VALIDITY:
Multiple Means to Provide Trustworthiness

David Mandel

Presentation at RILS 2013

Two Lines of Inquiry

- Standards – Fidelity in specifying what students need to know and be able to do

- Assessments – Trustworthiness in measuring student command of standards and in predicting the likelihood of their success in college and careers

But What Colleges & What Careers?

- Presumption that all colleges place equal demands on students – Highly Unlikely

- Careers that offer a ticket to the middle class (i.e., careers that require some postsecondary education)

How to Arrive at Valid Standards

- Examine exceptional high schools and poll their teachers
- Impanel sterling college faculty members
- Interview business executives

OR

- Systematically examine what actually happens

What Does it Really Mean to be College and Work Ready

- Secure syllabi, texts and graded student work from first-year classes in 7 randomly selected community colleges from 7 states

- Draw on faculty from 9 areas of study (e.g., autotech, business, criminal justice, IT, nursing) and from these programs' required English and mathematics courses

- Create an expert panel to rate and evaluate the empirical evidence

Mathematics Findings

Chart 1: Average Percent of Text Chapters and Exam Items Containing CCSS Domains for All Subject Matter (101) Courses
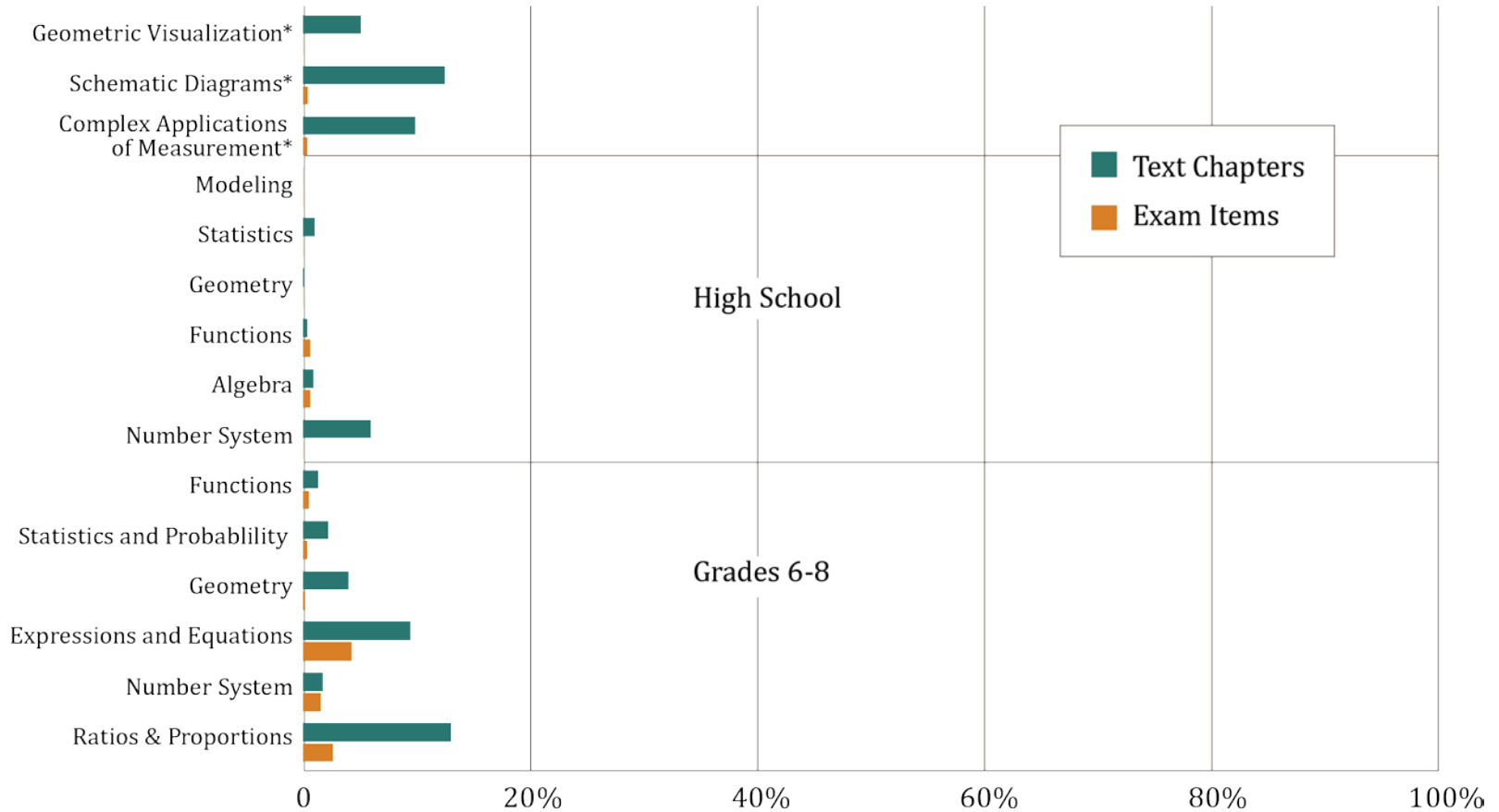
Chart 3: Average Percent of Text Chapters and Exam Items Rated at Each PISA Level for All Subject Matter (101) Courses
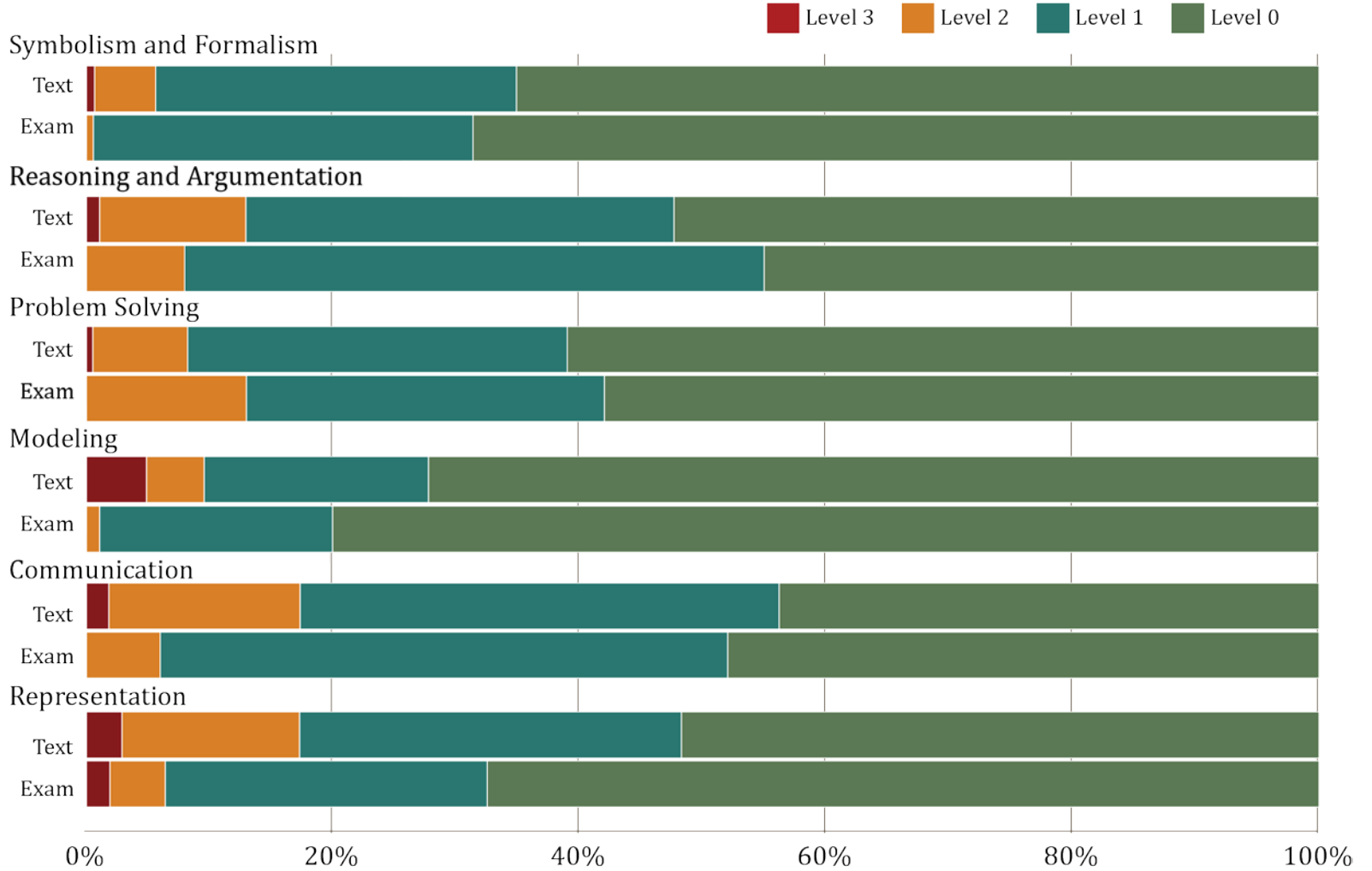
Legend: Level 3, Level 2, Level 1, Level 0

Symbolism and Formalism — Text, Exam
Reasoning and Argumentation — Text, Exam
Problem Solving — Text, Exam
Modeling — Text, Exam
Communication — Text, Exam
Representation — Text, Exam

Chart 6: Average Frequency of Prerequisites Called For by All Required Mathematics Courses: Based on the Percent of Text Chapters and Exam Items that Utilize Each CCSS Domain
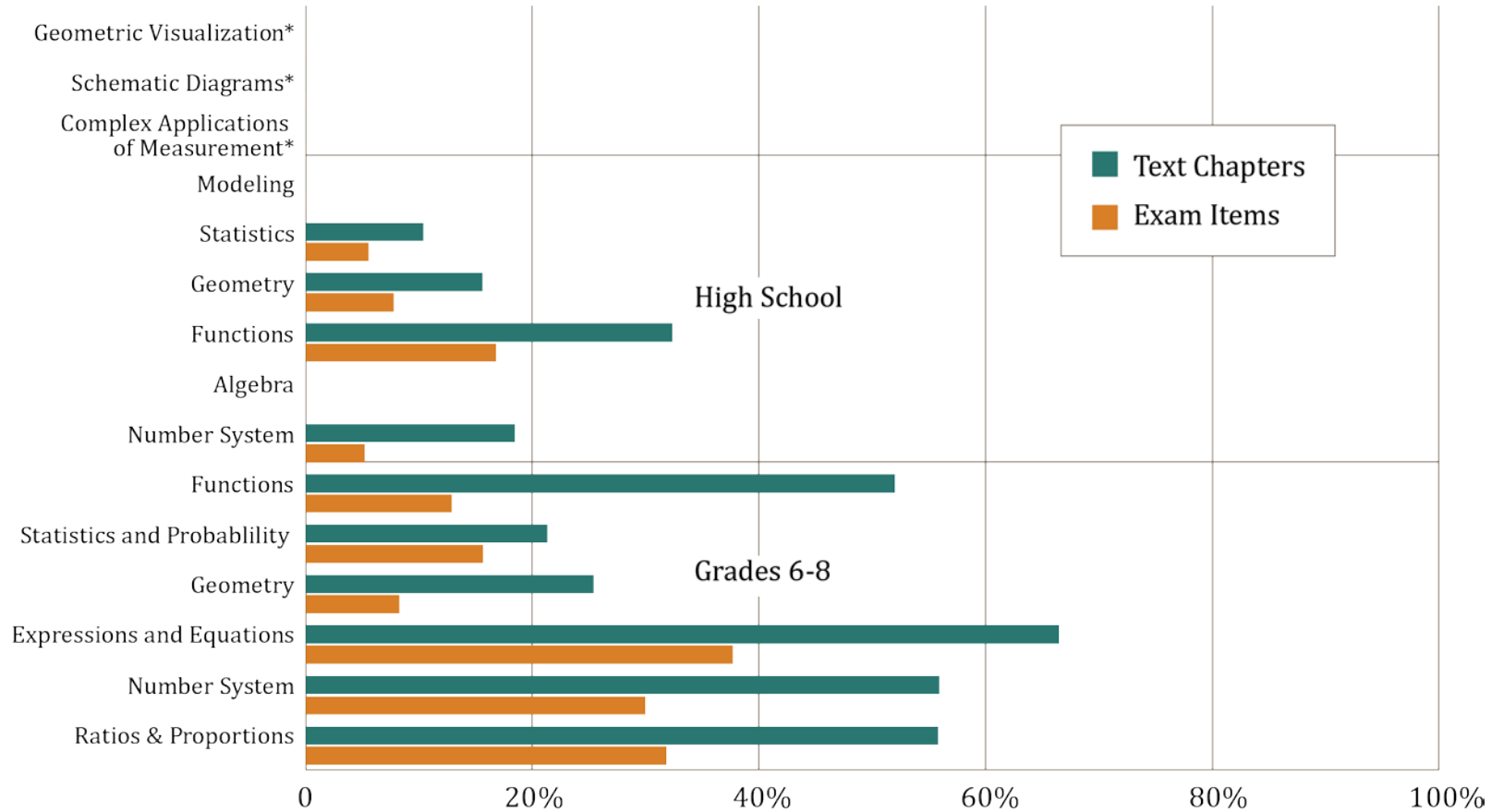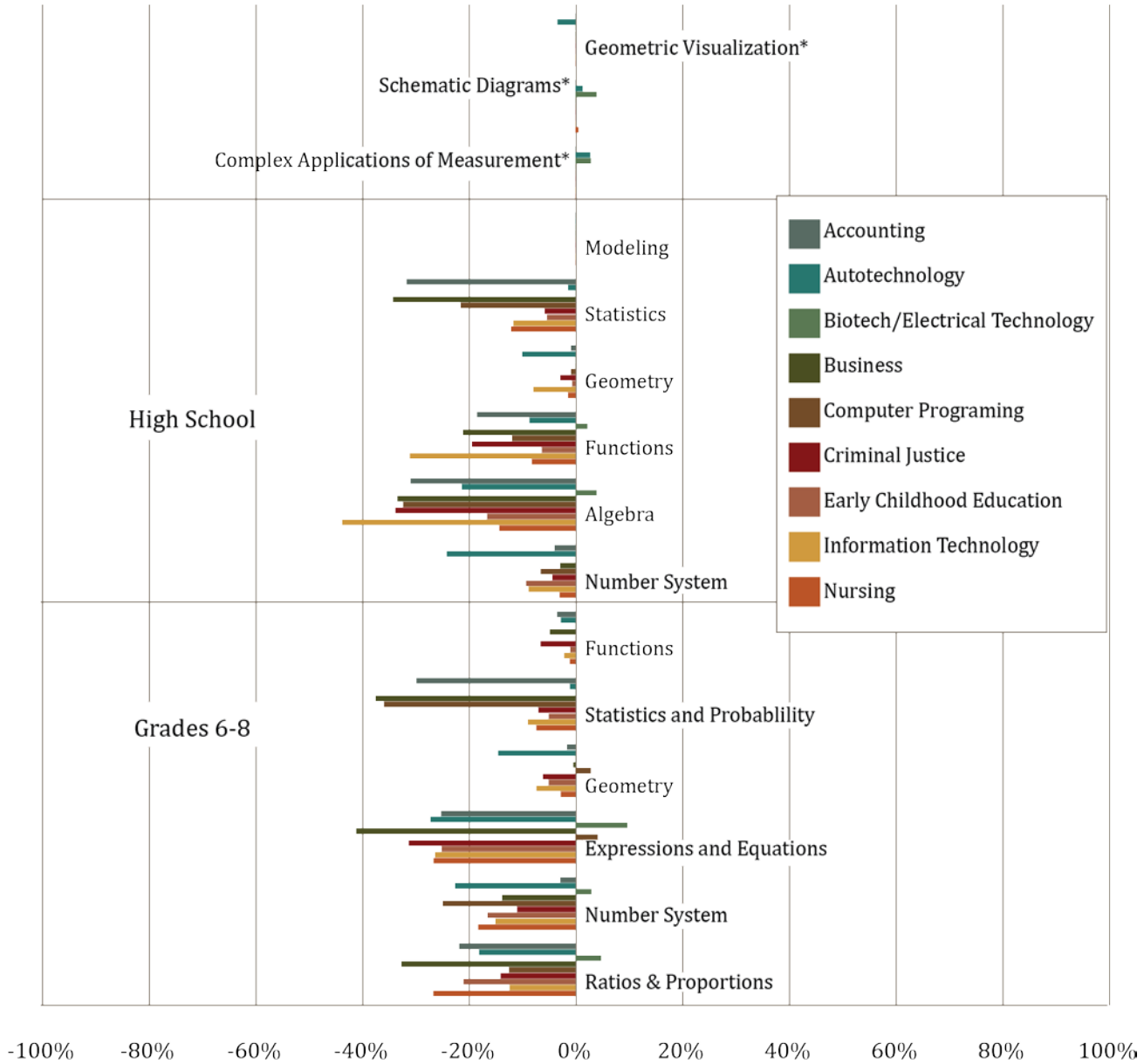
Chart 15: Discrepacy Between Average Percent of Exam Items Containing Mathematics Content in 101 Courses and Average Percent of Exam Items Containing Mathematics Content in Required Mathematics Courses, All Programs

English Findings

Chart 1: Text Difficulty Comparison: Grade Ratings for Texts in Introductory Courses in Select Programs of Study
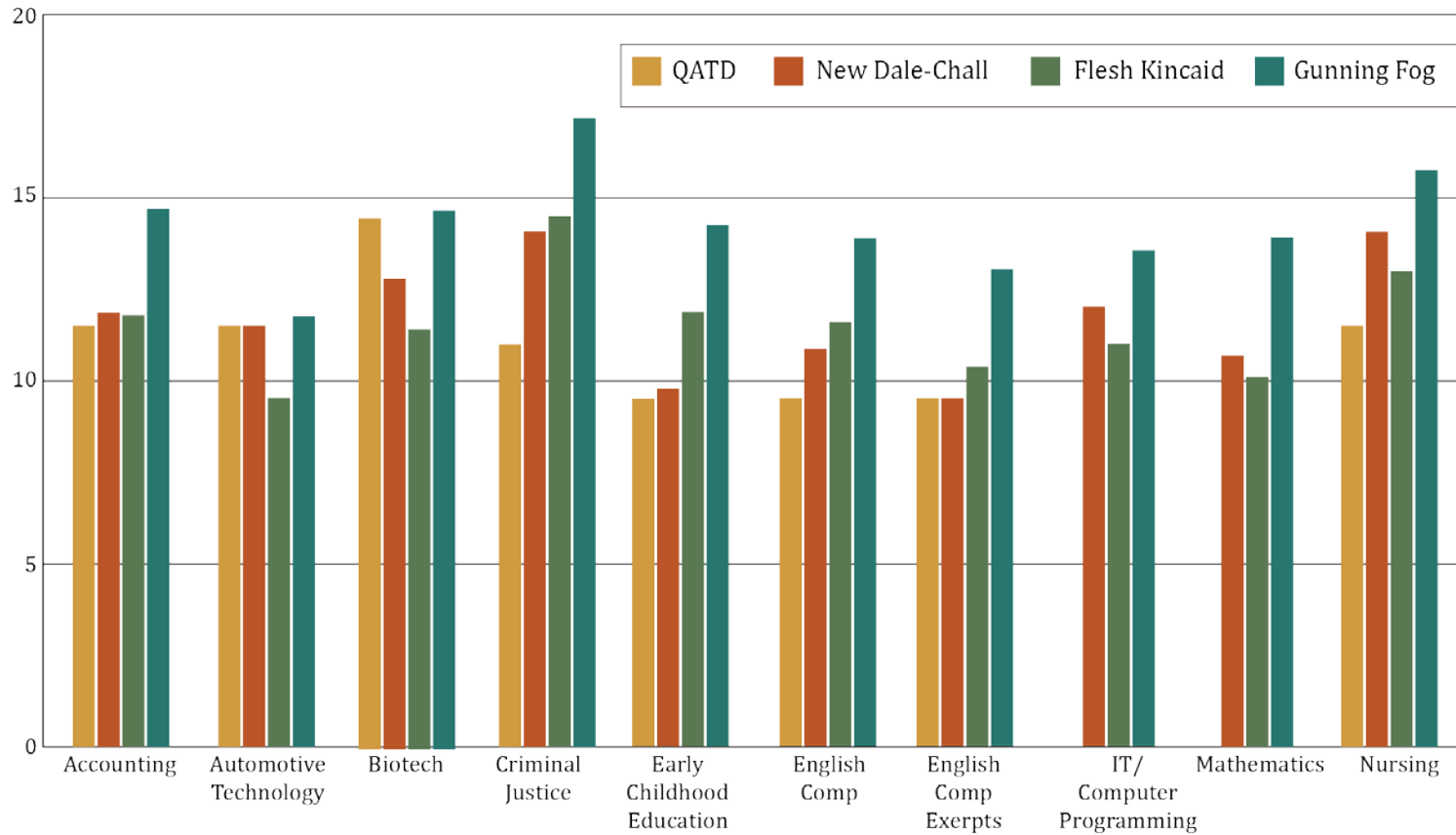
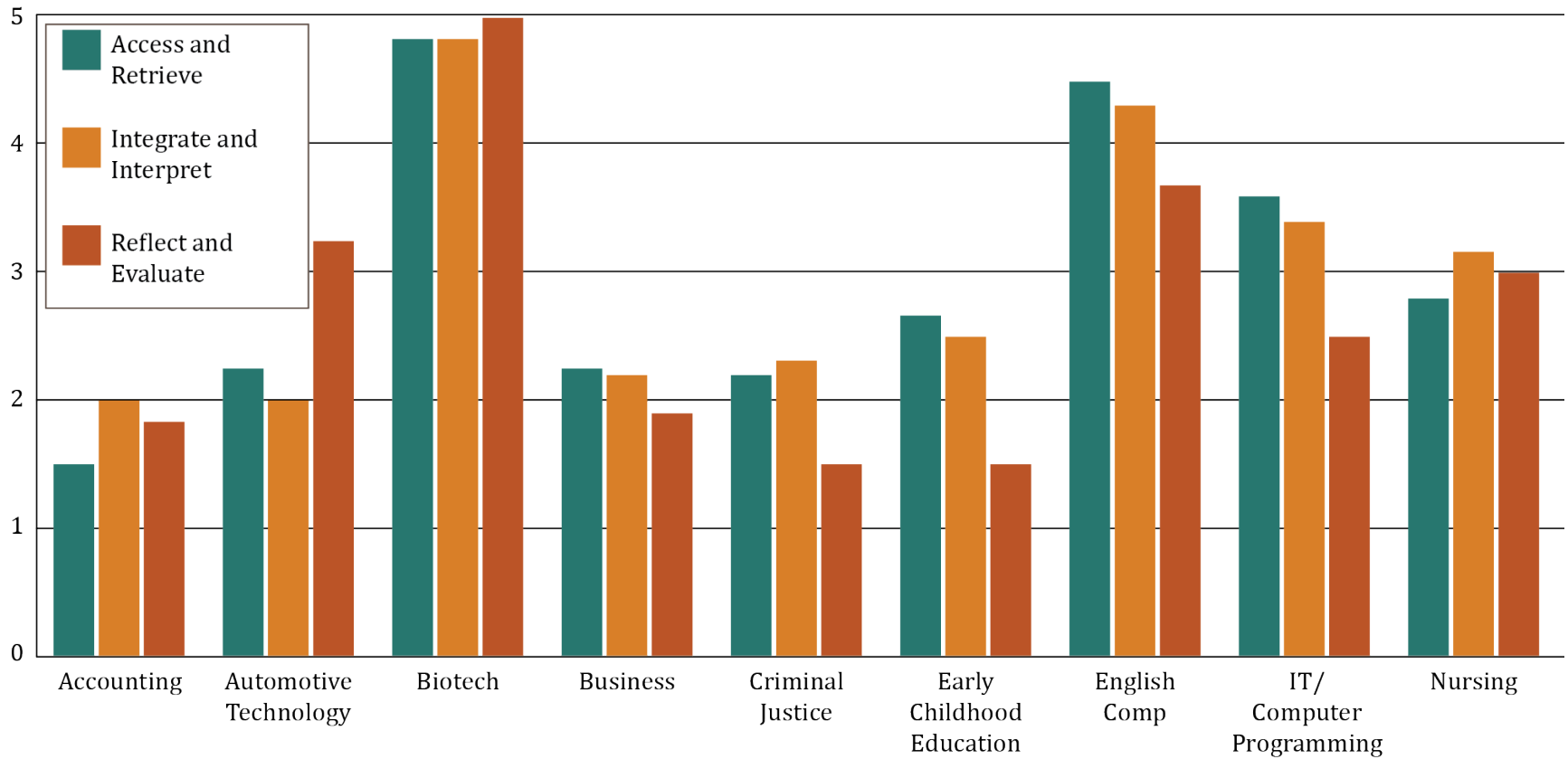# Chart 2: Average Reading Task Complexity Ratings by Type and Program

Chart 6: Panel Ratings of Essays Scored College Ready and Not College Ready by Instructors

Instructor Grade

Not College Ready

College Ready

0%   20%   40%   60%   80%   100%

Panelist Ratings

Not College Ready

Rating 1

Rating 1.5

Rating 2

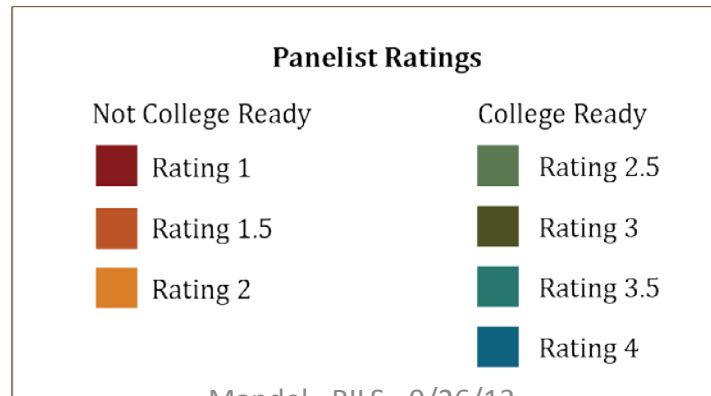College Ready

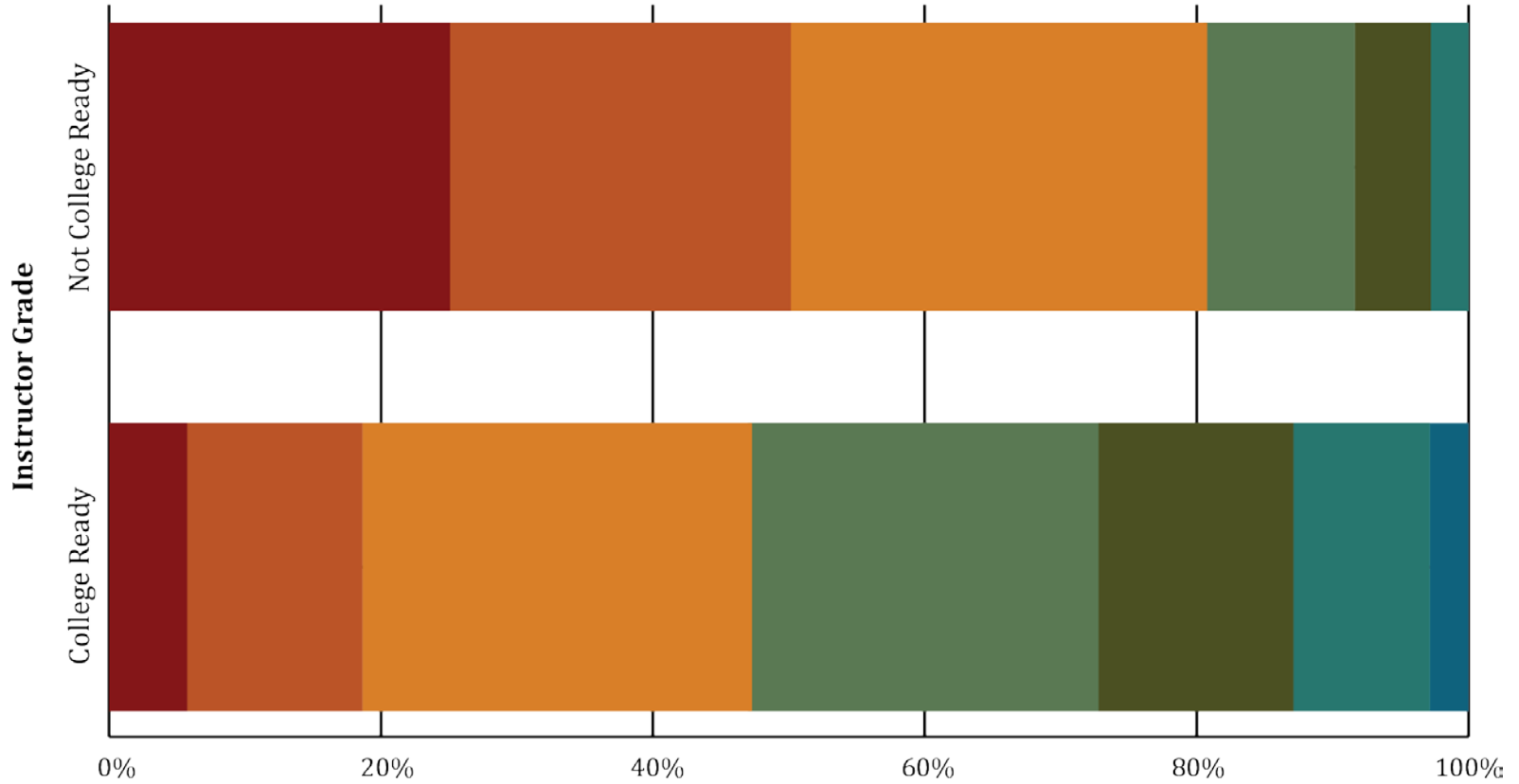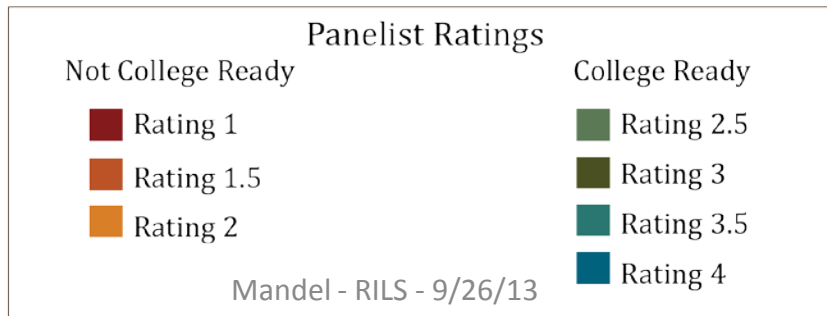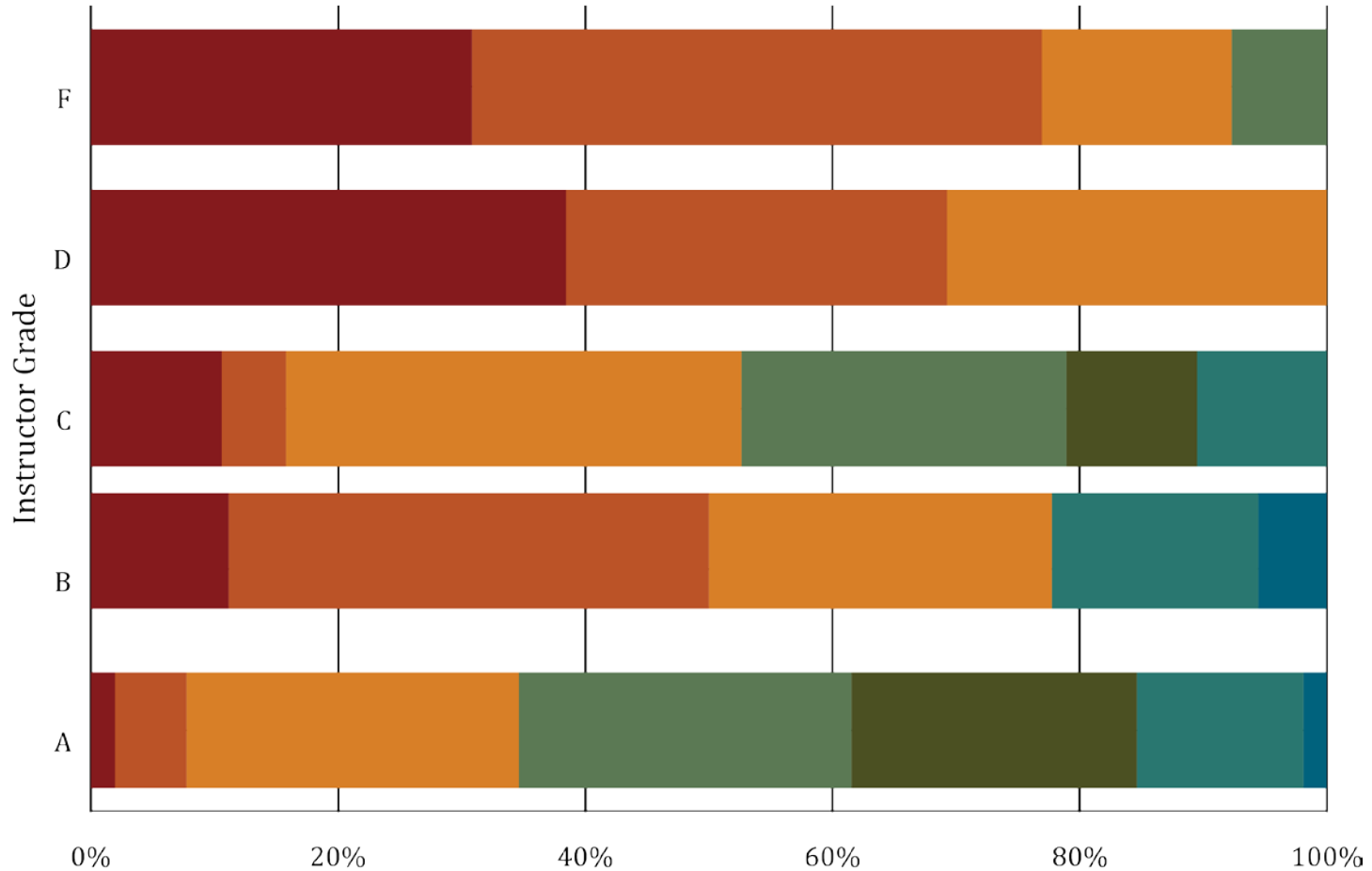Rating 2.5

Rating 3

Rating 3.5

Rating 4

Chart 7: The Distribution of Panel Ratings of Student Essays over Instructor Grades

The Assessment Challenge

- Alignment of Assessments to the Standards
  - Going beyond counting widgets
  - Consider not just the cognitive demands of the tasks, but of the stimuli as well
  - Joining careful mapping across multiple dimensions with holistic judgments

- Analyzing the Relationship between High School Performance and College Results

Analyzing the Relationship between High School Performance and College Results

- At the outset of a new system there is no longitudinal data to lean on as none of the initial HS students have attended college

- Options
  - Double-test students using predictors of college success – PSAT/SAT and Explore/PLAN
  - Begin collecting naturally occurring data points by tapping state longitudinal data systems OR sampling district student records systems

Analyzing the Relationship between High School Performance and College Results

- Once the initial HS students move on to college we can directly study the extent to which their HS exam results predict college success by examining their first year grades in credit bearing courses

- But should we be placing all of our bets on first year performance?

The Central Role of Year I College Grades

- Strong predictor of college completion

- Better than acceptance into college, an extraordinary low bar at many institutions

- But often awarded by junior faculty and teaching assistants – are they the best judges on which to rest validity determinations?

- Are they also a trustworthy predictor of career success?

The Central Role of Year I College Grades

- But what if initial college courses are not well aligned to current and future labor market requirements?

- We need the equivalent of CCSS for college fields of study

- If college curricula do not have a high fidelity connection to the world their students are about to enter, much of the effort to improve K-12 could be misdirected

The Challenges to Validity and How to Overcome Them

- A tradition of valuing reliability over validity
  - Other countries take a markedly different approach – e.g., Cambridge International Examinations

- Dominance of Multiple-Choice Exams
  - Greater focus on open-ended tasks and coursework

- Addiction to cheap tests over high quality tests
  - Trading off frequency for quality

- Especially crucial at a time when education is being urged to place more value on command of knowledge in depth and on cognitively demanding tasks

The Benefits of Raising the Validity Bar

- Increase the odds that the quality of education will rise as it must

- Free teachers from preparing students for tests that are at odds with accomplished practice

- Encourage teachers to practice in the best interests of students and thus make test prep an archaic practice