

NEPS *SURVEY PAPERS*

Autorenteam Kompetenzsäule  
**LÄNGSSCHNITTliche  
KOMPETENZMESSUNG  
IM NEPS: ANLAGE UND  
DESKRIPTIVE BEFUNDE**

NEPS *Survey Paper* No. 80  
Bamberg, Dezember 2020

**Survey Papers of the German National Educational Panel Study (NEPS)**

at the Leibniz Institute for Educational Trajectories (LifBi) at the University of Bamberg

The NEPS *Survey Paper* series provides articles with a focus on methodological aspects and data handling issues related to the German National Educational Panel Study (NEPS).

They are of particular relevance for the analysis of NEPS data as they describe data editing and data collection procedures as well as instruments or tests used in the NEPS survey. Papers that appear in this series fall into the category of 'grey literature' and may also appear elsewhere.

The NEPS *Survey Papers* are edited by a review board consisting of the scientific management of LifBi and NEPS.

The NEPS *Survey Papers* are available at [www.neps-data.de](http://www.neps-data.de) (see section "Publications") and at [www.lifbi.de/publications](http://www.lifbi.de/publications).

**Editor-in-Chief:** Thomas Bäumer, LifBi

**Review Board:** Board of Directors, Heads of LifBi Departments, and Scientific Management of NEPS Working Units

**Contact:** German National Educational Panel Study (NEPS) – Leibniz Institute for Educational Trajectories – Wilhelmsplatz 3 – 96047 Bamberg – Germany – [contact@lifbi.de](mailto:contact@lifbi.de)

# Längsschnittliche Kompetenzmessung im NEPS: Anlage und deskriptive Befunde

*Autorenteam Kompetenzsäule*

**E-Mail-Adresse des Erstautors/der Erstautorin:**

kathrin.lockl@lifbi.de

**Bibliographische Angaben bei deutschsprachigen Papers:**

Autorenteam Kompetenzsäule (2020). *Längsschnittliche Kompetenzmessung im NEPS: Anlage und deskriptive Befunde* (NEPS Survey Paper No. 80). Leibniz-Institut für Bildungsverläufe, Nationales Bildungspanel. <https://doi.org/10.5157/NEPS:SP80:1.0>

**Autorenteam Kompetenzsäule:**

Kathrin Lockl (*LifBi*), Cordula Artelt (*LifBi*), Claus Carstensen (*Universität Bamberg*), Tabea Durda (*LifBi*), Karin Gehrler (*LifBi*), Inga Hahn (*IPN Kiel*), Jan Marten Ihme (*IPN Kiel*), Jana Kähler (*IPN Kiel*), Anna-Lena Kock (geb. Gerken) (*IPN Kiel*), Olaf Köller (*IPN Kiel*), Kristin Litteck (*IPN Kiel*), Shally Novita (*LifBi*), Lara Aylin Petersen (*IPN Kiel*), Martin Senkbeil (*IPN Kiel*), Ilka Wolter (*LifBi*)

## Inhaltsverzeichnis

1. Einleitung.....	4
2. Mathematische Kompetenz .....	7
3. Lesekompetenz .....	22
4. Naturwissenschaftliche Kompetenz .....	35
5. Hörverstehen: Rezeptiver Wortschatz .....	50
6. ICT Literacy .....	58
7. Metakognition .....	73
8. Lesegeschwindigkeit.....	85
9. Kognitive Grundfähigkeiten .....	93

# Längsschnittliche Kompetenzmessung im NEPS: Anlage und deskriptive Befunde

## Zusammenfassung

In dem vorliegenden Beitrag wird ein Überblick über die Anlage und deskriptive längsschnittliche Befunde zu den Kompetenzmessungen im Rahmen des Nationalen Bildungspanels (NEPS) für den Zeitraum von 2010 bis 2020 gegeben. Basierend auf der ausführlichen Beschreibung der Rahmenkonzeption der Kompetenzmessungen im NEPS (Artelt, Weinert & Carstensen, 2013; Weinert et al., 2019) gewährt der Beitrag zunächst einen kurzen Einblick in die Anlage der längsschnittlich konzipierten Kompetenzmessungen. Im Anschluss werden jeweils für die einzelnen Kompetenzbereiche, die sich domänen-spezifischen Kompetenzen, Metakompetenzen sowie ergänzenden Konstrukten zuordnen lassen, Informationen zu den bisherigen Messungen gegeben und zentrale deskriptive Befunde dargestellt. Hierbei werden sowohl psychometrische Kennwerte als auch Mittelwertsunterschiede für zentrale Gruppierungsvariablen, wie z.B. sozio-ökonomischer Status oder Schultyp, berichtet.

## Abstract

This paper gives an overview of the design and descriptive longitudinal findings of the competence measurements within the National Educational Panel Study (NEPS) for the period from 2010 until 2020. Based on the detailed description of the framework of the competence measurements in the NEPS (Artelt, Weinert & Carstensen, 2013; Weinert et al., 2019), the paper first provides a brief insight into the design of the longitudinal competence measurements. Subsequently, for each of the individual competence areas which can be assigned to domain-specific competences, meta-competences, and supplementary constructs, information on the previous measurements is given and central descriptive findings are presented. Both psychometric parameters and mean differences are reported for central grouping variables such as socio-economic status or school type.

## Schlagworte

Kompetenzmessung, Längsschnitt, Nationales Bildungspanel (NEPS), domänenspezifische Kompetenzen

## 1. Einleitung

Kathrin Lockl

Ein zentrales Ziel des NEPS (Blossfeld, Roßbach & von Maurice, 2011) ist die Erfassung von Kompetenzen, die als bedeutsam für Bildungsverläufe und gesellschaftlichen Erfolg betrachtet werden (vgl. Artelt, Weinert & Carstensen, 2013; Weinert et al., 2019). Der Schwerpunkt der Instrumentenentwicklung zur Erfassung der Kompetenzentwicklung im Lebenslauf liegt darin, Strukturmodelle zu entwickeln, auf deren Basis bildungsrelevante Kompetenzen sowie deren Ausprägung und Erwerb über den gesamten Lebenslauf in konsistenter und kohärenter Weise abgebildet werden können. So können unterschiedliche Verläufe des Bildungserwerbs nachgezeichnet und in ihrer Abhängigkeit von bildungsrelevanten Einflussgrößen und ihrer Bedeutung für Bildungswege und berufliche Karrieren analysiert werden. Im Zentrum steht die Erfassung und Analyse der Entwicklung von bildungsrelevanten Kompetenzen, die zum Teil auch in den meist synchron (als Querschnitt) angelegten internationalen Schulleistungstudien gemessen werden.

Im Rahmen des NEPS werden längsschnittliche, über den Lebenslauf kohärente Messungen von Lesekompetenz, mathematischer und naturwissenschaftlicher Kompetenz vorgenommen. Ergänzt werden diese durch regelmäßige Messungen von Aspekten der Metakognition bzw. Selbstregulation sowie der Fähigkeit im Umgang mit Informations- und Kommunikationstechnologien (Information and Communication Technologies, ICT Literacy). Die Kompetenztests werden auf der Grundlage von theoretischen Rahmenkonzeptionen entwickelt, die es erlauben, die Kompetenzen kohärent über die Lebensspanne zu erfassen und die somit die Möglichkeit eröffnen, längsschnittliche Verläufe nachzuzeichnen. Indikatoren allgemeiner kognitiver Leistungsfähigkeit sowie sozialer Kompetenzen (z. B. zur Konfliktlösung und zur Teamfähigkeit in unterschiedlichen Gruppen) werden ebenfalls in die Kompetenzmessungen im NEPS einbezogen. Weiterhin werden etappenspezifische Kompetenzen erfasst, die stärker die Anforderungen berücksichtigen, die in bestimmten Lebensphasen gestellt werden (für detailliertere Informationen zur Auswahl der Kompetenzbereiche im NEPS vgl. Artelt, Weinert & Carstensen, 2013; Weinert et al., 2019).

Um Veränderungen im Erwerb der Kompetenzen modellieren zu können, sind wiederholte Messungen der einzelnen Konstrukte notwendig. Das Erhebungsprogramm zu den Kompetenzmessungen ist dabei so konzipiert, dass systematische intra- und interindividuelle Vergleiche über das Alter und teilweise über Kohorten hinweg möglich werden. Zu jedem Messzeitpunkt wird eine festgelegte Auswahl von Kompetenzbereichen getestet. Zudem werden feste Zeitintervalle zwischen den Erhebungswellen vorgegeben, insbesondere für die Domänen, die kohärent über die Lebensspanne gemessen werden. Diese Intervalle sind in den höheren Startkohorten, aufgrund der angenommenen geringeren Veränderungsrate, größer als in den jüngeren Kohorten, die sich noch in der formalen Ausbildung befinden. Neben den Konstrukten, die in mehr oder weniger engmaschigen Abständen erfasst werden, gibt es ergänzende Konstrukte, die als wichtige Kovariaten einbezogen und selbst nicht längsschnittlich modelliert werden (z. B. kognitive Grundfähigkeiten, Lesegeschwindigkeit). Für diese Konstrukte gilt, dass sie weniger häufig, aber zumindest einmal zu Beginn der Erhebungen in einer Startkohorte, erfasst werden sollen.

Die Verknüpfung der Daten über verschiedene Messzeitpunkte hinweg („Linking“), mit deren Hilfe die Daten auf einer gemeinsamen Skala abgebildet werden können, erfolgt auf der Basis zweier unterschiedlicher Strategien: Zum einen werden Ankeritems in Tests für unterschiedliche Altersgruppen bzw. Messzeitpunkte verwendet. Auf diese Weise werden manche der Testinstrumente in der Domäne Mathematik miteinander verlinkt. Zum anderen werden zusätzliche Linking-Studien (z. B. für die Domäne Lesen) durchgeführt, mit deren Hilfe die Testinstrumente für zwei Altersgruppen auf indirektem Weg verlinkt werden können (für eine ausführliche Beschreibung der Linking-Strategien im NEPS vgl. Fischer, Rohm, Gnamb & Carstensen, 2016). In den Scientific Use Files sind mittlerweile für alle Startkohorten verlinkte Daten verfügbar, die mehrere benachbarte Messzeitpunkte zueinander in Beziehung setzen und längsschnittliche Veränderungen auf Individualebene direkt abbildbar machen. Beispielsweise können in der Startkohorte 3 längsschnittliche Verläufe in Lesen und Mathematik über vier Messzeitpunkte hinweg (Klasse 5, 7, 9 und 12 inkl. Individualfeld) auf diese Art und Weise nachgezeichnet werden.

In diesem Bericht liegt der Fokus auf domänenspezifischen Kompetenzen und Metakompetenzen, für die die Verantwortung der Testentwicklung im Bereich der Kompetenzsäule des NEPS liegt. Zusätzlich werden ergänzende Konstrukte beschrieben, für die teilweise auf publizierte Verfahren zurückgegriffen werden kann. Im Einzelnen wird in diesem Bericht auf die folgenden Konstrukte eingegangen:

Domänenspezifische Kompetenzen:

- Mathematische Kompetenz
- Lesekompetenz
- Naturwissenschaftliche Kompetenz
- Hörverstehen (hier: rezeptiver Wortschatz)

Metakompetenzen:

- ICT-Literacy
- Metakognition

Ergänzende Konstrukte:

- Lesegeschwindigkeit
- Kognitive Grundfähigkeiten

Im Folgenden werden für die hier genannten Kompetenzbereiche Informationen zur Erfassung der Konstrukte sowie zentrale deskriptive Befunde aus den bisherigen Messungen dargestellt. Dabei werden für jeden Kompetenzbereich nach einer kurzen theoretischen Einführung Informationen zu den Zeitpunkten der Erfassung in den jeweiligen Startkohorten, zu Erhebungsdauer und Administrationsform gegeben sowie Beispielitems präsentiert. Daran schließen sich deskriptive Ergebnisse und psychometrische Kennwerte zu den einzelnen Erhebungen an. Für manche der Konstrukte werden zudem Korrelationen mit anderen Konstrukten im Sinne konvergenter bzw. diskriminanter Validität berichtet. Weiterhin wird für

jeden der Kompetenzbereiche ein Überblick gegeben, inwiefern sich Mittelwertsunterschiede für die zentralen Gruppierungsvariablen Geschlecht, sozio-ökonomischer Status, Migrationshintergrund, Schultyp und Bildungsabschluss finden lassen. Die beschriebenen Auswertungen beziehen sich auf den Stand der Scientific Use Files zu Beginn des Jahres 2020. In die Auswertungen gingen die bis zu diesem Zeitpunkt vorliegenden Daten ein; hierbei wurden keine Imputationen vorgenommen. Schließlich werden für jeden der Kompetenzbereiche ausgewählte Publikationen aufgelistet, die unter Einbeziehung von Kompetenzdaten der jeweiligen Domäne entstanden sind.

Es sei an dieser Stelle darauf hingewiesen, dass der Schwerpunkt dieses Berichts auf einer Darstellung deskriptiver Befunde zu den einzelnen Konstrukten liegt und nicht auf einer theoretischen Einbettung der zu messenden Kompetenzbereiche. Für weiterführende Details zur Auswahl und Konzeption der Kompetenzmessungen im NEPS vor dem Hintergrund bestehender theoretischer Modelle und vorliegender Large-Scale Assessments sei auf die Beiträge von Artelt, Weinert und Carstensen (2013) sowie Weinert et al. (2019) verwiesen.

Ein Gesamtüberblick, welche Kompetenzen in welcher Startkohorte zu welchem Zeitpunkt erfasst wurden, ist auf der Website des NEPS zu finden ([Fuß, Gnambs, Lockl & Attig, 2019](#)).

## Literatur

- Artelt, C., Weinert, S. & Carstensen, C. H. (2013). Assessing competencies across the life span within the German National Educational Panel Study (NEPS)—Editorial. *Journal of Educational Research Online (JERO)*, 5, 5-14.
- Blossfeld, H.-P., Roßbach, H.-G. & von Maurice, J. (Eds.). (2011). Education as a lifelong process: The German National Educational Panel Study (NEPS). *Zeitschrift für Erziehungswissenschaft* [Special Issue 14]. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Fischer, L., Rohm, T., Gnambs, T. & Carstensen, C. H. (2016). *Linking the data of the competence tests* (NEPS Survey Paper No. 1). Bamberg: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Fuß, D., Gnambs, T., Lockl, K. & Attig, M. (2019). *Competence data in NEPS: Overview of measures and variable naming conventions (Starting Cohorts 1 to 6)*. Bamberg: Leibniz Institute for Educational Trajectories, National Educational Panel Study. [https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/Kompetenzen/Overview\\_NEPS\\_Competence-Data.pdf](https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/Kompetenzen/Overview_NEPS_Competence-Data.pdf)
- 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 (Eds.), *Education as a lifelong process: The German National Educational Panel Study (NEPS)*, Edition ZfE (2. überarbeitete Auflage, S. 57-82). Wiesbaden: Springer. <https://doi.org/10.1007/978-3-658-23162-0>

## 2. Mathematische Kompetenz

Lara Aylin Petersen, Anna-Lena Kock (geb. Gerken) & Kristin Litteck (IPN Kiel)

### 2.1 Einleitung

Die mathematische Kompetenz zählt im NEPS zweifellos zu einer der zentralen Kompetenzen (Artelt et al., 2013; Ehmke et al., 2009; Weinert et al., 2019), der sowohl in der Bildungspraxis als auch in der Bildungsforschung eine hohe Bedeutung beigemessen wird. So gilt das Schulfach Mathematik traditionell als Hauptfach, dem innerhalb der Kultusministerkonferenz eine hohe Priorität im Bildungssystem eingeräumt wird (Sekretariat der ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 02.06.2006). Mathematischer Kompetenz wird dabei ein uniquer Beitrag zur Allgemeinbildung zugesprochen, da sie eine spezifische Modellierung realer Problemsituationen ermöglicht und dadurch einer Bearbeitung zugänglich macht (Winter, 1995). Diese aus normativer Perspektive begründete Relevanz mathematischer Kompetenz lässt sich auch empirisch stützen. So wurde etwa in bildungsökonomischen Analysen aufgezeigt, dass der mathematische Bildungsstand von Staaten substantiellen Einfluss auf deren langfristige wirtschaftliche Entwicklung hat (Hanushek & Wößmann, 2015).

Aufgrund ihrer großen Bedeutung wird die mathematische Kompetenz auch im nationalen und internationalen Bildungsmonitoring als eine der zentralen Kompetenzen adressiert (Mullis, Martin, Foy & Arora, 2012; OECD, 2014). Ein Großteil dieser Large Scale Assessment-Studien beschränkt sich dabei auf querschnittliche Untersuchungsdesigns, sodass bisher wenig über die kumulative Entwicklung mathematischer Kompetenz über die Lebensspanne und damit verbundener Bedingungsfaktoren bekannt ist (Blossfeld, Roßbach & von Maurice, 2011). Das Nationale Bildungspanel kann hier somit Antworten auf Fragestellungen nach dem Erfolg von Bildungsverläufen für den zentralen Bereich der mathematischen Kompetenz ermöglichen.

In Deutschland gibt es bereits einige Studien, welche die längsschnittliche Entwicklung von mathematischer Kompetenz analysieren. So wurde beispielsweise in der Studie „Kompetenzen und Einstellungen von Schülerinnen und Schülern“ (KESS) die Entwicklung der Kompetenz von Schülerinnen und Schülern in dem Stadtstaat Hamburg in den Jahrgängen 4, 7, 8, 10/11 und 13 längsschnittlich untersucht. Des Weiteren testete die PISA-I-Plus-Studie eine Teilstichprobe der PISA 2003 Studie erneut ein Jahr später am Ende der zehnten Klassenstufe (Deutsches PISA-Konsortium, 2006). Dieselben Klassenstufen untersuchte die PISA-Plus-Studie 2012, indem sie eine Teilstichprobe 2013 erneut in der zehnten Klassenstufe testeten (Lehner et al., 2017). Diese und weitere längsschnittliche Studien, welche die Entwicklung der mathematischen Kompetenz in Deutschland analysieren, decken oft nur eine begrenzte Lebensspanne (zumeist Schulalter) und/oder eine selektive Stichprobe ab. Das NEPS bietet hingegen erstmals die Möglichkeit die mathematische Kompetenzentwicklung vom Kindergarten bis zum Einstieg in das Berufsleben bzw. darüber hinaus innerhalb einer Kohorte zu untersuchen und dabei vielfältige Bedingungsfaktoren einzubeziehen.

## 2.2 Bisherige Erhebung des Konstrukts

Die mathematischen Kompetenztests des NEPS wurden bisher überwiegend in papierbasierter Form administriert. Neuere Erhebungen in den Startkohorten 4-6 wurden computerbasiert entwickelt. Die hier eingesetzten Multi-Stage-Tests, also das teilweise adaptive Vorgehen mit unterschiedlichen Schwierigkeitsstufen in Abhängigkeit von der gezeigten Leistung im Test, ermöglichen eine differenziertere Erfassung mathematischer Kompetenz. Zudem wird durch den computerbasierten Erhebungsmodus eine noch höhere Standardisierung ermöglicht, da die Interviewerinnen und Interviewer einen immer kleineren Teil der Kompetenztestung selbst mitgestalten (z. B. Durchführung der Instruktion). Die technische Umsetzung der computerbasierten Testungen eröffnet außerdem die Möglichkeit, Online-Testungen durchzuführen, was einen flexibleren und auf lange Sicht ökonomischeren Einsatz ermöglicht. In der Startkohorte 1 wurden bereits von Beginn an tabletbasierte Tests eingesetzt. Für Erhebungen ab einem Alter von vier Jahren wurde hierdurch die Entwicklung neuer Aufgabenformate ermöglicht. Eine altersgerechte und interaktive Darstellung der Aufgaben und der ergänzende Einsatz von Materialien ließen einen spielerischen Zugang zu den Aufgaben zu. In Zukunft soll die Entwicklung von tablet- bzw. computerbasierten Tests weiter ausgebaut werden, um die Möglichkeiten von technikbasierten Erhebungen in vollem Maße ausschöpfen zu können.

Die folgende Tabelle beschreibt sowohl die Erhebungszeiten als auch -modi der bisher im NEPS in der Domäne Mathematik gelaufenen bzw. aktuell laufenden Erhebungen. Die entweder gerade aktuell laufenden oder aber in ihrer Auswertung noch nicht abgeschlossenen Erhebungen sind zur Verdeutlichung in grauer Schrift gehalten.

Tabelle 1: Erhebungszeiten und -modi der in der Domäne Mathematik gelaufenen bzw. aktuell laufenden Erhebungen

<b>Startkohorte</b>	<b>Alter / Klasse</b>	<b>Jahr</b>	<b>Erhebungszeit (in Minuten)</b>	<b>Erhebungsmodus</b>
SC 1	4 Jahre	2016	25	Computer (Tablets)
SC 1	6 Jahre	2018	20	Computer (Tablets)
SC 1	8 Jahre	2020	20	Computer (Online)
SC 2	6 Jahre	2012	30	Papier
SC 2	Klasse 1	2013	30	Papier
SC 2	Klasse 2	2013/14	28	Papier
SC 2	Klasse 4	2015/16	28	Papier
SC 2	Klasse 7	2018/19	28	Papier
SC 3	Klasse 5	2010/11	28	Papier
SC 3	Klasse 7	2012/13	28	Papier

Startkohorte	Alter / Klasse	Jahr	Erhebungszeit (in Minuten)	Erhebungsmodus
SC 3	Klasse 9	2014/15	28	Papier
SC 3	Klasse 12	2017/18	28	Papier
SC 4	Klasse 9	2010/11	28	Papier
SC 4	Klasse 12 / Schulabgänger	2013/14	28	Papier
SC 4	21-Jährige	2016/17	28	Computer/ Online
SC 5	Studierende (Welle 1)	2011	28	Papier
SC 5	Studierende (Welle 12)	2017	28	Computer/ Online
SC 6	Erwachsene (Welle 3)	2010/11	28	Papier
SC 6	Erwachsene (Welle 9)	2016/17	28	Computer/ Online

### 2.3 Beispielitems

Tabelle 2 ordnet den folgenden Beispielitems sowohl die Inhaltsbereiche als auch die Fähigkeitsbereiche zu. Eine detailliertere Beschreibung der Beispielitems befindet sich in Schnittjer und Duchhardt (2015).

Tabelle 2: Inhaltsbereiche und Fähigkeitsbereiche der Beispielitems

Beispielitem	Altersstufe	Inhaltsbereich	Fähigkeiten
<b>1</b> Steine im Becher	Kindergarten	Quantität	Mathematische Probleme lösen Technische Fertigkeiten
<b>2</b> Der Zaun	Klasse 5-7	Raum und Form	Technische Fertigkeiten
<b>3</b> Im Zoo	Klasse 9	Veränderung und Beziehung	Repräsentieren (Darstellungen verwenden) Mathematisches Kommunizieren
<b>4</b> Nebenwirkung	Studierende	Daten und Zufall	Mathematisches Argumentieren Mathematisches Kommunizieren Technische Fertigkeiten

### Steine im Becher

In dieser Schüssel sind vier Steinchen. Jetzt lege ich noch drei Steinchen dazu.

*[Die Schüssel ist abgedeckt, so dass das Kind nicht hineinsehen kann.]*

Kannst du mir sagen, wie viele Steinchen **jetzt** in der Schüssel sind?

Abbildung 1. Beispielitem „Steine im Becher“.

### Der Zaun

#### DER ZAUN

Herr Braun besitzt ein rechteckiges Grundstück, das er einzäunen möchte. Nach einigen Berechnungen kauft er 40 m Zaun.

Das Grundstück hat eine Breite von 8 m.

#### Aufgabe 1:

Wie lang ist das Grundstück?

*Bitte kreuze die richtige Antwort an! Bitte kreuze nur ein Kästchen an!*

<input type="checkbox"/>	5 m
<input type="checkbox"/>	8 m
<input type="checkbox"/>	12 m
<input type="checkbox"/>	16 m

Abbildung 2. Beispielitem „Der Zaun“.

## Im Zoo

### IM ZOO

Im Sommer hat der Zoo von Astadt normalerweise mehr Besucher als im Winter. Aber in diesem Jahr wurde im Oktober ein Braunbär geboren, der mit seiner Niedlichkeit sehr viele Besucher anzog.

#### Aufgabe 1:

Welches der folgenden Diagramme stellt die Anzahl der Besucher im Zoo von Astadt korrekt dar?

*Bitte kreuze die richtige Antwort an! Bitte kreuze nur ein Kästchen an!*

<input type="checkbox"/>	<p>Anzahl Besucher</p> <p>Jan. — Dez.</p>
<input type="checkbox"/>	<p>Anzahl Besucher</p> <p>Jan. — Dez.</p>
<input type="checkbox"/>	<p>Anzahl Besucher</p> <p>Jan. — Dez.</p>
<input type="checkbox"/>	<p>Anzahl Besucher</p> <p>Jan. — Dez.</p>

Abbildung 3. Beispielitem „Im Zoo“.

**Nebenwirkung****NEBENWIRKUNG**

Ein Pharmazie-Unternehmen hat ein neues Mittel gegen Kopfschmerzen entwickelt. In einer Studie wurden zwei häufige Nebenwirkungen entdeckt: Juckreiz und Übelkeit. In der folgenden Tabelle ist die Anzahl der Versuchspersonen angegeben.

		Übelkeit	
		Ja	Nein
Juckreiz	Ja	50	70
	Nein	40	100

**Aufgabe 1:**

Sind die folgenden Aussagen über das Ergebnis der Studie korrekt?

*Bitte kreuzen Sie in jeder Zeile die richtige Antwort an!*

	Ja	Nein
Bei der Hälfte der teilnehmenden Personen trat zumindest eine Nebenwirkung auf, denn 50 ist die Hälfte von 100.	<input type="checkbox"/>	<input type="checkbox"/>
Übelkeit trat seltener auf als Juckreiz, denn $50+40$ ist kleiner als $50+70$ .	<input type="checkbox"/>	<input type="checkbox"/>
Bei etwa 53% der teilnehmenden Personen wurde mindestens eine Nebenwirkung beobachtet, denn $(50+40+70)/3 \approx 53\%$ .	<input type="checkbox"/>	<input type="checkbox"/>
Bei mehr als der Hälfte der teilnehmenden Personen, bei denen Übelkeit auftrat, wurde auch Juckreiz festgestellt, denn $50:90 > 50\%$ .	<input type="checkbox"/>	<input type="checkbox"/>

Abbildung 4. Beispielitem „Nebenwirkung“.

## 2.4 Empirische Analysen

### 2.4.1 Deskriptive Informationen

In Tabelle 3 sind die deskriptiven Angaben für die Datensätze der entsprechenden Erhebungen dargestellt. Auch hier sind entweder noch laufende Erhebungen oder aber noch nicht als Scientific Use Files (SUF) veröffentlichte Studien in grauer Schrift gehalten. Auffallend ist die geringe Ausfallquote mit weniger als 0.50% in allen Erhebungen mit zwei Ausnahmen in der SC1, da bei jüngeren Testpersonen häufiger Aufgaben unbearbeitet blieben. Durchschnittlich liegen die Reliabilitäten bei 0.78 und sind somit als durchgängig gut zu bewerten.

Tabelle 3: Deskriptive Informationen zu Erhebungen in der Domäne Mathematik

<b>Startkohorte</b>	<b>Alter / Klasse</b>	<b>N (gesamt)</b>	<b>N (gültig)</b>	<b>% fehlend</b>	<b>M</b>	<b>SD</b>	<b>Reliabilität</b>
SC 1	4 Jahre	2138	2024	5.33	-0.02	1.04	0.70
SC 1	6 Jahre	1989	1968	1.06	2.56	1.17	0.81
SC 2	6 Jahre	2727	2724	< 1.00	-0.01	1.17	0.82
SC 2	Klasse 1	6511	6489	< 1.00	1.71	1.13	0.77
SC 2	Klasse 2	6168	6167	< 1.00	2.23	1.17	0.81
SC 2	Klasse 4	6724	6718	< 1.00	4.52	1.15	0.75
SC 3	Klasse 5	5208	5193	< 1.00	-0.01	1.17	0.80
SC 3	Klasse 7	6194	6191	< 1.00	0.82	1.23	0.82
SC 3	Klasse 9	4889	4888	< 1.00	1.60	1.19	0.83
SC 3	Klasse 12/ Schulabgänger	3786	3785	< 1.00	2.60	1.11	0.77
SC 4	Klasse 9	14524	14523	< 1.00	0.02	1.22	0.81
SC 4	Klasse 12/ Schulabgänger	5734	5733	< 1.00	1.07	1.12	0.74
SC4	21-Jährige	7078	6909	2.38	1.21	1.20	0.75*
SC 5	Studierende (Welle 1)	5949	5939	< 1.00	-0.01	1.23	0.76
SC 5	Studierende (Welle 12)	4606	4323	6.14	0.14	1.18	0.75*
SC 6	Erwachsene (Welle 3)	5245	5221	< 1.00	-0.02	1.41	0.81
SC 6	Erwachsene (Welle 9)	4409	4373	< 1.00	0.03	1.23	0.75*

Bemerkung. N (gesamt) = Anzahl der Personen, denen der Test administriert wurde; N (gültig) = Anzahl der Personen mit gültigem Testwert, % fehlend = Anteil an Personen ohne gültigen Testwert, M = Mittelwert des Testwerts (bei erster Administration pro Startkohorte auf 0 fixiert), SD = Standardabweichung des Testwerts, Reliabilität = EAP Reliabilität.

\*Die Reliabilität wurde für die Erwachsenen der Startkohorten 4-6 gemeinsam berechnet, da alle denselben Test bearbeitet haben.

### 2.4.2 Psychometrische Informationen

Umfangreiche psychometrische Analysen zeigen, dass die restriktiven Kriterien von Pohl und Carstensen (2012) für Eindimensionalität von den NEPS-Mathematiktests nicht immer erfüllt werden und mehrdimensionale Modelle (nach den vier bzw. fünf Inhaltsgebieten) die Daten meistens besser abbilden. Selten wird das eindimensionale Modell bevorzugt (vgl. technischer Report von Schnittjer & Gerken, 2017, sowie Petersen, Litteck & Rohenroth, 2020). Jedoch werden zwischen den Dimensionen sehr hohe Korrelationen ( $.87 < r < .97$ , Duchhardt & Gerdes, 2012, 2013; Jordan & Duchhardt, 2013) gefunden, was auf einen erheblichen Anteil eines gemeinsamen Faktors mathematischer Kompetenz hinweist. Zudem weist das eindimensionale Rasch-Modell ebenfalls eine gute Passung zu den Daten auf. Dieses Ergebnis ist nicht überraschend und entspricht dem anderer LSA-Studien (z.B. PISA, IQB-Ländervergleich, Winkelmann, Robitzsch, Stanat & Köller, 2012; Brunner, 2006), sodass einer Nutzung eindimensional skalierten Daten nichts entgegensteht. Detaillierte Ergebnisse der Analysen sind in den technischen Reports nachzulesen (siehe Übersicht Tabelle 4).

Die umfassende Validierung der Testwertinterpretationen der NEPS-Mathematiktests war nicht Teil der Instrumentenentwicklung und würde zusätzliche Daten und Analysen erfordern. Zum Teil wurden solche Analysen jedoch extern durchgeführt (z.B. van den Ham, 2016). So konnte für den NEPS-K9-Mathematiktest (Startkohorte 4) erwartungsgemäß eine deutlich höhere Korrelation mit der gleichzeitig in NEPS erhobenen Naturwissenschaftskompetenz gefunden werden als mit der Lesekompetenz ( $r = .82$  zu  $r = .69$ ). Des Weiteren fiel der Zusammenhang der mathematischen Kompetenz mit der Mathematiknote deutlich stärker aus als mit der Deutschnote ( $r = -.28$  zu  $r = -.01$ ). Der Zusammenhang der in NEPS, PISA und dem IQB-Ländervergleich gemessenen mathematischen Kompetenz für die Klassenstufe 9 zeigte sehr hohe Korrelationen:  $r = .89$  für NEPS-PISA sowie  $r = .91$  für NEPS-Ländervergleich. Die Korrelationen vom NEPS-Mathematiktests zum Naturwissenschaftstest aus PISA bzw. des IQB-Ländervergleichs fielen erwartungsgemäß niedriger aus ( $r = .72$ ,  $r = .71$ ). Für den NEPS-K5-Mathematiktest konnten Nissen, Ehmke, Köller und Duchhardt (2015) einen sehr starken Zusammenhang ( $r = .90$ ) zu dem TIMSS-K4-Test nachweisen.

Tabelle 4: Veröffentlichte technische Reports der Domäne Mathematik nach Startkohorte

<b>Start- kohorte</b>	<b>Alter / Klasse</b>	<b>Jahr der Veröffentlichung</b>	<b>Autor/innen</b>
SC 1	4 Jahre	2018	Petersen & Gerken
SC 1	6 Jahre	2020	Kock, Litteck & Petersen
SC 2	6 Jahre	2018	Schnittjer
SC 2	Klasse 1	2018	Schnittjer & Fischer
SC 2	Klasse 2	2018	Schnittjer & Gerken
SC 2	Klasse 4	2020	Schnittjer, Gerken & Petersen
SC 3	Klasse 5	2012	Durchhardt & Gerdes
SC 3	Klasse 7	2017	Schnittjer & Gerken
SC 3	Klasse 9	2018	Van den Ham, Schnittjer & Gerken
SC 3	Klasse 12	2020	Petersen, Litteck & Rothenroth
SC 4	Klasse 9	2013	Durchhardt & Gerdes
SC 4	Klasse 12 / Schulabgänger	2017	Fischer, Rohm & Gnams
SC 4	21-Jährige	2020	Gnams
SC 5	Studierende (Welle 1)	2017	Gerken & Schnittjer
SC 5	Studierende (Welle 12)	2020	Gnams
SC 6	Erwachsene (Welle 3)	2013	Jordan & Durchhardt
SC 6	Erwachsene (Welle 9)	2020	Gnams

### 2.4.3 Gruppenunterschiede

Tabelle 5 fasst die standardisierten Mittelwertsunterschiede (Cohen's  $d$ ) der Testwerte für zentrale Gruppierungsvariablen zusammen. Auffällig ist hier, dass sich der Gruppenunterschied beim Geschlecht in der ersten Erhebung der SC 1 (4 Jahre) anders verhält als in den restlichen Erhebungen, wobei es als sehr geringer Gruppenunterschied (SC 1, 4 Jahre) interpretiert werden kann. Der Sozialstatus wirkt sich durchschnittlich über alle Startkohorten am zweitstärksten auf die Gruppenunterschiede im Vergleich zu den anderen Variablen aus ( $M_{\text{Sozial}} = -0.60$ ). Den größten durchschnittlichen Einfluss zeigt der Schultyp in der SC 3 und SC 4 ( $M_{\text{Schule}} = 1.03$ ). Erwartungsgemäß haben sich der Sozialstatus, der Migrationshintergrund, der Bildungsabschluss der Eltern und die besuchte Schulform auf die Gruppenunterschiede ausgewirkt. Personen mit hohem Sozialstatus ohne Migrationshintergrund auf einem Gymnasium und mit Abitur zeigen durchschnittlich bessere mathematische Fähigkeiten (*WLE-Schätzer*).

Tabelle 5: Standardisierte Mittelwertsunterschiede (Cohen's d) der Testwerte für zentrale Gruppierungsvariablen (Domäne Mathematik)

Startkohorte	Alter / Klasse	Geschlecht	Sozio-ökon. Status	Migrationshintergrund	Schultyp	Bildungsabschluss
		(weiblich vs. männlich)	(hoch vs. gering)	(ohne vs. mit)	(Gymnasium vs. Nicht-Gymn.)	(Abitur vs. kein Abitur)
SC 1	4 Jahre	0.16	0.47	0.49	-	0.55
SC 1	6 Jahre	-0.15	0.64	0.65	-	0.68
SC 2	6 Jahre	-0.17	0.64	0.60	-	0.53
SC 2	Klasse 1	-0.25	0.52	0.56	-	0.37
SC 2	Klasse 2	-0.37	0.56	0.61	-	0.37
SC 2	Klasse 4	-0.12	0.61	0.69	-	0.34
SC 3	Klasse 5	-0.29	0.65	0.62	1.17	0.68
SC 3	Klasse 7	-0.34	0.72	0.56	1.26	0.75
SC 3	Klasse 9	-0.27	0.73	0.46	1.24	0.74
SC 3	Klasse 12/ Schulabgänger	-0.49	0.59	0.48	0.91	0.59
SC 4	Klasse 9	-0.38	0.58	0.54	1.40	0.80
SC 4	Klasse 12/ Schulabgänger	-0.49	0.48	0.46	1.12	0.59
SC 4	21-Jährige	-0.53	0.47	0.51	0.97	0.59
SC 5	Studierende (Welle 1)	-0.81	0.32	0.33	-	0.26
SC 5	Studierende (Welle 12)	-0.72	0.24	0.23	-	0.30
SC 6	Erwachsene (Welle 3)	-0.82	0.82	0.18	-	1.17
SC 6	Erwachsene (Welle 9)	-0.68	0.66	0.12	-	0.91

Bemerkung. Positive Werte spiegeln einen höheren Mittelwert in der ersten Gruppe (z.B. „weiblich“) wider, während negative Werte einen höheren Mittelwert in der zweiten Gruppe (z.B. „Männer“) widerspiegeln. Sozioökonomischer Status = HISEI der Eltern bzw. eigener ISEI in SC6, Migrationshintergrund = ohne (= beide Elternteile in Deutschland geboren) vs. mit (= mind. ein Elternteil nicht in Deutschland geboren). Bildungsabschluss = Abitur der Eltern, in SC6 eigenes Abitur.

## 2.5 Ausgewählte Publikationen mit NEPS Daten zur mathematischen Kompetenz

- Brandt, N. D., Lechner, C. M., Tetzner, J. & Rammstedt, B. (2020). Personality, cognitive ability, and academic performance: Differential associations across school subjects and school tracks. *Journal of Personality*, 88(2), 249-265. <https://doi.org/10.1111/jopy.12482>
- Duchhardt, C., Jordan, A. K. & Ehmke, T. (2017). Adults' use of mathematics and its influence on mathematical competence. *International Journal of Science and Mathematics Education*, 15(1), 155-174. <https://doi.org/10.1007/s10763-015-9670-1>
- Ehmke, T., van den Ham, A. K., Sälzer, C., Heine, J. & Prenzel, M. (2020). Measuring mathematics competence in international and national large scale assessments: Linking PISA and the national educational panel study in Germany. *Studies in Educational Evaluation*, 65, 100847. <https://doi.org/10.1016/j.stueduc.2020.100847>
- Ehrmann, L. & Wolter, I. (2018). The impact of students' gender-role orientation on competence development in mathematics and reading in secondary school. *Learning and Individual Differences*, 61, 256-264. <https://doi.org/10.1016/j.lindif.2018.01.004>
- Gnambs, T. (2017). Human capital and reemployment success: The role of cognitive abilities and personality. *Journal of Intelligence*, 5, 9. <https://doi.org/10.3390/jintelligence5010009>
- Holtmann, A. C., Menze, L. & Solga, H. (2017). Persistent disadvantages or new opportunities? The role of agency and structural constraints for low-achieving adolescents' school-to-work transitions. *Journal of Youth and Adolescence*, 46(10), 2091-2113. <https://doi.org/10.1007/s10964-017-0719-z>
- Neumann, I., Duchhardt, C., Grüßing, M., Heinze, A., Knopp, E. & Ehmke, T. (2013). Modeling and assessing mathematical competence over the lifespan. *Journal for Educational Research Online*, 5(2), 80–109.
- Niehues, W., Kisbu-Sakarya, Y. & Selcuk, B. (2020). Family cohesion facilitates learning-related behaviors and math competency at the transition to elementary school. *Early Education and Development*, 1-14. <https://doi.org/10.1080/10409289.2020.1739418>
- Strobel, B. (2016). Does family language matter? The role of foreign language use and family social capital in the educational achievement of immigrant students in Germany. *Ethnic and Racial Studies*, 39(14), 2641-2663. <https://doi.org/10.1080/01419870.2016.1145712>
- Viesel-Nordmeyer, N., Ritterfeld, U. & Bos, W. (2020). Welche Entwicklungszusammenhänge zwischen Sprache, Mathematik und Arbeitsgedächtnis modulieren den Einfluss sprachlicher Kompetenzen auf mathematisches Lernen im (Vor-) Schulalter? *Journal für Mathematik-Didaktik*, 44, 125-155. <https://doi.org/10.1007/s13138-020-00165-0>

## Literatur

- Artelt, C., Weinert, S. & Carstensen, C. H. (2013). Assessing competencies across the life span within the German National Educational Panel Study (NEPS)—Editorial. *Journal of Educational Research Online (JERO)*, 5, 5-14.
- Blossfeld, H.-P., Roßbach, H.-G. & von Maurice, J. (Eds.). (2011). Education as a lifelong process: The German National Educational Panel Study (NEPS). *Zeitschrift für Erziehungswissenschaft* [Special Issue 14]. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Brunner, M. (2006). *Mathematische Schülerleistung: Struktur, Schulformunterschiede und Validität* (Dissertation, Humboldt-Universität, Berlin). Zugriff am 04.02.2016 auf [http://library.mpib-berlin.mpg.de/diss/Brunner\\_Dissertation.pdf](http://library.mpib-berlin.mpg.de/diss/Brunner_Dissertation.pdf)
- Deutsches PISA-Konsortium (2006). *PISA 2003: Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahres*. Münster u. a., Waxmann.
- Duchhardt, C. & Gerdes, A. (2012). *NEPS technical report for mathematics – scaling results of starting cohort 3 in fifth grade* (NEPS working paper no. 19). Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Duchhardt, C. & Gerdes, A. (2013). *NEPS technical report for mathematics – scaling results of starting cohort 4 in ninth grade* (NEPS working paper no. 22). Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Ehmke, T., Duchhardt, C., Geiser, H., Grüßing, M., Heinze, A. & Marschick, F. (2009). Kompetenzentwicklung über die Lebensspanne - Erhebung von mathematischer Kompetenz im Nationalen Bildungspanel. In A. Heinze & M. Grüßing (Hrsg.), *Mathematiklernen vom Kindergarten bis zum Studium. Kontinuität und Kohärenz als Herausforderung für den Mathematikunterricht* (S. 313–327). Münster: Waxmann.
- Fischer, L., Rohm, T. & Gnams, T. (2017). *NEPS technical report for mathematics: Scaling results of starting cohort 4 for grade 12* (NEPS Survey Paper No. 12). Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Gerken, A.-L. & Schnittjer, I. (2017): *NEPS technical report for mathematics: Scaling results of starting cohort 5 for first-year students* (NEPS Survey Paper No. 17). Bamberg: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Gnams, T. (2020): *NEPS technical report for mathematics - Scaling results of starting cohorts 4 (wave 10), 5 (wave 12), and 6 (wave 9)* (NEPS survey paper no. 72). Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Hanushek, E. A. & Wößmann, L. (2015). The knowledge capital of nations: Education and the economics of growth. CESifo Book Series. Cambridge (MA): MIT Press.
- Jordan, A.-K. & Duchhardt (2013). *NEPS technical report for mathematics—Scaling results of starting cohort 6–adults* (NEPS working paper no. 32). Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Kock, A.-L., Litteck, K. & Petersen, L. A. (2020): *NEPS technical report for mathematics - Scaling results of starting cohort 1 for six-year-old children* (NEPS survey paper no. 74). Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Lehner, M. C., Heine, J. H., Sälzer, C., Reiss, K., Haag, N. & Heinze, A. (2017). Veränderung der mathematischen Kompetenz von der neunten zur zehnten Klassenstufe. *Zeitschrift für Erziehungswissenschaft*, 20(2), 7-36.

- Mullis, I. V. S., Martin, M. O., Foy, P. & Arora, A. (2012). *TIMSS 2011 international results in mathematics*. Chestnut Hill, MA and Amsterdam: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA).
- Nissen, A., Ehmke, T., Köller, O. & Duchhardt, C. (2015). Comparing apples with oranges? An approach to link TIMSS and the National Educational Panel Study in Germany via equipercentile and IRT methods. *Studies in Educational Evaluation*, 47, 58–67.
- OECD (2014). *PISA 2012 results; What students know and can do - Student performance in mathematics, reading and science (Volume I, revised edition, February 2014), PISA*, OECD Publishing. <http://dx.doi.org/10.1787/9789264201118-en>
- Petersen, L. A. & Gerken, A.-L. (2018): *NEPS technical report for mathematics: Scaling results of starting cohort 1 for four-year old children* (NEPS survey paper no. 45). Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Petersen, L. A., Litteck, K. & Rohenroth, D. (2020). *NEPS technical report for mathematics: Scaling results of starting cohort 3 for grade 12* (NEPS survey paper no. 75). Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Pohl, S. & Carstensen, C. H. (2012). *NEPS technical report - Scaling the data of the competence tests: NEPS working paper no. 14*. Otto-Friedrich-Universität, Nationales Bildungspanel: Bamberg.
- Sekretariat der ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (02.06.2006). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring: (Beschluss der Kultusministerkonferenz vom 02.06.2006)*. Berlin.
- Schnittjer, I. (2018): *NEPS technical report for mathematics: Scaling results of starting cohort 2 in kindergarten* (NEPS survey paper no. 43). Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Schnittjer, I. & Duchhardt, C. (2015). *Mathematical competence: Framework and exemplary test items*. Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Schnittjer, I. & Fischer, L. (2018). *NEPS technical report for mathematics: Scaling results of starting cohort 2 for grade 1* (NEPS Survey Paper No. 46). Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Schnittjer, I. & Gerken, A.-L. (2017): *NEPS technical report for mathematics: Scaling results of starting cohort 3 in grade 7* (NEPS Survey Paper No. 16). Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Schnittjer, I. & Gerken, A.-L. (2018). *NEPS Technical Report for mathematics: Scaling results of Starting Cohort 2 for grade 2* (NEPS Survey Paper No. 47). Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Schnittjer, I., Gerken, A.-L. & Petersen, L. A. (2020): *NEPS technical report for mathematics - Scaling results of starting cohort 2 in fourth grade* (NEPS survey paper no. 69). Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Winkelmann, H., Robitzsch, A., Stanat, P. & Köller, O. (2012). Mathematische Kompetenzen in der Grundschule. *Diagnostica*, 58 (1), 15–30.
- 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 (Eds.), *Education as a lifelong*

process: The German National Educational Panel Study (NEPS), Edition ZfE (2. überarbeitete Aufl., S. 57-82). Wiesbaden: Springer.

Winter, H. (1995). Mathematikunterricht und Allgemeinbildung. *Mitteilungen der Gesellschaft für Didaktik der Mathematik*, 61, 37-46.

Van den Ham, A.-K. (2016). *Ein Validitätsargument für den Mathematiktest der National Educational Panel Study für die neunte Klassenstufe*. Unveröffentlichte Dissertation, Leuphana Universität Lüneburg, Lüneburg.

Van den Ham, A.-K., Schnittjer, I. & Gerken, A.-L. (2018). *NEPS technical report for mathematics: Scaling results of starting cohort 3 for grade 9* (NEPS survey paper no. 38). Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study.

### 3. Lesekompetenz

Karin Gehrler, Ilka Wolter & Shally Novita

#### 3.1 Einleitung

Die Lesekompetenz ist eine zentrale Schlüsselkompetenz für den langfristigen Bildungserfolg. Nicht nur im schriftsprachlichen Bereich ist diese Kompetenz bedeutsam, sondern auch für die Bewältigung der Bildungsanforderungen in anderen Domänen werden zumindest basale Lesefähigkeiten benötigt. Die Fähigkeit, geschriebene Texte zu verstehen und zu nutzen, stellt eine wesentliche Bedingung für die Weiterentwicklung eigenen Wissens und eigener Fähigkeiten dar und ist zugleich Voraussetzung für die Teilhabe am kulturellen und gesellschaftlichen Leben. Vielfältige Lebens- und Wissensbereiche werden über das Lesen eröffnet und erschlossen. Die Bandbreite von Leseanlässen ist sehr groß und das Lesen erfüllt gleichzeitig sehr unterschiedliche Funktionen (vgl. Groeben & Hurrelmann, 2004). Sie reichen von dem für die Weiterbildung und das lebenslange Lernen zentralen Lesen zur Wissenserweiterung bis hin zum literarisch-ästhetischen Lesen. Über Texte werden dabei nicht nur Informationen und Fakten vermittelt, sondern auch Ideen, Wertvorstellungen und kulturelle Inhalte transportiert. Die Konzeption von Lesekompetenz im Nationalen Bildungspanel legt entsprechend ein funktionales Verständnis der Lesekompetenz zugrunde, wie es sich auch im angelsächsischen Literacy-Konzept (s.a. OECD, 2009) widerspiegelt. Im Mittelpunkt steht der kompetente Umgang mit Texten in verschiedenen charakteristischen Alltagssituationen.

Um das Konzept der Lesekompetenz über die Lebensspanne möglichst kohärent abbilden zu können, wurden in der Rahmenkonzeption zum NEPS-Lesekompetenztest (Gehrler, Zimmermann, Artelt & Weinert, 2013) drei Merkmale spezifiziert, welche in den jeweils alters- bzw. etappenspezifischen Testformen berücksichtigt werden:

1. Textfunktionen, respektive Textsorten: Es werden fünf ausgewählte Textfunktionen und damit verknüpfte Textsorten in jedem Testheft über die Lebensspanne als längsschnittliches Konzept realisiert, d.h. jeder Test/jedes Testheft zur Messung der Lesekompetenz enthält insgesamt fünf Texte, die den folgenden fünf Textfunktionen entsprechen: a) Sachtexte, b) kommentierende Texte, c) literarische Texte, d) Anleitungen und e) Werbetexte (Gehrler & Artelt, 2013). Im Unterschied zu PISA werden in NEPS keine diskontinuierlichen Texte wie Grafiken, Tabellen, Straßenkarten u.ä. eingesetzt. Diskontinuierliche Texte fallen aus der NEPS-Konzeption heraus, da sie spezielle Anforderungen stellen und sie zudem nicht in jedem Alter, in dem die Lesekompetenz im NEPS getestet wird, bedeutsam sind.
2. Das zweite konsistente Merkmal der Rahmenkonzeption sind die kognitiven Verstehensanforderungen der Aufgaben: a) Informationen im Text ermitteln, b) textbezogene Schlussfolgerungen ziehen und c) Reflektieren und Bewerten.
3. Die verwendeten Aufgabenformate sind die in Large Scale Assessments üblichen Arten a) Multiple Choice und b) Entscheidungstabellen (stimmt-stimmt nicht) sowie ein eigen

entwickeltes Format von c) Zuordnungsaufgaben<sup>1</sup> und d) ergänzende Formate für computerbasierte Tests (siehe 3.2).

Durch die systematische Berücksichtigung verschiedener Textfunktionen, die in unterschiedlichen Altersstufen in jeweils lebensnahen und altersangemessenen Texten, Textthemen und unterschiedlichen Verstehensanforderungen der darauf bezogenen Aufgaben umgesetzt werden, ist es möglich, Lesekompetenz als ein breit angelegtes Fähigkeitskonstrukt zu operationalisieren (vgl. Abbildung 5).

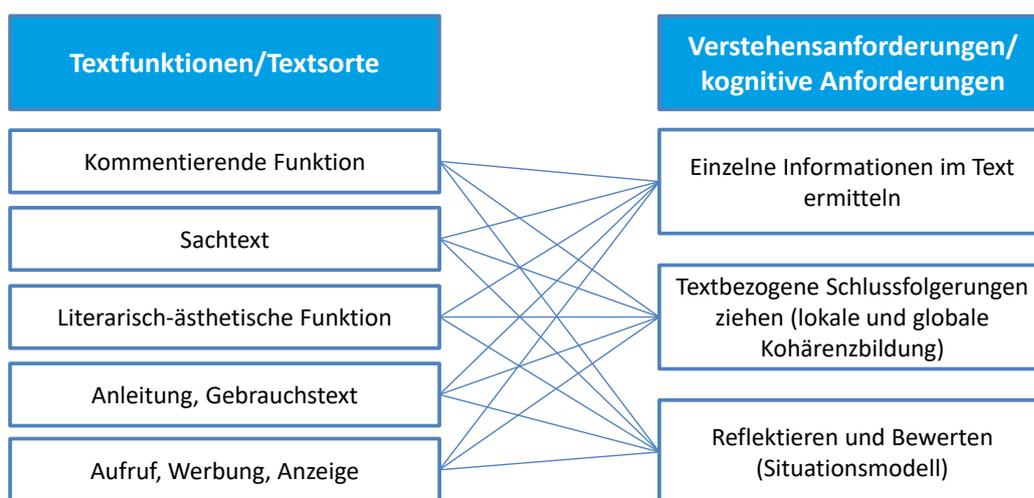


Abbildung 5. Textfunktionen und kognitive Anforderungen (Gehrer, Zimmermann, Artelt & Weinert, 2012, S.7; Gehrer, Zimmermann, Artelt & Weinert, 2013, S.63).

### 3.2 Bisherige Erhebung des Konstrukts

In der folgenden Tabelle 6 sind die bisherigen Erhebungen der Lesekompetenz im NEPS aufgelistet. In der 2. Klasse in der SC 1 und der SC 2 erfolgte die Erfassung des Leseverstehens nicht über ein in der Rahmenkonzeption entwickeltes Instrument, sondern über den ELFE-Test (Lenhard & Schneider, 2006).

Tabelle 6. Erhebungszeiten und -modi der in der Domäne Lesen gelaufenen bzw. aktuell laufenden Erhebungen

Startkohorte	Alter / Klasse	Jahr	Erhebungszeit (in Minuten) <sup>2</sup>	Erhebungsmodus
SC 2	Klasse 4	2015	28	Papier
SC 2	Klasse 7	2018	28	Papier

<sup>1</sup> Bei Aufgaben der Formate Entscheidungstabellen und Zuordnung werden Zusammenfassungen vorgenommen, so dass Antworten mit teilrichtigen Lösungen (partial credit items) entstehen; wir sprechen hier auch von komplexen Items oder Matching Items.

<sup>2</sup> Reine Testzeit ohne Instruktion und ohne Pause

Startkohorte	Alter / Klasse	Jahr	Erhebungszeit (in Minuten) <sup>3</sup>	Erhebungsmodus
SC 3	Klasse 5	2010	28	Papier
SC 3	Klasse 7	2012	28	Papier
SC 3	Klasse 9	2015	28	Papier
SC 3	Klasse 12 / Schulabgänger	2017	28	Papier
SC 4	Klasse 9	2011	28	Papier
SC 4	Klasse 12 / Schulabgänger	2013	28	Papier
SC 4	Junge Erwachsene	2016	28	Computer
SC 5	Studierende (Welle 1)	2011	28	Papier
SC 5	Studierende (Welle 12)	2017	28	Computer / Online
SC6	Erwachsene	2010 & 2012	28	Papier
SC6	Erwachsene	2016	28	Computer

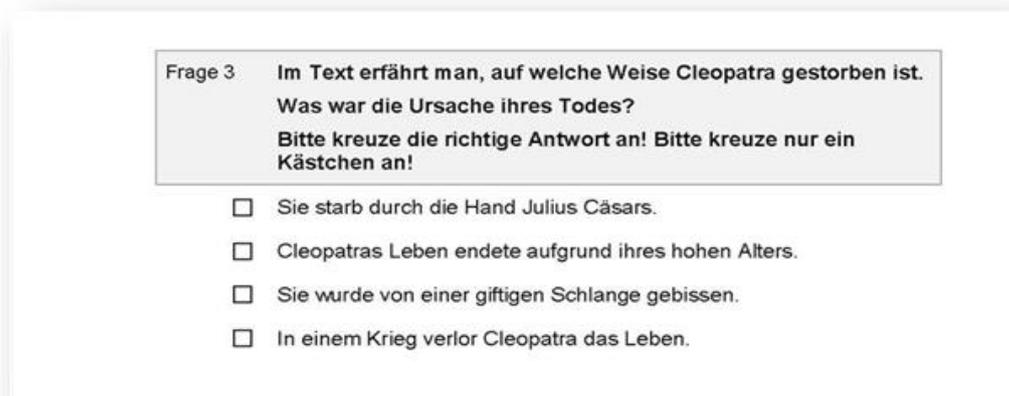
Während die Lesekompetenztests in den Anfangsjahren des NEPS papierbasiert eingesetzt wurden, findet die Administration der Tests seit 2016 weitgehend computerbasiert statt. Dies führt dazu, dass die Umsetzung komplexer Aufgabenformate zunehmend intuitiver wird für die Zielpersonen, da zum Beispiel Zuordnungsaufgaben über Drag-and-drop-Techniken umgesetzt werden. Auch die visuelle Darstellung der Aufgabenformate ist dadurch ansprechender und abwechslungsreicher, wodurch die Zielpersonen stärker in der Teilnahme motiviert sein sollten. Für die erweiterten Möglichkeiten der CBA-Testungen wurden auch zusätzliche Aufgabenformate entwickelt, die sich näher an (halb)offenen statt geschlossenen Formaten orientieren. So wurden bereits in den Haupterhebungen 2018/2019 in der SC4, SC6, SC5 komplexe Textanreicherungs-Aufgaben (TAA) administriert, welche an klassische Lückentexte erinnern; es müssen jedoch nicht einzelne Wörter in Sätze (Cloze-Tests), sondern ganze Sätze in den vorhandenen Text integriert werden, was den Aufbau eines komplexeren Situationsmodells erfordert (zur Kurzbeschreibung und Qualität siehe: Rohm et al., 2019). Das Format Highlighting-Aufgaben (HL), bei dem zur Lösungsabgabe mehrere Wörter, Satzteile oder ganze Sätze im gelesenen Text farbig markiert werden (vgl. Heyne et al., 2020), wird in den kommenden Wiederholungsmessungen (ab 2024) eingesetzt werden.

<sup>3</sup> Reine Testzeit ohne Instruktion und ohne Pause

Darüber hinaus ermöglicht die computergestützte Administrierung eine exakt standardisierte Zeitmessung und videobasierte Instruktion der Zielperson und bietet zudem durch prozessorientierte Logdaten zusätzliche Hintergrundvariablen über Nutzungs- und Strategieverhalten sowie weitere Prozesse während der Testbearbeitung. Diese zunehmende Erfassung der Lesekompetenz über computeradministrierte Tests im NEPS eröffnet zukünftig die Möglichkeit einer stärkeren Nutzung von Prozessdaten, die bspw. Aussagen über Lesestrategien und die Interaktion von Test- und Personenmerkmalen erlauben. Durch die Entwicklung neuer Aufgabenformate (Textanreicherungsaufgabe [umgesetzt in HE 2016] sowie Highlighting [in der Entwicklungsphase für HE 2024]) können am Computer motivierende Formate eingesetzt werden, welche die Aufgabenbearbeitung weiterhin attraktiv halten und zusätzliches Forschungspotenzial bieten.

### 3.3 Beispielitems

Folgende Beispielitems zeigen die in der Rahmenkonzeption beschriebenen kognitiven Anforderungen *Informationen entnehmen* (Abbildung 6) sowie *Reflektieren und Bewerten* (Abbildung 7), bzw. die beschriebenen Formate *Multiple Choice* (Abbildung 6) und *Zuordnungsaufgabe* (Abbildung 7)<sup>4</sup> in der Umsetzung für die Altersgruppe Klasse 5. Das gezeigte Textbeispiel (Abbildung 8) des Entwicklungspools ist mit 233 Wörtern etwas kürzer als der in der Haupterhebung für die fünften Klassen final eingesetzte Sachtext (307 Wörter).



Frage 3 Im Text erfährt man, auf welche Weise Cleopatra gestorben ist.  
Was war die Ursache ihres Todes?  
Bitte kreuze die richtige Antwort an! Bitte kreuze nur ein Kästchen an!

- Sie starb durch die Hand Julius Cäsars.
- Cleopatras Leben endete aufgrund ihres hohen Alters.
- Sie wurde von einer giftigen Schlange gebissen.
- In einem Krieg verlor Cleopatra das Leben.

Abbildung 6. Lesen – Multiple Choice Item Klasse 5; zu Sachtext „Ägypten“; Anforderung 1 Informationen entnehmen (Gehrer, Zimmermann, Artelt & Weinert, 2012, S. 9).

<sup>4</sup> Die Zuordnungsaufgaben für spätere Wellen wurden optimiert durch einen zweiten Distraktor-Buchstaben.

**Frage 7** Welche Teilüberschrift passt am besten zu welchem Textabschnitt?  
 Ordne die Buchstaben von A bis F den Textabschnitten zu!  
 Trage die passenden Buchstaben in die Tabelle ein! Ein Buchstabe bleibt übrig.

Abschnitt	Teilüberschrift
1	<input type="checkbox"/>
2	<input type="checkbox"/>
3	<input type="checkbox"/>
4	<input type="checkbox"/>
5	<input type="checkbox"/>

- A Fruchtbare Land
- B Gräber und Schätze
- C Werkzeuge für Pyramiden
- D Wichtige Mumienfunde
- E Cleopatra
- F Götter und Könige

Abbildung 7. Lesen – Zuordnungsaufgabe Klasse 5; Anforderung 3 Reflektieren und Bewerten (Gehrer, Zimmermann, Artelt & Weinert, 2012, S. 10).

## ÄGYPTEN - PYRAMIDEN AM NIL



(1) Entlang des Nils zieht sich ein schmaler Streifen fruchtbaren Landes. Der Rest von Ägypten ist Wüste. Doch dieser grüne Streifen war genug, um dort eine der ersten großen Kulturen der Menschheit entstehen zu lassen: Das Alte Ägypten.

(2) Den ägyptischen Staat führten die Pharaonen. Sie waren aber nicht nur Könige, sie wurden von den Ägyptern sogar als Götter verehrt. Sie wurden verehrt und nach ihrem Tode höchst prunkvoll beerdigt. Die Ägypter glaubten an ein Leben nach dem Tode und sie glaubten, dass der tote Mensch seinen Körper noch einmal benötigen würde.

(3) In den berühmtesten Gräbern, den Pyramiden, wurden keine Mumien gefunden. Wahrscheinlich waren sie, mitsamt ihren Schätzen, Opfer von Grabräubern geworden.

(4) Die Pyramiden von Gizeh sind die berühmtesten Überbleibsel des Pharaonenreiches. Fast unglaublich: Denn die alten Ägypter hatten damals noch nicht einmal Werkzeuge aus Eisen, mit denen sie die Steinblöcke hätten bearbeiten können. Ihre Beile, Hämmer und Meißel waren aus Kupfer oder Stein.



(5) Ägypten war sehr lange Zeit das mächtigste und reichste Land der Erde. Es überstand zahlreiche Eroberungen, Kriege und Fremdherrschaften, verlor aber doch nach und nach an Macht und Einfluss.  
 Nachdem die letzte ägyptische Herrscherin Cleopatra 30 vor Christus, also 3.000 Jahre nach der ersten Pharaonendynastie, an einem Schlangengift gestorben war, machte der römische Kaiser Augustus Ägypten zu einem Teil seines Reiches.  
 Damit war die Geschichte des Alten Ägypten endgültig zu Ende.

Abbildung 8. Sachtext (Auszug) „Ägypten“ aus Großpilot Klasse 5, 2009 (Gehrer, Zimmermann, Artelt & Weinert, 2012, S.7).

Das neu entwickelte Format *Textanreicherungsaufgabe* (TAA; engl. text enrichment task, TET) ist in der computerbasierten Testung für die höheren Startkohorten im Einsatz (Abbildung 9).

**Pilzzucht**

---

Textseite 1

! Frage 1

! Frage 2

! Frage 3

! Frage 4

**! Frage 5**

! Frage 5f

Frage 5:

**Wo können Sie die drei ergänzenden Sätze am sinnvollsten in den bereits gelesenen Text einfügen?**

*Ziehen Sie dazu das jeweilige Symbol an die gewünschte Stelle! Sie können die Symbole an allen Stellen ablegen, die mit einem Kreis gekennzeichnet sind. Beachten Sie bitte, dass der Text zweiseitig ist (Frage 5 und Frage 5f).*

▲ Und solche Pilze wollen Sie sicher nicht auf Ihrem Teller haben.

■ Allerdings haben nicht alle Laubhölzer den gleichen Effekt.

● Zerbrechen soll der Stamm dabei aber nicht!

**Pilzzucht auf Holz**

Für die Pilzzucht von Kräuterseitling, Shiitake oder Austernseitling verwenden Sie idealerweise Holz von Laubbäumen, welches von Herbst bis Frühling gefällt wurde.  Auf Weichhölzern wie Linde, Birke oder Weide beginnt die Ernte der Pilze noch im gleichen Jahr.  Auf Harthölzern wie Ahorn, Kastanie, Obstgehölzen, Buche oder Eiche währt die Durchwachsphase ein Jahr, dafür hält der Ertrag 5–7 Jahre lang an.

Bohren Sie ca. 30 Löcher mit 10cm Tiefe verteilt auf den ganzen Stamm und füllen Sie die Pilzsaat mit einem Trichter in die Löcher ein.  Verschließen Sie die Löcher mit Wachs oder Holzzapfen.  Die Pilze lassen Sie dann bei ca. 15 –24°C anwachsen.  Schützen Sie den Holzstamm vor trockener Zugluft und direkter Sonneneinstrahlung, um Austrocknung zu vermeiden.  Der Stamm sollte nicht mit Plastik umhüllt werden, da sich sonst Feuchtigkeit ansammelt und Schimmelpilze auftreten können.

←
→

Abbildung 9. Textanreicherungsaufgabe (Auszug) „Pilzzucht“, computerbasiert, aus Entwicklungsstudie B124, 2015 für SC 4/5/6 (Rohm, Scharl, Ettner & Gehrler, 2019, S.6).

Abbildung 10 zeigt die Videoinstruktion für die Technik Ziehen und Ablegen (drag & drop), welche für die Beantwortung der computerbasierten Versionen der *Zuordnungsaufgabe* (ZO) und TAA verwendet wird.

**Ziehen und Ablegen**

Frage :

---



---

Abbildung 10. Videobasierte Instruktion für die Technik Ziehen und Ablegen (Drag & Drop).

### 3.4 Empirische Analysen

#### 3.4.1 Deskriptive Informationen

In der nachfolgenden Tabelle 7 sind die deskriptiven Angaben für die bisher bereitgestellten Datensätze in SUF dargestellt. Ein interessantes Ergebnis dieser Daten ist die sehr geringe Ausfallquote innerhalb der Testdurchführungen, d. h. Versuchspersonen, die einmal mit der Bearbeitung der Lesetests begonnen haben, haben die Testung auch mit validen Angaben abgeschlossen, aus welchen Kompetenzwerte berechnet werden konnten. Außerdem zeigen sich die durchweg guten Reliabilitäten der Skalen, die in den verschiedenen Altersbereichen eingesetzt wurden.

Tabelle 7: Deskriptive Informationen zu Erhebungen in der Domäne Lesen

Startkohorte	Alter / Klasse	N (gesamt)	N (gültig)	% fehlend	M	SD	Reliabilität
SC 2	Klasse 4	6733	6700	< 1.00	-0.59	1.32	.82
SC 3 <sup>a</sup>	Klasse 5	5208	5193	< 1.00	-0.02	1.27	.81
SC 3 <sup>a</sup>	Klasse 7	6211	6186	< 1.00	0.79	1.36	.81
SC 3 <sup>a</sup>	Klasse 9	4624	4578	< 1.00	1.33	1.12	.81
SC 3 <sup>a</sup>	Klasse 12 / Schulabgänger	3663	3662	< 1.00	2.10	1.01	.80
SC 4	Klasse 9	13933	13897	< 1.00	-0.03	1.26	.81
SC 4	Klasse 12 / Schulabgänger	5805	5801	< 1.00	0.61	1.00	.80
SC 4 <sup>c</sup>	Erwachsene (21 Jahre)	6866	6863	< 1.00	0.57	0.96	.77 <sup>d</sup>
SC 5 <sup>b</sup>	Studierende (Welle 1)	5949	5945	< 1.00	-0.01	0.88	.64
SC 5 <sup>c</sup>	Studierende (Welle 12)	4730	4591	2.94	0.28	0.96	.77 <sup>d</sup>
SC 6	Erwachsene (Welle 3)	5349	5342	< 1.00	-0.05	1.33	.77
SC 6	Erwachsene (Welle 5)	3156	3154	< 1.00	-0.37	1.37	.80
SC 6 <sup>c</sup>	Erwachsene (Welle 9)	6460	6397	< 1.00	-0.02	1.04	.77 <sup>d</sup>

Bemerkungen. *N* (gesamt) = Anzahl der Personen, bei denen der Test administriert wurde; *N* (gültig) = Anzahl der Personen mit gültigem Testwert, % fehlend = Anteil an Personen ohne gültigen Testwert, *M* = Mittelwert des Testwerts (bei erster Administration pro Startkohorte auf 0 fixiert), *SD* = Standardabweichung des Testwerts, Reliabilität = EAP Reliabilität. <sup>a</sup> Die zusätzlich gesampelten Schülerinnen und Schüler an den Förderschulen (W1: *n*=570; W3: *n*=417) wurden in dieser Übersicht nicht berücksichtigt. <sup>b</sup> Die IRT Skalierung in der SC 5, Welle 1 erfolgte anhand der eingesetzten Stichprobe von *N*=7079. <sup>c</sup> Die Testung erfolgte 2016/2017 computerbasiert. <sup>d</sup> Die IRT Skalierung erfolgte für SC4, SC 5, SC6 2016/2017 gemeinsam.

### 3.4.2 Psychometrische Informationen

Umfangreiche psychometrische Analysen zeigen, dass das Instrument essentiell eindimensional misst und den Eigenschaften des Rasch-Modells entspricht. Detaillierte Ergebnisse dieser Analysen sind in folgenden technischen Reports nachzulesen: SC2 Survey Paper No. 30 (Rohm, Krohmer & Gnams, 2017); SC3 NEPS Working Paper No. 15 (Kranich et al., 2017) und Survey Paper No. 14 (Pohl, Haberkorn, Hardt & Wiegand, 2012), No. 20 (Scharl, Fischer, Gnams, & Rohm, 2017); SC4 NEPS Working Paper No. 16 (Haberkorn, Pohl, Hardt, & Wiegand, 2012) und Survey Paper No. 13 (Gnams, Fischer & Rohm, 2017), No. 62 (Rohm, Scharl, Ettner & Gehr, 2019); SC5 NEPS Working Paper No. 34 (Pohl, Haberkorn & Hardt, 2014), Nr. 48 (Koller, Haberkorn, & Rohm, 2014) und Survey Paper No. 62 (Rohm et al., 2019); SC6 NEPS Working Paper No. 25 (Hardt, Pohl, Haberkorn & Wiegand, 2013), und Survey Paper No. 62 (Rohm et al., 2019).

Im Folgenden sollen einige zentrale deskriptive Angaben zu den einzelnen Kohorten dargestellt werden. Die Mittelwertsunterschiede zwischen den Geschlechtern, Personen mit hohem und niedrigem sozioökonomischen Status sowie Personen mit und ohne Migrationshintergrund sind in Tabelle 8 dargestellt. Für die Qualität der Kompetenzmessung in der Domäne Lesen sprechen neben erwartungskonformen Mittelwertsunterschieden die stabilen Zusammenhänge mit inhaltlichen Konstrukten z. B. den Noten im Fach Deutsch oder motivational-affektiven Variablen, wie dem Selbstkonzept. Zu *Beginn der Sekundarstufe I* erzielen Kinder, die ein hohes Selbstkonzept im Bereich Lesen aufweisen (z.B. „Ich kann Texte sehr gut und schnell verstehen.“), auch höhere Kompetenzwerte in der Lesekompetenz zum gleichen Messzeitpunkt (Klasse 5:  $r = .36, p < .00$ ; Klasse 7:  $r = .29, p < .00$ ). Kinder mit höheren Noten im Fach Deutsch (Klasse 5:  $r = -.44, p < .00$ ; Klasse 7:  $r = -.36, p < .00$ ) schneiden ebenfalls besser in den Kompetenztestungen ab. Im Vergleich hierzu fallen die Zusammenhänge zu den Noten im Bereich Mathematik tendenziell, aber nicht statistisch bedeutsam geringer aus (Klasse 5:  $r = -.36, p < .00$ ; Klasse 7:  $r = -.28, p < .00$ ). Interessanterweise lassen sich in dieser Altersgruppe keine Zusammenhänge mit bestimmten Aspekten des Verhaltens beobachten: Die selbstberichtete Häufigkeit des Lesens ist zu beiden Messzeitpunkten nicht bzw. kaum mit der Kompetenz im Lesen (Klasse 5:  $r = -.01, p < .00$ ; Klasse 7:  $r = .06, p < .00$ ) verknüpft. Zum *Ende der Sekundarstufe I* zeigen sich in der Kohorte der Neuntklässler/innen analoge Zusammenhangsmuster mit den Noten im Fach Deutsch und affektiv-motivationalen Variablen. Die Lesekompetenz in der 9. Klasse hängt erwartungskonform mit den Noten der Schülerinnen und Schüler zum Ende der 8. Klasse ( $r = -.31, p < .00$ ) sowie mit den Noten zum Halbjahr der 9. Klasse ( $r = -.28, p < .00$ ), zum Halbjahr der 10. Klasse ( $r = -.27, p < .00$ ) und dem Ende der 11. Klasse ( $r = -.25, p < .00$ ) zusammen. Erwartungsgemäß fallen die Zusammenhänge geringer aus, wenn die Noten in der Mathematik ( $r_{G8} = -.19$ ;  $r_{G9HJ} = -.18$ ,  $r_{G10HJ} = -.15$ ,  $r_{G11} = -.12$ ) betrachtet werden; diese differenziellen Zusammenhänge zu den Noten in den beiden Fächern Deutsch und Mathematik sind in diesen höheren Altersbereichen nun auch statistisch bedeutsam mit einem kleinen Effekt. Auch zeigen sich positive Zusammenhänge mit dem Selbstkonzept der Schülerinnen und Schüler im Fach Deutsch in der 9. Klasse ( $r = .20, p < .00$ ) und dem Sachinteresse in der 10. Klasse ( $r = .20, p < .00$ ). Interessanterweise zeigen sich auch bereits in der 9. Klasse bei denjenigen Kindern höhere Lesekompetenzwerte, welche zwei Jahre später in der 11. Klasse einen Leistungskurs im Fach Deutsch wählen, als bei Schülerinnen und Schülern, die später einen Grundkurs im Lesen wählen ( $d = 0.15$ ). In der Gruppe der *Erwachsenen* korreliert die Lesekompetenz bedeutsam mit der Zeit, die die Person

täglich im Beruf und der Freizeit mit dem Lesen von Texten verbringt ( $r = .17, p < .00$ ). Dieser Effekt ist bedeutsam höher ausgeprägt, wenn nur die Zeit des beruflichen Lesens berücksichtigt wird ( $r = .21, p < .00$ ; Lesen in der Freizeit:  $r = .04, p < .00$ ). Dies kann ein Hinweis darauf sein, dass das Lesen im Beruf enger mit den komplexen Anforderungen des Textverständnisses verschiedener Textsorten zusammenhängt als das freizeitliche Lesen.

### 3.4.3 Gruppenunterschiede

Die nachfolgende Tabelle 8 fasst die standardisierten Mittelwertsunterschiede (Cohen's  $d$ ) der Testwerte in der Domäne Lesen für zentrale Gruppierungsvariablen zusammen.

Anzumerken ist zunächst, dass entgegen der Erwartung aufgrund bisheriger empirischer Forschungsarbeiten in allen Altersgruppen kaum bis geringe Geschlechtsunterschiede zugunsten der Mädchen bzw. Frauen in dieser Domäne vorliegen. In der Gruppe der Studierenden (SC 5) zeigen sich zu Studienbeginn und –ende keine Unterschiede in den Kompetenzwerten männlicher und weiblicher Teilnehmender. Eine mögliche Erklärung hierfür kann eventuell die etwas eingeschränkte Varianz in dieser Stichprobe im Vergleich zu allen anderen Startkohorten sein. Erwartungskonform fallen hingegen die Unterschiede von Personen mit hohem vs. niedrigem Sozialstatus (definiert über HISEI bzw. ISEI) aus. Personen mit hohem Sozialstatus weisen höhere Kompetenzen im Bereich Lesen aus. Beispielsweise zeigt sich in der Gruppe der Erwachsenen (SC 6, 1. MZP), dass ein hoher sozio-ökonomischer Hintergrund, operationalisiert über den ISEI des aktuellen Berufs, mit einer höheren Lesekompetenz einhergeht ( $r = .43, p < .00$ ). Ein weiteres konsistentes Ergebnis der NEPS-Kohorten ist, dass Personen ohne Migrationshintergrund höhere Werte in den Lesekompetenzen in allen Altersbereichen im Vergleich zu Personen mit Migrationshintergrund aufweisen. Erwachsene, die als höchsten Schulabschluss das Abitur angegeben haben, weisen ebenfalls erwartungsgemäß höhere Kompetenzwerte im Lesen auf als Personen mit niedrigeren Bildungsabschlüssen.

Tabelle 8: Standardisierte Mittelwertsunterschiede (Cohen's d) der Testwerte für zentrale Gruppierungsvariablen (Domäne Lesen)

Startkohorte	Alter / Klasse	Geschlecht	Sozioökonomischer Status	Migrationshintergrund	Schultyp	Bildungsabschluss
SC 2	Klasse 4	0.16	0.56	0.26		0.61
SC 3	Klasse 5	0.13	0.59	0.47	0.90	0.56
SC 3	Klasse 7	0.20	0.54	0.36	0.97	0.58
SC 3	Klasse 9	0.12	0.59	0.28	0.98	0.63
SC 3	Klasse 12 / Schulabgänger	0.28	0.45	0.32	0.76	0.48
SC 4	Klasse 9	0.26	0.71	0.49	1.14	0.59
SC 4	Klasse 12 / Schulabgänger	0.22	0.70	0.39	1.11	0.59
SC 4	21-Jährige	-0.01	0.54	0.42	0.93	0.58
SC 5	Studierende (Welle 1)	0.02	0.24	0.26		0.19
SC 5	Studierende (Welle 12)	0.06	0.29	0.22		0.32
SC 6 <sup>a</sup>	Erwachsene (Wellen 3, 5)	0.11	0.68	0.23		1.06
SC 6	Erwachsene (Welle 9)	0.05	0.69	0.10		1.02

Anmerkung. Positive Werte spiegeln einen höheren Mittelwert in der ersten Gruppe (z.B. „Frauen“) wider, während negative Werte einen höheren Mittelwert in der zweiten Gruppe (z.B. „Männer“) widerspiegeln. Sozioökonomischer Status = HISEI der Eltern bzw. eigener ISEI in SC6, Migrationshintergrund = ohne (= beide Elternteile in Deutschland geboren) vs. mit (= mind. ein Elternteil nicht Deutschland geboren), Bildungsabschluss = Abitur der Eltern, in SC6 eigenes Abitur. <sup>a</sup>In der SC 6 sind für den ersten Messzeitpunkt die Welle 3 sowie die Welle 5-Erstbefragten in den Analysen zusammengefasst.

### 3.5 Ausgewählte Publikationen mit NEPS-Daten zur Lesekompetenz

- Autorengruppe Bildungsberichterstattung (2020). Bildung in Deutschland 2020: Ein indikatorengestützter Bericht mit einer Analyse zu Bildung in einer digitalisierten Welt. Bielefeld, Deutschland: wbv Media.
- Durda, T., Artelt, C., Lechner, C., Rammstedt, B. & Wicht, A. (2020). Proficiency level descriptors for low reading proficiency: An integrative process model. *International Review of Education – Journal of Lifelong Learning (IRE)*, 66, 211–233. <https://doi.org/10.1007/s11159-020-09834-1>
- Ehrmann, L. & Wolter, I. (2018). The impact of students' gender-role orientation on competence development in mathematics and reading in secondary school. *Learning and Individual Differences*, 61, 256-264. <https://doi.org/10.1016/j.lindif.2018.01.004>
- Gehrer, K. & Nusser, L. (2020). Binnendifferenzierender Deutschunterricht und dessen Einfluss auf die Lesekompetenzentwicklung in der Sekundarstufe I. *Journal for Educational Research Online/ Journal für Bildungsforschung Online*, 12(2), 181–204.
- Gnambs, T., Stasielowicz, L., Wolter, I. & Appel, M. (2020). Do computer games jeopardize educational outcomes? A prospective study on gaming times and academic achievement. *Psychology of Popular Media Culture*, 9, 69-82. <https://doi.org/10.1037/ppm0000204>
- Lechner, C., Miyamoto, A. & Knopf, T. (2019). Should students be smart, curious, or both? Fluid intelligence, openness, and interest co-shape the acquisition of reading and math competence. *Intelligence*, 76. <https://doi.org/10.1016/j.intell.2019.101378>
- Linberg, T., Struck, O. & Bäumer, T. (2018). Vorzug Ganztagschule? Zusammenhänge mit der Kompetenzentwicklung im Bereich Lesen und Mathematik. *Zeitschrift für Erziehungswissenschaft*, 21(6), 1205–1227. <https://doi.org/10.1007/s11618-018-0830-2>
- Miyamoto, A., Pfof, M. & Artelt, C. (2018). Reciprocal relations between intrinsic reading motivation and reading competence: A comparison between native and immigrant students in Germany. *Journal of Research in Reading*, 41(1), 176–196. <https://doi.org/10.1111/1467-9817.12113>
- Thums, K., Artelt C. & Wolter, I. (2020). Reading for entertainment or information reception? Gender differences in reading preferences and their impact on text-type-specific reading competences in adult readers. *European Journal of Psychology of Education*. <https://doi.org/10.1007/s10212-020-00486-1>
- Wolter, I. & Seidel, T. (2017). Kompetent und beliebt? Der Zusammenhang von Kompetenzen und Selbstkonzepten in Mathematik und Lesen mit der wahrgenommenen Beliebtheit bei Peers [Competent and popular? The relationship of competencies and self-concepts in mathematics and reading to perceived peer popularity]. *Zeitschrift für Erziehungswissenschaft*, 20(3), 387-404. <https://doi.org/10.1007/s11618-017-0772-0>

## Literatur

- Gehrer, K. & Artelt, C. (2013). Literalität und Bildungslaufbahn: Das Bildungspanel NEPS. In C. Rosebrock & A. Bertschi-Kaufmann (Eds.), *Literalität erfassen: Bildungspolitisch, kulturell, individuell* (pp. 168–187). Weinheim, Deutschland: Juventa.
- Gehrer, K., Zimmermann, S., Artelt, C. & Weinert, S. (2012). *The assessment of reading competence (including sample items for grade 5 and 9). Status 2012* (NEPS Research Data Paper). Retrieved from University of Bamberg, National Educational Panel Study website: [https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC4/1-0-0/com\\_re\\_2012\\_en.pdf](https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC4/1-0-0/com_re_2012_en.pdf)
- Gehrer, K., Zimmermann, S., Artelt, C. & Weinert, S. (2013). NEPS framework for assessing reading competence and results from an adult pilot study. *Journal for Educational Research Online*, 5(2), 50–79. Retrieved from [www.j-e-r-o.com/index.php/jero/article/view/361/170](http://www.j-e-r-o.com/index.php/jero/article/view/361/170)
- Groeben, N. & Hurrelmann, B. (Hrsg.) (2004). *Lesesozialisation in der Mediengesellschaft: Ein Forschungsüberblick*. Weinheim: Juventa.
- Gnambs, T., Fischer, L. & Rohm, T. (2017). *NEPS technical report for reading: Scaling results of starting cohort 4 for grade 12* (NEPS Survey Paper No. 13). Bamberg: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Haberkorn, K., Pohl, S., Hardt, K. & Wiegand, E. (2012). *NEPS technical report for reading – Scaling results of starting cohort 4 in ninth grade* (NEPS Working Paper No. 16). Bamberg: Otto-Friedrich-Universität, Nationales Bildungspanel.
- Hardt, K., Pohl, S., Haberkorn, K. & Wiegand, E. (2013). *NEPS technical report for reading – Scaling results of starting cohort 6 for adults in main study 2010/11* (NEPS Working Paper No. 25). Bamberg: University of Bamberg, National Educational Panel Study.
- Heyne, N., Artelt, C., Gnambs, T., Gehrer, K. & Schoor, C. (2020). Instructed highlighting of text passages – Indicator of reading or strategic performance? *Lingua* Vol. 236. <https://doi.org/10.1016/j.lingua.2020.102803>
- Kintsch, W. (1998). *Comprehension. A paradigm for cognition*. Cambridge: University Press.
- OECD (2009). *PISA 2009 assessment framework – Key competencies in reading, mathematics, and science*. Paris: OECD
- Koller, I., Haberkorn, K. & Rohm, T. (2014). *NEPS technical report for reading: Scaling results of starting cohort 6 for adults in main study 2012* (NEPS Working Paper No. 48). Bamberg: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Krannich, M., Jost, O., Rohm, T., Koller, I., Pohl, S., Haberkorn, K., Carstensen, C. H., Fischer, L. & Gnambs, T. (2017). *NEPS technical report for reading: Scaling results of starting cohort 3 for grade 7* (NEPS Survey Paper No. 14). Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Lenhard, W. & Schneider, W. (2006). *ELFE 1-6. Ein Leseverständnistest für Erst- bis Sechstklässler*. Göttingen: Hogrefe.
- Pohl, S., Haberkorn, K. & Hardt, K. (2014). *NEPS technical report for reading – Scaling results of starting cohort 5 for first-year students* (NEPS Working Paper No. 34). Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study.

- Pohl, S., Haberkorn, K., Hardt, K. & Wiegand, E. (2012). *NEPS technical report for reading – Scaling results of starting cohort 3 in fifth grade* (NEPS Working Paper No. 15). Bamberg: Otto-Friedrich-Universität, Nationales Bildungspanel.
- Rohm, T., Krohmer, K. & Gnams, T. (2017). *NEPS technical report for reading: Scaling results of starting cohort 2 for grade 4* (NEPS Survey Paper No. 30). Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Rohm, T., Scharl, A., Ettner, J. & Gehrler, K. (2019). *NEPS technical report for reading: Scaling results of starting cohorts 4 (wave 10), 5 (wave 12), and 6 (wave 9)* (NEPS Survey Paper No. 62). Bamberg, Germany: Leibniz Institute for Educational Trajectories (LifBi), National Educational Panel Study.
- Scharl, A., Fischer, L., Gnams, T. & Rohm, T. (2017). *NEPS technical report for reading: Scaling results of starting cohort 3 for grade 9* (NEPS Survey Paper No. 20). Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study.

## 4. Naturwissenschaftliche Kompetenz

Inga Hahn & Jana Kähler

### 4.1 Einleitung

Die naturwissenschaftliche Kompetenz im NEPS basiert auf dem Literacy-Konzept (American Association for the Advancement of Science, 2009; Bybee, 1997a, 1997b; Bybee & PISA 2006 Science Expert Group, 2009) und auf dem Kompetenzbegriff von Weinert (2001) mit Erweiterung durch Klieme und Leutner (2006). Inhaltlich stellt die NEPS-Rahmenkonzeption im Bereich Naturwissenschaften eine Überschneidung der PISA-Rahmenkonzeption (OECD, 2006) und der Konzeption der Bildungsstandards (KMK, 2005a, 2005b, 2005c) dar. Dies zeigt sich zum einen in den Inhalten und zum anderen in den beiden Subfaktoren der inhaltsbezogenen (knowledge of science – KOS) und prozessbezogenen (knowledge about science – KAS) Komponenten. Die NEPS-Konzeption naturwissenschaftlicher Kompetenz folgt damit der Forderung, die Kompetenztests des NEPS mit anderen großen Large-Scale-Assessments zu verbinden. Vertiefende Informationen zur Rahmenkonzeption und zur Konzeption der Tests finden sich bei Hahn et al. (2013).

Naturwissenschaftliche Kompetenz ist ein bedeutsamer Faktor für die aktive Teilhabe an der modernen Wissensgesellschaft. Sie ist für viele Bereiche des täglichen Lebens bedeutsam, wenn es darum geht, natürliche oder technische Phänomene zu verstehen. Naturwissenschaftliche Kompetenz wird u.a. benötigt, wenn wir moderne Technologien nutzen, die Ausbreitung von Krankheiten oder den Klimawandel verstehen oder neue naturwissenschaftliche Erkenntnisse beurteilen wollen. Umweltschutz und ein nachhaltiger Umgang mit der Umwelt werden durch sie erst ermöglicht.

Die besondere Bedeutung naturwissenschaftlicher Kompetenz wird seit langem hervorgehoben. Studien wie TIMSS (Martin & Mullis, 2013), PISA (Prenzel, Artelt, Baumert, Blum, Hammann, Klieme & Pekrun, 2007) und die Prüfung der Erreichung der Bildungsstandards (KMK, 2005a, 2005b, 2005c) geben regelmäßig Auskunft über den Stand der naturwissenschaftlichen Kompetenz deutscher Schülerinnen und Schüler. Diese Erhebungen stellen keinen Selbstzweck dar. Zum einen geben Sie Aufschluss über das Bildungssystem, zum anderen ist der Stand der Kompetenzen in der Bevölkerung auch von ökonomischer Bedeutung. Naturwissenschaftliche Kompetenzen besitzen in den OECD-Staaten im Vergleich zu mathematischen Kompetenzen und Lesekompetenzen den stärksten Einfluss auf das Wirtschaftswachstum eines Landes (Hanushek & Wößmann, 2015). Es geht hier also im weiteren Sinne auch um die wirtschaftliche Konkurrenzfähigkeit eines Landes.

Anders als die genannten Querschnittstudien bietet das NEPS erstmalig in Deutschland die Möglichkeit, die naturwissenschaftliche Kompetenz längsschnittlich zu untersuchen. Dies stellt insbesondere für den frühkindlichen Bereich und für den Erwachsenenbereich ein Novum dar. Für diese Altersgruppen sind der Stand und die Entwicklung naturwissenschaftlicher Kompetenzen vergleichsweise unerforscht. Studien wie BIKS (Kotzerke, Ebert & Weinert, 2014) haben sich zwar einer ähnlichen Altersgruppe gewidmet, aber dabei nicht die naturwissenschaftliche Kompetenz der Kinder in den Blick genommen. Ebenso ist es mit der PIACC-Studie (Zabal et al., 2013), welche die Lesekompetenzen sowie mathematische und Problemlösekompetenzen Erwachsener untersucht, dabei aber die naturwissenschaftlichen

Kompetenzen außer Acht lässt. Inhaltlich besonders interessant sind die Übergangsbereiche vom Kindergarten zur Grundschule oder von der Ausbildung zum Arbeitsleben. Es ist zu erwarten, dass sich durch die Skizzierung der Entwicklung der naturwissenschaftlichen Kompetenz über die Lebensspanne sensible Phasen der Entwicklung und entsprechender Bildungsentscheidungen aufdecken lassen. Anhand dieser empirischen Belege können gezielt Fördermaßnahmen ergriffen werden, um den in der Bildungsforschung viel kritisierten Problematiken, wie z.B. dem anhaltenden mangelnden Interesse an den Naturwissenschaften, entgegen zu treten.

Vor dem Hintergrund des Klimawandels und der „Fridays for Future“-Initiative, in Fragen nachhaltiger Energien und angesichts ungelöster Fragen, die die Gesundheit der Menschen betreffen, wie z.B. multiresistente Keime und die Ausbreitung virusbedingter Erkrankungen, ist die naturwissenschaftliche Kompetenz der Menschen vielleicht wichtiger als je zuvor. Wir benötigen diese Kompetenz, um politische Entscheidungen, die diese Themenfelder betreffen, nachvollziehen und bewerten zu können und nicht zuletzt, um Schlussfolgerungen für das eigene Handeln ziehen zu können.

## 4.2 Bisherige Erhebung des Konstrukts

Die naturwissenschaftliche Kompetenz wurde im NEPS bisher überwiegend anhand von Papier- und Bleistifttests erhoben. Eine Ausnahme stellt hier die Startkohorte 1 dar, in der tabletbasierte Naturwissenschaftstests eingesetzt wurden. Diese wurden anhand einer Software erstellt, die eine sehr einfache Testentwicklung und ökonomische Testanwendung erlaubt. Die Administration ist intuitiv und sehr standardisiert. Gleichzeitig macht den Kindern das Lösen der Aufgaben am Tablet Spaß.

Im Rahmen der Digitalisierung und auch coronabedingt verändert sich der Erhebungsmodus in den anderen Startkohorten des NEPS nun zunehmend. In Zukunft wird die in Startkohorte 1 eingesetzte Software bei allen Naturwissenschaftstests der neuen Startkohorten des NEPS zum Einsatz kommen. Um die technischen Möglichkeiten in vollem Umfang zu nutzen, sind für die Zukunft auch simulationsbasierte Aufgaben geplant, die es ermöglichen sollen, auch komplexere Sachverhalte abzufragen.

Die folgende Tabelle gibt einen Überblick über die Dauer und die Erhebungsmodi der bisher in der Domäne Naturwissenschaften gelaufenen bzw. aktuell anstehenden Erhebungen (grau hervorgehoben).

Tabelle 9: Übersicht über Dauer und Zuständigkeiten der Naturwissenschaftserhebungen

Startkohorte	Alter / Klasse	Jahr	Erhebungsdauer (in Minuten)	Erhebungsmodus
SC 1	5 Jahre	2017	20	Computer (Tablets)
SC 1	7 Jahre*	2019	20	Computer (Tablets)
SC 1	9 Jahre	2021	20	Computer (Tablets)

Startkohorte	Alter / Klasse	Jahr	Erhebungsdauer (in Minuten)	Erhebungsmodus
SC 2	4 Jahre	2011	30	Papier
SC 2	Klasse 1	2013	30	Papier
SC 2	Klasse 3	2014/15	29	Papier
SC 2	Klasse 7*	2018/19	29	Papier
SC 3	Klasse 6	2011/12	29	Papier
SC 3	Klasse 9	2014/15	29	Papier
SC 3	Klasse 11	2016	29	Papier
SC 4	Klasse 9	2010/11	29	Papier
SC 4	Klasse 11	2012/13	29	Papier
SC 5	Studierende	2013	29	Papier / Computer / Online
SC 6	Erwachsene (Welle 5)	2012/13	25	Papier

Bemerkung. \* Diese Erhebungen wurden bisher noch nicht ausgewertet. Sie stehen für 2020 mit der Auswertung und SUF-Erstellung an.

### 4.3 Beispielitems

## Tierquiz

Hier siehst du einen jungen Vogel.  
Was meinst du, wer könnte seine Mutter sein?

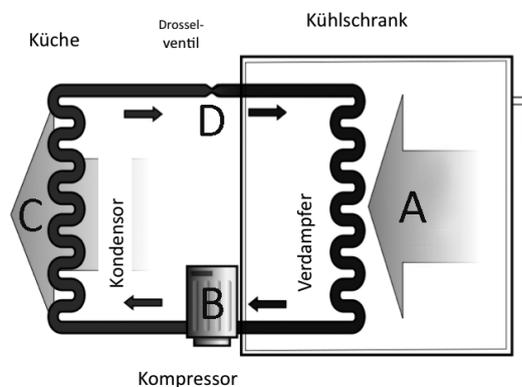


Abbildung 11. Beispiel eines Items für Kindergartenkinder (4 Jahre) (Kontext: Umwelt; Komponente: Entwicklung).

## ABKÜHLUNG GEFÄLLIG?

Kühlschränke in Privathaushalten basieren häufig auf dem Kompressorprinzip. Dabei gelangt ein Kühlmittel in flüssiger Form in den Verdampfer im Kühlraum (A) und verdampft dort. Die Energie für das Verdampfen wird dem Kühlschrankinneren in Form von Wärme entzogen. Das gasförmige Kühlmittel wird dann mit einem Kompressor abgesaugt und unter einem Druck von 8 bar verdichtet (B). Das unter hohem Druck stehende Gas gelangt in den Kondensator, gibt dort Wärme an die Umgebung ab und wird dort wieder flüssig (C). Ein Drosselventil reduziert den Druck auf 1 bar und das flüssige Kühlmittel gelangt erneut in den Kühlraum (D).

Der Siedepunkt des Kühlmittels in einem offenen Gefäß liegt bei ca.  $-30\text{ }^{\circ}\text{C}$ . Warum kondensiert es im Kondensator bei Raumtemperatur?



Bitte kreuzen Sie die richtige Antwort an! Bitte kreuzen Sie nur ein Kästchen an!

<input type="checkbox"/>	Der Siedepunkt hat nichts mit dem Kondensationspunkt zu tun.
<input type="checkbox"/>	Der Siedepunkt einer Flüssigkeit erhöht sich mit dem Druck.
<input type="checkbox"/>	Bei einem Druck von 8 bar wird jedes Gas flüssig.
<input type="checkbox"/>	Durch die Druckerhöhung kühlt das Gas auf $-30\text{ }^{\circ}\text{C}$ ab.

Abbildung 12. Beispiel eines Items für Studierende (Kontext: Technologie; Komponente: System).

## HERZ-KREISLAUF-SYSTEM

Kaffee wirkt aufgrund des Koffeingehaltes anregend auf den Kreislauf. Menschen, die bereits unter hohem Blutdruck leiden, sollten allerdings soweit wie möglich auf den Genuss von Kaffee verzichten und auf koffeinfreie Getränke ausweichen. Wissenschaftler vermuten nun aber, dass auch koffeinfreier Kaffee den Blutdruck eines Menschen erhöht.

Wie könnte eine entsprechende Studie angelegt sein?

	Testgruppe	Messung	Zeitpunkt der Messung
<b>Versuchsplan A</b>	Eine Testgruppe: trinkt zunächst koffeinfreien Kaffee, dann koffeinhaltigen Kaffee	Blutdruck	vor und nach dem Trinken
<b>Versuchsplan B</b>	Gruppe 1 trinkt koffeinfreien Kaffee, Gruppe 2 trinkt koffeinhaltigen Kaffee	Blutdruck	nach dem Trinken
<b>Versuchsplan C</b>	Gruppe 1 trinkt koffeinfreien Kaffee, Gruppe 2 trinkt Wasser	Blutdruck	nach dem Trinken
<b>Versuchsplan D</b>	Gruppe 1 hat Bluthochdruck, Gruppe 2 ist gesund; beide Gruppen trinken koffeinfreien Kaffee	Blutdruck	vor und nach dem Trinken

*Bitte kreuzen Sie die richtige Antwort an! Bitte kreuzen Sie nur ein Kästchen an!*

<input type="checkbox"/>	Versuchsplan A
<input type="checkbox"/>	Versuchsplan B
<input type="checkbox"/>	Versuchsplan C
<input type="checkbox"/>	Versuchsplan D

*Abbildung 13.* Beispiel eines Items für Erwachsene (Kontext: Gesundheit; Komponente: naturwissenschaftliche Denk- und Arbeitsweisen: Planung von Experimenten).

## 4.4 Empirische Analysen

### 4.4.1 Deskriptive Informationen

Die folgende Tabelle 10 gibt einen Überblick über die deskriptiven Ergebnisse der bisherigen Erhebungen in der Domäne Naturwissenschaften. Die Ergebnisse deuten auf eine gute bis sehr gute Qualität der Tests hin. Einschränkungen in der Varianz gibt es unter anderem bei den Vor- und Grundschultests. Dies mag darauf zurückzuführen sein, dass die Testaufgaben vorgelesen wurden und nicht individuell im eigenen Tempo bearbeitet werden konnten. Die Bearbeitungsdauer und die Anzahl der bearbeiteten Aufgaben waren abhängig vom Testleiter und der Testgruppe. Weitere Einschränkungen in der Varianz finden sich in der Klasse 11 und bei den Studierenden. Hier handelt es sich allerdings um selektierte Stichproben. Die Klasse 11-Schülerinnen und –Schüler bestehen lediglich aus Gymnasiasten und die Studierenden sind Personen mit Abitur, die zum Studium zugelassen wurden, was die Heterogenität der Gruppen einschränkt. Insgesamt ist der Prozentsatz der Personen mit zu wenig validen Antworten bei allen Tests mit maximal < 4.00% sehr gering.

Tabelle 10: Übersicht über deskriptive Kennwerte der bisherigen Erhebungen in der Domäne Naturwissenschaften

Startkohorte	Alter / Klasse	<i>N</i> (gesamt)	<i>N</i> (gültig)	% fehlend	<i>M</i>	<i>SD</i>	Reliabilität
SC 1	5 Jahre	2080	2060	< 1.00	-0.00	0.98	.67
SC 2	5 Jahre	2955	2947	< 1.00	-0.00	1.04	.75
SC 2	Klasse 1	6500	6496	< 1.00	1.38	0.93	.74
SC 2	Klasse 3	5800	5620	3.10	2.78	0.93	.70
SC 3	Klasse 6	4871	4871	< 1.00	0.00	1.18	.80
SC 3	Klasse 9	4882	4880	< 1.00	0.00	0.92	.80
SC 3	Klasse 11	1930	1930	< 1.00	0.01	0.92	.70
SC 4	Klasse 9	14475	14474	< 1.00	-0.00	1.00	.80
SC 4	Klasse 11	4417	4417	< 1.00	1.83	0.86	.64
SC 5	Studierende	8767	8442	3.71	-0.14	0.97	.65
SC 6	Erwachsene	6669	6656	< 1.00	-0.01	1.17	.72

Bemerkung. *N* (gesamt) = Anzahl der Personen, bei denen der Test administriert wurde; *N* (gültig) = Anzahl der Personen mit gültigem Testwert, % fehlend = Anteil an Personen ohne gültigen Testwert, *M* = Mittelwert des Testwerts (bei erster Administration pro Startkohorte auf 0 fixiert; für die SC3 stehen noch keine verlinkten WLEs zur Verfügung), *SD* = Standardabweichung des Testwerts, Reliabilität = EAP Reliabilität. Die Ergebnisse der Studierenden basieren auf den zusammengefassten Daten der papier- computerbasierten und Online-Erhebungen.

#### 4.4.2 Psychometrische Informationen

Die psychometrischen Analysen der Tests in den einzelnen Kohorten zeigen, dass die erfassten Daten zum einen – wie angelegt – ein eindimensionales Konstrukt erfassen bzw. sich durch ein eindimensionales Raschmodell abbilden lassen. Zum anderen gibt es Erhebungen, in denen ein zweidimensionales Modell die Daten besser erklärt, wobei die beiden Subdimension KOS und KAS (s. Einleitung) in diesen Fällen substantiell ( $r > .90$ ) korrelieren. Detaillierte Ergebnisse dieser Analysen sind in den jeweiligen Technical Reports der Erhebungen (z.B. Schöps, 2013; Schöps & Saß, 2013) nachzulesen.

Berechnet man latente Korrelationen der naturwissenschaftlichen Kompetenz mit dem Naturwissenschaftsinteresse und zur Abgrenzung mit dem Mathematikinteresse, so findet sich eine signifikante und in der Höhe angemessene Korrelation mit dem Naturwissenschaftsinteresse ( $r = .20, p < .01$ ) und eine nicht signifikante Korrelation mit dem Mathematikinteresse ( $r = -.01, p > .05$ ).

*Ergebnisse der Validierungsstudie „PISA, Bildungsstandards und das NEPS – Bildungsverläufe in Deutschland: Vergleich der Rahmenkonzepte und Validierung der NEPS-Testinstrumente in den Naturwissenschaften und in der Mathematik“*

Eine umfassende Validierung der Testwertinterpretationen des NEPS-Naturwissenschaftstests der Startkohorte 4 wurde 2012 für die neunte Klassenstufe durchgeführt. An dieser Untersuchung haben 1965 Schülerinnen und Schüler aus 80 Schulen teilgenommen, die an zwei unterschiedlichen Tagen die Naturwissenschaftstests von NEPS, PISA und des Ländervergleichs (LV) bearbeitet haben.

Die Ergebnisse der Studie lassen Rückschlüsse auf hohe Zusammenhänge in den Testwertinterpretationen der drei Tests zu. So konnte eine Korrelation von  $r = .85 (p < .01)$  zwischen den naturwissenschaftlichen Kompetenzen gemessen mit dem NEPS- und PISA-Test sowie eine Korrelation von  $r = .83 (p < .01)$  zwischen den naturwissenschaftlichen Kompetenzen gemessen mit dem NEPS- und LV-Test gefunden werden (Wagner, Schöps, Hahn, Pietsch & Köller, 2014). Die Untersuchung zeigt einen im Vergleich dazu etwas niedrigeren Zusammenhang des NEPS Naturwissenschaftstests für K9 mit dem NEPS Mathematiktest ( $r = .78, p < .01$ ) sowie den Mathematiktests aus PISA ( $r = .78, p < .01$ ) und LV ( $r = .77, p < .01$ ). Hier lässt sich die naturwissenschaftliche Kompetenz insofern von der mathematischen Kompetenz abgrenzen als diese Korrelationen deutlich geringer ausfallen, als die Korrelationen gleicher Konstrukte (Nawi) in den unterschiedlichen Erhebungen PISA, LV und NEPS. Des Weiteren korreliert der NEPS Naturwissenschaftstest für K9 zu  $r = .61 (p < .01)$  (Wagner et al., 2019) mit dem figuralen Teil des Berliner Tests zur Erfassung fluider und kristalliner Intelligenz (Wilhelm, Schroeders & Schipolowski, 2013), was eine in dieser Höhe übliche Korrelation zwischen Intelligenz- und Kompetenztests darstellt. Zur tiefer gehenden Überprüfung der Eindimensionalität des Konstrukts wurden weiterhin die partiellen Korrelationen (Huynh, Michaels & Ferrara, 1995) der Items des NEPS-Naturwissenschaftstests für die neunte Klasse analysiert. Nach der Herauspartialisierung des Konstrukts „naturwissenschaftliche Kompetenz“ liegen die paarweisen Korrelationen der Items des NEPS-Naturwissenschaftstests in einem Bereich von  $-.12$  bis  $.12$ . Gemäß Huynh et al. (1995) kann dann von der Eindimensionalität eines Konstrukts gesprochen werden, wenn die Partialkorrelationen in einem Bereich zwischen  $-.20$  bis  $.20$  liegen. Der vorliegende Befund spricht demnach dafür,

dass die mit dem NEPS-Test erfasste naturwissenschaftliche Kompetenz ein eindimensionales Konstrukt darstellt (Wagner et al., 2019). Die Ergebnisse der Validierungsstudie geben Hinweise darauf, dass der NEPS-Test zur Erfassung naturwissenschaftlicher Kompetenz für die neunte Klasse ähnliche Aussagen in Bezug auf die naturwissenschaftliche Kompetenz der Kinder treffen kann wie es PISA und der Ländervergleich tun. Hinzu kommt, dass der NEPS-Test effizient misst. Mit lediglich 30 Minuten Testzeit und lediglich 28 Items ermöglicht er valide und reliable Aussagen über die Kompetenz der Kinder.

#### *Ergebnisse aus dem Standardsetting/ den Erwachsenenendaten*

Die Erhebung der naturwissenschaftlichen Kompetenz der Erwachsenen wurde erstmalig im Jahr 2012 vorgenommen. Insgesamt nahmen 6669 Erwachsene an der Erhebung teil. Da es bisher international keine Befunde zu den naturwissenschaftlichen Kompetenzen Erwachsener gab, wurde im Rahmen eines Dissertationsprojekts eine Studie aufgelegt, die über die rein deskriptive Feststellung der Kompetenzen einen inhaltlichen und qualitativen Zugang zur Beschreibung naturwissenschaftlicher Kompetenzen Erwachsener in Deutschland ermöglichen sollte. In einem mehrstufigen Standard Setting Verfahren erarbeiteten insgesamt 10 Expertinnen und Experten einen Katalog an Fähigkeiten und Fertigkeiten, welche drei Kompetenzstufen differenzieren. Durch die Anwendung dieses Standards auf die in 2012 gewonnenen Daten konnte ermittelt werden, dass ungefähr 40.00% der Erwachsenen das Mindestniveau an naturwissenschaftlicher Kompetenz nicht erreichen (Haschke, Hahn, Bernholt & Köller, in Vorbereitung). Weiterführend konnte gezeigt werden, dass vor allem Bildungs- und Altersfaktoren maßgeblich zu Gruppenunterschieden in der erwachsenen Bevölkerung beitragen. Die naturwissenschaftlichen Fähigkeiten der Erwachsenen sind dabei nicht als isolierte Kompetenzen zu betrachten, sondern stehen mit anderen erhobenen Domänen des NEPS im engen Zusammenhang. Die latenten Korrelationen zeigen Zusammenhänge naturwissenschaftlicher Kompetenz mit Lesen von  $r = .79$  ( $p < .01$ ), mit ICT von  $r = .81$  ( $p < .01$ ) und mit Mathematik von  $r = .84$  ( $p < .01$ ). In darauf aufbauenden Analysen werden diese Zusammenhänge zukünftig detaillierter analysiert. Faktoren der Arbeitsanforderungen und das Weiterbildungsverhalten der Erwachsenen sind dabei zentral. Dafür werden weitere Messzeitpunkte in der Erwachsenenkohorte bedeutsam sein. Mit den Daten zur naturwissenschaftlichen Kompetenz von Erwachsenen in Verbindung mit dem gesetzten Standard wird dieses Forschungsfeld auch für andere Wissenschaftsbereiche geöffnet. Vor allem der Arbeitsmarktforschung und den Wissenschaftlerinnen und Wissenschaftlern der Erwachsenenbildung werden bislang unzugängliche Informationen geboten. Die Daten zur naturwissenschaftlichen Kompetenz Erwachsener sind weltweit einmalig und stellen ein Alleinstellungsmerkmal des NEPS dar.

### *Quer- und längsschnittliche Ergebnisse zum Effekt von Struktur- und Prozessvariablen auf die naturwissenschaftliche Kompetenz*

In einer ersten querschnittlichen Betrachtung der naturwissenschaftlichen Kompetenz von vier- bis sechsjährigen Kindern der Startkohorte 2 gingen Hahn und Schöps (2019) unter anderem der Frage nach, inwiefern sich migrationsbedingte Disparitäten in der naturwissenschaftlichen Kompetenz bereits bei so jungen Kindern zeigen. Weiterhin untersuchten die Autorinnen anhand von Regressionsmodellen, welche Struktur- und Prozessmerkmale einen Effekt auf die naturwissenschaftliche Kompetenz der Kinder haben. Die Ergebnisse verdeutlichen, dass der rezeptive Wortschatz in Deutsch den wichtigsten Prädiktor für die Ausprägung der naturwissenschaftlichen Kompetenz darstellt. Kinder mit Migrationshintergrund schnitten hinsichtlich ihrer Leistung im Naturwissenschaftstest signifikant schlechter ab als Gleichaltrige ohne Migrationshintergrund, was verdeutlicht, dass entsprechende migrationsbedingte Disparitäten bereits im Vorschulbereich existieren. Außerdem bestätigen die Analysen, dass Kinder mit Migrationshintergrund über einen signifikant geringeren rezeptiven Wortschatz in Deutsch verfügen als Kinder ohne Migrationshintergrund. Dies ist insofern besorgniserregend, da der Wortschatz nachweislich mit der Ausbildung von Konzepten verknüpft ist (Weinert, 2004) und Sprache eine Grundlage zum Erwerb naturwissenschaftlicher Kompetenzen darstellt (Sodian, 2002).

Aufbauend auf diesen Ergebnissen wurden in einer zweiten umfangreicheren Untersuchung der Daten der Startkohorte 2 die Effekte von Struktur- und Prozessmerkmalen aus Elternhaus und Kindergarten auf die naturwissenschaftliche Kompetenz erforscht (Kähler, Hahn, Ihme & Köller, 2020). In einem Mehrebenenmodell wurde untersucht, inwiefern sich die gefundenen sozial- und migrationsbedingten Disparitäten auch über den Wortschatz vermittelt auf die naturwissenschaftliche Kompetenz auswirken. Es zeigte sich, dass auch in dieser Untersuchung der Wortschatz der Kinder den stärksten Prädiktor für die naturwissenschaftliche Kompetenz darstellte. Darüber hinaus erwies er sich auch als Mediator für sozial- und migrationsbedingte Disparitäten: Kinder mit einem Migrationshintergrund bzw. Kinder aus einem nicht deutschsprachigen Zuhause zeigten einen geringeren Wortschatz und damit auch eine geringere naturwissenschaftliche Kompetenz. Ein naturwissenschaftlicher Schwerpunkt des Kindergartens trug unmittelbar zu einem höheren naturwissenschaftlichen Kompetenzniveau der Kinder bei.

Des Weiteren liegen erste Erkenntnisse aus einer längsschnittlichen Untersuchung der naturwissenschaftlichen Kompetenz junger Kinder vor (Kähler, Hahn & Köller, 2020). Auch hier wurden die Daten der Startkohorte 2 des NEPS untersucht. Insgesamt liegen aus drei Messzeitpunkten naturwissenschaftliche Kompetenzen vor: Kindergarten (4- bis 6-Jährige, Welle 1), Klasse 1 (7-Jährige, Welle 3) und Klasse 3 (9-Jährige, Welle 5). Mit Hilfe von latenten Wachstumskurvenmodellen wurde die Entwicklung der naturwissenschaftlichen Kompetenz untersucht. Zum einen zeigen die Ergebnisse erneut, dass es bereits im Kindergarten signifikante Kompetenzunterschiede zwischen den Kindern gibt. Vor allem sprachliche und soziale Nachteile wirken sich hierbei negativ auf die anfängliche naturwissenschaftliche Kompetenz aus. Des Weiteren zeigt sich, dass die naturwissenschaftliche Kompetenz (linear) wächst, es aber keine interindividuellen Unterschiede im Wachstum gibt. Das heißt, die im Kindergarten bestehenden Unterschiede bleiben über die Zeit in der Grundschule bestehen. Es gelingt also nicht, die bestehende Lücke zwischen sozial- und sprachlich benachteiligten

Kindern zu ihren Peers zu schließen. Es sind jedoch weitere längsschnittliche Untersuchungen nötig, um diese Befunde zu replizieren.

#### **4.4.3 Gruppenunterschiede**

Die nachfolgende Tabelle 11 fasst die standardisierten Mittelwertsunterschiede (Cohen's  $d$ ) der Testwerte für zentrale Gruppierungsvariablen zusammen. Die NEPS-Daten zur naturwissenschaftlichen Kompetenz bestätigen die in der Literatur beschriebenen Unterschiede zwischen Frauen und Männern zugunsten der Männer. Lediglich im Elementar- und Primarbereich unterscheiden sich Mädchen und Jungen noch nicht wesentlich in ihrer naturwissenschaftlichen Kompetenz. Die geringen Unterschiede fallen hier noch zugunsten der Mädchen aus.

Auch die übrigen Ergebnisse der Gruppenanalysen fallen erwartungskonform aus. Personen mit höherem sozioökonomischem Status (operationalisiert anhand des HISEI) besitzen eine höhere naturwissenschaftliche Kompetenz als Personen mit niedrigem Status. Ebenso erwartungskonform sind die Ergebnisse zugunsten der Personen ohne Migrationshintergrund und zugunsten von Personen auf dem Gymnasium bzw. zugunsten von Personen, deren Eltern als Bildungsabschluss „Abitur“ angeben.

Tabelle 11: Standardisierte Mittelwertsunterschiede der Testwerte für zentrale Gruppierungsvariablen (Domäne Naturwissenschaften)

Startkohorte	Alter / Klasse	Geschlecht	Sozio-ökonomischer Status	Migrationshintergrund	Schultyp	Bildungsabschluss
		(Frauen vs. Männer)	(hoch vs. gering)	(ohne vs. mit)	(Gymnasium vs. Nicht-Gymnasium)	(Abitur vs. kein Abitur)
SC 1	5 Jahre	0.09	0.51	0.68	-	0.58
SC 2	5 Jahre	0.07	0.53	0.57	-	0.52
SC 2	Klasse 1	0.04	0.50	0.39	-	0.55
SC 2	Klasse 3	-0.07	0.52	0.37	-	0.58
SC 3	Klasse 6	-0.18	0.61	0.62	0.92	0.60
SC 3	Klasse 9	-0.14	0.60	0.50	1.03	0.62
SC 3	Klasse 11	-0.44	0.42	0.27	0.50	0.45
SC 4	Klasse 9	-0.17	0.51	0.62	1.16	0.63
SC 4	Klasse 11	-0.43	0.17	0.41	0.69	0.24
SC 5	Studierende*	-0.53	0.17	0.24	-	0.22
SC 6	Erwachsene	-0.43	0.74	0.22	-	0.99

Bemerkung. Positive Werte spiegeln einen höheren Mittelwert in der ersten Gruppe (z.B. „Frauen“) wider, während negative Werte einen höheren Mittelwert in der zweiten Gruppe (z.B. „Männer“) widerspiegeln. \* Die Ergebnisse der Studierenden basieren auf den zusammengefassten Daten der papier- und der computerbasierten Erhebungen.

Sozioökonomischer Status = HISEI der Eltern bzw. eigener ISEI in SC6, Migrationshintergrund = ohne (= beide Elternteile in Deutschland geboren) vs. mit (= mind. ein Elternteil nicht Deutschland geboren).

## 4.5 Ausblick

Ökonomische Betrachtungen von Hanushek und Wößmann (2015) zeigen, dass naturwissenschaftliche Kompetenzen in den OECD-Staaten im Vergleich zu Lesekompetenzen und mathematischen Kompetenzen den stärksten Einfluss auf das Wirtschaftswachstum besitzen. Betrachtet man dazu den demographischen Wandel sowie den Fachkräftemangel, der sich auch auf naturwissenschaftlich-technische Berufe bezieht, so steht es außer Frage, dass es nötig ist, mehr über den Stand und die Entwicklung naturwissenschaftlicher Kompetenzen zu erfahren, um diese systematisch fördern zu können. Dazu gehört auch die genauere Erforschung der Einflüsse institutioneller und häuslicher Struktur- und Prozessmerkmale im Elementar- und Primarbereich sowie der Einflüsse der elterlichen Naturwissenschaftskompetenz auf die naturwissenschaftliche Kompetenz der Kinder. Dieser Frage wurde bereits für mathematische und sprachliche Kompetenzen nachgegangen (Ehmke & Siegle, 2008; Kluczniok, Lehrl, Kuger & Rossbach, 2013). Für die naturwissenschaftliche Kompetenz gibt es diesbezüglich erste Befunde auf Basis der Daten der Startkohorte 2. Allerdings mangelt es den bisherigen NEPS-Daten an Angaben zu naturwissenschaftsbezogenen häuslichen und institutionellen Aktivitäten und Ausstattungen der Lernumgebungen. Diese Informationen könnten in Teilen durch Daten einer neuen Startkohorte 1 zur Verfügung gestellt werden. Ein Oversampling von Kindern mit geringerem sozioökonomischem Status oder von Kindern mit Migrationshintergrund könnte Aufschluss darüber geben, ob und wie z.B. systematische (sprachliche) Fördermaßnahmen einen kompensatorischen Effekt auf die Entwicklung der naturwissenschaftlichen Kompetenz haben.

In Bezug auf den Fachkräftemangel wäre es weiterhin interessant, inwiefern Erwachsene auch nach dem Eintritt ins Rentenalter erwerbstätig bleiben und wie sich ganz allgemein-naturwissenschaftsbezogene Tätigkeiten sowie Fort- und Weiterbildungen auf die Ausprägung naturwissenschaftlicher Kompetenzen auswirken. Hierzu gibt es erste Ansatzpunkte in den bestehenden Daten, doch es wäre wichtig, den weiteren Verlauf zu betrachten und durch das Ziehen einer neuen Erwachsenenkohorte zu untersuchen, inwiefern aufgrund des fortschreitenden demographischen Wandels Unterschiede zwischen den Kohorten festzustellen sind. Für die zukünftigen Erhebungen der Naturwissenschaftskompetenz im Bereich der Erwachsenen wird ein Multistage-Test zur Verfügung stehen, um die Stichprobe in ihren Fähigkeiten noch besser abzubilden. Diese gestufte Erfassung wird erstmalig in der Erhebung in 2021 eingesetzt.

Weitere Fragestellungen im Bereich der Naturwissenschaften beschäftigen sich mit Prozessen der Testbearbeitung. Log- und Prozessdaten können dazu dienen, das Antwortverhalten der Testpersonen genauer zu untersuchen: Haben sie sich direkt für eine Antwortalternative entschieden oder waren sie unentschlossen? Lässt ihr Antwortverhalten auf Raten schließen? Diese Fragen könnten anhand computerbasierter Tests beantwortet werden und die Kontrolle dieser Faktoren könnte zur Validität der Antworten beitragen.

## 4.6 Ausgewählte Publikationen mit NEPS Daten zur naturwissenschaftlichen Kompetenz

- Hahn, I., Schöps, K., Rönnebeck, S., Martensen, M., Hansen, S. Saß, S., Dalehefte, I.M. & Prenzel, M. (2013). Assessing scientific literacy over the lifespan - A description of the NEPS science framework and the test development. *Journal for Educational Research Online* 5(2), 110-138.
- Hahn, I. & Schöps, K. (2019). Bildungsunterschiede von Anfang an? Die Bedeutung von Struktur- und Prozessmerkmalen für die naturwissenschaftliche Kompetenz von Vorschulkindern mit und ohne Migrationshintergrund. *Frühe Bildung*, 8(1), 3–12. <https://doi.org/10.1026/2191-9186/a000405>
- Haschke, L., Kampa, N., Hahn, I. & Köller, O. (2017). Setting standards to a scientific literacy test for adults using the Item-Descriptor (ID) Matching Method. In S. Blömeke & J.-E. Gustafsson (Eds.), *Methodology of Educational Measurement and Assessment (MEMA). Standard setting in education. The Nordic Countries in an international perspective* (pp. 319–339). Cham, Switzerland: Springer International Publishing.
- Haschke, L.I., Hahn, I., Bernholt, S. & Ihme, J.M. (2017). Zur Dimensionalität des PISA-Naturwissenschaftstests am Beispiel naturwissenschaftlicher Grundbildung und Fachkompetenzen. *Unterrichtswissenschaft*, 45(3), 254-268.
- Kähler, J., Hahn, I. & Köller, O. (2020). The development of early scientific literacy gaps in kindergarten children. *International Journal of Science Education*, Open Access. <https://doi.org/10.1080/09500693.2020.1808908>
- Kähler, J., Hahn, I., Ihme, J. M. & Köller, O. (2020). Naturwissenschaftliche Kompetenz von Vorschulkindern. Effekte von Struktur- und Prozessmerkmalen des Elternhauses und der Kindertagesstätte auf die naturwissenschaftliche Kompetenz von 4- bis 6-Jährigen. *Psychologie in Erziehung und Unterricht*, Open Access. <https://doi.org/10.2378/peu2020.art29d>
- Van den Ham, A.-K., Ehmke, T., Hahn, I., Wagner, H. & Schöps, K. (2016). Mathematische und naturwissenschaftliche Kompetenz in PISA, im IQB-Ländervergleich und in der National Education Panel Study (NEPS): Vergleich der Rahmenkonzepte und der dimensional Struktur der Testinstrumente. In Bundesministerium für Bildung und Forschung (Ed.), *Bildungsforschung: Vol. 44. Forschungsvorhaben in Anknüpfung an Large-Scale-Assessments* (pp. 140–160). Bielefeld, Deutschland: W. Bertelsmann Verlag.
- Wagner, H., Schöps, K., Hahn, I., Pietsch, M. & Köller, O. (2014). Konzeptionelle Äquivalenz von Kompetenzmessungen in den Naturwissenschaften zwischen NEPS, LV und PISA. *Unterrichtswissenschaft*, 42(4), 301–320. <https://doi.org/10.3262/UW1404301>
- Wagner, H., Hahn, I., Schöps, K., Ihme, J.M., Köller, O. (2018). Are the tests scores of the Programme for International Student Assessment (PISA) and the National Educational Panel Study (NEPS) science tests comparable? An assessment of test equivalence in

German Schools. *Studies in Educational Evaluation*, 59, 278-287. <https://doi.org/10.1016/j.stueduc.2018.09.002>

Wagner, H., Hahn, I., Schöps, K., Mahler, N. & Köller, O. (2019). Vergleichbarkeit der naturwissenschaftlichen Kompetenz in der neunten Klasse im Nationalen Bildungspanel und im IQB-Ländervergleich 2012. *Zeitschrift für Erziehungswissenschaft*, 22(4), 879–898. <https://doi.org/10.1007/s11618-019-00894-0>

## Literatur

American Association for the Advancement of Science (AAAS). (2009). *Benchmarks for science literacy: Project 2061*. New York: Oxford University Press. Retrieved from <http://www.project2061.org/publications/bsl/online/index.php>

Bybee, R. W. (1997a). *Achieving scientific literacy: From purposes to practices*. Portsmouth, NH: Heinemann Educational Books.

Bybee, R. W. (1997b). Towards an understanding of scientific literacy. In W. Gräber & C. Bolte (Ed.), *Scientific literacy – An international symposium* (pp. 37–68). Kiel.

Bybee, R.W. & PISA 2006 Science Expert Group. (2009). An assessment framework for scientific literacy. In R. W. Bybee & B.J. McCrae (Ed.), *PISA Science 2006. Implications for Science teachers and teaching* (pp. 15–26). Arlington/USA: NSTA press.

Ehmke, T. & Siegle, T (2008). Einfluss elterlicher Mathematikkompetenz und familialer Prozesse auf den Kompetenzerwerb von Kindern in Mathematik. *Psychologie in Erziehung und Unterricht*, 55, 253–264.

Hahn, I., Schöps, K., Rönnebeck, S., Martensen, M., Hansen, S., Saß, S., . . . Prenzel, M. (2013). Assessing scientific literacy over the lifespan – A description of the NEPS science framework and the test development. *Journal for Educational Research Online*, 5(2), 110–138.

Hahn, I. & Schöps, K. (2019). Bildungsunterschiede von Anfang an? Die Bedeutung von Struktur- und Prozessmerkmalen für die naturwissenschaftliche Kompetenz von Vorschulkindern mit und ohne Migrationshintergrund. *Frühe Bildung*, 8(1), 3–12. <https://doi.org/10.1026/2191-9186/a000405>

Hanushek, E. A. & Wößmann, L. (2015). *The knowledge capital of nations: education and the economics of growth*. Cambridge, MA: MIT Press.

Haschke, L.I., Hahn, I., Bernholt, S. & Köller, O. (in prep.). Scientific Literacy of German Adults.

Huynh, H., Michaels, H. R. & Ferrara, S. (1995). *Comparison of three statistical procedures to identify clusters of items with local dependancy*. San Francisco, CA.: Annual meeting of the National Council on Measurement in Education.

Kähler, J., Hahn, I., Ihme, J. M. & Köller, O. (2020). Naturwissenschaftliche Kompetenz von Vorschulkindern. Effekte von Struktur- und Prozessmerkmalen des Elternhauses und der Kindertagesstätte auf die naturwissenschaftliche Kompetenz von 4- bis 6-Jährigen. *Psychologie in Erziehung und Unterricht*, Open Access. <https://doi.org/10.2378/peu2020.art29d>

Kähler, J., Hahn, I., & Köller, O. (2020). The development of early scientific literacy gaps in kindergarten children. *International Journal of Science Education*, Open Access. <https://doi.org/10.1080/09500693.2020.1808908>

- Kluczniok, K., Lehl, S., Kuger, S. & Roszbach, H. G. (2013). Quality of the home learning environment during preschool age – Domains and contextual conditions. *European Early Childhood Education Research Journal*, 21(3), 420–438.
- KMK. (2005a). *Beschlüsse der Kultusministerkonferenz: Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss*. Beschluss vom 16.12.2004. München: Luchterhand.
- KMK. (2005b). *Beschlüsse der Kultusministerkonferenz: Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss*. Beschluss vom 16.12.2004. München: Luchterhand.
- KMK. (2005c). *Beschlüsse der Kultusministerkonferenz: Bildungsstandards im Fach Physik für den Mittleren Schulabschluss*. Beschluss vom 16.12.2004. München: Luchterhand.
- Kotzerke, M., Ebert, S. & Weinert, S. (2014). Wieso, weshalb, warum? Die Entwicklung des Grammatikverständnisses von der 1. bis zur 3. Klasse. In M. Mudiappa & C. Artelt (Eds.), *BiKS - Ergebnisse aus den Längsschnittstudien. Praxisrelevante Befunde aus dem Primar- und Sekundarschulbereich* (Vol. 15, pp. 73–86). Bamberg: University of Bamberg Press Bamberg.
- Martin, M. O. & Mullis, I. (Eds.). (2013). *TIMSS and PIRLS 2011: Relationships among reading, mathematics and science achievement at the fourth grade—Implications for early learning*. Chestnut Hill, MA.
- OECD. (2006). *Assessing scientific, reading and mathematical literacy: A framework for PISA 2006*. Paris: OECD.
- Prenzel, C. Artelt, J. Baumert, W. Blum, M. Hammann, E. Klieme & R. Pekrun (Ed.). (2007). *PISA 2006. Die Ergebnisse der dritten internationalen Vergleichsstudie*.
- Schöps, K. (2013). *NEPS technical report for science: Scaling results of starting cohort 2 in kindergarten* (NEPS working paper no. 24). Bamberg: University of Bamberg, National Educational Panel Study.
- Schöps, K. & Saß, S. (2013). *NEPS technical report for science: Scaling results of starting cohort 4 in ninth grade* (NEPS working paper no. 23).
- Sodian, B. (2002). Entwicklung begrifflichen Wissens. In R. Oerter & L. Montada (Hrsg.), *Entwicklungspsychologie* (S. 443 – 468). Weinheim: Beltz / PVU.
- Wagner, H., Schöps, K., Hahn, I., Pietsch, M. & Köller, O. (2014). Konzeptionelle Äquivalenz von Kompetenzmessungen in den Naturwissenschaften zwischen NEPS, LV und PISA. *Unterrichtswissenschaft*, 42(4), 301–320. <https://doi.org/10.3262/UW1404301>
- Wagner, H., Hahn, I., Schöps, K., Mahler, N. & Köller, O. (2019). Vergleichbarkeit der naturwissenschaftlichen Kompetenz in der neunten Klasse im Nationalen Bildungspanel und im IQB-Ländervergleich 2012. *Erziehungswissenschaft*, 22, 4, S. 878-898.
- Weinert, F. E. (2001). Concept of competence: A conceptual clarification. In S. Rychen, D. Salganik & L. Hersh (Eds.), *Defining and selecting key competencies* (pp. 54–65). Seattle: Hogrefe & Huber.
- Weinert, S. (2004). Wortschatzerwerb und kognitive Entwicklung. The development of vocabulary and cognitive development. *Sprache Stimme Gehör*, 28 (1), 20 – 28.
- Wilhelm O., Schroeders U. & Schipolowski S. (2013). *BEFKI: Berliner Test zur Erfassung fluider und kristalliner Intelligenz [Berlin test of fluid and crystallized intelligence]*. Göttingen: Hogrefe.
- Zabal, A., Martin, S., Klaukien, A., Rammstedt, B., Baumert, J. & Klieme, E. (2013). Grundlegende Kompetenzen der erwachsenen Bevölkerung in Deutschland im internationalen Vergleich. In B. Rammstedt (Ed.), *Grundlegende Kompetenzen Erwachsener im internationalen Vergleich: Ergebnisse von PIAAC 2012* (pp. 31–76). Münster: Waxmann.

## 5. Hörverstehen: Rezeptiver Wortschatz

Tabea Durda

### 5.1 Einleitung

Das Hörverstehen wird im NEPS auf Wort-, Satz- und Text-/Diskursebene erfasst. Da die Instrumente zum Hörverstehen auf Satzebene (rezeptive Grammatik) nicht in der Verantwortlichkeit der Kompetenzsäule liegen, wird an dieser Stelle auf dieses Konstrukt nicht weiter eingegangen. Im Folgenden wird daher lediglich die Erfassung des Hörverstehens auf Wortebene (rezeptiver Wortschatz) im Rahmen des NEPS thematisiert. Informationen zur Erfassung des Hörverstehens auf Text-/Diskursebene, welches bislang nur einmal in Klasse 9 (2015) eingesetzt wurde, sind unter [www.neps-data.de](http://www.neps-data.de) zu finden.

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), 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 kumulativ erworbener sprachlich-kognitiven Fähigkeiten erhoben.

In Deutschland gibt es bisher nur wenige Studien, welche die längsschnittliche Entwicklung des rezeptiven Wortschatzes analysieren, wie beispielsweise die ECCE-Studie und die BiKS-Studie (ECCE-Study Group, 1997; Weinert et al., 2012). Diese und weitere längsschnittliche Studien decken oft nur eine begrenzte Lebensspanne (zumeist Vor- und Schulalter) ab und/oder basieren auf eher selektiven Stichproben. Das NEPS bietet hingegen die Möglichkeit, den Wortschatzerwerb nicht nur im Kindergarten- und Grundschulalter zu analysieren, sondern auch in älteren Kohorten und dabei vielfältige Bedingungsfaktoren einzubeziehen.

Das international zur Erfassung des rezeptiven Wortschatzes am häufigsten eingesetzte Instrument ist sicher der inzwischen in verschiedenen Versionen vorliegende *Peabody Picture Vocabulary Test (PPVT)* (Dunn, 1959; Dunn & Dunn, 1981, 1997, 2004, 2007; Lenhard et al., 2015). Der PPVT ist grundsätzlich über einen sehr großen Altersbereich hinweg einsetzbar und zugleich einfach in der Durchführung und Auswertung. Eine deutsche Version des PPVT, die für Kinder ab 13 Jahren und für Erwachsene normiert ist, ist im Jahr 2004 erschienen (Dunn & Dunn, 2004) und beruht auf dem PPVT-III (Dunn & Dunn, 1997). Da 2011 eine publizierte deutsche Version des PPVT nur für ältere Kinder ab einem Alter von 13 Jahren vorlag (Dunn & Dunn, 2004), wurde für die Startkohorte 2 (Kindergarten) ein dem PPVT analoges Verfahren auf der Basis von Daten der BiKS-Studie erstellt. Im Rahmen der BiKS-Studie (Homuth et al., 2014) wurde im Längsschnitt BiKS-3-10 eine – auf der Basis von Daten der ECCE-Studie (European Child Care and Education (ECCE)- Study Group, 1997) – erstellte deutsche Forschungsversion des PPVT eingesetzt. In Startkohorte 1 wird eine deutschsprachige Forschungsversion des PPVT-IV eingesetzt (Dunn & Dunn, 2007; deutsche Bearbeitung von Lenhard et al., 2015).

## 5.2 Bisherige Erhebung des Konstrukts

Bisher wurde der rezeptive Wortschatz in fünf Altersgruppen eingesetzt.

Tabelle 12: Erhebungszeiten und -modi der für den Wortschatz (PPVT) gelaufenen bzw. aktuell laufenden Erhebungen

Startkohorte	Alter / Klasse	Erhebungsjahr	Erhebungszeit (in Minuten)	Erhebungsmodus
SC 1	3 Jahre	2015	15	Computer (Tablets)
SC 1	5 Jahre	2017	15	Computer (Tablets)
SC 1*	7 Jahre	2019	15	Computer (Tablets)
SC 2	5 Jahre	2011	20	Papier
SC 2	Klasse 1	2013	20	Papier
SC 2	Klasse 3	2014	20	Papier
SC 3	Klasse 6	2011	20	Papier
SC 4	Klasse 9	2010	20	Papier
SC 6	Erwachsene	2014	20	Computer

Anmerkung. \*Die Daten zu diesen Studien sind noch nicht im SUF verfügbar; die Daten werden deshalb in diesem Dokument noch nicht berichtet.

## 5.3 Beispielitems

Aufgabe der Zielperson ist es, zu jedem einzelnen vorgegebenen Wort aus jeweils vier Bildern das zu dem Wort gehörige Bild auszuwählen.

Der Test wird in der SC 2 im Kindergartenalter in einer spielerisch gestalteten Einzeltestsituation im Kindergarten durchgeführt. Um die Kinder bei schwächeren Leistungen nicht zu stark zu belasten, wird der Test nach sechs falsch beantworteten Items in Folge abgebrochen. In der SC 1 wird der PPVT in einer Einzeltestsituation im Haushalt der Zielfamilien durchgeführt. Dazu wird eine tabletbasierte Version eingesetzt. Durchführung, Boden- und Deckensetbestimmung folgen dabei den Vorgaben des PPVT-IV. Im Schulalter wird der Test im Klassenkontext durchgeführt, ein Abbruchkriterium gibt es hier nicht.

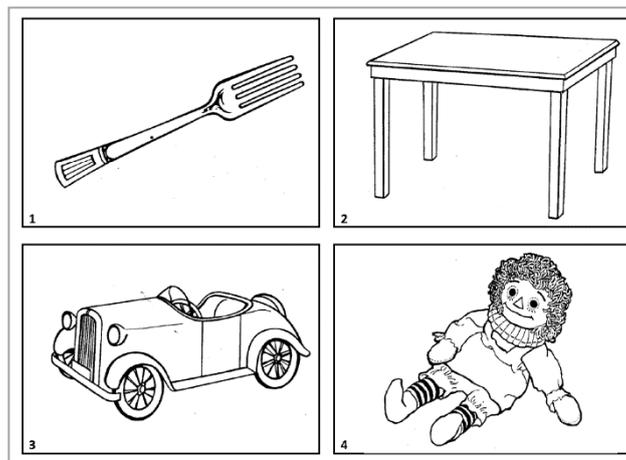


Abbildung 14. Beispielim "Puppe" in der SC 2, Kindergarten im Alter von ca. 5 Jahren.

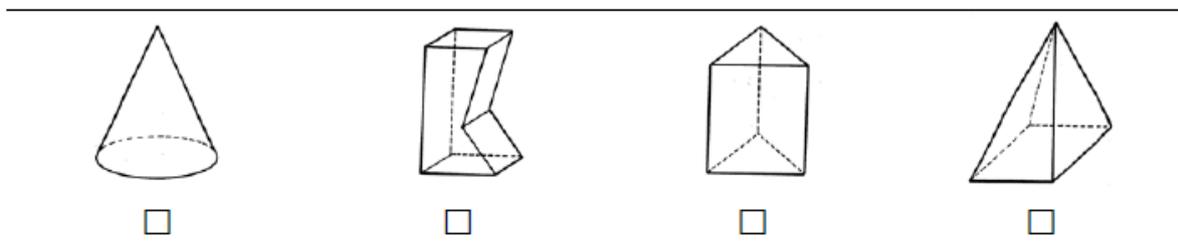


Abbildung 15. Beispielim "Pyramide" in der SC 3, Klasse 6.

## 5.4 Empirische Analysen

### 5.4.1 Deskriptive Informationen

In Tabelle 13 sind die deskriptiven Angaben für die bisher in den Scientific Use Files bereitgestellten Datensätze dargestellt. Ein erfreuliches Ergebnis dieser Daten ist die sehr geringe Ausfallquote innerhalb der Testdurchführungen in den jeweiligen Startkohorten, d.h. Personen, die den Wortschatztest begonnen haben, haben die Testung auch mit validen Angaben abgeschlossen, aus welchen Kompetenzwerte berechnet werden konnten. Außerdem zeigen sich durchweg gute Reliabilitäten der Skalen, die in den verschiedenen Altersbereichen eingesetzt wurden.

Tabelle 13: Deskriptive Informationen zu Erhebungen im Wortschatz (PPVT)

Startkohorte	Alter / Klasse	N (gesamt)	N (gültig)	% fehlend	M	SD	Reliabilität
SC 1	3 Jahre	2324	1843	20.70	46.84	28.16	0.97*
SC 1	5 Jahre	2080	2063	0.82	81.50	23.27	0.97*
SC 2	5 Jahre	4876	2886	1.27	46.86	14.73	0.81
SC 2	Klasse 1	4876	6477	5.25	39.85	10.18	0.86
SC 2	Klasse 3	4876	5603	3.86	45.18	9.47	0.85
SC 3	Klasse 6	4876	4876	0.00	52.10	10.30	0.88
SC 4	Klasse 9	15017	14500	3.44	57.00	10.67	0.86
SC 6	Erwachsene	8739	8731	0.09	73.63	9.66	0.87

Bemerkung. N (gesamt) = Anzahl der Personen, bei denen der Test administriert wurde; N (gültig) = Anzahl der Personen mit gültigem Testwert, % fehlend = Anteil an Personen ohne gültigen Testwert, M = Mittelwert des Testwerts, SD = Standardabweichung des Testwerts, Reliabilität = Cronbachs Alpha. Die Mittelwerte der Testwerte sind nur innerhalb der SC 1 miteinander vergleichbar, da ansonsten unterschiedliche Items mit jeweils einer unterschiedlichen Gesamtanzahl an Items in den Versionen eingesetzt wurden. \*gemäß Manual des PPVT-4 (Lenhard et al., 2015)

#### 5.4.2 Psychometrische Informationen

Mittlerweile liegen für den PPVT für SC 2 umfassende psychometrische Analysen vor, für die SC 3 bis SC 6 ist jeweils der Summenscore angegeben. Die Analysen für die SC 2 zeigen, dass der jeweilig eingesetzte PPVT-Test eine akzeptable Zuverlässigkeit und Modellanpassung aufweist. Weitere Informationen finden sich in Fischer und Durda (2020). Weiterhin wurden zur Bestimmung der konvergenten Validität die Ergebnisse mit anderen Konstrukten verglichen (vgl. Tabelle 14). Für die Qualität des PPVT sprechen die erwartungskonformen und stabilen Zusammenhänge mit inhaltlich ähnlichen Konstrukten. Zu *Beginn der ersten Klasse* beispielsweise erzielen Kinder, die einen höheren rezeptiven Wortschatz im Kindergartenalter und in der ersten Klasse aufweisen, auch höhere Kompetenzwerte in ihrem rezeptiven Grammatikverständnis. Kinder, Schülerinnen und Schüler sowie Erwachsene mit höheren Werten im PPVT schneiden ebenfalls besser in den Lesetests ab.

Tabelle 14: Korrelation des PPVT mit ähnlichen Konstrukten

	PPVT SC 2 5 J.	PPVT SC 2 K1	PPVT SC 2 K3	PPVT SC 3	PPVT SC 4	PPVT SC 6
<b>PPVT K1</b>	.66**					
<b>PPVT K3</b>	.65**	.72**				
<b>Grammatik 5 J.</b>	.73**	.56**	.50**			
<b>Grammatik K1</b>	.53**	.63**	.60**			
<b>Buchstaben- kenntnis</b>	.20**	.15**	.20**			
<b>Reime</b>	.44**	.30**	.31**			
<b>Identifikation von Phonemen</b>	.40**	.32**	.31**			
<b>Lese- geschwindig- keit</b>	.21**	.24**	.24**	.27**	.32**	.43** (W3) .49** (W5)
<b>Lese- kompetenz</b>	.23** (ELFE)	.33** (ELFE)	.35** (ELFE)	.55** (K5) .52** (K7)	.57** (K9) .50** (W10)	.52** (W3) .53** (W5)
	.44** (K4)	.45** (K4)	.45** (K4)	.52** (K9) .44** (K12)		.52** (W9)
<b>Orthographie (Set 1)</b>	.26**	.30**	.30**	.35** (K5) .36** (K7) .34** (K9)		
<b>Orthographie (Set 2)</b>	.26**	.31**	.31**	.37** (K5) .37** (K7) .33** (K9)		
<b>Hörverstehen</b>				.51**		

Bemerkung. \*\* Die Korrelation ist auf dem Niveau von 0.01 (zweiseitig) signifikant.

### 5.4.3 Gruppenunterschiede

Im Folgenden sollen einige zentrale deskriptive Angaben der einzelnen Kohorten dargestellt werden. Die Mittelwertsunterschiede zwischen den Geschlechtern, sowie Personen mit und ohne Migrationshintergrund, zwischen Schulabschlüssen und Schultypen sowie zwischen Personen mit hohem und niedrigem Sozialstatus sind in Tabelle 15 dargestellt. Anzumerken ist zunächst, dass entgegen der Erwartung aufgrund bisheriger empirischer Forschungsarbeiten in allen Altersgruppen keine Geschlechtsunterschiede zugunsten der Mädchen in dieser Domäne vorliegen. In der SC 2 fallen die Kompetenzwerte für Jungen und Mädchen dabei nahezu gleich aus. In der SC 3, SC 4 und SC 6 fallen die Kompetenzwerte im PPVT für Jungen bzw. Männer sogar bedeutsam höher aus als für Mädchen bzw. Frauen. Erwartungskonform fallen hingegen die Unterschiede von Personen mit hohem vs. niedrigem Sozialstatus aus. Personen mit hohem Sozialstatus weisen höhere Kompetenzen im rezeptiven Wortschatz aus. Ein weiteres konsistentes Ergebnis der NEPS Kohorten ist, dass Personen ohne Migrationshintergrund deutlich höhere Kompetenzwerte im PPVT in allen Altersbereichen im Vergleich zu Personen mit Migrationshintergrund aufweisen. Beim Vergleich verschiedener Schulabschlüsse und Schultypen zeigt sich erwartungskonform, dass Personen, die eine höhere Schulform besuchen und Personen, die selbst oder deren Eltern über einen höheren Schulabschluss verfügen, bessere Werte im Wortschatztest erzielen.

Tabelle 15: Standardisierte Mittelwertsunterschiede (Cohen's d) der Testwerte für zentrale Gruppierungsvariablen in den unterschiedlichen Startkohorten und Altersgruppen (rezeptiver Wortschatz)

Startkohorte	Alter / Klasse	Geschlecht	Sozio- ökonomischer Status	Migrations- hintergrund	Schultyp	Schulab- schluss
			(weiblich vs. männlich)	(hoch vs. gering)	(ohne vs. mit)	(Gymnasium vs. Nicht- Gymnasium)
SC 1	3 Jahre	0.05	0.40	0.63	-	0.46
SC 1	5 Jahre	-0.19	0.57	-1.01	-	0.76
SC 2	5 Jahre	-0.02	0.55	0.76	-	0.55
SC 2	Klasse 1	-0.08	0.49	0.58	-	0.54
SC 2	Klasse 3	-0.15	0.48	0.53	-	0.50
SC 3	Klasse 6	-0.27	0.52	0.95	0.70	0.45
SC 4	Klasse 9	-0.26	0.64	0.80	0.59	1.09
SC 6	Erwachsene	-0.18	0.67	0.70	-	0.78

Bemerkung. Positive Werte spiegeln einen höheren Mittelwert in der ersten Gruppe (z.B. „weiblich“) wider, während negative Werte einen höheren Mittelwert in der zweiten Gruppe (z.B. „männlich“) widerspiegeln. Sozioökonomischer Status = HISEI der Eltern bzw. eigener ISEI in SC6, Migrationshintergrund = ohne (= beide Elternteile in Deutschland geboren) vs. mit (= mind. ein Elternteil nicht Deutschland geboren), Schultyp = Schultyp, den der / die Schüler/in besucht, Schulabschluss = Höchster Schulabschluss der Eltern bei SC 1- SC 4, eigener höchster Schulabschluss in SC 6

## 5.5 Ausgewählte Publikationen mit NEPS-Daten zum rezeptiven Wortschatz

- Barthel, A. (2019). Soziale Ungleichheit in der Wortschatzentwicklung von der ersten zur dritten Jahrgangsstufe. *Zeitschrift für Grundschulforschung*, 12(1), 213–228. <https://doi.org/10.1007/s42278-019-00041-y>
- Gnambs, T. (2017). Human Capital and Reemployment Success: The role of cognitive abilities and personality. *Journal of Intelligence*, 5(1), 9. <https://doi.org/10.3390/jintelligence5010009>
- Gnambs, T. & Appel, M. (2017). Is computer gaming associated with cognitive abilities? A population study among German adolescents. *Intelligence*, 61, 18-28. <https://doi.org/10.1016/j.intell.2016.12.004>
- Kluczniok, K. & Mudiappa, M. (2019). Relations between socio-economic risk factors, home learning environment and children's language competencies: Findings from a German study. *European Educational Research Journal*, 18(1), 85–104. <https://doi.org/10.3389/fpsyg.2020.01425>
- Linberg, T., Baeumer, T. & Rossbach, H.-G. (2013). Data on early child education and care learning environments in Germany. *International Journal of Child Care and Education Policy*, 7, 24-42.
- Linberg, A., Lehrl, S. & Weinert, S. (2020). The early years home learning environment: Associations with parent-child-course attendance and children's vocabulary at age 3. *Frontiers in Psychology*, 11, Article 1425. <https://doi.org/10.3389/fpsyg.2020.01425>
- Relikowski, I., Schneider, T. & Linberg, T. (2015). Rezeptive Wortschatz- und Grammatikkompetenzen von Fünfjährigen mit und ohne Migrationshintergrund: Eine empirische Untersuchung aus bildungssoziologischer Perspektive. *Frühe Bildung*, 4(3), 135–143. <https://doi.org/10.1026/2191-9186/a000218>
- Strobel, B. & Seuring, J. (2016). Spracherhalt oder Sprachverlagerung? *KZfSS Kölner Zeitschrift Für Soziologie Und Sozialpsychologie*, 68(2), 309–339. <https://doi.org/10.1007/s11577-016-0360-y>

## 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.
- 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. (2004). *Peabody Picture Vocabulary Test (PPVT) (deutsche Version)*. Göttingen: Hogrefe.
- Dunn, L. M. & Dunn, L. M. (2007). *Peabody Picture Vocabulary Test. Fourth edition (PPVT-4)*. Upper Saddle River, NJ: Pearson.
- Fischer, L. & Durda, T. (2020). *NEPS technical report for receptive vocabulary: Scaling results of starting cohort 2 for kindergarten (wave 1), grade 1 (wave 3) and grade 3 (wave 5) (NEPS Survey Paper No. 65)*. Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- European Child Care and Education (ECCE)-Study Group (1997). *European childcare and education study. Cross national analyses of the quality and effects of early childhood programmes on children's development*. Berlin: Freie Universität Berlin. Fachbereich Erziehungswissenschaft. Psychologie und Sportwissenschaft. Institut für Sozial- und Kleinkindpädagogik.
- Homuth, C., Mann, D., Schmitt, M. & Mudiappa, M. (2014). Eine Forschergruppe, zwei Studien: BiKS-3-10 und BiKS-8-14. In M. Mudiappa & C. Artelt (Hg.). *BiKS – Ergebnisse aus den Längsschnittstudien*. Bamberg: University of Bamberg Press.
- Lenhard, A., Lenhard, W., Segerer, R. & Suggate, S. (2015). *Peabody Picture Vocabulary Test - Revision IV (Deutsche Adaption)*. Frankfurt a. M.: Pearson Assessment.
- Lipps, G. & Yiptong-Avila, J. (1999). *From Home to School – How Canadian Children Cope*. Ottawa: Statistics Canada.
- Weinert, S., Ebert, S., Lockl, K. & Kuger, S. (2012). Disparitäten im Wortschatzerwerb: Zum Einfluss des Arbeitsgedächtnisses und der Anregungsqualität in Kindergarten und Familie auf den Erwerb lexikalischen Wissens. *Unterrichtswissenschaft, 40*, 4-25.

## 6. ICT Literacy

Martin Senkbeil & Jan Marten Ihme

### 6.1 Einleitung

Ein kompetenter Umgang mit Informations- und Kommunikationstechnologien (Information and Communication Technologies, ICT Literacy) ist angesichts der fortschreitenden Digitalisierung in allen Lebensbereichen mittlerweile unverzichtbar und stellt eine Schlüssel- oder Metakompetenz für lebenslanges Lernen dar (ETS, 2002; Fraillon, Ainley, Schulz, Duckworth & Friedman, 2019). Die Fertigkeit, über digitale Medien vermittelte Informationen zu verstehen, zu nutzen und zu kommunizieren, spielt sowohl in vielen Berufsfeldern als auch bei der Bewältigung alltäglicher Probleme (z.B. internetgestützte Suche von Informationen) und der zwischenmenschlichen Kommunikation (z.B. über E-Mail oder soziale Netzwerke) eine zunehmend wichtige Rolle (Ferrari, 2013; Sefton-Green, Nixon & Erstad, 2009). Neuere Konzeptionen der ICT Literacy betonen dabei funktionale, d. h. von den Anforderungen der Lebens- und Arbeitswelt ausgehende Wissensbestände und Fertigkeiten, die neben technologischen Fertigkeiten (Funktionswissen über Programmanwendungen) vornehmlich Aspekte der Informationskompetenz (Fähigkeit, mit Hilfe digitaler Medien Informationen zu ermitteln, kritisch auszuwählen und sie effektiv für sich zu nutzen) umfassen (z.B. ETS, 2002; OECD, 2013). Diese Konzeptualisierung liegt auch den in NEPS eingesetzten Tests zur Erfassung von ICT Literacy zugrunde (siehe z.B. Senkbeil, Ihme & Wittwer, 2013a, b für eine ausführliche Darstellung der Rahmenkonzeption).

Da angesichts der zunehmend erforderlichen Flexibilisierung in Beruf und Gesellschaft große Bereiche des Wissens über die gesamte Lebensspanne weitgehend selbstgesteuert und vornehmlich über digitale Medien anzueignen sind (van Laar, van Deursen, van Dijk & de Haan, 2017), stellt ICT Literacy eine von acht Schlüsselkompetenzen dar, welche die Grundlage für ein lebenslanges Lernen bilden (Ferrari, 2012). ICT Literacy ist daher als wichtiges Bildungsziel im Sinne einer Kulturtechnik zu betrachten, deren Erwerb eine wesentliche Voraussetzung für die erfolgreiche Teilhabe in der Gesellschaft (z.B. Erfüllung persönlicher, beruflicher, sozialer und politischer Zielvorstellungen) darstellt (Blossfeld, 2010; KMK, 2016). Gemäß ihrem Stellenwert als gesellschaftlicher Schlüsselkompetenz wurde ICT Literacy in jüngster Zeit in internationalen Large-Scale-Assessments (LSA) wie ICILS (*International Computer and Information Literacy Study*) 2013/ 2018 oder PIAAC (*Programme for the International Assessment of Adult Competencies*) als eigenständige Metakompetenz erfasst (vgl. Senkbeil, Goldhammer, Bos, Eickelmann, Schwippert & Gerick, 2014). Zentrale Befunde dieser LSAs sind, dass ein Drittel der Schülerinnen und Schüler in Deutschland (33.2%) lediglich über rudimentäre Wissensbestände und Fertigkeiten im Umgang mit digitalen Medien verfügen (Eickelmann, Bos, Gerick & Labusch, 2019), sich erhebliche soziale Disparitäten in der ICT Literacy bereits in der Sekundarstufe I (Klasse 8) zeigen (Senkbeil, Drossel, Eickelmann & Vennemann, 2019) und Defizite in der ICT Literacy mit einem vergleichsweise geringem Einkommen im Erwachsenenalter assoziiert sind, wie anhand der Daten der PIAAC-Studie gezeigt werden konnte (Hanushek, Schwerdt, Wiederhold & Wößmann, 2015). Aufgrund der querschnittlich angelegten Untersuchungsdesigns von LSAs fehlt es jedoch bislang an empirisch belastbaren Befunden, z.B. zur oftmals herangezogenen Digital divide-Hypothese, wonach Unterschiede in der ICT Literacy über die Lebensspanne in Abhängigkeit von soziodemographischen Merkmalen (z.B. soziale Herkunft) kontinuierlich zunehmen und zu

Benachteiligungen in der Bildungsrendite führen (van Deursen & van Dijk, 2010; van Dijk, 2009). Das Nationale Bildungspanel bietet daher erstmals die Möglichkeit, die computer- und informationsbezogene Kompetenzentwicklung über einen längeren Zeitraum, d.h. über mehrere Bildungsetappen zu analysieren und die genannten Hypothesen anhand angemessener Untersuchungsdesigns empirisch zu prüfen.

## 6.2 Bisherige Erhebung des Konstrukts

Tabelle 16 informiert über die Erhebungszeiten und –modi bezüglich der bislang durchgeführten Erhebungen in NEPS für die Domäne ICT Literacy.

Tabelle 16: Erhebungszeiten und -modi der in der Domäne ICT Literacy gelaufenen bzw. aktuell laufenden Erhebungen

Startkohorte	Alter / Klasse	Erhebungsjahr	Erhebungszeit (in Minuten)	Erhebungsmodus
SC 2	Klasse 3	2014	28	Papier
SC 3	Klasse 6	2011	29	Papier
SC 3	Klasse 9	2014	28	Papier
SC 3	Klasse 12 / Schulabgänger	2017	28	Papier / Computer
SC 4	Klasse 9	2010	29	Papier
SC 4	Klasse 12 / Schulabgänger	2013	28	Papier
SC 5	Studierende	2013	28	Papier /Computer /Online
SC 6	Erwachsene	2012	28	Papier

## 6.3 Beispielitems

Tabelle 17 gibt einen Überblick über die nachfolgend dargestellten Beispielitems. Die Tabelle gibt für jedes Beispielitem die Alters- bzw. Klassenstufe an, in der es eingesetzt wurde, sowie den Inhaltsbereich und die Prozesskomponente aus der Rahmenkonzeption zu ICT Literacy (vgl. Senkbeil, Ihme & Wittwer, 2013a, b). ICT Literacy wurde bis 2017, wie die Beispielitems zeigen, anhand von Multiple Choice (MC)-Aufgaben erfasst. Anhand von MC-Aufgaben können deklarative und prozedurale Wissensbestände in Form von deklarativem Fakten- und Konzeptwissen sowie prozeduralem, anwendungsorientiertem Wissen erfasst werden. Die MC-Aufgaben beschreiben realistische Problemstellungen in einer Fülle authentischer Situationen

und erfassen, ob die Testpersonen angemessen mit bestimmten Aufgabenstellungen umgehen können, indem sie gefragt werden, was sie in der betreffenden Situation tun würden. Für eine möglichst realitätsnahe Gestaltung werden in den Aufgabenstimulus Screenshots integriert (z.B. von einem Internet-Browser oder Tabellenkalkulationen). Häufig werden als Distraktoren realistische Antwortalternativen in Form von Schaltflächen oder Menüs vorgegeben, die in die jeweiligen Screenshots integriert sind, oder es werden reale Programmanwendungen herangezogen und für die Konstruktion der Antwortmöglichkeiten verwendet (Senkbeil & Ihme, 2014). Seit 2018 werden zur Erfassung von ICT Literacy zusätzlich simulationsbasierte, d.h. computerisierte und interaktive Testaufgaben eingesetzt, die die internationale Anschlussfähigkeit der Kompetenzmessungen (Siddiq, Hatlevik, Olsen, Throndsen & Scherer, 2016) sowie eine verbesserte Konstruktrepräsentation gewährleisten (Senkbeil & Ihme, 2020; siehe auch Abschnitt 6.5).

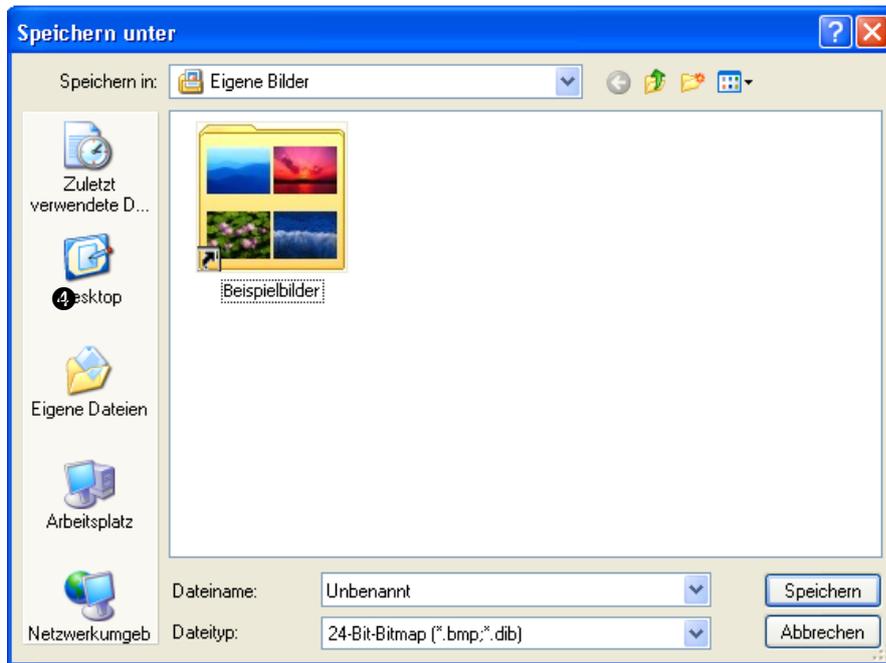
Tabelle 17: Beispielitems für die Domäne ICT Literacy für verschiedene Altersstufen

	<b>Beispielitem</b>	<b>Altersstufe</b>	<b>Inhaltsbereich</b>	<b>Prozesskomponente</b>
<b>1</b>	Suchbegriff für Blinde Kuh	Klasse 3	Internetgestützte Suchmaschinen und Datenbanken	Suchen und Organisieren
<b>2</b>	Datei speichern	Klasse 6	Betriebssystem	Anwenden und Verstehen
<b>3</b>	Tabelle ordnen	Klasse 9	Tabellenkalkulation	Suchen und Organisieren
<b>4</b>	Rückenschmerzen	Klasse 12	Internetgestützte Suchmaschinen und Datenbanken	Bewerten
<b>5</b>	Informationen suchen	Erwachsene	Internetgestützte Suchmaschinen und Datenbanken	Suchen und Organisieren

<p>In der Schule hat euch eure Lehrerin gesagt, ihr sollt mit Hilfe der Kindersuchmaschine „Blinde Kuh“ alles über die Tiere auf dem Bauernhof nachlesen. Über Google wollt ihr auf die Seite „Blinde Kuh“ gehen.</p> <p>Welchen Suchbegriff gebt ihr am besten bei Google ein?</p>	
Kindersuchmaschine	<input type="checkbox"/>
Kuh	<input type="checkbox"/>
Suchen	<input type="checkbox"/>
Bauernhof Kuh	<input type="checkbox"/>

Abbildung 16. Beispielitem („Suchbegriff für Blinde Kuh“) der Domäne ICT Literacy für Klassenstufe 3.

Peter möchte ein Bild in dem Bildbearbeitungsprogramm als TIFF-Datei abspeichern. Wie macht er das?



① Er geht auf „Eigene Dateien“.	<input type="checkbox"/>
② Er sucht unter „Dateityp“.	<input type="checkbox"/>
③ Er geht auf „Speichern“.	<input type="checkbox"/>
④ Er geht auf „Desktop“.	<input type="checkbox"/>

Abbildung 17. Beispielitem („Datei speichern“) der Domäne ICT Literacy für Klassenstufe 6.

Dies ist ein Ausschnitt aus einer Tabelle mit den längsten Flüssen der Erde. Welche Spalten muss Bianca in dieser Tabelle durchsuchen, um am einfachsten den längsten Fluss Asiens herauszufinden?

<b>Flüsse</b>			
<b>Name</b>	<b>Kontinent</b>	<b>Länge [km]</b>	<b>Mündung</b>
Amazonas	Südamerika	6448	Atlantik
Gelber Fluss	Asien	4845	Gelbes Meer
Jangtse	Asien	6380	Ostchinesisches Meer
Jenissei	Asien	5940	Nordpolarmeer
Kongo	Afrika	4835	Atlantik
Mekong	Asien	4500	Südchinesisches Meer
Mississippi	Nordamerika	6051	Golf von Mexico
Nil	Afrika	6671	Mittelmeer
Ob	Asien	5052	Ochotskisches Meer

- Name und Asien
- Länge
- Name und Länge
- Kontinent und Mündung
- Kontinent und Länge

Abbildung 18. Beispielitem („Tabelle ordnen“) der Domäne ICT Literacy für Klassenstufe 9.

Sie haben einen schmerzenden Rücken und wollen sich dazu erst einmal grundlegend informieren. Auf welche Seite sollten Sie gehen, wenn Sie möglichst seriöse und medizinisch abgesicherte Informationen finden wollen?

**[Rückenschmerz - 2x30 min Rückentraining pro Woche](#) ❶**

[www.kieser-training.de/DeinRuecken](http://www.kieser-training.de/DeinRuecken)

Testen Sie uns jetzt kostenlos.

+ [Karte von Ringstrasse 37-39, Kiel einblenden](#)

**[Hilfe bei Rückenschmerzen - Hexenschuss oder Ischias?](#) ❷**

[www.doc-gegen-schmerzen.de/pflichttext](http://www.doc-gegen-schmerzen.de/pflichttext)

Lesen Sie hier Wichtiges über die Ursachen!

**[BZgA: Leitbegriffe der Gesundheitsförderung: Leitbegriffe A - Z ...](#) ❸**

[www.leitbegriffe.bzga.de/bot\\_angebote\\_idx-151.html](http://www.leitbegriffe.bzga.de/bot_angebote_idx-151.html)

Infolgedessen treten insbesondere **Rückenschmerzen**, Kopfschmerzen, allgemeine nervöse Überlastung und Kreislaufbeschwerden auf. Auf der anderen Seite ...

**[Rückenschmerzen Hilfe - Übungen, Symptome, Behandlung für das ...](#) ❹**

[rueckenschmerzenhilfe.com/](http://rueckenschmerzenhilfe.com/)

**Rückenschmerzen:** Was hilft, die besten Übungen, Experten-Meinungen, Symptome. Tipps zum Thema **Rückenschmerzen** und Ursachen finden Sie bei uns!

↳ [Rückenschmerzen im ... - Ischiasnerv - Rückenschmerzen Übungen ...](#)

**[Rückenschmerzen](#) ❺**

[www.heilpraxisnet.de](http://www.heilpraxisnet.de) › Symptome

Ausführliches zu **Rückenschmerzen**, möglichen Ursachen, Hintergrund Erkrankungen und Symptomen. Jeder dritte leidet unter Schmerzen im Rücken.

**[Gegen Rückenschmerzen](#) ❻**

[www.officegym.com/Trainingsgeraete](http://www.officegym.com/Trainingsgeraete)

Das Trainingsgerät für jeden Stuhl!

Jetzt Sport mit Arbeit kombinieren

**[Gegen Rückenschmerzen](#)**

[www.tetesept.de/Rueckenbeschwerden](http://www.tetesept.de/Rueckenbeschwerden)

tetesept Natürlich-wirksame Hilfe.

Gegen Verspannungen und Schmerzen.

❶ Rückenschmerz	<input type="checkbox"/>
❷ Hilfe bei Rückenschmerzen	<input type="checkbox"/>
❸ BZgA: Leitbegriffe der Gesundheitsförderung	<input type="checkbox"/>
❹ Rückenschmerzen Hilfe – Übungen, Symptome,...	<input type="checkbox"/>
❺ Rückenschmerzen	<input type="checkbox"/>
❻ Gegen Rückenschmerzen	<input type="checkbox"/>

Abbildung 19. Beispielitem („Datei speichern“) der Domäne ICT Literacy für Klassenstufe 12.

Um Informationen über den aktuellen amerikanischen Präsidenten zu finden, haben Sie den Suchbegriff Barack Obama eingegeben. Welche Auswahllisten würden Sie benutzen, um Informationen in deutscher Sprache aus Deutschland zu bekommen?

<b>Ergebnisse finden</b>	mit <b>allen</b> Wörtern	Barack Obama	10 Ergebnisse
	mit der <b>genauen Wortgruppe</b>		
	mit <b>irgendeinem</b> der Wörter		
	<b>ohne</b> die Wörter		
<b>Sprache</b>	Antwortseiten, geschrieben in	beliebige Sprache	
<b>Region</b>	Gesuchte Seiten befinden sich in:	alle Regionen	
<b>Dateiformat</b>	Ausschließlich	Ausgabe von Ergebnissen des Dateiformats irgendein Format	
<b>Datum</b>	Suche nach zuerst gesichteten Webseiten	ohne Zeitbegrenzung	
<b>Position</b>	Antwortseiten, in denen meine Begriffe vorkommen	irgendwo auf der Seite	
		richtig	falsch
beliebige Sprache		<input type="checkbox"/>	<input type="checkbox"/>
alle Regionen		<input type="checkbox"/>	<input type="checkbox"/>
irgendein Format		<input type="checkbox"/>	<input type="checkbox"/>
ohne Zeitbegrenzung		<input type="checkbox"/>	<input type="checkbox"/>
irgendwo auf der Seite		<input type="checkbox"/>	<input type="checkbox"/>

Abbildung 20. Beispielitem („Informationen suchen“) der Domäne ICT Literacy für Erwachsene (Erhebungsjahr: 2012).

## 6.4 Empirische Analysen

### 6.4.1 Deskriptive Informationen

Tabelle 18 liefert einen Überblick über zentrale statistische Kennwerte der NEPS-Erhebungen, in denen die Domäne ICT Literacy erhoben wurde. Es zeigt sich, dass die geschätzte Standardabweichung in der Kohorte der Erwachsenen am größten ( $SD = 1.24$ ) und in der Kohorte der Studierenden am geringsten ( $SD = 0.76$ ) ist. Eine eher geringe Reliabilität der Tests ergibt sich Stichproben mit geringer Varianz, d.h. insbesondere in unteren Klassenstufen (Klasse 3) oder bei den Studierenden. Ansonsten beträgt die Reliabilität immer  $\geq .70$ .

Tabelle 18: Deskriptive Informationen zu Erhebungen in der Domäne ICT Literacy

Startkohorte	Alter / Klasse	N (gesamt)	N (gültig)	% fehlend	M	SD	Reliabilität
SC 2	Klasse 3	5620	5619	< 1.00	0.00	0.64	.59
SC 3	Klasse 6	4876	4872	< 1.00	-0.00	0.76	.70
SC 3	Klasse 9	4898	4876	< 1.00	0.79	0.89	.82
SC 3	Klasse 12 / Schulabgänger	3778	3748	< 1.00	1.27	0.81	.65
SC 4	Klasse 9	14540	14312	1.57	-0.01	0.94	.83
SC 4	Klasse 12 / Schulabgänger	5762	5761	< 1.00	1.03	0.80	.73
SC 5	Studierende	8767	8563	2.33	0.06	0.76	.57
SC 6	Erwachsene	6697	6167	7.91	-0.02	1.24	.81

Bemerkung. N (gesamt) = Anzahl der Personen, denen der Test administriert wurde; N (gültig) = Anzahl der Personen mit gültigem Testwert, % fehlend = Anteil an Personen ohne gültigen Testwert, M = Mittelwert des Testwerts (bei erster Administration pro Startkohorte auf 0 fixiert), SD = Standardabweichung des Testwerts, Reliabilität = EAP/PV-Reliabilität.

#### 6.4.2 Psychometrische Informationen

Die psychometrischen Analysen zu den Erhebungen (dokumentiert als Technical Reports der NEPS Working Papers und NEPS Survey Papers: Ihme & Senkbeil, 2018; Senkbeil & Ihme, 2012, 2015, 2017a,b; Senkbeil, Ihme & Adrian, 2013) zeigen, dass die in NEPS eingesetzten ICT Literacy-Tests zwar nicht die restriktiven Kriterien von Pohl und Carstensen (2012) für Eindimensionalität erfüllen, aber angesichts substanzieller Korrelationen der Subdimensionen (zwischen  $r = .70$  und  $r = .97$ ) eine eindimensionale Skalierung nach dem Raschmodell und die Annahme einer Globalskala rechtfertigen (vgl. hierzu auch Ihme, Senkbeil, Goldhammer & Gerick, 2017).

Eine umfassende Validierung der in NEPS eingesetzten ICT Literacy-Tests war nicht Teil der Instrumentenentwicklung und wurde anhand externer Studien durchgeführt. Dabei wurde der Test der Klassenstufe 9 hinsichtlich verschiedener Validitätsaspekte vergleichsweise ausführlich untersucht (Senkbeil, Ihme & Wittwer, 2013b; Senkbeil & Ihme, 2014, 2020). Für den Test der Klassenstufe 9 konnte empirisch gestützt werden, dass ICT Literacy ein eigenständiges Konstrukt im Sinne einer *new literacy* darstellt und sich von kognitiven Fähigkeiten abgrenzen lässt ( $r = .69$ ; Senkbeil & Ihme, 2014, 2020). Darüber hinaus stützen die theoretisch erwarteten konvergenten (z.B. instrumentelle Nutzungsmotivation, Selbstwirksamkeit) und

diskriminanten Zusammenhänge (z.B. hedonistische Nutzungsmotivation, Erfahrung im Umgang mit digitalen Medien) mit Merkmalen der Computervertrautheit die Gültigkeit der Konstruktinterpretation des Instruments (Senkbeil & Ihme, 2020; Senkbeil, Ihme & Wittwer, 2013a,b). Unterstützende Validitätsbelege konnten auch für die Testinstrumente der Schulabsolventinnen und –absolventen, der Studierenden und der Erwachsenen ermittelt werden (Senkbeil & Ihme, 2017c, d; Senkbeil, Ihme & Schöber, 2019).

Da ICT Literacy in NEPS bis 2017 anhand von Papier-und-Bleistift-Tests (PPT) erfasst wurde, wurde weiterführend die Validität der PPT-basierten ICT Literacy-Tests im Vergleich mit simulationsbasierten Testinstrumenten (SBT) überprüft. Diese beinhalteten reale oder simulierte Programmanwendungen mit interaktiven Aufgaben, um eine verhaltensnahe Erfassung des Konstrukts zu ermöglichen (Goldhammer, Naumann & Keßel, 2013). Die Ergebnisse deuten darauf hin, dass PPTs computer- und informationsbezogene Wissensbestände (deklaratives und prozedurales Fakten- und Konzeptwissen) valide erfassen, aber nicht das gleiche Konstrukt messen wie SBTs und eine geringere Konstruktbreite als SBTs aufweisen, da z.B. prozedurale Fertigkeiten unberücksichtigt bleiben (Senkbeil & Ihme, 2014, 2020). Gleichwohl zeigen sich zwischen statischen MC- und interaktiven Simulationsaufgaben substantielle und zum Teil sehr hohe Korrelationen, die je nach Testinstrument zwischen  $r = .68$  und  $r = .93$  variieren. Als Validitätsbelege für PPTs können zudem mit SBTs vergleichbar hohe konvergente und diskriminante Zusammenhänge mit Validierungsvariablen gelten (Senkbeil & Ihme, 2020). Aufgrund der bisherigen Validierungsstudien kann somit zusammengefasst werden, dass die in NEPS eingesetzten Papier-und-Bleistift-Tests deklarative und prozedurale Wissensbestände valide erfassen, wenngleich prozedurale und metakognitive Fertigkeiten anhand der Restriktionen im Aufgabenformat unberücksichtigt bleiben. Entsprechend werden seit 2018 zusätzlich simulationsbasierte Aufgaben zur Messung von ICT Literacy eingesetzt (siehe auch Abschnitt 6.5).

### 6.4.3 Gruppenunterschiede

Tabelle 19 fasst die standardisierten Mittelwertsunterschiede (Cohen's  $d$ ) der Testwerte für zentrale Gruppierungsvariablen zusammen. Dabei werden zum Teil beträchtliche soziale Disparitäten in der ICT Literacy (sozioökonomischer Status, elterlicher und eigener Bildungsabschluss) ermittelt, die sich während der Schulzeit zumindest nicht zu verringern und sich über die gesamte Lebensspanne (d.h. von der Grundschule bis ins Erwachsenenalter) sogar zu vergrößern scheinen. Mit Ausnahme der Studierenden entsprechen die sozial bedingten Disparitäten in den untersuchten Alterskohorten einem mittleren bis starken Effekt ( $|0.44| \leq d \leq |0.93|$ ). Ebenso zeigen sich große Effekte (Mittelwertsunterschiede) in Abhängigkeit der besuchten Schulform. Beim Geschlecht lassen sich nach der Sekundarstufe I zwar durchgehend Kompetenzvorteile für die Männer ermitteln, jedoch ist hier zu prüfen, inwieweit dieser Effekt auf das Ausscheiden leistungsschwächerer Männer nach Ablauf der Pflichtschulzeit (z.B. Schulabbruch, keine Aufnahme der beruflichen Erstausbildung) aus dem Nationalen Bildungspanel zurückzuführen ist.

Tabelle 19: Standardisierte Mittelwertsunterschiede der Testwerte für ausgewählte sozio-demographische Merkmale (Domäne ICT-Literacy)

Startkohorte	Alter / Klasse	Geschlecht	Sozio-ökonomischer Status	Migrationshintergrund	Schultyp	Bildungsabschluss
		(Frauen vs. Männer)	(hoch vs. gering)	(ohne vs. mit)	(Gymnasium vs. Nicht-Gymnasium)	(Abitur vs. Nicht-Abitur)
SC 2	Klasse 3	-0.07	0.449	0.23	—	0.48
SC 3	Klasse 6	-0.10	0.48	0.41	0.92	0.54
SC 3	Klasse 9	-0.05	0.54	0.43	1.01	0.55
SC 3	Klasse 12 / Schulabsolventen	-0.33	0.48	0.48	0.80	0.50
SC 4	Klasse 9	0.02	0.51	0.49	1.14	0.64
SC 4	Klasse 12 / Schulabsolventen	-0.27	0.44	0.53	1.07	0.55
SC 5	Studierende	-0.60	0.18	0.44	—	0.22
SC 6	Erwachsene	-0.36	0.75	0.14	—	0.93

Bemerkung. Positive Werte spiegeln einen höheren Mittelwert in der ersten Gruppe (z.B. „Frauen“) wider, während negative Werte einen höheren Mittelwert in der zweiten Gruppe (z.B. „Männer“) widerspiegeln. Sozioökonomischer Status = HISEI der Eltern bzw. eigener ISEI in SC6, Migrationshintergrund = mit (=mind. ein Elternteil nicht Deutschland geboren) vs. ohne (=beide Elternteile in Deutschland geboren).

## 6.5 Ausblick

Aufgrund der fortschreitenden Digitalisierung aller Lebensbereiche ist die längsschnittliche und valide Messung von ICT Literacy aufwendiger als in anderen Domänen (z.B. Kernkompetenzen wie Lesen oder Mathematik; Kuhlemeier & Hemker, 2007; Leu, Forzani et al., 2013). Konkret bedeutet dies, dass die Aufgaben zur Erfassung von ICT Literacy aufgrund technologischer Entwicklungen (z.B. Weiterentwicklung von Software, mobilen Endgeräten, Applikationen) und sich ändernder Nutzungsgewohnheiten (Czajka, 2011; van Eimeren & Frees, 2012; Medienpädagogischer Forschungsverbund Südwest, 2018) im Umgang mit digitalen Medien regelmäßig überprüft und gegebenenfalls erneuert werden müssen.

Eine in diesem Zusammenhang wichtige Neuerung in NEPS stellt seit 2018 der (zusätzliche) Einsatz von simulationsbasierten, d.h. von computerisierten und interaktiven Testaufgaben dar. Diese Erweiterung im Aufgabenformat kann als Übergang von einer wissensbasierten (anhand von MC-Aufgaben) zu einer wissens- und verhaltensbasierten Konstrukterfassung (anhand zusätzlicher Simulationsaufgaben) verstanden werden (Senkbeil & Ihme, 2020). Anhand von MC-Aufgaben können deklarative und prozedurale Wissensbestände in Form von deklarativem Fakten- und Konzeptwissen sowie prozeduralem, anwendungsorientiertem Wissen erfasst werden (Senkbeil & Ihme, 2014). Durch simulationsbasierte Aufgaben können zusätzlich multimediale Inhalte (z.B. Text, Grafiken, Video) in das Stimulusmaterial eingebunden werden, die ein hohes Maß an Interaktivität zwischen Antwortverhalten und Stimulusdarbietung gestatten. Hierbei ist das erworbene Fakten- und Konzeptwissen auf spezifische Problemstellungen in realitätsnahen Situationen anzuwenden, so dass anhand simulationsbasierter Aufgaben zusätzlich prozedurale Fertigkeiten und metakognitive Fertigkeiten (zielgerichtete Denk- und Handlungsabläufe, die bei komplexeren Tätigkeiten wie z.B. der Erstellung einer Präsentation benötigt werden) erfasst werden können. Da MC-Aufgaben und simulationsbasierte Aufgaben in den bisherigen NEPS-Studien sehr hoch miteinander korrelieren (bisherige Entwicklungsstudien und Haupterhebungen:  $.87 \leq r \leq .93$ ; vgl. auch Ihme et al., 2017; Senkbeil & Ihme, 2020), ist ICT Literacy in NEPS auch weiterhin als eindimensionales Konstrukt konzeptualisiert. Durch den kombinierten Einsatz von MC- und simulationsbasierten Aufgaben in NEPS ergeben sich eine Reihe von Vorteilen: (1) Die zusätzliche Darbietung von Simulationsaufgaben gewährleisten eine international anschlussfähige Konstruktrepräsentation (wissens- und verhaltensbasierte Konstrukterfassung; Senkbeil & Ihme, 2020; Siddiq et al., 2016). (2) Simulationsaufgaben ermöglichen die Analyse von Prozessdaten (z.B. Bearbeitungsdauer, Anzahl der Interaktionen pro Aufgabe), welche Hinweise auf (erfolgreiche) Bearbeitungsstrategien sowie spezifische Probleme beim Bearbeiten komplexer Aufgabenstellungen geben können (Goldhammer, Naumann, Stelter, Toth, Rölke & Klieme, 2014; Naumann, Goldhammer, Rölke & Stelter, 2014). (3) Die Darbietung älterer (d.h. in früheren Kohorten eingesetzt) und weiterhin geeigneter MC-Aufgaben gewährleistet die Anschlussfähigkeit an bisherige Kompetenzmessungen in NEPS und ermöglicht beispielsweise die Analyse von Kohorteneffekten. (4) Die enge Koppelung von wissens- und verhaltensbasierter Kompetenzmessung ermöglicht die Fortschreibung längsschnittlicher Kompetenzmodellierung in den bisherigen Startkohorten (z.B. SC3, SC4).

## 6.6 Ausgewählte Publikationen mit NEPS-Daten zu ICT Literacy

- Ihme, J. M. & Senkbeil, M. (2017). Warum können Jugendliche ihre eigenen computerbezogenen Kompetenzen nicht realistisch einschätzen? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 49, 24–37.
- Ihme, J. M., Senkbeil, M., Goldhammer, F. & Gerick, J. (2017). Assessment of computer and information literacy in ICILS 2013: Do different item types measure the same construct? *European Educational Research Journal*, 16, 716–732.
- Senkbeil, M. & Ihme, J. M. (2014). Wie valide sind Papier-und-Bleistift-Tests zur Erfassung computerbezogener Kompetenzen? *Diagnostica*, 60, 22–34.
- Senkbeil, M. & Ihme, J. M. (2017c). Motivational factors predicting ICT literacy: First evidence on the structure of an ICT motivation inventory. *Computers & Education*, 108, 145–158.
- Senkbeil, M. & Ihme, J. M. (2017d). Entwicklung und Validierung eines Kurzfragebogens zur Erfassung computerbezogener Anreizfaktoren bei Erwachsenen. *Diagnostica*, 63, 87–98.
- Senkbeil, M. & Ihme, J. M. (2020). Diagnostik von ICT Literacy: Messen Multiple-Choice-Aufgaben und simulationsbasierte Aufgaben vergleichbare Konstrukte? *Diagnostica*, 66, 147–157.
- Senkbeil, M., Ihme, J. M. & Schöber, C. (2019). Wie gut sind angehende und fortgeschrittene Studierende auf das Leben und Arbeiten in der digitalen Welt vorbereitet? Ergebnisse eines Standard-Setting-Verfahrens zur Beschreibung von ICT-bezogenen Kompetenzniveaus. *Zeitschrift für Erziehungswissenschaft*, 22, 1359–1384.
- Senkbeil, M., Ihme, J. M. & Schöber, C. (2020; Open Access). Schulische Medienkompetenzförderung in einer digitalen Welt: Über welche digitalen Kompetenzen verfügen angehende Lehrkräfte? *Psychologie in Erziehung und Unterricht*.
- Senkbeil, M., Ihme, J.M. & Wittwer, J. (2013a). The test of Technological and Information Literacy (TILT) in the National Educational Panel Study: Development, empirical testing, and evidence for validity. *Journal for Educational Research Online*, 5, 139–161.
- Senkbeil, M., Ihme, J. M. & Wittwer, J. (2013b). Entwicklung und erste Validierung eines Tests zur Erfassung technologischer und informationsbezogener Literacy (TILT) für Jugendliche am Ende der Sekundarstufe I. *Zeitschrift für Erziehungswissenschaft*, 16, 671–691.

## Literatur

- Blossfeld, H.-P. (2010). Education across the life course. In German Data Forum (RatSWD) (Ed.). *Building or progress. Expanding the research infrastructure for the social, economic, and behavioral sciences* (Vol. 2, S. 825–840). Opladen: Budrich UniPress.
- Czajka, S. (2011). Internetnutzung in privaten Haushalten in Deutschland. Ergebnisse der Erhebung 2010. In Statistisches Bundesamt (Hrsg.), *Wirtschaft und Statistik* (S. 709–728). Wiesbaden: Statistisches Bundesamt.
- Eickelmann, B., Bos, W., Gerick, J. & Labusch, A. (2019). Computer- und informationsbezogenen Kompetenzen von Schülerinnen und Schülern der 8. Jahrgangsstufe in Deutschland im zweiten internationalen Vergleich. In B. Eickelmann, W. Bos, J. Gerick, F. Goldhammer, H. Schaumburg, K. Schwippert, M. Senkbeil & J. Vahrenhold (Hrsg.), *ICILS 2018 #Deutschland – Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im zweiten internationalen Vergleich und Kompetenzen im Bereich Computational Thinking* (S. 113–135). Münster: Waxmann.
- ETS [Educational Testing Service] (2002). *Digital transformation. A framework for ICT literacy*. Princeton, NJ: ETS.
- Ferrari, A. (2012). *Digital competence in practice: An analysis of frameworks*. European Union.
- Ferrari, A. (2013). *DIGCOMP: A framework for developing and understanding. Digital competence in Europe*. Luxembourg.
- Fraillon, J., Ainley, J., Schulz, W., Duckworth, D. & Friedman, T. (2019). *IEA International computer and information literacy study 2018: Assessment framework*. Amsterdam: IEA.
- Goldhammer, F., Naumann, J & Keßel, Y. (2013). Assessing individual differences in basic computer skills: Psychometric characteristics of an interactive performance measure. *European Journal of Psychological Assessment*, 29, 263–275.
- Goldhammer, F., Naumann, J., Stelter, A., Toth, K., Rölke, H. & Klieme, E. (2014). The time on task effect in reading and problem solving is moderated by item difficulty and ability: Insights from computer-based large-scale assessment. *Journal of Educational Psychology*, 106, 608–626.
- Hanushek, E. A., Schwerdt, G., Wiederhold, S. & Wößmann, L. (2015). Returns to skills around the world: Evidence from PIAAC. *European Economic Review*, 73, 103–130.
- Ihme, J. M. & Senkbeil, M. (2018). *NEPS technical report for computer literacy: Scaling results of starting cohort 2 for grade 3* (NEPS Survey Paper No. 42). Bamberg: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Ihme, J. M., Senkbeil, M., Goldhammer, F. & Gerick, J. (2017). Assessment of computer and information literacy in ICILS 2013: Do different item types measure the same construct? *European Educational Research Journal*, 16, 716–732.
- KMK [Kultusministerkonferenz] (2016). *Bildung in der digitalen Welt. Strategie der Kultusministerkonferenz*. Bonn: KMK.
- Kuhlemeier, H. & Hemker, B. (2007). The impact of computer use at home on students` internet skills. *Computers & Education*, 49, 460–480.
- Leu, D. J., Forzani, E., Burlingame, C., Kukikowich, J., Sedransk, N., Coiro, J. & Kennedy, C. (2013). The new literacies of online research and comprehension: Assessing and preparing students for the 21st century with common core standards. In S. B. Neuman & L. B. Gambrell (Eds.), *Quality reading instruction in the age of common core standards*. International Reading Association.

- Medienpädagogischer Forschungsverbund Südwest (2018). *JIM-Studie 2018. Jugend, Information, Medien*. Stuttgart: Medienpädagogischer Forschungsverbund Südwest.
- Naumann, J., Goldhammer, F., Rölke, H. & Stelter, A. (2014). Erfolgreiches Problemlösen in technologiebasierten Umgebungen: Wechselwirkungen zwischen Interaktionsschritten und Aufgabenanforderungen. *Zeitschrift für Pädagogische Psychologie*, 28, 193–203.
- OECD (2013). *OECD skills outlook 2013. First results from the survey of adult skills*. Paris: OECD.
- Pohl, S. & Carstensen, C. H. (2012). *NEPS technical report – Scaling the data of the competence tests: NEPS working paper no. 14*. Otto-Friedrich-Universität, Nationales Bildungspanel: Bamberg.
- Sefton-Green, J., Nixon, H. & Erstad, O. (2009). Reviewing approaches and perspectives on „Digital Literacy“. *Pedagogies: An International Journal*, 4, 107–125.
- Senkbeil, M., Drossel, K., Eickelmann, B. & Vennemann, M. (2019). Soziale Herkunft und computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im zweiten internationalen Vergleich. In B. Eickelmann, W. Bos, J. Gerick, F. Goldhammer, H. Schaumburg, K. Schwippert, M. Senkbeil & J. Vahrenhold (2019). *ICILS 2018 # Deutschland* (S. 310–333). Münster: Waxmann.
- Senkbeil, M., Goldhammer, F., Bos, W., Eickelmann, B., Schwippert, K. & Gerick, J. (2014). Das Konstrukt der computer- und informationsbezogenen Kompetenzen. In W. Bos, B. Eickelmann, J. Gerick, F. Goldhammer, H. Schaumburg, K. Schwippert et al. (Hrsg.), *ICILS 2013. Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern in der 8. Jahrgangsstufe im internationalen Vergleich* (S. 83–112). Münster: Waxmann.
- Senkbeil, M. & Ihme, J. M. (2012). *NEPS technical report for computer literacy – Scaling results of starting cohort 4 in ninth grade* (NEPS working paper no. 17). Bamberg: Otto-Friedrich-Universität, Nationales Bildungspanel.
- Senkbeil, M. & Ihme, J. M. (2014). Wie valide sind Papier-und-Bleistift-Tests zur Erfassung computerbezogener Kompetenzen? *Diagnostica*, 60, 22–34.
- Senkbeil, M. & Ihme, J. M. (2015). *NEPS technical report for computer literacy: Scaling results of starting cohort 6 - adults* (NEPS working paper no. 61). Bamberg: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Senkbeil, M. & Ihme, J. M. (2017a). *NEPS technical report for computer literacy: Scaling results of starting cohort 4 for grade 12* (NEPS survey paper no. 25). Bamberg: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Senkbeil, M. & Ihme, J. M. (2017b). *NEPS technical report for computer literacy: Scaling results of starting cohort 3 for grade 9* (NEPS survey paper no. 29). Bamberg: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Senkbeil, M. & Ihme, J. M. (2017c). Motivational factors predicting ICT literacy: First evidence on the structure of an ICT motivation inventory. *Computers & Education*, 108, 145–158.
- Senkbeil, M. & Ihme, J. M. (2017d). Entwicklung und Validierung eines Kurzfragebogens zur Erfassung computerbezogener Anreizfaktoren bei Erwachsenen. *Diagnostica*, 63, 87–98.
- Senkbeil, M. & Ihme, J. M. (2020). Diagnostik von ICT Literacy: Messen Multiple-Choice-Aufgaben und simulationsbasierte Aufgaben vergleichbare Konstrukte? *Diagnostica*, 66, 147–157.

- Senkbeil, M., Ihme, J. M. & Adrian, E. D. (2013). *NEPS technical report for computer literacy – Scaling results of starting cohort 3 in grade 6 (wave 2)* (NEPS Working Paper No. 39). Bamberg: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Senkbeil, M., Ihme, J. M. & Schöber, C. (2019). Wie gut sind angehende und fortgeschrittene Studierende auf das Leben und Arbeiten in der digitalen Welt vorbereitet? Ergebnisse eines Standard-Setting-Verfahrens zur Beschreibung von ICT-bezogenen Kompetenzniveaus. *Zeitschrift für Erziehungswissenschaft*, 22, 1359–1384.
- Senkbeil, M., Ihme, J.M. & Wittwer, J. (2013a). The test of Technological and Information Literacy (TILT) in the National Educational Panel Study: Development, empirical testing, and evidence for validity. *Journal for Educational Research Online*, 5, 139–161.
- Senkbeil, M., Ihme, J. M. & Wittwer, J. (2013b). Entwicklung und erste Validierung eines Tests zur Erfassung technologischer und informationsbezogener Literacy (TILT) für Jugendliche am Ende der Sekundarstufe I. *Zeitschrift für Erziehungswissenschaft*, 16, 671–691.
- Siddiq, F., Hatlevik, O. E., Olsen, R. V., Throndsen, I. & Scherer, R. (2016). Taking a future perspective by learning from the past – A systematic review of assessment instruments that aim to measure primary and secondary school students' ICT literacy. *Educational Research Review*, 19, 58–84.
- van Deursen, A. & van Dijk, J. (2010). Internet skills and the digital divide. *New Media & Society*, 13, 893–911.
- van Dijk, J. (2009). One Europe, digitally divided. In Chadwick, A. & Howard, P. N. (Eds.), *Routledge handbook of internet politics* (pp. 288–304). New York: Taylor & Francis.
- van Eimeren, B. & Frees, B. (2012). 76 Prozent der Deutschen online – neue Nutzungssituationen durch mobile Endgeräte. *Media Perspektiven*, 7–8, 362–379.
- Van Laar, E., van Deursen, A.J.A.M., van Dijk, J.A.G.M. & de Haan, J. (2017). The relation between 21st-century skills and digital skills: A systematic literature review. *Computers in Human Behavior*, 72, 577–588.

## 7. Metakognition

Kathrin Lockl

### 7.1 Einleitung

Neben Kernkompetenzen wie Lesen und Mathematik werden im NEPS metakognitive Kompetenzen erfasst, die die Fähigkeit betreffen, das eigene Lernen erfolgreich regulieren zu können (vgl. Artelt et al., 2013; Weinert et al., 2019). Metakognition gilt als zentrale Komponente selbstregulierten Lernens und ist als fächerübergreifende Kompetenz für den schulischen und außerschulischen Wissenserwerb gerade auch vor dem Hintergrund lebenslangen Lernens von großer Bedeutung (Boekaerts, 1999; Efklides, 2008; Pintrich, 2004; Winne & Hadwin, 1998). Es wird angenommen, dass Personen, die ihre eigenen kognitiven Kapazitäten richtig einschätzen können, die erkennen, welche Anforderungen eine Lernaufgabe beinhaltet, und die in der Lage sind, daraufhin geeignete Strategien auszuwählen, effektiver lernen, als Personen, die nicht über entsprechende Ressourcen verfügen. Empirisch bestätigt wurden diese Annahmen beispielsweise im Rahmen der PISA-Studie, in der sich das Lernstrategiewissen nach den kognitiven Grundfähigkeiten als zweitwichtigster Prädiktor für die Lesekompetenz erwies (Artelt, Schiefele, Schneider & Stanat, 2002).

Auf der Basis gängiger Konzeptionen (z.B. Brown, 1987; Flavell, 1979) werden im Nationalen Bildungspanel zwei Aspekte von Metakognition unterschieden. Die deklarative Metakognition bezieht sich auf die Wissenskomponente, welche beispielsweise das Wissen darüber beinhaltet, wo die eigenen Stärken und Schwächen bei Gedächtnis- und Lernaufgaben liegen, welche Merkmale eine Aufgabe leichter oder schwieriger machen oder welche Strategien bei welcher Aufgabenstellung nützlich sind. Die prozedurale Metakognition betrifft die Fähigkeit, eigene Lernaktivitäten überwachen und steuern zu können (Regulationskomponente). Hierzu zählt, den eigenen Lernstand richtig einzuschätzen und daraufhin angemessene Konsequenzen für das weitere Lernverhalten zu ziehen und beispielsweise geeignete Lernstrategien auszuwählen. Im Rahmen des Nationalen Bildungspanels werden Indikatoren für beide Aspekte der Metakognition erhoben.

In bisherigen Forschungsarbeiten zur deklarativen und prozeduralen Metakognition wurden vorwiegend querschnittliche Ansätze verfolgt (z.B. Artelt et al, 2002; Schneider, Schlagmüller & Visé, 1998). Längsschnittliche Studien stellen eher die Ausnahme dar (vgl. z.B. Artelt, Neuenhaus, Lingel & Schneider, 2012) und fokussierten häufig auf die Entwicklung von Metakognition im frühen Kindesalter (Annevirta, Laakkonen, Kinnunen & Vauras, 2007; Ebert, 2011; Lockl & Schneider, 2007). Um jedoch ein besseres Verständnis darüber zu erhalten, wie sich Metakognition entwickelt und von welchen kognitiven und motivationalen Faktoren die Entwicklung beeinflusst wird, sind Längsschnittstudien notwendig. Mit Hilfe der längsschnittlich angelegten Erhebungen in NEPS können neue Erkenntnisse über die Entwicklung von Metakognition in verschiedenen Phasen des Lebenslaufs gewonnen werden. Zudem ermöglichen es die längsschnittlichen Studien mehr über die Bedeutung der Metakognition für andere Kompetenzbereiche zu erfahren. Insofern gehen die im NEPS durchgeführten Erhebungen zur Metakognition über die bisher in der Literatur berichteten Studien hinaus und bieten ein vielversprechendes Analysepotential.

Im Folgenden wird zunächst auf die Erfassung der deklarativen Metakognition eingegangen, bevor im Anschluss einige zentrale Befunde zur prozeduralen Metakognition dargestellt werden.

## 7.2 Bisherige Erhebung des Konstrukts deklarative Metakognition

Die folgende Tabelle beschreibt sowohl die Erhebungszeiten als auch die Erhebungsmodi der bisherigen Erhebungen zur deklarativen Metakognition. Alle bisherigen Erhebungen fanden im Klassenkontext statt.

Tabelle 20: Erhebungszeiten und –modi der für die deklarative Metakognition gelaufenen Erhebungen

Startkohorte	Alter/ Klasse	Erhebungsjahr	Erhebungszeit (in Minuten)	Erhebungsmodus
SC 2	Klasse 1	2013	15	Papier
SC 2	Klasse 3	2014	15	Papier
SC 3	Klasse 6	2011	15	Papier
SC 3	Klasse 9	2014	15	Papier
SC 4	Klasse 9	2010	15	Papier

## 7.3 Beispielitems

Die deklarative Metakognition wird im NEPS über neu konstruierte Wissenstests erfasst, die auf den Konstruktionsprinzipien bewährter Testverfahren beruhen, wie z.B. dem metakognitiven Wissenstest zu Lesestrategien im Rahmen von PISA (Artelt et al., 2009; Schlagmüller & Schneider, 2007). Der Fokus liegt auf der Erfassung des konditionalen Strategiewissens nach Paris, Lipson und Wixson (1983), welches sich auf das Wissen über die Nützlichkeit von Strategien in bestimmten Situationen bezieht und eine wichtige Grundlage für den Einsatz von Strategien in konkreten Lernsituationen darstellt. Im Hinblick auf die längsschnittliche Erhebung von deklarativer Metakognition erschien es sinnvoll, metakognitives Wissen in allen Altersgruppen domänenübergreifend, d.h. losgelöst von bestimmten Fächern oder Inhaltsbereichen (wie z.B. Lesen oder Mathematik), zu erfassen (Händel, Artelt & Weinert, 2013).

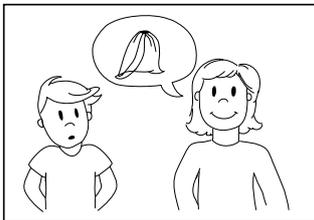
Eva hat am nächsten Tag Sport-Unterricht. Ihre Turnsachen hat sie am Abend vorher schon fertig in ihre Tasche gepackt. Was kann Eva tun, damit sie die Turnsachen am nächsten Morgen nicht vergisst?

Die drei Bilder zeigen, was Eva machen kann, um ihre Turnsachen nicht zu vergessen.



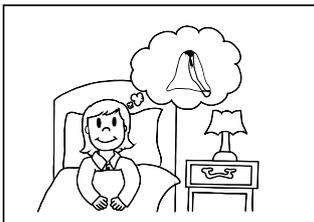
Eva hängt ihre Tasche an die Wohnungstür.

☆	☆☆	☆☆☆
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



Eva sagt ihrem kleinen Bruder Bescheid. Er soll sie an die Tasche erinnern.

☆	☆☆	☆☆☆
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



Eva denkt vor dem Schlafengehen ganz fest an die Tasche.

☆	☆☆	☆☆☆
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

(Anmerkung: In den Testheften der Kinder sind nur die Bilder abgedruckt; die dazugehörigen Texte werden vom Testleiter vorgelesen.)

Abbildung 21. Beispielaufgabe Klasse 1.

Die Testverfahren enthalten mehrere Szenarien, in denen eine Auswahl von Strategien vor dem Hintergrund ganz konkreter Anforderungen vorgegeben wird. Innerhalb eines Szenarios werden paarweise Vergleiche zwischen mehreren Handlungsoptionen vorgenommen, die zuvor anhand von Experteneinschätzungen validiert wurden. Aufgabe der teilnehmenden Personen ist es, für die jeweils skizzierten Kontexte die Nützlichkeit verschiedener Strategien zu bewerten. Bei der Konstruktion der Tests wird besonders darauf geachtet, altersangemessene Situationen und Strategien auszuwählen und das Anforderungsniveau des Tests altersgemäß zu gestalten. So werden Schülerinnen und Schülern der ersten Klasse z.B. weniger Strategien vorgegeben und diese werden in den Testheften bildhaft dargestellt.

Peter hat diese Woche viel zu tun: Er soll zwei Mal im Verein schwimmen, hat viele Hausaufgaben und muss ein Geburtstagsgeschenk für seinen Freund besorgen.  
Wie sollte er vorgehen, um alles zu schaffen?

*Bitte beurteile die Nützlichkeit aller Vorschläge.*

	gar nicht nützlich	kaum nützlich	etwas nützlich	sehr nützlich
Er macht einen Plan für die Woche und teilt sich die Zeit für die Aufgaben ein. Seinen Plan hält er sehr genau ein.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Er verbindet verschiedene Aufgaben miteinander. Das Geschenk kauft er auf dem Weg zum Schwimmverein.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Er lässt sich von anderen helfen. Er bittet seinen Bruder, das Geschenk zu kaufen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Er macht nur die Hausaufgaben, die schnell gehen. Dann macht er die anderen Dinge.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Er besorgt zuerst das Geburtstagsgeschenk. Wenn das zu lange dauert, lässt er die Hausaufgaben oder das Schwimmen weg.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Er macht immer genau die Dinge, auf die er gerade am meisten Lust hat.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Abbildung 22. Beispielaufgabe Klasse 9.

Da sich die existierenden Testinstrumente zur deklarativen Metakognition eher auf Lernstrategien in einem schulischen oder universitären Kontext beziehen und sich zudem Strategien in Situationen, die im alltäglichen Leben von Erwachsenen eine Rolle spielen, weniger gut im Sinne eines Expertenurteils als richtig oder falsch bewerten lassen, bieten sich bei den höheren Startkohorten Verfahren zur Selbsteinschätzung an. Geeignete Items zur Erfassung des selbstregulierten Lernens und der Selbstkontrolle sollen in zukünftigen Erhebungen eingebracht werden.

## 7.4 Empirische Analysen

### 7.4.1 Deskriptive Informationen

Tabelle 21 gibt einen Überblick über zentrale statistische Kennwerte der NEPS-Erhebungen, in denen Indikatoren der deklarativen Metakognition erhoben wurden. Insgesamt lassen sich gute Reliabilitäten der Skala in den verschiedenen Altersbereichen nachweisen.

Tabelle 21: Deskriptive Informationen zu Erhebungen der deklarativen Metakognition

Startkohorte	Alter / Klasse	N (gesamt)	N (gültig)	% fehlend	M	SD	Reliabilität
SC 2	Klasse 1	6733	6291	6.60	0.56	0.20	0.73
SC 2	Klasse 3	5798	5597	3.50	0.59	0.17	0.67
SC 3	Klasse 6	4927	4876	1.00	0.72	0.15	0.89
SC 3	Klasse 9	4624	4589	< 1.00	0.72	0.15	0.89
SC 4	Klasse 9	13980	13925	< 1.00	0.80	0.14	0.89

Bemerkung. N (gesamt) = Anzahl der Personen, bei denen der Test administriert wurde; N (gültig) = Anzahl der Personen mit gültigem Testwert, % fehlend = Anteil an Personen ohne gültigen Testwert, M = Mittelwert des Testwerts; die Mittelwerte sind aufgrund von unterschiedlichen Items pro Erhebung nicht miteinander vergleichbar, SD = Standardabweichung des Testwerts, Reliabilität = Cronbachs alpha.

### 7.4.2 Psychometrische Informationen

Eine ausführliche Darstellung der psychometrischen Kennwerte ist in den Dokumentationen auf der Website einsehbar (vgl. Lockl, 2012, 2013, 2015a, 2017, 2018).

Im Hinblick auf die Validitätsprüfung bestand, basierend auf einer Reihe von Studien (Ebert, 2011, 2015; Lockl & Schneider, 2006, 2007), die Erwartung, dass sprachliche Fähigkeiten und Fertigkeiten mit dem metakognitiven Wissen verknüpft sein sollten, sich aber dennoch davon abgrenzen lassen. In Einklang damit ergeben sich in der ersten Klasse relativ niedrige Zusammenhänge zwischen metakognitivem Wissen und sprachlichen Maßen, die mit Koeffizienten zwischen  $r = .25$  und  $r = .29$  ähnlich hoch ausfallen wie in der BIKS-Studie (Haberkorn, Lockl, Pohl, Ebert & Weinert, 2014). Somit sprechen die Befunde im Sinne der Validitätsprüfung dafür, dass bei der Messung metakognitiven Wissens bei jüngeren Kindern keine starke Konfundierung mit den sprachlichen Kompetenzen gegeben ist.

Weiterhin zeigen sich erwartungsgemäß bedeutsame Zusammenhänge zwischen metakognitivem Wissen einerseits und verschiedenen Indikatoren für schulische Leistungen andererseits. So ist das metakognitive Wissen bei den Erhebungen in der Sekundarstufe sowohl in Klasse 6 als auch in Klasse 9 deutlich mit den Leistungen in den anderen Kompetenzdomänen (u.a. Naturwissenschaften, ICT-Literacy, Lesen) assoziiert. Die Korrelationen liegen

dabei in einem mittleren Bereich zwischen  $r = .38$  und  $r = .57$ . Etwas niedrigere Zusammenhänge mit mathematischen und naturwissenschaftlichen Kompetenzen (mit Korrelationen zwischen  $r = .31$  und  $r = .33$ ) werden in Klasse 1 berichtet. In allen Klassenstufen sind die Leistungen im Test zur deklarativen Metakognition zudem mit kompetenzbezogenen Lehrereinschätzungen bzw. in höheren Klassen mit den Noten z.B. in Deutsch und Mathematik verknüpft.

Betrachtet man die Prädiktivität des metakognitiven Wissens für die Leistung in späteren Kompetenztests, so zeigen Befunde von Miyamoto, Pfof und Artelt (2019), dass das metakognitive Wissen in Klasse 6 die Lesekompetenz der Schülerinnen und Schüler in Klasse 7 signifikant vorhersagt. Dem metakognitiven Wissen kommt zudem die Rolle eines Mediators zu, wobei signifikante indirekte Effekte der intrinsischen Motivation auf die Lesekompetenz vermittelt über das metakognitive Wissen nachzuweisen sind. Interessanterweise ist das metakognitive Wissen bei Schülerinnen und Schülern mit niedrigerer Lesekompetenz vergleichsweise bedeutsamer als bei den Gleichaltrigen mit höherer Lesekompetenz, um die Beziehung zwischen intrinsischer Motivation und Lesekompetenz zu erklären. Damit unterstreichen die in der NEPS-Studie gewonnenen Ergebnisse die querschnittlichen Befunde der PISA-Studien 2000 und 2009 (vgl. Artelt et al., 2002; Artelt & Schneider, 2015) und erweitern diese auch für andere Altersgruppen unter einer längsschnittlichen Perspektive.

### **7.4.3 Gruppenunterschiede**

Die nachfolgende Tabelle fasst die standardisierten Mittelwertsunterschiede (Cohen's  $d$ ) der Testwerte für zentrale Gruppierungsvariablen zusammen.

In Einklang mit der Literatur zeigen sich in den höheren Klassenstufen leichte Mittelwertsunterschiede im metakognitiven Wissen zugunsten von Mädchen sowie deutlichere Vorteile für Schülerinnen und Schüler, die ein Gymnasium oder eine Realschule besuchen (vgl. Artelt, Beinicke, Schlagmüller & Schneider, 2009; Artelt, Neuenhaus, Lingel & Schneider, 2012). Weiterhin lassen sich erwartungsgemäß Vorteile für Schülerinnen und Schüler mit höherem Sozialstatus und ohne Migrationshintergrund nachweisen.

Tabelle 22: Mittelwertsunterschiede (Cohen's d) der Testwerte für zentrale Gruppierungsvariablen (deklarative Metakognition)

<b>Startkohorte</b>	<b>Alter / Klasse</b>	<b>Geschlecht</b>	<b>Sozio-ökonomischer Status</b>	<b>Migrationshintergrund</b>	<b>Schultyp</b>	<b>Bildungsabschluss</b>
		(weiblich vs. männlich)	(hoch vs. gering)	(ohne vs. mit)	(Gymnasium vs. Nicht-Gymnasium)	(Abitur vs. kein Abitur)
SC 2	Klasse 1	-0.07	0.23	0.14	-	0.30
SC 2	Klasse 3	-0.03	0.36	0.15		0.40
SC 3	Klasse 6	0.22	0.52	0.42	0.82	0.52
SC 3	Klasse 9	0.42	0.45	0.32	0.82	0.49
SC 4	Klasse 9	0.35	0.37	0.37	0.82	0.41

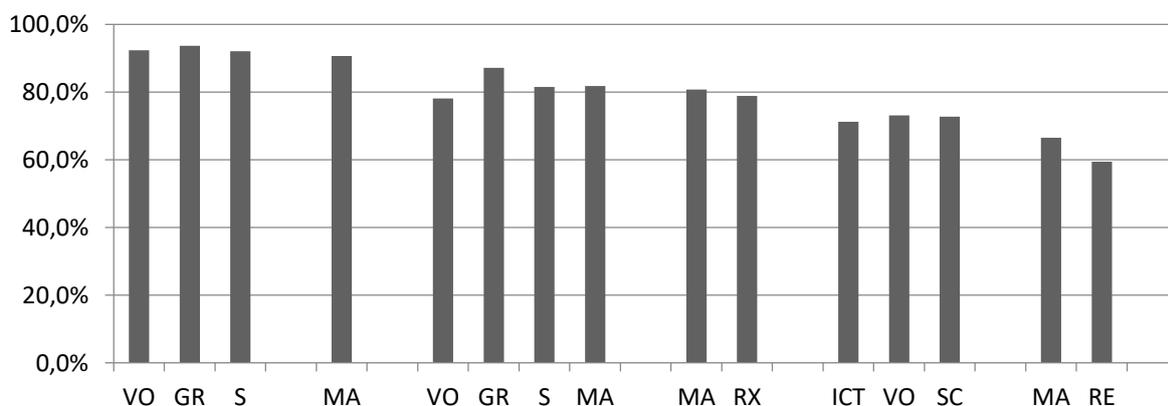
Bemerkung. Positive Werte spiegeln einen höheren Mittelwert in der ersten Gruppe (z.B. „weiblich“) wider, während negative Werte einen höheren Mittelwert in der zweiten Gruppe (z.B. „männlich“) widerspiegeln. Sozioökonomischer Status = HISEI der Eltern, Migrationshintergrund = ohne (= beide Elternteile in Deutschland geboren) vs. mit (= mind. ein Elternteil nicht Deutschland geboren).

## 7.5 Prozedurale Metakognition

Zur prozeduralen Metakognition gehört die Regulation des Lernprozesses durch Aktivitäten der Planung, Überwachung und Kontrolle. Der prozedurale Aspekt der Metakognition wird im Rahmen von NEPS in Kombination mit den Kompetenztests der einzelnen Domänen dabei nicht als direktes Maß derartiger Planungs-, Überwachungs- und Kontrollaktivitäten gemessen, sondern als metakognitives Urteil, das sich auf die Überwachung der Lernleistung während (bzw. kurz nach) der Lernphase bezieht (s.a. Nelson & Narens, 1990). Hierzu werden die Studienteilnehmerinnen und -teilnehmer nach Bearbeitung der jeweiligen Kompetenztests gebeten, ihre eigene Leistung in dem gerade bearbeiteten Test einzuschätzen. Erfragt wird ab Klassenstufe 5 die Anzahl der vermutlich richtig gelösten Aufgaben; Kindergarten- und Grundschul Kinder sollen ihre Leistung anhand einer 5-Punkte Smileyskala einschätzen.

Pro Domäne wird hierzu in der Regel eine Frage eingesetzt. Bei Kompetenzdomänen, die sich in zusammenhängende einzelne Teile gliedern lassen (z.B. Lesekompetenz bezogen auf unterschiedliche Texte), wird die Abfrage der prozeduralen Metakognition entsprechend auch auf diese Teile bezogen. Die Erhebungsdauer beträgt im Klassenkontext für die allgemeine Abfrage 1 Minute, bei zusätzlichen textspezifischen Abfragen 2 Minuten. Bei Individualerhebungen ist mit durchschnittlich weniger als 30 Sekunden für die allgemeine Abfrage und mit 1,50 Minuten für die allgemeine Abfrage inklusive textspezifischer Fragen zu rechnen. In den Scientific Use Files werden für jede eingesetzte Frage zwei Werte zur Verfügung gestellt: einerseits der geschätzte Anteil der richtig gelösten Aufgaben und andererseits der geschätzte Anteil der richtig gelösten Aufgaben relativiert an der tatsächlichen Leistung im Kompetenztest (Lockl, 2015b).

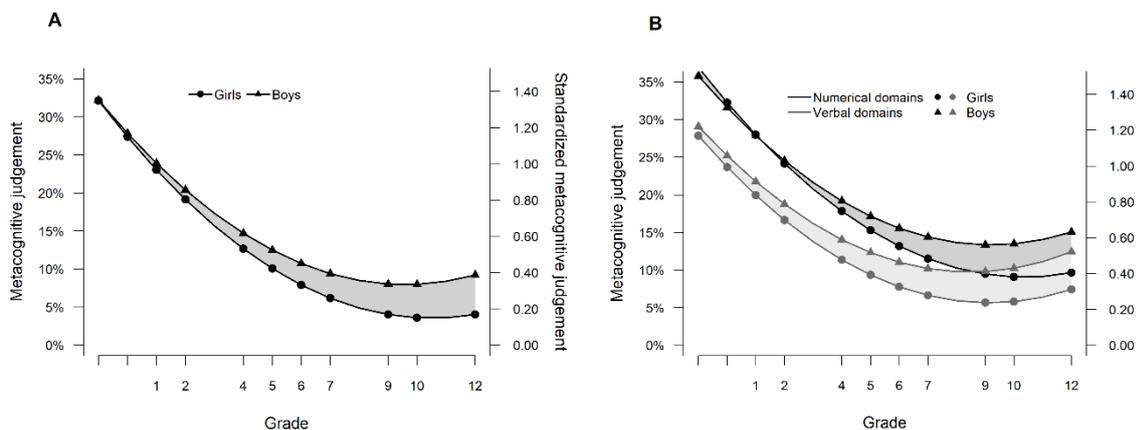
Abbildung 23 gibt am Beispiel der Startkohorte 2 einen Einblick in die deskriptiven Ergebnisse zu dem von den Schülerinnen und Schülern eingeschätzten Anteil der richtig gelösten Aufgaben. Während Kinder im Kindergartenalter im Mittel noch der Auffassung sind, fast alle (d.h. mehr als 90% der) Aufgaben richtig gelöst zu haben, nimmt dieser Anteil mit steigendem Alter ab und liegt am Ende der Grundschulzeit bei etwa 60%.



**Abbildung 23.** Einschätzungen zu den durchgeführten Kompetenztests (in Prozent).

Anmerkung: VO: Wortschatz, GR: Grammatik, S: Naturwissenschaften, MA: Mathematik, RX: frühe Lesekompetenz, RE: Lesekompetenz, ICT: ICT-Literacy

Wird die Differenz der vermuteten Anzahl der richtigen Aufgaben und der tatsächlich richtigen Anzahl der Aufgaben berechnet, zeigt sich in Einklang mit Befunden zur prozeduralen Metakognition (Schneider & Lockl, 2008), dass die Einschätzungen mit zunehmendem Alter realistischer werden, auch wenn sich selbst Erwachsene im Mittel eher über- als unterschätzen. Interessante Befunde ergeben sich im Hinblick auf den Zusammenhang von Überschätzung und Geschlecht. So zeigen Wolter, Lockl, Artelt und Gnambs (in prep.) anhand von Mehrebenen-Wachstumsmodellen, dass sich Jungen stärker überschätzen als Mädchen und dass diese Geschlechtsunterschiede im Verlauf der Entwicklung größer werden. Weiterhin hängt das Ausmaß der Überschätzung von der Domäne ab, wobei sich Personen im mathematisch-naturwissenschaftlichen Bereich stärker überschätzen als im sprachlichen Bereich (vgl. Abbildung 24). Entgegen den Erwartungen lässt sich hierbei jedoch kein differenzieller Effekt beobachten, d.h. das Ausmaß der Überschätzung ist bei Jungen im Vergleich zu Mädchen im mathematisch-naturwissenschaftlichen Bereich nicht höher als im sprachlichen Bereich.



**Abbildung 24.** Geschlechts- und domänenspezifische Unterschiede im längsschnittlichen Verlauf der Überschätzungen.

Anmerkung: A: Mittlere Veränderungen in den Überschätzungen über die Domänen hinweg; B: Verläufe für mathematisch-naturwissenschaftliche und sprachliche Domänen (Wolter, Lockl, Artelt & Gnambs, in prep.).

Generell bieten die Daten zur prozeduralen Metakognition eine ausgezeichnete Gelegenheit, die Fähigkeit zur Einschätzung der eigenen Leistung anhand des gleichen Indikators über die gesamte Lebensspanne hinweg zu untersuchen, was in der Literatur einzigartig ist. In diesem Kontext sind auch Zusammenhänge mit dem Selbstkonzept, Merkmalen der Persönlichkeit oder mit Motivation von großem Interesse. Beispielsweise stellt sich die Frage, welche Effekte eine überaus optimistische Selbsteinschätzung auf die Motivation und zukünftigen Kompetenzerwerb einer Person hat. Bezogen auf junge Kinder gehen verschiedene Autoren davon aus, dass die sehr positiven Selbsteinschätzungen von Kindern unter gewissen Umständen als durchaus adaptiv betrachtet werden können (Bjorklund, Gaultney & Green, 1993). Ob sich dies auch für die Daten im NEPS nachweisen lässt, ob dies für verschiedene Altersgruppen zutrifft und ob sich eine übermäßige Überschätzung ab einem gewissen Ausmaß als negativ für zukünftige Kompetenzentwicklungen herausstellt, sind interessante Forschungsfragen, die im Rahmen von NEPS untersucht werden können.

## 7.6 Ausgewählte Publikationen mit NEPS-Daten zur deklarativen und prozeduralen Metakognition

- Händel, M., Artelt, C. & Weinert, S. (2013). Assessing metacognitive knowledge: Development and evaluation of a test instrument. *Journal of Educational Research Online*, 5, 162–188.
- Händel, M., Lockl, K., Heydrich, J., Weinert, S. & Artelt, C. (2014). Assessment of metacognitive knowledge in students with special educational needs. *Metacognition and Learning*, 9(3), 333-352. <https://doi.org/10.1007/s11409-014-9119-x>
- Händel, M., Lockl, K., Heydrich, J., Weinert, S. & Artelt, C. (2015). Kompetenztestung bei Schülern und Schülerinnen mit Förderschwerpunkt Lernen – Effekte unterschiedlicher Testbedingungen am Beispiel eines Tests zum metakognitiven Wissen. In P. Kuhl, P. Stanat, H. A. Pant, M. Prenzel, B. Lütje-Klose & C. Gresch (Hrsg.), *Einbeziehung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen: Forschungsstand, Herausforderungen & Perspektiven* (S. 221-242). Wiesbaden: Springer VS.
- Lockl, K., Händel, M. & Artelt, C. (2018). Kompetenztestung bei Grundschulkindern: Differenzielle Effekte unterschiedlicher Testbedingungen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 50(1), 33-43.
- Lockl, K., Händel, M., Haberkorn, K. & Weinert, S. (2016). Metacognitive knowledge in young children: Development of a new test procedure for first graders. In H.P. Blossfeld, H.-P., J. von Maurice, M. Bayer & J. Skopek (Eds.), *Methodological Issues of Longitudinal Surveys: The Example of the National Educational Panel Study* (pp.465-484). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Miyamoto, A., Pfof, M. & Artelt, C. (2019). The relationship between intrinsic motivation and reading comprehension: Mediating effects of reading amount and metacognitive knowledge of strategy use. *Scientific Studies of Reading*, 23, 445-460, <https://doi.org/10.1080/10888438.2019.1602836>

## Literatur

- Annevirta, T., Laakkonen, E., Kinnunen, R. & Vauras, M. (2007). Developmental dynamics of metacognitive knowledge and text comprehension skill in the first primary school years. *Metacognition and Learning*, 2, 21–39.
- Artelt, C., Beinicke, A., Schlagmüller, M. & Schneider, W. (2009). Diagnose von Strategiewissen beim Textverstehen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 41(2), 96-103.
- Artelt, C., Neuenhaus, N., Lingel, K. & Schneider, W. (2012). Entwicklung und wechselseitige Effekte von metakognitiven und bereichsspezifischen Wissenskomponenten in der Sekundarstufe. *Psychologische Rundschau*, 63, 18-25. <https://doi.org/10.1026/0033-3042/a000106>
- Artelt, C., Schiefele, U., Schneider, W. & Stanat, P. (2002). Leseleistungen deutscher Schülerinnen und Schüler im internationalen Vergleich (PISA): Ergebnisse und Erklärungsansätze. *Zeitschrift für Erziehungswissenschaft*, 5, 6-27.
- Artelt, C. & Schneider, W. (2015 online). Cross-country generalizability of the role of metacognitive knowledge for students' strategy use and reading competence. *Teachers College Record*, 117 (1), 010304, 1-32.
- Artelt, C., Weinert, S. & Carstensen, C. H. (2013). Assessing competencies across the life span within the German National Educational Panel Study (NEPS)—Editorial. *Journal of Educational Research Online (JERO)*, 5, 5-14.
- Bjorklund, D. F. Gaultney, J. F. & Green, B. L. (1993). I watch, therefore I can do: The development of meta-imitation during the preschool years and the advantage of optimism about one's imitative skills. In R. Pasnak & M. L. Howe (Eds.), *Emerging themes in cognitive development* (Vol. 2, pp. 79-102). New York: Springer.
- Boekaerts, M. (1999). Self-regulated learning: Where we are today. *International Journal of Educational Research*, 31, 445-457.
- Brown, A. L. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In F. E. Weinert and R. H. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 65-116). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Ebert, S. (2011). *Was Kinder über die mentale Welt wissen - Die Entwicklung von deklarativem Metagedächtnis aus der Sicht der "Theory of Mind"*. Hamburg: Dr. Kovač.
- Ebert, S. (2015). Longitudinal relations between theory of mind and metacognition and the impact of language. *Journal of Cognition and Development*, 16, 559-586. <https://doi.org/10.1080/15248372.2014.926272>
- Efklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist*, 13, 277-287.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring. A new area of cognitive-developmental inquiry. *American Psychologist*, 34, 906-911.
- Haberkorn, K., Lockl, K., Pohl, S., Ebert, S. & Weinert, S. (2014). Metacognitive knowledge in children at early elementary school. *Metacognition and Learning*, 9, 239-263. <https://doi.org/10.1007/s11409-014-9115-1>
- Händel, M., Artelt, C. & Weinert, S. (2013). Assessing metacognitive knowledge: Development and evaluation of a test instrument. *Journal of Educational Research Online*, 5, 162-188.
- Lockl, K. (2012). *Assessment of declarative metacognition: Starting cohort 4 – ninth grade*. Bamberg: Otto-Friedrich-Universität, Nationales Bildungspanel.

- Lockl, K. (2013). *Assessment of declarative metacognition: Starting cohort 3 – sixth grade*. Bamberg: Otto-Friedrich-Universität, Nationales Bildungspanel.
- Lockl, K. (2015a). *Assessment of declarative metacognition: Starting cohort 2 – grade 1*. Bamberg: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Lockl, K. (2015b). *Assessment of procedural metacognition: Scientific use file 2015*. Bamberg: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Lockl, K. (2017). *Assessment of declarative metacognition: Starting cohort 2 – grade 3*. Bamberg: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Lockl, K. (2018). *Assessment of declarative metacognition: Starting cohort 3 – grade 9*. Bamberg: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Lockl, K. & Schneider, W. (2006). Precursors of metamemory in young children: The role of theory of mind and metacognitive vocabulary. *Metacognition and Learning, 1*, 15-31.
- Lockl, K. & Schneider, W. (2007). Knowledge about the mind: Links between theory of mind and later metamemory. *Child Development, 78*, 148-167.
- Miyamoto, A., Pfost, M. & Artelt, C. (2019). The relationship between intrinsic motivation and reading comprehension: Mediating effects of reading amount and metacognitive knowledge of strategy use. *Scientific Studies of Reading, 23*, 445-460, <https://doi.org/10.1080/10888438.2019.1602836>
- Nelson, T. O. & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 26, 125-173). New York: Academic Press.
- Paris, S. G., Lipson, M. Y. & Wixson, K. K. (1983). Becoming a strategic reader. *Contemporary Educational Psychology, 8*, 293-316. [https://doi.org/10.1016/0361-476X\(83\)90018-8](https://doi.org/10.1016/0361-476X(83)90018-8)
- Pintrich, P. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review, 16*, 385–407.
- Schlagmüller, M. & Schneider, W. (2007). *WLST 7-12 - Würzburger Lesestrategie-Wissenstest für die Klassen 7-12*. Göttingen: Hogrefe.
- Schneider, W. & Lockl, K. (2008). Procedural metacognition in children: Evidence for developmental trends. In J. Dunlosky & R. A. Bjork (Eds.), *A handbook of memory and metamemory* (391-409). Mahwah, NJ: Erlbaum.
- Schneider, W., Schlagmüller, M. & Visé, M. (1998). The impact of metamemory and domain-specific knowledge on memory performance. *European Journal of Psychology of Education, 13*, 91-103. <https://doi.org/10.1007/BF03172815>
- 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 (Eds.), *Education as a lifelong process: The German National Educational Panel Study (NEPS), Edition ZfE* (2. überarbeitete Aufl., S. 57-82). Wiesbaden: Springer. <https://doi.org/10.1007/978-3-658-23162-0>
- Winne, P. H. & Hadwin, A. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (277-304). Mahwah, NJ: Erlbaum.
- Wolter, I., Lockl, K., Artelt, C. & Gnambs, T. (in prep.). *The development of gender differences in task-related self-evaluations in language and mathematical-science domains*.

## 8. Lesegeschwindigkeit

Ilka Wolter & Karin Gehrler

### 8.1 Einleitung

Flankierend zum Lesekompetenztest, bei dem das verstehende Lesen im Vordergrund steht, wird im NEPS ein Indikator der Lesegeschwindigkeit erhoben, bei dem primär basale Lese-prozesse bzw. deren Automatisierung im Vordergrund stehen.

Die Erhebung der Lesegeschwindigkeit im NEPS erfasst demnach vorrangig basale Lese-prozesse, wie die Dekodiergeschwindigkeit, und fokussiert daher eher auf automatischen Lese-prozessen. Eine geringe automatische Dekodierfähigkeit führt zu einem schlechteren Textverständnis, vor allem bei Novizinnen und Novizen (Rosebrock & Nix, 2006). Aber auch in höheren Altersgruppen bestehen interindividuelle Unterschiede in der Lesegeschwindigkeit. Zudem sagt die Lesegeschwindigkeit Unterschiede in der Lesekompetenz, beispielsweise in der Klasse 9, vorher (Artelt, Stanat, Schneider & Schiefele, 2001).

Diese Befunde machen deutlich, dass es wichtig ist, im Rahmen des NEPS ein lesebezogenes Kontrollmaß bei Personen unterschiedlichen Alters zu erheben. Der Einbezug dieses Konstrukts erlaubt die Analyse von Entwicklungsverläufen in der domänenspezifischen Kompetenz unabhängig spezifischer Ausgangsvoraussetzungen in der Dekodiergeschwindigkeit. Aus diesem Grund wird dieses Konstrukt möglichst zu Beginn aller Startkohorten einmalig erhoben und dann ggf. im weiteren Verlauf nur noch in größeren Abständen wiederholt.

### 8.2 Bisherige Erhebung des Konstrukts

Tabelle 23 gibt einen Überblick über die Erhebungszeiten und –modi bezüglich der bislang durchgeführten Erhebungen zur Lesegeschwindigkeit.

Tabelle 23. Erhebungszeiten und –modi der für die Lesegeschwindigkeit gelaufenen bzw. aktuell laufenden Erhebungen

Startkohorte	Alter / Klasse	Erhebungsjahr	Erhebungszeit (in Minuten)	Erhebungsmodus
SC 1*	Klasse 2	2020	3.00	Computer (Online**)
SC 2	Klasse 2	2013	3.00	Papier
SC 3	Klasse 5	2010	2.00	Papier
SC 3	Klasse 9	2015	2.00	Papier
SC 4	Klasse 9	2010	2.00	Papier

Startkohorte	Alter / Klasse	Erhebungsjahr	Erhebungszeit (in Minuten)	Erhebungsmodus
SC 5	Studierende (1. Jahr)	2011	2.00	Papier
SC 6	Erwachsene	2010 & 2012	2.00	Papier

Anmerkungen. \*Diese Studie ist zum jetzigen Zeitpunkt noch nicht in den SUFs verfügbar; die Daten werden deshalb in diesem Dokument noch nicht berichtet. \*\* Coronabedingt musste von CBA auf telefonisch begleitete Online-Testung umgestellt werden.

Aktuell befindet sich das Konstrukt in der Umstellung von papier- auf computerbasierte Erhebung. Um Deckeneffekte zu vermeiden, wurde eine Erweiterung des bisherigen Instruments von 51 auf 72 Sätze eigenentwickelt. Dieses steht in der Computerversion für die höheren Kohorten für die Wiederholungsmessungen ab 2020 zur Verfügung.

### 8.3 Beispielitems

Die Lesegeschwindigkeit wird ab einem gefestigten Lesealter kohorteninvariant erhoben, d. h. über alle Startkohorten in den Altersgruppen ab Klasse 5 wird dasselbe Testinstrument eingesetzt. Dem eingesetzten Test, der in zwei Minuten von den Studienteilnehmerinnen und Studienteilnehmern bearbeitet wird, liegen die Testkonstruktionsprinzipien der beiden Salzburger Lesescreenings (z.B. Auer, Gruber, Mayringer & Wimmer, 2005) zugrunde. Das Testmaterial wurde für die Zwecke des Nationalen Bildungspanels jedoch neu konzipiert. Insgesamt werden den Studienteilnehmerinnen und Studienteilnehmern 51 bzw. 72 Sätze (erweiterte Version 2020) vorgelegt, die in der Regel allein auf Basis allgemeinen Weltwissens beantwortet werden können, also kein spezifisches inhaltliches Vorwissen voraussetzen (z.B. „Mäuse können fliegen“). Nach jedem Satz muss angekreuzt werden, ob der Satz inhaltlich zutreffend ist („richtig“) oder nicht („falsch“) (siehe Abb. 25). In der erweiterten computerbasierten Version wird die Antwort über die Tastatur eingegeben (siehe Abb. 26), die Instruktion erfolgt mittels eines Videos und es stehen Übungsisems zur Verfügung.

Im frühen Lesealter wurde in der SC 2, Klasse 2 das Testverfahren Salzburger Lese-Screening für die Schulstufen 1–4 von Mayringer und Wimmer (2003, mit freundlicher Genehmigung des Verlags Hans Huber) papierbasiert administriert (siehe Abb. 27). Für die SC 1 (Neugeborene), Klasse 2, wurde die verbesserte Nachfolgeversion, das Salzburger Lese-Screening für die Schulstufen 2–9 von Wimmer und Mayringer (2014/2016, mit freundlicher Genehmigung des Verlags Hogrefe) in Eigenregie in einer Computerversion erstellt und auf Tablet administriert. Die Antwort wird am Tablet vom Kind selbst angetippt (siehe Abb. 28), die Instruktion erfolgt mittels eines Videos und es stehen mehrere Übungsisems zur Verfügung. Coronabedingt wurde auf eine Online-Version umgestellt, sodass das Kind den Test auf dem haushaltseigenen Endgerät, also auch auf einem Laptop über die Tastatur, beantworten konnte. Das Kind wurde telefonisch von einer geschulten Interviewerin begleitet.

Bei der Bearbeitung des Tests unterscheiden sich Personen vorrangig danach, wie viele Sätze sie in der vorgegebenen Zeit bearbeiten können. Unterschiede zwischen Personen im Anteil falsch bearbeiteter Sätze sind aufgrund des inhaltlich wenig anspruchsvollen Materials zu vernachlässigen. Als Maß der Lesegeschwindigkeit wird die Zahl der innerhalb der zweiminütigen Bearbeitungszeit bzw. dreiminütigen Bearbeitungszeit (frühes Lesealter) richtig beurteilten Sätze ermittelt.

**Instruktion**  
 Auf den nächsten Seiten findest du eine Reihe von Sätzen. Der Inhalt der Sätze stimmt aber nicht immer. Deine Aufgabe ist es, bei jedem Satz durch Ankreuzen zu markieren, ob er wahr oder falsch ist. Die Sätze kommen dir vielleicht recht leicht, teilweise auch lustig vor. Es geht bei diesen Sätzen eher um Schnelligkeit und weniger um dein Wissen. Auf dieser Seite findest du zwei Beispielsätze, so dass du dich mit den Aufgaben vertraut machen kannst. Blättere erst weiter auf die nächste Seite, wenn du ein Zeichen vom Testleiter erhältst!

	richtig	falsch
In einer Garage findet man immer eine Badewanne.	<input type="checkbox"/>	<input type="checkbox"/>
Die Person, die bei einem Fußballspiel auf die Einhaltung der Regeln achtet, nennt man Schiedsrichter.	<input type="checkbox"/>	<input type="checkbox"/>

Abbildung 25. Beispielaufgaben zur eigenentwickelten, papierbasierten Erfassung der Lesegeschwindigkeit z.B. in SC3 (Zimmermann, Gehrler, Artelt & Weinert, 2012).

Beispiel:

Katzen gehören zu der Familie der Fische.



Falsch / nein





Ja / richtig

Abbildung 26. Beispielaufgabe zur computerbasierten Erfassung der Lesegeschwindigkeit (erweiterte Version) z.B. in SC8.

	richtig	falsch
Tee kann man trinken.	<input type="checkbox"/>	<input type="checkbox"/>
In der Wüste regnet es oft.	<input type="checkbox"/>	<input type="checkbox"/>
Erdbeeren sind ganz blau.	<input type="checkbox"/>	<input type="checkbox"/>

Abbildung 27. Übungsaufgaben zur papierbasierten Erfassung der Lesegeschwindigkeit (SLS 1–4) in SC2 (©Verlag Hans Huber, Hogrefe AG, Bern • Nachdruck und jegliche Art der Vervielfältigung verboten • Mit freundlicher Genehmigung des Verlages Hans Huber).



Abbildung 28. Beispielaufgabe zur tabletbasierten Erfassung der Lesegeschwindigkeit in SC1 (©Verlag Hans Huber, Hogrefe AG, Bern • Nachdruck und jegliche Art der Vervielfältigung verboten • Mit freundlicher Genehmigung der Hogrefe AG).

## 8.4 Empirische Analysen

### 8.4.1 Deskriptive Informationen

In der folgenden Tabelle 24 sind die deskriptiven Angaben über die Daten der Lesegeschwindigkeit für die bisher bereitgestellten Datensätze in SUF dargestellt. Es zeigt sich, dass dieses Konstrukt in den Erhebungen eine sehr geringe Ausfallquote innerhalb der Testdurchführungen aufweist, d. h. Versuchspersonen, die mit der Bearbeitung dieses zweiminütigen Tests begonnen haben, schließen die Testung mit validen Angaben ab. Außerdem lassen sich gute Reliabilitäten der Skala in den verschiedenen Altersbereichen nachweisen.

Tabelle 24: Deskriptive Informationen zu Erhebungen der Lesegeschwindigkeit

Startkohorte	Alter / Klasse	N (gesamt)	N (gültig)	% fehlend	M	SD	Reliabilität
SC 2	Klasse 2	6340	6114	3.60	21.36	9.26	.98
SC 3	Klasse 5	5740	5208	9.30	21.24	7.00	.98
SC 3	Klasse 9	4898	4888	< 1.00	34.37	8.15	.98
SC 4	Klasse 9	14539	14524	< 1.00	34.06	8.68	.98
SC 5	Studierende (1. Jahr)	5951	5946	< 1.00	42.01	6.75	.98
SC6	Erwachsene (W3)	7256	7246	< 1.00	35.97	9.29	.99 <sup>a</sup>
SC6	Erwachsene (W5-Erstbefr.)	3205	3205	0.00	34.58	9.10	.99 <sup>a</sup>

Bemerkung. N (gesamt) = Anzahl der Personen, bei denen der Test administriert wurde; N (gültig) = Anzahl der Personen mit gültigem Testwert, % fehlend = Anteil an Personen ohne gültigen Testwert, M = Mittelwert des Testwerts, SD = Standardabweichung des Testwerts, Reliabilität = Split-Half (gerade versus ungerade Items; Missings als falsch). <sup>a</sup> Die Reliabilität der Skala wurde in der SC 6 über beide Teilstichproben (ALWA+W1 Erstbefragte sowie W3-Erstbefragte) berechnet.

### 8.4.2 Psychometrische Informationen

Die Skala der Lesegeschwindigkeit zeigt in allen Alterskohorten eine gute Reliabilität und aufgrund erwartungskonformer Mittelwertsunterschiede für verschiedene Gruppen und Korrelationen mit inhaltlich relevanten Konstrukten kann eine gute Validität der Skala geschlossen werden.

Es zeigen sich erwartungsgemäß moderate bis hohe Korrelationen der Lesegeschwindigkeit mit der Lesekompetenz (im Sinne von Leseverständnis) einer Person. Die Lesegeschwindigkeit in der SC 3 korreliert moderat mit der Lesekompetenz zum selben Messzeitpunkt ( $r = .34$ ,  $p <$

.00). Ein vergleichbares Muster zeigt sich in der 9. Klasse in der SC 4 ( $r = .35, p < .00$ ), in der Gruppe der Studierenden in der SC 5 ( $r = .25, p < .00$ ) und auch noch einmal deutlich ausgeprägter im Erwachsenenalter in der SC 6 ( $r = .50, p < .00$ ).

In der Betrachtung der Lesegeschwindigkeit zeigen sich erwartungskonform zu *Beginn der Sekundarstufe I* in der 5. Klasse mittlere Korrelationen mit den Noten im Fach Deutsch ( $r = -.31, p < .00$ ) sowie mit den Noten im Fach Mathematik ( $r = -.22, p < .00$ ). Auch hängt die Lesegeschwindigkeit der Schülerinnen und Schüler positiv mit deren Selbstkonzepten im Bereich Lesen ( $r = .31, p < .00$ ) und im Fach Deutsch ( $r = .21, p < .00$ ) zusammen. Kinder mit höheren Werten in der Lesegeschwindigkeit weisen auch ein Jahr später zum Ende der 6. Klasse bessere Noten im Fach Deutsch ( $r = -.21, p < .00$ ) und im Fach Mathematik ( $r = -.12, p < .00$ ) auf. Zum *Ende der Sekundarstufe I* zeigen sich ebenfalls erwartungskonform positive Zusammenhänge zu den Noten im Fach Deutsch ( $r_{G8} = -.22, p < .00$ ), dem Selbstkonzept im Fach Deutsch ( $r = .21, p < .00$ ) und dem Interesse im Fach Deutsch ( $r = .14, p < .00$ ). Zudem weisen Schülerinnen und Schüler, die später einen Leistungskurs im Fach Deutsch wählen, in der 9. Klasse höhere Werte in der Lesegeschwindigkeit auf, als Schülerinnen und Schüler, die später einen Grundkurs im Fach Deutsch wählen ( $d = 0.22$ ). Wie sich bereits schon in der Lesekompetenz gezeigt hat, lassen sich auch für die Lesegeschwindigkeit kaum oder nur geringe Zusammenhänge zur Häufigkeit des Lesens der Schülerinnen und Schüler nachweisen (Klasse 5:  $r = .03, p < .05$ ; Klasse 9:  $r = .14, p < .00$ ). Im *Erwachsenenalter* zeigt sich ein positiver, wenn auch nicht allzu großer Zusammenhang der Lesegeschwindigkeit einer Person mit der Zeit, die sie täglich im Beruf und der Freizeit mit dem Lesen von Texten verbringt ( $r = .16, p < .00$ ). Dieser Effekt ist ebenso wie in den Ergebnissen zum Leseverständnis tendenziell höher ausgeprägt, wenn nur die Zeit des beruflichen Lesens berücksichtigt wird ( $r = .17, p < .00$ ), im Vergleich zum Lesen in der Freizeit ( $r = .05, p < .00$ ).

### 8.4.3 Gruppenunterschiede

Die nachfolgende Tabelle fasst die standardisierten Mittelwertsunterschiede (Cohen's  $d$ ) der Testwerte für zentrale Gruppierungsvariablen zusammen.

Es zeigt sich in allen Altersbereichen ein Vorteil für Personen mit hohem Sozialstatus (HISEI bzw. ISEI) im Vergleich zu Personen mit niedrigerem Sozialstatus (siehe Tabelle 25). Die Lesegeschwindigkeit korreliert bedeutsam mit dem sozio-ökonomischen Status der Person ( $r = .33, p < .00$ ; operationalisiert über den ISEI des aktuellen Berufs). Ebenfalls erwartungskonform zeigen sich Vorteile für Personen ohne Migrationshintergrund gegenüber Personen mit einem Zuwanderungshintergrund. Dieser Unterschied fällt in den Schulkohorten noch sehr gering aus und nimmt dann in den Erwachsenenstichproben deutlich zu. Insgesamt fallen die Unterschiede in den Gruppen in Bezug auf Sozialstatus und Schulform geringer aus als in der ebenfalls zur Sprachkompetenz gehörenden Lesekompetenz. Lediglich der Geschlechtsunterschied zugunsten der weiblichen Teilnehmerinnen fällt in diesem Konstrukt (d.h. Lesegeschwindigkeit) höher aus als in der Lesekompetenz (im Sinne von Leseverständnis).

Tabelle 25: Standardisierte Mittelwertsunterschiede (Cohen's d) der Testwerte für zentrale Gruppierungsvariablen in den unterschiedlichen Startkohorten und Altersgruppen (Lesegeschwindigkeit)

Startkohorte	Alter / Klasse	Geschlecht	Sozio-ökonomischer Status	Migrationshintergrund	Schultyp	Bildungsabschluss
		(Frauen vs. Männer)	(hoch vs. gering)	(ohne vs. mit)	(Gymnasium vs. Nicht-Gymnasium)	(Abitur vs. kein Abitur)
SC 2	Klasse 2	0.09	0.26	0.05		0.29
SC 3	Klasse 5	0.13	0.29	0.20	0.61	0.29
SC 3	Klasse 9	0.43	0.33	0.13	0.87	0.37
SC 4	Klasse 9	0.34	0.26	0.19	0.79	0.33
SC 5	Studierende	0.23	0.15	0.46		0.12
SC 6	Erwachsene <sup>a</sup>	0.29	0.53	0.40		0.65

Bemerkung. Positive Werte spiegeln einen höheren Mittelwert in der ersten Gruppe (z.B. „Frauen“) wider, während negative Werte einen höheren Mittelwert in der zweiten Gruppe (z.B. „Männer“) widerspiegeln. Sozioökonomischer Status = HISEI der Eltern bzw. eigener ISEI in SC6, Migrationshintergrund = ohne (= beide Elternteile in Deutschland geboren) vs. mit (= mind. ein Elternteil nicht Deutschland geboren) <sup>a</sup> In der SC 6 sind für den ersten Messzeitpunkt die Welle 3 sowie die Welle 5-Erstbefragten in den Analysen zusammengefasst. Interpretation der d-Werte: .2 = schwach, .5 = mittel, .8 = große Effekte (Cohen, 1988).

## Literatur

- Artelt C., Stanat P., Schneider W. & Schiefele U. (2001) Lesekompetenz: Testkonzeption und Ergebnisse. In: Baumert J. et al. (Hrsg.) *PISA 2000*. VS Verlag für Sozialwissenschaften. [https://doi.org/10.1007/978-3-322-83412-6\\_4](https://doi.org/10.1007/978-3-322-83412-6_4)
- Auer, M., Gruber, G., Mayringer, H. & Wimmer, H. (2005). *Salzburger Lesescreening für die Klassenstufen 5–8*. Göttingen: Hogrefe. [www.testzentrale.de](http://www.testzentrale.de)
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New Jersey: Lawrence Erlbaum Associates.
- Mayringer, H. & Wimmer, H. (2003). *Salzburger Lesescreening für die Schulstufen 1–4 (SLS 1–4)*. Bern: Hans Huber.
- Rosebrock, C. & Nix, D. (2006). Forschungsüberblick: Leseflüssigkeit (Fluency) in der amerikanischen Leseforschung und -didaktik. *Didaktik Deutsch*, 20, 90–112.
- Zimmermann, S., Gehrler, K., Artelt, C. & Weinert, S. (2012). *The assessment of reading speed in grade 5 and grade 9. Status: 2012* (NEPS research data paper). Retrieved from University of Bamberg, National Educational Panel Study website: [https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC4/1-0-0/com\\_rs\\_2012\\_en.pdf](https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC4/1-0-0/com_rs_2012_en.pdf)
- Wimmer, H. & Mayringer, H. (2014/2016). *SLS 2–9. Salzburger Lesescreening für die Schulstufen 2–9*. Bern: Hogrefe.

## 9. Kognitive Grundfähigkeiten

Shally Novita

### 9.1 Einleitung

Zusätzlich zu den domänenspezifischen Kompetenzen (bspw. in den Domänen Lesen, Mathematik und Naturwissenschaften) werden im Rahmen des NEPS auch domänenübergreifende, möglichst bildungsunabhängige Indikatoren der fluiden Intelligenz bzw. der kognitiven Mechanik bei Personen sehr unterschiedlichen Alters erhoben (Blossfeld, 2008). Dabei besteht das Ziel im NEPS nicht darin, kognitive Grundfähigkeiten im Sinne einer bedeutsamen Outcome-Variable zu erfassen; eher dienen die kognitiven Grundfähigkeiten als wichtige Kovariate, die selbst nicht längsschnittlich modelliert wird. Dennoch ist es wichtig, in allen Startkohorten regelmäßig, aber in größeren Abständen, Indikatoren domänenübergreifender intellektueller Fähigkeiten zu erheben, um die Entwicklung domänenspezifischer und bildungsabhängiger Kompetenzen angemessen modellieren zu können und dabei den inkrementellen Einfluss bildungsbezogener Faktoren von dem Einfluss weniger bildungsabhängiger, kulturunabhängiger Faktoren abgrenzen zu können.

Kognitive Grundfähigkeiten werden im Nationalen Bildungspanel auf der Grundlage der von Baltes, Staudinger und Lindenberger (1999) etablierten Unterscheidung von „kognitiver Mechanik“ und „kognitiver Pragmatik“ erfasst. Während erstere über möglichst bildungsunabhängige, neuartige und domänenunspezifische Aufgabeninhalte gemessen wird, bauen Aufgaben zur Messung der kognitiven Pragmatik auf erworbenen Fertigkeiten und erworbenem Wissen auf (Ackerman, 1987).

In Abgrenzung hiervon sollen die in diesem Abschnitt vorgestellten Tests grundlegende kognitive Fähigkeiten im Sinne der kognitiven Mechanik erfassen (Brunner, Lang & Lüdtker, 2014). Diese unterliegen zwar ebenfalls alterstypischen Veränderungen; im Unterschied zu stärker bildungs- und wissensbezogenen Kompetenzen erweisen sie sich aber als weniger kultur-, erfahrungs- und sprachabhängig und als stärker biologisch verankert. Sie bilden eine individuelle Grundlage und differenzierende Basisfunktion für den Erwerb bildungsabhängiger Kompetenzen.

Aus den Facetten der kognitiven Mechanik stechen zwei gängige Markiertvariablen besonders hervor: Wahrnehmungsgeschwindigkeit (WG) und schlussfolgerndes Denken (SF).

### 9.2 Bisherige Erhebung des Konstrukts

In der Erhebung der kognitiven Grundfähigkeiten werden drei strukturäquivalente, aber altersangepasste Testversionen differenziert: (1) Eine Version ist für Vorschülerinnen und Vorschüler, (2) eine Version für Schülerinnen und Schüler der 2. Jahrgangsstufe und (3) eine Version ist für Personen ab zehn Jahren (5. Klasse plus oder Erwachsene) vorgesehen (Lang, Kamin, Rohr, Stünkel & Williger, 2014).

Die Wahrnehmungsgeschwindigkeit wird in NEPS über den Bilder-Zeichen-Test (NEPS-BZT) und das schlussfolgernde Denken wird über den Matrizentest (NEPS-MAT) erfasst. Beim NEPS-

BZT geht es darum, in einer vorgegebenen Zeit einen visuellen Reiz wahrzunehmen und in eine motorische Reaktion umzusetzen. Die zeitliche Begrenzung der Bearbeitung führt dazu, dass bei schnellerer Wahrnehmungsgeschwindigkeit eine höhere Zahl korrekter Antworten produziert wird (Lang et al., 2014). Der NEPS-MAT erfasst in der Tradition der RAVEN-Matrizen (Raven, 1977) schlussfolgerndes Denken (reasoning). Die Regeln, nach denen die Aufgaben erstellt wurden, sind für alle Altersgruppen gleich. Allerdings war die Erstellung altersangepasster Testvarianten aufgrund der unterschiedlichen Leistungsfähigkeit der Altersgruppen notwendig (Lang et al., 2014). Die Ergebnisse beider Tests ergeben einen Schätzer für kognitive Grundfähigkeiten, der jedoch nicht mit dem Gesamtergebnis eines herkömmlichen Intelligenztests (IQ) gleichgesetzt werden kann.

Tabelle 26 informiert über die Erhebungszeiten und –modi bezüglich der bislang durchgeführten Erhebungen der kognitiven Grundfähigkeiten (siehe Haberkorn & Pohl, 2013) im NEPS.

Tabelle 26: Erhebungszeiten und –modi der für die kognitiven Grundfähigkeiten gelaufenen bzw. aktuell laufenden Erhebungen

<b>Startkohorte</b>	<b>Alter / Klasse</b>	<b>Subtest</b>	<b>Anzahl der Items</b>	<b>Erhebungszeit (in Minuten)</b>	<b>Erhebungsmodus</b>
SC 1	6 Jahre	NEPS-BZT	42	1.50	Computer (Tablet)
		NEPS-MAT	12	6.00	
SC 2	5 Jahre	NEPS-BZT	42	1.50	Papier
		NEPS-MAT	12	6.00	
SC 2	Klasse 2	NEPS-BZT	42	1.00	Papier
		NEPS-MAT	12	6.00	
SC 3	Klasse 5	NEPS-BZT	93	1.50	Papier
		NEPS-MAT	12	9.00	
SC 3	Klasse 9	NEPS-BZT	93	1.50	Papier
		NEPS-MAT	12	9.00	
SC 4	Klasse 9	NEPS-BZT	93	1.50	Papier
		NEPS-MAT	12	9.00	
SC 5	Studierende	NEPS-BZT	93	1.50	Papier / Computer / Online
		NEPS-MAT	12	9.00	
SC 6	Erwachsene	NEPS-BZT	93	1.50	Computer
		NEPS-MAT	12	9.00	

## 9.3 Beispielitems

### Wahrnehmungsgeschwindigkeit

Im NEPS wird die Wahrnehmungsgeschwindigkeit über den Bilder-Zeichen-Test (NEPS-BZT) erfasst. Dieser basiert auf einer Weiterentwicklung des Digit-Symbol-Tests (DST) aus den Intelligenztests der Wechsler-Familie durch Lang et al. (2007). Analog zu dieser Weiterentwicklung erfordert der NEPS-BZT die umgekehrte Leistung: nach einem Lösungsschlüssel zu den jeweils vorgegebenen Symbolen die richtigen Zeichen bzw. Zahlen einzutragen.

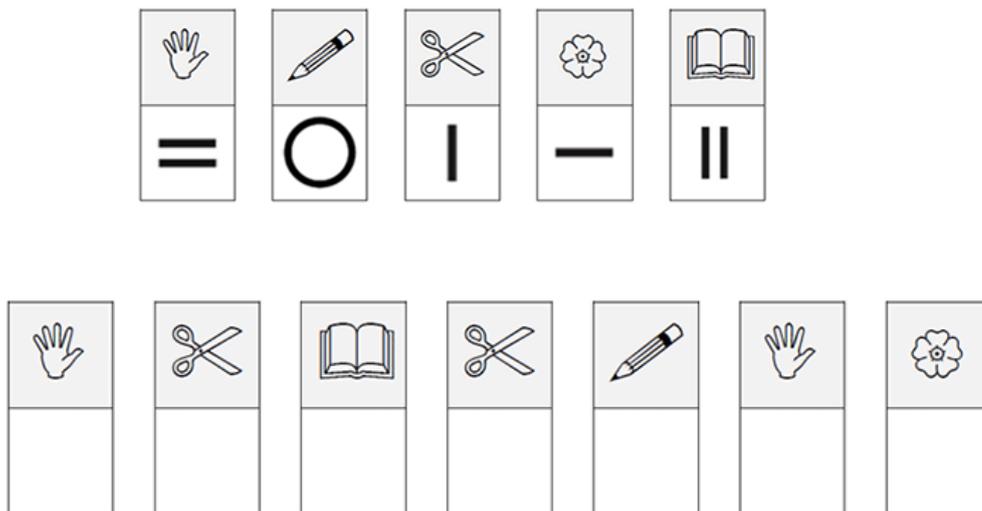


Abbildung 29. Beispielitem NEPS-BZT (Computerversion) für Vorschülerinnen und Vorschüler.

Am Anfang des Tests steht eine Beispielseite zum Einüben des Testprinzips, danach folgen zwei Testdurchgänge. Da für jede dieser Testseiten (außer der Beispielseite) vom Testleiter/ von der Testleiterin ein strenges Zeitlimit vorgegeben wird, werden diese als „Blöcke“ bezeichnet. Es wird jeweils klar instruiert, wann mit dem nächsten Aufgabenblock fortzufahren ist. Für die jüngsten Altersgruppen werden einfache geometrische Formen als Stimuli und einfache Zeichen als Targets gewählt (siehe Abbildung 29). Jeder Aufgabenblock umfasst in drei Zeilen insgesamt 21 Aufgaben. Es gibt zwei Aufgabenblöcke, also insgesamt 42 Items (siehe auch Lang et al., 2014).

### Schlussfolgerndes Denken

Zur Erfassung des schlussfolgernden Denkens werden im NEPS Matrizentests (NEPS-MAT) eingesetzt. Jede Matrizenaufgabe besteht aus mehreren waagrecht und senkrecht angeordneten Feldern, in denen verschiedene geometrische Elemente abgebildet sind – nur eines bleibt frei. Die logischen Regeln, denen die Anordnung der geometrischen Elemente folgt, müssen erschlossen werden, um aus den angebotenen Lösungen die richtige Ergänzung für das frei gebliebene Feld auswählen zu können.

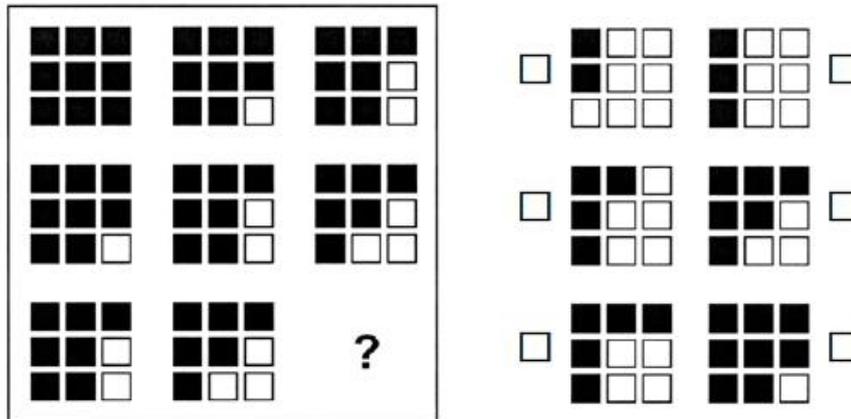


Abbildung 30. Beispielitem NEPS-MAT (Paper-Pencil-Version) für Personen ab zehn Jahren.

Am Anfang des Tests ist eine Seite mit zwei Beispielaufgaben enthalten, anhand derer der Testleiter bzw. die Testleiterin die Aufgabenstellung erklärt und einüben lässt. Da der NEPS-MAT auf die maximale Leistungsfähigkeit im Bereich des schlussfolgernden Denkens zielt und nicht auf die Schnelligkeit, mit der Lösungen produziert werden können, sind die Zeitlimits bei diesem Test möglichst großzügig bemessen. Erst nach deren jeweiligem Ablauf fordert der Testleiter bzw. die Testleiterin zur Bearbeitung des nächsten Aufgabenblocks auf. Es wird klar instruiert, wann mit der Aufgabe begonnen werden soll bzw. wann mit dem nächsten Aufgabenblock fortzufahren ist. Im NEPS-MAT für Erwachsene sind die Aufgaben auf drei Blöcke mit jeweils vier Items verteilt.

Die kognitiven Grundfähigkeiten werden ab SC 2 (Kindergarten) überwiegend über einen Paper-Pencil-Test erhoben. Ausgenommen hiervon wurden in der SC 6 (Erwachsene) und in einer randomisiert zugewiesenen Substichprobe der SC 5 (Studierende) diese Verfahren erstmalig computerbasiert im NEPS administriert. In der SC 1 wurden die kognitiven Grundfähigkeiten tabletbasiert im Individualfeld administriert.

## 9.4 Empirische Analysen

### 9.4.1 Deskriptive Informationen

Tabelle 27 berichtet die deskriptiven Informationen für die jeweiligen Erhebungen der kognitiven Grundfähigkeiten. Die Angaben für die SC 5 (Studierende) zum schlussfolgernden Denken (NEPS-MAT) beruhen auf der Erhebung dieser Subskala in beiden Modi: dem Paper-Pencil-Test und dem computerbasierten Test. Die Reliabilitäten zum schlussfolgernden Denken (NEPS-MAT) sind in dieser Gruppe sowohl getrennt ( $\alpha = .45$  für Paper-Pencil und  $\alpha = .37$  für CBA) als auch gemeinsam für beide Modi relativ gering. In Verbindung mit dem hohen Mittelwert und der daraus resultierenden geringeren Varianz in dieser Gruppe lassen diese Ergebnisse auf einen Deckeneffekt schließen.

Tabelle 27: Deskriptive Informationen zu Erhebungen der kognitiven Grundfähigkeiten

Startkohorte	Alter / Klasse		N (gesamt)	N (gültig)	% fehlend	M	SD	Reliabilität
SC 1	6 Jahre	NEPS-BZT	1989	1980	< 1.00	12.00	4.73	.78
		NEPS-MAT	1989	1987	< 1.00	6.43	2.64	.74
SC 2	5 Jahre	NEPS-BZT	2727	2707	< 1.00	17.81	6.09	.83
		NEPS-MAT	2727	2664	2.30	5.33	2.37	.63
SC 2	Klasse 2	NEPS-BZT	6122	6112	< 1.00	26.88	7.84	.71
		NEPS-MAT	6122	6102	< 1.00	6.77	2.62	.74
SC 3	Klasse 5	NEPS-BZT	5208	5204	< 1.00	44.03	13.42	.80
		NEPS-MAT	5208	5186	< 1.00	6.89	2.63	.66
SC 3	Klasse 9	NEPS-BZT	4599	4594	< 1.00	63.23	13.32	.79
		NEPS-MAT	4599	4593	< 1.00	9.21	2.25	.65
SC 4	Klasse 9	NEPS-BZT	14031	13975	< 1.00	59.13	13.99	.80
		NEPS-MAT	14031	13953	< 1.00	8.65	2.46	.67
SC 5	Studierende	NEPS-BZT (PAPI)	1374	1374	0.00	60.24	9.03	.76
		NEPS-BZT (CBA)	624	622	< 1.00	44.03	6.96	.79
		NEPS-MAT	1998	1996	< 1.00	10.36	1.48	.46
SC 6	Erwachsene	NEPS-BZT	8864	8815	< 1.00	32.07	9.39	.92
		NEPS-MAT	8864	8812	< 1.00	8.19	2.77	.74

Bemerkung. N (gesamt) = Anzahl der Personen, bei denen der Test administriert wurde; N (gültig) = Anzahl der Personen mit gültigem Testwert, % fehlend = Anteil an Personen ohne gültigen Testwert, M = Mittelwert des Testwerts, SD = Standardabweichung des Testwerts, Reliabilität = Cronbachs Alpha.

Aufgrund von unterschiedlichen Versionen des DGCF sind die statistischen Kennwerte über manche der Kohorten nicht in absoluten Werten vergleichbar. Ausführliche Informationen bezüglich der Itemanzahl und Erhebungszeit siehe Haberkorn und Pohl (2013).

In der Betrachtung der Ergebnisse fällt weiterhin auf, dass sich die Mittelwerte der kognitiven Grundfähigkeiten in der 9. Klasse in der SC 3 und der SC 4 bedeutsam unterscheiden (WG:  $t(df = 18517) = 17.43$ ,  $d = 0.26$ ; SD:  $t(18495) = 13.66$ ,  $d = 0.20$ ). Dieses Ergebnis kann darauf hinweisen, dass die Startkohorten über die Messzeitpunkte positiv selektiert sind, also die Stichprobenausfälle nicht zufällig und unabhängig von der Kompetenz der Personen erfolgen. Zu diesem Punkt gilt es zukünftige Panelausfälle genauer zu analysieren und einer zunehmenden Selektivität der Startkohorten über die Zeit bestmöglich entgegenzuwirken.

#### **9.4.2 Psychometrische Informationen**

In einer Validierungsstudie (Lang et al., 2007) wurde in einer Stichprobe von insgesamt 409 Teilnehmerinnen und Teilnehmern aus fünf Altersgruppen (Vorschule, 2. Klasse, 5. und 9. Klasse sowie Erwachsene) die konvergente und divergente Validität mit verschiedenen Untertests klassischer Skalen geprüft. Zu diesem Zweck wurden Untertests des HAWIVA-III und HAWIK-IV (Petermann & Petermann, 2008), der Raven Matrizen (SPM-C; Raven, 1977), dem CFT-1 (Cattell, Weiß & Osterland, 1997), dem Mehrfachwahl-Wortschatz-Intelligenz-Test (MWT-B; Lehrl, 1977) sowie dem Wortflüssigkeitstest des Leistungsprüfsystems (Horn, 1983) und den beiden Merkfähigkeitstests des IST-2000-R (Amthauer, Brocke, Liepmann & Beauducel, 2001) in Verbindung mit den NEPS-Tests zur Erfassung der kognitiven Grundfähigkeiten eingesetzt. Es konnte über alle Altersbereiche die konvergente und divergente Validität (Korrelationen zwischen  $r = -.08$  und  $.43$  für divergente und  $r = .42$  und  $.65$  für konvergente Validität) für die Erfassung der kognitiven Grundfähigkeiten im NEPS nachgewiesen werden.

Im NEPS wurden in der SC 3 die Wahrnehmungsgeschwindigkeit (NEPS-BZT) und der Test zum schlussfolgernden Denken (NEPS-MAT) bereits wiederholt in der Welle 1 (5. Klasse) und Welle 5 (9. Klasse) eingesetzt. Die Wahrnehmungsgeschwindigkeit für beide Messzeitpunkte korreliert zu  $r = .31$  ( $p < .01$ ) und das schlussfolgernde Denken zu  $r = .50$  ( $p < .01$ ). Die Ergebnisse zeigen, dass obwohl beide Konstrukte relativ stabil über einen Zeitraum von vier Jahren bleiben, das schlussfolgernde Denken im Vergleich zur Wahrnehmungsgeschwindigkeit tendenziell eine etwas höhere Stabilität aufweist.

#### **9.4.3 Gruppenunterschiede**

Die nachfolgende Tabelle zeigt anhand der Mittelwertsunterschiede, dass sich in der allgemeinen kognitiven Grundfähigkeit in Bezug auf das Geschlecht, den Migrationshintergrund und den Sozialstatus nur geringe, teilweise etwas inkonsistente Unterschiede zwischen Personengruppen nachweisen lassen. Die deutlichsten und konsistentesten Ergebnisse zeigen sich vor allem im schlussfolgernden Denken in Bezug auf verschiedene Schultypen: Hier lassen sich die größten Unterschiede zwischen Schülerinnen und Schülern der Hauptschule vs. des Gymnasiums (für NEPS-MAT liegen alle Effektstärken über 1) aufzeigen. Der negative Wert bedeutet hier, dass die Schülerinnen und Schüler am Gymnasium eine höhere Fähigkeit zum schlussfolgernden Denken aufweisen als Schülerinnen und Schüler an den Hauptschulen. In der Wahrnehmungsgeschwindigkeit fallen die Schulformeffekte erwartungsgemäß geringer aus. Auch zeigt sich in dieser Komponente der kognitiven Grundfähigkeit erwartungskonform kein Unterschied zwischen Personen mit und ohne Migrationshintergrund.

Tabelle 28: Standardisierte Mittelwertsunterschiede (Cohen's d) der Testwerte für zentrale Gruppierungsvariablen (kognitive Grundfähigkeiten)

Startkohorte	Alter / Klasse		Geschlecht	Sozio-ökon. Status	Migrationshintergrund	Schultyp	Bildungsabschluss
			(Frauen vs. Männer)	(hoch vs. gering)	(ohne vs. mit)	(Gymnasium vs. Nicht-Gymn.)	(Abitur vs. kein Abitur)
SC 1	6 Jahre	NEPS-BZT	.29	-.10	-.20	-	-.09
		NEPS-MAT	.09	.26	.15	-	.22
SC 2	5 Jahre	NEPS-BZT	.26	.11	.05	-	.07
		NEPS-MAT	.13	.14	.12	-	.22
SC 2	Klasse 2	NEPS-BZT	.12	.02	-.04	-	.01
		NEPS-MAT	.13	.25	.06	-	.27
SC 3	Klasse 5	NEPS-BZT	.19	-.01	-.09	0.20	.03
		NEPS-MAT	-.08	.35	.33	0.74	.39
SC 3	Klasse 9	NEPS-BZT	.29	.11	-.01	0.43	.20
		NEPS-MAT	.02	.37	.30	0.77	.41
SC 4	Klasse 9	NEPS-BZT	.34	.04	-.08	0.22	.05
		NEPS-MAT	-.05	.35	.35	0.82	.43
SC 5	Studierende	NEPS-BZT	.17	.00	.02	-	.01
		NEPS-MAT	-.34	.01	.03	-	.10
SC 6	Erwachsene	NEPS-BZT	.04	.44	.10	-	.45
		NEPS-MAT	-.27	.49	.13	-	.69

Bemerkung. Positive Werte spiegeln einen höheren Mittelwert in der ersten Gruppe (z.B. „Frauen“) wider, während negative Werte einen höheren Mittelwert in der zweiten Gruppe (z.B. „Männer“) widerspiegeln. Sozioökonomischer Status = HISEI der Eltern bzw. eigener ISEI in SC6, Migrationshintergrund = ohne (= beide Elternteile in Deutschland geboren) vs. mit (= mind. ein Elternteil nicht Deutschland geboren). Bildungsstatus: Bildungsstatus der Eltern (SC 1- SC 5) und der eigene (SC 6). Interpretation der d-Werte: .2 = schwach, .5 = mittel, .8 = große Effekte (Cohen, 1988).

## 9.5 Publikationen mit NEPS-Daten zu kognitiven Grundfähigkeiten

Gnambs, T. & Nusser, L. (2019). The longitudinal measurement of reasoning abilities in students with special educational needs. *Frontiers in Psychology, 10*. <https://doi.org/10.3389/fpsyg.2019.00232>

Nusser, L. & Messingschlager, M. (2018). *Erfassung kognitiver Grundfähigkeiten bei Schülerinnen und Schülern an Förderschulen in Startkohorte 4 (Klasse 9)* (NEPS Survey Paper, 33). Bamberg: Leibniz-Institut für Bildungsverläufe, Nationales Bildungspanel.

### Literatur

Ackerman, P. L. (1987). Individual differences in skill learning: An integration of psychometrics and information processing perspectives. *Psychological Bulletin, 102*, 3-27.

Amthauer, R., Brocke, B., Liepmann, D. & Beauducel, A. (2001). *Intelligenz-Struktur-Test 2000 R (I-S-T 2000 R)*. Göttingen: Hogrefe.

Baltes, P. B., Staudinger, U. M. & Lindenberger, U. (1999). Lifespan psychology: Theory and application to intellectual functioning. *Annual Review of Psychology, 50*, 471-507.

Blossfeld, H.-P. (2008). A proposal for a National Educational Panel Study (NEPS) in Germany. Part A: Overview. Bamberg: Bamberg University.

Brunner, M., Lang, F. R. & Lüdtke, O. (2014). Erfassung der fluiden kognitiven Leistungsfähigkeit über die Lebensspanne im Rahmen der National Educational Panel Study: Expertise (*NEPS Working Paper No. 42*). Bamberg: Leibniz-Institut für Bildungsverläufe, Nationales Bildungspanel.

Cattell, R. B., Weiß, R. H. & Osterland, J. (1997). *Grundintelligenztest Skala 1 (CFT 1)* (5. Auflage). Göttingen: Hogrefe.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New Jersey: Lawrence Erlbaum Associates.

Haberkorn, K. & Pohl, S. (2013). Cognitive basic skills – Data in the scientific use file. Bamberg: University of Bamberg, National Educational Panel Study.

Horn, W. (1983). *Leistungsprüfsystem (L-P-S)*. Göttingen: Hogrefe.

Lang, F. R., Kamin, S., Rohr, M., Stünkel, C. & Williger, B. (2014). Erfassung der fluiden kognitiven Leistungsfähigkeit über die Lebensspanne im Rahmen des Nationalen Bildungspanels: Abschlussbericht zu einer NEPS-Ergänzungsstudie (*NEPS Working Paper No. 43*). Bamberg: Leibniz-Institut für Bildungsverläufe, Nationales Bildungspanel.

Lang, F. R., Weiss, D., Stocker, A. & von Rosenblatt, B. (2007). Assessing cognitive capacities in computer-assisted survey research: Two ultra-short tests of intellectual ability in the Germany Socio-Economic Panel (SOEP). *Schmollers Jahrbuch. Journal of Applied Social Science Studies, 127*, 183-192.

Lehrl, S. (1977). *Mehrfachwahl-Wortschatz-Intelligenztest (MWT-B)*. Erlangen: Verlag Dr. med. D. Straube.

Petermann, F. & Petermann, U. (2008). *Hamburg-Wechsler-Intelligenztest für Kinder – IV (HAWIK-IV)* (2. Auflage). Bern: Huber.

Raven, J. C. (1977). *Standard progressive matrices: Sets A, B, C, D & E*. San Antonio, TX: Harcourt.