



# PPEM Overview

May 2018



# Contents

- Overview
- Calculation technicalities
- Data used
- Dashboard dive

[www.gapsc.com/Commission/Media/DocsPresentations.aspx](http://www.gapsc.com/Commission/Media/DocsPresentations.aspx)



# Overview

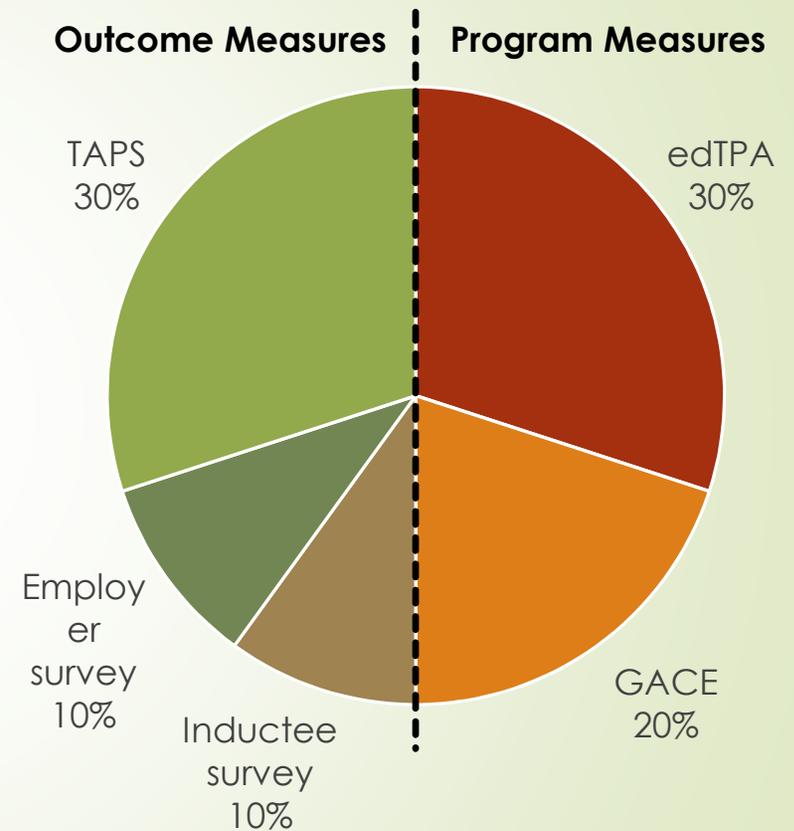


# What is a PPEM?

- ▶ Preparation Program Effectiveness Measures
  - ▶ TPPEM
  - ▶ LPPEM
- ▶ Purposes
  - ▶ Improvement
  - ▶ Transparency
  - ▶ Accountability

# TPPEM Components

- Contains both **in-program** and **outcome** measures
  - GACE** content assessment scores
  - edTPA** classroom performance assessment scores
  - TAPS** classroom observation scores from first teaching year
  - Surveys of **inductee** teachers and their **employers** from first teaching year

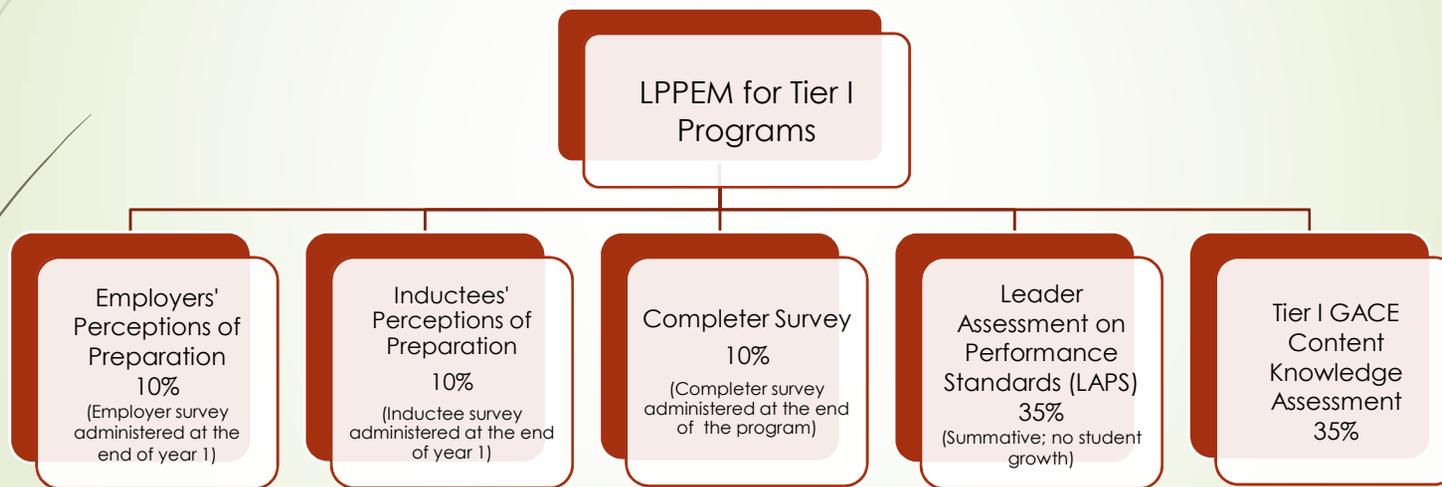




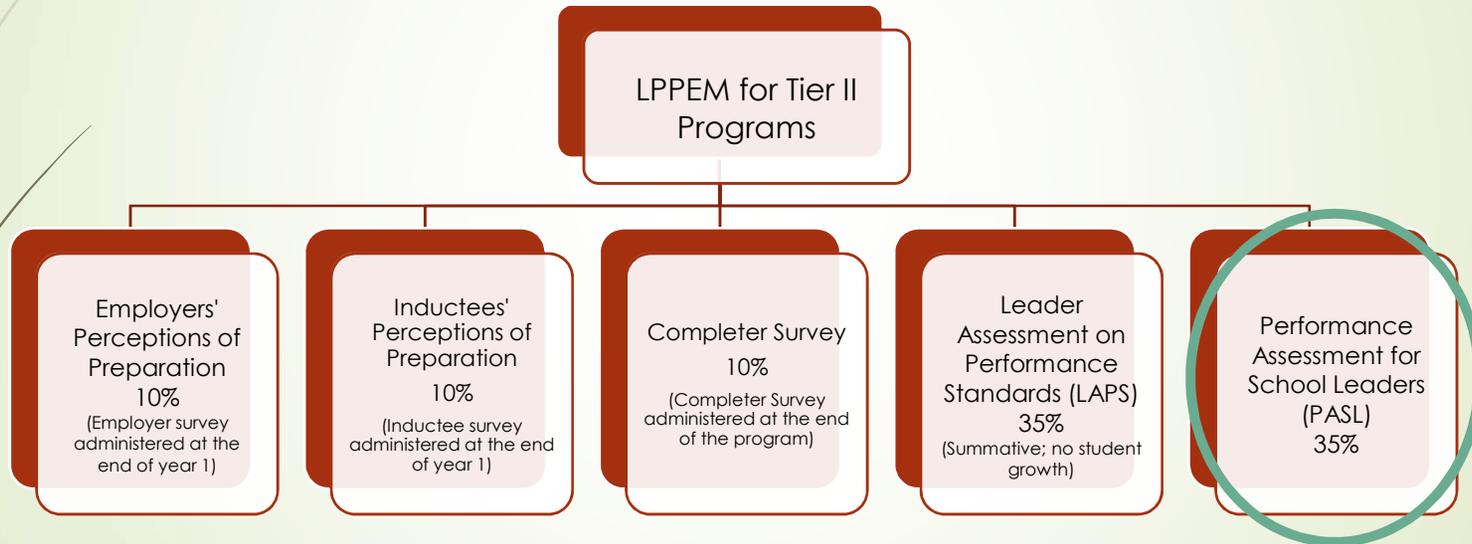
# Ratings

- ▶ Level IV: Exemplary
- ▶ Level III: Effective
- ▶ Level II: At Risk of Low Performing
- ▶ Level I: Low Performing

# LPPEM Tier I components



# LPPEM Tier II components





# Calculation technicalities



# Raw scores

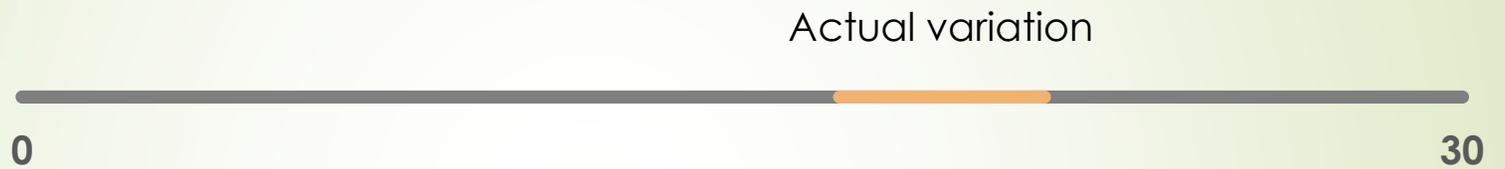
- ▶ Captures greater variability than pass rate
- ▶ Allows us to distinguish among providers, esp. at high end
  - ▶ Most pass rates cluster above 90% per provider, but raw scores vary significantly
- ▶ Allows us to avoid a ceiling effect



# Scaling using ranges

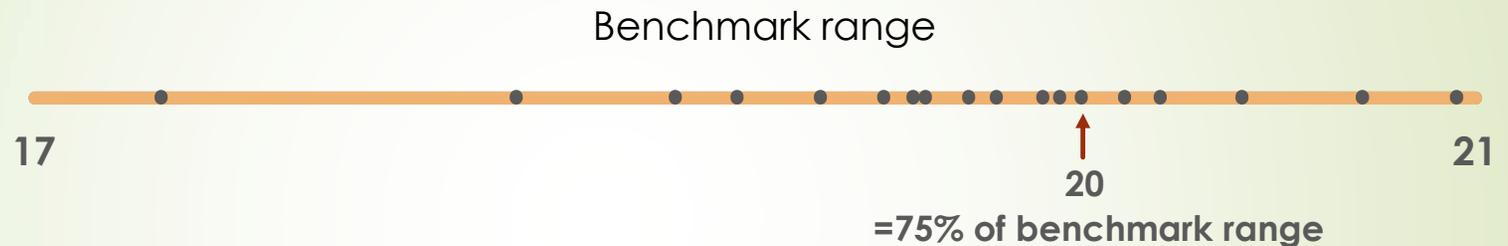
- ▶ When awarding points for each measure, we use the actual range of provider scores in a baseline year rather than the total range of hypothetical scores
- ▶ “Zooming in” on the range of actual scores allows us to better see distinctions
- ▶ Equalizes the amount of variation among measures
  - ▶ Effective weight of a measure reflects its assigned weight
  - ▶ Without scaling, measures with more variation would disproportionately influence the overall PPEM rating

# Scaling using ranges: TAPS example



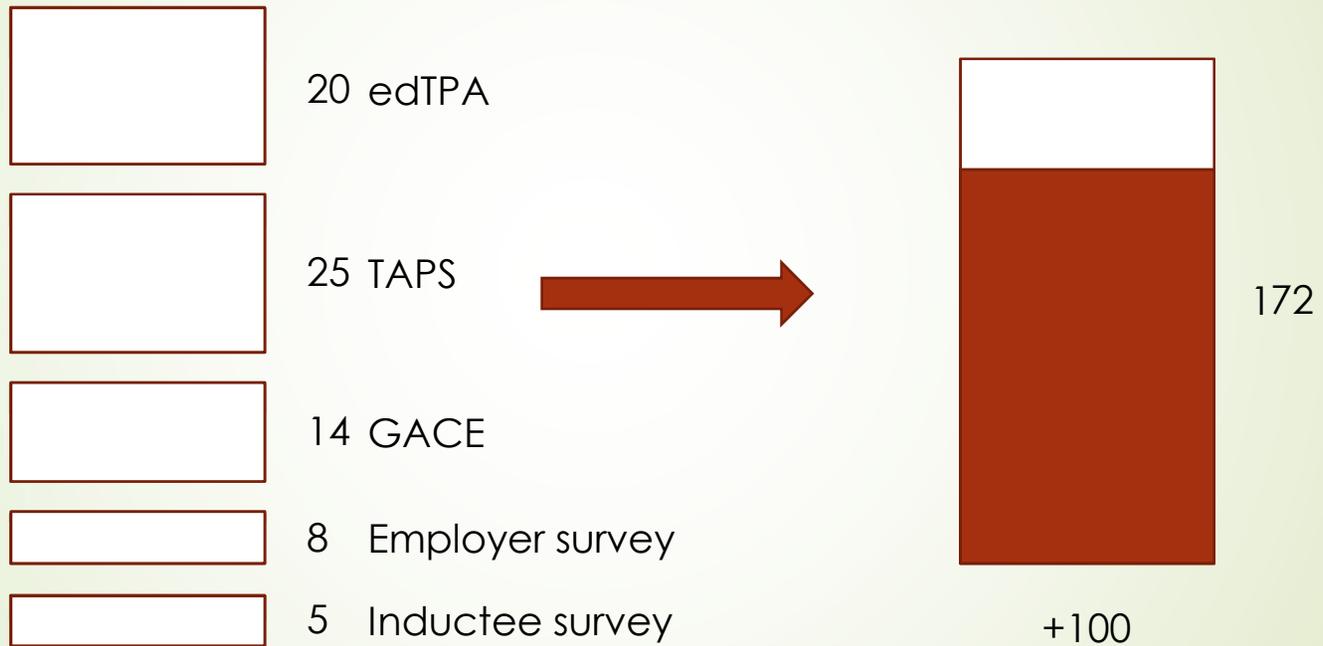
- ▶ About 94% of Georgia teachers receive a Level III (i.e. proficient) summative TAPS rating
- ▶ Most provider averages fall between 17 and 21

# Scaling using ranges: TAPS example



- Setting a benchmark range allows us to better distinguish between scores
- A provider's TAPS average is calculated as a proportion of the benchmark range
- This proportion is multiplied by the weight assigned to the measure, generating the earned PPEM points for that measure

# Overall index score





## Overall PPEM rating ranges

- ▶ Consensus of PPEM Advisory Group based on standard-setting activity

Rating	Index Range	Percent of Providers
Exemplary	180 to 200	25%
Effective	160 to <180	71%
At Risk of Low Performing	140 to <160	4%
Low Performing	<140	0%



# Measure indicators

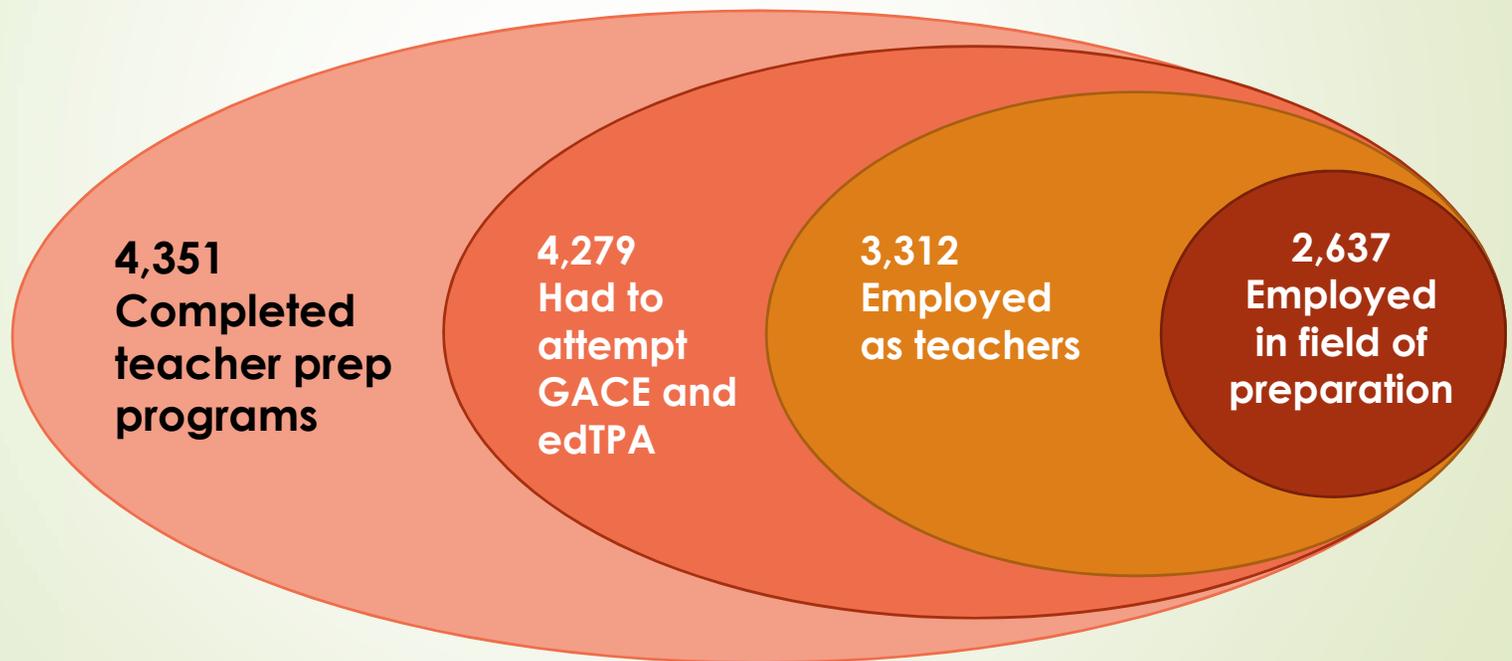
- ▶ Same levels as overall ratings: Levels 1-4 with 4 being highest
- ▶ Unlike overall rating, **not** consequential, nor used in the calculation of the overall rating
- ▶ Acts as an aid to providers in interpreting their score and identifying strengths and opportunities to improve
- ▶ Helps providers see where they are with reference to statewide scores in the baseline years
- ▶ Highlights score disparities
  - ▶ Ex. A provider's overall score and most measures are rated Effective, but TAPS is At-Risk



Data Used

# Data – Measurements Depend on Assessments and Employment

- Teacher preparation program completers in 2016, employed in 2017\*



\*Diagram not to scale

# The N Size Challenge

- ▶ Meeting N size **minimum** is important for data reliability sufficient to use for evaluation.
- ▶ We set the minimum N size at **10 completers' data** for each measure
- ▶ Modeling with 2016 TAPS showed only **56% of providers** met the minimum N
- ▶ Solution: Aggregation across **three years**
  - ▶ Allowed **89%** of providers to meet the minimum N (with limited Y1 data)
  - ▶ Also increases the statistical **reliability** of the PPEM

**56%** → **89%**



# Data Years Used

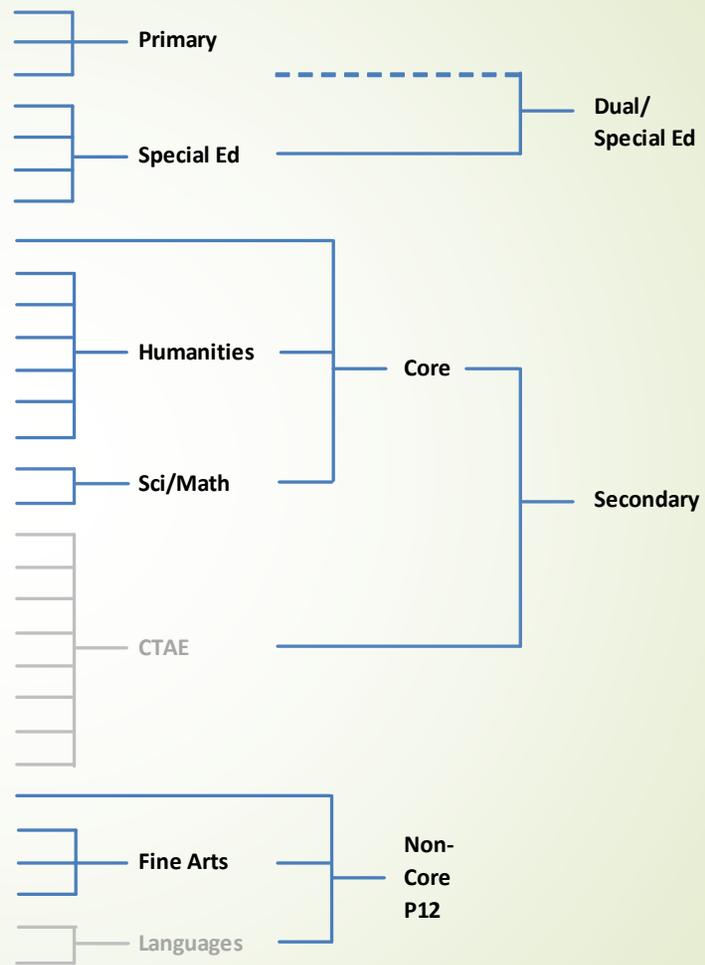
- ▶ All data: collected during the prior three reporting years (Sept-Aug)
  - ▶ Calculated in 2017-18, collected in 2016-17, 2015-16, and 2014-15
- ▶ **GACE and edTPA:** collected in program
  - ▶ Window starts when (traditional) EPP grants eligibility, ends on Aug 31 of completion year
  - ▶ Completion years 2016-17, 2015-16, and 2014-15
- ▶ **TAPS, inductee survey, employer survey:** collected during employment
  - ▶ First available result during employment within the first three years following completion
  - ▶ Completion years 2015-16, 2014-15, and 2013-14



# Grouping like programs

- ▶ N size issue remained at program level
  - ▶ Lack of accountability/transparency if not reported
  - ▶ Program receives no data for improvement
- ▶ Even after 3 year aggregation, only 1/3 of programs met N size minimums to receive a PPEM rating
- ▶ Combining like programs allows us to rate more than half of remaining programs
  - ▶ Drawback – data received is from several programs; combining in stages from most-like to broader groupings helps ameliorate this
  - ▶ Alternative is little/incomplete/no data

Birth through Kindergarten Education  
 Early Childhood Education  
 Special Education General Curriculum/ECE  
 Special Education Adapted Curriculum  
 Special Education Deaf Education  
 Special Education General Curriculum  
 Special Education Physical And Health Disabilities  
 Middle Grades Education  
 English Education  
 Economics Education  
 Geography Education  
 History Education  
 Political Science Education  
 Social Science Education  
 Mathematics Education  
 Science Education  
 Agriculture Education  
 Business Education  
 Family And Consumer Science Education  
 Health Occupations  
 Healthcare Science And Technology Education  
 Marketing Education  
 Technology Education  
 Trade And Industrial Education  
 Health And Physical Education  
 Art Education  
 Music Education  
 Drama Education  
 English to Speakers Of Other Languages Education  
 Foreign Language Education



— has been used   
 — has not been used   
 - - - secondary aggregation



# Dashboard dive

# Dashboard roll-out

**2018-19**

**Provider dashboard**

- non-consequential
- AY2018-19 calculation
- detail view
- [www.gapsc.org](http://www.gapsc.org)

**Early summer 2019**

**Provider dashboard**

- consequential
- AY2019-20 calculation
- detail view
- [www.gapsc.org](http://www.gapsc.org)

**Summer 2019**

**Public dashboard**

- consequential
- AY2019-20 calculation
- overall rating and  
measure indicators only
- [www.gapsc.com](http://www.gapsc.com)



Let's dive into the dashboard ...



For additional info:  
[www.gapsc.com](http://www.gapsc.com)  
ppems@gapsc.com

**#GaPSCDriveIn18**  
**@Ga\_PSC**

