



investing  
in  
innovation fund



# Learning from the National Evaluation of i3: Challenges, responses, and future plans

i3 Learning Community Meeting  
September 11-12, 2014



# What We Have Learned



- A grant program/agency can incentivize high-quality evaluations
- Across the first 3 cohorts, the majority of the studies have the potential to meet NEi3 standards with or without reservations
  - 89% of all grants have potential to meet standards
  - 83% of development grants using RCTs, QED (twice as many QEDs)
  - 100% of validation and scale-up grants using RCTs or QEDs (about half and half)

# What We Have Learned



- At the same time, this emphasis and the success of helping evaluators design strong studies does not eliminate risks that can undermine what can be learned
  - Low power: 17 development grants (as designed), 3 validation grants (as implemented)
  - Serious confounds: 6 development grants, 2 validation grants, 1 scale-up grant (1 of multiple studies)
  - Only studying part of the full i3-supported intervention or sample: 2 development grants, 2 validation grants, 1 scale-up grant
  - Dropping down in rigor of design: 6 development grants proposed QEDs but ended up implementing promise evaluations

# What We Have Learned



- Institutionalizing rewards for prior evidence has increased attention on/efforts to implement strong studies
  - OII and other federal agencies are increasingly focused on evidence—as rationale for funding and as a result of funding
  - OII/IES are increasingly clear about the evidence requirements for winning another i3 grant—bar has been set higher
- In earlier meeting, this link to evidence cited as one definition for the “innovation” in Investing in Innovation

# What We Have Learned



- A tiered evidence program such as i3 faces issues about appropriate design standards for lower level grants
  - Many of the development grants are developing, investigating innovative approaches
  - Guidance on rigorous designs may be misaligned with stage of intervention
    - Focus on implementation may be more aligned?
    - How to weigh innovation versus need for believable evidence

# What We Can Learn



- Effectiveness of each individual intervention/approach at improving student achievement
  - Which interventions have been shown to be “effective” and “promising”
- This is the smallest ‘grain size’

# What Else Can We Learn



- Can we learn something more systematic about educational interventions, can we go beyond “one-off” evaluations
- Is there a way to group interventions and summarize findings within groups
  - WWC has moved away from topic area summaries and towards intervention reports
  - Dangers of summarizing effects of groups of i3 interventions

# What Else We Can Learn



- Group the i3 interventions by
  - Absolute Priority?
  - Outcome domains?
  - Intervention components?



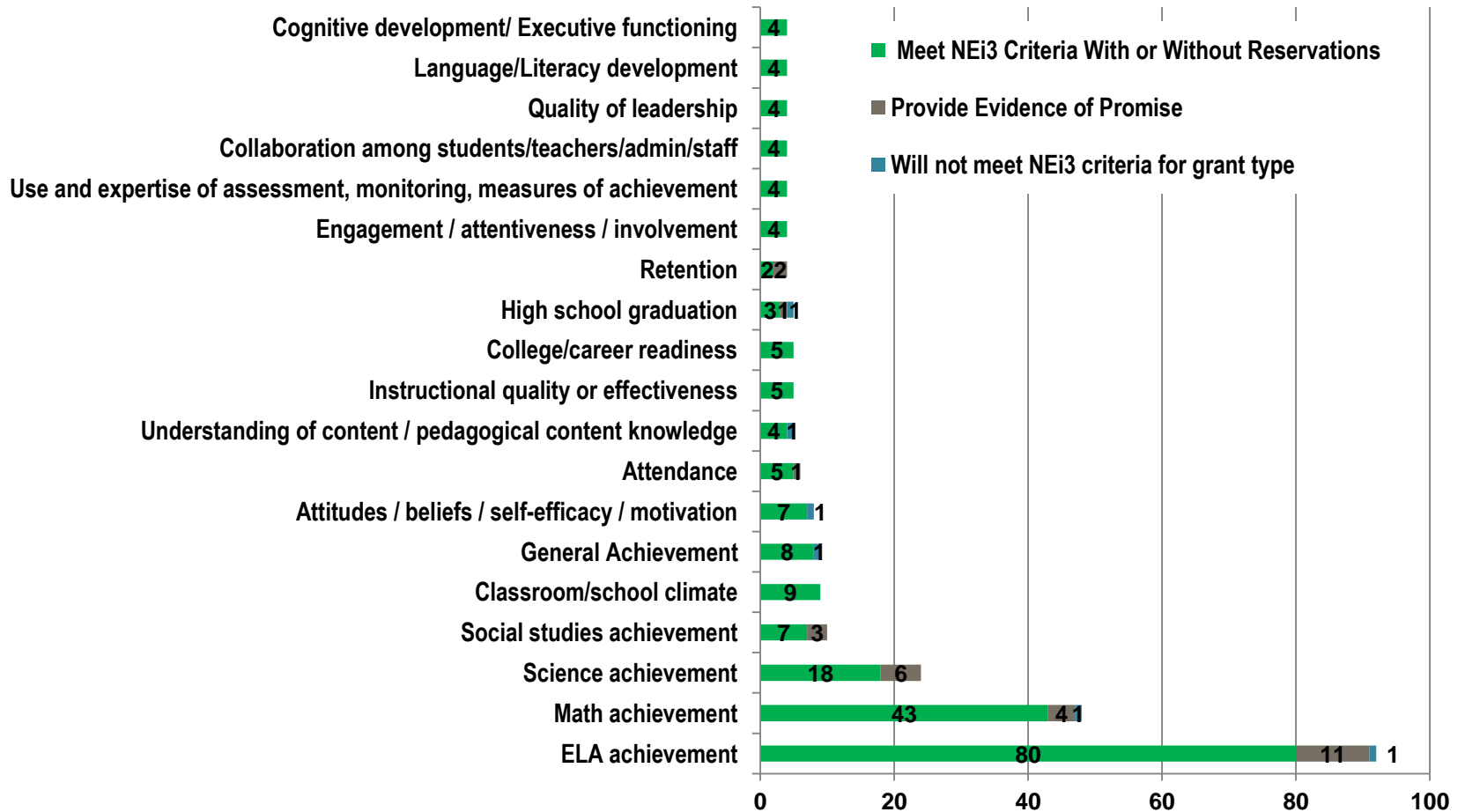
# Group by broad policy area/goal



## FY2010-2013 i3 Grants by Absolute Priority

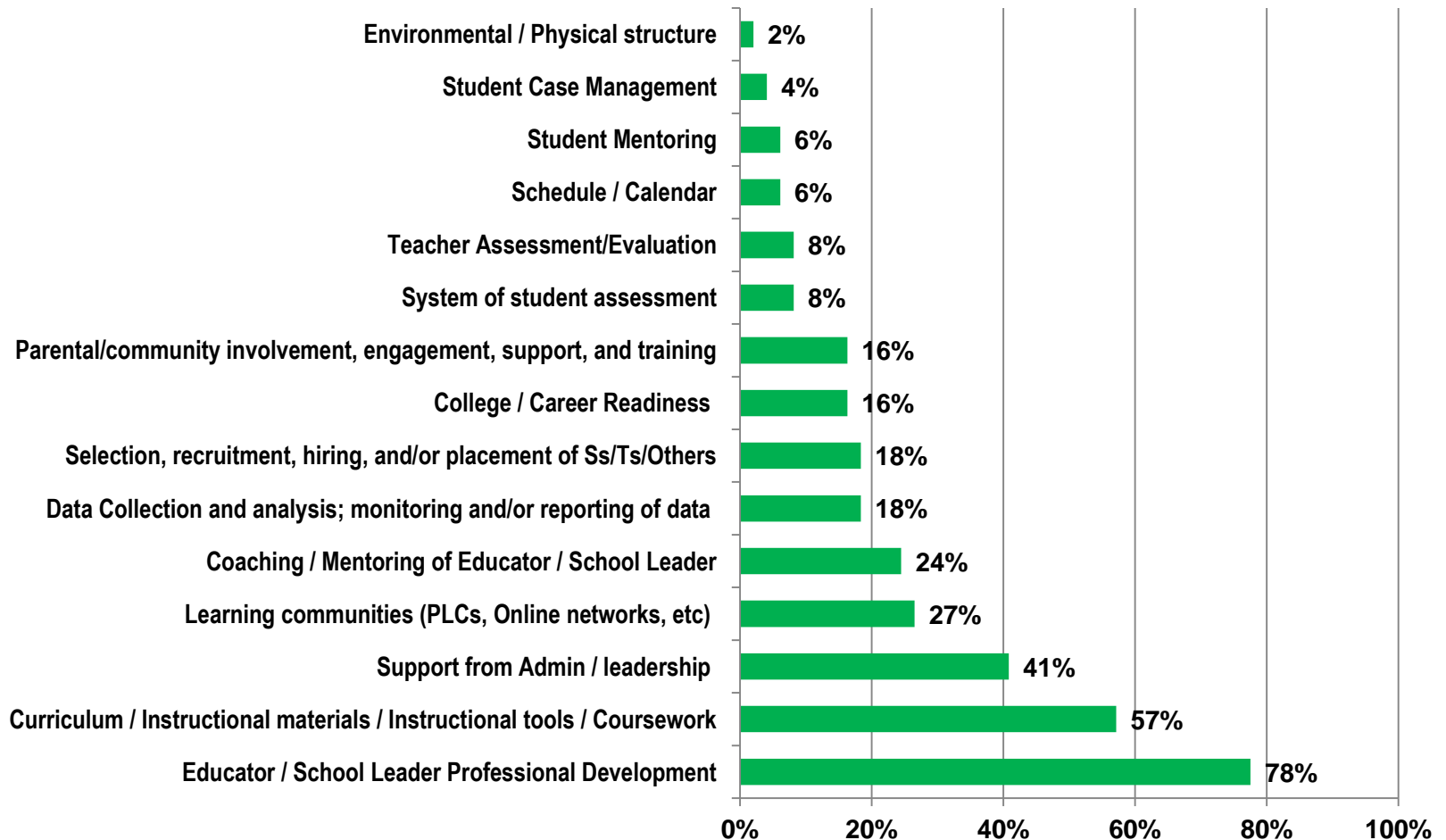
AP1: Effective Teachers and Principals	20
AP2: STEM	10
AP2: Use of Data	9
AP3: High Standards and High-Quality Assessments	21
AP3: Parent and Family	4
AP4: Persistently Low-Performing Schools	13
AP4: Turnaround	7
AP5: Rural Achievement	8

# Group by intended outcome?



\*nearly all grants are conducting multiple contrasts, some across multiple domains

# Group by approach to affecting change?



\*Many grants utilize multiple key components

# Revisiting Original Goal of Learning Community



- Trying to identify site-level contextual factors that help us understand variation in effects across sites
  - Site-level variation can be studied only in evaluations where each site has an effect (i.e., each site has treatment and comparison units)
- Grantees selected for this learning community based on having evaluation designs that would allow for studying site-level variation
- The field has not moved far in measuring common contextual factors