NEPS
National Educational Panel Study

Research Data

Information on Competence Testing

NEPS Starting Cohort 8 — Grade 5

Education for Tomorrow's World

Wave 1: Grade 5

LIfBi
LEIBNIZ INSTITUTE FOR
EDUCATIONAL TRAJECTORIES

Leibniz Institute for Educational Trajectories (LIfBi)
Wilhelmsplatz 3, 96047 Bamberg
Director: Prof. Dr. Cordula Artelt
Administrative Director: Dr. Stefan Echinger
Bamberg; April 10, 2025

Information on testing	
Sample	Study A104, Starting Cohort 8, Grade 5, Survey wave 1, Year 2022
Test situation	Group testing, normally taking place in the classroom with two test instructors for each group.
Test sequence	The survey took place on one day. The students completed a paper-based NEPS test booklet on reading competence and a paper-based NEPS test booklet on mathematical competence. After a 15-minute break, the students worked on a test booklet based on the Educational Standards (BiSta) for the primary sector (either mathematics or reading). The rotation of which NEPS test booklet each student started with was randomly assigned. Prior to the start of the tests, the students were given general instructions on paper about the task formats. After completing the tasks, there was a 15-minute break. Finally, the students completed a computer-assisted self-administered questionnaire (CASI). The test sequence was randomly assigned. **Test sequence regular school** - Rotation A test booklet: 1. NEPS reading competence; 2. NEPS mathematical competence; 3. BiSta mathematics or BiSta reading competence - Rotation B Test Booklet: 1. NEPS mathematical competence; 2. NEPS reading competence; 3. BiSta mathematics or BiSta reading - Student questionnaire **Test sequence special school** - Test booklet: NEPS reading competence - Student questionnaire
Test duration (net processing time)	Regular school: 96 min (including student questionnaire 40 min) Special school: 58 min (including student questionnaire 30 min)

Breaks	2x 15 min
Administration time	Regular school: approx. 186 min; special school: approx. 138 min

Information on the individual tests

Construct	Number of Items	Allowed Processing Time	Survey Mode	Next Measurement
NEPS Reading competence	33	28 min	paper-pencil	2024
NEPS Mathematical competence	24	28 min	paper-pencil	2024
BiSta The educational standards for reading comprehension in primary education	81	40 min	paper-pencil	2027
BiSta The educational standards for mathematics for primary education	75	40 min	paper-pencil	2027

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.

The NEPS test booklets for reading competence and mathematical competence in Grade 5 of Starting Cohort 8 (Study A104) were identical to those in Grade 5 of Starting Cohort 3 (Study A28).

Reading competence

The ability to understand and use written texts is an important precondition for further developing personal knowledge and personal skills and a prerequisite for participating in cultural and social life. Manifold areas of knowledge and life are made accessible through reading. The range of reading occasions is very wide, and reading fulfills many different functions (cf. Groeben & Hurrelmann, 2004). They range from reading for expanding knowledge, which is crucial for further education, to lifelong learning as well as literary-esthetic reading. Not only do texts convey information and facts, but they also transfer ideas, moral concepts, and cultural contents. Accordingly, the concept of reading competence in the National Education Panel incorporates functional understanding as a basis for reading competence, as is also reflected in the Anglo-Saxon *Literacy* Concept (see also OECD, 2009), with a focus on competent handling of texts in different typical everyday situations.

In order to represent the concept of reading competence over the entire life span as coherently as possible, three characteristic features are specified in the framework concepts of the NEPS reading competence tests. They are considered in the following age- and stage-specific test forms:

1. text functions, text types,
2. comprehension requirements,
3. task formats.

1. Text functions/text types

The NEPS distinguishes between five text functions and associated text types, which are represented in each version of the test: a) factual texts, b) commenting texts, c) literary texts, d) instructions, and e) advertising texts (Gehrer, Zimmermann, Artelt, & Weinert, 2013). This selection is based on the assumption that these five text functions have practical relevance for the various age backgrounds of the participants. The text functions and/or text types (see Gehrer & Artelt, 2013) can be characterized as follows:

Texts conveying factual information represent basic texts for learning, fundamental acquisition of knowledge, and extraction of information; examples of these are: articles, reports, reportages, and announcements. Texts with a commenting function are texts in which a stand is taken or contradictive arguments are discussed and in which reflection is integrated. Examples of such texts are cleverly worded essays or humorous comments, which are implemented in tests for college students and adult cohorts. In school cohorts, a text with a discussion about the pleasures and disadvantages of smoking may be used, for example. The literary-esthetic function of texts is included in the third category, which encompasses short stories and extracts from novels or stories. Specific literary text types such as stage plays, satires, or poems are excluded as a result of their specific reception, which is presumably strongly dependent on educational track and curriculum. The fourth category comprises text types that are product inserts such as building and assembly instructions, package inserts for medication, work instructions, and cooking recipes. The fifth category (appeals, advertisements, notifications) includes text types such as job advertisements and recreation programs.

The five selected text functions and their associated text types are implemented in each test booklet over the life span as a longitudinal concept, which means that each test/each test booklet for measuring reading competence contains five texts corresponding to the five text functions. Unlike the PISA studies, the NEPS does not include discontinuous texts such as graphs, tables, and road maps. Discontinuous texts are excluded from the NEPS concept as they place special demands on readers, which are not always meaningful for each age group in which reading competence is measured.

Age-specific selection (text complexity, topic selection/task requirements):

For each age cohort, texts are selected according to their thematic orientation as well as their lexical, semantic, and grammatical properties which have to be appropriate for the respective group of readers.

The growth of reading competence from childhood to early adulthood is taken into account by increasing the text complexity (larger vocabulary, longer words, foreign words, higher complexity of sentence structures) and the basic length of texts. In addition, texts are selected on topics that correspond to and are appropriate for the environment of the respective age group. They cover a wide spectrum of topics ranging from animals (for children) to social and philosophical questions related to the meaning of life for adults. Additionally, the test material is adjusted to the respective age group through age-adapted phrasing of the questions, the answer options, and the comprehension requirements of the tasks.

2. Comprehension requirements / task types

From the literature on reading competence and text comprehension (e.g., Kintsch, 1998; Richter & Christmann, 2002), it is possible to derive different types of comprehension requirement which are reflected in the NEPS concept in three specific requirement types of tasks (task types). The variants are called *types* as there is no explicit assumption that the tasks of one type are necessarily more difficult or easier than tasks of another type (Gehrer, Zimmermann, Artelt, & Weinert, 2013).

For tasks of the first type ("finding information in the text"), detailed information must be identified at sentence level; in other words, the reader is required to decipher words and recognize statements or propositions. For tasks on this requirement cluster, the wording of the information needed to solve the respective tasks is either contained in the text and identical with the task itself, or the phrasing varies slightly.

In the case of the second task type ("drawing text-related conclusions"), conclusions have to be drawn from several sentences that have to be related to each other in order to extract local or global coherence. In some cases, the relevant sentences are located closely together. In others, several sentences are spread over entire sections. In another form of this task type, the reader has to understand the thoughts expressed in the entire text, which requires the comprehension and integration of larger and more complex text portions.

For the third type, the main requirement involves "reflecting and assessing", which is often linked to the mental representation of the text in a situation model in literature. In one version of this task type, the task is to understand the central idea, the main events, or the core

message of text, whereas in another version the purpose and intention of a text have to be recognized or the readers are asked to assess the credibility of a text.
The different comprehension requirements can be found in all text functions and are considered in the respective test versions in a well-proportioned ratio. (cf. Fig. 1.).

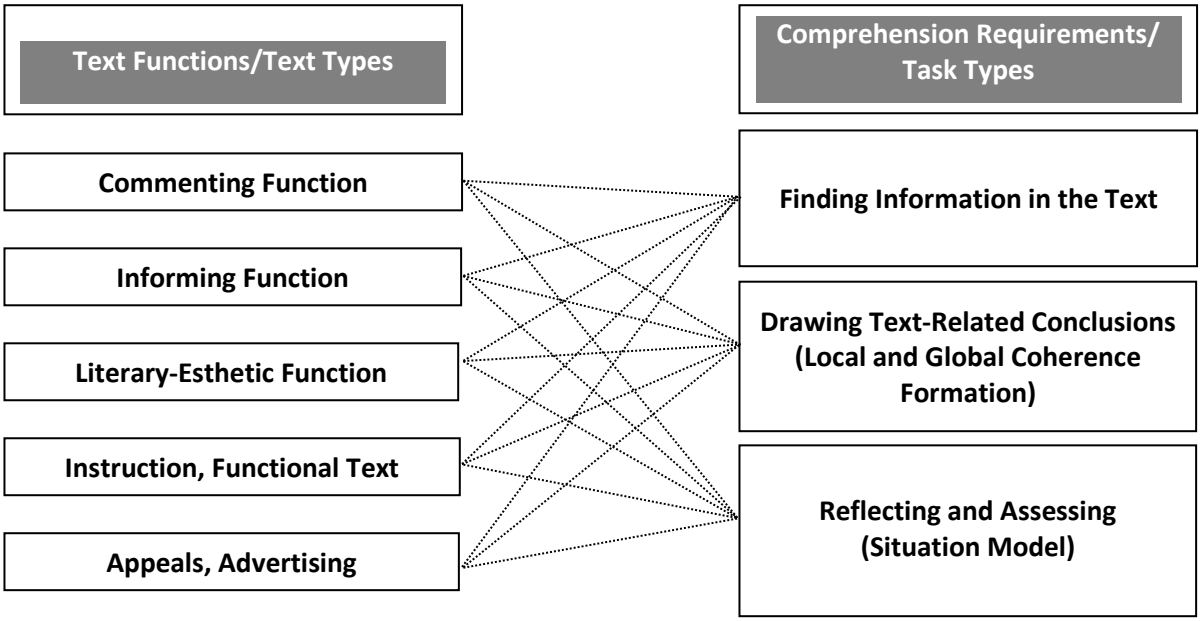

Fig. 1: Text functions and comprehension requirements (cf. Gehrer, Zimmermann, Artelt, & Weinert, 2013, p. 63)

3. Task formats

The majority of tasks have a multiple-choice format. This tasks format consists of a question/assignment about a text for which four answers are offered, one of which is the correct answer. As another task format, decision-making tasks are used, which require readers to judge individual statements and state whether they are right or wrong according to the text. So-called matching tasks represent a third format in which, for example, a subtitle must be chosen and assigned to different sections of a text. For tasks of the second and third formats, summaries are made, if necessary, thus creating answers with partly correct solutions (partial-credit items). Because surveys have been implemented repeatedly since 2016, further formats are administered within computer-based tests, for example, for college students (SC5), adults (SC6), and young adults (SC4). One of these formats is text enrichment tasks, in which the subjects have to insert three or four additional sentences into appropriate places in the given texts (for description, see: Rohm, Scharl, Ettner, & Gehrer, 2019). Furthermore, highlighting tasks are in preparation (Heyne, Artelt, Gnambs, Gehrer, & Schoor, 2020), in which subjects have to mark text passages in order to answer given questions about the texts.
By systematically considering different text functions which are implemented in different age groups in realistic and age-adapted texts with appropriate text themes and different

comprehension requirements, it is possible to operationalize reading competence as a comprehensive ability construct.

4. Scaling of items

Items of several task formats have been Rasch-scaled and longitudinally linked (Fischer, Rohm, Gnambs, & Carstensen, 2016). In addition, partial-credit items have been calculated based on the answers on decision-making tasks and matching tasks. Therefore, subjects´ answers to the tasks are aggregated in one score and are not used as single items. The quality criteria and psychometric characteristics of the items are presented in the technical reports of the different starting cohorts (*SC3*: Krannich, Jost, Rohm, Koller, Carstensen, Fischer & Gnambs, 2017; Pohl, Haberkorn, Hardt & Wiegand, 2012; Scharl, Fischer, Gnambs, & Rohm, 2017).

Bibliography

Fischer, L., Rohm, T., Gnambs, T., & Carstensen, C. H. (2016). *Linking the data of the competence tests* (NEPS Survey Paper No. 1). Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study. https://www.neps-data.de/Portals/0/Survey%20Papers/SP_I.pdf

Gehrer, K. & Artelt, C. (2013*).* Literalität und Bildungslaufbahn: Das Bildungspanel NEPS*.* In A. Bertschi-Kaufmann, & C. Rosebrock (Hrsg.). *Literalität erfassen: bildungspolitisch, kulturell, individuell* (S. 168–187). Weinheim, Germany: Juventa.

Gehrer, K., Zimmermann, S., Artelt, C. & Weinert, S. (2012). *The assessment of reading competence (including sample items for grade 5 and 9).* Scientific Use File 2012, Version 1.0.0. Bamberg: University of Bamberg, National Educational Panel Study. 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. https://www.waxmann.com/artikelART102722

Groeben, N. & Hurrelmann, B. (Hrsg.) (2004). *Lesesozialisation in der Mediengesellschaft: Ein Forschungsüberblick*. Weinheim: Juventa.

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.

Richter, T. & Christmann, U. (2002). Lesekompetenz: Prozessebenen und interindividuelle Unterschiede. In N. Groeben, B. Hurrelmann (Hrsg.), *Lesekompetenz: Bedingungen, Dimensionen, Funktionen* (S. 25–58). Weinheim: Juventa.

Technical Reports

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.

Mathematical competence

In the National Education Panel Study, the construct of *mathematical competence* is based on the idea of *mathematical literacy* as was defined, for example, in PISA. Thus, the construct describes "[…] an individual's capacity to identify and understand the role that mathematics plays in the world, to make well-founded mathematical judgments and to use and engage with mathematics in ways that meet the needs of that individual's life as a constructive, concerned and reflective citizen." (OECD, 2003, 24). Regarding younger children, this idea refers to competent handling of mathematical problems in *age-specific contexts.*

Accordingly, mathematical competence in NEPS is operationalized by items assessing more than pure mathematical knowledge; instead, solving the items requires recognizing and flexibly applying mathematics in realistic, mainly extra-mathematical situations.

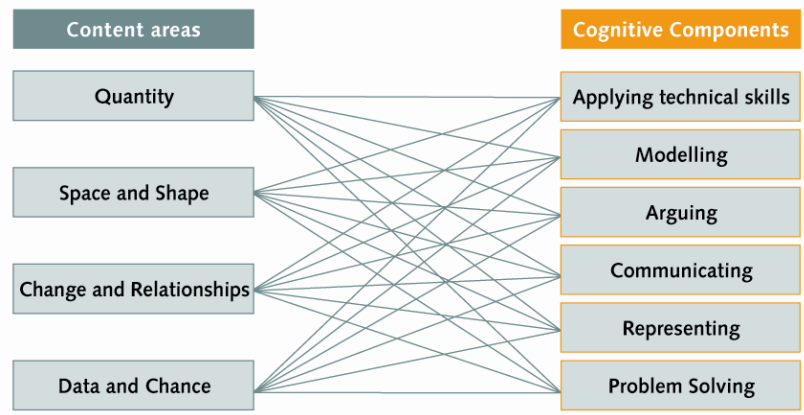

Fig. 1: Framework of mathematical competence in NEPS

The NEPS framework of mathematical competence distinguishes between content-related and process-related components (cf. Fig. 1). In detail, the content areas are characterized as follows:

- **Quantity** comprises all kinds of quantifications when numbers are used to organize and describe situations.
  Examples from the *elementary sector*: comparisons of sets, counting (ordinal/cardinal aspects of numbers), simple operations (e.g., adding)
  Examples from the *adult sector*: calculations of percentages and interests, calculations of area and volume, use of different units, simple equation systems
- **Space and Shape** includes all types of planar and spatial configurations, shapes or patterns.
  Examples from the *elementary sector*: recognizing geometric shapes, simple properties of shapes, perspective
  Examples from the *adult sector*: three-dimensional mathematical objects, geometric mappings, elementary geometric theorems
- **Change and Relationships** includes all kinds of (functional) relationships and patterns.
  Examples from the *elementary sector*: recognizing and continuing patterns, relationships

among numbers, proportionality
Examples from the *adult sector*: interpreting curves or function graphs, properties of linear, quadratic, and exponential functions, extremum problems

- **Data and Chance** comprises all situations involving statistical data or chance.
  Examples from the *elementary sector*: intuitively assessing probabilities, collecting and structuring data
  Examples from the *adult sector*: interpreting statistics, basic statistical methods, calculating probabilities

The cognitive components of mathematical thinking processes are distinguished as follows:

- **Applying technical skills** includes using known algorithms and remembering mathematical knowledge or calculation methods.
- **Modelling** includes the representation in a situation model and in a mathematical model as well as interpreting and validating results in real-life situations.
- **Arguing** includes assessing explanations and proofs, but also developing own explanations or proofs.
- **Communicating** requires communication on mathematical contents and includes, among other things, the correct and adequate use of mathematical technical terms.
- **Representing** comprises the use and interpretation of mathematical representations such as tables, charts or graphs.
- **Problem Solving** takes place, when there is no obvious approach, and, therefore, includes systematic trying, generalizing or examining special cases.

The test items used in NEPS refer to one content area that is mainly addressed by the item but may well contain several cognitive components (further description of the framework in Neumann et al., 2013). This differentiation renders the framework concept of mathematical competence in NEPS compatible with both the PISA studies and the German National Mathematics Education Standards. Some literature also show a high correlation between NEPS, the PISA studies and federal states comparisons from the Institute of Educational Quality Improvement (IQB): r = .89 for NEPS-PISA and r = .91 for NEPS-IQB (van den Ham, 2016).

Bibliography

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. Retrieved http://journal-for-educational-research-online.com/index.php/jero/article/view/362.

Organisation for Economic Co-Operation and Development [OECD] (2003). *The PISA 2003 assessment framework – mathematics, reading, science and problem solving knowledge and skills*. Paris: OECD.

Van den Ham, A.-K. (2016). *Ein Validitätsargument für den Mathematiktest der National Educational Panel Study für die neunte Klassenstufe.* Unpublished doctoral dissertation, Leuphana University Lüneburg, Lüneburg.

German Reading comprehension based on the educational standards for primary education

In 2004, binding educational standards for the subject German in primary education were introduced by the Conference of Ministers of Education (KMK 2005). The following competency areas for the subject German were distinguished[1]:

- Speaking and Listening: Engaging in conversations, listening comprehension, acting out scenes, and talking about learning.
- Writing: Possessing writing skills, writing correctly, planning, writing, and revising texts.
- Reading – Engaging with Texts and Media: Possessing reading skills and experiences, understanding and presenting texts.
- Language and Language Use: Knowing and using basic linguistic structures and terms, investigating linguistic communication, working on words, sentences, and texts, and discovering similarities and differences between languages.

For the implementation in Starting Cohort 8, only tasks related to the competency area of *reading* were used, as reading comprehension in German is considered a key qualification. Therefore, all further explanations refer to this domain. The following presentation focuses on a brief description of the tasks used. Explanations of the educational standards, the underlying concept of proficiency, and proficiency levels can be found in the reports for the *IQB trends in student achievement studies* (see particularly Pant et al. 2017; Stanat et al. 2017).

Operationalisation

Reading competencies were assessed using tasks from the *IQB trends in student achievement studies*. The tasks aim to assess the extent to which the participating pupils understand texts in an age-appropriate manner. The tasks consist of a stimulus and corresponding questions (items). The stimulus is typically half to one and a half pages long, with both literary and non-literary texts (non-fiction texts) being used. Both continuous texts (e.g., descriptions) and non-continuous texts (e.g., tables) were presented for non-fiction texts. The items use closed, semi-open, and open response formats. In closed formats, answers must be marked, underlined, or ordered. Semi-open and open formats are similar. Both require children to formulate their own responses

[1] In 2022, the KMK adopted further developed educational standards for the subject of German, but these must first be implemented in schools. The SC8 surveys are therefore based exclusively on the standards from 2004.

without being given options. While some questions require just a single word as an answer, others may require several sentences. All tasks were developed at the Institute for Educational Quality Improvement (IQB) by teachers under the guidance of experts in didactics and were thoroughly tested before use.

Scaling of the items

Eight test booklets were used in a multi-matrix design. After recoding and aggregation, a total of 81 items were included in the analyses. The preparation of the data and the estimation of the item and person parameters, based on a Rasch model, follows the procedure for *IQB trends in student achievement studies* (see Sachse et al., 2022, for a detailed description). First, the item parameters were estimated (calibration of the items) and then the students were placed on the common difficulty and competence scale (estimation of the person parameters). Students ' background characteristics were not considered when estimating the Weighted Likelihood Estimates (WLEs) provided in the data set.

Literature

KMK - Secretariat of the Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany. (2005). *Educational standards in the subject German for primary education.* Resolution of 15 October 2004. Munich: Luchterhand.

Pant, H.A., Böhme, K., Stanat, P., Schipolowski, S. & Köller, O. (2017). *Competence level models for the primary sector.* In P. Stanat, S. Schipolowski, C. Rjosk, S. Weirich & N. Haag (Eds.) IQB- Bildungstrend 2016 Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich (pp. 45-52). Waxmann.

Sachse, K.A., Weirich, S., Becker, B., Hafiz, N.J., Schneider R. & Schipolowsk, S. (2022). *Methodological principles: Design, implementation and evaluation of the IQB Education Trend 2021.* In P. Stanat, S. Schipolowski, R. Schneider, K.A. Sachse, S. Weirich & S. Henschel (Eds.) IQB-Bildungstrend 2021 Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im dritten Ländervergleich (pp. 233-258). Waxmann.

Stanat, P., Schipolowski, S., Rjosk, C. & Pant, H.A. (2017). Conceptual foundations of the IQB Education Trend 2016. In P. Stanat, S. Schipolowski, C. Rjosk, S. Weirich & N. Haag (Eds.) IQB- Bildungstrend 2016 Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich (pp. 13-19). Waxmann.

Mathematics based on the educational standards for primary education

The educational standards for mathematics for primary education (KMK, 2005) define the mathematical competencies that pupils in Germany should generally have acquired by the end of Grade 4.[2] The educational standards describe three main dimensions: the process dimension, the content dimension and the requirements dimension. Six central mathematical competencies are defined within the process dimension. The content dimension describes five key ideas or domains, while the requirements dimension describes the cognitive demand of a task. The domains (key ideas) described in the content dimension include 1) *numbers and operations*, 2) *space and shape*, 3) *patterns and structures*, 4) *sizes and measurement* and 5) *data, frequency and probability*.

The following explanations focus on the description of the types of tasks used for the assessment. For a more comprehensive description of the educational standards and the underlying concept of proficiency and proficiency levels in mathematics, please refer to the reports for the *IQB trends in student achievement studies* (see Stanat et al., 2017; Pant et al., 2017).

Operationalisation

Mathematical competencies were assessed using tasks from the *IQB trends in student achievement studies*. Most tasks consist of a brief instruction or problem statement that comprises one to five lines and introduces students to the context. Some stimuli also include diagrams, drawings, or illustrations. Up to seven related sub-tasks follow these stimuli, using different formats. While most tasks use closed response formats, some require short answers. Multiple sentences are occasionally expected as answers, for example, to describe the calculation process. Regarding content, all five aforementioned key ideas are covered by the tasks. These include applying basic arithmetic operations, completing patterns, mirroring figures, completing calculation chains, and handling place value charts. Other tasks focus on estimating, converting, and assigning various units and dimensions. Finally, students are expected to assess the probability of certain events or gather and interpret information from diagrams. All tasks were developed at the Institute for Educational Quality Improvement (IQB) by teachers under the guidance of experts in didactics and were empirically tested before use.

Scaling

Ten test booklets were used in a multi-matrix design. After recoding and aggregation, a total of 75 items were included in the analyses. The preparation of the data and the estimation of the item and person parameters, based on a Rasch model, is based on the procedure for educational trends (see Sachse et al., 2022, for a detailed description). First, the item parameters were estimated (calibration of the items) and then the students were located on the common difficulty and competence scale (estimation of the person parameters). Students'

[2] In 2022, the KMK adopted further developed educational standards for the subject of mathematics, but these must first be implemented in schools. The SC8 surveys are therefore based exclusively on the standards from 2004.

background characteristics were not considered when estimating the Weighted Likelihood Estimates (WLEs) provided in the data set.

Literature

KMK - Secretariat of the Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany. (2005). *Educational standards in the subject German for primary education.* Resolution of 15 October 2004. Munich: Luchterhand.

National Council of Teachers of Mathematics (Eds.). (2000*).* Principles and standards for school mathematics. Reston, VA: National Council of Teachers of Mathematics OECD (Eds.). (2003). *The PISA 2003 assessment framework mathematics, reading, science and problem solving knowledge and skills*. Paris: OECD.

Pant, H.A., Böhme, K., Stanat, P., Schipolowski, S. & Köller, O. (2017). Competence level models for the primary sector. In P. Stanat, S. Schipolowski, C. Rjosk, S. Weirich & N. Haag. (Eds.) *IQB- Bildungstrend 2016 Kompetenzen in den Fächer Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich* (pp. 45-52). Waxmann.

Sachse, K.A., Weirich, S., Becker, B., Hafiz, N.J., Schneider R. & Schipolowsk, S. (2022). Methodological principles: Design, implementation and evaluation of the IQB-Bildungstrend 2021. In P. Stanat, S. Schipolowski, R. Schneider, K.A. Sachse, S. Weirich & S. Henschel (Eds.) *IQB-Bildungstrend 2021 Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im dritten Ländervergleich* (pp. 233-258). Waxmann.

Stanat, P., Schipolowski, S., Rjosk, C. & Pant, H.A. (2017). Conceptual foundations of the IQB-Bildungstrend 2016. In P. Stanat, S. Schipolowski, C. Rjosk, S. Weirich & N. Haag (Eds.) *IQB-Bildungstrend 2016 Kompetenzen in den Fächer Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich* (pp. 13-19). Waxmann.