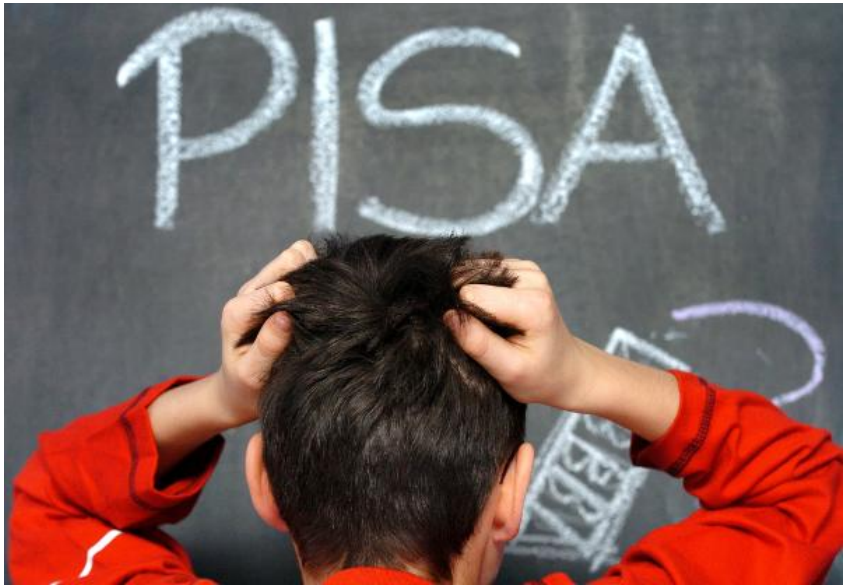
Institut zur Qualitätsentwicklung
im Bildungswesen

Student Assessment in Germany: Present and Future Initiatives

Prof. Petra Stanat, Ph.D.

Presentation at the
National Assessment
Governing Board Meeting
Washington, DC
November 17, 2017

The so-called „PISA shock“



Results of PISA 2000 for Germany

- mean reading literacy significantly below the mean of OECD member states
- large variance of achievement scores
- particularly poor results at the lower end of the achievement distribution
- pronounced disparities associated with students´
  - socio-economic background
  - migration background
- large achievement differences among the 16 states

Comprehensive strategy for educational monitoring of the 16 states

KULTUSMINISTER KONFERENZ

1. Participation in international large-scale assessments of student achievement (PISA, PIRLS, TIMSS)

2. Testing and implementation of educational standards for primary school, secondary level I, and secondary level II
   - national assessment studies at the system level in primary schools (grade 4) and secondary level I (grade 9)
   - pool of tasks for school-leaving exam qualifying for university admission (grade 12/13)

3. Tools for quality ensurance at the school level (e.g., comparison tests VERA in grades 3 and 8)

4. Comprehensive educational reporting (every 2 years)

Comprehensive strategy for educational monitoring of the 16 states

KULTUSMINISTER KONFERENZ

1. Participation in international large-scale assessments of student achievement (PISA, PIRLS, TIMSS)

2. Testing and implementation of educational standards for primary school, secondary level I, and secondary level II
   - national assessment studies at the system level in primary schools (grade 4) and secondary level I (grade 9)
   - pool of tasks for school-leaving exam qualifying for university admission (grade 12/13)

3. Tools for quality ensurance at the school level (e.g., comparison tests VERA in grades 3 and 8)

4. Comprehensive educational reporting (every 2 years)

IQB

Comprehensive strategy for educational monitoring of the 16 states

KULTUSMINISTER KONFERENZ

1. Participation in international large-scale assessments of student achievement (PISA, PIRLS, TIMSS)

2. Testing and implementation of educational st ndards for primary school, secondary level I, and s level II
   - national assessment studies evel in primary schools (grad dary level I (grade 9)
   - pool of tas eaving exam qualifying for univ on (grade 12/13)

3. Tools for quality ensurance at the school level (e.g., comparison tests VERA in grades 3 and 8)

4. Comprehensive educational reporting (every 2 years)

Largely low stakes

IQB

The Institute for Educational Quality Improvement

- founded in 2004
- independent academic institute, located at the Humboldt University Berlin
- financed by the 16 federal states („Länder“) in Germany
- interdisciplinary team
  psychologists, educational researchers, teachers, psychometricians
- expertise on subject-matter content: cooperation with other universities and research institutes

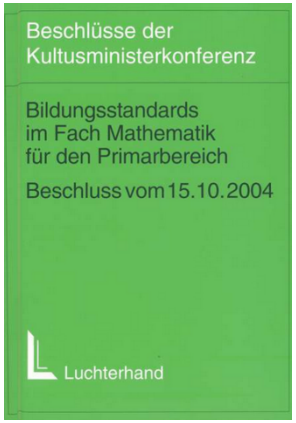
Basis of the national assessment system: National Educational Standards

- Adopted for core subjects by the Standing Conference of the Ministers of Education and Cultural Affairs (KMK)
- Describe core elements of knowledge and skills students should, on average, have acquired by the end of a certain stage in their educational career
- Implementation is mandatory for all 16 states
- Alignment of state-specific curricula and central exams with the national standards

KULTUSMINISTER KONFERENZ

Educational Standards in Germany

	Primary school	Secondary level I		Secondary level II
		Lower Degree	Intermediate Degree	Academic Degree
German	2004	2004	2003	2012
Mathematics	2004	2004	2003	2012
First foreign language (English, French)	-	2004	2003	2012
Biology, Chemistry, Physics	-	-	2004	2020



From standards to feedback

Output Domain (e.g., skills, cognitive processes)

Performance Standards: Structure and levels of domain-specific competence

Content Standards (e.g., state curriculum framework)

Content Domain

From standards to feedback

Evidential Aspects of Validation
(Procedural, Internal, and External Aspects)

Output Domain (e.g., skills, cognitive processes)

Item Domain

Performance Standards: Structure and levels of domain-specific competence

Content Standards (e.g., state curriculum framework)

Content Domain

Test Specifications

Test (Tasks /Item Bank)

Standard Setting (Cut Scores)

Test domains

German	reading listening orthography language and reflection on language (school-level assessment only) writing (school-level assessment only)
English/French	reading listening
Mathematics	five main ideas (e.g., measurement, data and odds)
Biology **Chemistry** **Physics**	four competence domains (e.g., using scientific knowledge, generating scientific knowledge)

Proficiency levels: English listening comprehension, intermediate degree (excerpts)

Based on the Common European Framework of Reference for Languages (2001)

V Optimal Standard (B2.2, C1)	Can understand enough to follow longer input on unfamiliar and abstract topics. Understands a broad spectrum of idiomatic phrases and colloquial expressions. Can follow longer monologues and conversations even if they are not clearly structured. [C1]
IV Norm Standard Plus (B2.1)	Can understand the main propositions of standard input on concrete and abstract topics, even if content and language are complex. Can follow longer input and complex argumentations if the topic is to some extent familiar and it is structured by explicit signals. [B2.1]
III Norm Standard (B1.2)	Can understand factual information with low complexity on common everyday and work-related topics, is able to understand main propositions and individual pieces of information if the input is clearly articulated and the accent is familiar. [B1.2]
II Minimal Standard (A2.2, B1.1)	Can understand main points of clearly articulated standard input on familiar matters regularly encountered in school, leisure, etc. [B1.1]
I Below Minimal Standard (A1, A2.1)	Can understand expressions and words related to matters of immediate relevance. [A2.1]

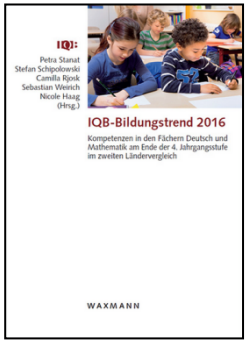
From standards to feedback

Evidential Aspects of Validation
(Procedural, Internal, and External Aspects)

Consequential Aspects of Validation
(Utility and Impact Aspects)

Output Domain (e.g., skills, cognitive processes)

Item Domain



Performance Standards: Structure and levels of domain-specific competence

Content Standards (e.g., state curriculum framework)

Content Domain

Test Specifications

Test (Tasks /Item Bank)

Standard Setting (Cut Scores)

Test Use and Interpreta-tions (Intended & Actual)

Test Impact

Assessments in Germany

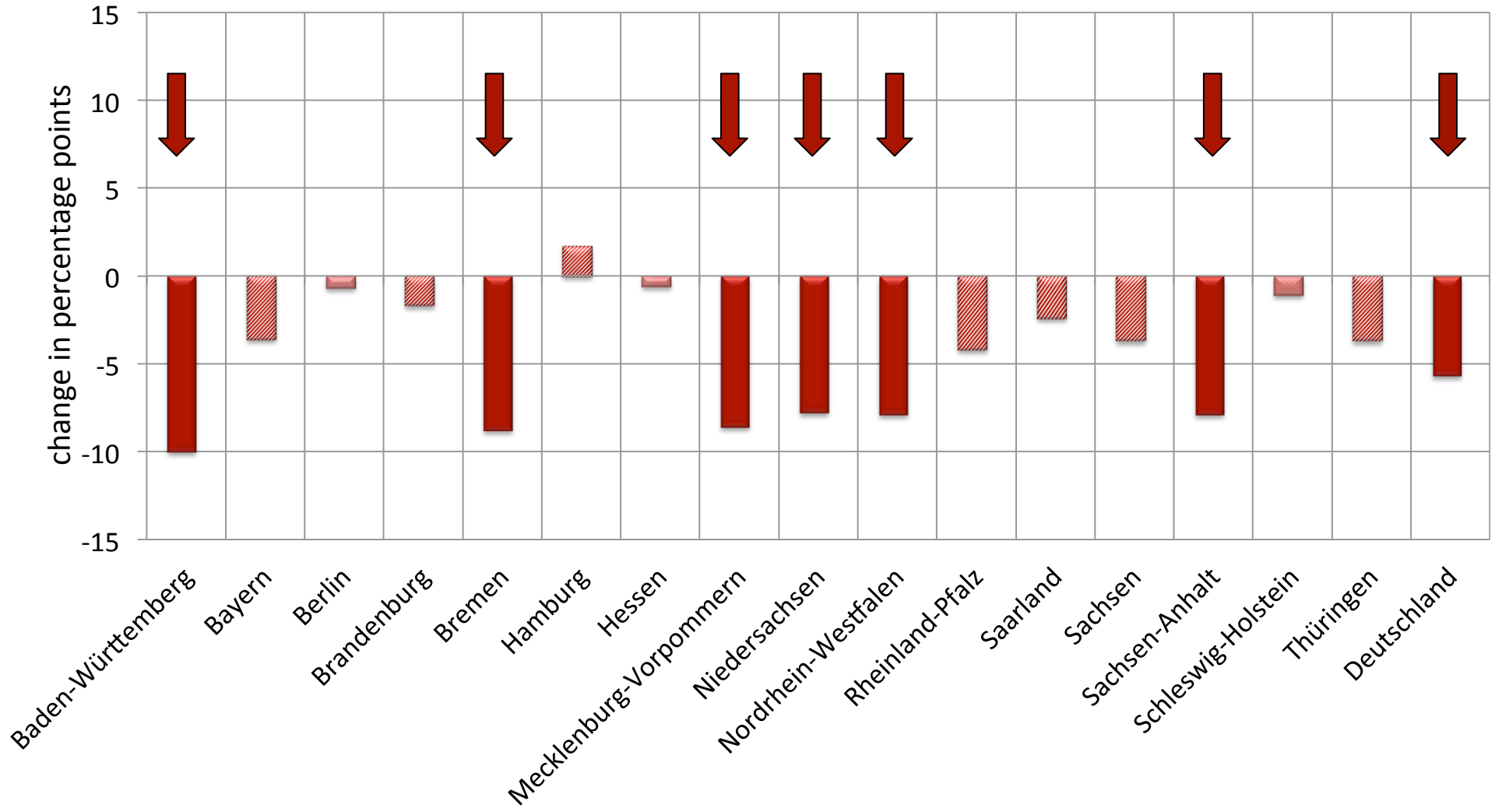
	International *PISA, PIRLS, TIMSS*	***National IQB Assessments*** *Across the 16 states*	**School-level IQB Assessments (VERA)** *Within the 16 states*
Standards-based tests?	No	Yes	Yes
Who is tested?	Sample (approx. 4,000-5,000)	Sample (approx. 30-40,000)	Population
Frequency	3 years, 5 years	3 years, 5 years	Every year
Main Purpose	System monitoring	System monitoring	School & teaching improvement
Who is accountable?	Federal Ministry of Education; 16 State Ministries of Education	16 State Ministries of Education and their school authorities	Principals, teachers

Assessments in Germany

	International *PISA, PIRLS, TIMSS*	***National IQB Assessments*** *Across the 16 states*	**School-level IQB Assessments (VERA)** *Within the 16 states*
Standards-based tests?	No	Yes	Yes
Who is tested?	Sample (approx. 4,000-5,000)	Sample (approx. 30-40,000)	Population
Frequency	3 years, 5 years	3 years, 5 years	Every year
Main Purpose	System monitoring	System monitoring	School & teaching improvement
Who is accountable?	Federal Ministry of Education; 16 State Ministries of Education	16 State Ministries of Education and their school authorities	Principals, teachers

National IQB Assessments, 2nd Cycle:
“IQB Educational Trends“

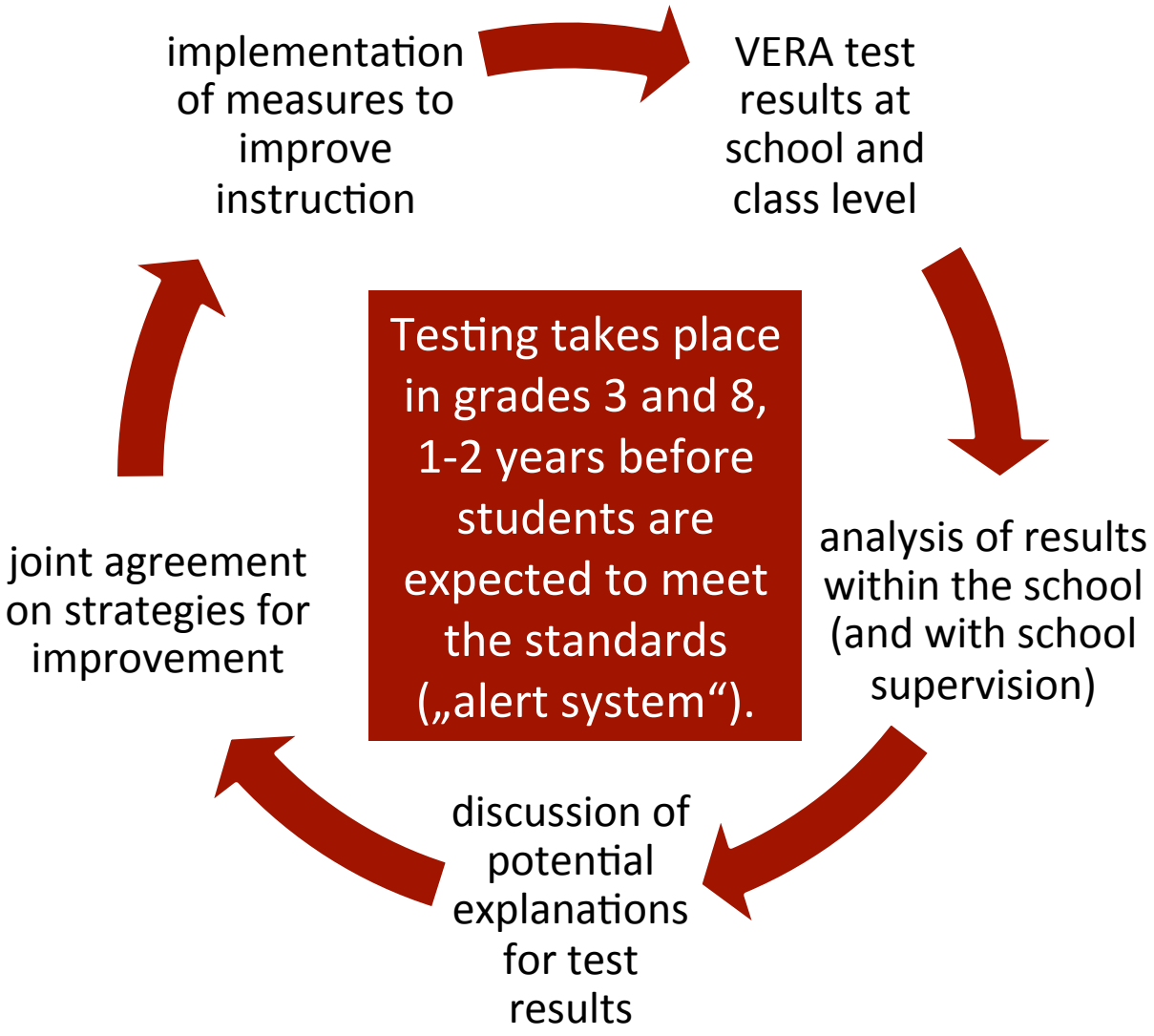
Changes in the proportion of students (grade 4) reaching at least the norm standard („Regelstandard“) **between 2011 und 2016**: **mathematics**



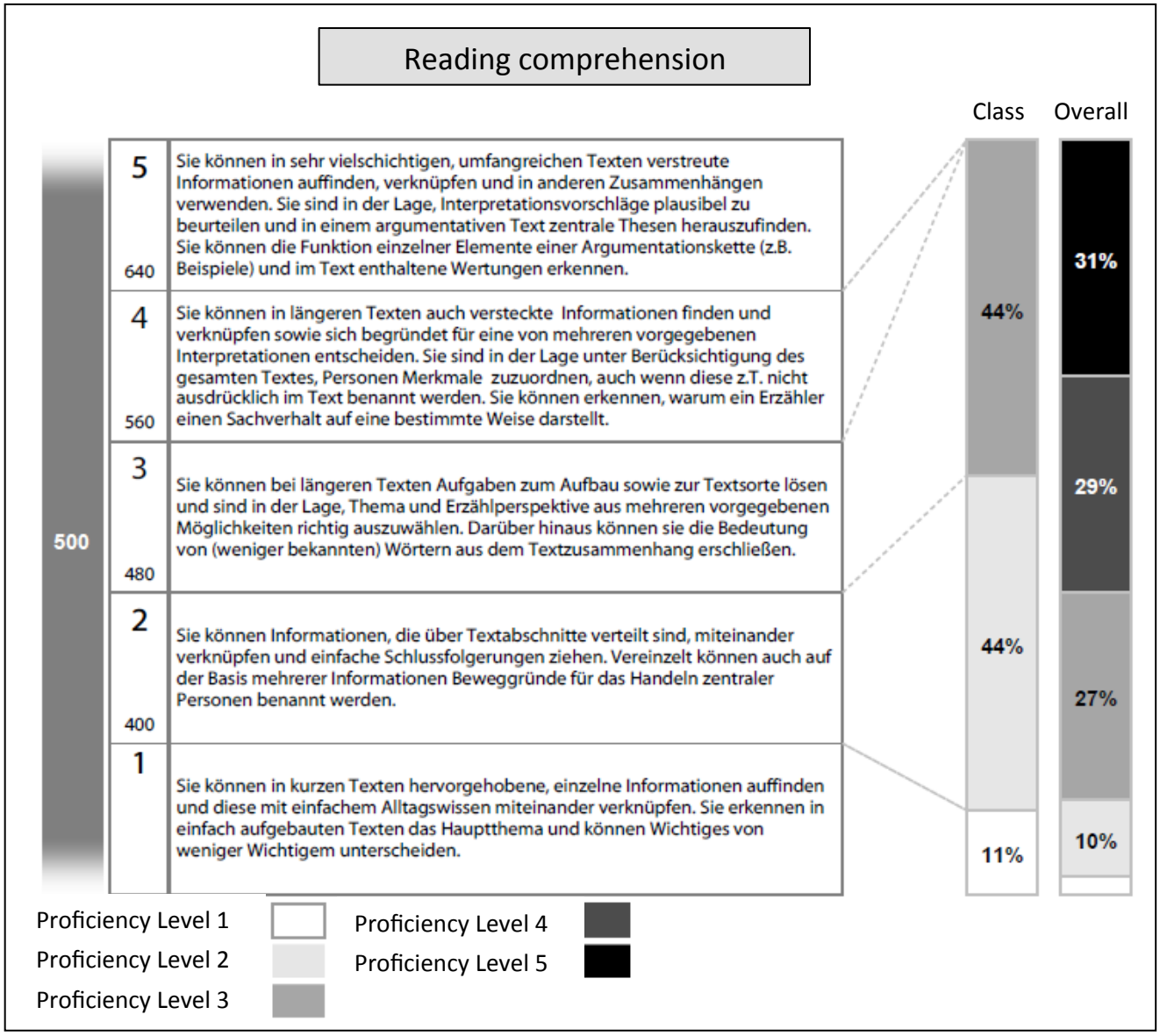
Assessments in Germany

	International *PISA, PIRLS, TIMSS*	*National IQB Assessments* *Across the 16 states*	**School-level IQB Assessments (VERA)** *Within the 16 states*
Standards-based tests?	No	Yes	Yes
Who is tested?	Sample (approx. 4,000-5,000)	Sample (approx. 30-40,000)	Population
Frequency	3 years, 5 years	3 years, 5 years	Every year
Main Purpose	System monitoring	System monitoring	School & teaching improvement
Who is accountable?	Federal Ministry of Education; 16 State Ministries of Education	16 State Ministries of Education and their school authorities	Principals, teachers

Goals of the school-level IQB assessments



VERA-feedback for schools (example) – approaches differ between states



Student Assessment in Germany: Future Initiatives

- Feasibility study for computer-based assessment (starting with VERA).
- Development of modules for each test domain from which states/ schools/teachers can choose (mandatory core module + optional modules for different levels of proficiency).
- Initiatives to improve usage of test results as a tool for developing instructional quality (important focus).
- Continued and improved use of large-scale assessments for further research (e.g., effects of different schooling models on students with special education needs, longitudinal addition to PISA 2012/IQB-Bildungstrend 2012).

Institut zur Qualitätsentwicklung
im Bildungswesen

iqboffice@iqb.hu-berlin.de

https://www.iqb.hu-berlin.de/

Thank you for your attent on!

Examples for educational standards: English listening comprehension, intermediate degree

Students are able to…

- follow the main points made in longer conversations,
- understand announcements on concrete issues that are spoken at normal speed in standard language,
- understand presentations if they are clearly structured and their complexity is limited and if students are familiar with the topic,
- understand the main information presented on the radio and on television regarding topics that are of personal interest.

Differences in the proportion of students (grade 4) reaching at least the norm standards in **2016** between the 16 states and Germany overall: **mathematics**