DIDACTICA MATHEMATICAE
Vol. 38, 2016, p. 175–202

Bruno Scheja (Köln, Germany)

Cognitive activation through mathematics tasks in the context of centralised examinations on ISCED 2 Level, using the example of North-Rhine Westphalia

Abstract: This study analyses the cognitive activation through mathematics tasks in the context of the centralised examination- Zentrale Prüfung 10 (ZP 10) – administered in North-Rhine Westphalia. Cognitive activation, here, is understood as being one characteristic within a range of curricular, content-related and cognitive categories that can potentially be activated in the course of the solution process of a given task. Here, it is fundamental to first consider obligatory curricular standards as well as to categorise mathematics tasks in terms of their function. This first section is then followed by the analysis of the cognitive requirements of the complete set of test tasks used between 2007 (introduction of ZP 10) and 2016 ($N = 1706$).

1 Introduction

One of the major elements of the changes of educational systems that started internationally in the 1980s is the implementation of centralised (final) examinations, which are in most cases administered at the end of a specific grade (cf. Krüger, 2015; Zabala et al., 2008). The sets of tasks used for these examinations as well as the standards that are at their basis reflect consequences expected by education policy with regard to educational processes in the context of individual subjects. One of the countries that began testing its pupils' performance in the wake of the educational reforms that started in 2001 on the basis of the complex interaction of the factors mentioned above is the Federal

Key words: secondary education, examinations, task design, educational research

Republic of Germany (cf. KMK, 2001). The centralised examinations are gradually administered on both national and federal levels. In the years following the successively implemented reform, Germany's performance in mathematics, as it has been gauged, for example, in the context of the international PISA test (2003-2012), shows an above-average increase (cf. OECD, 2004; OECD, 2014). Against the backdrop of this, it is of special subject-didactic interest to focus on instruments that both capture as well as influence quality, as for example the "centralised examination 10 (ZP 10)", which is obligatorily administered at the end of ISCED (*International Standard Clasiffication of Education*) level 2 (*secondary education*) in North-Rhine Westphalia. It can be assumed that, due to both its function as well as the specific point in time in learners' education biography at which the test is taken, the ZP 10 is a major factor of influence that affects German mathematics education.[1] In the few studies on this controlling instrument that have been published thus far, the focus is most frequently put on systemic, i.e. organisational and administrative considerations (cf. Klein et al., 2014). However, subject-internal aspects that reflect, for example, the new post-reform understanding of modes of thinking and working in mathematics education, are rarely considered (cf. Büchter & Pallack, 2012; Kühn & Drüke-Noe, 2013). A pivotal instrument for standardisation, which can be identified as being a carrier of these modes of thinking and working, are the *tasks* used to test learners. Against this backdrop, the characteristics of the requirements of said tasks, which define their level of difficulty/complexity, and their orchestration as a reflection of their quality are of major importance (cf. Neubrand et al., 2013). The category-system, which is based on the cognitivist-based groundwork done by Neubrand et al. (cf. 2002) and which was developed in the context of the COACTIV-project (cf. Neubrand et al., 2013) and then modified by Drüke-Noe (cf. 2014), is a comprehensive tool that captures the potential for cognitive activation of tasks. On the basis of a *rational task analysis* (cf. Resnik & Ford, 1981), using such a classification scheme allows for a faceted development of the main objective of this paper, which is to capture, anaylse and evaluate the level of cognitive activation of mathematics tasks of the North-Rhine Westphalian ZP 10.

By choosing this objective, the author takes up the desideratum of research outlined above in order to demonstrate and compare a comprehensive, content-related and conceptual profile of test tasks of a educational system that has recently been reformed. The realization of the chosen objective first introduces the discussion on the curricular standards of the ZP 10, or, the

[1] 23% of all German pupils who attended a general education school did so in North-Rhine Westphalia.

scholastic standards for Mathematics for the intermediate secondary school level certificate in the theroetical part of this paper. In the following chapter, tasks, which can be understood as being pivotal vehicles for both the implementation and examination of such standards, will be discussed in terms of their type and their quality as carriers of the potential for cognitive activation. The data collection scheme presented in the chapter on methodology, i.e. chapter 3, which is composed of categories and their respective properties and which is used for the subsequent analysis of the quality of cognitive activation, then allows for the implementation of the main objective of this paper in chapter 4. Finally, the results are summarised and dicsussed in chapter 5. This discussion also includes suggestions which are supposed to contribute to the development of task culture, since test tasks should be closely connected to classroom tasks and exam tasks.

2 Theoretical Background

2.1 Curricular Standards for the Centralised Examination ZP 10 in Mathematics

As part of the seven-step reformation scheme of the Conference of the Ministers of Education and Cultural Affairs – German abbreviation: KMK – nationwide effective scholastic standards that are differentiated for different grades were mandatorily implemented between 2004 and 2005 as primary measures of the fifth field of educational action (cf. KMK, 2001). Connected with their implementation was the goal to "establish a transparent system of Quality Assurance in Germany. Furthermore, they are supposed to improve teaching processes in order to attain higher educational returns" (Blum at al., 2006, p. 9; author's translation). For the intermediate secondary school level certificate (ISCED 2), they were first limited to the subjects German, Mathematics, L2 (English/ French) as well as the scientific domains. Irrespective of the subject, the scholastic standards are based on "educational goals that are broadly understood, embedded in subject matter and reflecting the structure of the respective subject. They are supposed to be attained in school and are furthermore embedded in the framework of interdisciplinary educational goals" (ibid., p. 15; author's translation). In this respect, the scholastic standards are *Performance Standards,* which list precisely those subject-specific competences at the end of secondary level I that are considered pivotal for a seamless continuation of any further educational as well as professional qualification.

With the obligations assumed by the federal states to implement and execute the scholastic standards, the latter affect everyone involved in education

on all levels of the education system (macro-level (e.g. federal state level), meso-level (school level) as well as micro-level (school class level)). Here, these standards fulfil a guiding function for the different subjects but are also decisive for internal as well as external evaluations and have evaluation function and monitoring function respectively (Klieme et al., 2003, p. 99f.). The latter imply, for example, standardised tests, which have certification function, at the end of secondary level I in 15 federal states[2]. The concept of the *scholastic standards for Mathematics for the intermediate secondar y school level certificate* is in part based on already established competence models. Here, as with PISA, three different dimensions are distinguished:

1. Comprising what is called *general mathematical knowledge*, the first dimension implies pivotal mathematical activities, meaning the performance while dealing with mathematical content, namely "mathematical argumentation", "solving problems mathematically", "modeling mathematically", "using mathematical representations", "using mathematics on a symbolical/formal/technical level" as well as "communicating mathematically". A comparison with the *Process Standards* as the respective standards of the *Principle and Standards for School Mathematics* of the NCTM as the most prominent of its kind shows that *Modeling* is incorporated while *Connections* is not (cf. Reiss, 2004).
2. *Mathematical content competences*, which are distinguished with regard to the following central ideas, i.e. subject matters, "numbers", "measuring", "space and form", "functional correlation" as well as "data and contingency". Concerning this aspect, these competences and the *Content Standards* of the NCTM are congruent, although the scholastic standards provide requirements that are phrased more explicitly and thus more concrete (ibid.).
3. In regard to the *requirement levels* (*Anforderungsbereiche*), three levels of cognitive requirement are distinguished, which can be identified by referring to the following, already established taxonomies of educational goals when it comes to solving problems: "reproducing", "establishing connections/correlations" as well as "generalising and reflecting" (cf. Anderson & Krathwohl et al., 2001). From one level to the next, there is an increase in cognitive complexity, which manifests itself in the respective empirical difficulty (cf. KMK 2003, p. 6ff.). Similar nuances regarding the level of requirements cannot be found in the standards of the NCTM.

[2]Only Rhineland-Palatinate refrains from the implementation of such tests (cf. Kühn, 2013).

> Instead, they name goals and contents of an "outstanding mathematics education" as ideal standards (Reiss, 2004, p. 637).

This concept of different dimensions of the scholastic standards makes it possible to make the general educational goals of mathematics education, which in part trace back to Freudenthal (1983) but mostly Winter (1995), more concrete. Thus, the function of mathematics to broaden general knowledge as presented by Freudenthal, which, besides formal educational goals, especially emphasizes the didactic principle of connectivity (*Beziehungshaltigkeit*) is explicated. In his explication, Winter assumes that the mathematics education that broadens general knowledge should enable all learners to make three well-balanced *basic experiences* listed and explained in the following table:

Type of basice experience	**Characterisation of basic experience**
Application-Oriented-Learning (Anwendungsorientierung)	Learners should be enabled to "perceive and understand technical, social and cultural phenomena and processes with the help of mathematics and to evaluate them with the aid of mathematical principles" (KMK, 2003, p. 6; author's translation).
Structure-Oriented-Learning (Strukturorientierung)	Learners should be enabled to "know and understand mathematics and its language, symbols, images and formulas and their significance for the description and performance/solution of tasks and problems within and outside the field of mathematics" (ibid.; author's translation).
Problem-Based-Learning (Problemorientierung)	Learners should be enabled to "acquire a general capacity to solve problems while performing tasks and solving problems with the aid of mathematical means" (ibid.; author's translation).

Table 1. The three basic experiences according to Winter (cf. 1995).

In terms of concept and design, the fundamental significance of these basic experiences for the scholastic standards and thus the respective processes in mathematics education is reflected in the orchestration of the competences imparted to learners. Hence, a unilateral orientation towards rules and algorithms in the context of mathematics education carries the risk of learners "developing a unilateral or even contorted conception of mathematics since the three *basic experiences* as put forward by Winter are not sufficiently taken into account in a balanced manner" (Drüke-Noe, 2014, p. 25, author's translation). With regard to the tasks used in mathematics education as pivotal means to control teaching-processes in the classroom (reflecting standards) as well as the quality of teaching (capturing the implementation of standards),

this means: As key elements of control, regardless of their function, – i.e. both as learning tasks as well as assessment tasks – they should appropriately reflect the given standards for mathematics education. This is primarily reflected in a well-balanced distribution of the three basic experiences in the established requirement profile.

2.2 Types of Mathematical Tasks Against the Backdrop of Cognitive Activation

In the context of everyday teaching, tasks always play an important role, although in mathematics education they mostly dominate in situations of learning and testing. Here, they always ask learners "to specifically deal with a clearly defined mathematical issue" (Neubrand, 2002, p. 16). Whether they are learning tasks or assessment tasks, they have various structural similarities and it is thus appropriate to categorize them in terms of their intended purpose (cf. Drüke-Noe, 2014, p. 8f), i.e. in terms of their function.

In situations of learning, *classroom tasks* (*Unterrichtsaufgaben*) are an interface of the activity and communication between teachers and learners and thus a pivotal tool for the preparation and conduct of lessons. Teachers use them to control and manage their lessons and hence their learners' activities. In the context of this process, tasks carry mathematical contents and can therefore be understood as being the main means of control for the progression of the development of competences. This is mostly ensured by tasks, which, according to Jordan et al. (2006, p. 13), can be used as flexible, widely applicable and actively controllable means of controlling didactic aspects and content. Given an appropriate level of awareness, teachers can also influence learners' understanding of a mathematical term or process, the structure of a dense network of terms and ultimately their overall conception of mathematics (Neubrand et al., 2013, p. 126). The tasks used in class are hence often a means of identifying the requirements set out for and experienced by learners. These tasks set a framework for the formation of mathematical skills, which can be consciously experienced by learners while being actively involved in the solution process. In the process of building knowledge, classroom tasks are thus a vehicle of learners' cognitive activities.

Assessment tasks, on the other hand, are a means to attest "which skills have been acquired in the course of the education process" (Terhart et al., 2009, p. 24; author's translation). They are usually used for examinations or other forms of assessment in the context of internal and external evaluation of learner-performance. Internal evaluation comprises, for example, class-tests or informal tests that are used for individual classes or forms as a whole

respectively; external forms of evaluation are then based on/guided by overarching standards (used for the assessment of learners of one federal state or assessment on a national level respectively as well as final examinations). One major characteristic of assessment tasks of the first category is that the teachers, who base their decision on the type of classroom tasks used prior to the assessment of performance, always choose and design them. Hence, these tasks make it possible to draw valid conclusions about the main emphasis chosen during as well as the quality of the mathematical requirements that characterises the lessons taught prior to the respective assessment (as can be seen, for example, in the context of the COACTIV-study (cf. Kunter et al., 2013)). Furthermore, they provide suggestions for possible action or behaviour in regard to the respective group of learners. Studies on the implementation of the diagnostic potential of these tasks, however, reveal that this potential only selectively leads to consequences on teaching. Assessment tasks that are part of external forms of assessment - mostly *Large Scale Assessments* – however, are designed by exteranl institutions and should fulfil the quality criteria of test theory (objectivity, validity, reliability). Depending on their function, the development of tasks requires the variable operationalisation of requirements as presented in the standards (cf. Büchter & Pallack 2012). In the case of German, national forms of assessment, these tasks are based on standards (under the overall control of the IQB/ Institute for Quality Development), i.e. individually and in combination as tests, they are in line with the competence standards outlined above. Assessments conducted in individual federal states, however, reflect the requirements of the respective curricula (in North Rhine-Westphalia these standards are under the overall control of QUA-LIS/ Agency for Quality and Support of the State Institute for Schools). Following the resulution of the KMK, these should be in line with the educational overarching standards. Thus, the tasks used for these forms of assessment are indirectly based on standards. Through thorough analyses of the structure and requiremens of tasks, Drüke-Noe has shown that, contrary to expectations, classroom tasks as well as assessment tasks "mostly have identical cognitive charactersitics" (Drüke-Noe, 2014, p. 9), so that the difference only becomes apparent in how teachers deal with and assess learners' results. Here, classroom tasks allow for deviations and divergenes to happen while assessment tasks are designed with a focus on performance and results and thus require convergence (cf. Büchter & Leuders, 2005, ibid). In summary, it can be said that the assessment tasks to be analysed in this study can be understood as also being carriers of cognitive activation. With regard to the task type analysed in this study, cognitive activation means the facets of cognitive processes that unfold while solving the task at hand within the "anticipated effect of teaching" (Neubrand et al. 2013, p. 129).

As a means to answer the pivotal research question of this study, one objective is to classify the tasks of ZP 10 in a way that their potential for cognitive activation is revealed. The evaluation categories used to realise this objective are empirically tried and tested and are taken from relevant pedagogical/pschological literature. These categories make it possible to and analyse cognitive activation guided by three questions:

- How are the ZP 10 tasks distributed across the different subject areas and how can the curricular level of requirement be evaluated?
- How are the ZP 10 tasks distributed across the different types of mathetmatical activities?
- How can the level of cognitive requirement of the ZP 10 tasks be evaluated?

3 Method

3.1 Methodological aspects of this study

By focussing on the idea of cognitive activation as well as by employing a certain method of classifying tasks, this study links to essential characteristics of the COACTIV project and its conceptional extensions.Thus, after presenting the data sets in 4.1, the conceptional basics and the method of classifying tasks will be explained in 4.2. Finally, the process of coding is will be reported in 4.3.

3.1.1 Sample and Dataset

The dataset analysed within the scope of this study is the set of tasks used for the centralised examination that is part of tenth grade (ZP 10) and conducted in all of North Rhine-Westphalia between 2007 (introduction of the ZP 10) and 2016. It comprises the tasks for all ISCED 2 school types (N=1706) and consists of the respective task sets used for *Hauptschule* (consisting of subsets of tasks for "Hauptschulabschluss" after year 10 (secondary school qualification) and "Realschulabschluss" (secondary school certificate)), *Gesamtschule* ("Elementary Course" and "Advanced Course"), *Realschule*, *Abendrealschule* (night school) as well as *Gymnasium* (here, this test was only conducted from 2007 to 2010).

3.1.2 Task Categories

3.1.2.1 Task Categories Based on the COACTIV-Study

In order to capture the level of cognitive activation of the tasks used, the author draws on a range of empirically established categories for classification drafted in the context of the COACTIV-project (Professional Competence of Teachers, Cognitively Activating Instruction, and the Development of Students' Mathematical Literacy) (cf. Kunter et al., 2013) and developed further in the context of the analyses of tasks by Drüke-Noe (cf. 2014).

Successively extending the conceptional groundwork done by Neubrand et al. (2002), both studies extensively analyse, among other things, classroom tasks and classtest tasks respectively, which were used in grades nine and ten at German schools. The classification scheme chosen for this paper subsumes assessment criteria of tasks under three dimensions/ frameworks: content framework, cognitive framework as well as cognitive elements of the mathematical modeling cycle. These dimensions are further differentiated into categories with, mostly, four different property values (cf. Table 2). The values of these categories, as for example in the case of "topic area," reach from highly inferential to lowly inferential ratings (cf. Rosenshine, 1970).

In the context of the dimension "Content framework", a task is allocated one or more *topic areas*, which constitute the framework of content knowledge in the course of its processing. The category *Curricular knowledge level*, then, captures cognitive activation so that, through the respective allocation, the periods in which a topic is covered, the range of different topic areas and thus the curricular requirement level are explicitly taken into account (Jordan et al., 2008, p. 87). This criterion is captured – and this should be noted explicitly – by using German mathematics curricula as a guidance for rating (cf. Jordan et al., 2006, p. 29).

Falling under the category "Type of mathematical activity", the subsequent dimension, "Cognitive framework", captures the task type by showing the type of the corresponding solution process. First, there is a disticntion made between different types of solutions, which are characterised by the employment of one the following forms of modeling: 1. *intra-mathematical modeling*; 2. *extra-mathematical modeling*. The first merely requires the contextless application of skills or factual knowledge to already given assumptions (cf. Fig. 1), whereby processing forms of performance such as modeling or structuring are not required for the solution.

Dimension	Category	Properties
Content framework	Topic area	1 – arithmetic, 2 – algebra, 3 – geometry, 4 – stochastics
	Curricular knowlegde level	1 – elementary knowledge (level I), 2 – basic knowledge at lower secondary level (level II), 3 – advanced knowledge at lower secondary level (level III)
Cognitive framework	Type of mathematical activity	1 – technical task, 2 – procedural task, 3 – conceptual task
Elements of the modeling cycle	Extra-mathematical modeling	0 – not required, 1 – standard modeling, 2 – multistep modeling, 3 – reflection on a model, development and validation of complex models
	Intra-mathematical modeling	0 – not required, 1 – standard modeling, 2 – multistep modeling, 3 – reflection on a model, validation, strategy development
	Processing mathematical texts	0 – not required, 1 – direct text comprehension, 2 – text comprehension with reorganisation, 3 – comprehension of logically complex texts
	Mathematical argumentation	0 – not required, 1 – standard reasoning, 2 – multistep argumentation, 3 – development of complex argumentation, proofs, evaluation of arguments
	Using mathematical representaions	0 – not required, 1 – standard represenataions, 2 – switching between different representations, 3 – evaluating representations

Table 2. Pivotal assessment criteria of this study (cf. Neubrand et al., 2013, p. 134).

Task 3	**Task 4**
Solve the following equation:	Solve the following equation system:
$2x + 1.8 = 3.6 - x$	(I) $2x + y = 2$ (II) $x - 0.5y = 2$

Figure 1. Examples of a technical tasks taken from the examination administered in 2015 (MSW, 2015).

Mathematical modeling tasks, by contrast, require learners to transfer a problem to a processing ansatz and to refer back the solutions to the given initial situation. Thus structured solution processes – exemplarily understood as modeling cycle (cf. Jordan et al., 2006, p. 31) – are reflected in procedural and conceptual modeling tasks. The solving of procedural modeling tasks primarily requires procedural and algorithmic thinking/knowledge, while the processing

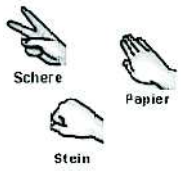
of conceptual modeling tasks mainly requires conceptual thinking/knowledge (ibid). Here, both types of tasks can be embedded into a context located either inside or outside the field of mathematics. Thus, the three task types cover different fields of mathematical competence. In accordance with section 2, it is, among other things, the goal of German mathematics education to ensure that the different ways of thinking/types of knowledge that underlie these three task types are adequately reflected in the tasks used in class.

The categories of the third dimension "Cognitive elements of the mathematical modeling cycle" operationalise the requirement level of the cognitive processes and activities corresponding to the five types of mathematical activities considered in this study. For this, it is of fundamental importance that the task context presenting a mathematical problem – whether it lies inside or outside the field of mathematics – is transfered into a processing model with the same structure (*intra-mathematical modeling* and *extra-mathematical modeling*), in which processes such as translating/ transforming, processing or interpreting/validating can occur in a cognitively analogous manner, although with different characteristics (ibid., p. 32). The performance of these two activities can occur – analogous to *mathematical argumentation* as well as *using mathematical representations* – on four different levels of complexity: 0. no allocation of a property value 1. low, 2. moderate, 3. high (cf. Neubrand et al., 2013, p. 133). For the mathematical activity "mathematical argumentation", which, in the anlysis of PISA tasks, has also been verified to be a characteristic that generates complexity, these are operationalised as follows:

- 0. no argumentation required;
- 1. Standard reasoning: "Simple reproduction of standard forms of argumentation; perform forms of argumentation that only require common knowledge; developing one-step or purely procedural forms of argumentation" (Jordan et al., 2006, p. 40/ author's translation);
- 2. Multistep argumentation: "Developing straightforward multistep and also conceptual mathematical forms of argumentation and presenting them in writing or, if required, understanding them" (ibid./ author's translation);
- 3. Devoloping of complex argumantation, proofs: "Developing complex mathematical forms of argumentation (explanations, proofs, strategies, generalisations) and presenting them in writing or, if required, understanding them; comparing or evaluating different types of mathematical argumentation (ibid./ author's translation).

The category *Processing mathematical texts* captures the skill level that is required to extract information from the text of a given task, which are necessary for structuring the solution process. Here, test tasks 3 and 4 (Fig. 1), for example, requires the lowest level of verbal argumentation (property = 0) while the level that is required for the task shown in Fig. 2 can be classified as moderate.

"**Part 2, task 3**
Anne and Paul are playing "Rock-Paper-Scissors". After counting to three, both form one of the three signs with their hands. The following applies:



- Rock beats scissors because it blunts the scissors.
- Scissors beats paper because they cutt he paper.
- Paper beats rock because it wraps the rock.
- If both players form the same sign, the round ends with a draw.

In the following, it is assumed that both players make their decisions by pure chance. (. . .)

b) 1) Fill in the table below: Who wins which game? Where does the game end in a draw?"

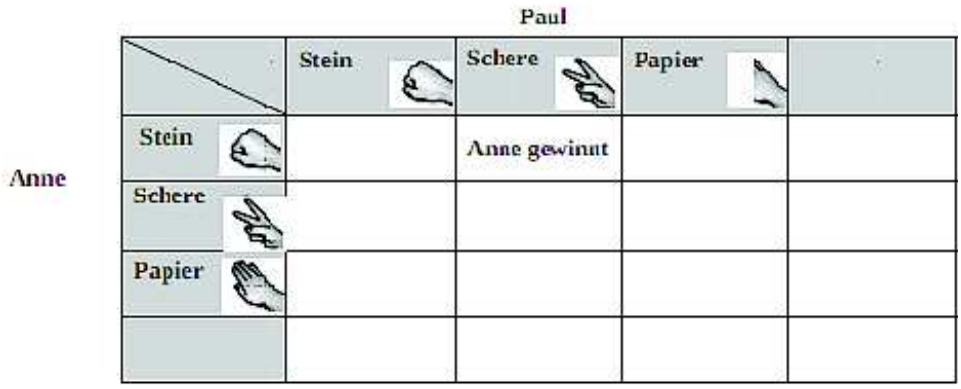


Figure 2. Example of a task taken from the task set used in 2013 (cf. MSW, 2013, p. 4).

An analysis of the task reveals that the sequence of the (parts of the) sentences, in contrast to a task with a "low requirement level", are not in accordance with the steps necessary to process the task mathematically. The logical connections are hidden under the linguistic structure, i.e., the main and subordinate clauses used in the task have linguistic references. Here,

"if-then-connectives" (conditionals) and universal quantifications, respectively, are fundamental logical functions.

3.1.2.2 On the Necessity of Extending the Assessment Scheme

A comparison of the test task as presented in Fig. 3, in which learners are asked to solve a general quadratic equation and the task 3 presented in Fig. 1, shows that they, due to their similarity with regard to the type of mathematical activity required for the respective solution, can both be allocated to the same task type, namely "technical task", although the process of finding a solution in accordance with the requirements of the task is more "difficult".

Find the solutions to the following equation:

$4x^2 + 4 = 8x$

Figure 3. Example for a technical task (Drüke-Noe, 2014, p. 69).

In the scheme presented above, their potential for cognitive activation, unfolding in the course of the solution process, is uniformly captured with congruent property values: 2 (topic area), 3 (curricular knowledge level), 1 (Type of mathematical activity), 0 (Extra-mathematical modeling), 0 (Intra-mathematical modeling), 0 (Processing mathematical texts), 0 (Mathematical argumentation), 0 (Using mathematical representaions) (cf. Table 3).

The analytical category *working technically* ("Technisches Arbeiten") developed by Drüke-Noe amends the scheme by explicating necessary activities and their requirements in terms of technical skills. Since both technical tasks as well as tasks, whose solution process can be described by means of a modeling cycle, require learners to deal with calculi of varying complexity (cf. Fig. 4.4 in Drüke-Noe, 2014, p. 71) it is possible to apply the category to all three types of tasks. Here, the category also captures the requirement level of dealing with calculi, while considering different steps of the solution process (cf. Fig. 4.3 . 70 and Fig. 4.4 p. 71, Drüke-Noe, 2014). In accordance with the classification scheme as presented in Table 2, this dealing with calculi is required in the context of one of the four topic areas (Arithmetic, Algebra, Geometry, Stochastics). Here, Drüke-Noe designates four different levels:

Activity/ Category	Topic area	Level	Properties and examples
Working technically	Arithmetic	0	Not required, e.g. a
		1	Only multiplication and division calculation/ only addition and subtraction calculation, e.g. simple exponentiations (positive base and power)
		2	Simple hierarchical operations (multiplication and division calculation before addition and subtraction calculation, exponentiation before multiplication and division calculation), e.g. multiplying/dividing exponentiations (positive base and power), extracting the root of a given number
		3	Complex hierarchical operations (Exponentiation with multiplication and division calculation/ addition and subtraction calculation)
	Algebra	0	Not required, e.g. Is the mapping. . . a function?
		1	Only multiplication and division calculation/ only addition and subtraction ccalculation, basic multiplication and division calculation before addition and subtraction calculation, e.g. calculating with a linear equation
		2	Simple hierarchical operations (complex multiplication and division calculation before addition and subtraction calculation, exponentiation before multiplication and division calculation) e.g. complex multiplication and division calculation before addition and subtraction calculation, solving quadratic equations, squaring a sum/factorising, calculating with trigonometric functions
		3	Complex hierarchical operations (Exponentiation with multiplication and division calculation/ addition and subtraction calculation), e.g. solving mixed quadratic equations or exponential equations
	Geometry	0	Not required, e.g. what is the measure of a right angle?
		1	Only multiplication and division calculation/ only addition and subtraction ccalculation, e.g. calculating sin(23) , calculating by using the intercept theorem
		2	Simple hierarchical operations (i.a. exponentiation before multiplication and division calculation), e.g. calculating angle size by using the law of sines
		3	Complex hierarchical operations (exponentiation and multiplication and division calculation combined), e.g. calculating the face or volume of solid figures
	Stochastics	0	Not required, e.g. noting all combinations
		1	Only multiplication and division calculation or only addition and subtraction ccalculation, e.g. calculating simple probabilities
		2	Simple hierarchical operations (multiplication and division calculation before addition and subtraction calculation, exponentiation before multiplication and division calculation), e.g. calculating the arithmetic mean
		3	Complex hierarchical operations (exponentiation combined with multiplication and division calculation and addition and subtraction calculation), e.g. calculating probability in the context of the multistep Bernoulli experiment

Table 3. Descriptions of the different levels of working technically (taken from Drüke-Noe, 2014, p. 78/ author's translation).

Exceeding the the limits of the COACTIV approach, a thus operationalised mathematical activity now makes it possible to locate, for example, the requirement level of the two technical tasks presented at the beginning with regard to the "technical work" required by said tasks. Hence, they are both first allocated to the topic area "algebra" and the category "Advanced knowledge at lower secondary level (level III)". By reference to the corresponding, ideal solutions, it becomes clear that, in the first example (cf. Fig. 3), the hierarchical structure of the terms on both sides of the equation can be identified first. Dealing with the mixed quadratic equation requires a) different mathematical operations (general: $ax^2 + bx + c = d$, with a, b, c, d being rational numbers and a$\neq$0) and, irrespective of the solution procedure chosen, b) the understanding use of powers and roots (of sums). In this respect, these are "Complex hierarchical operations of (exponentiation combined with multiplication and division calculation and addition and subtraction calculation) – level 3", since finding the solution requires the use of priority rules, such as "exponentiation before multiplication and division calculation". In the second example (task 3, Fig. 1), the multiplicative-additive structure of the terms on both sides of the equation can be identified first, i.e. before appropriately using sums and products and before finally isolating x by means of dividng two numbers. This designates precisely the technical requirements of level 2.

3.1.3 Coding

The subsequent coding of the ZP10-tasks was done independently by the author (coder 1) and an experienced maths teacher (coder 2), who was thoroughly introduced to the conception of the study beforehand. Since the reliability of the assessment is of major importance, both coders practiced coding with tasks taken from the COACTIV-study. In the context of the COACTIV-project, these taks had already been allocated to COACTIV- assessement categories (cf. Table 2). Moreover, both coders also had the coding of "working technically", done by Drüke-Noe and her colleagues, as a basis.

In a further step, it was of particular interest to first identify and quantify the variability of the *interrater agreements* in order to obtain indications from the agreements as to the stability and reliability of the coding-trial. The Cohen's kappa coefficient k has been applied to measure the quality of agreement. $0.7 < k$ has been defined as the minimum value of reliability and thus for proceeding to the coding of the tasks used in the test conducted at ZP 10. Following Wirtz and Caspar (2002), $0.75 < k$ is understood to be a "very good", $0.6 < k < 0.75$ a "good" and $0.4 < k < 0.6$ an "acceptable" delineation regarding the concordance of the respective assessments. Table 4 shows the

Kappa values for the different assessment categories and their properties used in this study that were determined in the context of the coding-trial.

	Kappa values k ($N = 291$)		
Category/property	Coder 1/ COACTIV-Coder	Coder 2/ COACTIV-Coder	Coder 1/ Coder 2
Arithmetic	0.71	0.73	0.7
Algebra	0.74	0.69	0.71
Geometry	0.89	0.91	0.88
Curricular knowledge level	0.93	0.86	0.83
Type of mathematical activity	0.79	0.77	0.82
Extra-mathematical modeling	0.94	0.93	0.94
Intra-mathematical modeling	0.92	0.95	0.9
Processing mathematical texts	0.84	0.71	0.73
Mathematical argumentation	0.95	0.9	0.93
Using mathematical representaions	0.98	0.99	0.99
Working technically[3]	0.89	0.76	0.82

Table 4. Measure of coding-agreement within the framework of the assessment categories of this study.

The results in the respective columns show that the quality of agreement of the coding-trial ($N = 291$) for all categories/properties can be described as being at least "good". The $k - v$alues for the subjetc areas "Arithmetic" and "Algebra" are good, while the values for "Geometry" suggest a very good quality of agreement. The latter is also true for the allocations in the category "Curricular knowledge level". For both types of mathematical activity – "extra-mathematical modeling" and "intra-mathematical modeling" – the $k - v$alues are consistently very good, while those for the category "Processing mathematical texts" suggest a good quality of agreement. The category "working technically", which extends the assessment sheme, is characterised by a quality of agreement that can be described as "very good" since here, the $k - v$alues are consistently higher than 0.75. The coding of the last two categories of the dimension "Cognitive elements of the mathematical modeling cycle" can even be described as being almost matching. Overall, it can be noted that the coding of the North-Rhine Westphalian set of tasks can sufficiently be connected to the coding of the categories of the COATIV-scheme and the category "working technically".

[3]This ist the measure of agreement between coder 1 and 2, respectively, and the assessments of Drüke-Noe and her colleagues.

4 Results

Guided by the three pivotal questions of this study (cf. 3.), the following chapter presents the results of this study on cognitive activation through North Rhine-Westphalian ZP 10 tasks used between 2007 and 2016 ($N = 1706$).

4.1 Cognitive activation through ZP 10 tasks

First, the topic areas and the curricular level of requirement will be analysed in section 5.1. This analysis is then followed by an examination of the types of mathematical activities and hence the cognitive framework of the tasks (5.2). The chapter is then completed by an in-depth look at the cognitive requirements of the test tasks (5.3).

4.1.1 Content-related activation of ZP 10 tasks

A major goal of central examinations is to gather information about tenth graders' knowledge levels and skills with regard to the standards of the scholastic standards for the *Mittleren Schulabschluss* (General Certificate of Secondary Education (GCSE)). The "cenral ideas" listed within the dimension "content-related competences" (cf. section 2.1) have been captured in terms of the following topic areas: Arithmetic, Algebra, Geometry und Stochastics. First, table 4 shows that the tasks are unevenly distributed across all four topic areas. At percentages between 26.7% and 31.6%, it is thus primarily the topic areas of Arithmetic, Algebra and Geometry that are tested predominantely as well as equally frequent. With 14.9% of the tasks, the topic area of Stochastics constitues – as expected – the smallest percentage.

When analysing these topic areas with regard to the curricular requirement level of the mathematical knowledge tested in said areas, it becomes evident that the intermediate level with $M = 2.43$ is between the "basic knowledge at lower secondary level (Level II)" and the "advanced knowledge at lower secondary level (level III)". In particular, it is furthermore significant that tasks, which exclusively require the elementary knowledge of primary school (with four forms), barely carry any weight. Moreover, the third curricular knowledge level is covered slightly less than the second level. This indicates that – contrary to expectations[4] – mostly tasks whose curricular requirements can be classified as belonging to the periods of teaching before 10th grade are

[4]According to MSW, the ministry in charge for NRW, "the second, more extensive part of the test should include tasks that require competences of forms 9/10 and should be based on major topics relevant for said forms" (https://www.standardsicherung.schulministerium.nrw.de/zp10/fragen-und-antworten/, 21.10.2016, author's translation).

used for the test. Overall, about six out of ten test tasks can be allocated to the category "Basic knowledge of lower secondary education (Level II)". However, table 4 shows that, in the context of the different topic areas, there are considerable differences as to their curricular requirement level. Here, the results show that tasks that belong to the topic areas Arithmetic ($M = 2.11$, $SD = 0.30$) and Stochastics ($M = 2.29$, $SD = 0.37$) rarely reach a level higher than that of "lower secondary level", while the tasks that belong to the topic areas Algebra ($M = 2.52$, $SD = 0.57$) and Geometry ($M = 2.51$, $SD = 0.51$) do test knowledge of a wider range of abstraction. Irrespective of the curricular requirement levels, no form of anticipation within the scope of the curriculum could be noted in any of the topic areas. So far, the descriptions show that the average curricular requirement level of the North-Rhine Westphalian test tasks ranks between the low characteristic values of PISA 2003-national and the values achieved for class tests conducted in tenth grade (type of school "Gymnasium": $M = 2.80$; type of school "not Gymnasium": $M = 2.67$) in the context of the COACTIV-study (cf. Drüke-Noe, 2014, p. 103). This also shows one of the functions of this instrument within the system of external performance assessment: without a clear focus, the ZP 10 tests the knowledge and skill level of the education period between fifth and tenth grade (= second and third level of curricular knowledge).

Topic area	**Number of items ($N = 1706$)**	**Percentage**	**Curricular knowledge level M**	**Standard deviation SD**
Arithmetic	492	31.6	2.11	0.30
Algebra	462	29.7	2.52	0.57
Geometry	416	26.7	2.51	0.51
Stochastics	233	14.9	2.29	0.37
	Sum of allocations with $N = 1706$: 1766[5]	104	2.43	

Table 5. Distribution of the ZP 10 tasks across the four topic areas and the curricular knowledge levels.

[5]This number ist the result of multiple allocations within the different categories.

4.1.2 Type of mathematical activity

When looking at the cognitive processes that characterise the type of mathematical activity in the course of the solution process, it can be noted that procedural modeling tasks constitute a clear focus of testing (82.4%). At 9.0%, conceptual modeling tasks that require learners to find connections/ correlations between/of mathematical concepts only constitute a small proportion. Conceptual-procedural modeling tasks constitute a total of 91.4% of all test tasks. Here, application-oriented, problem-solving contexts constitute a clear majority. Only 134 of the 1706 test tasks require (cf. Table 6) learners to apply algorithms/ numeric methods of calculation to a given approach, i.e. to simply "handle them technically".

	Type of mathematical activity		
Property	Technical task	Procedural/procedural modeling task	Conceptual modeling task
Percentage with $N = 1706$	8.6	82.4	9.0

Table 6. Distribution of tasks across the different types of mathematical activity.

The distribution of the tasks across the different types of tasks indicates the focus on cognitive activation that seemed desirable for the Ministry of Education in the years following the introduction of the ZP 10. As part of centralised examinations, then, an apparent departure from a traditional self-image of mathematics education took place. It was considered to "be focussing to much on the acquisition of calculating skills and on mastering standard procedures (...) while often neglecting the development of flexibly applicable concepts and skills, which are especially required when handling transformation processes in new contexts (*Sachsituationen*)" (Blum et al., 2004, p. 156/ author's translation). Within the scope of the set of tasks analysed here, mathematical terms and procedures are almost exclusively used as "tools" for various, mostly application-oriented context problems. However, the comparison of this distribution of tasks (see Table 5) with that of both PISA tests *(PISA 2000 international (Literacy-Concept)*: technical task: 3.2%; procedural modeling task: 45; conceptual modeling task: 52%/ *PISA 2000 national*: technical task: 26.7%; procedural modeling task: 38.4%; conceptual modeling task: 34.9%) reveals the partial existence of considerable differences in terms of their respective concept (cf. Neubrand et al., 2002, p. 101).

In comparison to the two test concepts, the conceptual modeling tasks are underrepresented while procedural modeling tasks are overrepresented. The proportion of technical tasks is only small when compared to the national PISA tasks. Such a focus on algorithmic-procedural modes of thinking in the time

"after PISA", hence, preserves an imbalanced, although currently different, range of thinking skills within this category. Considering the approved three *basic experiences* by Winter (cf. Section 2), which should be the basis of the scholastic standards and thus of all teaching and learning processes, it becomes clear that neither the traditional/process-oriented focus nor the algorithmic-procedural focus are really desirable. Given the strong focus on *Application-Orientation* in the context of the ZP 10 tasks, it can be assumed that the instrument does not sufficiently capture the *basic experiences* in mathematics education

4.1.3 Cognitive Requirements of ZP 10 tasks

The types of mathematical activity considered earlier (see 5.2) show the characteristics of the solution process. However, they do not show the level of cognitive requirement that is necessary to perform said process. In order to present the analytic dimesion in a nuanced manner, both the attained average level of requirement as well as the distribution across the corresponding levels (0-3) are considered for every category. Table 7 shows the distribution of these levels across the mathematical activities within the scope of the North-Rhine Westphalian set of tasks.

	Level of the respective mathematical activity				
Mathematical activity	0	1	2	3	arithmetic mean
Extra-mathematical modeling	24.8	32.5	41.9	0.8	1.187
Intra-mathematical modeling	83.8	4.3	11.2	0.7	0.288
Mathematical argumentation	90.1	5.1	4.8	0.000	0.147
Using mathematical representaions	35.3	56.5	8.2	0.000	0.729
Working technically	19.5	46.3	19.0	15.2	1.299
Processing mathematical texts	21.5	63.0	15.5	0.000	0,94

Table 7. Distribution of the levels of requirement across the mathematical activities.

The table above illustrates that the analysis of transition processes between reality and methematics – i.e. *extra-mathematical modeling* – constitutes the majority (74.4%) of all tasks used for the North-Rhine Westphalian examination ($M = 1.187$). Here, intermediate mathematical modeling tasks constitute a majority (41.9%), while basic as well as challenging extra-mathematical modeling tasks constitute 32.5% and, repsectively, 0.8% of all tasks. Due to the results for the type of task addressed in the previous section - where it became apparant that only 8.86% of the tasks fall under the category *technical task* – the difference of 16.2% is attributable to the category *intra-mathematical modeling* ($M = 0.288$). Analogous to the category "extramathematical mo-

deling" the following is true: If a task requires mathematical modeling, the requirement level is most frequenlty "intermediate" and only rarely "elementary" or "advanced". On this basis, looking at just the cognitive processing of the test tasks (modeling task vs. technical task) while also considering the required levels of cognitive activation, it becomes apparant that even 87.8% of all tasks require either an elementary (36.8%) or an intermediate (53.1%) level of modeling.

Due to the fact that most of the tasks can be categorized as "closed tasks", only 9.9% of all tasks require learners to name closed chains of argumentation or to evaluate already given *argumentations* respectively. The 169 (9.9% of all tasks) tasks are distributed almost evenly across the elementary and the intermediate requirement level (standard and multistep argumentations). In 64.7% of all test tasks, *mathematical representations* are used ($M = 0.729$). They are mostly iconic representations of a low requirement level (56.5%). Here, similar to all the categories considered so far, it can also be noted that the concept does not incorportate cognitively challenging mathematical representations. On average, *working technically* is not required in one out of five tasks ($M = 1.299$). However, if it is needed for the solution process, it most frequently occurs on an elementary/basic level (46.3%) while the intermediate and the low requirement level are allocated to 19.0% and 15.2% of the tasks respectively. With regard to the requirement levels of the category *Processing mathematical texts*, the tasks are distributed similarly. Apart from "extra-mathematical modeling" and "using mathematical representations", they show the highest average in this category ($M = 0,94$). Furthermore, a comparison of the corresponding arithmetic means of the categroy with those of comparable sets of test tasks shows that theese arithmetic means (cf. Drüke-Noe, 2014, p. 160f; Neubrand et al., 2013, p. 128) can be classified as cognitively activating. This is surprising insofar as the tasks used only require, on average in eight ot of nine cases, the "elementary level", which imposes rather low requirements on learners.

5 Discussion and outlook

5.1 Summary and discussion of findings

Overall, the results of this study on cognitive activation show that North-Rhine Westphalian mathematics education of the ISCED level 2 between 2007 and 2016 have a comparatively high potential for cognitive activation, which is, however, relatively inhomogeneous. With regard to curricular characteristics of the test tasks analysed (first research question) it is characteristic that it

is mainly the topic areas algebra, arithmetic and geometry that are tested, while stochastics is only considered marginally, as is rue for the centralised examinations in other federal states (cf. Kühn & Drüke-Noe 2013). Against the backdrop of the defined goal of the ZP, the average level of curricular requirement within the realm of the four topic areas can be said to be rather low. The analyses in the context of the second question of this paper, in which the types of mathematical activities in the solution process have been examined, mostly showed that tasks whose solution require modeling processes, predominate considerably. Here, especially procedural and – with certain limitations – conceptual ways of thinking are used as a "tool" for inner-mathematical and especially various extra-mathematical contexts. In contrast, technical tasks, i.e. the task type that is used/employed most frequently in everyday mathematics lessons, are only considered marginally (cf. Jordan et al. 2008). Finally, the analysis of the levels of cognitive requirement in the context of the third research question showed that, with regard to the distribution of relevant test contents and the existing range of mathematical activities (intra-and extra - mathematical modeling, processing mathematical texts, mathematical argumentations, mathematical representations and working technically), the tasks show distinctive characteristics of broadly activating Large Scale Assessment-tasks (of which the international PISA test is one example/ cf. Neubrand et al., 2013).

Against the backdrop of these insights, it seems reasonable to suggest that both the unfolding of effects of any changing mathematics education that is embedded in the scholastic standards/core curriculum – e.g. with regard to cognitive activation (cf. Maier et al., 2011) – as well as the connection of test results with qualification systems concerning the subsequent ISCED 3 level, are a possible explanation for the increased mathematics performance in the PISA test.

However, it must be critically stated that the ascertained focus of the conception of ZP 10 tasks only slightly reflects the fundamental standards designated by the *scholastic standards*, which can be connected to already existing findings of subject-didactic research (cf. Kühn & Drüke-Noe, 2013) and implementation research (cf. Gräsel, 2010). This is, for example, reflected by the fact that the breadth of the three *basic experiences* of mathematics education that broadens general knowledge, as put forward by Winter, is disproportionately embedded in several respects. Considering the high proportion of extra-mathematical modeling, the basic experience of "Application-Orientation" can be described as being excessively implemented. In contrast, both intra-mathematical modeling as well as contextless technical tasks contribute insufficiently to the capture of the basic experience "structure orien-

tation". An analogous limitation of the construct validity becomes apparent in the low proportion of Winter's "Reflexion on one's *Denkhandeln* (*internal acting*)" (1995, p. 41) in the form of mathematical argumentation, which is in turn a substantial aspect of the basic experience of "problem-orientation".

With regard to the implementation of the scholastic standards in the conception of the tasks analysed in this paper, it must be critically stated that the requirement levels of all mathematical avtivities are unevenly distributed across their complete range of requirements. Primarily, this becomes clear when considering the fact that tasks that can be allocated to the highest, reflecting and generaliszing requirement level, are missing. From a diagnostic point of view, this also raises the following question: To what extent does the ZP 10 adequately fulfil its main purpose - i.e. testing skills and knowledge in accordance with the scholastic standards - especially when it comes to learners who show a high level of performance? This obvious lack of veridicality – one of the main catalysing variables for leraning success (cf. Helmke, 2004, p. 94) – thus, seems to encompass the feedback currently provided by the ministry, which reflects teaching success according to set competence standards, shows necessities for affirmative action and helps teachers to deduce measures for individualized education. A question that is both open as well as of great interest is the one concerning the effect of a thus limited requirement distribution in the context of mathematical activities on the controlling of teaching on the meso- and micro level of the North-Rhine Westphalian education system. Here, it seems first of all possible to interpret test tasks themselves, rather than curricular standards, as summative tools used to standardise teaching and that washback effects with regard to lesson design will be identifiable. Clear indications for the effect of this decoupling are, on the one hand, provided by already existing research fidings (cf. Maier et al., 2011) and, on the other hand, by the existence of an already developed, diverse supplement industry, which provides teachers with i.a. workbooks, practice tests as well as material for planning und teaching lessons.

5.2 Outlook

From the perspective of the basic experiences according to Winter as well as from the perspective of a more comprehensive assessment of learners' performance, there emerge possible approaches to an improvement of ZP 10 task culture. These approaches include the establishment of a higher level of cognitive challenges through tasks as well as a form of learning that is directed towards conncetion and comprehension. Implementing these approaches would enable mathematics to contribute profoundly to general education/knowledge

(*Allgemeinbildung*), which should be given expression in the designing of lessons. Because, against the background of the well-documented positive effect of final examinations on the conception of class tests with regard to content-related *and* cognitive characteristics of tasks, the conception of centralised tests according to standards is pivotal (cf. Drüke-Noe 2014; Meier et al. 2011). The adoptions of concrete tasks from cenralised examinations suggest that the amount of tasks that require more challenging activities should also be carefully increased for the ZP10. This conscious effect of control could also be adapted to a consideration of the topic area of stochastics, which has not yet been considered adequaltely in the desinging of ZP 10 until now.

Against this background, continued education and training for teachers in general as well as for eachers designing the tasks for such examinations would be desirable. In the context of cognitive activation, the goal of such continued education and training should be to make teachers aware of different task characteristics as well as the consious consideration of said characteristics in the process of choosing tasks for ZP 10, class tests and the everyday mathematics lesson.

References

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., Wittrock, M. C.: 2001, *A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of Educational Objectives*, New York: Longman.

Blum, W., vom Hofe, R., Jordan, A. Kleine, M.: 2004, Grundvorstellungen als aufgabenanalytisches und diagnostisches Instrument bei PISA, in: *Mathematische Kompetenzen von Schülerinnen und Schülern in Deutschland*, 145–157, Wiesbaden: VS Verlag für Sozialwissenschaften.

Blum, W., Drüke-Noe, C. R., Hartung, C. R., Köller, O.: 2006, *Bildungsstandards Mathematik: konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichtsanregungen, Fortbildungsideen,* Berlin: Cornelsen Verlag Scriptor.

Büchter, A. Leuders, T.: 2005; *Mathematikaufgaben selbst entwickeln. Lernen fördern-Leistung überprüfen,* Berlin: Cornelsen Scriptor.

Büchter, A., Pallack, A.: 2012, Methodische Überlegungen und empirische Analysen zur impliziten Standardsetzung durch zentrale Prüfungen, *Journal für Mathematikdidaktik* **33**, 59–85.

Drüke-Noe, Ch.: 2014, *Aufgabenkultur in Klassenarbeiten im Fach Mathematik. Empirische Untersuchungen in neunten und zehnten Klassen*, Wiesbaden: Springer Spektrum.

F r e u d e n t h a l, H.: 1983, *Didactical phenomenology of mathematical structures*, Dordrecht: Reidel.

H o f f a r t, E.: 2011, *Mathematische Vergleichsarbeiten in der Grundschule: Zum diagnostischen Potential von Aufgaben und deren Bearbeitungen einer landesweiten Vergleichsarbeit für dritte Klassen*, Universität Giessen.

J o r d a n, A., R o s s, N., K r a u s s, S., B a u m e r t, J., B l u m, W., N e u b r a n d, M., L ö w e n, K., B r u n n e r, M., K u n t e r, M.: 2006, *Klassifkationsschema für Mathematikaufgaben: Dokumentation der Aufgabenkategorisierung im COACTIV-Projekt*, Berlin: Max-Planck-Institut für Bildungsforschung.

J o r d a n, A., K r a u s s, S., L ö w e n, K., B l u m, W., N e u b r a n d, M., B r u n n e r, M., K u n t e r, M., B a u m e r t, J.: 2008, Aufgaben im COACTIV-Projekt: Zeugnisse des kognitiven Aktivierungspotentials im deutschen Mathematikunterricht, *Journal für Mathematikdidaktik* **29**, 83–107.

K l e i n, E. D., K r ü g e r, M., K ü h n, S., M., A c k e r e n, I.: 2014, Wirkungen zentraler Abschlussprüfungen im Mehrebenensystem Schule. Eine Zwischenbilanz internationaler und nationaler Befunde und Forschungsdesiderata, *Zeitschrift für Erziehungswissenschaft* **17**, 7–33.

K l i e m e, E., A v e n a r i u s, H., B l u m, W., D ö b r i c h, P., G r u b e r, H., P r e n z e l, M., R e i s s, K., R i q u a r t s, K., R o s t, J., T e r n o r t h, H. - E. V o l l m e r, H. J: 2003, *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise*, Bonn: Bundesministerium für Bildung und Forschung.

KMK: 2001, *Erste Konsequenzen aus der PISA Studie. Beschluss vom 05./06. 12.2001*, www.kmk.org/aktuell/pm011206.htm#ref1, 01.11.16.

KMK: 2003, *Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss–Beschluss der Kultusministerkonferenz vom 04.12.2003*, www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2003/2003_12_04-Bildungsstandards-MatheMittleren-SA.pdf, 10.09.2016.

K r ü g e r, M.: 2015, *Aufgabenkultur in zentralen Abschlussprüfungen: Exploration und Deskription naturwissenschaftlicher Aufgabenstellungen im internationalen Vergleich*, Münster: Waxmann.

K u n t e r, M., B a u m e r t, J., B l u m, W., K l u s m a n n, U., K r a u s s, S., N e u b r a n d, M.: 2013, *Cognitive activation in the mathematics classroom and professional competence of teachers. Results from the COACTIV project*, New York, Springer.

K ü h n, S., D r ü k e - N o e, C h.: 2013, Qualität und Vergleichbarkeit durch Bildungsstandards und zentrale Prüfungen? – Ein bundesweiter Vergleich von Prüfungsanforderungen im Fach Mathematik zum Erwerb des Mittleren Schulabschlusses, *Zeitschrift für Pädagogik* **6**, 912–932.

K ü h n, S.: 2013, Vergleichbarkeit beim Mittleren Schulabschluss? Ein Überblick über die Vielfalt schulstrukturell möglicher Bildungswege und Prüfungsverfahren in den deutschen Ländern, *Die Deutsche Schule* **105**, 87–101.

M a i e r, U., B o h l, T., M e t z, K. K l e i n k n e c h t, M.: 2011, Einflüsse von Merkmalen des Testsystems und Schulkontextfaktoren auf die Akzeptanz und Rezeption von zentralen Testrückmeldungen durch Lehrkräfte, *Journal of Educational Research Online* **3**, 62–93.

MSW: 2013, *Zentrale Prüfung 2013, Mathematik, Realschule*, Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen.

MSW: 2015, *Zentrale Prüfung 2015, Mathematik, Realschule*, Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen.

N e u b r a n d, M., K l i e m e, E., L ü d t k e, O. N e u b r a n d, J.: 2002, Kompetenzstufen und Schwierigkeitsmodelle für den PISA-Test zur mathematischen Grundbildung, *Unterrichtswissenschaft* **30**, 100–119.

N e u b r a n d, M., J o r d a n, A., K r a u s s, S., B l u m, W. L ö w e n, K.: 2011, Aufgaben im COACTIV-Projekt: Einblicke in das Potenzial für kognitive Aktivierung im Mathematikunterricht, in: *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*, 115–132, Waxmann.

N e u b r a n d, M., J o r d a n, A. (†), K r a u s s, S., B l u m, W., L ö w e n, K.: 2013, Task Analysis in COACTIV: Examining the Potential for Cognitive Activation in German Mathematics Classrooms (Chapter 7), in: *Cognitive Activation in the Mathematics Classroom and Professional Competence of Teachers*, 125–144, Springer.

OECD: 2004, *Learning for Tomorrow's World. First Results from PISA 2003*, OECD Publishing.

OECD: 2014, *PISA 2012 Results: What Students Know and Can Do*: *Student Performance in Mathematics, Reading and Science*, OECD Publishing.

R e i s s, K.: 2004, Bildungsstandards und die Rolle der Fachdidaktik am Beispiel der Mathematik, *Zeitschrift für Pädagogik* **50**, 635–649.

R e s n i c k, L. F o r d, W.: 1981, *The Psychology of Mathematics for Instruction*, Hillsdale: Lawrence Erlbaum Associates.

R o s e n s h i n e, B.: 1970, Evaluation of instruction, *Review of Educational Research* **40**, 279–300.

S c h e j a, B.: 2007, *Der Wandel von Bildungssystemen am Beispiel von Deutschland und Polen unter besonderer Berücksichtigung der PISA-Studie,* Dissertation, Universität zu Köln.
T e r h a r t, E., B a u m g a r t, F., M e d e r, N. v o n S y c h o w s k i, G.: 2009, Standardisierte Prüfungsverfahren in der Erziehungswissenschaft: Kontext, Formen, Konsequenzen, *Erziehungswissenschaft* **20**, 9–36.
W i n t e r, H.: 1995, Mathematikunterricht und Allgemeinbildung, *Mitteilungen der Gesellschaft für Didaktik der Mathematik* **61**, 37–46.
Z a b a l a, D., M i n n i c i, A., M c M u r r e r, J., B r i g g s, L.: 2008, *State high school exit exams: A movetoward end-of-course exams*, Washington, D. C., Center on Education Policy (CEP).

Kognitywna aktywizacja przez zadania matematyczne pochodzące z centralnych egzaminów poziomu ISCED 2 na przykładzie Nadrenii Północnej-Westfalii

S t r e s z c z e n i e

Egzaminy zewnętrzne sa powszechnymi elementami monitoringu i sterowania systemem szkolnym. Współczesne badania pokazują wyraźny związek między formą zadań egzaminacyjnych a sposobami pracy i treściami poruszanymi podczas prowadzenia lekcji matematyki. Zentrale Prüfung – ZP 10 – jest centralnym egzaminem przeprowadzonym na koniec nauki na stopniu secondary education w Nadrenii Pomocnej-Westfalii, najbardziej licznym landzie w Niemczech. Wyniki tych egzaminów, jak i tych osiągniętych w badaniach PISA w okresie miedzy 2003 i 2015 pokazują znaczny przyrost umiejętności niemieckich uczniów. Z tego powodu interesujące staja się przeanalizowanie zadań z egzaminów zewnętrznych na poziomie ISCED 2, aby móc określić elementy, których wpływ na takie postępy możemy przyjąć jako znaczny. Artykuł poświęcony jest badaniu kognitywnej aktywizacji przez zadania matematyczne pochodzące z ZP 10. Przez kognitywną aktywizację rozumiane są kategorie programowe, treściowe (content-related) i kognitywne, które potencjalnie mogą być zidentyfikowane w procesie rozwiązania zadania. Te analizę poprzedza przedstawienie standardów programowych i dyskusja dotycząca funkcji zadań matematycznych. Pomiar kognitywnej aktywizacji został przeprowadzony poprzez analizę pewnego zbioru zadań egzaminacyjnych, użyte we wszystkich typach szkol Nadrenii Północnej-Westfalii na poziomie ISCED 2 w okresie pomiędzy 2007 – (rok wprowadzenia narzędzia) i 2016 ($N = 1706$).

University of Cologne
Institut für Mathematikdidaktik
Gronewaldstraße 2
50931 Köln GERMANY
e-mail: *schejab@uni-koeln.de*