NEPS
National Educational Panel Study

Research Data

Information on Competence Testing

NEPS Starting Cohort 1 — Newborns

Education From the Very Beginning

Wave 10: 9 years

LIfBi
LEIBNIZ INSTITUTE FOR
EDUCATIONAL TRAJECTORIES

Information on testing	
Sample	Study B154, students in third grade (9 years), Starting Cohort 1, wave 10, year 2021. The survey started due to the pandemic at the beginning of March 2021 as a CAPI-by-Phone[1] interview. The interviewer conducted the biographical interview by phone from home. From April 2021, testing and initial interviewing of the child was administrated as CAPI[2] interview in the children's homes in the presence of their anchor person and the interviewer.
Test situation	**Computer-assisted telephone interview (CAPI-by-Phone) with technology based testing (TBT[3]) and a computer assisted self/web interview for the child (CASI[4]/CAWI[5])**
Test sequence	At the end of the biographical telephone interview (part one) with a legal guardian of the target child, consent was sought for the target child to complete the task and be interviewed in the household. If this was given, the testing and interviewing of the target child in the household was carried out on a tablet PC under strict hygiene conditions at a further appointment (part two). The target children themselves processed the competence tests and the survey directly on the tablet PC. The interviewer was responsible for the administration of the test transitions and partly for carrying out the instructions if these were not video-based. If there was no consent to visit the household, the children's survey could be completed online as a CAWI.
	Rotations
	The testing took place in the following order: 1. Vocabulary: Listening comprehension at word level[6] + procedural metacognition 2. Scientific literacy + procedural metacognition 3. Basic cognitive skills (nonverbal; NEPS-MAT)
Test duration (net processing time)	Approx. 29 minutes (45 minutes incl. instructions)
Administration time (incl. survey)	Approx. 70 minutes (approx. 45 minutes TBT-testing; approx. 10 minutes children survey, approx. 15 minutes preparation test situation) The biographical interview was conducted at a separate appointment before the child was tested and surveyed.

[1] CAPI-by-Phone = Computer Assisted Telephone Interview by CAPI interviewer
[2] CAPI = Computer Assisted Personal Interview
[3] TBT = Technology Based Testing
[4] CASI = Computer Assisted Self Interview
[5] CAWI = Computer Assisted Web Interview
[6] Lenhard, A., Lenhard, W., Segerer, R., & Suggate, S. (2015). *Peabody Picture Vocabulary Test-Revision IV German Adaption*, PPVT-IV. Frankfurt am Main, Germany: Pearson Assessment

Information on constructs				
Constructs	**Number of items**	**Allowed processing time**	**Survey mode**	**Next measurement (expected)**
Vocabulary: Listening comprehension at word level	Max. 19 sets with 12 tasks each (with termination criterion)	Approx. 20 min	CAPI (TBT)	tba
Scientific literacy	19	20 min	CAPI (TBT)	tba
Basic cognitive skills (nonverbal)				
Reasoning (NEPS-MAT)	2 x 6 = 12	2 x 3 min	CAPI (TBT)	tba
Domain-specific procedural metacognition regarding vocabulary: Listening comprehension at word level	1	1 min	CAPI (TBT)	tba
Domain-specific procedural metacognition regarding scientific Literacy	1	1 min	CAPI (TBT)	tba

Preliminary note

The development of the individual tests is based on framework concepts. They constitute overarching concepts on the basis of which education-relevant competences are to be shown consistently and coherently over the entire personal history. Therefore, the following framework concepts that served as a basis for the development of the test tools to measure the above-mentioned constructs are identical in the different studies.

Vocabulary: Listening comprehension at word level

Listening comprehension at word, sentence and text/discourse level as indicators of linguistic competence in German

The importance of linguistic competences for school learning and for explaining social disparities in school careers is largely undisputed.

In the National Educational Panel Study (NEPS), German linguistic competences are captured, on the one hand, via listening comprehension at word, sentence and text/discourse level and, on the other hand, (from 2nd primary school class onwards) via indicators of reading skills (reading competence [text comprehension], reading speed). In Starting Cohort 1 of the NEPS, from children aged 3 years, listening comprehension is solely captured at word level and later on, in primary school, via indicators of reading ability.

Listening comprehension at word level (receptive vocabulary)

Measures of receptive vocabulary are good and internationally applicable indicators of language skills and abilities acquired by children and adults. In numerous large international panel studies, passive vocabulary is collected as the central and sometimes sole indicator of cumulatively acquired linguistic-cognitive abilities taking into consideration individual basic skills (e.g., working memory capacity, speed variables) and environmental stimuli. Examples of such studies are the Head Start Family and Child Experiences Survey – FACES (USA)[7], the National Longitudinal Survey of Children and Youth – NLCSY (Canada; among others Lipps & Yiptong-Avila, 1999), the British Cohort Study – BCS70 (e.g., Bynner, 2004) or the European Child Care and Education (ECCE) Study, which is conducted in Germany, Austria, Spain and Portugal (e.g., European Child Care and Education (ECCE) Study Group, 1997).

The internationally most frequently used instrument for assessing receptive vocabulary is the Peabody Picture Vocabulary Test (PPVT), which is available in various versions (Dunn, 1959; Dunn & Dunn, 1981, 1997, 2007). Generally, the PPVT can be used for a very large age range (from 2.5 years to late adulthood) and is both easy to carry out and to analyze. A German version of the PPVT-IV (Dunn & Dunn, 2007; German version by Lenhard, Lenhard, Segerer, & Suggate, 2015) was used in Starting Cohort 1.

In the NEPS, the test was administered via a tablet PC. The children's task was to select one picture out of four, matching it to the word presented as an auditory cue by the tablet.

According to the guidelines of the PPVT-IV, the level of difficulty (test entry and termination) varies depending on the age and performance of the children. The practice unit at the start of the test also varies depending on the age and performance of the children. Once a child has solved at least two tasks correctly during the practice phase, he or she moves on to the test phase. The test consists of a total of 19 sets with staggered levels of difficulty, each set consisting of 12 items.

Testing procedure in this wave: The test begins with a practice phase consisting of at least two and a maximum of six tasks. The starting set depends on performance in the practice phase as well as on the age of the children. If the child makes more than one mistake in the starting set, the next lower set follows until a maximum of one mistake is made in one set (basal set). The test is then carried out until

the ceiling set – the set in which the child has made more than seven mistakes – has been identified; the sets that have already been processed are left out.

The Scientific Use File contains the number of administered practice items, all scored test items (correct/false), the basal set, and the ceiling set. In addition, the sum score is included, indicating the number of correctly solved items. All items that are in lower sets than the basal set are assumed to be correct.

References

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, D. M. (1981). *Peabody Picture Vocabulary Test-Revised (PPVT-R).* Circle Pines, MN: American Guidance Service.

Dunn, L. M. & Dunn, D. M. (1997). *Peabody Picture Vocabulary Test, Third Edition (PPVT-III).* Circle Pines, MN: American Guidance Service.

Dunn, L. M. & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4).* Upper Saddle River, NJ: Pearson.

European Child Care and Education (ECCE) Study Group. (1997). *European Child Care and Education.* Berlin, Germany: Freie Universität Berlin, Department of Education.

Lenhard, A., Lenhard, W., Segerer, R., & Suggate, S. (2015). *Peabody Picture Vocabulary Test-4.* Frankfurt am Main, Germany: Pearson Assessment.

Lipps, G. & Yiptong-Avila, J. (1999). From home to school: how Canadian children cope. *Education Quarterly Review, 6*(2), 51–57.

Scientific literacy

Scientific literacy is the precondition for participating in world affairs marked by science and technology (Prenzel, 2000; Prenzel et al., 2001; Rost et al., 2004) and is viewed as a predictor for an economically, socially and culturally successful life. Many problems and issues we encounter in our daily life require an understanding of natural sciences and technology. Scientific topics and problems affect all people. Therefore, current discussions on the goals of scientific education focus on the concept of scientific literacy for all people (Osborne & Dillon, 2008). Such literacy is the basis for lifelong learning, serves as a connection for further learning (OECD, 2006; Prenzel et al., 2007) and, thus, also influences professional careers.

Based on this, the NEPS definition of scientific literacy follows the Anglo-Saxon literacy concept (Bybee, 1997; Gräber, Nentwig, Koballa & Evans, 2002; OECD, 2006) which does not regard scientific competence as the simple reproduction of acquired knowledge but rather as the flexible use of acquired knowledge in different situations and contexts of daily life.

In the NEPS, scientific literacy is understood as the use of scientific knowledge in the environmental, technological and health contexts (Hahn et al., 2013). In addition, the concept distinguishes between content-related and process-related elements (see Fig. 1). Knowledge of science comprises content-related *matter*, *systems*, *development* and *interactions*. Knowledge about science includes *enquiry and scientific reasoning* which involve, among other things, checking hypotheses, interpreting findings as well as measuring principles and measuring error control.

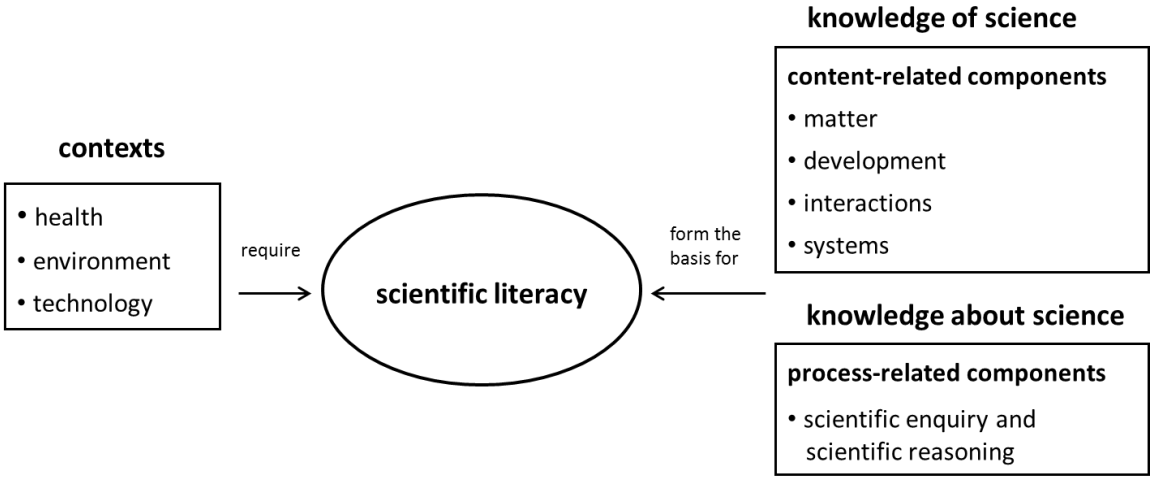

Fig.1. Application contexts as well as content-related and process-related elements of the NEPS scientific literacy test (Hahn et al., 2013).

To select its contexts as well as the content-related and process-related elements, the NEPS uses PISA (OECD, 2006), the *Benchmarks for Scientific Literacy* of the *American Association for the Advancement of Science* (AAAS, 2009) and the education standards of the Conference of Ministers of Education for the medium-level school-leaving qualification (KMK, 2005a, 2005b, 2005c) as a guideline. The selected contexts are of personal, social and global relevance. Considering current scientific research and the general events of the day, it is assumed that they will remain important across the entire life span of the test persons. Figure 2 provides an overview of the overlap of content-related components between PISA, the German educational standards and the NEPS. The selected content-related and process-related elements cover central concepts of all scientific disciplines.

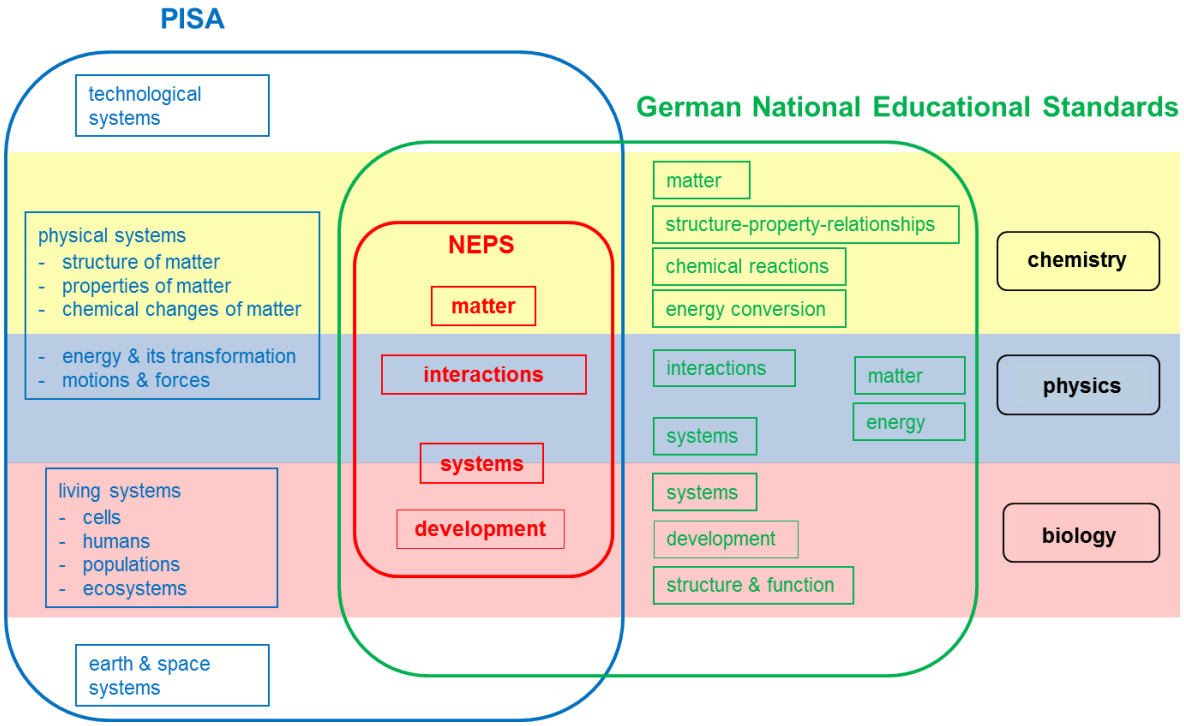

Fig.2. Overview of the overlap of content-related components between PISA, the German educational standards and the NEPS (Hahn et al., 2013).

The scientific literacy of 9-year-old children in Starting Cohort 1 of the NEPS is assessed using a tablet-based test. The test items are embedded in a "science and technology game", and a little dragon called "Nepsi" guides the children through the game. He reads the items and possible answers to them and then asks the children either to pick the right answer out of four pictures (multiple choice) or to judge whether successively shown pictures present a right or wrong answer (multiple true-false). The items consist of pictures and text so that the children can also read them themselves while Nepsi is presenting them.

In the end, one scientific literacy score is computed and published in the Scientific Use File for Starting Cohort 1.

References

American Association for the Advancement of Science. (1989). *Science for all Americans: A Project 2061 Report on goals in science, mathematics and technology*. Washington, DC: American Association for the Advancement of Science.

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

Bybee, R. W. (1997). Towards an understanding of scientific literacy. In W. Gräber & C. Bolte (Eds.). *Scientific literacy – An international symposium* (pp. 37–68). Kiel, Germany: Institute for Science and Mathematics Education (IPN).

Gräber, W., Nentwig, P., Koballa, T. & Evans, R. (Eds.). (2002). *Scientific Literacy. Der Beitrag der Naturwissenschaften zur Allgemeinen Bildung*. Opladen: Leske + Budrich.

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.

KMK (2005a). *Resolutions of the Conference of Ministers of Education: Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss*. Resolution of 16.12.2004. Munich, Germany: Luchterhand.

KMK (2005b). *Resolutions of the Conference of Ministers of Education: Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss*. Resolution of 16.12.2004. Munich, Germany: Luchterhand.

KMK (2005c). Resolutions of the Conference of Ministers of Education*: Bildungsstandards im Fach Physik für den Mittleren Schulabschluss.* Resolution of 16.12.2004. Munich, Germany: Luchterhand

OECD (2006). *Assessing scientific, reading and mathematical literacy. A framework for PISA 2006*. Paris: Organisation for Economic Co-Operation and Development.

Osborne, J. & Dillon, J. (2008). *Science education in Europe: Critical reflections. A report to The Nuffield Foundation*. London, United Kingdom: King's College.

Prenzel, M. (2000). Lernen über die Lebensspanne aus einer domänenspezifischen Perspektive: Naturwissenschaften als Beispiel. In F. Achtenhagen & W. Lempert (Eds.), *Lebenslanges Lernen im Beruf - seine Grundlegung im Kindes- und Jugendalter. Volume IV. Formen und Inhalte von Lernprozessen* (pp. 175–192). Opladen, Germany: Leske + Budrich.

Prenzel, M., Rost, J., Senkbeil, M., Häußler, P., & Klopp, A. (2001). Naturwissenschaftliche Grundbildung: Testkonzeption und Ergebnisse. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann & Weiß, M. (Eds.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (pp. 191–248). Opladen, Germany: Leske + Budrich.

Prenzel, M., Schöps, K., Rönnebeck, S., Senkbeil, M., Walter, O., Carstensen, C. H., & Hammann, M. (2007). Naturwissenschaftliche Kompetenz im internationalen Vergleich. In M. Prenzel, C. Artelt, J. Baumert, W. Blum, M. Hammann, E. Klieme & R. Pekrun (Eds.), *PISA 2006 – Die Ergebnisse der dritten internationalen Vergleichsstudie* (pp. 63–105). Münster, Germany: Waxmann.

Rost, J., Prenzel, M., Carstensen, C.-H., Senkbeil, M., & Groß, K. (Ed.). (2004). *Naturwissenschaftliche Bildung in Deutschland. Methoden und Ergebnisse von PISA 2000*. Wiesbaden, Germany: VS Verlag für Sozialwissenschaften.

Basic cognitive skills (nonverbal) – perceptual speed and reasoning

In the NEPS, basic cognitive skills are measured based on the differentiation between "cognitive mechanics" and "cognitive pragmatics" in accordance with Baltes, Staudinger, and Lindenberger (1999). While the former is measured using task contents with an approach that is as education-independent, new and domain-unspecific as possible, the tasks for measuring cognitive pragmatics are based on acquired skills and knowledge (Ackerman, 1987). Consequently, some of the domain-specific performance tests used within the framework of the NEPS may serve as indicators of pragmatics.

In contrast to this, the tests of basic cognitive skills aim at assessing individual differences in fluid cognitive abilities. While these abilities are subject to age-related changes, in comparison to education- and knowledge-related competences they have been proven to be less dependent on culture, experience and language. In this context, these tests provide an individual basis and a fundamental differentiating function for the acquisition of education-dependent competences.

Among the facets of cognitive mechanics, two common marker variables stand out: **perceptual speed** and **reasoning**.

Perceptual speed reflects the basal speed of information processing (*"speed"*). In the NEPS, this is measured using the **Picture Symbol Test (NEPS-BZT)** which is based on an improved version of the Digit-Symbol Test (DST) from the tests of the Wechsler family by Lang, Weiss, Stocker, and von Rosenbladt (2007). Analogously to this improved version, the NEPS-BZT requires the test person to enter the correct figures for the preset symbols according to an answer key.

Reasoning serves as a key marker of mental performance (Baltes et al., 1999). The **NEPS reasoning test (NEPS-MAT)** is designed as a matrices test, in line with the tradition of typical reasoning tests. Each item of the matrices test consists of several horizontally and vertically arranged fields in which different geometrical elements are shown – with only one field remaining free. The test person has to deduce the logical rules on which the pattern of the geometrical elements is based in order to be able to select the correct element for the free field from the solutions provided.

In this survey (wave 10 of starting cohort 1) only the reasoning test (NEPS-MAT) was used. The perceptual speed test is part of the survey program in wave 11.

Both tests were designed in such a way that they can be effectively used without requiring changes to the item sets across as many age groups as possible and relatively independently from the subjects' mother tongue. In Starting Cohort 1 the tests were administered as a computer-based assessment. The computer-based assessment was administered for the first time in Starting Cohort 1, Wave 7.
The results of both tests provide an estimator of basic cognitive skills which, however, is not directly comparable to the overall result of a traditional intelligence test (IQ). It can be used to control for differential initial capacities in the competence acquisition process.

References

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

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

Lang, F. R., Weiss, D., Stocker, A., & Rosenbladt, B. v. (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.

Metacognition

Metacognition is the knowledge and control of the own cognitive system. According to Flavell (1979) und Brown (1987), declarative and procedural aspects of metacognition are differentiated which are both covered in the National Education Panel.

Procedural metacognition

Procedural metacognition includes the regulation of the learning process through activities of planning, monitoring and controlling. Within the framework of NEPS in combination with the competence tests of the individual domains, the procedural aspect of metacognition is not assessed as a direct measure of such planning, monitoring and controlling activities but as a metacognitive judgement that refers to the control of the learning performance during (and/or shortly after) the learning phase (also see Nelson & Narens, 1990). After the study participants have taken their competence tests, they are requested to rate their own performance. They are asked to state the portion of questions presumably answered correctly. Kindergarten and elementary school children are shown a 5-point smiley scale to give their judgments.

Usually, one question is asked per domain. For competence domains that can be divided into coherent individual parts (e.g. reading competence referring to different texts), the inquiry of procedural metacognition is referred to these parts as well, which, of course, leads to a longer processing time.

References

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.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34*, 906-911.

Nelson, T.O. & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G.H. Bower (Ed.), *The psychology of learning and motivation* (pp. 125-141). New York: Academic Press.