

The logo for NEPS (Nationales Bildungspanel) features the acronym 'NEPS' in a bold, blue, sans-serif font. To the left of the text is a stylized orange bracket shape that partially encloses the letters.

NEPS

Nationales Bildungspanel

Informationen zu den direkten Maßen und zur Kompetenztestung

NEPS Startkohorte 1 — Neugeborene
Bildung von Anfang an

4. Welle: 37-39 Monate

The logo for LifBi (Leibniz-Institut für Bildungsverläufe e.V.) consists of the letters 'LifBi' in a bold, black, sans-serif font. A vertical blue bar is positioned to the left of the 'i', and a vertical pink bar is positioned to the left of the 'B'.

LifBi

**LEIBNIZ-INSTITUT FÜR
BILDUNGSVERLÄUFE e.V.**

Urheberrechtlich geschütztes Material
Leibniz-Institut für Bildungsverläufe e.V. (LifBi)
Wilhelmsplatz 3, 96047 Bamberg
Direktorin: Prof. Dr. Cordula Artelt
Wissenschaftlich-kordinierende Geschäftsführerin: Dr. Jutta von Maurice
Kaufmännischer Geschäftsführer: Dr. Robert Polgar
Bamberg; 5. Februar 2020

Informationen zur Testung				
Testsituation	Einzeltestung von dreijährigen Kindern im Haushalt der Familie, Ankerperson und Interviewerin anwesend.			
Ablauf der Testung	<p>Die vier Kompetenztests wurden in folgender Reihenfolge administriert:</p> <ol style="list-style-type: none"> 1. Wortschatz: rezeptives Hörverständnis auf Wortebene (Lenhard, A., Lenhard, W., Segerer, R., Suggate, S. (2015). <i>Peabody Picture Vocabulary Test-Revision IV German Adaption</i>, PPVT-IV. Frankfurt, Germany: Pearson) 2. Zahlenspanne: Phonologisches Arbeitsgedächtnis 3. Kategorisierung (nonverbal): SON-R Subtest (Tellegen, P. J., Winkel, M., Wijnberg-Williams, B. J., & Laros, J. A. (2007). <i>Snijders-Oomen Non-verbaler Intelligenztest (SON-R2 1/2-7; German version)</i>. Göttingen, Germany: Hogrefe) 4. Belohnungsaufschub: Exekutive Kontrolle <p>Die Response-Eingabe der Kompetenztests erfolgte entweder durch das Kind selbst (Wortschatz und SON-R Subtest Kategorisierung) oder durch die Interviewerinnen (Zahlenspanne und Belohnungsaufschub). Die Instruktionen wurden ebenso wie die Items vom Tablet vorgegeben.</p>			
Testdauer (ohne Aufbau)	ca. 35 Minuten			
Informationen zu den einzelnen Tests				
Konstrukt	Anzahl der Items	Dauer (ca.)	Erhebungsmodus	Nächste Messung(en)
Wortschatz: rezeptives Hörverständnis auf Wortebene	max. 19 Sets mit jeweils 12 Aufgaben (mit Abbruchkriterium)	15 Minuten	Je Aufgabe (= je vorgegebenem Wort): 4 Auswahlbilder; auf Tablet-PC administriert	Welle 6 (2017) Welle 8 (2019)
Zahlenspanne: Phonologisches Arbeitsgedächtnis	max. 15 Aufgaben (mit Abbruchkriterium)	5 Minuten	Wiedergabe mündlich; auf Tablet-PC administriert	Welle 7 (2018)
Kategorisierung: SON-R Subtest	max. 2 Sets mit insgesamt 15 Aufgaben (mit Abbruchkriterium)	10 Minuten	Items bildbasiert (nonverbal); Vorgabe: auf Tablet-PC administriert	-
Belohnungsaufschub: Exekutive Kontrolle	-	5 Minuten	materialbasiert; auf Tablet-PC administriert	Welle 6 (2017) Welle 8 (2019)

Vorbemerkung

Der Entwicklung der einzelnen Tests liegen Rahmenkonzeptionen zugrunde. Dabei handelt es sich um übergeordnete Konzeptionen, auf deren Basis bildungsrelevante Kompetenzen größtenteils über den gesamten Lebenslauf in konsistenter und kohärenter Weise abgebildet werden sollen. Die Rahmenkonzeptionen, auf deren Grundlage die Testinstrumente entwickelt wurden, sind deshalb in verschiedenen Studien identisch.

Zusätzlich zu Kompetenzmaßen, die kohärent über den Lebenslauf erfasst werden, werden etappenspezifische Maße zu bestimmten Zeitpunkten im Lebensverlauf erhoben, zu denen sie besonders aussagefähig sind (vgl. Berendes, Weinert, Zimmermann & Artelt, 2013¹). In der Regel erfolgt hierbei keine Messwiederholung.

¹ Berendes, K., Weinert, S., Zimmermann, S., & Artelt, C. (2013). Assessing language indicators across the lifespan within the German National Educational Panel Study (NEPS). *Journal for Educational Research Online/Journal für Bildungsforschung Online*, 5(2), 15–49.

Wortschatz: rezeptives Hörverständnis auf Wortebene

Hörverstehen auf Wort-, Satz- und Text-/Diskursebene als Indikatoren der Sprachkompetenz im Deutschen

Die Bedeutung sprachlicher Kompetenzen für schulisches Lernen sowie für die Erklärung sozialer Disparitäten in den Schulkarrieren ist weitgehend unbestritten.

Die sprachlichen Kompetenzen im Deutschen werden in NEPS einerseits über das Hörverstehen auf Wort-, Satz- und Text-/Diskursebene sowie andererseits – ab der 2. Grundschulklasse – über Indikatoren der Lesefähigkeiten (Lesekompetenz [Textverständnis], Lesegeschwindigkeit) erfasst. In Startkohorte 1 des NEPS erfolgt ab dem Alter von 3 Jahren ausschließlich das Erfassen des Hörverstehens auf Wortebene und später - ab der 2. Grundschulklasse (Welle 9) – die Erhebung von Indikatoren der Lesefähigkeit.

Hörverstehen auf Wortebene (rezeptiver Wortschatz)

Maße des rezeptiven Wortschatzes stellen einen guten, international anschlussfähigen Indikator für die erworbenen sprachlichen Fähigkeiten und Fertigkeiten von Kindern und Erwachsenen dar. In zahlreichen großen internationalen Panelstudien wie zum Beispiel dem Head Start Family and Child Experiences Survey – FACES (USA)², dem National Longitudinal Survey of Children and Youth – NLCSY (Kanada; u.a. Lipps & Yiptong-Avila, 1999), der British Cohort Study – BCS70 (z.B. Bynner, 2004) oder der European Child Care and Education (ECCE)-Study, die in Deutschland, Österreich, Spanien und Portugal durchgeführt wurde (z.B. European Child Care and Education (ECCE)-Study Group, 1997), wird der passive Wortschatz als zentraler, manchmal sogar als alleiniger Indikator der vor dem Hintergrund individueller Grundfähigkeiten (z.B. Arbeitsgedächtniskapazität, Geschwindigkeitsvariablen) und Umweltanregung kumulativ erworbenen sprachlich-kognitiven Fähigkeiten erhoben.

Das international zur Erfassung des rezeptiven Wortschatzes am meisten eingesetzte Instrument ist der inzwischen in verschiedenen Versionen vorliegende Peabody Picture Vocabulary Test (PPVT; Dunn, 1959; Dunn & Dunn, 1981, 1997, 2007). Der PPVT ist im Grundsatz über einen sehr großen Altersbereich (von 2,5 Jahren bis ins hohe Erwachsenenalter) hinweg einsetzbar und zugleich einfach in der Durchführung und Auswertung.

In Startkohorte 1 wird eine deutschsprachige Version des PPVT-IV (Dunn & Dunn, 2007; deutsche Bearbeitung von Lenhard, Lenhard, Segerer & Suggate, 2015) eingesetzt.

Die Administrierung erfolgte im NEPS über einen Tablet-PC. Aufgabe der Kinder war es, zu jedem einzeln vorgegebenen Wort, welches auditiv vom Tablet vorgegeben wird, aus jeweils vier Bildern das zu dem Wort gehörige Bild auszuwählen.

Nach den Vorgaben des PPVT-IV variiert der Schwierigkeitsgrad (Testeinstieg und -abbruch) in Abhängigkeit vom Alter und der Leistung der Kinder. Der Test beginnt mit einer Übungseinheit, die je nach Alter und Leistung des Kindes gesteuert wird. Hat das Kind in der Übungsphase mindestens 2 Aufgaben richtig gelöst, beginnt die Testphase. Insgesamt besteht der Test aus 19 schwierigkeitsgestaffelten Sets mit jeweils 12 Items.

Testablauf in dieser Welle: Dem Alter der Kinder entsprechend beginnt der Test mit einer Übungsphase, die aus mind. 2 und max. 4 Aufgaben besteht. In der Testphase beginnen alle Kinder mit

² <http://www.acf.hhs.gov/programs/opre/hs/faces/>

Set 1 (Bodenset). Der Test wird so lange durchgeführt, bis das Deckenset, also das höchste bewältigte Schwierigkeitsniveau, identifiziert wird. Das Deckenset ist dann erreicht, wenn in einem Set mehr als 7 Fehler gemacht wurden.

Im Scientific Use File finden sich die Anzahl administrierter Übungsaufgaben, Korrektheit der Antworten für jedes Testitem (richtig, falsch), das Bodenset und das Deckenset. Zudem wird der Summenwert angegeben, der die Anzahl der richtig gelösten Items angibt. Dabei werden alle Items, die sich in niedrigeren Sets als das Bodenset befinden, als richtig gelöst angenommen.

Literatur

- Bynner, J. (2004). Participation and progression: Use of British Cohort Study data in illuminating the role of basic skills and other factors (Nuffield Review of 14–19 Education and Training Working Paper 9). Adelaide, Australia: National Centre for Vocational Education Research.
- Dunn, L. M. (1959). *Peabody Picture Vocabulary Test (PPVT): Manual of directions and forms*. Nashville, TN: American Guidance Service.
- Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test-Revised (PPVT-R)*. Circle Pines, MN: American Guidance Service.
- Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test, Third Edition (PPVT-III)*. Circle Pines, MN: American Guidance Service.
- Dunn, L. M., & Dunn, L. M. (2007). *Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4)*. Upper Saddle River, NJ: Pearson.
- European Child Care and Education (ECCE)-Study Group. (1997). *European Child Care and Education*. Berlin, Germany: Freie Universität Berlin, Fachbereich Erziehungswissenschaft, .
- Lenhard, A., Lenhard, W., Segerer, R., & Suggate, S. (2015). *Peabody picture vocabulary test-4*. Frankfurt, Germany: Pearson.
- Lipps, G., & Yiptong-Avila, J. (1999). From home to school: how Canadian children cope. *Education Quarterly Review*, 6(2), 51–57.

Zahlenspanne – phonologisches Arbeitsgedächtnis

Das Kurzzeit- oder Arbeitsgedächtnis gilt als das kapazitätsbegrenzte Nadelöhr der Informationsverarbeitung. Auf der einen Seite können Menschen eine nahezu unbegrenzte Anzahl von Informationen langfristig speichern; auf der anderen Seite erweist sich die Fähigkeit, unverbundene Informationen (z. B. eine Telefonnummer) nach einmaligem Hören unmittelbar wiederzugeben als begrenzt. Entsprechende Leistungen des Kurzzeit- bzw. Arbeitsgedächtnisses (funktionale Kapazität) erweisen sich als interindividuell unterschiedlich und nehmen während der Kindheit bis ins Jugendalter zu (für einen Überblick, vgl. Weinert, 2010).

Im Nationalen Bildungspanel wird das Konstrukt „Zahlenspanne“, basierend auf dem theoretischen Rahmenmodell des Arbeitsgedächtnisses, z. B. von Baddeley und Hitch (1974), erhoben. Als Indikatoren für die Kapazität des phonologischen Arbeitsgedächtnisses gelten die Leistungen in sogenannten Spannenaufgaben (Baddeley, 1992). In Spannenaufgaben werden z. B. Folgen von Zahlen (bzw. Ziffern) auditiv präsentiert, die in gleicher Reihenfolge reproduziert werden sollen (vgl. „digit

span“). Die Spannenaufgaben werden dabei in ansteigender Länge, d. h. zunehmender Anzahl von Ziffern, präsentiert, bis die fehlerfreie Reproduktion nicht mehr gelingt. Dies stellt die maximale Länge der Sequenz von Ziffern dar, die von einem Individuum nach einmaligem Hören unmittelbar korrekt reproduziert werden kann (Baddeley, Gathercole & Papagno, 1998). Die kurzfristige Speicherung und unmittelbare Wiedergabe der auditiven Informationen beanspruchen die sogenannte phonologische Schleife, welche ein passives Subsystem im Arbeitsgedächtnis-Modell darstellt (Baddeley & Hitch, 1974). Aufgrund einer schnellen Präsentationsrate wird die Nutzung (bzw. Nutzungsunterschiede) bezogen auf Merkstrategien minimiert, weshalb die individuelle Leistung als Indikator für die jeweilige Kapazität des phonologischen Arbeitsgedächtnisses interpretieren werden kann. Die Reproduktion der Ziffernfolgen wird allerdings nicht nur von der strukturellen Kapazität der phonologischen Schleife, sondern auch von Geschwindigkeitsfaktoren (Artikulationsgeschwindigkeit, Geschwindigkeit der Itemidentifikation) beeinflusst, die wiederum mit Vorwissensaspekten, wie z. B. sprachlichem Wissen, zusammenhängen.

Die individuell unterschiedliche Kapazität des phonologischen Kurzzeit- bzw. Arbeitsgedächtnisses erweist sich in zahlreichen Studien als bedeutsam für die Entwicklung sprachlicher Fertigkeiten, insbesondere den Wortschatzerwerb (vgl. für Überblicke z. B. Baddeley, Gathercole, & Papagno, 1998; Gathercole & Baddeley, 1993; Weinert, 2010) und den Erwerb von Lesekompetenz (z. B. Berendes, Weinert, Zimmermann, & Artelt, 2013; Gathercole & Baddeley, 1993). Darüber hinaus weist u. a. eine Untersuchung von Krajewski und Schneider (2009) auch auf einen Zusammenhang zwischen der Leistungsfähigkeit der phonologischen Schleife vor der Einschulung und der mathematischen Entwicklung in der Schule hin.

In Startkohorte 1 des NEPS erfolgt die Durchführung der Zahlenspanne angelehnt an die deutsche Version der „Kaufman Assessment Battery for Children“ (K-ABC; Melchers & Preuß, 2009). Geprüft wird die Fähigkeit eine verbal vorgegebene Ziffernreihe in gehörter Reihenfolge unmittelbar wiederzugeben (Zahlenfolgegedächtnis). Verwendet werden Ziffern zwischen 1 und 10, wobei auf die Ziffer 7 aufgrund der Mehrsilbigkeit verzichtet wurde (vgl. Melchers & Preuß, 2009). Die Steuerung der Aufgabe und die auditive Vorgabe der Zahlenspannen erfolgen standardisiert und spielerisch von einem Tablet-PC; die Administration erfolgte auf Deutsch. Aufgabe der Kinder ist es, die jeweiligen Ziffernfolgen unmittelbar nach ihrer Vorgabe in derselben Reihenfolge wiederzugeben.

Der Test besteht aus einer Übungsphase und anschließenden Lern- und Testitems. Die Übungsphase umfasst 1 Item, das bei falscher oder fehlender Antwort wiederholt wird, um sicherzustellen, dass die Instruktion verstanden wurde; sie geht nicht in den Summenscore ein. Darauf folgen fünf Aufgabensets, welche aus jeweils 3 Items bestehen. Die ersten beiden Items des ersten Sets stellen noch Lernitems in dem Sinne dar, dass die Kinder vom Tablet-PC Feedback zur Richtigkeit ihrer Antwort erhalten und das Item gegebenenfalls nochmals vorgegeben wird. Die Lernitems gehen nur dann in den Summenscore ein, wenn sie bereits bei der ersten Vorgabe richtig reproduziert wurden. Die Items enthalten zu Beginn zwei Ziffern und die Zifferanzahl nimmt pro Aufgabenset um eine Ziffer zu, d. h. Set 5 enthält Items mit jeweils sechs Ziffern. Für den Abbruch des Tests sind die Übungsitems nicht entscheidend; bei den Testitems wird der Test bis einschließlich Aufgabenset 3 nur abgebrochen, wenn alle Items innerhalb eines Sets falsch beantwortet wurden. Danach wird das Spiel bereits nach einer falschen Antwort abgebrochen. Die maximal erreichbare Punktezahl liegt bei 15 Punkten (13 Testitems und 2 Lernitems), wobei jede korrekt reproduzierte Folge mit einem Punkt bewertet wird.

Im Scientific Use File³ ist Folgendes veröffentlicht: die Anzahl der insgesamt administrierten Übungsitems; die Korrektheit der Lösung jedes einzelnen Lern- und Testitems; der Summenwert, der aus der Anzahl aller richtig gelösten Lern- und Testitems besteht; die maximale Ziffernspanne, bei der mind. ein Item richtig beantwortet wurde; sowie eine Variable, welche anzeigt, bei welcher Ziffernspanne der Test abgebrochen wurde. Bei Lernitems wurde sowohl für die Bewertung der Korrektheit, als auch für den Summenscore nur der erste Versuch des Kindes gewertet.

Literatur

- Baddeley, A. D. (1992). Working memory. *Science*, 255(5044), 556–559. <https://doi.org/10.1126/science.1736359>
- Baddeley, A., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation*, 8, 47–89. [https://doi.org/10.1016/S0079-7421\(08\)60452-1](https://doi.org/10.1016/S0079-7421(08)60452-1)
- Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, 105(1), 158–173. <https://doi.org/10.1037/0033-295X.105.1.158>
- Berendes, K., Weinert, S., Zimmermann, S., & Artelt, C. (2013). Assessing language indicators across the lifespan within the German National Educational Panel Study (NEPS). *Journal for Educational Research Online*, 5(2), 15–49.
- Gathercole, S. E., & Baddeley, A. D. (1993). Phonological working memory: A critical building block for reading development and vocabulary acquisition? *European Journal of Psychology of Education*, 8(3), 259–272. <https://doi.org/10.1007/BF03174081>
- Krajewski, K., & Schneider, W. (2009). Exploring the impact of phonological awareness, visual-spatial working memory, and preschool quantity-number competencies on mathematics achievement in elementary school: Findings from a 3-year longitudinal study. *Journal of experimental child psychology*, 103(4), 516–531. <https://doi.org/10.1016/j.jecp.2009.03.009>
- Melchers, P., & Preuß, U. (2009). *Kaufman Assessment Battery for Children (K-ABC), dt. Version (8., unveränderte Aufl.)*. Frankfurt, Germany: Pearson Assessment.
- Weinert, S. (2010). Beziehungen zwischen Sprachentwicklung und Gedächtnisentwicklung. In H.-P. Trollenier, W. Lenhard & P. Marx (Hrsg.), *Brennpunkte der Gedächtnisforschung: Entwicklungs- und pädagogisch-psychologische Perspektiven* (147–170). Göttingen, Germany: Hogrefe.

Kategorisierung: SON-R Subtest

Der Snijders-Oomen non-verbale Intelligenztest (SON-R 2½-7; Tellegen, Winkel, Wijnberg-Williams, & Laros, 2007) ist ein standardisierter und normierter Individualtest zur Erfassung der nonverbalen Intelligenz von Kindern im Alter zwischen 2 ½ und 7 Jahren, der aus verschiedenen Untertests besteht. In der Neugeborenenkohorte des Nationalen Bildungspanels (Startkohorte 1) wurde nur der Untertest „Kategorien“ des SON-R 2½-7 durchgeführt, mit dem Leistungen des abstrakten Denkens, bzw. schlussfolgernden Denkens („reasoning“), erfasst werden (einen Überblick bietet u. a. Quinn, 2011). In

³ Hinweis: Die hier als veröffentlicht beschriebenen Daten beziehen sich auf Version SC1:7.0.0.

Kategorisierungsaufgaben geht es v. a. darum, Ordnungsprinzipien abzuleiten oder Verbindungen zwischen (abstrakten) Begriffen oder Objekten herzustellen. Die Fähigkeit zur Abstraktion, d. h. zum Erkennen von Zusammenhängen und zum Ziehen von Schlussfolgerungen, ist dabei in nahezu allen Konzepten von Intelligenz zentral vertreten (u. a. Cattell, 1963; Spearman, 1923; Wechsler, 1950). Der Untertest „Kategorien“ misst eine Facette entsprechender Fähigkeiten und erwies sich auch deshalb als geeignet für NEPS, da er vergleichsweise einfach auf eine Administrierung per Tablet-Computer angepasst werden konnte. Er kann als ein Kurzindikator nicht-sprachlicher kognitiver Grundfähigkeiten im frühen Kindesalter betrachtet werden.

Der Untertest „Kategorien“ besteht aus insgesamt 15 Testaufgaben, die sich in zwei Teile gliedern, sowie einer zusätzlichen Übungsaufgabe, über die das Verständnis für die Testaufgaben im zweiten Testteil hergestellt wird. Im ersten Teil (Aufgaben 1 bis 7) mussten die Kinder jeweils vier bis sechs Bilder anhand bestimmter Merkmale sortieren. So mussten z. B. Blumen und Bonbons zueinander sortiert werden. Bei den Aufgaben des zweiten Teils (Aufgaben 8 bis 15) wurden dem Kind jeweils drei Bilder von Objekten gezeigt, die ein gemeinsames Merkmal aufweisen, z. B. wurden drei verschiedene Fruchtarten gezeigt. Aus einer Reihe von fünf weiteren Bildern mussten diejenigen zwei gewählt werden, die das gleiche Merkmal haben wie die ersten drei Bilder.

Die spielerisch gestaltete Durchführung in Startkohorte 1 des NEPS erfolgte über ein Tablet mit Touchscreen. Wie im Testmanual des SON-R 2½-7 vorgesehen, wurden die Aufgaben als Lerntest umgesetzt, d. h. nach jeder Aufgabe bekamen die Kinder Feedback über die Richtigkeit ihrer Antwort (Tellegen et al., 2007). Nach 3 Fehlern wurde der Untertest automatisch abgebrochen.

Im Scientific Use File ist die Aufgabenbewertung für jedes Einzelitem (nicht gelöst/gelöst) der 15 Testitems sowie für das Übungsitem veröffentlicht. Basierend auf den Testitems wurde zudem ein Schätzwert zur Personenfähigkeit (Weighted Likelihood Estimator, WLE) und dessen Standardfehler berechnet, um die Fähigkeit der Einzelperson in Relation zur Gesamtgruppe zu setzen (Hojtink & Boomsma, 1995; Warm, 1989). In das Schätzmodell gingen keine weiteren Hintergrundvariablen ein.

Literatur

- Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, 54(1), 1–22.
- Hojtink, H., & Boomsma, A. (1995). On Person Parameter Estimation in the Dichotomous Rasch Model. In G. H. Fischer, & I. W. Molenaar, *Rasch Models* (53–68). New York, NY: Springer-Verlag.
- Quinn, P. C. (2011). Born to categorize. In U. Goswami, *The Wiley-Blackwell Handbook of Childhood Cognitive Development* (129–153). Chichester, England: Wiley-Blackwell.
- Spearman, C. (1923). *The nature of "intelligence" and the principles of cognition*. Oxford, England: Macmillan.
- Tellegen, P. J., Winkel, M., Wijnberg-Williams, B. J., & Laros, J. A. (2007). *Snijders-Oomen Non-verbaler Intelligenztest (SON-R2 1/2-7; German version)*. Göttingen, Germany: Hogrefe.
- Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, 54(3), 427–450.

Wechsler, D. (1950). Cognitive, conative, and non-intellective intelligence. *American Psychologist*, 5(3), 78–83.

Belohnungsaufschub: Exekutive Kontrolle

„Selbstregulation ist die Fähigkeit, das eigene Denken, Fühlen und Handeln zu kontrollieren und zu steuern“ (Neubauer, Gawrilow & Hasselhorn, 2011, S. 203), um persönlich gesteckte Ziele zu planen, zu verfolgen und folglich zu erreichen (Zimmermann, 2000). Fähigkeiten der Selbstregulation umfassen verschiedene, durchaus unterschiedliche Facetten, die über höchst verschiedene Maße (Selbst-/Fremdurteil, direkte standardisierte Beobachtungen, experimentelle Aufgaben etc.) gemessen werden. Im NEPS werden unterschiedliche Aspekte und Facetten erfasst – einerseits Aspekte der kognitiven Selbstregulation im Sinne prozeduraler und deklarativer Metakognition (Weinert et al., 2019) sowie verschiedene Befragungssitems bezogen auf emotionale und verhaltensbezogene Selbstkontrolle.

In der Startkohorte 1 (Säuglingskohorte) und Startkohorte 2 (Kindergartenkohorte) des NEPS werden zudem Aufgaben zum Belohnungsaufschub eingesetzt (für das Vorgehen in Startkohorte 2 vgl. Luplow, Schönmoser, Lorenz & Schmitt, 2019). Zur Erfassung des Belohnungsaufschubs kann insbesondere zwischen zwei Verfahren unterschieden werden: dem Warte- und dem Wahlparadigma (Mischel, 1974, 2015). Beim Warte-Paradigma, welches in Startkohorte 1 im Alter von 3 sowie im Alter von 5 Jahren Verwendung findet, wird die Fähigkeit untersucht, eine sofortige kleine Belohnung zugunsten einer größeren zeitlich verzögerten Belohnung aufzuschieben (Mischel & Gilligan, 1964). Damit stellt das Paradigma eine Reaktionsinhibitions-Aufgabe (Garon, Bryson & Smith, 2008) dar, die konzeptuell den exekutiven Funktionen zugeordnet werden kann und für die Steuerung und Planung von Handlungen und Verhalten bedeutsam sind (Neubauer et al., 2011). In Startkohorte 1 werden exekutive Funktionen zudem im Alter von 4 Jahren über eine Flankeraufgabe gemessen.

Die Fähigkeit zur Selbstregulation, zu kognitiven Abwägungsprozessen bezüglich potentieller Verhaltensweisen (Mischel, 1974) gilt als bedeutsam für die kindliche Entwicklung. Eine Reihe von Studien deutet darauf hin (u.a. Watts, Duncan, & Quan, 2018), dass sich gut entwickelte Fähigkeiten der Selbstregulation im Vorschulalter als prädiktiv für spätere akademische Leistungen, den Umgang mit Stress (Stressresistenz) sowie für die Entwicklung sozio-emotionaler Kompetenzen und der Konzentrationsfähigkeit erweisen (Baumeister & Vohs, 2004; Tangney, Baumeister & Boone, 2004; Wulfert, Block, Ana, Rodriguez & Colman, 2002; Kochanska, Murray & Coy, 1997). Darüber hinaus gibt es Hinweise, dass die Wahlentscheidung für ein sofortiges kleines Geschenk oder ein größeres Geschenk zu einem späteren Zeitpunkt mit dem Alter der Kinder, aber auch mit sozialen Hintergrundmerkmalen und anderen Kindmaßen in Zusammenhang steht (Lemmon & Moore, 2007; Thompson, Barresi & Moore, 1997; Watts et al., 2018). So steigt die „Tendenz zum Warten [...] mit zunehmenden Alter der Kinder signifikant an“ (Lorenz, Schmitt, Luplow & Schönmoser, 2016, S. 71).

In dieser Erhebung waren die „Situations- und Wartebedingungen“ (Mischel, 1974), wie im Folgenden beschrieben, standardisiert. Dem Kind, sitzend an einem Tisch, werden zwei Geschenke präsentiert, ein großes und ein kleines verpacktes Geschenk. Zwischen den Geschenken wird eine USB-Taste platziert. Die Aufgabe besteht darin, entweder eine bestimmte, für das Kind unbekannte, Zeitspanne abzuwarten, um das große Geschenk zu erhalten oder um die Wartezeit abubrechen. Letzteres war durch Betätigung der USB-Taste ("Button") möglich. Dadurch erhält das Kind sofort das kleine Geschenk. Die einmalige Wahlentscheidung, welche vom Kind getroffen wird, bestimmt somit die

Geschenkauswahl und die Wartezeit. In dieser Erhebung konnte eine maximale Wartezeit von 3 Minuten (181 Sekunden) erreicht werden; nach dieser Wartezeit erhielt das Kind das größere Geschenk. Im Scientific Use File sind die Wartezeit und die Auswahl (großes, kleines Geschenk) des Kindes als Variablen enthalten.

Literatur

- Baumeister, R. F., & Vohs, K. D. (2004). *Handbook of self-regulation: Research, theory, and applications*. New York, NY: The Guilford Press.
- Garon, N., Bryson S. E., Smith I. M. (2008). Executive Function in Preschoolers: A Review Using an Integrative Framework. *Psychological Bulletin*, 134 (1), 31–60.
- Kochanska, G., Murray, K. T., & Coy, K. C. (1997). Inhibitory control as a contributor to conscience in childhood: From toddler to early school age. *Child Development*, 68(2), 263–277.
- Lemmon, K., & Moore, C. (2007). The development of prudence in the face of varying future rewards. *Developmental Science*, 10(4), 502–511.
- Lorenz, C., Schmitt, M., Luplow, N., & Schönmoser, C. (2016). Soziale Disparitäten im Vorschulalter und der Einfluss der Selbstregulation. *Zeitschrift für Grundschulforschung*, 9(1), 65–77.
- Luplow, N., Schönmoser, C., Lorenz, C., & Schmitt, M. (2019). *Die Messung des Belohnungsaufschubes in der Startkohorte 2 des Nationalen Bildungspanels (NEPS) im Kindergarten und der Grundschule (NEPS Survey Paper No. 54)*. Bamberg, Germany: Leibniz-Institut für Bildungsverläufe, Nationales Bildungspanel.
- Mischel, W. (1974). Processes in delay of gratification. In L. Berkowitz, *Advances in Experimental Social Psychology* (Bd. 7, 249–292). New York, NY: Academic Press.
- Mischel, W. (2015). *Der Marshmallow Test*. München, Germany: Siedler Verlag.
- Mischel, W., & Gilligan, C. (1964). Delay of gratification, motivation for the prohibited gratification, and responses to temptation. *The Journal of Abnormal and Social Psychology*, 69(4), 411–417.
- Neubauer, A., Gawrilow, C., & Hasselhorn, M. (2011). Belohnungsaufschub: Ein Ansatz zur Frühprognose volitionaler Kompetenzen. In M. Hasselhorn, & W. Schneider, *Frühprognose schulischer Kompetenzen* (202–220). Göttingen, Germany: Hogrefe.
- Tangney, J. P., Baumeister, R. F., & Boone, A. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality*, 72(2), 271–324.
- Thompson, C., Barresi, J., & Moore, C. (1997). The development of future-oriented prudence and altruism in preschoolers. *Cognitive Development*, 12(2), 199–212.
- Watts, T. W., Duncan, G. J., & Quan, H. (2018). Revisiting the Marshmallow Test: A Conceptual Replication Investigating Links Between Early Delay of Gratification and Later Outcomes. *Psychological Science*, 29(7), 1159–1177.
- Weinert, S., Artelt, C., Prenzel, M., Senkbeil, M., Ehmke, T., Carstensen, C. H., & Lockl, K. (2019). Development of Competencies Across the Life Course. In H.-P. Blossfeld, & H.-G. Roßbach,

Education as a Lifelong Process: The German National Educational Panel Study (NEPS) (57–82). Wiesbaden, Germany: Springer VS.

Wulfert, E., Block, J. A., Ana, E. S., Rodriguez, M. L., & Colman, M. (2002). Delay of Gratification: Impulsive Choices and Problem Behaviors in Early and Late Adolescence. *Journal of Personality, 70*(4), 533–552.

Zimmermann, B. J. (2000). Attaining Self-Regulation: A Social Cognitive Perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner, *Handbook of Self-Regulation* (13–39). San Diego, CA: Academic Press.