

Method for analysis:

The revised Teacher Work Sample was piloted by the following programs during the Spring 2022 semester:

- Early Childhood Education
- Elementary Education
- MAT English
- MAT Mathematics
- MAT Science
- MAT Social Studies
- Special Education

A total of 95 candidates across the seven programs completed the revised Teacher Work Sample.

An independent, two-tailed t-test was done for each comparison group (gender, first generation status, race/ethnicity). Confidence levels were set for 0.05; statistically significant differences are highlighted in red. Demographic data were pulled from candidate demographic surveys; data were excluded for candidates who selected “Prefer not to answer” or left their response for a category blank; therefore, the sample size is different across comparison groups. All qualifiers (gender, race, first generation status) were self-selected by candidates:

Table 1. Demographics as Identified by Participants

Program	Gender	Ethnicity	First Generation
Early Childhood Education	37 females 1 male	5 African American 1 Hispanic/Latino 1 Multiracial 31 White	12 First Generation 25 Not First Generation 1 Preferred not to answer
Elementary Education	23 Females 2 Male	4 African American 21 White	6 First Generation 19 Not First Generation
MAT English	1 Female 2 Did not answer	1 White 2 Did not answer	1 First Generation 2 Did not answer
MAT Math	1 Female 1 Male	2 White	1 First Generation 1 Not First Generation
MAT Science	1 Female	1 White	1 Not First Generation
MAT Social Studies	7 Female 8 Male	1 African American 17 White 1 Did not answer	5 First Generation 8 Not First Generation 1 Preferred not to answer 1 Did not answer
Special Education	10 Female 1 Male	11 White	2 First Generation 9 Not First Generation

For the comparison data for race/ethnicity, the term “Underrepresented” is used for students that self-identified as Black, Hispanic/Latino, or Multiracial. Because there were so few students that identified as Hispanic/Latino (n=1) and Multiracial (n=1), and ANOVA could not be done to analyze the differences amongst underrepresented subgroups, as a sample size of at least five candidates is required. Comparison data may be found in Table 2:

Table 2. Revised TWS comparisons by subgroup – Spring 2022

Revised Teacher Work Sample - Spring 2022									
TWS Indicator	Comparison by Gender			Comparison by Generational Status			Comparison by Race/Ethnicity		
	Mean Score for Females (n=80)	Mean Score for Males (n=13)	p-value	Mean Score for First Generation College Students (n=27)	Mean Score for Non-first Generational College Students (n=63)	p-value	Mean Score for Underrepresented Students (n=12)	Mean score for White Students (n=80)	p-value
CF 1 School Information	3.09	3.54	0.007	3.19	3.1	0.486	3.17	3.14	0.869
CF 2 Knowledge of Classroom Information	3.25	3.62	0.03	3.19	3.32	0.308	3.25	3.3	0.777
CF 3 Knowledge of Student Characteristics	3.18	3.62	0.0004	3.15	3.24	0.344	3.25	3.23	0.849
LGA 1 List 2 to 3 learning goals	3.14	3.62	0.00004	3.26	3.16	0.27	3.08	3.21	0.298
LGA 2 Levels of learning goals	3.1	3.38	0.006	3.22	3.11	0.174	3	3.16	0.135
LGA 3 Alignment of Learning Goals with standards	3.06	3.54	0.0005	3.11	3.13	0.884	3.17	3.11	0.709
LGA 4 Appropriateness of Learning Goals	2.98	3.08	0.662	2.89	3	0.534	3	2.98	0.917
LGA 5 Mastery levels for each Learning Goal	3.04	3.54	<0.00001	3.15	3.08	0.42	3.08	3.1	0.884
LGA 6 Pre-post Assessment Blueprint: Learning Goals	3.06	3.38	0.023	3.04	3.11	0.492	3.08	3.1	0.884
LGA 7 Pre-post Assessment Blueprint: Adaptations	2.99	3.15	0.46	3.19	2.92	0.126	2.92	3.03	0.643
LGA 8 Pre-post Assessment Blueprint: Modes of Assessments	3.19	3.23	0.79	3.33	3.15	0.09	3.08	3.23	0.337

LGA 9 Pre-post Assessment Blueprint: Scoring Criteria	3.03	3.31	0.107	3.07	3.03	0.752	3	3.06	0.75
DI 1 Results of pre-assessment	3.2	3.54	0.026	3.19	3.25	0.597	3.25	3.24	0.94
DI 2 Unit Overview	3.26	3.15	0.508	3.19	3.27	0.504	3.17	3.25	0.622
DI 3 Integration of Technology	3.2	3.54	0.026	3.33	3.19	0.228	3.08	3.27	0.243
DI 4 Instructional Strategies	3.2	3.46	0.179	3.11	3.25	0.338	3.08	3.25	0.409
DI 5 Formative Assessments	2.92	3.38	0.007	2.74	3.06	0.025	3.08	2.96	0.539
ASL 1 Visual Representation of Student Performance	3.2	3.69	0.0003	3.33	3.22	0.324	3.08	3.29	0.179
ASL 2 Analysis of Student Performance	3.13	3.31	0.135	3.15	3.13	0.867	3.08	3.15	0.696
ASL 3 Instructional Implications from Data	3.13	3.23	0.259	3.04	3.16	0.327	3.08	3.15	0.696
ASL 4 Analysis of an Individual Student	3.24	3.54	0.012	3.26	3.29	0.779	3.18	3.29	0.477
R 1 Self-assessment of SCTS 4.0	3.28	3.77	0.003	3.22	3.38	0.267	2.83	3.41	0.002
R 2 Identify Teaching Strengths	3.25	3.69	0.003	3.26	3.32	0.648	3	3.35	0.039
R 3 Identify areas of Professional Development	3.09	3.62	0.01	3.19	3.13	0.726	2.91	3.19	0.253

Overall, there was a statistically significant difference between male and female performance on 16 of the 24 TWS indicators, with males outperforming females on all 16 indicators. However, these differences should be considered with caution, given the small sample size of males in the EPP. In addition, there are more males in the secondary MAT programs than the undergraduate licensure programs; therefore, an ANOVA was completed to determine the effect of program on candidate scores. The sample sizes for each program were as follows: 38 Early Childhood Education candidates, 25 Elementary Education candidates, 21 MAT Candidates, and 11 Special Education candidates. Because the MAT program had one science candidate, two mathematics candidates, and three English candidates, and 15 social studies candidates, all MAT candidates were grouped together to have a large enough sample size for the ANOVA. The ANOVA was run on the first two sections of the TWS (Contextual Factors and Learning Goals/Assessment), to look for patterns.

Based on the ANOVA, candidates in the MAT programs consistently scored significantly higher than their counterparts in all programs:

Table 3. Comparing TWS Scores by Program Using ANOVA/Post Hoc Tukey HSD – Contextual Factors and Learning Goals

Contextual Factors Indicator 1	Comparison Groups	Mean Scores	p-value
	Early Childhood Elementary	3.11 2.76	0.113
	Early Childhood MAT	3.11 3.71	0.001
	Early Childhood Special Education	3.11 3.27	0.691
	Elementary MAT	2.76 3.71	<0.00001
	Elementary Special Education	2.76 3.27	0.006
	MAT Special Education	3.71 3.27	0.024
Contextual Factors Indicator 2	Comparison Groups	Mean Scores	p-value
	Early Childhood Elementary	3.26 3.04	0.487
	Early Childhood MAT	3.26 3.86	0.002
	Early Childhood Special Education	3.26 3.09	0.691
	Elementary MAT	3.04 3.86	<0.00001
	Elementary Special Education	3.04 3.09	0.988
	MAT Special Education	3.86 3.09	0.00003
Contextual Factors Indicator 3	Comparison Groups	Mean Scores	p-value
	Early Childhood Elementary	3.11 3.08	0.993
	Early Childhood MAT	3.11 3.86	<0.00001
	Early Childhood Special Education	3.11 3.00	0.684
	Elementary MAT	3.08 3.86	<0.00001
	Elementary Special Education	3.08 3.00	0.833
	MAT Special Education	3.86 3.00	<0.00001

Design for Instruction Indicator 1	Comparison Groups	Mean Scores	p-value
	Early Childhood	3.03	0.942
	Elementary	3.00	
	Early Childhood	3.03	<0.00001
	MAT	3.95	
	Early Childhood	3.03	0.942
	Special Education	3.00	
Design for Instruction Indicator 2	Comparison Groups	Mean Scores	p-value
	Early Childhood	3.00	-
	Elementary	3.00	
	Early Childhood	3.00	<0.00001
	MAT	3.71	
	Early Childhood	3.00	-
	Special Education	3.00	
Design for Instruction Indicator 3	Comparison Groups	Mean Scores	p-value
	Early Childhood	3.05	0.480
	Elementary	2.88	
	Early Childhood	3.05	<0.00001
	MAT	3.71	
	Early Childhood	3.05	0.972
	Special Education	3.00	
Design for Instruction Indicator 4	Comparison Groups	Mean Scores	p-value
	Early Childhood	3.26	0.0005
	Elementary	2.36	
	Early Childhood	3.24	0.807
	MAT	3.43	
	Early Childhood	3.24	0.423
	Special Education	2.91	
Design for Instruction Indicator 5	Comparison Groups	Mean Scores	p-value
	Elementary	2.36	0.00002
	MAT	3.43	
	Elementary	2.36	0.057
	Special Education	2.91	
	MAT	3.43	0.79
	Special Education	2.91	

	Early Childhood MAT	3.00 3.60	<i><0.00001</i>
	Early Childhood Special Education	3.00 2.82	0.200
	Elementary MAT	3.04 3.60	<i><0.00001</i>
	Elementary Special Education	3.04 2.82	0.082
	MAT Special Education	3.60 2.82	<i><0.00001</i>
Design for Instruction Indicator 6	Comparison Groups	Mean Scores	p-value
	Early Childhood Elementary	3.03 2.80	0.192
	Early Childhood MAT	3.03 3.71	<i><0.00001</i>
	Early Childhood Special Education	3.03 3.00	0.995
	Elementary MAT	2.80 3.71	<i><0.00001</i>
	Elementary Special Education	2.80 3.00	0.291
	MAT Special Education	3.71 3.00	<i><0.00001</i>
Design for Instruction Indicator 7	Comparison Groups	Mean Scores	p-value
	Early Childhood Elementary	3.26 2.24	<i><0.00001</i>
	Early Childhood MAT	3.26 3.43	0.813
	Early Childhood Special Education	3.26 3.00	0.498
	Elementary MAT	2.24 3.43	<i><0.00001</i>
	Elementary Special Education	2.24 3.00	0.0006
	MAT Special Education	3.43 3.00	0.107
Design for Instruction Indicator 8	Comparison Groups	Mean Scores	p-value
	Early Childhood Elementary	3.24 2.96	0.303
	Early Childhood MAT	3.24 3.43	0.620
	Early Childhood Special Education	3.24 3.00	0.442
	Elementary MAT	2.96 3.43	0.020
	Elementary Special Education	2.96 3.00	0.994
	MAT Special Education	3.43 3.00	0.039
Design for Instruction Indicator 9	Comparison Groups	Mean Scores	p-value
	Early Childhood Elementary	3.11 2.56	0.0003
	Early Childhood MAT	3.11 3.81	<i><0.00001</i>
	Early Childhood	3.11	0.121

	Special Education	2.82	
	Elementary	2.56	<i><0.00001</i>
	MAT	3.81	
	Elementary	2.56	<i>0.191</i>
	Special Education	2.82	
	MAT	3.81	<i><0.00001</i>
	Special Education	2.82	

Across the data sources analyzed with the ANOVA, it became apparent that the MAT candidates are regularly scoring significantly higher than their colleagues in other programs. Ideally, the EPP would be able to use a regression model to determine the effect of program and gender on candidate scores; unfortunately, having so few candidates in some of the MAT programs (English, Math, and Science), as well as so few males in other programs (Early Childhood, Elementary Education, Special Education) made it impossible to do a regression. Instead, the EPP opted to then conduct an independent, two-tailed t-test comparing male and female scores in the MAT program, to determine if there were any differences in candidate performance that would be attributed to gender versus program placement:

Table 4. Revised TWS comparisons by gender, MAT Candidates only – Spring 2022

Indicator	Mean Score for MAT Females (n=10)	Mean Score for MAT Males (n=9)	p-value
CF 1 School Information	3.6	3.78	<i>0.434</i>
CF 2 Knowledge of Classroom Information	2.92	3.22	<i>0.273</i>
CF 3 Knowledge of Student Characteristics	3.03	3.11	<i>0.766</i>
LGA 1 List 2 to 3 learning goals	2.91	3.11	<i>0.174</i>
LGA 2 Levels of learning goals	3.8	3.56	<i>0.277</i>
LGA 3 Alignment of Learning Goals with standards	3.5	3.89	<i>0.216</i>
LGA 4 Appropriateness of Learning Goals	3.4	3.33	<i>0.884</i>
LGA 5 Mastery levels for each Learning Goal	3.5	3.78	<i>0.233</i>
LGA 6 Pre-post Assessment Blueprint: Learning Goals	3.8	3.67	<i>0.620</i>
LGA 7 Pre-post Assessment Blueprint: Adaptations	3.7	3.33	<i>0.200</i>
LGA 8 Pre-post Assessment Blueprint: Modes of Assessments	3.7	3.33	<i>0.312</i>
LGA 9 Pre-post Assessment Blueprint: Scoring Criteria	3.9	3.67	<i>0.236</i>
DI 1 Results of pre- assessment	4.0	3.89	<i>0.305</i>
DI 2 Unit Overview	3.6	3.44	<i>0.595</i>
DI 3 Integration of Technology	3.8	3.78	<i>0.912</i>
DI 4 Instructional Strategies	3.8	3.89	<i>0.711</i>
DI 5 Formative Assessments	2.9	3.89	<i>0.029</i>
ASL 1 Visual Representation of Student Performance	3.9	4.0	<i>0.357</i>

ASL 2 Analysis of Student Performance	3.9	3.56	0.098
ASL 3 Instructional Implications from Data	3.7	3.56	0.541
ASL 4 Analysis of an Individual Student	3.8	3.78	0.912
R 1 Self-assessment of SCTS 4.0	3.8	4.0	0.357
R 2 Identify Teaching Strengths	4.0	4.0	-
R 3 Identify areas of Professional Development	4.0	4.0	-

Overall, it appears that program, rather than gender, impacts the differences in scores; when comparing male and female scores for MAT candidates on the TWS, there was only one data point with a statistically significant difference in mean score (Design for Instruction Indicator 5: Formative Assessments).

Between first-generation and non-first-generation students there was only one statistically significant difference in scoring, with non-first-generation students earning higher mean scores on their inclusion of formative assessments and scoring guidelines for the formative assessments ($p=0.025$). Finally, for unrepresented students and White students, there were two indicators with statistically significant differences in scoring (Reflection Indicator 1, Self-Assessment Using the SCTS 4.0 and Reflection Indicator 2, Identify Teaching Strengths).