



BAYLEY-III

BAYLEY SCALES OF INFANT AND TODDLER DEVELOPMENT – THIRD EDITION

EINFÜHRUNG IN DIE BAYLEY-III SKALEN UND DEN BAYLEY-III SCREENING TEST
INKLUSIVE EINES VERGLEICHS MIT DER VORGÄNGERVERSION (BAYLEY-II)

WHITE PAPER

BAYLEY-III

BAYLEY SCALES OF INFANT AND TODDLER DEVELOPMENT – THIRD EDITION

EINFÜHRUNG IN DIE BAYLEY-III SKALEN UND DEN BAYLEY-III SCREENING TEST
INKLUSIVE EINES VERGLEICHS MIT DER VORGÄNGERVERSION (BAYLEY-II)

Daniela Schultheiss, Mag.
im Auftrag der Pearson Assessment & Information GmbH

Februar 2015

PEARSON

Executive Summary

Ziel dieses White Papers ist es, Ärzte und Psychologen in pädiatrischen Einrichtungen mit der dritten Auflage der Bayley Scales of Infant and Toddler Development (Bayley-III) in deutscher Übersetzung (Reuner & Rosenkranz, 2014) vertraut zu machen. Im Vergleich zu ihrer Vorgängerversion (Bayley-II) verfügt die Revision über wichtige Neuerungen, die das Testverfahren zur frühkindlichen Entwicklungsdiagnostik noch messgenauer, differenzierter und praktikabler machen.

Die neue Skalenstruktur der Bayley-III ermöglicht eine getrennte Erfassung des kognitiven, sprachlichen und motorischen Entwicklungsstands von sehr jungen Kindern zwischen einem und 42 Monaten. Die Unterteilung in die fünf Untertests Kognition, Sprache Rezeptiv, Sprache Expressiv, Feinmotorik und Grobmotorik garantiert eine differenzierte und präzise Diagnostik einzelner Leistungsbereiche. Inhaltlich wurde die kognitive Skala um das Konzept „Spielverhalten“ erweitert, was als wichtige Information zur Einschätzung des Entwicklungsniveaus von Kleinkindern dient. Neben der rechtzeitigen Diagnostik von Entwicklungsverzögerungen eignet sich die Bayley-III darüber hinaus zur Testung von klinischen Patientengruppen wie etwa ehemals Frühgeborenen, Kindern mit angeborenen Herzfehlern, tiefgreifenden Entwicklungsstörungen oder chronischen Erkrankungen. Die wichtigste Neuerung der dritten Version der Bayley-Skalen besteht jedoch in einer erstmals deutschen Normierung. Aufgrund der großen Altersspanne eignet sich der Individualtest auch ideal für Langzeitbeobachtungen. Seine internationale Verbreitung gewährleistet zudem die Vergleichbarkeit von Forschungsarbeiten zur frühkindlichen Entwicklungsdiagnostik. Als zeitökonomische Alternative bietet sich zudem der dazu gehörige Bayley-III Screening Test an. Mit der Kurzversion des Tests wird eine erste Grobklassifizierung vorgenommen, die als Entscheidungshilfe für weitere Untersuchungen dienen kann.

Inhaltlich und messtheoretisch entspricht die Bayley-III dem aktuellen Expertenwissen aus Forschung und Praxis. Somit trägt der Entwicklungstest zur Verbesserung und Weiterentwicklung der Entwicklungsdiagnostik sehr junger Kinder bei und bietet neben einer präzisen Einschätzung des individuellen Entwicklungsstatus auch Empfehlungen zu Fördermaßnahmen.

Inhaltverzeichnis

Executive Summary	3
1. Einleitung	5
2. Zweck des Verfahrens	6
3. Patientengruppe	6
4. Anwender des Verfahrens	7
5. Durchführungsdauer	7
6. Unterschiede zwischen Bayley-III und Bayley-II	8
7. Bayley-III Screening Test	10
8. Die Untertests	11
9. Klinische Gruppen	14
10. Auswertung	15
11. Anwendungsgrenzen	16
12. Zusammenfassung	17
13. Literatur	18

I. Einleitung

Um die schwerwiegenden Folgen von frühkindlichen Entwicklungsabweichungen so gering wie möglich zu halten, sind nicht nur geeignete Frühfördermaßnahmen unabdingbar, sondern allem voran sind präzise diagnostische Verfahren gefragt. Gerade bei Säuglingen und sehr jungen Kindern fällt die Erfassung des aktuellen Entwicklungsniveaus besonders schwer. Mit der 2014 erschienenen dritten Auflage der Bayley Scales of Infant and Toddler Development (Bayley-III) in deutscher Fassung wird Kinderärzten und Psychologen ein Instrument zur Verfügung gestellt, welches das kognitive, sprachliche und motorische Funktionsniveau von Säuglingen und Kleinkindern messgenau, differenziert und praktikabel diagnostiziert.

Gerade im Bereich der frühkindlichen Entwicklungsdiagnostik ist es unabdingbar, bereits vorhandene Testverfahren kontinuierlich dem aktuellen Forschungsstand anzupassen sowie die testtheoretischen Charakteristika weiterhin zu optimieren. Für die Diplompsychologen Dr. Gitta Reuner und Joachim Rosenkranz hatte es daher oberste Priorität, bei der Übersetzung der Manuale den kulturellen und sprachlichen Besonderheiten gerecht zu werden sowie die Durchführung einer Normierungserhebung so präzise wie möglich zu planen. Das Ergebnis ist ein komplett überarbeitetes Testverfahren, das sich nicht nur durch inhaltliche und strukturelle Neuerungen, sondern vor allem auch durch eine erstmals komplett deutsche Normierung auszeichnet.

Was ist nun neu an der dritten Revision der Bayley-Skalen? Wodurch ist die Bayley-III ihrer Vorgängerversion überlegen? Eine wesentliche Veränderung betrifft die neu angelegte, dreigliedrige Skalenstruktur, die mit fünf Untertests die Leistungsbereiche Kognition, Sprache (Rezeptiv und Expressiv) und Motorik (Fein- und Grobmotorik) messgenau und differenziert erfasst. So werden in den Bayley-III Skalen erstmals kognitive und sprachliche Leistungen getrennt voneinander erhoben, was die Eindeutigkeit der Skalen erhöht. Inhaltlich wurde die kognitive Skala durch das Hinzufügen von Items um das Konstrukt des Spielverhaltens ergänzt, welches ein wichtiges diagnostisches Kriterium für die Erfassung des frühkindlichen Entwicklungsstands darstellt.

Weiterhin eignet sich der enorm große Altersbereich von einem bis 42 Monaten besonders gut für Langzeitbeobachtungen, die die Wirksamkeit von Fördermaßnahmen überprüfen können. Die Bayley-III dient nicht nur zur frühzeitigen Diagnostik von Entwicklungsabweichungen, sondern auch zur Beobachtung von klinischen Patientengruppen wie ehemals Frühgeborenen, Kindern mit angeborenem Herzfehler, tiefgreifender Entwicklungsstörung oder Trisomie 21. Die herausragendste Leistung der deutschsprachigen Neuauflage besteht jedoch in der aktuellen und erstmaligen deutschen Normierung, die die Diagnostik mit den Bayley-III Skalen in den deutschen Bevölkerungs- und Kulturraum einbettet.

Durch die inhaltlichen, strukturellen und messtheoretischen Verbesserungen liegt mit der Bayley-III nun für Kinderärzte und Psychologen in pädiatrischen Einrichtungen ein frühkindlicher Entwicklungstest vor, der eine präzise, differenzierte Leistungsbereiche abdeckende und leicht durchführbare Messung des Entwicklungsniveaus von sehr jungen Kindern ermöglicht. Zusätzlich zur Bayley-III steht ein Screening Test zur Verfügung, der zeitsparend eine erste Grobklassifizierung erlaubt und als Entscheidungshilfe dafür dienen kann, ob eine umfassende Untersuchung empfehlenswert ist.

Ziel dieses Papers ist es, dem Leser einen genauen Überblick über den Inhalt, die Struktur und wichtige Neuerungen bei der Bayley-III zu geben. So werden in einzelnen Abschnitten der Zweck des Verfahrens sowie Details zur Patientengruppe, zu den Anwendern des Individualtests und zur Durchführungsdauer vorgestellt. Des Weiteren werden die zentralen Unterschiede zwischen der Bayley-II und Bayley-III näher beleuchtet und der Bayley-III Screening Test vorgestellt. Ab S. 12 werden die fünf Untertests detaillierter umrissen und es werden wichtige Informationen zu den klinischen Gruppen sowie der Auswertung des Entwicklungstests gegeben.

2. Zweck des Verfahrens

Bei der Bayley-III handelt es sich um die deutsche Adaptation der Bayley Scales of Infant and Toddler Development, die von den Diplompsychologen Dr. Gitta Reuner und Joachim Rosenkranz für den deutschsprachigen Markt herausgegeben wird. Die Bayley-III folgt im Aufbau dem Prinzip eines Stufenleiterverfahrens (Petermann & Macha, 2005), indem die Aufgaben der insgesamt fünf Skalen jeweils nach ihrem Schwierigkeitsgrad in einer Reihenfolge angeordnet sind. Der Individualtest dient der frühkindlichen Entwicklungsdiagnostik bei Säuglingen und Kleinkindern zwischen einem und 42 Monaten. Anhand dreier Skalen und fünf Untertests werden die Leistungen von Kindern im kognitiven, sprachlichen (rezeptiv und expressiv) sowie motorischen (fein- und grobmotorisch) Bereich gemessen. Mithilfe des international verbreiteten Verfahrens werden Entwicklungsverzögerungen bei Säuglingen und Kleinkindern ökonomisch und detailliert identifiziert. Neben der präzisen Einschätzung des frühkindlichen Entwicklungsniveaus können des Weiteren durch die umfassende Entwicklungsdiagnostik Förderungsmöglichkeiten bei Entwicklungsabweichungen abgeleitet werden.

Aufgrund der differenzierten Diagnostik mittels der drei Skalen Kognition, Sprache und Motorik sowie der großen Altersspanne zählt der Test zu einem der wenigen adäquaten entwicklungsdiagnostischen Verfahren zur Testung von Säuglingen und Kleinkindern auf dem deutschen Markt. Der große Altersrange zwischen einem und 42 Monaten eignet sich darüber hinaus sehr gut für Langzeitbeobachtungen der Entwicklungsverläufe von Kindern.

Als Grundlage für die Einschätzung der Testgüte der Bayley-III dient die deutsche Normstichprobe. Die Reliabilitätskoeffizienten der Untertests liegen zwischen $r = .77$ (Sprache rezeptiv) und $r = .89$ (Grobmotorik). Für die Skalen fallen die Werte höher aus: Hier liegen die Reliabilitätskoeffizienten zwischen $r = .86$ (Sprache) und $r = .88$ (Motorik). Die interne Konsistenz der Bayley-III Skalen kann somit als gut erachtet werden. Messwiederholungserhebungen zur Feststellung der Retest-Reliabilität für die deutsche Normierung sind bereits in Planung. Auch die Frage, ob der Test wirklich das misst, was gemessen werden soll (Validität), kann positiv beantwortet werden: So gelten sowohl die Inhalts- als auch die Konstruktvalidität als nachgewiesen und für die Kriteriumsvalidität liegen ebenfalls erste Befunde vor.

3. Patientengruppe

Die Bayley-III richtet sich an Säuglinge und Kleinkinder im Alter zwischen einem und 42 Monaten, deren Entwicklungsniveau in den Bereichen Kognition, Sprache und Motorik eingeschätzt werden soll. Das entwicklungsdiagnostische Verfahren kann bei dieser Patientengruppe Entwicklungsauffälligkeiten feststellen und frühzeitige Interventionen empfehlen. Speziell die dritte Auflage der Bayley Scales dient darüber hinaus zur Untersuchung bestimmter klinischer Gruppen wie ehemals Frühgeborene oder Kinder mit angeborenem Herzfehler, die sich durch ein erhöhtes Risiko für Entwicklungsstörungen auszeichnen (vgl. Kapitel 9).

4. Anwender des Verfahrens

Zu den primären Anwendern der Bayley-III zählen Psychologen sowie Kinder- und Jugendpsychotherapeuten/-psychiater. Darüber hinaus kann das Verfahren auch von anderen Berufsgruppen, wie (Kinder-)Ärzten, Pädagogen, Ergotherapeuten, Logopäden, Physiotherapeuten etc. durchgeführt werden, wenn sie ausreichende Kenntnisse in den Bereichen Psychometrie, frühkindliche kognitive, sprachliche und motorische Entwicklung nachweisen können.

Alle Anwender der Bayley-III sollten aufgrund der komplexen Testdurchführung und Interpretation der Testergebnisse unbedingt im Umgang mit dem Testverfahren geschult sein und über Expertise in der Säuglings- und Kleinkinddiagnostik sowie fundierte Kenntnisse der normalen frühkindlichen Entwicklung verfügen.

5. Durchführungsdauer

Die Neuauflage der Bayley-Skalen ist nicht nur komfortabel durchzuführen, sondern erstreckt sich auch über einen sehr überschaubaren Zeitraum bei der Durchführung. Die gesamte Testbatterie der Bayley-III mit den drei Skalen und fünf Untertests in den Bereichen Kognition, Sprache - Rezeptiv, Sprache - Expressiv, Fein- und Grobmotorik umfasst 324 Items, die immer nur einmal vorgegeben und nicht wiederholt dargeboten werden dürfen. Bei Säuglingen und Kleinkindern zwischen einem und 12 Monaten beträgt die Durchführungsdauer des Tests etwa 50 Minuten. Bei Klein- und Vorschulkindern zwischen 13 und 42 Monaten können circa 90 Minuten veranschlagt werden.

Zu beachten gilt, dass möglichst alle Untertests in einer Sitzung bearbeitet werden sollten. Die Reihenfolge der Untertests kann bei der Durchführung beliebig variiert werden. Ist ein Kind jedoch enorm unruhig oder zeigt sich auffällig unmotiviert, kann durchaus eine kurze Pause zwischen den Untertests eingelegt werden. Die erwähnten Zeitangaben dienen als grobe Richtwerte. Generell kann die Durchführungsdauer je nach Konzentration, Ausdauer, Motivation des Kindes und den äußeren Umständen individuell variieren. Bei der Testdurchführung sollte daher immer bedürfnisorientiert auf das jeweilige Kind eingegangen werden.

Geht es jedoch vorerst nur um eine erste Überprüfung des Entwicklungsstands eines Kindes, empfiehlt sich zur groben Ersteinschätzung der Bayley-III Screening Test, der mit 15 bis 30 Minuten Dauer noch zeitökonomischer durchzuführen ist (vgl. Kapitel 7).

Durchführungsdauer:

Bayley III:

Säuglinge und Kleinkinder zwischen 1 und 12 Monaten: ca. 50 Minuten
Kleinkinder zwischen 13 und 42 Monaten: ca. 90 Minuten

Bayley III Screening Test:

Säuglinge und Kleinkinder zwischen 1 und 12 Monaten: 15 - 20 Minuten
Kleinkinder zwischen 13 und 42 Monaten: ca. 30 Minuten

6. Unterschiede zwischen Bayley-III und Bayley-II

Als Anwender, der mit der deutschen Version der Bayley-II von 2008 vertraut ist, kann man sich berechtigterweise fragen, was für die Anschaffung der Neuauflage des Entwicklungstests spricht und ob sich die Investition für die eigene Berufspraxis lohnt. Im Folgenden soll erläutert werden, was dringend für eine Überarbeitung des Bayley-II sprach, was die Revisionsziele für die neue Fassung waren und durch welche zentralen Merkmale sich die Bayley-III Skalen von ihrer Vorgängerversion unterscheiden.

Revisionsziele

Nach intensivem Austausch zwischen Experten und Expertinnen sowohl aus der Forschung als auch aus der Praxis einigte man sich auf folgende Revisionsziele, die die Messgenauigkeit, die Aktualität sowie die Praktikabilität des Individualtests für Säuglinge und Kleinkinder weiter optimieren sollten.

An erster Stelle stand die Notwendigkeit, die bestehenden Normen zu aktualisieren. Die 2008 in Deutschland erschienene Übersetzung der Bayley Scales of Infant Development – Second Edition (BSID-II; Bayley, 1993) hatte sowohl die Belege für die Gütekriterien der Reliabilität und Validität als auch die Normierung der Originalversion adaptiert. Daher wurde es notwendig, sich um eine Normierung anhand einer repräsentativen deutschen Stichprobe zu bemühen und dahingehend auch die psychometrische Gütequalität der Bayley-Skalen zu verbessern.

Ein weiteres wichtiges Ziel bestand in der Steigerung des klinischen Nutzens. Konzentrierte sich die zweite Fassung der Bayley-Skalen vornehmlich auf die Diagnostik von Entwicklungsabweichungen bei Säuglingen und Kleinkindern, zeichnet sich die Neuauflage in der dritten Version zusätzlich dadurch aus, dass mit ihr auch spezielle klinische Patientengruppen wie ehemals Frühgeborene, Kinder mit Trisomie 21, Kinder mit angeborenem Herzfehler oder tiefgreifender Entwicklungsstörung beobachtet und behandelt werden können (Details siehe Kapitel 9).

Weiterhin sollte durch eindeutig festgelegte (altersbezogene) Startpunkte und einfache Einstiegs-, Umkehr- und Ausstiegsregeln die Durchführung des Testverfahrens auf Seiten der Anwender vereinfacht werden. Mithilfe der Aktualisierung der Item-Vorgaben sowie des Testmaterials soll ferner die Motivation seitens der Testteilnehmer (Säuglinge und Kleinkinder) gesteigert werden.

Überblick über die Revisionsziele der Bayley-III:

- Aktualisierung der Normen
- Verbesserung der psychometrischen Güte
- Erhöhung des klinischen Nutzens
- Vereinfachung der Testdurchführung
- Aktualisierung der Item-Vorgabe
- Aktualisierung der Stimulus-Materialien
- Erhalt der grundlegenden Qualitäten der Bayley-Skalen

Wesentliche Neuerungen

Basierend auf den oben genannten Revisionszielen zeichnet sich die Neuauflage der deutschen Version der Bayley-III durch drei zentrale Neuerungen aus, die die Qualität des Entwicklungstests wesentlich verbessern.

1. Neue Skalenstruktur

Die Bayley-III besteht, wie bereits in den vorherigen Kapiteln angesprochen, in ihrer neuen Struktur aus fünf Untertests (Kognition, Sprache Rezeptiv, Sprache Expressiv, Feinmotorik, Grobmotorik) mit insgesamt 324 Items, von denen etwa ein Drittel neu hinzugefügt wurden. Die fünf Untertests können wiederum den drei Skalen Kognition, Sprache und Motorik zugeordnet werden (vgl. Kapitel 8). Der wesentliche Vorzug dieser neuen Skalenstruktur im Vergleich zur Bayley-II besteht darin, dass nun sprachliche von nicht-sprachlichen kognitiven Leistungen differenziert erfasst werden können. In der Vorgängerversion flossen beide Facetten gleichermaßen in den Standardwert der kognitiven Skala ein, was sich vor allem bei Kleinkindern ab 18 Monaten als schwierig erwies. Durch die klare Trennung des kognitiven vom sprachlichen Bereich können nun Konfundierungen vermieden werden und die Skalen gewinnen an Eindeutigkeit, was zielführendere sowie präzisere Diagnosen ermöglicht.

Weiterhin ermöglichen in der Bayley-III neu hinzugefügte Items auf der kognitiven Skala die Erfassung des Konzepts „Spielverhalten“, welches gerade bei Kleinkindern als wertvolle diagnostische Information dient.

2. Aktuelle deutsche Normierung

Die Normierung der Bayley-III Skalen erfolgte erstmals anhand einer nach Geschlecht ausgewogenen und für Deutschland repräsentativen Stichprobe (N = 1.009), in der auch der große Altersrange von 42 Monaten berücksichtigt und auf 17 Altersgruppen verteilt wurde.

Charakteristika des Bayley-III:

- Erstmals deutsche Normierung
- Differenziertere Skalenstruktur und Erweiterung um Spielverhalten
- Vereinfachung der Durchführungsregeln
- Hohe Motivation der Kinder durch überarbeitetes Testmaterial
- Enorm großer Altersbereich von einem bis 42 Monate abgedeckt
- Geeignet für klinische Patientengruppen

3. Verbesserte Anwendbarkeit

Aufgrund der sukzessiven Itemvorgabe, eindeutig festgelegter Startpunkte für jede Altersstufe sowie einfacher Ein- und Ausstiegsregeln (vgl. Kapitel 8) sind die Skalen und Untertests der Bayley-III nun wesentlich leichter zu erlernen und anzuwenden als in der Vorgängerversion. Auch für die Kinder ist der revidierte Test angenehmer durchzugehen: Das zeitgemäß angepasste Testmaterial und die für jede Altersstufe leicht verständlichen Anweisungen steigern die Motivation der Kleinkinder sich aktiv mit den Items auseinanderzusetzen.

Mit der revidierten Neuauflage der Bayley-Skalen ist nun ein Instrument zur frühkindlichen Entwicklungsdiagnostik entstanden, das einfach in der Anwendung ist und präzise Messungen ermöglicht. Diese Vorzüge der Bayley-III-Skalen in Verbindung mit der internationalen Verbreitung des Tests sowie seines internationalen Validierungsstands bietet derzeit kein anderes Testverfahren zur frühkindlichen Diagnostik im deutschsprachigen Raum.

7. Bayley-III Screening Test

Beim Bayley-III Screening Test handelt es sich um eine Kurzform der Bayley-III Skalen, die einer schnellen Untersuchung des kognitiven, sprachlichen und motorischen Funktionsniveaus dient. Diese Kurzvariante, die ebenfalls als Individualtest konzipiert ist, eignet sich insbesondere für Fälle, in denen ein schnelles Screening des Entwicklungsstands gefragt ist. Die Ergebnisse des Screening Tests geben dem Anwender Aufschluss darüber, ob eine weitere umfassendere Untersuchung mit der Langform der Bayley-III Skalen notwendig ist. Besonders gut eignet sich der Bayley-III Screening Test zur Anwendung bei großen Risikogruppen, wo aller Wahrscheinlichkeit nach viele Kinder als „unauffällig“ klassifiziert werden. Hier wäre es sehr zeitaufwendig und unökonomisch mit jedem Patienten die Langform des frühkindlichen Entwicklungstests durchzuführen.

Der Screening Test dient somit einer ersten Grobklassifizierung und stellt für Ärzte und Psychologen eine Entscheidungshilfe darüber dar, ob eine umfassende Testung empfehlenswert ist oder nicht. Mit wenig Zeitaufwand erhält man ausreichende Informationen für eine erste Einschätzung, was das Verfahren sehr zeitökonomisch macht. Die Durchführungsdauer beträgt bei Säuglingen und Kleinkindern zwischen einem und 12 Monaten etwa 15 bis 20 Minuten sowie bei Kindern zwischen 13 und 42 Monaten circa 30 Minuten.

Durchführungsdauer des Bayley-III Screening Tests:

Säuglinge und Kleinkinder zwischen 1-12 Monaten:	15 - 20 Minuten
Kleinkinder zwischen 13 und 42 Monaten:	ca. 30 Minuten

Inhaltlich besteht der Bayley-III Screening Test aus insgesamt 142 Items der Bayley-III Skalen und enthält vom Umfang her somit weniger als die Hälfte der Langversion, ohne dabei jedoch an Messgenauigkeit zu verlieren (s. u.). Die Itemauswahl erfolgte ebenfalls wie bei den Bayley-III Skalen nach inhaltlichen Kriterien, psychometrischen Charakteristika sowie anhand der klinischen Relevanz. Sowohl die Skalenstruktur (vgl. Kapitel 8) als auch die Itemvorgabe und die Auswertungskriterien (vgl. Kapitel 10) sind mit der Langversion identisch. Für jeden Untertest wird ein eigener Cut-Off Wert ermittelt. Anders als bei der Bayley-III erfolgt bei der Auswertung des Screening Tests in jedem der fünf Bereiche eine Klassifizierung des Entwicklungsstands anhand der Kategorien „Auffällig“, „Gefährdet“ oder „Unauffällig“.

Da das Screening Säuglinge und Kleinkinder mit Werten in den unteren 25 Prozenträngen identifizieren soll, wurden für den Bayley-III Screening Test bewusst leichte Items aus der Langversion ausgewählt. Konkret waren dies Items, die durchschnittlich von 25 Prozent der Kinder nicht gelöst werden.

Itemverteilung über die Untertests beim Bayley-III Screening Test:

- Kognition: 34 Items
- Sprache Rezeptiv: 26 Items
- Sprache Expressiv: 25 Items
- Feinmotorik: 28 Items
- Grobmotorik: 29 Items

Die Qualität der Gütekriterien des Screening Tests ist vergleichbar mit der der Bayley-III Skalen. So verfügt der Bayley-III Screening Test über eine hohe interne Konsistenz (verteilt über die fünf Untertests zwischen 0.82 und 0.88) sowie über eine hohe zeitüberdauernde Stabilität, die sich mit zunehmenden Alter noch leicht erhöht (Retest-Reliabilitätskoeffizienten zwischen 0.80 und 0.83). Die Kürzung der Items führt somit nicht zu Einbußen bei der Messqualität. Weiterhin liegen auch Angaben zur Sensitivität und Spezifität des Screening Tests vor – also zum positiven und negativen Vorhersagewert der einzelnen Untertests. Die Genauigkeit der Diagnose liegt im moderaten Bereich, während die Falsch-Negativ-Rate (sprich falsch identifizierte) hervorragend ist. Diese messtheoretischen Kriterien dienen als wichtige Belege für die diagnostische Genauigkeit des Screening Tests im Bereich erster grober Klassifizierungen.

8. Die Untertests

Die dritte Version der Bayley-Skalen vermag es aufgrund ihrer neuen Skalenstruktur nun zwischen sprachlichen und nicht-sprachlichen kognitiven Leistungen zu differenzieren. Die insgesamt 324 Items der Bayley-III verteilen sich auf die drei Skalen der Kognition, Sprache und Motorik. Die Sprach-Skala wurde in der überarbeiteten Version noch einmal in Sprache-Rezeptiv und Sprache-Expressiv unterteilt. Gleiches gilt für die Skala Motorik: Hier wird nun zwischen Feinmotorik und Grobmotorik differenziert. Somit enthält der Individualtest zur frühkindlichen Entwicklungsdiagnostik in der dritten Auflage fünf Untertests, die in beliebiger und an individuelle Bedürfnisse angepasster Reihenfolge bearbeitet werden können. Aufgrund der sukzessiven Itemvorgabe sowie eindeutig festgelegter Startpunkte und einfach anzuwendenden Einstiegs- und Ausstiegs-Regeln ist die neue Skalenstruktur leicht erlernbar.

Regeln bei der Testdurchführung:

Einstiegs-Regel:

Der Einstieg in den Test erfolgt an dem am Alter des Kindes entsprechenden Startpunkt. Ab diesem muss das Kind für jedes der ersten 3 aufeinanderfolgenden Items 1 Punkt erzielen. Ist dieses Einstiegs-Kriterium erfüllt bzw. die Leistungsbasis erreicht, kann der Test solange fortgeführt werden bis die Ausstiegs-Regel erfüllt ist (s. u.). Die Einstiegs-Regel ist für jeden der fünf Untertests identisch.

Umkehr-Regel:

Erreicht ein Kind bei einem der ersten drei aufeinanderfolgenden Items 0 Punkte, gilt das Einstiegs-Kriterium als nicht erfüllt. In diesem Fall geht der Anwender zurück zum Startpunkt der nächst niedrigeren Altersstufe und bietet die Items in aufsteigender Reihenfolge dar. Die Umkehr-Regel wird so lange angewendet, bis ein Kind jeweils bei den ersten drei Items eines Startpunktes 1 Punkt erzielt.

Ausstiegs-Regel:

Ein Untertest wird dann beendet, wenn bei fünf aufeinanderfolgenden Items jeweils 0 Punkte erzielt werden. In diesem Fall gilt die oberste Leistungsgrenze als erreicht. Alle weiteren Items werden nun mit 0 bewertet.

Im Folgenden werden die fünf Untertests schematisch vorgestellt:

Kognitive Skala (Kog)

Die kognitive Skala der Bayley-III soll erfassen wie der Säugling oder das Kleinkind seine Umgebung wahrnimmt, wie er/es denkt und wie er/es auf Objekte reagiert. Es wird der sensu-motorische Entwicklungsstand, aber auch das Explorations- und Manipulationsverhalten sowie die Objektbezogenheit und Begriffbildungsfähigkeit gemessen. Bei Säuglingen wird etwa ihr Interesse an neuen Dingen oder das Spielen mit verschiedensten Spielsachen beobachtet. Kleinkindern werden unter anderem Puzzle vorgegeben, die sie ergänzen sollen oder man legt ihnen einfache Problemlösungsaufgaben vor. Vorschulkinder beobachtet man beim Phantasiespiel oder etwa dabei, wie sie Gruppen bilden.

Im Vergleich zur Bayley-II wurde in der neuen Testrevision der Einfluss motorischer Fertigkeiten sowie der Einfluss der rezeptiven Sprachfähigkeit zur Itemlösung reduziert, was die Eindeutigkeit der kognitiven Skala in der Bayley-III weiter erhöht. Ein Großteil der Items aus den Bayley-II-Skalen wurde für die dritte Version übernommen und es wurden zusätzlich 19 neue Items hinzugefügt. Insgesamt besteht die kognitive Skala nun aus 91 Items.

Sprach-Skala (Spr)

Die Sprach-Skala der Bayley-III untergliedert sich in die zwei Untertests Sprache-Rezeptiv und Sprache-Expressiv. Viele Items wurden aus der Mental Scale der 2. Auflage der Originalversion (BSID-II; Bayley, 1993) übernommen, die vornehmlich Sprache erfasst. Insgesamt besteht die Sprach-Skala aus 95 Items.

Sprachverständnis – Rezeptiv (SprR)

Der Untertest Sprache Rezeptiv setzt sich aus 49 Items zusammen, von denen im Vergleich zur Vorgängerversion 38 neu sind. Diese Skala erfasst das Sprachverständnis von Säuglingen und Kleinkindern und misst, wie gut sie Geräusche zu erkennen vermögen sowie ihr Wort- und Anweisungsverständnis. Bei Vorschulkindern wird zusätzlich auch das Grammatikverständnis erhoben. Weiterhin wird mit diesem Untertest etwa die passive Wortschatzentwicklung, die rezeptive morphologische Sprachentwicklung sowie das Verstehen von Tätigkeitswörtern diagnostiziert.

Sprachproduktion – Expressiv (SprE)

Beim Untertest Sprache Expressiv geht es um die Erfassung der aktiven Sprachproduktion. Er besteht aus 46 Items, von denen 19 in der Revision neu hinzugefügt wurden. Hier geht es um das Verständlichmachen mit Lauten, Gesten und Wörtern, so dass die Items sowohl präverbale Kommunikation (Brabbeln, Gestikulieren) als auch die aktive Wortschatzentwicklung beinhalten. Während bei Säuglingen die non-verbale Verständigung erfasst wird, sollen Kleinkinder etwa konkret Objekte und Bildinhalte benennen oder einfache Fragen beantworten. Bei Vorschulkindern sind die Fragen dann etwas komplexer aufgebaut und sie werden durch das Zeigen von Bildern zum Sprechen animiert. Die aktive Sprache wird weiterhin durch Zwei-Wort-Äußerungen, der Verwendung von Mehrzahl und Verbformen erhoben.

Motorik-Skala (Mot)

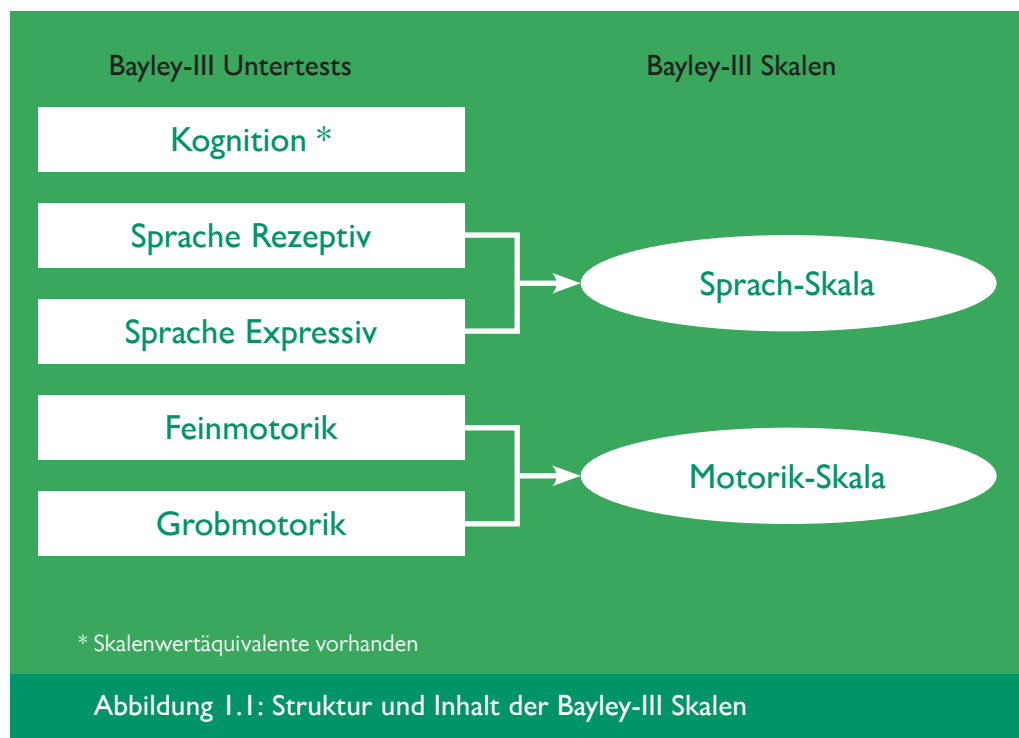
Die Motorik-Skala der Bayley-III setzt sich aus den beiden Untertests Feinmotorik und Grobmotorik zusammen. Insgesamt besteht sie aus 138 Items, die zum Einen aus der BSID-II Motor Scale und Mental Scale entnommen wurden, zum Anderen aber auch neu hinzugefügt wurden. Ein wichtiger Schwerpunkt der Skala liegt auf der Bewegungsqualität.

Feinmotorik (MotF)

Der Untertest Feinmotorik beinhaltet 66 Items und misst, wie Säuglinge und Kleinkinder ihre Finger und Hände einsetzen, um Aufgaben auszuführen. So erfassen die Items das Greifverhalten, die Bewegungsplanung und Bewegungsgeschwindigkeit, aber auch das visuelle Verfolgen von Objekten und Objektmanipulationen. Bei Säuglingen geht es primär um die Erhebung der Muskelkontrolle, während Kleinkinder in den einzelnen Aufgaben dazu aufgefordert werden, Stapel zu bilden, Umrisse zu zeichnen oder kleinere Objekte zu bewegen. Bei Vorschulkindern wird beispielsweise beobachtet, wie sie Papier schneiden oder komplexere Formen malen.

Grobmotorik (MotG)

Die 72 Items des Untertests Grobmotorik werfen einen Blick darauf, wie gut Säuglinge oder Kleinkinder ihren Körper (Extremitäten und Rumpf) bewegen können. Es wird die Entwicklung des Sich-Aufrichtens, der dynamischen Bewegung, der Balance und der Bewegungsplanung beobachtet. So wird etwa beim Säugling erfasst, wie gut er seinen Kopf kontrollieren oder wie geschickt er sich rollen, sitzen und krabbeln kann. Kleinkinder werden in Items dazu aufgefordert, Gewichte zu tragen, und es werden ihre Fähigkeiten beim Stehen und Gehen ohne fremde Hilfe beurteilt. Vorschulkindern sollen beispielsweise gegen Bälle treten, Treppen steigen oder das Gleichgewicht halten.



9. Klinische Gruppen

Ein wichtiger Einsatzbereich der Bayley-III Skalen liegt in der Untersuchung von klinischen Patientengruppen, die einer erhöhten Gefahr von Entwicklungsabweichungen ausgesetzt sind. Für folgende Kinder ist die Beobachtung mit dem Bayley-III besonders geeignet:

- Frühgeborene Kinder oder Kinder mit geringem Geburtsgewicht
- Kinder mit einer angeborenen Herzerkrankung
- Kinder mit Trisomie 21
- Kinder mit tiefgreifender Entwicklungsstörung
- Kinder mit Infantiler Zerebralparese
- Kinder mit oder mit Verdacht auf eine spezifische Sprachentwicklungsstörung
- Kinder mit Risiko einer Entwicklungsstörung
- Kinder mit Asphyxie während der Geburt
- Kinder mit pränataler Exposition mit Alkohol

Gerade bei diesen klinischen Risikogruppen ist eine messgenaue und differenzierte Diagnostik unabdingbar. Aufgrund ihrer sehr zufriedenstellenden psychometrischen Eigenschaften sowie der Feingliederung ihrer (Sub-)Skalen in verschiedene Entwicklungsfacetten eignet sich die Bayley-III ideal für die Anwendung innerhalb dieser Patientengruppen. Neben der präzisen Diagnostik lassen sich mit dem Individualtest auch Empfehlungen zur Frühförderung ableiten. Der große Altersbereich ermöglicht zudem Langzeitbeobachtungen, die die Wirksamkeit von Interventionen untersuchen können.

10. Auswertung

Die Auswertung am Ende des Tests erfolgt anhand des dazu gehörigen Protokollbogens, in dem Antworten, Beobachtungen und die Punktvergabe für jedes Item der jeweiligen Untertests eingetragen werden. Im Protokollbogen der Bayley-III Skalen befindet sich auch ein Abschnitt, in dem Verhaltensbeobachtungen während der Untersuchung vermerkt werden können. Während beim Bayley-III Screening Test anhand der Rohwerte Cut-Off Werte gebildet werden, anhand derer die Ergebnisse aus den einzelnen Untertests in die drei Kategorien „Auffällig“, „Gefährdet“ und „Unauffällig“ eingeordnet werden (vgl. Kapitel 7), findet eine solche Grobklassifizierung bei der Auswertung der Bayley-III Skalen nicht statt. Alternativ werden die Rohwerte in standardisierte Werte umgewandelt. In jedem Untertest können Werte zwischen 1 und 19 erreicht werden, wobei Werte zwischen 8 und 12 als durchschnittlich betrachtet werden.

Für die Auswertung werden zunächst die Rohwerte für jeden Untertest berechnet. Wird das Kriterium bei einem Item erfüllt, wird 1 Punkt eingetragen. Gilt das Kriterium als nicht erfüllt, notiert man 0 Punkte. Detaillierte Instruktionen zur Punktevergabe für jedes einzelne Item der fünf Untertests befinden sich im Durchführungsmanual der Bayley-III Skalen und dienen zur Orientierung für den Testleiter. Die Rohwerte ergeben sich aus dem Zusammenzählen aller Items, für die ein Punkt vergeben wurde, sowie aus dem Hinzuaddieren aller Items, die vor dem Startpunkt lagen. Die Startpunkte für die verschiedenen Altersgruppen sind im Protokollbogen abzulesen und mit den Buchstaben A-Q markiert.

Um nun die verschiedenen Untertest- und Skalenwerte miteinander vergleichen zu können, ermittelt man anhand von nach Altersstufen unterteilten Tabellen im Manual und den errechneten Rohwerten die standardisierten **Untertestwerte** ($M = 10, SD = 3$) und **Skalenwerte** ($M = 100, SD = 15$). Neben einem numerischen Vergleich ist mit den ermittelten Skalen- und Untertestwerten auch das Erstellen von **Profildigrammen** möglich, mit denen die Testergebnisse grafisch veranschaulicht werden.

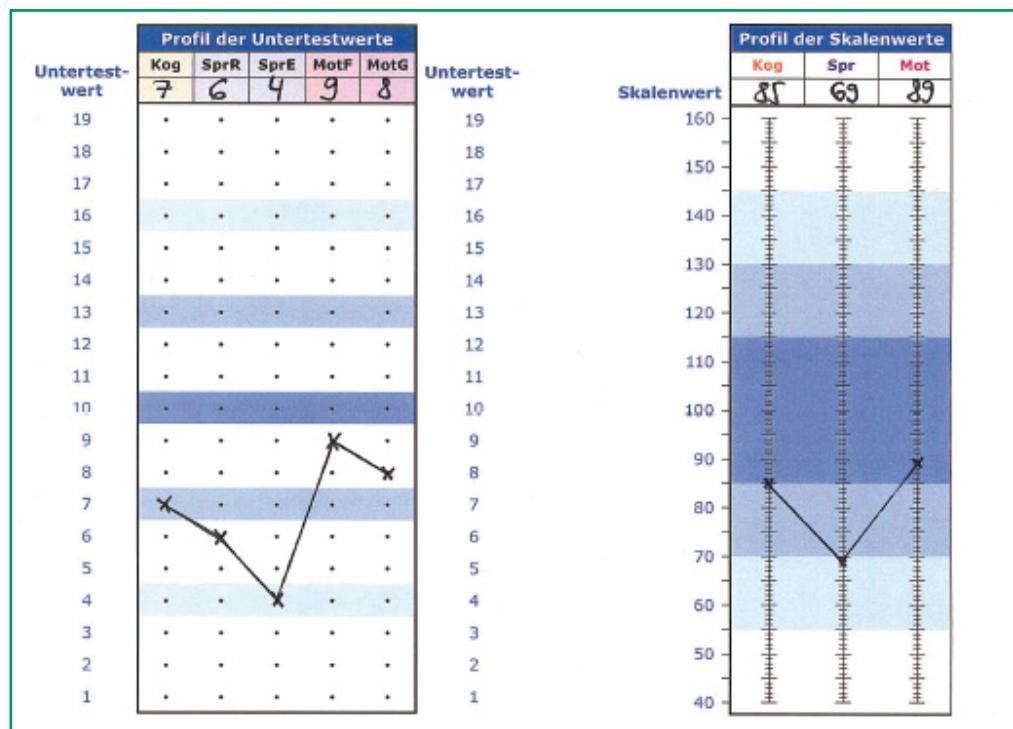


Abbildung 2.10: Erstellen des Profildigramms

Weiterhin lassen sich anhand von Tabellen der jeweilige *Prozentrang* und das *Konfidenzintervall* ermitteln. Zusätzlich besteht die Möglichkeit zu *Diskrepanzvergleichen*, die angeben, ob der Unterschied zwischen zwei Untertestwerten statistisch signifikant ist. Bei Langzeitbeobachtungen und mehrfacher Testung empfiehlt sich zudem die Berechnung von *Wachstumswerten*, die das relative Wachstum für die einzelnen Untertests – sprich den Leistungszuwachs über die Zeit – angeben.

Im Elternbericht werden sowohl die wichtigsten Ergebnisse zu den Skalen und Untertests eingetragen und zusammengefasst als auch Beobachtungen während der Untersuchung vermerkt und Vorschläge zur Förderung / Intervention kommuniziert.

II. Anwendungsgrenzen

So präzise die Bayley-III Skalen das Entwicklungsniveau von enorm kleinen Kindern auch erfassen können, so erlaubt die Anwendung des Tests keine vorschnellen Diagnosen bei auffälligen Ergebnissen in den verschiedenen Unterbereichen. Die Ergebnisse der Bayley-III sollten stets in einem ganzheitlichen Zusammenhang mit anderen Patientinformationen wie etwa Verhaltensbeobachtungen, weiteren Untersuchungen oder Anamnesebögen betrachtet werden. Weiterhin wird dringend empfohlen, bei Auffälligkeiten in den Testergebnissen Rücksprache mit Experten wie Psychologen, Ergotherapeuten oder Logopäden zu halten.

Die Ergebnisse der Bayley-III Skalen erlauben zwar die Ableitung von Empfehlungen für Fördermaßnahmen, jedoch stellen sie keine Einschätzung zur Art der Verzögerung bereit. Weiterhin sollten bei Kindern mit schwerwiegenden physischen oder sensorischen Beeinträchtigungen nicht die normorientierten Werte der Standardtestversion zur Beurteilung des Funktionsniveaus herangezogen werden. Im Anhang des Durchführungsmanuals befinden sich detaillierte Informationen zur Testhandhabung bei dieser Patientengruppe.

Vorsichtig sollten Anwender der Bayley-III auch bei der Interpretation von Veränderungen im Falle von Testwiederholungen sein. Veränderte Werte sind keineswegs mit einem Behandlungserfolg oder -misserfolg gleichzusetzen. Auch gilt es zwischen den Skalen- und Rohwerten zu differenzieren, denn beide Angaben enthalten unterschiedliche Informationen in Bezug auf das Entwicklungsniveau. Während die Rohwerte als Bezugspunkt für individuelle Entwicklungen bei wiederholten Messungen dienen, setzen die Skalenwerte die Ergebnisse eines Patienten in Bezug zu seiner Altersgruppe.

Die Bayley-III Skalen dienen der Erfassung des Entwicklungsniveaus von Säuglingen und Kleinkindern. Aus den entsprechenden Diagnosen dürfen keine Rückschlüsse auf die Intelligenz der getesteten Kinder gezogen werden. Dieses Konstrukt wird durch die Bayley-III ausdrücklich nicht erfasst.

Der Entwicklungstest kann auch bei Kindern angewendet werden, die außerhalb des vorgegebenen Altersbereichs liegen, sollten diese enorme Entwicklungsstörungen aufweisen. In einem solchen Fall muss auf die Altersäquivalente zurückgegriffen werden. Hierbei gilt es zu beachten, dass diese nicht über verschiedene Untersuchungsbereiche hinweg miteinander vergleichbar sind. Da die Altersäquivalente zudem nicht intervallskaliert sind, können bereits minimale Veränderungen bei den Rohwerten zu größeren Verschiebungen bei den Altersäquivalenten führen.

12. Zusammenfassung

Mit der deutschsprachigen Revision der Bayley-Skalen von Reuner und Rosenkranz und dem dazugehörigen Screening Test liegt nun ein messgenaues, differenziertes und praktikables Messinstrument zur Erfassung des frühkindlichen Entwicklungsstands im kognitiven, sprachlichen und motorischen Leistungsbereich vor.

Die wesentlichen Revisionsziele konnten mit der Bayley-III umgesetzt werden: Es liegt nun erstmals eine deutsche Normierung vor, die psychometrische Güte des Testverfahrens konnte weiter verbessert werden und durch die besondere Eignung für klinische Gruppen wie Frühgeborene, Kinder mit angeborenen schweren Herzfehlern oder chronischen Erkrankungen konnte der klinische Nutzen der Bayley-III Skalen gesteigert werden.

Durch die Einführung altersbezogener Startpunkte sowie festgelegter Einstiegs-, Umkehr- und Ausstiegsregeln ist die Durchführung des Individualtests leichter erlernbar und die Aktualisierung der Items sowie des Testmaterials erhöht die Teilnahmemotivation der Kinder. Auch kann die Reihenfolge der Untertests individuell an die Bedürfnisse des Kleinkindes angepasst werden. Die differenzierte Skalenstruktur ermöglicht eine genauere Unterscheidung der kognitiven, sprachlich-rezeptiven/sprachlich-expressiven sowie der fein- und grobmotorischen Leistungsbereiche, was präzise Diagnosen ermöglicht. Die Einführung des Konzepts „Spielverhalten“ auf der kognitiven Skala gewährleistet weitere wichtige diagnostische Informationen für die frühkindliche Entwicklung.

Der besonders große Altersbereich zwischen einem und 42 Monaten ermöglicht ein breites Anwendungsspektrum der Bayley-III Skalen und eignet sich vorzüglich für die Durchführung von Langzeitbeobachtungen. Mit dem neu herausgegebenen Entwicklungstest lassen sich nicht nur fundierte Diagnosen zum Entwicklungsstand von Säuglingen und Kleinkindern stellen, sondern auch fördernde Maßnahmen und Interventionsvorschläge ableiten.

Die zusätzliche Bereitstellung des Bayley-III Screening Tests bietet eine zeitökonomische Alternative zur Langversion, mit der eine grobe Klassifizierung des Entwicklungsniveaus möglich ist. Auf Basis dieser Ergebnisse kann dann entschieden werden, ob eine umfassende Untersuchung notwendig ist.

Die deutsche Bayley-III ist ein solide konstruierter Breitband-Entwicklungstest, der bei geschulter Anwendung für viele entwicklungsdiagnostische Fragestellungen herangezogen werden kann. Aufgrund ihrer internationalen Verbreitung wird auch die Vergleichbarkeit in Forschungsarbeiten zur frühkindlichen Entwicklungsdiagnostik gewährleistet.

Die dritte Auflage der Bayley Scales of Infant and Toddler Development in deutscher Übersetzung entspricht dem aktuellen Expertenstand aus Forschung und Praxis und leistet somit einen wichtigen Beitrag zur Weiterentwicklung der diagnostischen Möglichkeiten im frühkindlichen Entwicklungsbereich.

13. Literatur

Bayley, N. (1933). The California first-year mental Scale.
Berkeley: University of California Press.

Bayley, N. (1936). The California infant scale of motor development.
Berkeley: University of California Press.

Bayley, N. (1969). Manual for the Bayley scales of infant development.
San Antonio: Psychological Corporation.

Bayley, N. (1993). Bayley Scales of Infant Development (2nd ed.).
San Antonio: Psychological Corporation.

Bayley, N. (2006a). Bayley Scales of Infant and Toddler Development (3rd ed.).
San Antonio: Psychological Corporation.

Bayley, N. (2006b). Bayley Scales of Infant and Toddler Development.
Administration Manual (3rd ed.). San Antonio: Psychological Corporation.

Bayley, N. (2006c). Bayley Scales of Infant and Toddler Development.
Technical Manual (3rd ed.). San Antonio: Psychological Corporation.

Bayley, N. (2014). Bayley-Scales of Infant and Toddler Development, Third Edition –
Deutsche Version. Hrsg. von Reuner, G. & Rosenkranz, J. Frankfurt/M.: Pearson Assessment.

Petermann, F. & Macha, T. (2005). Entwicklungsdiagnostik.
Kindheit und Entwicklung, 14, 131-139. Göttingen: Hogrefe.



Pearson Clinical and Talent Assessment
Baseler Straße 35-37
60329 Frankfurt am Main

Telefon: +49 69 756146-0
Fax: +49 69 756146-10
E-Mail: info.de@pearson.com
www.pearsonassessment.de
Twitter: [@PearsonAssess](https://twitter.com/PearsonAssess)