología clínica y Bioestadística. Complejo Hospitalario-Universitario Juan Canalejo (serial online) 2003; 10:120-124. Disponible en: URL:

[http://www.Fisterra.com/mbe/investga/pruebas\\_diagnosticas/pruebas\\_diagnosticas.asp](http://www.Fisterra.com/mbe/investga/pruebas_diagnosticas/pruebas_diagnosticas.asp).

42. Giannini H, Moroldo B. Relevancia médica, científica y social de la validación, estandarización y adaptación cultural de las pruebas de pesquisa. Arch. Pediatr 2002; 100(5).
43. Glascoe P. Are overreferrals on developmental screening tests really a problem? Pediatrics 2001; 155 (1):54-59.

## **ANEXOS**

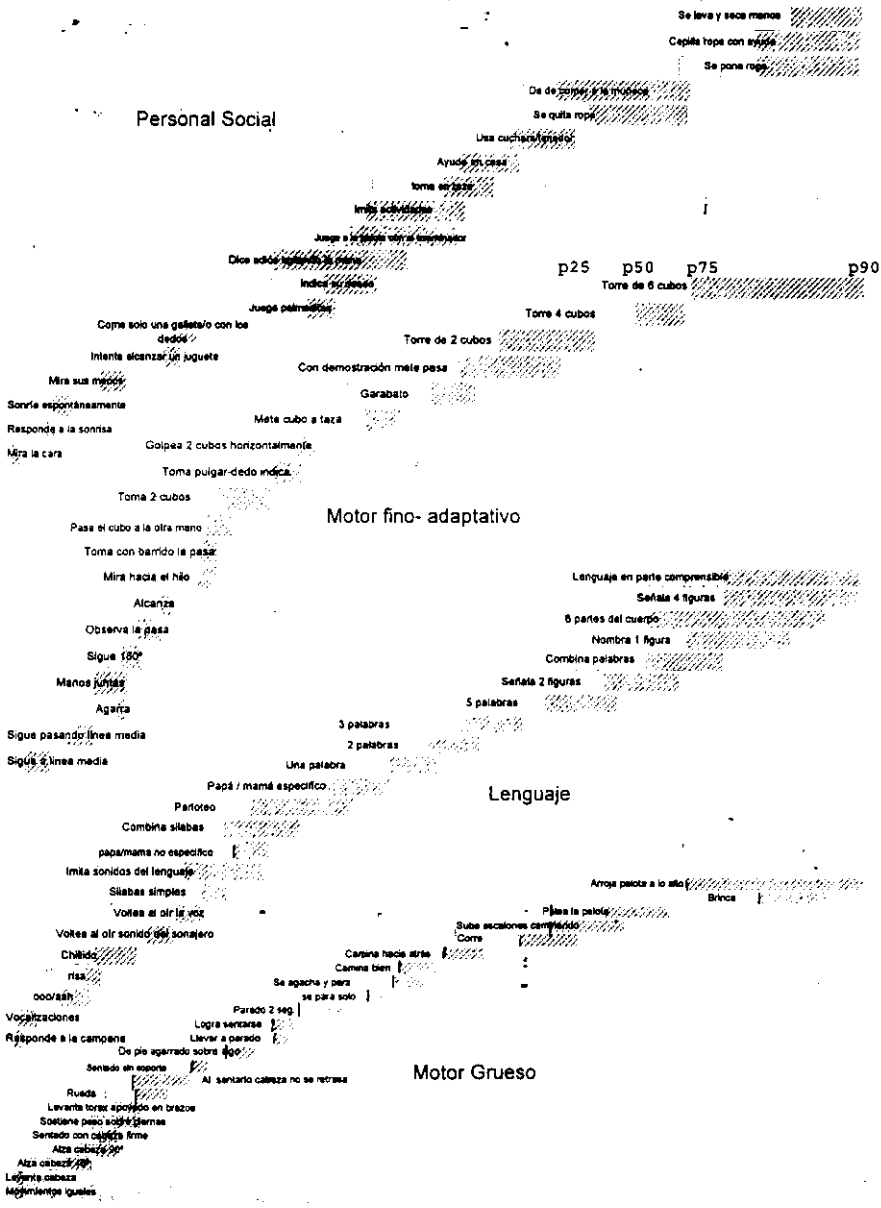
**ANEXO 1. Formato de calificación de la Prueba de tamizaje del desarrollo Denver II**

Laboratorio de Seguimiento del Neurodesarrollo  
formato de registro Denver II

nombre: .....

F. Nac. ....

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30



0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30

para varios registros cambiar color  
indicar la fecha al margen de la vertic

**ANEXO 2. Formato de calificación de la Escala de desarrollo infantil de Gesell**



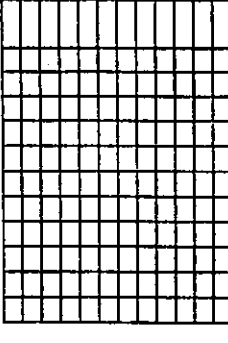
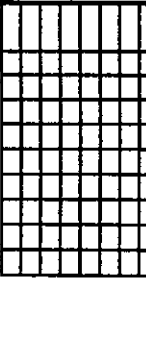
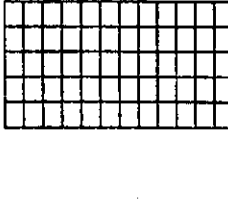
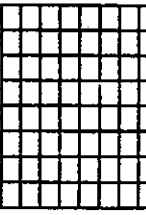


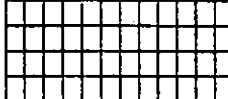
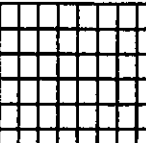
ESQUEMA EVOLUTIVO  
GESELL

4 semanas motriz	8 Semanas motriz
<p><b>Adaptativa</b></p> <p>1 Su: Predomina posición lateral de cabeza (*12 s)            2 Su: predomina posición asimétrica, reflejo tónico - nugal (*16s)            3 Su: Rueda parcialmente sobre el costado (*8s)            4 Tent. Sent: Al intentarlo, caída hacia atrás de la cabeza (*8s)            5 Sent: Caída de la cabeza (*8s)            6 Pro: en suspensión ventral, la cabeza cuelga hacia adelante            7 Pro: colocado en esta posición, la cabeza gira al costado (*8 s)            8 Pro: levanta la cabeza a Zona 1, momentáneamente.            9 Pro: movimientos de arrastre (*8 s)            10 Su: ambas manos fuertemente cerradas y apretadas (*12 s)            11 Son: la mano se cierra al contacto (*8 s).</p>	<p>12 Sent: cabeza predominantemente erecta, aunque bam.            13 Pro: cabeza compensa suspensión ventral            14 Pro: cabeza en línea media            15 Pro: intermitentemente, levanta la cabeza a Zona II</p>
<p><b>Lenguaje</b></p> <p>1 Arcg: Son: mira únicamente cuando están en línea de visión (*8 s)            2 Arcg: Lo sigue con la mirada hacia línea media            3 Son: cae inmediatamente de la mano (*8 s) -            4 Camp: escucha, actividad disminuye (*24s).</p>	<p><b>Adaptativa</b></p> <p>8 Arcg: la mirada demora en línea media (*12 s).            9 Arcg: mira la mano del examinador.            10 Arcg: lo sigue más allá de línea media            11 Son: retiene brevemente.            12 Camp: respuesta facial (*24s).</p>
<p><b>Personal - Social</b></p> <p>1 Expr: faz inexpresiva (*8 s)            2 Expr: mirada vaga, indirecta (*8 s)            3 Voz: pequeños ruidos guturales (*8 s)</p>	<p><b>Lenguaje</b></p> <p>4 Expr: sonrisa (social).            5 Epr: mirada directa            6 Expr: expresión despierta, viva            7 Voz: a-e-u (*56 s)</p>
<p><b>Personal - Social</b></p> <p>1 Soc: mira al Ex., actividad disminuye (*8 s)            2 Su: mira indefinidamente a su contorno (*8 s)</p>	<p><b>Personal- Social</b></p> <p>3 Soc: respuesta facial social            4 Soc: sigue a persona en movimiento.            5 Su: mira al Experimentador</p>

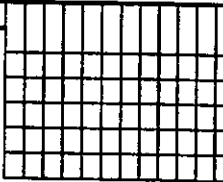
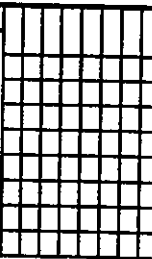
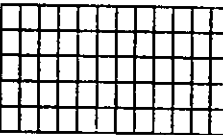
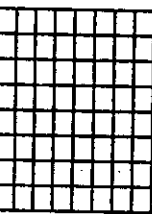


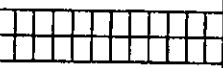

ESQUEMA EVOLUTIVO  
GESELL

20 semanas	24 Semanas
<p><b>motriz</b></p> <p>36 Tentativa a sentar: Cabeza no péndula            37 Sent: Cabeza erecta, firme            38 Pro: Brazos extendidos            39 Pro/Mes: Araña la superficie de la mesa (*28s)            40 Cu: Prensión precaria (*24s)</p>	<p><b>motriz</b></p> <p>41 Su: Eleva los pies en extensión            42 Su: Rueda a posición prona            43 Tent. sent: Levanta la cabeza, ayuda (*40s)            44 Sent/Silla: Tronco erecto (*36s)            45 Cu: Agarra, prensión palmar (*36s)            46 Son: Retiene</p>
<p><b>Adaptativa</b></p> <p>26 Son/Camp: Aproximación bimanual (*28s)            27 Son/Arcg: Lo ase sólo si está próximo a la mano (*24s)            28 Son: Sigue con la vista sonajero caído            29 Cu: Mantiene el primero y mira el segundo            30 CCu: Al tocarlos, agarra uno (*24s)</p>	<p><b>Adaptativa</b></p> <p>31 Arcg/Son/Cu/Camp: Se aproxima y agarra            32 Son: Intenta agarrarlo cuando se le cae            33 Cu: Mira al tercer cubo inmediatamente            34 Cu y Camp: Lleva a la boca (*18s)            35 Cu: Rescata el cubo caído            36 CCu: Sostiene uno se aproxima a otro</p>
<p><b>Lenguaje</b></p> <p>13 Voz: Chillidos (*36s)</p>	<p><b>Lenguaje</b></p> <p>14 Comp: Vuelve la cabeza hacia lugar del sonido            15 Voz: Gruñidos (*36s)            16 Voz: Parloteo espontáneo (incluyendo juguetes)</p>
<p><b>Personal - Social</b></p> <p>16 Soc: Sonríe a la imagen del espejo            17 Ali: Acaricia la botella (*36s)</p>	<p><b>Personal- Social</b></p> <p>18 Soc: Distingue extraños            19 Jue: Agarra el pie (supino) (*36s)            20 Jue: Sentado con soporte, 30 minutos (*40s)            21 Expr: Sonríe y parlotea</p>

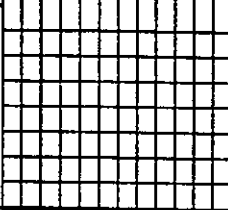
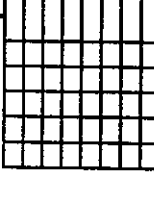
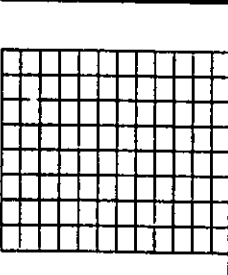
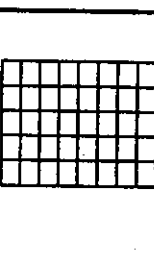
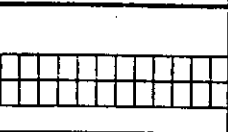
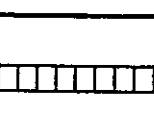
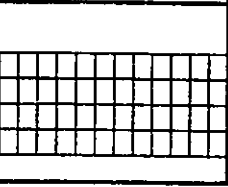
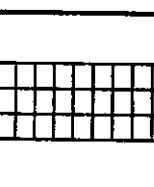
**ESQUEMA EVOLUTIVO  
GESELL**

<p align="center"><b>12 semanas motriz</b></p> <p>16 Su: Cabeza predominantemente a un lado (r.t.n )          17 Su: Cabeza en posición mediana, posición simétrica.          18 Sent: Cabeza dirigida adelante, bamboleo (*18s)          19 Pa: Brevemente, pequeña fracción del peso corporal          20 Pa: Levanta el pie (*24s)          21 Pro: Cabeza sostenida en Zona II          22 Pro: Sobre antebrazos (*20s)          23 Pro: Cadena baja (piernas flexionadas) (*40s)          24 Su: Manos abiertas o laxamente cerradas          25 Son: Sostiene activamente          26 Ta: La toca</p>		<p align="center"><b>16 Semanas motriz</b></p> <p>27 Su: Predomina posición mediana de la cabeza          28 Su: Predomina postura simétrica          29 Su: Las manos se encuentran (*24s)          30 Sent: Cabeza firme dirigida adelante (*20s)          31 Pro: Cabeza sostenida en Z III          32 Pro: Piernas extendidas o semiextendidas (*40s)          33 Pro: Tendencia a rodar (*20s)          34 Arcg: Lo retiene          35 Su: Araña, rasca, ase (*24s)</p>	
<p align="center"><b>Adaptativa</b></p> <p>13 Arcg/ Su: Dirige mirada a línea media          14 Arcg/su: Sigue 180°          15 Son/Su: La mira manteniéndola en la mano          16 Cu,Ta: Los mira, más que momentáneamente          17 Arcg/Su: Sigue 90°</p>		<p align="center"><b>Adaptativa</b></p> <p>18 Arcg/Son: Mira inmediatamente          19 Arcg/Son,Cu,Ta: provocan actividad en miembros sup.          20 Arcg/Son: Mira, sosteniéndolos en la mano          21 Arcg: Lleva a la boca          22 Arcg: Mano libre a línea media (*28s)          23 Mes: Baja la mirada a la mesa o a las manos          24 Cu/Ta: Mirada de la mano a los objetos (*20s)          25 Bol: Observa</p>	
<p align="center"><b>Lenguaje</b></p> <p>8 Voz: Murmullo (*36s)          9 Voz: Cloqueo          10 Soc: Respuesta vocal - social</p>		<p align="center"><b>Lenguaje</b></p> <p>11 Expr: Se entusiasma, respira fuertemente (*32s)          12 Voz: Ríe fuertemente</p>	
<p align="center"><b>Personal - Social</b></p> <p>6 Soc: respuesta vocal - social          7 Su: Observa predominantemente al examinador          8 Jue: Observa sus manos (*24s)          9 Jue: Tira de la ropa (*24s)</p>		<p align="center"><b>Personal - Social</b></p> <p>10 Soc: Sonríe espontáneamente          11 Soc: Al intentar sentarlo coicaliza ó sonríe (*24s)          12 Ali: Anticipación a la vista de alimentos          13 Jue: Con ayuda permanece sentado 10 a 15 min. (*40s)          14 Jue: Juego manual, dedos entre sí (*24s)          15 Jue: Tira de la ropa sobre la cara</p>	

**ESQUEMA EVOLUTIVO  
GESELL**

<b>36 semanas</b>		<b>40 Semanas</b>	
<p align="center"><b>motriz</b></p> <p>59 Sent: Firmemente más de 10 minutos 60 Sent: Se inclina hacia delante y recobra posición anterior 61 Pa: Sostenido de la baranda, mantiene todo su peso 62 Cu: Presión digito - radial 63 Bol: Presión tipo tijera (*40s)</p>		<p align="center"><b>motriz</b></p> <p>64 Sent: Permanece firme, indefinidamente 65 Sent: Pasa a la posición prona 66 Pa: Se para apoyado en la barandilla (*15m) 67 Pro: Gatea (*15m) 68 Cu: Suelta crudamente (*15m) 69 Bol: Agarra rápidamente 70 Bol: Presión pinza tipo inferior (*48s) 71 Arci: Agarra la cinta con facilidad</p>	
<p align="center"><b>Adaptativa</b></p> <p>50 Cu: ase tercer cubo (*40s) 51 Cu: Golpea, sacude cubo contra cubo (*15m) 52 Ta/ Cu: Lleva el cubo contra la taza (*44s) 53 Bol/bot: Se dedica primero a la botella (*40s) 54 Arci: Manipula la cinta</p>		<p align="center"><b>Adaptativa</b></p> <p>55 Cu: Aparea dos cubos (*15m) 56 Cu/Ta: Toca cubos dentro de la taza 57 Bol: Se aproxima con el índice 58 Bol en bot: Observa bolita si sale de botella 59 Bol y Bot: Atiende antes a bolita 60 Bol y Bot: Agarra bolita 61 Camp: Agarra por el mango 62 Camp: Sacude espontáneamente</p>	
<p align="center"><b>Lenguaje</b></p> <p>20 Voz: da-da (o equivalente) 21 Voz: Imita sonidos 22 Comp: Responde al nombre y a no-no</p>		<p align="center"><b>Lenguaje</b></p> <p>23 Voz: da-da y mamá 24 Voz: Una palabra 25 Comp: Adiós y aplmaditas</p>	
<p align="center"><b>Personal - Social</b></p> <p>29 Ali: Sostiene el biberón (*15m) 30 Ali: Come, solo, una galletita</p>		<p align="center"><b>Personal- Social</b></p> <p>31 Soc: Movimientos de adiós y palmaditas (*44s)</p>	

ESQUEMA EVOLUTIVO  
GESELL

28 semanas		32 Semanas	
<p><b>motriz</b></p> <p>47 Su: Levanta la cabeza (*40s)            48 Sent: Brevemente, inclinado hacia delante, sobre las manos            49 Sent: Momentáneamente erecto            50 Pa: Sostiene gran parte del peso (*36s)            51 Pa: Salta activamente (*32s)            52 Cu: Presión palmar radial (*36s)            53 Bol: Intenta asirla con toda la mano, la toca (*32s)</p>		<p><b>motriz</b></p> <p>54 Sent: Un minuto erecto, inseguro (*36s)            55 Pa: Se mantiene brevemente sostenido de la mano            56 Pro: Gira (*40s)            57 Bol: Barrido radial (*36s)            58 Bol: Ineficaz presión tipo inferior tijera (*36s)</p>	
<p><b>Adaptativa</b></p> <p>37 Son/Camp: Intenta acercarse, asir con una mano            38 Cu: Sostiene uno, agarra otro            39 CCu: Mantiene dos más que momentáneamente            40 Camp: Sacude contra la mesa (*40s)            41 Son: Sacude definitivamente            42 Arcg/Cu: Pasa de una mano a la otra            43 Camp: Transfiere de una mano a la otra, correctamente            44 Camp: Retiene</p>		<p><b>Adaptativa</b></p> <p>45 Cu: Agarra segundo cubo            46 Cu: Retiene segundo cubo se aproxima al tercero            47 Cu: Mantiene dos, prolongadamente            48 Ta/Cu: Sostiene el cubo, mira la taza            49 Arci: Apriciona el aro</p>	
<p><b>Lenguaje</b></p> <p>17 Voz: m-m-m llorando (*40s)            18 Voz: Sonidos vocales polisilábicos (*36s)</p>		<p><b>Lenguaje</b></p> <p>19 Voz: Sílabas simples tales como da, ba, ta, ca</p>	
<p><b>Personal - Social</b></p> <p>22 Ali: Toma bien los sólidos            23 Jue: Lleva el pie a la boca (supina) (*36s)            24 Expr: Toca, acaricia imagen            25 Arci: Calma o abandona esfuerzo (32s)**</p>		<p><b>Personal-Social</b></p> <p>26 Jue: Muerde, chupa los juguetes (*18s)            27 Jue: busca persist. juguetes fuera de su alcance (*40s)            28 Arci: Persistente</p>	

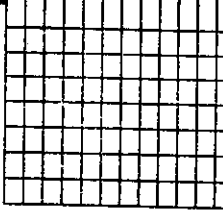
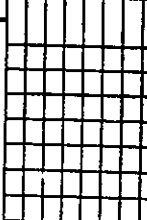
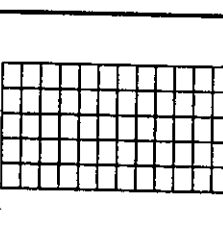

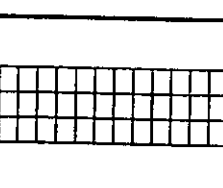
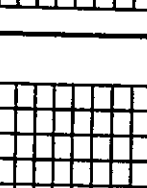
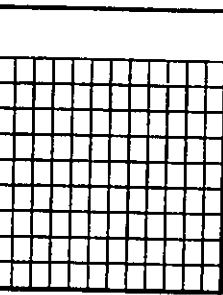
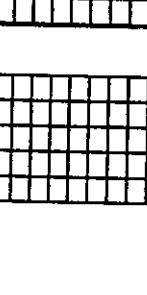
ESQUEMA EVOLUTIVO  
GESELL

52 Semanas	56 Semanas
<p><b>motriz</b></p> <p>78 Mar: Necesita ser sostenido de una sola mano (*15m)</p>	<p><b>motriz</b></p> <p>79 Pa: Momentáneamente solo 80 Cu: Agarra dos en una mano</p>
<p><b>Adaptativa</b></p> <p>71 Cu: (dem) intenta construir torre fracasa (*15m) 72 Cu y Ta: (Dem) deja caer cubo dentro de la taza (*56s) 73 Arc: Sacude el aro por la cirta 74 Tab: Mira selectivamente el agujero redondo</p>	<p><b>Adaptativa</b></p> <p>75 Cu y Ta: (No dem.) Un cubo dentro de la taza 76 Dib: Vigorosos garabateos Imitativos 77 Tab: (Dem) inserta bloque redondo (*15m)</p>
<p><b>Lenguaje</b></p> <p>26 Voz: Dos palabras además de mamá y papá 27 Comp: Alcanza un juguete (pedido y gesto)</p>	<p><b>Lenguaje</b></p> <p>28 Voz: 3-4 palabras 29 Voz: Jerga incipiente (*24m) 30 Comp: Unos cuantos objetos por el nombre</p>
<p><b>Personal - Social</b></p> <p>37 Esp: Acerca pelota al espejo 38 Vestido: Coopera en vestirse (*48m)</p>	<p><b>Personal- Social</b></p> <p>39 Pel: Suelta con débil impulso hacia el examinador</p>

ESQUEMA EVOLUTIVO  
GESELL

44 semanas	48 Semanas
<p><b>motriz</b></p> <p>72 Pa: Se eleva (apoyado), vuelve a su lugar (*48s)</p> <p>73 Camp: Agarra por la punta del mango</p>	<p><b>motriz</b></p> <p>74 Sent: Gira o se tuerce</p> <p>75 Pa: Trasládase apoyado en la barandilla (*15m)</p> <p>76 Mar: Necesita ser sostenido de ambas manos (*52s)</p> <p>77 Bol: Neta prensión en pinza</p>
<p><b>Adaptativa</b></p> <p>63 Ta/Cu: Saca el cubo de la taza</p> <p>64 Ta/Cu: (dem) cubo dentro de la taza sin soltarlo (*52s)</p> <p>65 Bol en Bot: Señala bolita a través del vidrio (*18m)</p> <p>66 Camp: Mira y hurga el badajo</p> <p>67 Arci: Se dedica primero a la cinta</p>	<p><b>Adaptativa</b></p> <p>68 Cu: Juego sucesivo (*36M)</p> <p>69 Bol y Bot: Ase la bolita únicamente</p> <p>70 Tab: Remueve fácilmente el bloque redondo</p>
<p><b>Lenguaje</b></p>	<p><b>Lenguaje</b></p>
<p><b>Personal - Social</b></p> <p>32 Soc: Extiende el juguete a otra persona, sin soltarlo (*52s)</p> <p>33 Ali: Toma de la taza (en parte)</p> <p>34 Esp: Busca la imagen de la pelota en la mano (*52s)</p>	<p><b>Personal- Social</b></p> <p>35 Jue: Juguetea al lado de la barandilla (*15m)</p> <p>36 Jue: Juega en mesa de examen(*52s)</p>

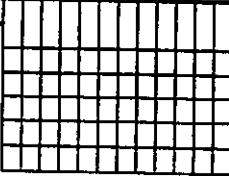
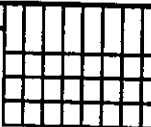
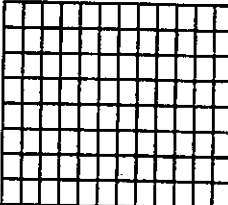
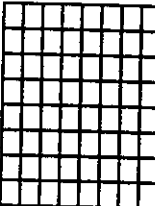
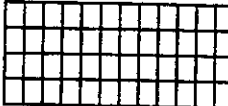
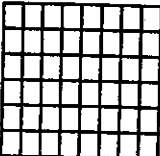
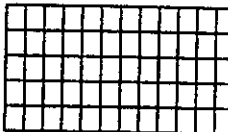
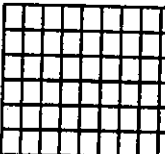
ESQUEMA EVOLUTIVO  
GESELL

15 Meses		16 meses	
<p><b>motriz</b></p> <p>81 Mar: Pocos pasos, inicia, se para                      82 Mar: Cae por derrumbe(*18m)                      83 Mar: Abandona el gateo                      84 Esc: Sube gateando (*18m)                      85 CCu: Torre de 2                      86 Bol: (Sin dem.) introduce en botella                      87 Lib: Ayuda volver páginas</p>		<p><b>motriz</b></p> <p>88 Mar: Cae rara vez                      89 Mar: Ligera, corre tieso (*24m)                      90 Esc: Sube, sostenido de una mano (*21m)                      91 Sp: Se sienta solo                      92 Sg: Se trepa                      93 Pel: Empuja con el pie (*21m)                      94 Lib: Vuelve páginas, 2-3 por vez (*24m)</p>	
<p><b>Adaptativa</b></p> <p>78 Cu: torre 2                      79 Cu y Ta: Coloca y saca 6 cubos (*18m)                      80 Dib: Incipiente imitación del trazo (*18m)                      81 Tab: (Sim dem.) Coloca bloque redondo                      82 Tab: Adapta bloque redondo rápidamente</p>		<p><b>Adaptativa</b></p> <p>83 CCu: Torre de 3-4                      84 Ta/Cu: 10 dentro de la taza                      85 Bol y Bol: Extracción                      86 Dib: Garabateo espontáneo (*36m)                      87 Dib: Imitando hace un trazo                      88 Tab: Apila 3 bloques (*24m)</p>	
<p><b>Lenguaje</b></p> <p>31 Voz: 4 o 5 palabras incluyendo nombre                      32 Voz: Usa jerga (*24m)                      33 Lib: Acaricia dibujos (*18m)</p>		<p><b>Lenguaje</b></p> <p>34 Lib: Mira selectivamente                      35 Voz: 10 palabras incluyendo nombre                      36 Lam: Nombra o señala un dibujo                      37 Obj: Nombra la pelota                      38 Pel: 2 órdenes</p>	
<p><b>Personal - Social</b></p> <p>40 Ai: Descarta biberón                      41 Ai: Inhibe agarrar el plato                      42 Control de esfínteres: Regulación parcial                      43 Control de esfínteres: Control rectal                      44 Control de esfínteres: Indica ropas mojadas (*18m)                      45 Común: Dice ta-ta o equivalente                      46 Común: Señala, parlotea, cuando desea algo (*21m)                      47 Jue: Muestra u ofrece juguetes (*21m)                      48 Jue: Tira objetos de juego o rechaza (*28m)</p>		<p><b>Personal- Social</b></p> <p>49 Ai: Agarra plato vacío                      50 Ai: Come solo, en parte derrama (*36m)                      51 Control de esfínteres: Regulado durante el día                      52 Jue: Arrastra un juguete (*30m)                      53 Jue: Lleva o abraza muñeca (*24m)</p>	

ESQUEMA EVOLUTIVO  
GESELL

21 Meses		24 meses	
<p><b>motriz</b></p> <p>95 Mar: Se agacha jugando 96 Esc: Baja, sostenido de una mano (*24m) 97 Esc: Sube, sostenido del pasamanos (*24m) 98 Pel. g: Patea (dem) (*24m)</p>		<p><b>motriz</b></p> <p>99 Mar: Corre bien, no cae 100 Esc: Sube y baja sólo 101 Pel.g: (Sin dem) patea 102 CCu: Torre de 6 o 7 103 Lib: Vuelve páginas de una en una</p>	
<p><b>Adaptativa</b></p> <p>89 CCu: Torre de 5 o 6 90 CCu: Imita empujar tren (*24m) 91 Tab: Ubica 2-3 bloques 92 Ca. Pr: Inserta esquina del bloque cuadrado (*24m) 93 Ca. Pr: Recupera la pelota</p>		<p><b>Adaptativa</b></p> <p>94 CCu: Torre de 6 o 7 95 CCu: Alinea 2 o más, para tren (*36m) 96 Dib: imita trazo en V 97 Dib: imita trazos circulares 98 Tab: Ubica bloques aislados (*30m) 99 Tab: Adapta después de 4 ensayos (*30m) 100 Ca. Pr: Inserta trozo cuadrado</p>	
<p><b>Lenguaje</b></p> <p>39 Voz: 20 palabras 40 Cnv: Combina 2-3 palabras espontáneamente 41 Pel: 3 órdenes</p>		<p><b>Lenguaje</b></p> <p>42 Conv: Descarta la jerga 43 Conv: Frase de 3 palabras 44 Conv: Usa yo, mi, tu 45 Lam: Nombra 3 o más dibujos 46 Lam: Identifica 5 o más dibujos 47 Obj: Nombra 2 48 Pel: 4 órdenes</p>	
<p><b>Personal - Social</b></p> <p>54 Ali: Maneja la taza 55 Común: Pide comida, bebida e ir al baño 56 Común: Repite 2 o más últimas palabras (*24m) 57 Común: Empuja personas para mostrar (*24m)</p>		<p><b>Personal- Social</b></p> <p>58 Ali: Inhibe girar la cuchara 59 Control de esf: Seco de noche, levantandolo (*36m) 60 Control de esf: Verbaliza necesidades regularmente (*42m) 61 Vest: Se pone prenda simples 62 Común: Manifiesta experiencias inmediatas 63 Común: Se refiere a sí mismo por el nombre (*30m) 64 Común: Comprende y pregunta por otro 65 Jue: Alcanza taza llena de cubos 66 Jue: Imitación doméstica 67 Jue: Predomina juego paralelo (*42m)</p>	

**ESQUEMA EVOLUTIVO  
GESELL**

30 Meses		36 Meses	
<p align="center"><b>motriz</b></p> <p>104 Mar: (Dem) En punta de pie 105 Salt: Con ambos pies 106 Pa: Intenta pararse sobre un solo pie 107 CCu: Torre de 8 108 Dib: Sostiene el lápiz con los dedos</p>		<p align="center"><b>motriz</b></p> <p>109 Esc: Alterna los pies al subir 110 Salt: Del último escalón 111 Pa: Sobre un pie, equilibrio momentáneo</p>	
<p align="center"><b>Adaptativa</b></p> <p>101 CCu: Torre de 8 102 CCu: Agrega chimenea al tren 103 Dib: Dos o más trazos para una cruz (*36m) 104 Dib: Imita V y H 105 For. Col: Ubica uno 106 Tab: Inserta 3 bloques 107 Tab: Adapta repetidamente, error (*36m) 108 Dígitos: Repite 2 ejem. 3-7 (1 de 3 ensayos)</p>		<p align="center"><b>Adaptativa</b></p> <p>109 CCu: Torre de 9 (10 en 3 ensayos) 110 CCu: Imita puente (*42m) 111 Dib: Nombra sus dibujos 112 Dib: Copia círculo 113 Dib: Imita cruz (*48m) 114 For. Col: Aparea tres 115 Tab: Adapta sin error o con corrección esp.e inmediata 116 Dígitos: repite 3 Ejem. 2-4-1 (1 en 3 ensayos)</p>	
<p align="center"><b>Lenguaje</b></p> <p>49 Nombre: Dice su nombre completo 50 Lam: Nombra 5 dibujos 51 Lam: Identifica 7 dibujos 52 Obj: Indica el uso</p>		<p align="center"><b>Lenguaje</b></p> <p>53 Lib: Explica acción 54 Conv: Usa plurales 55 Lam: Nombra 8 imágenes 56 Sexo: Dice su sexo 57 Comprensión: Contesta una pregunta 58 Ordenes: Obedece 2 órdenes, pelota y silla</p>	
<p align="center"><b>Personal - Social</b></p> <p>68 Común: Se refiere a sí mismo por el pronombre 69 Común: Repite en conversación y otras actividades (*36m) 70 Jue: Empuja juguete, buena dirección 71 Jue: Ayuda a llevar de vuelta cosas a su lugar 72 Jue: Puede trasportar objeto frágil</p>		<p align="center"><b>Personal- Social</b></p> <p>73 Ali: Come solo, derrama poco 74 Ali: Vierte bien en un jarrito 75 Vest: Se pone los zapatos 76 Vest: Desabrocha botones accesibles 77 Común: Comprende hacer turnos 78 Común: Conoce unos pocos versos</p>	



**ANEXO 3. Indicadores de validez por áreas y por grupos de edad, comparando los resultados de la prueba de tamizaje Denver II y la escala de desarrollo infantil de Gesell (estándar)**

**Tabla 1. Análisis del Diagnóstico Global del área motora de Gesell PC90 por el Diagnóstico Global del área motora de Denver II según criterio de Glascoe. (Toda la población)**

Nivel de confianza: 95.0%

Prueba diagnóstica	Prueba de referencia		Total
	alterado	salteración	
alterado	646	166	812
sin alteración	370	739	1109
Total	1016	905	1921

	Valor	IC (95%)	
Sensibilidad (%)	63.58	60.57	66.59
Especificidad (%)	81.66	79.08	84.23
Índice de validez (%)	72.10	70.07	74.13
Valor predictivo + (%)	79.56	76.72	82.39
Valor predictivo - (%)	66.64	63.82	69.46
Prevalencia (%)	52.89	50.63	55.15

**Tabla 2. Análisis del Diagnóstico Global del área motora de Gesell PC90 por el Diagnóstico Global del área motora de Denver II según criterio de Glascoe. (Grupo de edad de 0-6 meses)**

Nivel de confianza: 95.0%

Prueba diagnóstica	Prueba de referencia		Total
	alterado	salteración	
alterado	206	25	231
sin alteración	88	37	125
Total	294	62	356

	Valor	IC (95%)	
Sensibilidad (%)	70.07	64.66	75.47
Especificidad (%)	59.68	46.66	72.69
Índice de validez (%)	68.26	63.28	73.23
Valor predictivo + (%)	89.18	84.95	93.40
Valor predictivo - (%)	29.60	21.20	38.00
Prevalencia (%)	82.58	78.50	86.66

**Tabla 3. Análisis del Diagnóstico Global del área motora de Gesell PC90 por el Diagnóstico Global del área motora de Denver II según criterio de Glascoe. (Grupo de edad de 7-12 meses)**

Nivel de confianza: 95.0%

Prueba diagnóstica	Prueba de referencia		Total
	alterado	salteración	
alterado	168	22	190
sin alteración	130	153	283
Total	298	175	473

	Valor	IC (95%)	
Sensibilidad (%)	56.38	50.58	62.17
Especificidad (%)	87.43	82.23	92.63
Índice de validez (%)	67.86	63.55	72.18
Valor predictivo + (%)	88.42	83.61	93.23
Valor predictivo - (%)	54.06	48.08	60.05
Prevalencia (%)	63.00	58.55	67.46

**Tabla 4. Análisis del Diagnóstico Global del área motora de Gesell PC90 por el Diagnóstico Global del área motora de Denver II según criterio de Glascoe. (Grupo de edad de 13-18 meses)**

Nivel de confianza: 95.0%

Prueba diagnóstica	Prueba de referencia		Total
	alterado	salteración	
alterado	79	49	128
sin alteración	31	203	234
Total	110	252	362

	Valor	IC (95%)	
Sensibilidad (%)	71.82	62.96	80.68
Especificidad (%)	80.56	75.47	85.64
Índice de validez (%)	77.90	73.49	82.31
Valor predictivo + (%)	61.72	52.91	70.53
Valor predictivo - (%)	86.75	82.19	91.31
Prevalencia (%)	30.39	25.51	35.26

**Tabla 5. Análisis del Diagnóstico Global del área motora de Gesell PC90 por el Diagnóstico Global del área motora de Denver II según criterio de Glascoe. (Grupo de edad de 19-24 meses)**

Nivel de confianza: 95.0%

Prueba diagnóstica	Prueba de referencia		Total
	alterado	salteración	
alterado	86	58	144
sin alteración	30	118	148
Total	116	176	292

	Valor	IC (95%)	
Sensibilidad (%)	74.14	65.74	82.54
Especificidad (%)	67.05	59.82	74.27
Índice de validez (%)	69.86	64.43	75.30
Valor predictivo + (%)	59.72	51.36	68.08
Valor predictivo - (%)	79.73	72.92	86.54
Prevalencia (%)	39.73	33.94	45.51

**Tabla 6. Análisis del Diagnóstico Global del área motora de Gesell PC90 por el Diagnóstico Global del área motora de Denver II según criterio de Glascoe. (Grupo de edad de 25-30 meses)**

Nivel de confianza: 95.0%

Prueba diagnóstica	Prueba de referencia		Total
	alterado	salteración	
alterado	90	9	99
sin alteración	53	109	162
Total	143	118	261

	Valor	IC (95%)	
Sensibilidad (%)	62.94	54.67	71.20
Especificidad (%)	92.37	87.16	97.59
Índice de validez (%)	76.25	70.89	81.60
Valor predictivo + (%)	90.91	84.74	97.08
Valor predictivo - (%)	67.28	59.75	74.82
Prevalencia (%)	54.79	48.56	61.02

**Tabla 7. Análisis del Diagnóstico Global del área adaptativa de Gesell PC90 por el Diagnóstico Global del área adaptativa de Denver II según criterio de Glascoe. (Toda la población)**

Nivel de confianza: 95.0%

Prueba diagnóstica	Prueba de referencia		Total
	alterado	salteración	
alterado	140	27	167
sin alteración	840	889	1729
Total	980	916	1896

	Valor	IC (95%)	
Sensibilidad (%)	14.29	12.04	16.53
Especificidad (%)	97.05	95.90	98.20
Índice de validez (%)	54.27	52.00	56.54
Valor predictivo + (%)	83.83	77.95	89.72
Valor predictivo - (%)	51.42	49.03	53.80
Prevalencia (%)	51.69	49.41	53.96

**Tabla 8. Análisis del Diagnóstico Global del área adaptativa de Gesell PC90 por el Diagnóstico Global del área adaptativa de Denver II según criterio de Glascoe. (Grupo de edad de 0-6 meses)**

Nivel de confianza: 95.0%

Prueba diagnóstica	Prueba de referencia		Total
	alterado	salteración	
alterado	25	1	26
sin alteración	258	47	305
Total	283	48	331

	Valor	IC (95%)	
Sensibilidad (%)	8.83	5.35	12.32
Especificidad (%)	97.92	92.83	100.00
Índice de validez (%)	21.75	17.16	26.35
Valor predictivo + (%)	96.15	86.84	100.00
Valor predictivo - (%)	15.41	11.19	19.63
Prevalencia (%)	85.50	81.55	89.44

**Tabla 9. Análisis del Diagnóstico Global del área adaptativa de Gesell PC90 por el Diagnóstico Global del área adaptativa de Denver II según criterio de Glascoe. (Grupo de edad de 7-12 meses)**

Nivel de confianza: 95.0%

Prueba diagnóstica	Prueba de referencia		Total
	alterado	salteración	
alterado	89	23	112
sin alteración	213	148	361
Total	302	171	473

	Valor	IC (95%)	
Sensibilidad (%)	29.47	24.16	34.78
Especificidad (%)	86.55	81.14	91.96
Índice de validez (%)	50.11	45.49	54.72
Valor predictivo + (%)	79.46	71.54	87.39
Valor predictivo - (%)	41.00	35.79	46.21
Prevalencia (%)	63.85	59.41	68.28

**Tabla 10. Análisis del Diagnóstico Global del área adaptativa de Gesell PC90 por el Diagnóstico Global del área adaptativa de Denver II según criterio de Glascoe. (Grupo de edad de 13-18 meses)**

Nivel de confianza: 95.0%

Prueba diagnóstica	Prueba de referencia		Total
	alterado	salteración	
alterado	16	3	19
sin alteración	110	233	343
Total	126	236	362

	Valor	IC (95%)	
Sensibilidad (%)	12.70	6.49	18.91
Especificidad (%)	98.73	97.09	100.00
Índice de validez (%)	68.78	63.87	73.70
Valor predictivo + (%)	84.21	65.18	100.00
Valor predictivo - (%)	67.93	62.84	73.02
Prevalencia (%)	34.81	29.76	39.85

**Tabla 11. Análisis del Diagnóstico Global del área adaptativa de Gesell PC90 por el Diagnóstico Global del área adaptativa de Denver II según criterio de Glascoe. (Grupo de edad de 19-24 meses)**

Nivel de confianza: 95.0%

Prueba diagnóstica	Prueba de referencia		Total
	alterado	salteración	
alterado	5	0	5
sin alteración	69	218	287
Total	74	218	292

	Valor	IC (95%)	
Sensibilidad (%)	6.76	0.36	13.15
Especificidad (%)	100.00	99.77	100.00
Índice de validez (%)	76.37	71.33	81.41
Valor predictivo + (%)	100.00	90.00	100.00
Valor predictivo - (%)	75.96	70.84	81.08
Prevalencia (%)	25.34	20.18	30.50

**Tabla 12. Análisis del Diagnóstico Global del área adaptativa de Gesell PC90 por el Diagnóstico Global del área adaptativa de Denver II según criterio de Glascoe. (Grupo de edad de 25-30 meses)**

Nivel de confianza: 95.0%

Prueba diagnóstica	Prueba de referencia		Total
	alterado	salteración	
alterado	3	0	3
sin alteración	118	140	258
Total	121	140	261

	Valor	IC (95%)	
Sensibilidad (%)	2.48	0.00	5.66
Especificidad (%)	100.00	99.64	100.00
Índice de validez (%)	54.79	48.56	61.02
Valor predictivo + (%)	100.00	83.33	100.00
Valor predictivo - (%)	54.26	47.99	60.54
Prevalencia (%)	46.36	40.12	52.60

**Tabla 13. Análisis del Diagnóstico Global del área de lenguaje de Gesell PC90 por el Diagnóstico Global del área de lenguaje de Denver II según criterio de Glascoe. (Toda la población)**

Nivel de confianza: 95.0%

Prueba diagnóstica	Prueba de referencia		Total
	alterado	salteración	
alterado	411	134	545
sin alteración	739	637	1376
<b>Total</b>	<b>1150</b>	<b>771</b>	<b>1921</b>

	Valor	IC (95%)	
Sensibilidad (%)	35.74	32.93	38.55
Especificidad (%)	82.62	79.88	85.36
Índice de validez (%)	54.55	52.30	56.81
Valor predictivo + (%)	75.41	71.71	79.12
Valor predictivo - (%)	46.29	43.62	48.96
Prevalencia (%)	59.86	57.65	62.08

**Tabla 14. Análisis del Diagnóstico Global del área de lenguaje de Gesell PC90 por el Diagnóstico Global del área de lenguaje de Denver II según criterio de Glascoe. (Grupo de edad de 0-6 meses)**

Nivel de confianza: 95.0%

Prueba diagnóstica	Prueba de referencia		Total
	alterado	salteración	
alterado	85	72	157
sin alteración	51	148	199
<b>Total</b>	<b>136</b>	<b>220</b>	<b>356</b>

	Valor	IC (95%)	
Sensibilidad (%)	62.50	54.00	71.00
Especificidad (%)	67.27	60.85	73.70
Índice de validez (%)	65.45	60.37	70.53
Valor predictivo + (%)	54.14	46.03	62.25
Valor predictivo - (%)	74.37	68.05	80.69
Prevalencia (%)	38.20	33.01	43.39

**Tabla 15. Análisis del Diagnóstico Global del área de lenguaje de Gesell PC90 por el Diagnóstico Global del área de lenguaje de Denver II según criterio de Glascoe. (Grupo de edad 7-12 meses)**

Nivel de confianza: 95.0%

Prueba diagnóstica	Prueba de referencia		Total
	alterado	salteración	
alterado	178	60	238
sin alteración	132	103	235
Total	310	163	473

	Valor	IC (95%)	
Sensibilidad (%)	57.42	51.75	63.08
Especificidad (%)	63.19	55.48	70.90
Índice de validez (%)	59.41	54.88	63.94
Valor predictivo + (%)	74.79	69.06	80.52
Valor predictivo - (%)	43.83	37.27	50.39
Prevalencia (%)	65.54	61.15	69.93

**Tabla 16. Análisis del Diagnóstico Global del área de lenguaje de Gesell PC90 por el Diagnóstico Global del área de lenguaje de Denver II según criterio de Glascoe. (Grupo de edad 13-18 meses)**

Nivel de confianza: 95.0%

Prueba diagnóstica	Prueba de referencia		Total
	alterado	salteración	
alterado	73	1	74
sin alteración	224	64	288
Total	297	65	362

	Valor	IC (95%)	
Sensibilidad (%)	24.58	19.51	29.64
Especificidad (%)	98.46	94.70	100.00
Índice de validez (%)	37.85	32.71	42.98
Valor predictivo + (%)	98.65	95.34	100.00
Valor predictivo - (%)	22.22	17.25	27.20
Prevalencia (%)	82.04	77.95	86.14

**Tabla 17. Análisis del Diagnóstico Global del área de lenguaje de Gesell PC90 por el Diagnóstico Global del área de lenguaje de Denver II según criterio de Glascoe. (Grupo de edad 19-24 meses)**

Nivel de confianza: 95.0%

Prueba diagnóstica	Prueba de referencia		Total
	alterado	salteración	
alterado	26	0	26
sin alteración	159	107	266
Total	185	107	292

	Valor	IC (95%)	
Sensibilidad (%)	14.05	8.78	19.33
Especificidad (%)	100.00	99.53	100.00
Índice de validez (%)	45.55	39.66	51.43
Valor predictivo + (%)	100.00	98.08	100.00
Valor predictivo - (%)	40.23	34.14	46.31
Prevalencia (%)	63.36	57.66	69.05

**Tabla 18. Análisis del Diagnóstico Global del área de lenguaje de Gesell PC90 por el Diagnóstico Global del área de lenguaje de Denver II según criterio de Glascoe. (Grupo de edad 25-30 meses)**

Nivel de confianza: 95.0%

Prueba diagnóstica	Prueba de referencia		Total
	alterado	salteración	
alterado	38	0	38
sin alteración	112	111	223
Total	150	111	261

	Valor	IC (95%)	
Sensibilidad (%)	25.33	18.04	32.63
Especificidad (%)	100.00	99.55	100.00
Índice de validez (%)	57.09	50.89	63.28
Valor predictivo + (%)	100.00	98.68	100.00
Valor predictivo - (%)	49.78	42.99	56.56
Prevalencia (%)	57.47	51.28	63.66

**Tabla 19. Análisis del Diagnóstico Global del área personal-social de Gesell PC90 por el Diagnóstico Global del área personal-social de Denver II según criterio de Glascoe. (Toda la población)**

Nivel de confianza: 95.0%

Prueba diagnóstica	Prueba de referencia		Total
	alterado	salteración	
alterado	428	102	530
sin alteración	796	595	1391
<b>Total</b>	<b>1224</b>	<b>697</b>	<b>1921</b>

	Valor	IC (95%)	
Sensibilidad (%)	34.97	32.25	37.68
Especificidad (%)	85.37	82.67	88.06
Índice de validez (%)	53.25	51.00	55.51
Valor predictivo + (%)	80.75	77.30	84.21
Valor predictivo - (%)	42.77	40.14	45.41
Prevalencia (%)	63.72	61.54	65.89

**Tabla 20. Análisis del Diagnóstico Global del área personal-social de Gesell PC90 por el Diagnóstico Global del área personal-social de Denver II según criterio de Glascoe. (Grupo de edad de 0-6 meses)**

Nivel de confianza: 95.0%

Prueba diagnóstica	Prueba de referencia		Total
	alterado	salteración	
alterado	34	2	36
sin alteración	228	92	320
<b>Total</b>	<b>262</b>	<b>94</b>	<b>356</b>

	Valor	IC (95%)	
Sensibilidad (%)	12.98	8.72	17.24
Especificidad (%)	97.87	94.42	100.00
Índice de validez (%)	35.39	30.29	40.50
Valor predictivo + (%)	94.44	85.57	100.00
Valor predictivo - (%)	28.75	23.63	33.87
Prevalencia (%)	73.60	68.88	78.32

**Tabla 21. Análisis del Diagnóstico Global del área personal-social de Gesell PC90 por el Diagnóstico Global del área personal-social de Denver II según criterio de Glascoe. (Grupo de edad de 7-12 meses)**

Nivel de confianza: 95.0%

Prueba diagnóstica	Prueba de referencia		Total
	alterado	salteración	
alterado	92	30	122
sin alteración	162	189	351
Total	254	219	473

	Valor	IC (95%)	
Sensibilidad (%)	36.22	30.11	42.33
Especificidad (%)	86.30	81.52	91.08
Índice de validez (%)	59.41	54.88	63.94
Valor predictivo + (%)	75.41	67.36	83.46
Valor predictivo - (%)	53.85	48.49	59.20
Prevalencia (%)	53.70	49.10	58.30

**Tabla 22. Análisis del Diagnóstico Global del área personal-social de Gesell PC90 por el Diagnóstico Global del área personal-social de Denver II según criterio de Glascoe. (Grupo de edad de 13-18 meses)**

Nivel de confianza: 95.0%

Prueba diagnóstica	Prueba de referencia		Total
	alterado	salteración	
alterado	132	27	159
sin alteración	108	95	203
Total	240	122	362

	Valor	IC (95%)	
Sensibilidad (%)	55.00	48.50	61.50
Especificidad (%)	77.87	70.09	85.65
Índice de validez (%)	62.71	57.59	67.83
Valor predictivo + (%)	83.02	76.87	89.17
Valor predictivo - (%)	46.80	39.69	53.91
Prevalencia (%)	66.30	61.29	71.31

**Tabla 23. Análisis del Diagnóstico Global del área personal-social de Gesell PC90 por el Diagnóstico Global del área personal-social de Denver II según criterio de Glascoe. (Grupo de edad de 19-24 meses)**

Nivel de confianza: 95.0%

Prueba diagnóstica	Prueba de referencia		Total
	alterado	salteración	
alterado	65	1	66
sin alteración	168	58	226
Total	233	59	292

	Valor	IC (95%)	
Sensibilidad (%)	27.90	21.92	33.87
Especificidad (%)	98.31	94.16	100.00
Índice de validez (%)	42.12	36.29	47.96
Valor predictivo + (%)	98.48	94.78	100.00
Valor predictivo - (%)	25.66	19.75	31.58
Prevalencia (%)	79.79	75.02	84.57

**Tabla 24. Análisis del Diagnóstico Global del área personal-social de Gesell PC90 por el Diagnóstico Global del área personal-social de Denver II según criterio de Glascoe. (Grupo de edad de 25-30 meses)**

Nivel de confianza: 95.0%

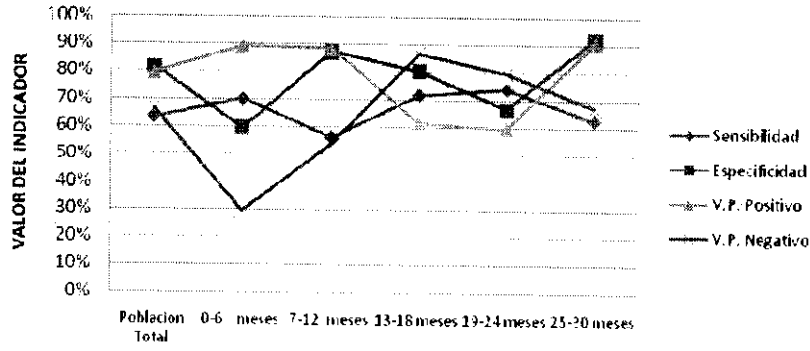
Prueba diagnóstica	Prueba de referencia		Total
	alterado	salteración	
alterado	58	26	84
sin alteración	81	96	177
Total	139	122	261

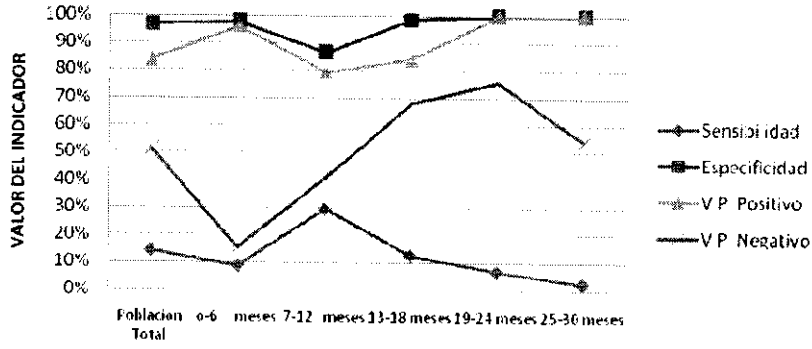
	Valor	IC (95%)	
Sensibilidad (%)	41.73	33.17	50.28
Especificidad (%)	78.69	71.01	86.36
Índice de validez (%)	59.00	52.85	65.16
Valor predictivo + (%)	69.05	58.57	79.53
Valor predictivo - (%)	54.24	46.62	61.86
Prevalencia (%)	53.26	47.01	59.50

**ANEXO 4. Gráficas de los indicadores de validez de la prueba  
Denver II según área del desarrollo evaluada para la población  
total y por rangos de edad.**

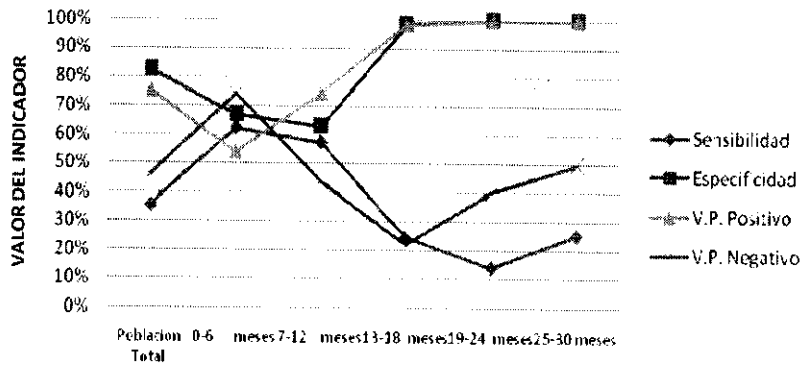
**GRAFICA 1.- INDICADORES DE VALIDEZ EN EL AREA MOTORA. POBLACIÓN TOTAL Y POR RANGOS DE EDAD**



**GRAFICA 2.- INDICADORES DE VALIDEZ EN EL AREA ADAPTATIVA. POBLACIÓN TOTAL Y POR RANGOS DE EDAD**



**GRAFICA 3.-INDICADORES DE VALIDEZ EN EL AREA DE LENGUAJE. POBLACIÓN TOTAL Y POR RANGOS DE EDAD**



**GRAFICA 4.-INDICADORES DE VALIDEZ EN EL AREA PERSONAL SOCIAL. POBLACIÓN TOTAL Y POR RANGOS DE EDAD**

