Battle of the conjunctions: Disjunctive vs. compensatory course placement

Terrence Willett, Cabrillo College
Mallory Newell, De Anza College
Loris Fagioli, Irvine Valley College

November 17, 2016

http://www.rpgroup.org/projects/multiple-measures-assessment-project

KEEP
CALM
AND
RESEARCH
ON

Agenda

- Project Overview
- Research Basis
  - Impact Analysis and Relationship to Equity
- Comparing different ways to combine test and non-test data and information in placement Systems
- Integration with Common Assessment Initiative (CAI)

Project Overview

Collaboration

- CAI
- CCCCO
- Cal-PASS+
- RP Group
- 60 CCCs

Model Development

- English
- Math
- ESL
- Reading
- Non-cognitive Variables
- Self-reported transcript data

Engagement

- Local replication
- Webinars
- Professional development
- Support
- Pilot results inform statewide implementation

Growing body of evidence

- Weak relationship between assessment tests and college course outcomes: bit.ly/CCRCAssessment

- Incredible variability in cut scores; CCCs often use HIGHER cut scores than 4-year institutions: bit.ly/NAGB2012

- Underestimates students of color, women, first generation college students, low SES: bit.ly/DefiningPromise

- Long thread of research in the CCCs
  - Willett, Hayward, & Dahlstrom, 2008 http://bit.ly/WIllett2008
  - Hetts, Fuenmayor, & Rothstein, 2012 http://www.lbcc.edu/PromisePathways
  - Willett & Karanjeff, 2014 http://bit.ly/RPSTEPS

Why Multiple Measures?

- Tests used in isolation have been under-placing students
- Multiple measures
  - provides a more complete picture of student ability
  - provides a way to increase the accuracy of placement, particularly reducing underplacement http://bit.ly/CCRCPlacementAccuracy
  - are required by law (Title V)
  - supported by statewide senate

Methods

- Matched data from high schools and community colleges in CalPASS Plus
- Recursive decision trees with Poisson model
- Rules and R code:
  http://rpgroup.org/projects/multiple-measures-assessment-project/decision-rules

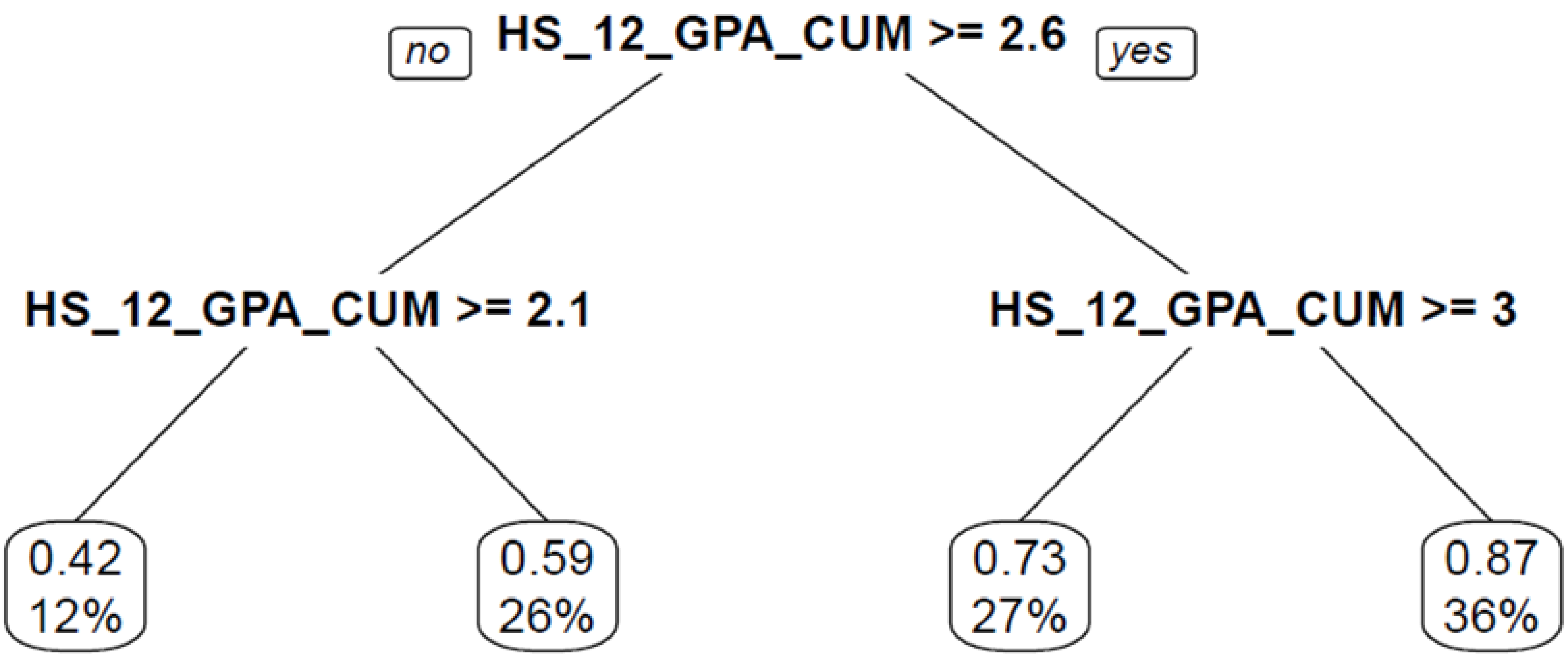
```
#load a data file into R
setwd("C:/Users/MyName/Documents/ProjectFolder")

MyData <- read.delim(“DataFile.txt", quote = "", row.names = NULL,
stringsAsFactors = FALSE)
```

Variables Explored in the Models

- High School Cumulative GPA (primary predictor)
- Grades in high school courses
- CST scores
- Advanced Placement course taking
- Taking higher level courses (math)
- Delay between HS and CCC (math)
- Type of English or Math course

Transfer Level English



Transfer Level MMAP Rule Sets

Transfer Level Course	Direct Matriculant	Non-Direct Matriculant
College Algebra (STEM) *(and high school Alg II recommended)*	HS 11 GPA >=3.2 OR HS 11 GPA >=2.9 AND Pre-Calculus C (or better)	HS 12 GPA >=3.2 OR HS 12 GPA >=3.0 AND Pre-Calculus or Statistics (C or better)
Statistics (Non STEM) *(and high school Alg I recommended)*	HS 11 GPA >=3.0 OR HS 11 GPA >=2.3 AND Pre-Calculus C (or better)	HS 12 GPA >=3.0 OR HS 12 GPA >=2.6 AND Pre-Calculus C (or better)
English	HS 11 GPA >=2.6	HS 12 GPA >=2.6
Reading	HS 11 GPA >=2.7	HS 12 GPA >=2.8
ESL	HS 11 GPA >=2.7	HS 12 GPA >=2.6

Various Placement Systems and Their Impact on Student Equity

What are some possible placement systems?

Disjunctive placement:

Take the highest placement of multiple measures

i.e. Test or High School (HS) Transcripts or AP score or EAP or…

Recommended by MMAP

Compensatory placement:

Combination of all multiple measures with equal or varying weights

i.e. Placement = Test + HS GPA + HS Course + AP score + …

Conjunctive placement:

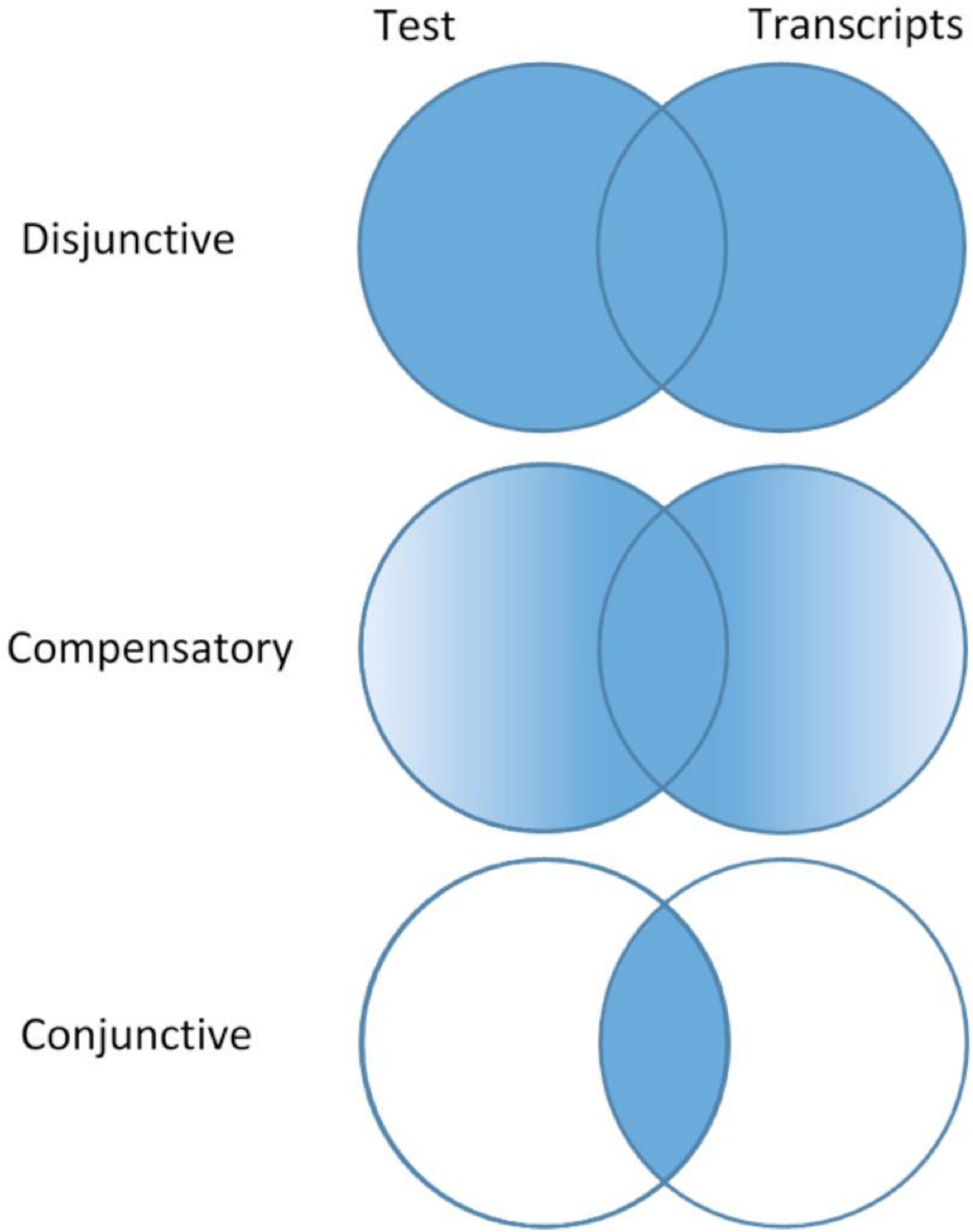
Lowest placement where all measures agree

i.e. exceed both Test score threshold and HS GPA criteria

Highly restrictive

Not recommended by the CCCCO

Special thanks to Dr. Barry Gribbons of College of the Canyons for first highlighting these systems to the MMAP Team


Test
Transcripts
Disjunctive
Compensatory
Conjunctive

How can we compare these systems?

- **Accuracy**: The proportion of students who are correctly predicted to be successful or to be unsuccessful.
- **Other Classification Metrics**: Positive predictive value, Sensitivity, Specificity, etc.
- **1 year throughput rate:** The number of students successfully completing the gatekeeper course at the end of a course sequence divided by the number of students in the initial cohort within 1 year.
- **Underrepresented Minority Placement Rate:** Equity and disproportionate impact are major considerations when evaluating the performance of placement systems

Classification metrics

- **Accuracy:** proportion of students correctly predicted to be successful or to be unsuccessful = (TP+TN)/(TP+FP+TN+FN)
- **PPV:** Positive predictive value, the number of passing students (i.e., true positives) divided by the number of students predicted to succeed = TP/(TP+FP)
- **NPV:** Negative predictive value = TN/(TN+FN)
- **Specificity:** TN/(TN+FP) = 1 - Type I error = True Positive Rate
- **Sensitivity:** TP/(TP+FN) = 1 - Type II error = Power = 1 - False Positive Rate

TP=True Positive, FP=False Positive

TN=True Negative, FN=False Negative

Information that can be used to evaluate placement systems. Two way continency table or "confusion matrix".

	Predicted to Fail	Predicted to Pass
Actually Failed	**True Negative**	*False Positive*
Actually Passed	*False Negative*	**True Positive**

Cal-PASS Plus
A System of Data

the RP group
the Research & Planning Group for California Community Colleges

Information that can be used to evaluate placement systems

Accuracy = (TP+TN)/(TP+FP+TN+FN)

	Predicted to Fail	Predicted to Pass
Actually Failed	**True Negative**	*False Positive*
Actually Passegd	*False Negative*	**True Positive**

Accuracy

Information that can be used to evaluate placement systems

PPV = TP/(TP+FP)

	Predicted to Fail	Predicted to Pass
Actually Failed	**True Negative**	*False Positive*
Actually Passed	*False Negative*	**True Positive**

PPV

Cal-PASS Plus
A System of Data

the RP group
the Research & Planning Group for California Community Colleges

Information that can be used to evaluate placement systems

NPV = TN/(TN+FN)

NPV	Predicted to Fail	Predicted to Pass
Actually Failed	**True Negative**	*False Positive*
Actually Passed	*False Negative*	**True Positive**

Information that can be used to evaluate placement systems

Specificity = TN/(TN+FP)

Specificity	Predicted to Fail	Predicte to Pass
Actually Failed	**True Negative**	*False Positive*
Actually Passed	*False Negative*	**True Positive**

Cal-PASSPlus
A System of Data

theRPgroup
the Research & Planning Group for California Community Colleges

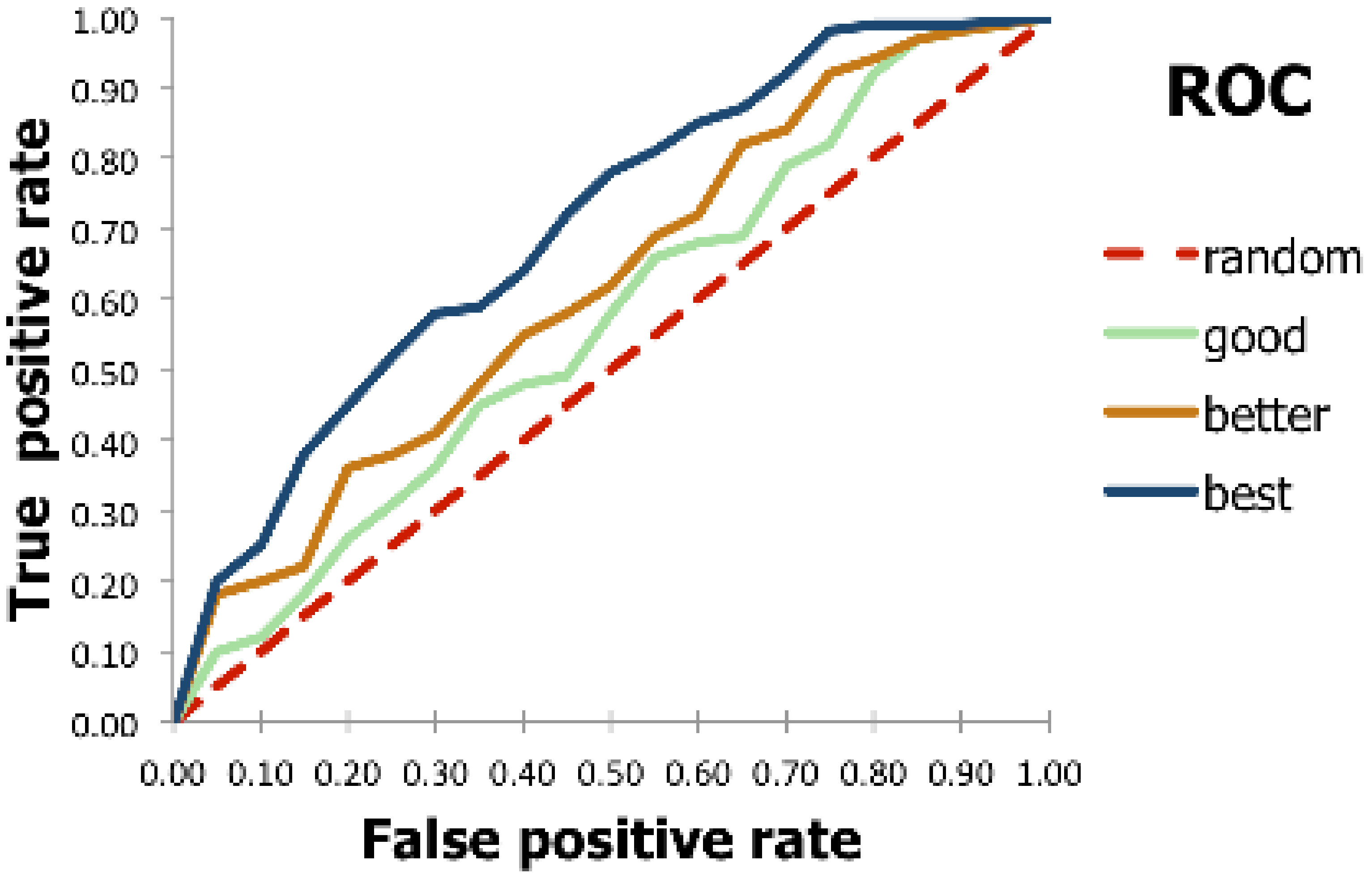
Information that can be used to evaluate placement systems

Sensitivity= TP/(TP+FN)

	Predicted to Fail	Predicted to Pass
Actually Failed	**True Negative**	*False Positive*
Actually Passed	*False Negative*	**True Positive**

Sensitivity

Response Operating Curve



R Resources for Model Metrics

- “caret” package for testing many models
- "e1071" corrects some errors in caret
- "pROC" draws ROC curves
- Max Kuhn caret site and webinar slides
  - http://topepo.github.io/caret/index.html
  - http://bit.ly/2eIz7Rz
  - http://bit.ly/2f2W78V

Managing Errors

- Typically have trade off in specificity v. sensitivity and must consider consequences of false positives v. false negatives.
- High Sensitivity
  - Airport security
  - Allow students a chance to pass a course
- High Specificity
  - Convicting someone of a serious crime
  - Protect students from failing a course

Types of Placement Error

- **Overplacement**: Student is placed above their ability to succeed. Highly visible.
- **Underplacement**: Student could have been successful at a higher level than where placed. Tends to be invisible.
- Current placement systems tend to result in much greater underplacement error.
- Total placement error is minimized when over- and underplacement are balanced.
- Consequences to students of each error not equal

Evaluating Placement Systems

Disjunctive placement:

Take the highest placement (Test or MMAP)

Recommended by MMAP

Compensatory placement:

Logistic regression (combines Test, MMAP simultaneously)

Run with two cut-values: 0.70, 0.50

Conjunctive placement:

Only if Test and MMAP in agreement

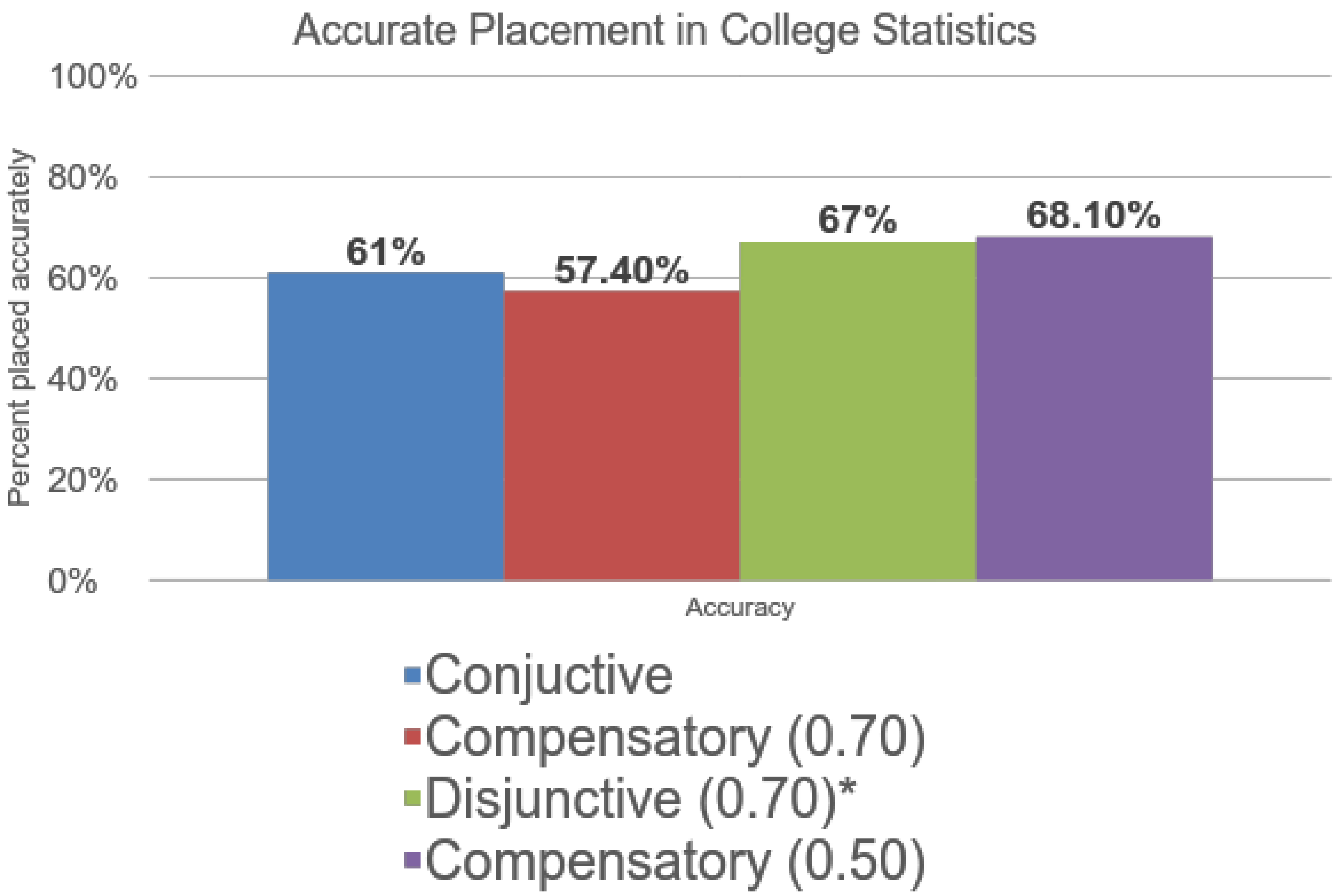
Highly restrictive

Not recommended by the CCCCO

Cal-PASS Plus
A System of Data

the RP group
the Research & Planning Group for California Community Colleges

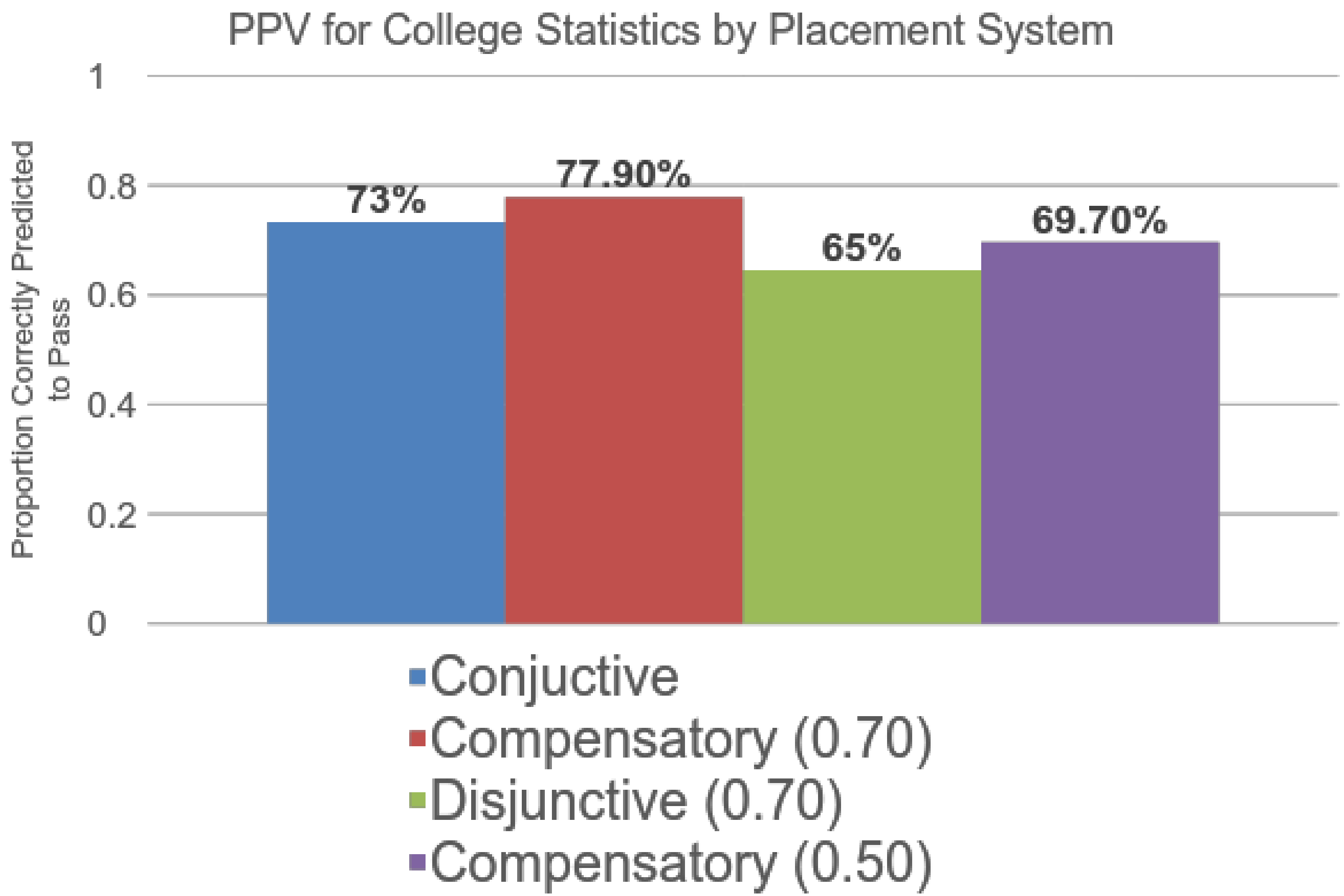
Accuracy: Statistics Course



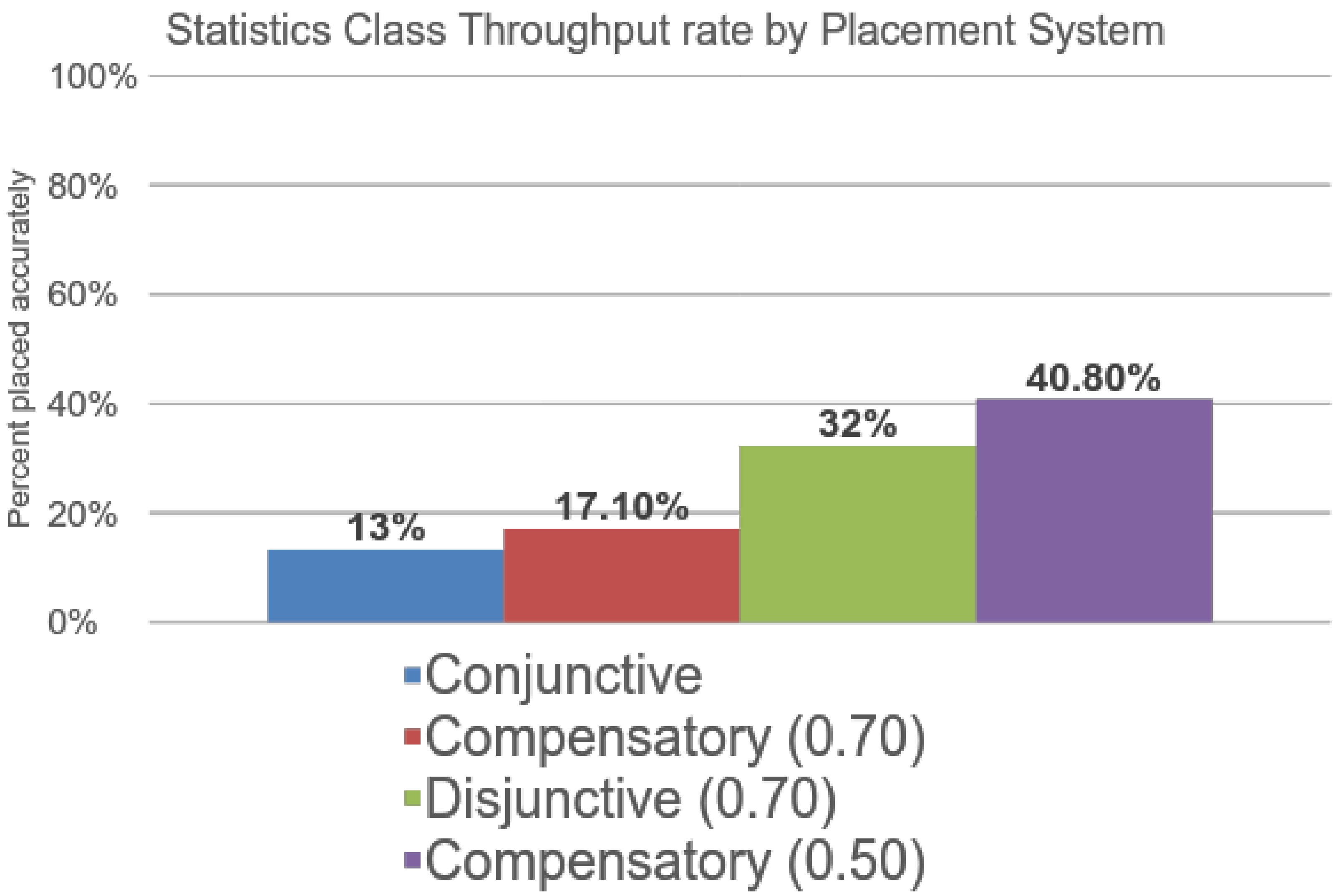
Cal-PASSPlus
A System of Data

theRPgroup
the Research & Planning Group
for California Community Colleges

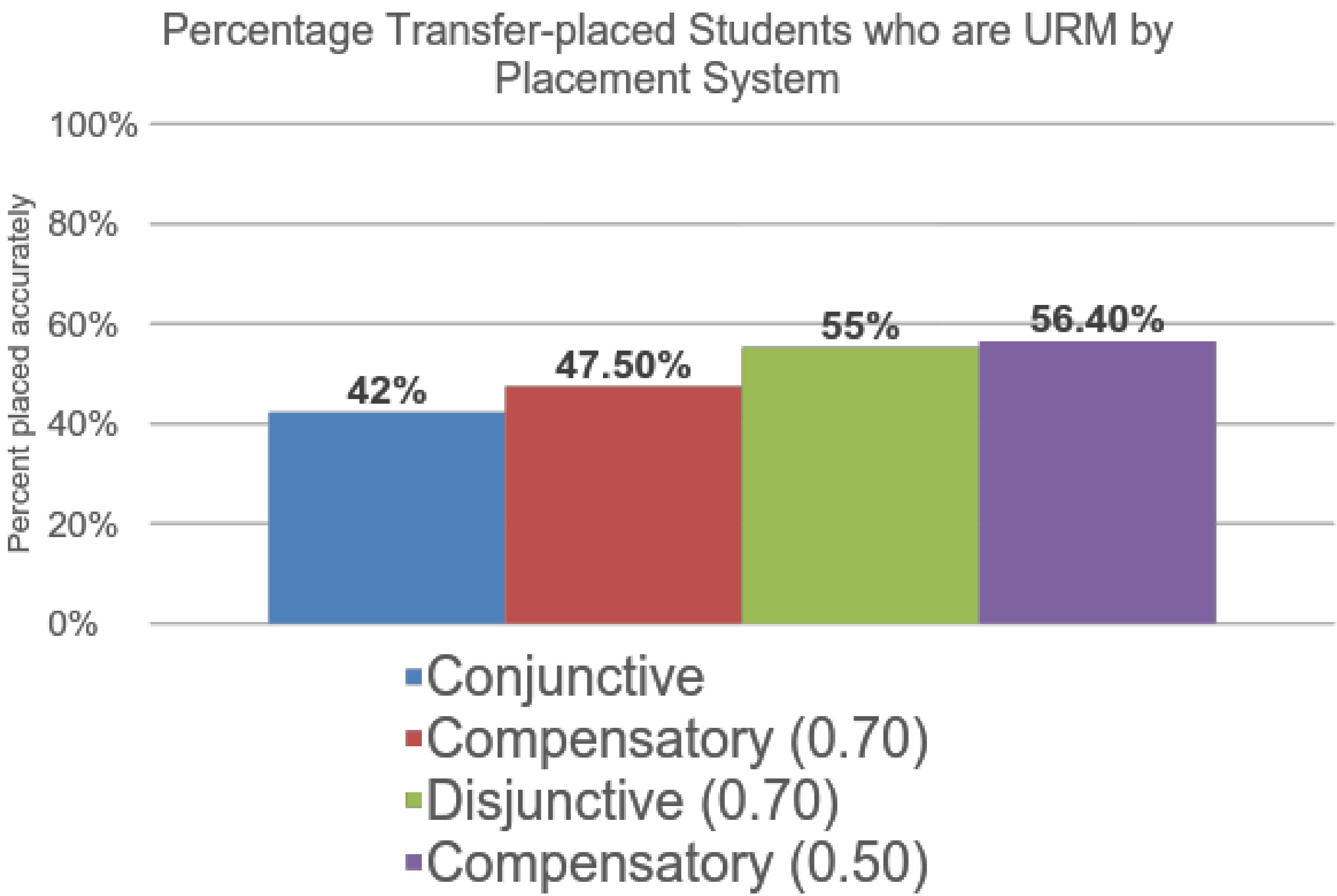
Positive Predicted Value (PPV): Statistics Course



One Year Throughput Rate: College Statistics Course

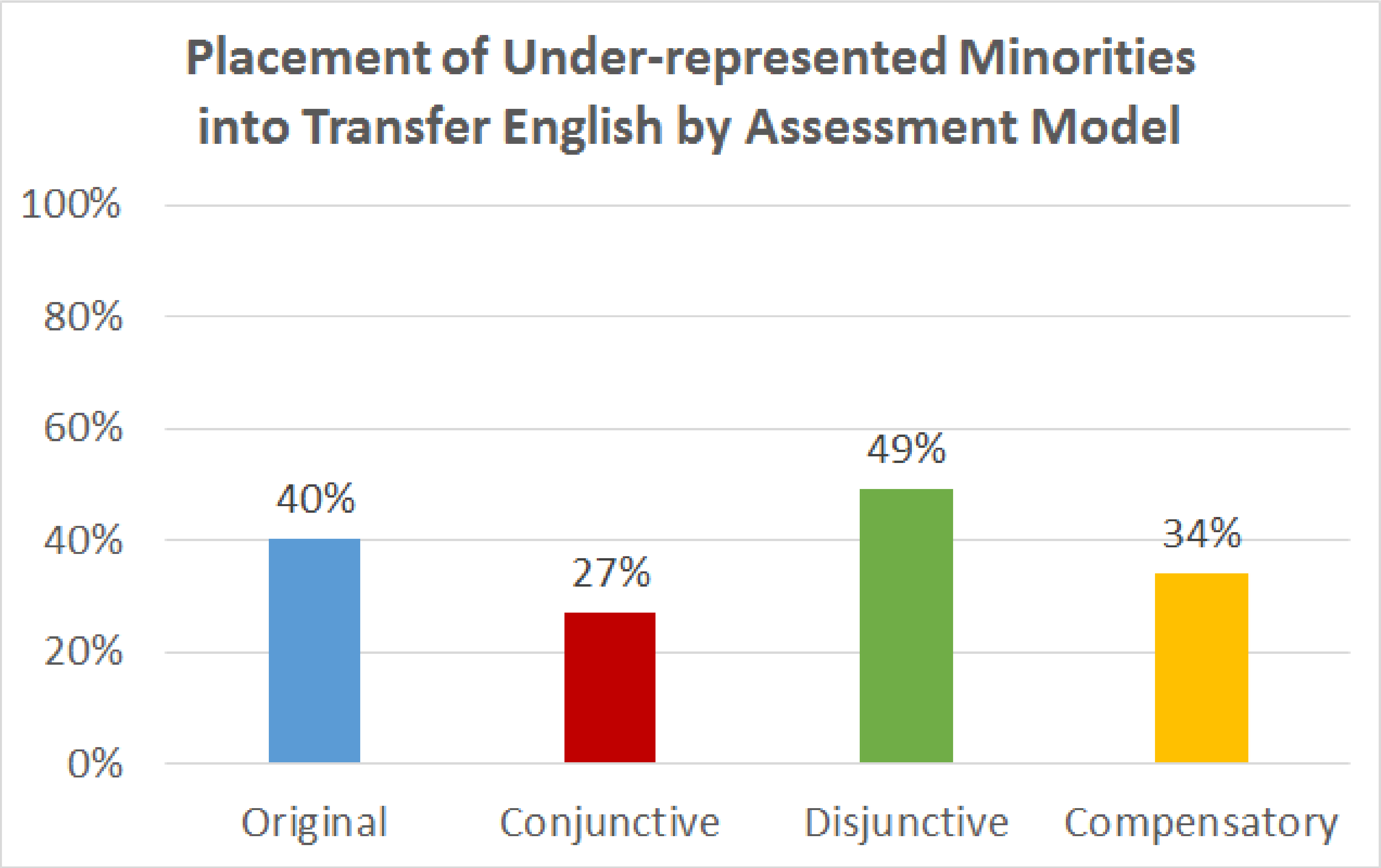


Percentage of Transfer-placed Students who are URM



Cal-PASS Plus
A System of Data

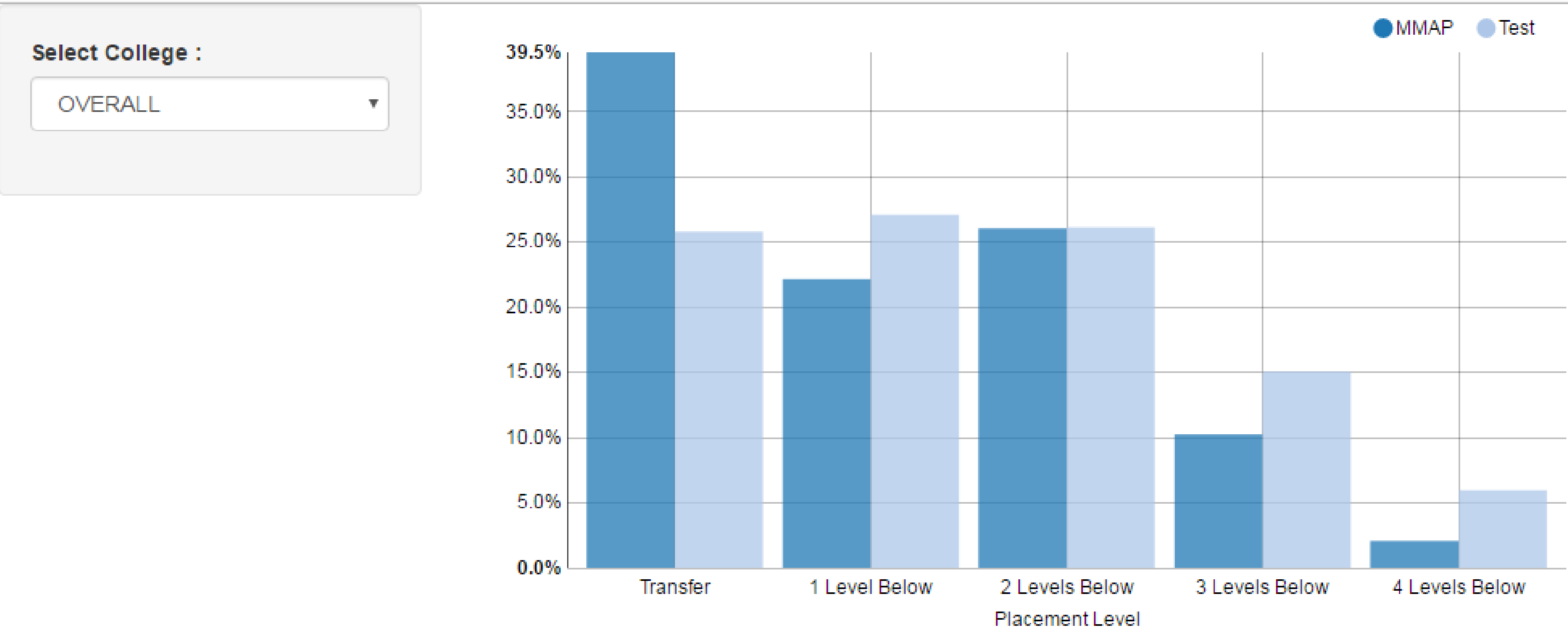
the RP group
the Research & Planning Group for California Community Colleges



Summary of Placement Models

- No single metric is sufficient but several well-chosen metrics can allow for a more informed decision
- Throughput is an important metric to consider
- PPV can be calculated for all placement systems; metrics that require a True Negative cannot be calculated for disjunctive placement systems.
- When requiring >70% probability of passing transfer-level course, disjunctive models have higher access and throughput than compensatory models
- If compensatory model is set to a 0.50 criterion or cut-value, it can outperform a disjunctive model (with a .70 criterion) in terms of accuracy, access, PPV and throughput
- The conjunctive model was very restrictive and had the lowest throughput rates and URM placement rates

Interactive rCharts



http://ramnathv.github.io/rCharts/

Integration of MMAP with CAI

Integration of MMAP with CAI

- Note: Common Assessment updates currently on pause
- Common Assessment platform will house a transcript data repository
  - repository will be source-agnostic & store transcript data from variety of sources, including CalPASS & self-report via CCC Apply
  - statewide decision trees programmed into platform, for internally generated Multiple Measures placement recommendation
  - expect data points used in MM placement recommendation
- Students will receive a single placement recommendation created from a disjunctive placement model

Contacts:

John Hetts
Educational Results Partnership
jhetts@edresults.org

Ken Sorey
Educational Results Partnership
ksorey@edresults.org

Dan Lamoree
Educational Results Partnership
dlamoree@edresults.org

Rachel Baker
UC Irvine
rachelbb@uci.edu

Mallory Newell
The RP Group
newellmallory@fhda.edu

Loris Fagioli
The RP Group
lfagioli@rpgroup.org

Terrence Willett
The RP Group
twillett@rpgroup.org

Craig Hayward
The RP Group
chayward@rpgroup.org