Technical Appendices: Table of Contents

Appendix A: School District Dashboards

See separate file of school district dashboards.

Appendix B:

2015 DPAS-II Evaluation Survey: Teachers and Specialists

The DPAS-II Evaluation Survey was conducted by Research for Action for the Delaware Department of Education. Surveys for Teachers and Specialists were conducted online from May 5 to June 5, 2015. The survey was designed to probe teachers' and specialists' views of the DPAS-II Evaluation System.

The following tables include whole counts for each question and response item, as well as the corresponding percentage. Each question captures the total responders for that specific question. The responses include both completed and partial respondents.[1]

	Completed Interviews	**Partial Interviews**	**Total Responses**
Teachers	4,080	1,291	5,371
Specialists	659	266	925

Screener Questions

S1. What is your title?

(Asked Teachers)
S2. What subject(s) do you teach?

(Asked Teachers)
S3. What grade do you teach?

(Asked Specialists)
S4. Which of the following Specialist categories does your position fall in?

S5. Given your perspective as an educator, what do you think is the most important factor influencing student academic achievement?

[Open-Ended Responses Given]

[1] Partial responses consist of any respondent who began the survey and did not complete the last question. This includes individuals who completed most of the survey but not the last question, and those who answered only a few questions.

Section I: General DPAS-II Perceptions

Q1. DPAS-II is fair and equitable.

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	35	0.8	5	0.7
Agree	756	17.0	95	12.7
Neither agree nor disagree	1,349	30.3	237	31.6
Disagree	1,353	30.4	230	30.7
Strongly disagree	817	18.4	138	18.4
Don't know	139	3.1	44	5.9
Total	**4,449**	**100**	**749**	**100**

Q2. DPAS-II is one of the top three drivers of student achievement gains in your school or work location.

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	134	3.0	16	2.1
Agree	653	14.6	54	7.2
Neither agree nor disagree	1,010	22.6	172	23.0
Disagree	1,267	28.4	219	29.2
Strongly disagree	1,191	26.7	205	27.4
Don't know	211	4.7	83	11.1
Total	**4,466**	**100**	**749**	**100**

Q3. Do you understand how you are evaluated as an educator under DPAS-II?

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very much	2,120	47.4	293	39.2
Somewhat	1,953	43.7	340	45.5
Slightly	321	7.2	74	9.9
Not at all	77	1.7	41	5.5
Total	**4,471**	**100**	**748**	**100**

(Asked Teachers)
Q4. How much do you think DPAS-II improves instructional practice?

(Asked Specialists)
Q4. How much do you think DPAS-II improves practice?

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very much	109	2.4	5	0.7
Somewhat	1,204	27.0	158	21.2
Slightly	1,437	32.2	216	29.0
Not at all	1,594	35.7	322	43.2
Don't know	118	2.6	44	5.9
Total	**4,462**	**100**	**745**	**100**

Q5. How much do you think DPAS-II informs professional development?

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very much	193	4.3	16	2.1
Somewhat	1,208	27.1	167	22.4
Slightly	1,318	29.6	208	27.8
Not at all	1,471	33.0	273	36.5
Don't know	264	5.9	83	11.1
Total	**4,454**	**100**	**747**	**100**

Q6. Does DPAS-II provide accurate ratings of educators? Please indicate for each type of rating.

a. Criterion-level ratings

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very accurate	138	3.2	89	12.6
Somewhat accurate	1,513	35.1	333	47.1
Neither accurate nor inaccurate	1,080	25.0	109	15.4
Somewhat inaccurate	527	12.2	53	7.5
Very inaccurate	455	10.5	56	7.9
Don't know	601	13.9	67	9.5
Total	**4,314**	**100**	**707**	**100**

b. Component-level ratings (I – IV)

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very accurate	218	5.2	31	4.6
Somewhat accurate	1,773	42.3	287	42.5
Neither accurate nor inaccurate	928	22.2	141	20.9
Somewhat inaccurate	532	12.7	64	9.5
Very inaccurate	409	9.8	61	9.0
Don't know	328	7.8	91	13.5
Total	**4,188**	**100**	**675**	**100**

c. Component V, Measure A (State Assessment Scores)

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very accurate	71	1.6	5	0.7
Somewhat accurate	734	16.9	106	15.0
Neither accurate nor inaccurate	824	18.9	142	20.0
Somewhat inaccurate	795	18.3	94	13.3
Very inaccurate	1,336	30.7	160	22.6
Don't know	590	13.6	202	28.5
Total	4,350	100	709	100

d. Component V, Measure B (Content Assessments)

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very accurate	107	2.5	5	0.7
Somewhat accurate	1,108	25.5	135	19.2
Neither accurate nor inaccurate	877	20.2	144	20.5
Somewhat inaccurate	813	18.7	90	12.8
Very inaccurate	1,058	24.4	110	15.6
Don't know	380	8.7	219	31.2
Total	**4,343**	**100**	**703**	**100**

e. Component V, Measure C (Growth Goals)

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very accurate	145	3.3	15	2.1
Somewhat accurate	1,115	25.6	175	24.8
Neither accurate nor inaccurate	904	20.8	168	23.8
Somewhat inaccurate	777	17.9	112	15.8
Very inaccurate	985	22.6	161	22.8
Don't know	423	9.7	76	10.7
Total	**4,349**	**100**	**707**	**100**

f. Summative Rating

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very accurate	158	3.7	36	5.2
Somewhat accurate	1,510	35.7	244	35.3
Neither accurate nor inaccurate	1,089	25.7	167	24.2
Somewhat inaccurate	643	15.2	72	10.4
Very inaccurate	506	12.0	75	10.9
Don't know	324	7.7	97	14.0
Total	**4,230**	**100**	**691**	**100**

Q7. Overall, what grade would you give DPAS-II?

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
A	40	0.9	2	0.3
B	668	15.1	79	10.9
C	1,605	36.4	249	34.4
D	1,387	31.4	244	33.7
F	713	16.2	150	20.7
Total	**4,413**	**100**	**724**	**100**

Section II: DPAS-II Implementation

Q8. Who was your primary evaluator this year?

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
The principal	2,124	43.7	376	45.2
The assistant or vice principal	2,542	52.3	330	39.7
District administrator	93	1.9	94	11.3
Other	102	2.1	32	3.8
Total	**4,861**	**100**	**832**	**100**

Q9. Please indicate how useful the following steps were during your evaluation this year.

a. Component V Conference

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very useful	759	15.9	111	13.9
Somewhat useful	1,505	31.6	233	29.1
Slightly useful	1,105	23.2	180	22.5
Not at all useful	1,077	22.6	181	22.6
Not applicable	317	6.7	95	11.9
Total	**4,763**	**100**	**800**	**100**

b. Professional Responsibilities Conference

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very useful	669	14.1	117	14.6
Somewhat useful	1,416	29.7	258	32.3
Slightly useful	1,097	23.0	167	20.9
Not at all useful	1,183	24.8	155	19.4
Not applicable	396	8.3	102	12.8
Total	**4,761**	**100**	**799**	**100**

c. Short observations

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very useful	971	20.3	110	13.7
Somewhat useful	1,391	29.1	227	28.2
Slightly useful	1,047	21.9	154	19.1
Not at all useful	775	16.2	110	13.7
Not applicable	591	12.4	204	25.3
Total	**4,775**	**100**	**805**	**100**

d. Full observations

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very useful	1,340	28.1	155	19.3
Somewhat useful	1,842	38.6	277	34.5
Slightly useful	1,020	21.4	164	20.4
Not at all useful	448	9.4	113	14.1
Not applicable	124	2.6	94	11.7
Total	**4,774**	**100**	**803**	**100**

e. Post-observation conferences

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very useful	1,587	33.1	191	23.6
Somewhat useful	1,641	34.3	251	31.1
Slightly useful	918	19.2	157	19.4
Not at all useful	457	9.5	110	13.6
Not applicable	187	3.9	99	12.3
Total	**4,790**	**100**	**808**	**100**

(Asked Teachers)

Q10. During your Component V Conference this year, did you and your evaluator determine how student growth measures would be used in evaluating your instructional practice?

(Asked Specialists)

Q10. During your Component V Conference this year, did you and your evaluator determine how student growth measures would be used in evaluating your practice?

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Yes	3,656	77.1	576	72.5
No	597	12.6	102	12.8
Did not have a Component V Conference this year	486	10.3	117	14.7
Total	**4,739**	**100**	**795**	**100**

(*Asked if respondent and evaluator determined student growth measures during your Component V Conference)*

Q11. Were the goals mutually established between you and your evaluator?

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Yes	3,183	88.0	531	93.3
No	432	12.0	38	6.7
Total	**3,615**	**100**	**569**	**100**

(*Asked If respondent and evaluator determined student growth measures during your Component V Conference)*

Q12. How much do measures align with school goals?

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very much	1,809	49.6	244	42.8
Somewhat	1,158	31.8	160	28.1
Slightly	338	9.3	70	12.3
Not at all	106	2.9	43	7.5
Don't know	235	6.4	53	9.3
Total	**3,646**	**100**	**570**	**100**

Q13. During your Professional Responsibilities Conference this year, did you and your evaluator plan professional growth activities?

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Yes	1,941	40.9	369	46.5
No	2,129	44.9	297	37.5
Did not have a Professional Responsibilities Conference this year	675	14.2	127	16.0
Total	4,745	100	793	100

Q14. Please enter how many of the following observations you received this year:
Here we are defining observations as both short and full observations. For our purposes, walkthroughs are not included within our definition of observations.

a. Full observations

	Teachers		Specialists	
	Mean	**N**	**Mean**	**N**
	1.3	4616	1.1	700

b. Short observations

	Teachers		Specialists	
	Mean	**N**	**Mean**	**N**
	3.0	4023	1.3	587

Q15. Following each type of observation, how often did you receive actionable and specific feedback?

a. Full observations

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Always	2,576	58.2	269	38.6
Often	779	17.6	122	17.5
Sometimes	531	12.0	86	12.4
Rarely	228	5.2	35	5.0
Never	147	3.3	61	8.8
Not applicable	165	3.7	123	17.7
Total	**4,426**	**100**	**696**	**100**

b. Short observations

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Always	1,132	28.7	118	18.1
Often	640	16.2	90	13.8
Sometimes	530	13.4	74	11.4
Rarely	250	6.3	30	4.6
Never	232	5.9	53	8.1
Not applicable	1,160	29.4	286	43.9
Total	**3,944**	**100**	**651**	**100**

Q16. Please provide two examples of feedback you received that were actionable and specific:

[Open-Ended Responses Given]

Q16a. **Is your school/administrator clear about the difference between short observations for DPAS-II and walkthroughs or other observations?**

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Yes	2,977	64.9	393	53.0
No	565	12.3	60	8.1
Don't know	1,043	22.7	289	38.9
Total	**4,585**	**100**	**742**	**100**

Q17. How useful was your Summative Evaluation Conference during your evaluation last year?

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very useful	693	15.4	109	15.3
Somewhat useful	1,217	27.1	160	22.4
Slightly useful	761	16.9	133	18.6
Not at all useful	459	10.2	99	13.9
Not applicable	1,368	30.4	213	29.8
Total	**4,498**	**100**	**714**	**100**

Q18. Please indicate whether or not your evaluator did the following during your Summative Evaluation Conference last year:

(Asked Teachers)

a. My evaluator shared his or her overall impression of my (instructional) practice

(Asked Specialists)

a. My evaluator shared his or her overall impression of my practice

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Yes, my evaluator did this	2,992	70.5	470	70.3
No, my evaluator did not do this	226	5.3	45	6.7
Not applicable	1,026	24.2	154	23.0
Total	**4,244**	**100**	**669**	**100**

(Asked Teachers)

b. My evaluator provided recommendations designed to improve my instructional practice

(Asked Specialists)

b. My evaluator provided recommendations designed to improve my practice

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Yes, my evaluator did this	2,493	59.2	343	51.6
No, my evaluator did not do this	587	13.9	134	20.2
Not applicable	1,132	26.9	188	28.3
Total	**4,212**	**100**	**665**	**100**

(Asked Teachers)

c. My evaluator provided expectations designed to improve specific aspects of my instructional practice

(Asked Specialists)

c. My evaluator provided expectations designed to improve specific aspects of my practice

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Yes, my evaluator did this	1,556	37.0	224	33.7
No, my evaluator did not do this	1,219	29.0	214	32.2
Not applicable	1,429	34.0	226	34.0
Total	**4,204**	**100**	**664**	**100**

(Asked if evaluator provided expectations)

d. My evaluator provided a timeline for when I need to meet expectations

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Yes, my evaluator did this	621	51.4	96	56.1
No, my evaluator did not do this	418	34.6	49	28.7
Not applicable	170	14.1	26	15.2
Total	**1,209**	**100**	**171**	**100**

Q18a. Please indicate how many hours you have spent on the following DPAS-II tasks during the 2014-2015 school year:

a. Preparing for observations

	Teachers		Specialists	
	Mean	**N**	**Mean**	**N**
	5.5	**4315**	**3.6**	**650**

b. Being observed

	Teachers		Specialists	
	Mean	**N**	**Mean**	**N**
	2.8	**4321**	**1.5**	**644**

c. Receiving and reviewing feedback from evaluator

	Teachers		Specialists	
	Mean	N	Mean	N
	2.3	4308	1.4	654

d. Other

	Teachers		Specialists	
	Mean	N	Mean	N
	7.0	1224	3.8	275

Q18b. On a 0 to 10 scale, please indicate whether you think DPAS-II is an exercise in compliance.

	Teachers		Specialists	
	Mean	N	Mean	N
	8.5	4,425	7.7	692

Q18c. One a 0 to 10 scale, please indicate whether you think DPAS-II is an exercise in evaluation.

	Teachers		Specialists	
	Mean	N	Mean	N
	6.9	4,414	6.0	691

Q18d. On a 0 to 10 scale, please indicate whether you think DPAS-II is an exercise in instructional improvement.

	Teachers		Specialists	
	Mean	N	Mean	N
	5.8	4,425	4.9	694

(Asked Teachers)
Q19. What aspects of DPAS-II are most useful to you in improving your instructional practice?
(Asked Specialists)
Q19. What aspects of DPAS-II are most useful to you in improving your practice?

[Open-Ended Responses Given]

Q20. Over the past two years, the following changes to DPAS-II have been implemented. Do you think these changes enhance DPAS-II?

a. Changes to Component II and III: Evaluators may use short observations, which must be at least 10-minutes, after at least one full observation has occurred for Components II & III only.

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very much	436	10.1	60	8.7
Somewhat	1115	25.8	160	23.3
Slightly	784	18.1	70	10.2
Not at all	883	20.4	102	14.9
Don't know	1111	25.7	294	42.9
Total	**4329**	**100**	**686**	**100**

b. Changes to Component IV: Districts/charters can opt to strengthen Component IV, for example by substituting a collaboratively developed Component.

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very much	283	6.6	48	7.0
Somewhat	852	19.8	123	18.0
Slightly	699	16.3	72	10.6
Not at all	675	15.7	87	12.8
Don't know	1791	41.7	352	51.6
Total	**4300**	**100**	**682**	**100**

c. Credentialed Observers: Districts can credential additional observers to assist with the DPAS-II process.

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very much	258	6.0	50	7.4
Somewhat	707	16.5	103	15.2
Slightly	674	15.7	72	10.6
Not at all	1175	27.4	139	20.5
Don't know	1481	34.5	315	46.4
Total	**4295**	**100**	**679**	**100**

d. Criterion-level Ratings: All educators are required to receive ratings on each of the criteria in the DPAS-II for Teachers and Specialists rubric.

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very much	310	7.2	37	5.4
Somewhat	1129	26.3	141	20.6
Slightly	986	23.0	119	17.4
Not at all	935	21.8	156	22.8
Don't know	931	21.7	231	33.8
Total	**4291**	**100**	**684**	**100**

Q20a. The following changes to DPAS-II have been proposed. Do you think these changes enhance DPAS-II?

a. Changes in Weighting: Components I through IV will receive greater emphasis, as evaluators will have more discretion in using Component V scores when Components I through IV are strong.

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very much	1073	25.2	123	18.0
Somewhat	1341	31.5	177	26.0
Slightly	763	17.9	114	16.7
Not at all	475	11.2	91	13.3
Don't know	608	14.3	177	26.0
Total	**4260**	**100**	**682**	**100**

b. Increasing the Number of Rating Categories for Components I through IV: Each of Components I through IV will be assigned a score along a 4-point scale rather than a binary ("Satisfactory"/"Unsatisfactory") scale.

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very much	717	16.9	85	12.6
Somewhat	1376	32.4	179	26.4
Slightly	869	20.4	107	15.8
Not at all	737	17.3	138	20.4
Don't know	552	13.0	168	24.8
Total	**4251**	**100**	**677**	**100**

c. Annual Appraisals: Beginning in 2016-2017, Annual Summative Appraisals would be required of all teachers.

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very much	329	7.7	38	5.6
Somewhat	917	21.5	124	18.3
Slightly	819	19.2	96	14.2
Not at all	1546	36.2	242	35.7
Don't know	657	15.4	178	26.3
Total	**4268**	**100**	**678**	**100**

Q20b. Are you aware districts can be granted a waiver to implement an alternative evaluation system?

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Yes	667	15.5	81	11.8
No	3635	84.5	606	88.2
Total	**4302**	**100**	**687**	**100**

Q20c. Would you be interested in your district implementing an alternative evaluation system?

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Yes	1850	42.9	339	49.3
No	262	6.1	27	3.9
Don't know	2196	51.0	322	46.8
Total	**4308**	**100**	**688**	**100**

Section III: DPAS-II Utility

Q21. Of the five major components (as defined in the DPAS-II Guide) used in teacher evaluations, which do you believe are accurate indicators of teacher performance?

	Teachers; N=4,304	
	Total	**%**
Component I: Planning and Preparation		
Yes	2,841	66.0
No	1,463	34.0
Component II: Classroom Environment		
Yes	2,828	65.7
No	1,476	34.3
Component III: Instruction		
Yes	3,680	85.5
No	624	14.5
Component IV: Professional Responsibilities		
Yes	1,450	33.7
No	2,853	66.3
Component V: Student Improvement		
Yes	1,270	29.5
No	3,034	70.5
None of the above		
Yes	168	3.9
No	4,136	96.1
Don't know		
Yes	121	2.8
No	4,183	97.2

	Specialists; N=689	
	Total	**%**
Component I: Planning and Preparation		
Yes	309	44.9
No	380	55.1
Component II: Professional Practice and Delivery of Service		
Yes	466	67.6
No	223	32.4
Component III: Professional Consultation and Collaboration		
Yes	383	55.6
No	306	44.4
Component IV: Professional Responsibilities		
Yes	382	55.4
No	307	44.6
Component V: Student Improvement		
Yes	126	18.3
No	563	81.7
None of the above		
Yes	69	10.0
No	620	90.0
Don't know		
Yes	43	6.2
No	646	93.8

Q22. Thinking about the 2014-2015 school year, please indicate whether you agree or disagree with the following statements about DPAS-II:

a. DPAS-II is being implemented appropriately at my school.

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	725	17.0	117	17.0
Agree	1888	44.2	262	38.1
Neither agree nor disagree	910	21.3	138	20.1
Disagree	316	7.4	47	6.8
Strongly disagree	180	4.2	15	2.2
Don't know	253	5.9	109	15.8
Total	**4272**	**100**	**688**	**100**

b. I contributed to the changes in the DPAS-II process

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	99	2.3	25	3.7
Agree	348	8.2	54	7.9
Neither agree nor disagree	1150	27.1	162	23.8
Disagree	1041	24.5	150	22.0
Strongly disagree	922	21.7	180	26.4
Don't know	685	16.1	110	16.2
Total	**4245**	**100**	**681**	**100**

c. Educators have been adequately involved in improving DPAS-II.

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	116	2.7	25	3.6
Agree	599	14.0	92	13.4
Neither agree nor disagree	1084	25.4	156	22.8
Disagree	779	18.2	111	16.2
Strongly disagree	686	16.1	92	13.4
Don't know	1008	23.6	209	30.5
Total	**4272**	**100**	**685**	**100**

Q23. (DDOER) I am able to use the following components to extract information that improves my instruction:

a. Component I: Planning and Preparation

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	626	14.8	66	10.0
Agree	1988	47.1	226	34.1
Neither agree nor disagree	890	21.1	169	25.5
Disagree	402	9.5	80	12.1
Strongly disagree	245	5.8	90	13.6
Not applicable	69	1.6	32	4.8
Total	**4220**	**100**	**663**	**100**

(Asked Teachers)

b. Component II: Classroom Environment Instruction

(Asked Specialists)

b. Component II: Professional Practice and Delivery of Services

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	574	13.7	75	11.3
Agree	1998	47.6	273	41.0
Neither agree nor disagree	925	22.1	150	22.5
Disagree	377	9.0	63	9.5
Strongly disagree	241	5.7	80	12.0
Not applicable	80	1.9	25	3.8
Total	**4195**	**100**	**666**	**100**

(Asked Teachers)

c. Component III: Instruction

(Asked Specialists)

c. Component III: Professional Consultation and Collaboration

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	720	17.2	66	9.9
Agree	2197	52.4	264	39.7
Neither agree nor disagree	737	17.6	162	24.4
Disagree	262	6.3	65	9.8
Strongly disagree	210	5.0	78	11.7
Not applicable	65	1.6	30	4.5
Total	**4191**	**100**	**665**	**100**

d. Component IV: Professional Responsibilities

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	326	8.0	67	10.4
Agree	1297	31.9	235	36.3
Neither agree nor disagree	1324	32.5	167	25.8
Disagree	659	16.2	70	10.8
Strongly disagree	406	10.0	80	12.4
Not applicable	59	1.4	28	4.3
Total	**4071**	**100**	**647**	**100**

e. Component V, Measure A: State Assessment Scores

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	140	3.4	14	2.1
Agree	637	15.3	67	10.2
Neither agree nor disagree	949	22.8	121	18.4
Disagree	839	20.1	85	12.9
Strongly disagree	1015	24.4	165	25.1
Not applicable	584	14.0	205	31.2
Total	**4164**	**100**	**657**	**100**

f. Component V, Measure B: Content Assessments

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	209	5.0	16	2.5
Agree	1141	27.4	78	12.1
Neither agree nor disagree	1027	24.7	140	21.7
Disagree	695	16.7	78	12.1
Strongly disagree	835	20.1	126	19.5
Not applicable	256	6.1	207	32.1
Total	**4163**	**100**	**645**	**100**

g. Component V, Measure C: Growth Goals

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	239	5.7	26	3.9
Agree	1158	27.7	161	24.4
Neither agree nor disagree	1014	24.3	147	22.3
Disagree	741	17.7	111	16.8
Strongly disagree	804	19.2	172	26.1
Not applicable	221	5.3	43	6.5
Total	**4177**	**100**	**660**	**100**

Q24. Thinking about the 2014-2015 school year, please indicate whether you agree or disagree with the following statements about **your evaluator**:

a. My evaluator handles the workload pertaining to educator evaluations effectively.

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	1077	25.7	163	24.4
Agree	1806	43.1	258	38.7
Neither agree nor disagree	576	13.7	102	15.3
Disagree	381	9.1	41	6.1
Strongly disagree	210	5.0	20	3.0
Not applicable	145	3.5	83	12.4
Total	**4195**	**100**	**667**	**100**

(Asked Teachers)

b. My evaluator provides specific and actionable feedback about ways to improve my instructional practice.

(Asked Specialists)

b. My evaluator provides specific and actionable feedback about ways to improve my practice.

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	1047	25.1	142	21.5
Agree	1902	45.5	251	37.9
Neither agree nor disagree	633	15.2	143	21.6
Disagree	382	9.1	59	8.9
Strongly disagree	170	4.1	32	4.8
Not applicable	44	1.1	35	5.3
Total	**4178**	**100**	**662**	**100**

(Asked Teachers)

c. My evaluator has a background in or knowledge of the content area(s) I teach.

(Asked Specialists)

c. My evaluator has a background in or knowledge of the content area(s) I support.

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	945	22.5	113	16.9
Agree	1482	35.3	195	29.2
Neither agree nor disagree	580	13.8	104	15.6
Disagree	597	14.2	128	19.2
Strongly disagree	444	10.6	100	15.0
Not applicable	149	3.6	27	4.0
Total	**4197**	**100**	**667**	**100**

(Asked Teachers)

d. My evaluator has a background in or knowledge of the grade level(s) I teach.

(Asked Specialists)

d. My evaluator has a background in or knowledge of the grade level(s) I support.

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	1063	25.4	162	24.5
Agree	1790	42.8	282	42.6
Neither agree nor disagree	531	12.7	90	13.6
Disagree	376	9.0	41	6.2
Strongly disagree	227	5.4	39	5.9
Not applicable	199	4.8	48	7.3
Total	**4186**	**100**	**662**	**100**

e. I trust my evaluator.

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	1477	35.7	241	36.7
Agree	1583	38.3	240	36.5
Neither agree nor disagree	625	15.1	111	16.9
Disagree	195	4.7	19	2.9
Strongly disagree	202	4.9	24	3.7
Not applicable	50	1.2	22	3.3
Total	**4132**	**100**	**657**	**100**

(Asked Teachers)

f. My evaluator and I agree on what good teaching looks like in the classroom.

(Asked Specialists)

f. My evaluator and I agree on what good educator support looks like in the classroom

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	1360	32.6	155	23.6
Agree	1825	43.7	248	37.8
Neither agree nor disagree	575	13.8	116	17.7
Disagree	164	3.9	23	3.5
Strongly disagree	113	2.7	14	2.1
Not applicable	136	3.3	100	15.2
Total	**4173**	**100**	**656**	**100**

g. My evaluator has worked with me to set ambitious goals for student performance.

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	961	23.0	130	19.6
Agree	1602	38.3	199	30.0
Neither agree nor disagree	923	22.1	204	30.8
Disagree	425	10.2	59	8.9
Strongly disagree	224	5.4	39	5.9
Not applicable	48	1.1	32	4.8
Total	**4183**	**100**	**663**	**100**

(Asked Teachers)
Q25. This year, how often did you change your instructional practices based on feedback related to DPAS-II?
(Asked Specialists)
Q25. This year, how often did you change your practices based on feedback related to DPAS-II?

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Did not change my instructional practice	1797	43.0	443	66.2
Once this year	730	17.5	105	15.7
2-3 times this year	1092	26.1	82	12.3
About once every 2-3 months	299	7.2	23	3.4
About once a month or more	263	6.3	16	2.4
Total	**4181**	**100**	**669**	**100**

(Asked if changed practices based on feedback related to DPAS-II)
Q26. Please give an example of the most recent time you used the feedback.

[Open-Ended Responses Given]

Q27. Based on this year's feedback from your evaluation, how likely is it that you will change aspects of your instructional practice based on feedback from DPAS-II?

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very likely	586	14.0	66	9.9
Somewhat likely	1222	29.2	139	20.9
Slightly likely	1041	24.8	113	17.0
Not at all likely	999	23.8	238	35.7
Don't know	342	8.2	110	16.5
Total	**4190**	**100**	**666**	**100**

Q28. Thinking about the 2014-2015 school year, please indicate whether you agree or disagree with the following statements **about your professional development opportunities**.

a. I have access to Professional Development that incorporates DPAS-II measures.

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	346	8.3	49	7.4
Agree	1671	40.1	193	29.1
Neither agree nor disagree	964	23.1	131	19.8
Disagree	625	15.0	134	20.2
Strongly disagree	272	6.5	91	13.7
Don't know	294	7.0	65	9.8
Total	**4172**	**100**	**663**	**100**

b. I have access to Professional Development specifically related to my areas for growth indicated by DPAS-II.

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	281	6.8	40	6.1
Agree	1289	31.1	182	27.8
Neither agree nor disagree	1057	25.5	121	18.5
Disagree	791	19.1	134	20.5
Strongly disagree	397	9.6	117	17.9
Don't know	324	7.8	60	9.2
Total	**4139**	**100**	**654**	**100**

c. I have colleagues and/or or administrators at my school that create opportunities to learn about DPAS-II.

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	330	7.9	48	7.2
Agree	1508	36.2	214	32.2
Neither agree nor disagree	1084	26.0	170	25.6
Disagree	669	16.1	87	13.1
Strongly disagree	295	7.1	75	11.3
Don't know	282	6.8	71	10.7
Total	**4168**	**100**	**665**	**100**

d. Professional Development in my district is aligned with DPAS-II.

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	320	7.7	37	5.6
Agree	1454	35.0	173	26.0
Neither agree nor disagree	1119	27.0	170	25.6
Disagree	543	13.1	76	11.4
Strongly disagree	266	6.4	67	10.1
Don't know	448	10.8	142	21.4
Total	**4150**	**100**	**665**	**100**

Q29. Please indicate whether you agree or disagree with the following statements about student academic achievement and your instructional practice:

(Asked Teachers)

a. State Assessment Scores (Measure A) are an appropriate measure of my instructional practice

(Asked Specialists)

a. State Assessment Scores (Measure A) are an appropriate measure of my practice

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	64	1.5	4	0.6
Agree	346	8.3	29	4.4
Neither agree nor disagree	666	16.1	58	8.8
Disagree	906	21.9	109	16.5
Strongly disagree	1417	34.2	219	33.1
Don't know	747	18.0	242	36.6
Total	**4146**	**100**	**661**	**100**

(Asked Teachers)

b. Content Assessments (Measure B) are an appropriate measure of my instructional practice

(Asked Specialists)

b. Content Assessments (Measure B) are an appropriate measure of my practice

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	132	3.2	7	1.1
Agree	1013	24.5	57	8.6
Neither agree nor disagree	842	20.3	78	11.8
Disagree	855	20.7	95	14.4
Strongly disagree	1021	24.7	187	28.4
Don't know	275	6.6	235	35.7
Total	**4138**	**100**	**659**	**100**

(Asked Teachers)

c. Growth Goals (Measure C) are an appropriate measure of my instructional practice

(Asked Specialists)

c. Growth Goals (Measure C) are an appropriate measure of my practice

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	184	4.4	19	2.9
Agree	1024	24.7	151	22.8
Neither agree nor disagree	938	22.7	124	18.7
Disagree	750	18.1	129	19.5
Strongly disagree	864	20.9	211	31.8
Don't know	380	9.2	29	4.4
Total	**4140**	**100**	**663**	**100**

Q30. Please indicate whether you agree or disagree with the following statements.

a. (DDOER) I understand conceptually how Measure A (Student Growth Targets) are established

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	241	5.8	26	3.9
Agree	1693	40.7	191	28.9
Neither agree nor disagree	719	17.3	126	19.1
Disagree	541	13.0	65	9.8
Strongly disagree	371	8.9	44	6.7
Don't know	597	14.3	209	31.6
Total	**4162**	**100**	**661**	**100**

b. (DDOER) I understand conceptually how Measure B (Student Assessments) goals are established

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	305	7.3	24	3.6
Agree	2028	48.9	185	27.9
Neither agree nor disagree	749	18.0	129	19.5
Disagree	520	12.5	65	9.8
Strongly disagree	333	8.0	44	6.6
Don't know	215	5.2	216	32.6
Total	**4150**	**100**	**663**	**100**

Section IV: Fairness and Perceived Accuracy of DPAS-II

(Asked Teachers)
Q31. Are DPAS-II ratings an accurate representation of your instructional practice?
(Asked Specialists)
Q31. Are DPAS-II ratings an accurate representation of your practice?

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very accurate	162	3.9	14	2.1
Somewhat accurate	1550	37.2	188	28.1
Neither accurate nor inaccurate	852	20.4	139	20.8
Somewhat inaccurate	736	17.7	88	13.2
Very inaccurate	687	16.5	194	29.0
Don't know	180	4.3	46	6.9
Total	**4167**	**100**	**669**	**100**

Q32. Are the tasks required to complete DPAS-II easy to understand?

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very easy to understand	396	9.5	38	5.7
Somewhat easy to understand	1763	42.4	199	30.1
Neither easy nor difficult to understand	804	19.3	139	21.0
Somewhat difficult to understand	912	21.9	212	32.0
Very difficult to understand	283	6.8	74	11.2
Total	**4158**	**100**	**662**	**100**

Q33. In your DPAS-II evaluation, please indicate whether or not you feel that you were held to the same standards as other educators <u>in the following groups</u>.

a. Within your content area(s) at your school

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	733	17.7	92	14.0
Agree	1925	46.4	222	33.7
Neither agree nor disagree	554	13.4	124	18.8
Disagree	350	8.4	30	4.6
Strongly disagree	199	4.8	30	4.6
Don't know	387	9.3	160	24.3
Total	**4148**	**100**	**658**	**100**

b. Your school overall

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	608	14.6	82	12.3
Agree	1667	40.1	205	30.7
Neither agree nor disagree	624	15.0	134	20.1
Disagree	519	12.5	48	7.2
Strongly disagree	285	6.9	30	4.5
Don't know	449	10.8	168	25.2
Total	**4152**	**100**	**667**	**100**

c. Your district

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	456	11.0	61	9.2
Agree	1319	31.8	186	28.0
Neither agree nor disagree	773	18.6	123	18.5
Disagree	450	10.8	49	7.4
Strongly disagree	294	7.1	41	6.2
Don't know	862	20.8	205	30.8
Total	**4154**	**100**	**665**	**100**

Section V: Student Characteristics, Teaching Practices, and School Culture

Q35. I would like to continue working as an educator as long as I am able.

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	2161	51.9	321	48.3
Agree	1215	29.2	215	32.3
Neither agree nor disagree	361	8.7	78	11.7
Disagree	291	7.0	34	5.1
Strongly disagree	134	3.2	17	2.6
Total	**4162**	**100**	**665**	**100**

Q36. Overall, my school is a good place to work.

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	1750	42.0	307	46.1
Agree	1481	35.5	248	37.2
Neither agree nor disagree	512	12.3	63	9.5
Disagree	292	7.0	35	5.3
Strongly disagree	135	3.2	13	2.0
Total	**4170**	**100**	**666**	**100**

Q37. Now, thinking about **your students** during the 2014-2015 school year, please indicate whether you agree or disagree with the following statements.

a. The amount a student can learn is primarily related to family background.

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	450	10.8	59	8.9
Agree	1091	26.2	161	24.2
Neither agree nor disagree	1130	27.1	201	30.2
Disagree	1174	28.2	200	30.0
Strongly disagree	320	7.7	45	6.8
Total	**4165**	**100**	**666**	**100**

b. I can get through to the most difficult student.

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	489	11.8	81	12.2
Agree	2264	54.4	358	54.0
Neither agree nor disagree	1024	24.6	188	28.4
Disagree	326	7.8	30	4.5
Strongly disagree	56	1.3	6	0.9
Total	**4159**	**100**	**663**	**100**

c. I can help all students make at least one year's growth in academic achievement during the school year.

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Strongly agree	480	11.5	47	7.1
Agree	1846	44.4	181	27.5
Neither agree nor disagree	1102	26.5	302	45.9
Disagree	610	14.7	73	11.1
Strongly disagree	120	2.9	55	8.4
Total	**4158**	**100**	**658**	**100**

Q38. What percent were eligible for free and reduced-priced lunch?

	Teachers		Specialists	
	Mean	**N**	**Mean**	**N**
	63.3	**3807**	**65.3**	**65**

Q39. What percent were English language learners?

	Teachers		Specialists	
	Mean	**N**	**Mean**	**N**
	22.3	**3823**	**25.9**	**551**

Q40. What percent received special education supports?

	Teachers		Specialists	
	Mean	N	Mean	N
	30.7	3835	43.5	563

Q41. How many hours per week do you spend on the following:

a. Instruction

	Teachers		Specialists	
	Mean	N	Mean	N
	24.2	3897	10	570

b. Discipline

	Teachers		Specialists	
	Mean	N	Mean	N
	4.6	3964	3.4	568

c. Lesson planning

	Teachers		Specialists	
	Mean	N	Mean	N
	8.2	3965	3.1	577

d. Administration

	Teachers		Specialists	
	Mean	N	Mean	N
	1.5	3682	4.2	556

e. Grading

	Teachers		Specialists	
	Mean	N	Mean	N
	5.3	3973	.9	555

f. Meetings and professional development

	Teachers		Specialists	
	Mean	N	Mean	N
	3.2	3980	6.2	581

g. Other

	Teachers		Specialists	
	Mean	N	Mean	N
	3.7	1614	13.8	366

Q42. How many hours per week do you spend on tasks:

a. <u>With</u> students

	Teachers		Specialists	
	Mean	N	Mean	N
	26.7	3796	24.7	617

b. <u>Without</u> students

	Teachers		Specialists	
	Mean	N	Mean	N
	13.0	3808	16.4	608

Q43. During the 2014-2014 school year, how many hours per week do you spend with the following groups?

a. Data teams

	Teachers		Specialists	
	Mean	N	Mean	N
	0.9	3747	1.1	554

b. Lesson planning groups

	Teachers		Specialists	
	Mean	N	Mean	N
	1.5	3808	.5	547

c. Whole school

	Teachers		Specialists	
	Mean	N	Mean	N
	0.9	3687	1.7	536

d. Teacher associations/Union meetings

	Teachers		Specialists	
	Mean	N	Mean	N
	0.2	3657	.2	538

e. School-level committees

	Teachers		Specialists	
	Mean	N	Mean	N
	1.1	3706	1.4	547

f. District-wide professional development

	Teachers		Specialists	
	Mean	N	Mean	N
	0.7	3616	.8	541

g. Professional Learning Communities

	Teachers		Specialists	
	Mean	**N**	**Mean**	**N**
	1.9	**3812**	**1.3**	**561**

Q44. Does DPAS-II improve Professional Learning Communities?

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very much	88	2.1	7	1.1
Somewhat	701	17.0	89	13.6
Slightly	820	19.9	106	16.2
Not at all	1980	48.0	227	34.6
Don't know	533	12.9	227	34.6
Total	**4122**	**100**	**656**	**100**

(Asked Teachers)
Q45. At your school, is DPAS-II used to highlight strong educators and instructional practices?
(Asked Specialists)
Q45. At your school, is DPAS-II used to highlight strong specialists and practices?

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very much	214	5.2	20	3.1
Somewhat	830	20.3	80	12.2
Slightly	664	16.3	74	11.3
Not at all	1481	36.3	221	33.7
Don't know	892	21.9	260	39.7
Total	**4081**	**100**	**655**	**100**

Q46. At your school, is DPAS-II used to compare educators?

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Very much	231	5.6	12	1.8
Somewhat	600	14.6	54	8.2
Slightly	460	11.2	41	6.2
Not at all	1311	31.9	154	23.4
Don't know	1508	36.7	396	60.3
Total	**4110**	**100**	**657**	**100**

Q47. At your school, is DPAS-II a positive, negative or mixed influence on school culture?

	Teachers		Specialists	
	Total	**%**	**Total**	**%**
Positive influence	182	4.4	21	3.2
Negative influence	1448	35.0	263	39.7
Mixed influence	1971	47.7	248	37.4
Don't know	533	12.9	131	19.8
Total	**4134**	**100**	**663**	**100**

Section VIII: Understanding of DPAS-II

Q48. Please list the **main** area of DPAS-II that you would like to learn more about/receive specific training on:

[Open-Ended Responses Given]

Q49. Please list the **main** area that you feel could improve your growth and development as an educator:

[Open-Ended Responses Given]

Q50. You stated that you [INSERT ANSWER FROM Q3] that DPAS-II is fair and equitable. Please explain your response:

[Open-Ended Responses Given]

Q51. Do you have any suggestions about ways that the DPAS-II system can improve?

[Open-Ended Responses Given]

2015 DPAS-II Evaluation Survey: Administrators

The DPAS-II Evaluation Survey was conducted by Research for Action for the Delaware Department of Education. The survey for Administrators was conducted online from May 5 to June 15, 2015. The survey was designed to probe administrators' views of the DPAS-II Evaluation System.

The following tables include whole counts for each question and response item, as well as the corresponding percentage. Each question captures the total responders for that specific question. The responses include both completed and partial[2] respondents.

	Completed Interviews	**Partial Interviews**	**Total Responses**
Administrators	288	145	433

Screener Questions

S1. What is your title?

S5. Given your perspective as a leader, what do you think is the most important factor influencing student academic achievement?

[Open-Ended Responses Given]

Section I: General DPAS-II Perceptions

Q1. DPAS-II <u>for Administrators</u> is fair and equitable.

	Administrators	
	Total	**%**
Strongly agree	17	5.1
Agree	111	33.0
Neither agree nor disagree	110	32.7
Disagree	63	18.8
Strongly disagree	15	4.5
Don't know	20	6.0
Total	**336**	**100**

[2] Partial responses consist of any respondent who began the survey and did not complete the last question. This includes individuals who completed most of the survey but not the last question, and those who answered only a few questions.

Q2. DPAS-II for Administrators is one of the top three drivers of student achievement gains in your school or work location.

	Administrators	
	Total	**%**
Strongly agree	11	3.3
Agree	55	16.4
Neither agree nor disagree	77	22.9
Disagree	115	34.2
Strongly disagree	66	19.6
Don't know	12	3.6
Total	**336**	**100**

Q3. Do you understand how you are evaluated as a leader under DPAS-II for Administrators?

	Administrators	
	Total	**%**
Very much	184	54.9
Somewhat	121	36.1
Slightly	16	4.8
Not at all	14	4.2
Total	**335**	**100**

Q4. How much do you think DPAS-II for Administrators improves leadership performance?

	Administrators	
	Total	**%**
Very much	17	5.2
Somewhat	108	32.8
Slightly	115	35.0
Not at all	70	21.3
Don't know	19	5.8
Total	**329**	**100**

Q4a. How much do you think DPAS-II for Teacher/Specialist improves instructional/educator practice?

	Administrators	
	Total	**%**
Very much	17	5.1
Somewhat	169	50.6
Slightly	103	30.8
Not at all	40	12.0
Don't know	5	1.5
Total	**334**	**100**

Q5. How much do you think DPAS-II for Administrators informs professional development?

	Administrators	
	Total	**%**
Very much	15	4.6
Somewhat	112	34.3
Slightly	86	26.3
Not at all	92	28.1
Don't know	22	6.7
Total	**327**	**100**

Q5a. How much do you think DPAS-II for Teachers/Specialists informs professional development?

	Administrators	
	Total	**%**
Very much	32	9.7
Somewhat	136	41.2
Slightly	99	30.0
Not at all	57	17.3
Don't know	6	1.8
Total	**330**	**100**

Q6. Does DPAS-II for Administrators provide accurate ratings of leaders? Please indicate for each type of rating.

a. Criterion-level ratings

	Administrators	
	Total	**%**
Very accurate	12	4.0
Somewhat accurate	154	50.8
Neither accurate nor inaccurate	68	22.4
Somewhat inaccurate	32	10.6
Very inaccurate	10	3.3
Don't know	27	8.9
Total	**303**	**100**

b. Component-level ratings (I – IV)

	Administrators	
	Total	**%**
Very accurate	13	4.4
Somewhat accurate	154	51.7
Neither accurate nor inaccurate	69	23.2
Somewhat inaccurate	24	8.1
Very inaccurate	12	4.0
Don't know	26	8.7
Total	**298**	**100**

c. Component V, Part A (State Assessment Scores)

	Administrators	
	Total	**%**
Very accurate	5	1.6
Somewhat accurate	86	28.3
Neither accurate nor inaccurate	70	23.0
Somewhat inaccurate	54	17.8
Very inaccurate	60	19.7
Don't know	29	9.5
Total	**304**	**100**

d. Component V, Part B, Section 1 (Student Growth Measures)

	Administrators	
	Total	**%**
Very accurate	6	2.0
Somewhat accurate	103	33.6
Neither accurate nor inaccurate	74	24.1
Somewhat inaccurate	50	16.3
Very inaccurate	47	15.3
Don't know	27	8.8
Total	**307**	**100**

e. Component V, Part B, Section 2 (District Priority Achievement Measures)

	Administrators	
	Total	**%**
Very accurate	9	2.9
Somewhat accurate	97	31.5
Neither accurate nor inaccurate	79	25.6
Somewhat inaccurate	51	16.6
Very inaccurate	40	13.0
Don't know	32	10.4
Total	**308**	**100**

f. Summative Rating

	Administrators	
	Total	**%**
Very accurate	4	1.5
Somewhat accurate	102	37.9
Neither accurate nor inaccurate	77	28.6
Somewhat inaccurate	41	15.2
Very inaccurate	24	8.9
Don't know	21	7.8
Total	**269**	**100**

Q6a. Does DPAS-II for Teachers/Specialists provide accurate ratings of educators? Please indicate for each type of rating.

a. Criterion-level ratings

	Administrators	
	Total	**%**
Very accurate	29	9.3
Somewhat accurate	197	62.9
Neither accurate nor inaccurate	34	10.9
Somewhat inaccurate	33	10.5
Very inaccurate	13	4.2
Don't know	7	2.2
Total	**313**	**100**

b. Component-level ratings (I – IV)

	Administrators	
	Total	**%**
Very accurate	21	6.8
Somewhat accurate	209	67.2
Neither accurate nor inaccurate	26	8.4
Somewhat inaccurate	34	10.9
Very inaccurate	14	4.5
Don't know	7	2.3
Total	**311**	**100**

c. Component V, Measure A (State Assessment Scores)

	Administrators	
	Total	**%**
Very accurate	8	2.6
Somewhat accurate	113	36.1
Neither accurate nor inaccurate	45	14.4
Somewhat inaccurate	78	24.9
Very inaccurate	53	16.9
Don't know	16	5.1
Total	**313**	**100**

d. Component V, Measure B (Content Assessments)

	Administrators	
	Total	**%**
Very accurate	7	2.2
Somewhat accurate	133	42.1
Neither accurate nor inaccurate	40	12.7
Somewhat inaccurate	75	23.7
Very inaccurate	49	15.5
Don't know	12	3.8
Total	316	100

e. Component V, Measure C (Growth Goals)

	Administrators	
	Total	**%**
Very accurate	7	2.2
Somewhat accurate	131	42.0
Neither accurate nor inaccurate	42	13.5
Somewhat inaccurate	72	23.1
Very inaccurate	52	16.7
Don't know	8	2.6
Total	**312**	**100**

f. Summative Rating

	Administrators	
	Total	**%**
Very accurate	9	2.9
Somewhat accurate	161	51.9
Neither accurate nor inaccurate	47	15.2
Somewhat inaccurate	53	17.1
Very inaccurate	34	11.0
Don't know	6	1.9
Total	**310**	**100**

Q7. Overall, what grade would you give DPAS-II for Administrators?

	Administrators	
	Total	**%**
A	7	2.2
B	75	23.7
C	123	38.9
D	78	24.7
F	33	10.4
Total	**316**	**100**

Q7a. Overall, what grade would you give DPAS-II for Teachers/Specialists?

	Administrators	
	Total	**%**
A	8	2.5
B	96	30.2
C	112	35.2
D	73	23.0
F	29	9.1
Total	**318**	**100**

Section II: DPAS-II for Administrators Implementation

Q8. During the 2014-2015 school year, were you evaluated as a principal or assistant principal?

	Administrators	
	Total	**%**
Principal	129	37.1
Assistant Principal	162	46.6
District leaders	57	16.4
Total	**348**	**100**

Q8a. Who was your primary evaluator this year?

	Administrators	
	Total	**%**
Superintendent	59	16.8
Assistant Superintendent	23	6.5
District Administrator	125	35.5
Other	145	41.2
Total	**352**	**100**

Q9. Please indicate how useful the following steps were during <u>your</u> evaluation this year. *If you did not complete one of the following, please indicate that the item was "Not applicable".*

a. Goal Setting Conference

	Administrators	
	Total	**%**
Very useful	81	23.7
Somewhat useful	121	35.4
Slightly useful	76	22.2
Not at all useful	40	11.7
Not applicable	24	7.0
Total	**342**	**100.0**

b. Student Performance Goal-Setting Form

	Administrators	
	Total	**%**
Very useful	44	13.1
Somewhat useful	116	34.6
Slightly useful	81	24.2
Not at all useful	66	19.7
Not applicable	28	8.4
Total	**335**	**100.0**

c. Priority Leadership Area Form

	Administrators	
	Total	**%**
Very useful	59	17.3
Somewhat useful	113	33.1
Slightly useful	77	22.6
Not at all useful	46	13.5
Not applicable	46	13.5
Total	**341**	**100.0**

d. Direct observations

	Administrators	
	Total	**%**
Very useful	75	22.3
Somewhat useful	104	31.0
Slightly useful	68	20.2
Not at all useful	40	11.9
Not applicable	49	14.6
Total	**336**	**100.0**

e. Evidence collection

	Administrators	
	Total	**%**
Very useful	57	16.8
Somewhat useful	116	34.1
Slightly useful	91	26.8
Not at all useful	42	12.4
Not applicable	34	10.0
Total	**340**	**100.0**

f. Mid-Year Conference

	Administrators	
	Total	**%**
Very useful	66	19.5
Somewhat useful	111	32.8
Slightly useful	74	21.9
Not at all useful	40	11.8
Not applicable	47	13.9
Total	**338**	**100.0**

g. DPAS-II for Administrators Principal Practice Rubric

	Administrators	
	Total	**%**
Very useful	47	13.9
Somewhat useful	118	34.9
Slightly useful	80	23.7
Not at all useful	49	14.5
Not applicable	44	13.0
Total	**338**	**100.0**

Q10. In completing the Student Performance Goal-Setting Form, did you and your evaluator set targets on student academic achievement measures?

	Administrators	
	Total	**%**
Yes	277	82.9
No	24	7.2
Did not complete the Student Performance Goal-Setting Form this year	33	9.9
Total	**334**	**100.0**

(Asked if respondent and evaluator set targets on student academic achievement measures)
Q11. Were the goals mutually established between you and your evaluator?

	Administrators	
	Total	**%**
Yes	255	92.4
No	21	7.6
Total	**276**	**100.0**

(Asked if respondent and evaluator set targets on student academic achievement measures)
Q12. How much do measures align with school and LEA/district goals?

	Administrators	
	Total	**%**
Very much	200	72.5
Somewhat	61	22.1
Slightly	11	4.0
Not at all	2	0.7
Don't know	2	0.7
Total	**276**	**100.0**

Q13. In completing the Priority Leadership Area Form, did you and your evaluator identify areas to contribute to your growth as a leader?

	Administrators	
	Total	**%**
Yes	245	73.4
No	32	9.6
Did not complete the Priority Leadership Area Form this year	57	17.1
Total	**334**	**100.0**

Q13a. Was the mid-year conference helpful in making mid-year course corrections?

	Administrators	
	Total	**%**
Very helpful	45	13.5
Somewhat helpful	109	32.6
Slightly helpful	68	20.4
Not at all helpful	42	12.6
Not applicable	70	21.0
Total	**334**	**100.0**

Q13b. Was the mid-year conference productive for your leadership growth?

	Administrators	
	Total	**%**
Very productive	49	14.8
Somewhat productive	103	31.1
Slightly productive	77	23.3
Not at all productive	40	12.1
Not applicable	62	18.7
Total	**331**	**100.0**

Q17. How useful was your Summative Evaluation Conference during your evaluation last year?

	Administrators	
	Total	**%**
Very useful	41	12.7
Somewhat useful	80	24.7
Slightly useful	62	19.1
Not at all useful	40	12.3
Not applicable	101	31.2
Total	**324**	**100.0**

Q18. Please indicate whether or not your evaluator did the following during your Summative Evaluation Conference last year:

a. My evaluator shared his or her overall impression of my leadership performance

	Administrators	
	Total	**%**
Yes, my evaluator did this	212	67.9
No, my evaluator did not do this	28	9.0
Not applicable	72	23.1
Total	**312**	**100.0**

b. My evaluator provided recommendations designed to improve my leadership performance

	Administrators	
	Total	**%**
Yes, my evaluator did this	173	55.8
No, my evaluator did not do this	59	19.0
Not applicable	78	25.2
Total	**310**	**100.0**

c. My evaluator provided expectations designed to improve specific aspects of my leadership performance

	Administrators	
	Total	**%**
Yes, my evaluator did this	72	23.2
No, my evaluator did not do this	132	42.6
Not applicable	106	34.2
Total	**310**	**100.0**

(Asked if evaluator provided expectations)

d. My evaluator provided a timeline for when I need to meet expectations

	Administrators	
	Total	**%**
Yes, my evaluator did this	29	50.9
No, my evaluator did not do this	20	35.1
Not applicable	8	14.0
Total	**57**	**100.0**

Q18a. (Adapted from TN) Please indicate how many hours you have spent on the following DPAS-II for Administrators tasks during the 2014-2015 school year.

a. Preparing and attending the Goal-Setting Conference

	Administrators	
	Mean	**N**
	3.5	**284**

b. Preparing and participating in direct observations

	Administrators	
	Mean	N
	5.5	258

c. Collecting evidence

	Administrators	
	Mean	N
	6.3	272

d. Preparing and attending the Mid-Year Conference

	Administrators	
	Mean	N
	2.9	280

e. Preparing and attending the Summative Conference

	Administrators	
	Mean	N
	3.2	264

f. Receiving and reviewing feedback from your evaluator

	Administrators	
	Mean	N
	2.7	279

g. Other

	Administrators	
	Mean	N
	4.3	59

Q18b. On a 0 to 10 scale, please indicate whether you think DPAS-II for Administrators **is** an exercise in compliance?

	Administrators	
	Mean	N
	8.9	311

Q18c. On a 0 to 10 scale, please indicate whether you think DPAS-II for Administrators **is** an exercise in evaluation?

	Administrators	
	Mean	N
	6.4	313

Q18d. On a 0 to 10 scale, please indicate whether you think DPAS-II for Administrators **is** an exercise in professional growth?

	Administrators	
	Mean	**N**
	6.0	**312**

Q19. What aspects of DPAS-II for Administrators are most useful to you in improving your leadership performance?

[Open-Ended Responses Given]

Q20. Over the past two years, the following changes to DPAS-II for Teachers/Specialists have been implemented. Do you think these changes enhance DPAS-II?

e. Changes to Component II and III: Evaluators may use short observations, which must be at least 10-minutes, after at least one full observation has occurred for Components II & III only.

	Administrators	
	Total	**%**
Very much	58	19.0
Somewhat	104	34.1
Slightly	64	21.0
Not at all	54	17.7
Don't know	25	8.2
Total	**305**	**100.0**

f. Changes to Component IV: Districts/charters can opt to strengthen Component IV, for example by substituting a collaboratively developed Component.

	Administrators	
	Total	**%**
Very much	26	8.5
Somewhat	60	19.7
Slightly	49	16.1
Not at all	60	19.7
Don't know	110	36.1
Total	**305**	**100.0**

g. Credentialed Observers: Districts can credential additional observers to assist with the DPAS-II process.

	Administrators	
	Total	**%**
Very much	52	17.0
Somewhat	74	24.3
Slightly	45	14.8
Not at all	78	25.6
Don't know	56	18.4
Total	**305**	**100.0**

h. Criterion-level Ratings: All educators are required to receive ratings on each of the criteria in the DPAS-II for Teachers and Specialists rubric.

	Administrators	
	Total	**%**
Very much	43	14.1
Somewhat	112	36.7
Slightly	68	22.3
Not at all	66	21.6
Don't know	16	5.2
Total	**305**	**100.0**

Q20a. The following changes to DPAS-II for Teachers/Specialists have been proposed. Do you think these changes enhance DPAS-II?

d. Changes in Weighting: Components I through IV will receive greater emphasis, as evaluators will have more discretion in using Component V scores when Components I through IV are strong.

	Administrators	
	Total	**%**
Very much	104	34.4
Somewhat	101	33.4
Slightly	51	16.9
Not at all	26	8.6
Don't know	20	6.6
Total	**302**	**100.0**

e. Increasing the Number of Rating Categories for Components I through IV: Each of Components I through IV will be assigned a score along a 4-point scale rather than a binary ("Satisfactory"/"Unsatisfactory") scale.

	Administrators	
	Total	**%**
Very much	53	17.5
Somewhat	114	37.7
Slightly	52	17.2
Not at all	50	16.6
Don't know	33	10.9
Total	**302**	**100.0**

f. Annual Appraisals: Beginning in 2016-2017, Annual Summative Appraisals would be required of all teachers.

	Administrators	
	Total	**%**
Very much	23	7.6
Somewhat	71	23.5
Slightly	34	11.3
Not at all	152	50.3
Don't know	22	7.3
Total	**302**	**100.0**

Q20b. Are you aware districts can be granted a waiver to implement an alternative evaluation system?

	Administrators	
	Total	**%**
Yes	171	55.9
No	135	44.1
Total	**306**	**100.0**

Q20c. Would you be interested in your district implementing an alternative evaluation system?

	Administrators	
	Total	**%**
Yes	158	51.8
No	42	13.8
Don't know	105	34.4
Total	**305**	**100.0**

Section III: DPAS-II for Administrators' Utility

Q21. Of the five major components (as defined in the DPAS-II Guide) used in administrator evaluations, which do you believe are accurate indicators of leadership performance?

	Administrators; N=304	
	Total	**%**
Component I: Vision and Goals		
Yes	197	64.8
No	107	35.2
Component II: Teaching and Learning		
Yes	223	73.3
No	81	26.7
Component II: Teaching and Learning		
Yes	200	65.8
No	104	34.2
Component IV: Professional Responsibilities		
Yes	140	46.1
No	164	53.9
Component V: Student Improvement		
Yes	137	45.1
No	167	54.9
None of the above		
Yes	10	3.3
No	294	96.7
Don't know		
Yes	18	5.9
No	286	94.1

Q22. Thinking about the 2014-2015 school year, please indicate whether you agree or disagree with the following statements about DPAS-II for Administrators:

a. DPAS-II for Administrators is being implemented appropriately at my school.

	Administrators	
	Total	**%**
Strongly agree	71	23.5
Agree	137	45.4
Neither agree nor disagree	53	17.5
Disagree	15	5.0
Strongly disagree	17	5.6
Don't know	9	3.0
Total	**302**	**100.0**

b. I contributed to the changes in the DPAS-II for Administrators process.

	Administrators	
	Total	**%**
Strongly agree	20	6.7
Agree	34	11.3
Neither agree nor disagree	76	25.3
Disagree	65	21.7
Strongly disagree	65	21.7
Don't know	40	13.3
Total	**300**	**100.0**

c. Leaders have been adequately involved in improving DPAS-II for Administrators.

	Administrators	
	Total	**%**
Strongly agree	20	6.6
Agree	54	17.9
Neither agree nor disagree	77	25.6
Disagree	47	15.6
Strongly disagree	37	12.3
Don't know	66	21.9
Total	**300**	**100.0**

Q23. I am able to use the following components to extract information that improves my leadership performance:

a. Component I: Vision and Goals

	Administrators	
	Total	**%**
Strongly agree	20	6.8
Agree	151	51.4
Neither agree nor disagree	73	24.8
Disagree	23	7.8
Strongly disagree	18	6.1
Don't know	9	3.1
Total	**294**	**100.0**

b. Component II: Teaching and Learning

	Administrators	
	Total	**%**
Strongly agree	24	8.2
Agree	156	53.6
Neither agree nor disagree	70	24.1
Disagree	18	6.2
Strongly disagree	14	4.8
Don't know	9	3.1
Total	**291**	**100.0**

c. Component III: People, Systems and Operations

	Administrators	
	Total	**%**
Strongly agree	23	7.8
Agree	154	52.6
Neither agree nor disagree	76	25.9
Disagree	19	6.5
Strongly disagree	13	4.4
Don't know	8	2.7
Total	**293**	**100.0**

d. Component IV: Professional Responsibilities

	Administrators	
	Total	**%**
Strongly agree	21	7.2
Agree	143	49.1
Neither agree nor disagree	81	27.8
Disagree	24	8.2
Strongly disagree	14	4.8
Don't know	8	2.7
Total	**291**	**100.0**

e. Component V, Part A: State Assessment Scores

	Administrators	
	Total	**%**
Strongly agree	11	3.9
Agree	92	32.3
Neither agree nor disagree	65	22.8
Disagree	58	20.4
Strongly disagree	42	14.7
Don't know	17	6.0
Total	**285**	**100.0**

f. Component V, Part B, Section 1: Student Growth Measures

	Administrators	
	Total	**%**
Strongly agree	14	4.8
Agree	113	39.0
Neither agree nor disagree	64	22.1
Disagree	51	17.6
Strongly disagree	36	12.4
Don't know	12	4.1
Total	**290**	**100.0**

g. Component V, Part B, Section 2: District Priority Achievement Measures

	Administrators	
	Total	**%**
Strongly agree	13	4.5
Agree	92	31.9
Neither agree nor disagree	92	31.9
Disagree	39	13.5
Strongly disagree	32	11.1
Don't know	20	6.9
Total	**288**	**100.0**

Q24. Thinking about the 2014-2015 school year, please indicate whether you agree or disagree with the following statements about **your evaluator**:

a. My evaluator handles the workload pertaining to educator evaluations effectively.

	Administrators	
	Total	**%**
Strongly agree	78	26.8
Agree	127	43.6
Neither agree nor disagree	40	13.7
Disagree	17	5.8
Strongly disagree	20	6.9
Don't know	9	3.1
Total	**291**	**100.0**

b. My evaluator provides specific and actionable feedback about ways to improve my leadership performance.

	Administrators	
	Total	**%**
Strongly agree	75	26.0
Agree	117	40.5
Neither agree nor disagree	42	14.5
Disagree	28	9.7
Strongly disagree	20	6.9
Don't know	7	2.4
Total	**289**	**100.0**

c. I trust my evaluator.

	Administrators	
	Total	**%**
Strongly agree	130	45.0
Agree	95	32.9
Neither agree nor disagree	30	10.4
Disagree	17	5.9
Strongly disagree	12	4.2
Don't know	5	1.7
Total	**289**	**100.0**

d. My evaluator and I agree on what good leadership looks like in a LEA/district.

	Administrators	
	Total	**%**
Strongly agree	102	35.1
Agree	120	41.2
Neither agree nor disagree	36	12.4
Disagree	12	4.1
Strongly disagree	11	3.8
Don't know	10	3.4
Total	**291**	**100.0**

e. My evaluator has worked with me to set ambitious goals for student performance.

	Administrators	
	Total	**%**
Strongly agree	89	30.7
Agree	104	35.9
Neither agree nor disagree	50	17.2
Disagree	21	7.2
Strongly disagree	22	7.6
Don't know	4	1.4
Total	**290**	**100.0**

Q25. This year, how often did you change your leadership performance based on feedback related to DPAS-II for Administrators?

	Administrators	
	Total	**%**
Did not change my leadership performance	173	60.5
Once this year	35	12.2
2-3 times this year	62	21.7
About once every 2-3 months	8	2.8
About once a month or more	8	2.8
Total	**286**	**100.0**

(Asked if changed practices based on feedback related to DPAS-II)

Q26. Please give an example of the most recent time you used the feedback.

[Open-Ended Responses Given]

Q27. Based on this year's feedback from your evaluation, how likely is it that you will change aspects of your leadership performance based on feedback from DPAS-II for Administrators?

	Administrators	
	Total	**%**
Very likely	50	17.2
Somewhat likely	89	30.7
Slightly likely	56	19.3
Not at all likely	59	20.3
Don't know	36	12.4
Total	**290**	**100.0**

Q29. Please indicate whether you agree or disagree with the following statements about student academic achievement and your leadership performance:

a. State Assessment Scores (Part A) are an appropriate measure of my leadership performance.

	Administrators	
	Total	**%**
Strongly agree	8	2.8
Agree	44	15.2
Neither agree nor disagree	71	24.5
Disagree	73	25.2
Strongly disagree	81	27.9
Don't know	13	4.5
Total	**290**	**100.0**

b. Student Growth Measures (Part B, Section 1) are an appropriate measure of my leadership performance.

	Administrators	
	Total	**%**
Strongly agree	11	3.8
Agree	89	30.7
Neither agree nor disagree	65	22.4
Disagree	59	20.3
Strongly disagree	56	19.3
Don't know	10	3.4
Total	**290**	**100.0**

c. District Priority Achievement Measures (Part B, Section 2) are an appropriate measure of my leadership performance.

	Administrators	
	Total	**%**
Strongly agree	11	3.8
Agree	87	30.0
Neither agree nor disagree	83	28.6
Disagree	42	14.5
Strongly disagree	48	16.6
Don't know	19	6.6
Total	**290**	**100.0**

Section IV: Fairness and Perceived Accuracy of DPAS-II for Administrators

Q31. Are DPAS-II ratings for Administrators an accurate representation of your leadership performance?

	Administrators	
	Total	**%**
Very accurate	10	3.4
Somewhat accurate	118	40.4
Neither accurate nor inaccurate	65	22.3
Somewhat inaccurate	46	15.8
Very inaccurate	35	12.0
Don't know	18	6.2
Total	**292**	**100.0**

Q32. Are the tasks required to complete DPAS-II for Administrators, as an administrator being evaluated, easy to understand?

	Administrators	
	Total	**%**
Very easy	33	11.5
Somewhat easy	145	50.5
Neither easy nor difficult	59	20.6
Somewhat difficult	43	15.0
Very difficult	7	2.4
Total	**287**	**100.0**

Q33. In your DPAS-II evaluation, please indicate whether or not you feel that you were held to the same standards as other leaders in the following groups.

a. Your LEA/district

	Administrators	
	Total	**%**
Strongly agree	43	15.0
Agree	117	40.9
Neither agree nor disagree	40	14.0
Disagree	24	8.4
Strongly disagree	13	4.5
Don't know	49	17.1
Total	**286**	**100.0**

b. Delaware

	Administrators	
	Total	**%**
Strongly agree	25	9.0
Agree	73	26.2
Neither agree nor disagree	52	18.6
Disagree	33	11.8
Strongly disagree	12	4.3
Don't know	84	30.1
Total	**279**	**100.0**

Section V: Student Characteristics, Teaching Practices, and School Culture

Q35. I would like to continue working as an administrator as long as I am able.

	Administrators	
	Total	**%**
Strongly agree	181	62.6
Agree	74	25.6
Neither agree nor disagree	21	7.3
Disagree	10	3.5
Strongly disagree	3	1.0
Total	**289**	**100.0**

Q36. Overall, my LEA/district is a good place to work.

	Administrators	
	Total	**%**
Strongly agree	151	52.4
Agree	92	31.9
Neither agree nor disagree	20	6.9
Disagree	20	6.9
Strongly disagree	5	1.7
Total	**288**	**100.0**

Q37. The amount a student can learn is primarily related to family background.

	Administrators	
	Total	**%**
Strongly agree	9	3.1
Agree	42	14.4
Neither agree nor disagree	62	21.2
Disagree	108	37.0
Strongly disagree	71	24.3
Total	**292**	**100.0**

Q41. How many hours per week do you spend on the following:

a. Discipline

	Administrators	
	Mean	**N**
	8.4	277

b. General administrative tasks

	Administrators	
	Mean	**N**
	17.8	264

c. Meetings

	Administrators	
	Mean	**N**
	10.4	278

d. Professional development

	Administrators	
	Mean	**N**
	4.0	271

e. Other

	Administrators	
	Mean	**N**
	10.5	159

Q43. During the 2014-2015 school year, how many hours per week do you spend with the following groups?

a. Data teams

	Administrators	
	Mean	**N**
	3.1	260

b. Lesson planning groups

	Administrators	
	Mean	**N**
	2.4	256

c. Whole school

	Administrators	
	Mean	**N**
	3.7	248

d. Teacher associations/Union meetings

	Administrators	
	Mean	**N**
	.7	253

e. School-level committees

	Administrators	
	Mean	**N**
	3.6	261

f. District-wide professional development

	Administrators	
	Mean	**N**
	1.6	248

Q45. In your LEA/district, is DPAS-II for Administrators used to highlight strong leaders and leadership performances?

	Administrators	
	Total	**%**
Very much	18	6.4
Somewhat	94	33.5
Slightly	61	21.7
Not at all	108	38.4
Total	**281**	**100.0**

Q46. In your LEA/district, is DPAS-II for Administrators used to compare leaders?

	Administrators	
	Total	**%**
Very much	13	4.7
Somewhat	61	21.9
Slightly	44	15.8
Not at all	160	57.6
Total	**278**	**100.0**

Q47. In your LEA/district, is DPAS-II for Administrators a positive, negative or mixed influence on school or district culture?

	Administrators	
	Total	**%**
Positive influence	48	16.7
Negative influence	36	12.5
Mixed influence	111	38.5
Don't know	93	32.3
Total	**288**	**100.0**

Section VI: Understanding of DPAS-II for Administrators

Q48. Please list the **main** area of DPAS-II for Administrators that you would like to learn more about/receive specific training on:

[Open-Ended Responses Given]

Q49. Please list the **main** factor that you feel could improve your growth and development as a leader:

[Open-Ended Responses Given]

Q50. You stated that you [INSERT ANSWER FROM Q3] that DPAS-II for Administrators is fair and equitable. Please explain your response:

[Open-Ended Responses Given]

Q51. Do you have any suggestions about ways that the DPAS-II for Administrators system can improve?

[Open-Ended Responses Given]

Appendix C. Multivariate Statistical Analysis

To examine the factors that influence understanding and views of DPAS-II, fidelity of implementation, and changes in practice, we estimate a series of multi-level multivariate statistical models using data from the statewide teacher, specialist, and administrator surveys. In this appendix, we discuss the variables and methods used in this analysis.

Dependent Variables

Below, we examine the seven dependent variables used in this analysis. We estimated a set of statistical models for each of these outcomes. We performed separate statistical models for teachers (Tables 1C-7C), specialists (Tables 9C-10C), and administrators (Tables 12C-19C).

Goal One Measures of Understanding and Views of DPAS-II

1. Overall Views of DPAS-II (measured on a 4.0 GPA scale) [see tables 1C, 9C, and 12C]
2. Views of Fairness of DPAS-II (measured as a 1 if the respondent said that they strongly agree or agree that DPAS-II is fair and equitable, and as a 0 if the respondent stated that they are neutral, disagree, or strongly disagree that DPAS-II is fair and equitable) [see tables 2C, 9C, and 13C]
3. Understanding of DPAS-II (measured as a 1 if the respondent said that they very much or somewhat understand the DPAS-II evaluation system, and as a 0 if the respondent stated that they slightly or do not at all understand the DPAS-II evaluation system) [see tables 3C, 9C,
and 14C]

Goal Two Measure of Implementation of DPAS-II

4. Fidelity of Implementation (This is a 0 to 1 continuous scale that represents the average number of fidelity of implementation indicators experienced by a teacher, specialist, or administrator. Refer to the main report for a list of the questions used to collapse this scale. All scales held together with a reliability [Cronbach's alpha] of .6 or greater.) [see table 4C, 9C, and 15C]

Goal Three Measures of Intermediate Outcomes

5. Change in Practice (This is a variable that measures if a teacher, specialist, or administrator changed their practice at least once a year based on feedback from DPAS-II. A value of 1 indicates that the teacher, specialist, or administrator changed her or his practice at least one time during the past year. A value of 0 indicates that she or he did not change practice during the past year.) [see table 5C, 10C, and 16C]
6. School Culture Indicator #1: Quality of Workplace (This variable is based on responses to the question: "Is your school [or district] a good place to work?" A value of 1 indicates that a respondent answered that they strongly agree or agree that her or his school is a good place to work. A value of 0 indicates an answer of neutral, disagree, or strongly disagree.) [see table 6C, 10C, and 17C]
7. School Culture Indicator #2: Job Satisfaction and Desire to Continue Teaching/Working (This variable is based on responses to the statement: "I wish to continue working as [a teacher, an educator, or an administrator] as long as I am able." A value of 1 indicates that a respondent

answered that they strongly agree or agree with this statement. A value of 0 indicates an answer of neutral, disagree, or strongly disagree.) [see table 7C, 10C, and 18C]

Independent Variables and Controls

Below we list the key independent variables and controls used in our analysis. For the administrator and specialist analysis, we included a smaller subset of controls because the total sample size was only about 600 respondents and there were fewer degrees of freedom than in the teacher sample that had a sample size of over 4000 respondents. If one of the variables below was used as the dependent variable it was not included as an independent or control variable.

Independent Variables

- Overall Views of DPAS-II
  - Grade given to DPAS-II (measured on a 4.0 scale)
- Views of the Purpose of DPAS-II
  - Do you view DPAS-II as a tool to improve teaching/practice? (Often/ Somewhat/ Slightly/ Not at all scale; in most analyses, "not at all" is omitted)
  - Do you view DPAS-II as a tool to inform professional development? (Often/ Somewhat/ Slightly/ Not at all scale; in most analyses, "not at all" is omitted)
- Understanding of DPAS-II
  - Do you understand the DPAS-II evaluation system? (Very much/ Somewhat/ Slightly/ Not at all scale; in most cases, "not at all" is omitted)
- Views of Components I-V (Are these components accurate measures of practice?)
  - Component I (1 if agree is accurate, 0 if not accurate)
  - Component II (1 if agree is accurate, 0 if not accurate)
  - Component III (1 if agree is accurate, 0 if not accurate)
  - Component IV (1 if agree is accurate, 0 if not accurate)
  - Component V (1 if agree is accurate, 0 if not accurate)
- Fidelity of Implementation (0 to 1 scale)
- School Culture
  - School Culture Indicator #1: Quality of Workplace (This variable is based on responses to the question: "Is your school [or district] a good place to work?" A value of 1 indicates that a respondent answered that they strongly agree or agree that her or his school is a good place to work. A value of 0 indicates an answer of neutral, disagree, or strongly disagree.)
  - School Culture Indicator #2: Job Satisfaction and Desire to Continue Teaching/Working (This variable is based on responses to the statement: "I would like to continue working as an [educator or administrator] as long as I am able." A value of 1 indicates that a respondent answered that they strongly agree or agree with this question. A value of 0 indicates an answer of neutral, disagree, or strongly disagree.)
  - School Culture Indicator #3: Influence of DPAS-II on school culture (This variable is based on responses to the question: At your school, is DPAS-II a positive, negative, or mixed influence on school culture? To represent this question use two variables to represent this question: positive (i.e. 1 if the respondent answered positive and 0 otherwise) and mixed (1 if the respondent answered mixed and 0 otherwise). These

coefficients should be interpreted as the difference between holding a positive versus holding a negative view and the difference between holding mixed versus a negative view.)

Controls (some of these only apply to the teacher survey)

- Educator or Administrator Efficacy
    - Does family background determine achievement? (We asked the respondent: "Do you agree or disagree that the amount a student can learn is primarily related to family background?" The answers were on an agree/disagree scale.)
    - Can you help all students achieve one year of academic growth? (We asked the respondent: "Do you agree or disagree that you can help all students make at least one year's growth in academic achievement during the school year?" The answers were on an agree/disagree scale.)
- School Poverty (measured by percent free and reduced lunch)
- Sex of the Respondent (1=female, 0= male)
- Years of Experience (years of experience)
- Subject
    - Does the respondent teach English/Language Arts?
    - Does the respondent teach Mathematics?
- Grade Level (Is the respondent an elementary teacher or a middle school/high school teacher?)
- Question order effect (Were the DPAS-II evaluation questions placed at the beginning or end of the survey? This has a value of 1 if DPAS-II evaluation questions were placed at the end of the survey, and a value of 0 if they were placed at the end of the survey.)

See tables 8C, 11C, and 19C for the descriptive statistics for the dependent variables, independent variables, and controls for teacher, specialist, and administrator surveys.

Methodology

We estimated the factors that influence each of the above dependent variables with a series of multilevel regressions. If the dependent variable was continuous (i.e., GPA and implementation variables), we use a Hierarchical Linear Model; if the dependent variable was binary (0 or 1) such as the fairness, understanding, change in practice, and school culture variables, we used a multilevel logistic regression.

The coefficients in the linear models show the change in the dependent variable if the independent variable increases by one unit. The coefficients in the logistic model show the change in the log-odds of the dependent variable if the independent variable changes by one unit. A positive number means a positive correlation and a negative number means a negative correlation. If a variable has a statistically significant effect, it is identified with a *, **, or *** depending on the level of statistical significance. If a coefficient does not have a *,**, or *** it is not statistically significant. Lack of statistical significance means that an effect could have occurred by chance and the coefficient shown is not statistically significantly different from zero. We use the Bayesian Information Criteria (bic) to select the best fit model. The BIC is a global test of fit that accounts for the effects of sample size and the number of parameters in the model. A smaller BIC means that the model has a better fit.

The coefficient in the logistic regression shows the change in log-odds of the dependent variable for a one unit change in the independent variable. The dependent variable is in log-odds to avoid unrealistic estimates of predicted values above 1 and below 0. A log odds is the natural log of the odds ratio. The odds ratio is the proportion of a dichotomous (0/1) variable that have a value of 1 over the proportion that have a value of 0. In the body of the report, we present the logistic regression as predicted probabilities. To convert the log odds into a predicted probabilities, we calculated the predicted value (Y) in log-odds where Y=Constant + Coef#1*Variable#1 + … + CoefN*VariableN. Next we transformed Y into a proportion with the following equation: Predicted Probability = (e^Y/(1+e^Y)). We describe the predicted probability in the text as the percentage of individuals who would have responded with a 1 to the question of interest for a given set of independent variables.

We used multilevel models to account for the nested nature of teachers and specialists within schools within school districts. For administrators, we only account for administrators nested within districts. Our multilevel analysis partially controls for the fact that there might be a number of unmeasured influences on the dependent variable that are common to all teachers or specialists in a given school or in a given district.

We use list wise deletion to address missing data. We assume that missing responses to an individual question are missing at random. To partially test this assumption, we examined if the average responses to the set of questions about knowledge of DPAS-II differed between the first 1000 respondents, the first 2000 respondents, and the full sample. We found no statistically significant differences in the responses between these groups. This provides evidence that missing data probably has a similar pattern as the available data. We also randomly varied the placement of questions about views of DPAS-II, and found that there was no difference in the responses to this question based on placement in the completed surveys versus the partially completed surveys. This suggests that individuals who did not complete the entire survey were not notably different in their views from individuals who completed the entire survey.

Table 1C: Factors that Influence Views of DPAS-II for Administrators (on a 4.0 GPA scale), A Multilevel Regression

	Model1	Model2	Model3	Model4
	coef/se	coef/se	coef/se	coef/se
DPAS-II improves practice – "Often"/"Somewhat"	1.099***			0.802***
	(0.034)			(0.036)
DPAS-II improves practice – "Slightly"	0.570***			0.434***
	(0.029)			(0.029)
DPAS-II informs Professional Development – "Often"/"Somewhat"	0.508***			0.395***
	(0.033)			(0.034)
DPAS-II informs Professional Development – "Slightly"	0.233***			0.181***
	(0.029)			(0.030)
Teacher understands DPAS-II evaluation system – "Very much"/"Somewhat"		0.777***		0.228*
		(0.116)		(0.094)
Teacher understands DPAS-II evaluation system – "Slightly"		0.379**		0.105
		(0.127)		(0.101)
Teacher understands Measure A – "Strongly Agree"/"Agree"		0.136***		0.023
		(0.037)		(0.029)
Teacher understands Measure B – "Strongly Agree"/"Agree"		0.186***		0.088**
		(0.036)		(0.029)
Question order effect				0.061**
				(0.023)
Component I is an accurate measure of practice			0.307**	0.028
			(0.106)	(0.027)
Component II is an accurate measure of practice			-0.051	0.040
			(0.106)	(0.027)
Component III is an accurate measure of practice			0.099	0.120**
			(0.150)	(0.037)
Component IV is an accurate measure of practice			0.269**	-0.036
			(0.100)	(0.026)
Component V is an accurate measure of practice			1.009***	0.220***
			(0.091)	(0.027)
Hours of Instruction				-0.003*
				(0.001)

Table 1C (Continued): Factors that Influence Views of DPAS-II for Administrators (on a 4.0 GPA scale), A Multilevel Regression

	Model1 coef/se	Model2 coef/se	Model3 coef/se	Model4 coef/se
Teacher's school is a good place to work – "Strongly Agree"/"Agree"				0.146***
				(0.029)
Student's family background determines achievement – "Strongly Agree"/"Agree"				-0.008
				(0.023)
Teacher can help all students achieve academic growth – "Strongly Agree"/"Agree"				0.084***
				(0.023)
DPAS-II is a positive influence on school culture				0.772***
				(0.063)
DPAS-II is a mixed influence on school culture				0.423***
				(0.027)
Don't know type of influence DPAS-II has on school culture				0.515***
				(0.039)
Sex of Respondent: Female				0.038
				(0.029)
Total Years of Experience				-0.008***
				(0.001)
School Poverty Measure: Students eligible for Free and Reduced-Price Lunch				-0.001
				(0.000)
Teaches English or Language Arts				-0.056*
				(0.028)
Teaches Math				-0.021
				(0.027)
Elementary Teacher				0.095**
				(0.029)

Table 1C (Continued): Factors that Influence views of DPAS-II for Administrators (on a 4.0 GPA scale), A Multilevel Regression

	Model1 coef/se	Model2 coef/se	Model3 coef/se	Model4 coef/se
_cons	0.872***	0.749***	-2.262***	0.218
	(0.035)	(0.121)	(0.142)	(0.112)
Number of observations	4,331	4,068	4,027	3,543
Log-Likelihood	-4,817.26	-5,413.81	-1,813.13	-3,600.78
bic	9,701.515	10,894.112	3,684.363	7,454.908

note: *** p<.001, ** p<.01, * p<.05

note: Standard Error in Parentheses

Table 2C: Factors that Influence Teachers' Views of fairness of DPAS-II, A Multilevel Logistic Regression

	Model1F coef/se	Model2F coef/se	Model3F coef/se	Model4F coef/se	Model5F coef/se
DPAS-II improves practice – "Often"/"Somewhat"	2.214***				1.567***
	(0.150)				(0.179)
DPAS-II improves practice – "Slightly"	0.871***				0.561**
	(0.155)				(0.180)
DPAS-II informs Professional Development – "Often"/"Somewhat"	1.123***				0.844***
	(0.136)				(0.164)
DPAS-II informs Professional Development – "Slightly"	0.459**				0.382*
	(0.143)				(0.169)
Teacher understands DPAS-II evaluation system – "Very much"/"Somewhat"		2.288*			0.971
		(1.014)			(1.043)
Teacher understands DPAS-II evaluation system – "Slightly"		0.872			0.172
		(1.056)			(1.091)
Teacher understands Measure A – "Strongly Agree"/"Agree"		0.691***			0.386**
		(0.108)			(0.133)
Teacher understands Measure B – "Strongly Agree"/"Agree"		0.163			0.134
		(0.107)			(0.131)
Component I is an accurate measure of practice			0.307**		0.183
			(0.106)		(0.130)
Component II is an accurate measure of practice			-0.051		-0.008
			(0.106)		(0.131)
Component III is an accurate measure of practice			0.099		-0.007
			(0.150)		(0.190)
Component IV is an accurate measure of practice			0.269**		-0.026
			(0.100)		(0.124)
Component V is an accurate measure of practice			1.009***		0.485***
			(0.091)		(0.115)
DPAS-II is a positive influence on school culture				3.723***	2.105***
				(0.215)	(0.252)
DPAS-II is a mixed influence on school culture				1.759***	1.072***
				(0.145)	(0.168)

Table 2C (Continued): Factors that Influence Teachers' Views of fairness of DPAS-II, A Multilevel Logistic Regression

	Model1F coef/se	Model2F coef/se	Model3F coef/se	Model4F coef/se	Model5F coef/se
Don't know type of influence DPAS-II has on school culture				2.000***	1.532***
				(0.172)	(0.203)
Teacher's school is a good place to work – "Strongly Agree"/"Agree"				0.975***	0.637***
				(0.147)	(0.170)
Question order effect					-0.270*
					(0.108)
Hours of Instruction					0.004
					(0.005)
Student's family background determines achievement – "Strongly Agree"/"Agree"					-0.057
					(0.112)
Teacher can help all students achieve academic growth – "Strongly Agree"/"Agree"					0.434***
					(0.115)
Sex of Respondent: Female					-0.073
					(0.136)
Total Years of Experience					-0.017**
					(0.006)
School Poverty Measure: Students eligible for Free and Reduced-Price Lunch					-0.001
					(0.002)
Teaches English or Language Arts					-0.098
					(0.131)
Teaches Math					-0.149
					(0.127)
Elementary Teacher					-0.125
					(0.130)
_cons	-3.387***	-4.202***	-2.262***	-3.870***	-5.597***
	(0.139)	(1.013)	(0.142)	(0.183)	(1.092)
Number of observations	4,245	3,946	4,027	3,940	3,438
Log-Likelihood	-1,614.46	-1,789.64	-1,813.13	-1,595.67	-1,159.66
bic	3,279.045	3,628.970	3,684.363	3,241.018	2,555.456

note: *** p<.001, ** p<.01, * p<.05

note: Standard Error in Parentheses

Table 3C: Factors that Influence Teachers' Understanding of DPAS-II, A Multilevel Logistic Regression

	Model1U coef/se	Model2U coef/se	Model3U coef/se	Model4U coef/se
DPAS-II improves practice – "Often"/"Somewhat"	0.083***			0.061***
	(0.013)			(0.015)
DPAS-II improves practice – "Slightly"	0.051***			0.036**
	(0.011)			(0.012)
DPAS-II informs Professional Development – "Often"/"Somewhat"	0.075***			0.064***
	(0.013)			(0.014)
DPAS-II informs Professional Development – "Slightly"	0.044***			0.038**
	(0.011)			(0.012)
Grade given to DPAS-II		0.050***		
		(0.005)		
Teacher understands Measure A – "Strongly Agree"/"Agree"			0.067***	0.050***
			(0.012)	(0.012)
Teacher understands Measure B – "Strongly Agree"/"Agree"			0.047***	0.042***
			(0.012)	(0.012)
Question order effect				-0.036***
				(0.009)
Component I is an accurate measure of practice				0.024*
				(0.011)
Component II is an accurate measure of practice				-0.005
				(0.011)
Component III is an accurate measure of practice				0.066***
				(0.015)
Component IV is an accurate measure of practice				0.007
				(0.011)
Component V is an accurate measure of practice				-0.021
				(0.011)
Hours of Instruction				0.000
				(0.000)
Teacher's school is a good place to work – "Strongly Agree"/"Agree"				0.010
				(0.012)

Table 3C (Continued): Factors that Influence Teachers' Understanding of DPAS-II, A Multilevel Logistic Regression

	Model1U coef/se	Model2U coef/se	Model3U coef/se	Model4U coef/se
Student's family background determines achievement – "Strongly Agree"/"Agree"				-0.011
				(0.010)
Teacher can help all students achieve academic growth – "Strongly Agree"/"Agree"				-0.008
				(0.010)
DPAS-II is a positive influence on school culture				0.015
				(0.026)
DPAS-II is a mixed influence on school culture				-0.007
				(0.011)
Don't know type of influence DPAS-II has on school culture				-0.035*
				(0.016)
Sex of Respondent: Female				0.028*
				(0.012)
Total Years of Experience				0.002**
				(0.001)
School Poverty Measure: Students eligible for Free and Reduced-Price Lunch				-0.000*
				(0.000)
Teaches English or Language Arts				0.019
				(0.011)
Teaches Math				-0.005
				(0.011)
Elementary Teacher				0.007
				(0.012)
_cons	0.830***	0.832***	0.846***	0.731***
	(0.008)	(0.009)	(0.008)	(0.027)
Number of observations	4,410	4,394	4,146	3,563
Log-Likelihood	-706.59	-713.44	-831.00	-448.80
bic	1,480.320	1,468.813	1,711.982	1,134.763

note: *** p<.001, ** p<.01, * p<.05

note: Standard Error in Parentheses

Table 4C: Factors that Influence Fidelity of Implementation of DPAS-II, A Multilevel Logistic Regression

	Model1I coef/se	Model2I coef/se	Model3I coef/se	Model4I coef/se	Model5I coef/se
DPAS-II improves practice – "Often"/"Somewhat"	0.047***				0.012
	(0.011)				(0.013)
DPAS-II improves practice – "Slightly"	0.022*				0.011
	(0.009)				(0.010)
DPAS-II informs Professional Development – "Often"/"Somewhat"	0.070***				0.037**
	(0.010)				(0.011)
DPAS-II informs Professional Development – "Slightly"	0.033***				0.019
	(0.009)				(0.010)
Teacher understands DPAS-II evaluation system – "Very much"/"Somewhat"		0.051			0.028
		(0.038)			(0.041)
Teacher understands DPAS-II evaluation system – "Slightly"		0.007			0.001
		(0.041)			(0.043)
Teacher understands Measure A – "Strongly Agree"/"Agree"		0.044***			0.030**
		(0.009)			(0.010)
Teacher understands Measure B – "Strongly Agree"/"Agree"		0.002			-0.002
		(0.009)			(0.009)
Question order effect		0.011			0.007
		(0.007)			(0.008)
Component I is an accurate measure of practice			-0.011		-0.012
			(0.009)		(0.009)
Component II is an accurate measure of practice			-0.015		-0.017
			(0.009)		(0.009)
Component III is an accurate measure of practice			-0.005		0.001
			(0.012)		(0.013)
Component IV is an accurate measure of practice			0.038***		0.030***
			(0.008)		(0.009)
Component V is an accurate measure of practice			0.026**		0.005
			(0.008)		(0.009)

Table 4C (Continued): Factors that Influence Fidelity of Implementation of DPAS-II, A Multilevel Logistic Regression

	Model1l coef/se	Model2l coef/se	Model3l coef/se	Model4l coef/se	Model5l coef/se
Grade given to DPAS-II				0.041***	0.011
				(0.004)	(0.006)
Hours of Instruction					0.000
					(0.000)
Teacher's school is a good place to work – "Strongly Agree"/"Agree"					0.070***
					(0.010)
Student's family background determines achievement – "Strongly Agree"/"Agree"					0.031***
					(0.008)
Teacher can help all students achieve academic growth – "Strongly Agree"/"Agree"					-0.011
					(0.008)
DPAS-II is a positive influence on school culture					0.054**
					(0.019)
DPAS-II is a mixed influence on school culture					-0.032***
					(0.010)
Don't know type of influence DPAS-II has on school culture					0.003
					(0.013)
Sex of Respondent: Female					-0.067***
					(0.010)
Total Years of Experience					0.000
					(0.000)
School Poverty Measure: Students eligible for Free and Reduced-Price Lunch					0.000*
					(0.000)
Teaches English or Language Arts					0.002
					(0.009)
Teaches Math					-0.023**
					(0.009)
Elementary Teacher					-0.001
					(0.010)
_cons	0.674***	0.654***	0.731***	0.661***	0.629***
	(0.010)	(0.039)	(0.013)	(0.010)	(0.046)

Table 4C (Continued): Factors that Influence Fidelity of Implementation of DPAS-II, A Multilevel Logistic Regression

	Model1I coef/se	Model2I coef/se	Model3I coef/se	Model4I coef/se	Model5I coef/se
Number of observations	2,926	2,830	2,950	2,955	2,460
Log-Likelihood	633.37	570.11	584.32	628.34	634.91
bic	-1,202.884	-1,068.696	-1,096.739	-1,216.725	-1,019.958
note: *** p<.001, ** p<.01, * p<.05					
note: Standard Error in Parentheses					

Table 5C: Factors that Influence Teachers' Change in Practice Based on DPAS-II Input, A Multilevel Logistic Regression

	ModelC1 coef/se	ModelC2 coef/se	ModelC3 coef/se	ModelC4 coef/se	ModelC5 coef/se
Grade given to DPAS-II	0.550***				0.111
	(0.039)				(0.069)
Fidelity of Implementation Index		1.913***			1.523***
		(0.205)			(0.246)
DPAS-II improves practice – "Often"/"Somewhat"			1.175***		0.989***
			(0.111)		(0.153)
DPAS-II improves practice – "Slightly"			0.897***		0.717***
			(0.091)		(0.121)
DPAS-II informs Professional Development – "Often"/"Somewhat"			0.590***		0.614***
			(0.107)		(0.137)
DPAS-II informs Professional Development – "Slightly"			0.380***		0.421***
			(0.093)		(0.119)
Component I is an accurate measure of practice				-0.024	-0.032
				(0.082)	(0.111)
Component II is an accurate measure of practice				0.009	-0.066
				(0.083)	(0.111)
Component III is an accurate measure of practice				0.302**	0.224
				(0.110)	(0.153)
Component IV is an accurate measure of practice				0.369***	0.177
				(0.081)	(0.108)
Component V is an accurate measure of practice				0.320***	0.028
				(0.081)	(0.110)
Teacher understands DPAS-II evaluation system – "Very much"/"Somewhat"					-0.495**
					(0.182)
Teacher understands Measure A – "Strongly Agree"/"Agree"					-0.046
					(0.117)
Teacher understands Measure B – "Strongly Agree"/"Agree"					-0.026
					(0.115)
Question order effect					-0.057
					(0.093)
Hours of Instruction					-0.006
					(0.005)

Table 5C (Continued): Factors that Influence Teachers' Change in Practice Based on DPAS-II Input, A Multilevel Logistic Regression

	ModelC1 coef/se	ModelC2 coef/se	ModelC3 coef/se	ModelC4 coef/se	ModelC5 coef/se
Teacher's school is a good place to work – "Strongly Agree"/"Agree"					0.109
					(0.123)
Student's family background determines achievement – "Strongly Agree"/"Agree"					0.267**
					(0.096)
Teacher can help all students achieve academic growth – "Strongly Agree"/"Agree"					-0.048
					(0.096)
DPAS-II is a positive influence on school culture					-0.424
					(0.265)
DPAS-II is a mixed influence on school culture					-0.156
					(0.115)
Don't know type of influence DPAS-II has on school culture					-0.305
					(0.171)
Sex of Respondent: Female					-0.146
					(0.120)
Total Years of Experience					-0.044***
					(0.005)
School Poverty Measure: Students eligible for Free and Reduced-Price Lunch					0.002
					(0.002)
Teaches English or Language Arts					0.112
					(0.115)
Teaches Math					-0.077
					(0.110)
Elementary Teacher					0.082
					(0.118)
_cons	-0.500***	-0.959***	-0.580***	-0.131	-0.739*
	(0.076)	(0.158)	(0.073)	(0.104)	(0.342)
Number of observations	3,664	2,881	3,628	3,718	2,464
Log-Likelihood	-2,354.89	-1,868.51	-2,259.68	-2,463.49	-1,450.53
bic	4,734.402	3,760.908	4,568.536	4,984.527	3,135.343

note: *** p<.001, ** p<.01, * p<.05

note: Standard Error in Parentheses

Table 6C: Factors that Influence Teachers' Views of the School as a Good Place to Work, A Multilevel Logistic Regression

	ModelSG1	ModelSG2	ModelSG3	ModelSG4	ModelSG5
	coef/se	coef/se	coef/se	coef/se	coef/se
Grade given to DPAS-II	0.671***				0.406***
	(0.048)				(0.085)
Fidelity of Implementation Index		2.367***			2.189***
		(0.268)			(0.310)
DPAS-II improves practice – "Often"/"Somewhat"			0.865***		0.241
			(0.135)		(0.205)
DPAS-II improves practice – "Slightly"			0.316**		-0.192
			(0.104)		(0.152)
DPAS-II informs Professional Development – "Often"/"Somewhat"			0.779***		0.327
			(0.129)		(0.182)
DPAS-II informs Professional Development – "Slightly"			0.375***		0.043
			(0.106)		(0.150)
Component I is an accurate measure of practice				-0.059	-0.155
				(0.096)	(0.142)
Component II is an accurate measure of practice				0.043	0.059
				(0.098)	(0.142)
Component III is an accurate measure of practice				0.285*	0.212
				(0.123)	(0.188)
Component IV is an accurate measure of practice				0.235*	0.147
				(0.096)	(0.140)
Component V is an accurate measure of practice				0.338***	-0.020
				(0.099)	(0.147)
Teacher understands DPAS-II evaluation system – "Very much"/"Somewhat"					-0.015
					(0.216)

Table 6C (Continued): Factors that Influence Teachers' Views of the School as a Good Place to Work, A Multilevel Logistic Regression

	ModelSG1	ModelSG2	ModelSG3	ModelSG4	ModelSG5
	coef/se	coef/se	coef/se	coef/se	coef/se
Teacher understands Measure A – "Strongly Agree"/"Agree"					0.331*
					(0.151)
Teacher understands Measure B – "Strongly Agree"/"Agree"					0.099
					(0.148)
Question order effect					0.049
					(0.120)
Hours of Instruction					0.001
					(0.006)
Student's family background determines achievement – "Strongly Agree"/"Agree"					-0.179
					(0.123)
Teacher can help all students achieve academic growth – "Strongly Agree"/"Agree"					0.140
					(0.122)
Sex of Respondent: Female					0.150
					(0.153)
Total Years of Experience					0.004
					(0.007)
School Poverty Measure: Students eligible for Free and Reduced-Price Lunch					-0.011***
					(0.003)
Teaches English or Language Arts					0.119
					(0.147)
Teaches Math					-0.045
					(0.141)
Elementary Teacher					0.044
					(0.168)
_cons	0.495***	0.006	0.819***	1.047***	-0.475
	(0.107)	(0.204)	(0.102)	(0.131)	(0.435)

Table 6C (continued): Factors that Influence Teachers' Views of the School as a Good Place to Work, A Multilevel Logistic Regression

	ModelSG1	ModelSG2	ModelSG3	ModelSG4	ModelSG5
	coef/se	coef/se	coef/se	coef/se	coef/se
Number of observations	4,107	2,873	4,081	4,156	2,481
Log-Likelihood	-1,930.19	-1,268.46	-1,934.83	-2,041.83	-1,035.26
bic	3,885.341	2,560.810	3,919.554	4,141.992	2,273.739

note: *** p<.001, ** p<.01, * p<.05

note: Standard Error in Parentheses

Table 7C: Factors that Influence Teachers' Interest in Continuing to Teach, A Multilevel Logistic Regression

	ModelT1 coef/se	ModelT2 coef/se	ModelT3 coef/se	ModelT4 coef/se	ModelT5 coef/se
Grade given to DPAS-II	0.597***				0.324***
	(0.046)				(0.090)
Fidelity of Implementation Index		1.274***			0.216
		(0.250)			(0.313)
DPAS-II improves practice – “Often”/”Somewhat”			0.908***		-0.072
			(0.135)		(0.204)
DPAS-II improves practice – “Slightly”			0.366***		0.001
			(0.101)		(0.152)
DPAS-II informs Professional Development – “Often”/”Somewhat”			0.399**		0.022
			(0.126)		(0.183)
DPAS-II informs Professional Development – “Slightly”			0.143		-0.155
			(0.104)		(0.150)
Component I is an accurate measure of practice				0.125	0.095
				(0.094)	(0.140)
Component II is an accurate measure of practice				0.035	0.110
				(0.096)	(0.141)
Component III is an accurate measure of practice				0.104	0.016
				(0.122)	(0.189)
Component IV is an accurate measure of practice				0.306**	0.145
				(0.097)	(0.142)
Component V is an accurate measure of practice				0.439***	0.163
				(0.100)	(0.150)
Teacher understands DPAS-II evaluation system – “Very much”/”Somewhat”					0.212
					(0.211)
Teacher understands Measure A – “Strongly Agree”/”Agree”					-0.073
					(0.154)
Teacher understands Measure B – “Strongly Agree”/”Agree”					-0.117
					(0.152)

Table 7C (Continued): Factors that Influence Teachers' Interest in Continuing to Teach, A Multilevel Logistic Regression

	ModelT1 coef/se	ModelT2 coef/se	ModelT3 coef/se	ModelT4 coef/se	ModelT5 coef/se
Question order effect					-0.203
					(0.121)
Hours of Instruction					0.006
					(0.006)
Teacher's school is a good place to work – "Strongly Agree"/"Agree"					1.791***
					(0.139)
Student's family background determines achievement – "Strongly Agree"/"Agree"					-0.277*
					(0.122)
Teacher can help all students achieve academic growth – "Strongly Agree"/"Agree"					0.571***
					(0.122)
DPAS-II is a positive influence on school culture					0.868
					(0.504)
DPAS-II is a mixed influence on school culture					0.296*
					(0.142)
Don't know type of influence DPAS-II has on school culture					0.246
					(0.219)
Sex of Respondent: Female					0.023
					(0.153)
Total Years of Experience					-0.017*
					(0.007)
School Poverty Measure: Students eligible for Free and Reduced-Price Lunch					0.004
					(0.002)
Teaches English or Language Arts					-0.079
					(0.148)
Teaches Math					0.014
					(0.143)
Elementary Teacher					-0.251
					(0.151)
_cons	0.673***	0.703***	1.018***	1.095***	-0.894*
	(0.075)	(0.185)	(0.072)	(0.111)	(0.416)

Table 7C (Continued): Factors that Influence Teachers' Interest in Continuing to Teach, A Multilevel Logistic Regression

	ModelT1 coef/se	ModelT2 coef/se	ModelT3 coef/se	ModelT4 coef/se	ModelT5 coef/se
Number of observations	4,098	2,866	4,074	4,148	2,462
Log-Likelihood	-1,883.89	-1,298.46	-1,898.20	-1,978.27	-961.51
bic	3,792.744	2,620.793	3,846.284	4,014.846	2,157.277

note: *** p<.001, ** p<.01, * p<.05

note: Standard Error in Parentheses

Table 8C: Descriptive Statistics for Teachers

Variable	Obs	Mean	Std. Dev	Min	Max
Grade given to DPAS-II	4,413	1.532064	0.964078	0	4
DPAS-II is fair and equitable – "Strongly Agree"/"Agree"	4,310	0.183527	0.387143	0	1
Teacher understands DPAS-II evaluation system – "Very much"/"Somewhat"	4,471	0.910981	.2848017	0	1
Fidelity of Implementation Index	3,138	0.736735	0.201426	0.125	1
Teacher changed practice based on feedback from DPAS-II	4,181	0.570199	0.495107	0	1
Teacher's school is a good place to work – "Strongly Agree"/"Agree"	4,170	0.77482	0.417751	0	1
Would like to continue working as a Teacher – "Strongly Agree"/"Agree"	4,162	0.811149	0.391438	0	1
DPAS-II improves practice – "Often"/"Somewhat"	4,462	0.294263	0.455762	0	1
DPAS-II improves practice – "Slightly"	4,462	0.322053	0.467316	0	1
DPAS-II informs Professional Development – "Often"/"Somewhat"	4,454	0.314549	0.464388	0	1
DPAS-II informs Professional Development – "Slightly"	4,454	0.295914	0.456504	0	1
Teacher understands DPAS-II evaluation system – "Very much"/"Somewhat"	4,471	0.071796	0.258179	0	1
Teacher understands Measure A – "Strongly Agree"/"Agree"	4,162	0.46468	0.498811	0	1
3					
Teacher understands Measure B – "Strongly Agree"/"Agree"	4,150	0.562169	0.496180	0	1
Question order effect	9,160	0.501201	0.500026	0	1
Component I is an accurate measure of practice	4,304	0.660316	0.473657	0	1
Component II is an accurate measure of practice	4,304	0.657063	0.474746	0	1
Component III is an accurate measure of practice	4,304	0.854786	0.352357	0	1
Component IV is an accurate measure of practice	4,304	0.336896	0.472704	0	1

Table 8C: Descriptive Statistics for Teachers (continued)

Component V is an accurate measure of practice	4,304	0.295074	0.456129	0	1
Hours of Instruction	3,993	24.61182	10.04631	0	40
Teacher's school is a good place to work – "Strongly Agree"/"Agree"	4,170	0.77482	0.417751	0	1
Student's family background determines achievement – "Strongly Agree"/"Agree"	4,165	0.369988	0.482859	0	1
Teacher can help all students achieve academic growth – "Strongly Agree"/"Agree"	4,158	0.559404	0.496518	0	1
DPAS-II is a positive influence on school culture	4,134	0.044025	0.205176	0	1
DPAS-II is a mixed influence on school culture	4,134	0.476778	0.499521	0	1
Don't know type of influence DPAS-II has on school culture	4,134	0.128931	0.335164	0	1
Sex of Respondent: Female	9,149	0.761504	0.426187	0	1
Total Years of Experience	9,149	12.40835	9.107616	0	46
School Poverty Measure: Students eligible for Free and Reduced-Price Lunch	3,811	63.34532	30.12073	0	100
Teaches English or Language Arts	5,178	0.486288	0.49986	0	1
Teaches Math	5,178	0.472383	0.499285	0	1
Elementary Teacher	5,187	0.415269	0.492816	0	1

Table 9C: Factors that Influence Specialists Views of DPAS-II, Fairness of DPAS-II, Understanding of DPAS-II, and Fidelity of Implementation, Multilevel Linear and Logistic Regressions

	Dependent Variable: Views of GPA	Dependent Variable: Views of Fairness	Dependent Variable: Understanding of DPAS-II	Dependent Variable: Fidelity of Implementation
	ModelA	ModelB	ModelC	ModelD
	coef/se	coef/se	coef/se	coef/se
Specialist understands DPAS-II evaluation system – "Very much"/"Somewhat"	0.213*	1.386		0.045
	(0.085)	(0.929)		(0.038)
DPAS-II improves practice – "Often"/"Somewhat"	0.620***	1.543***	1.073	-0.009
	(0.093)	(0.453)	(0.618)	(0.032)
DPAS-II informs Professional Development – "Often"/"Somewhat"	0.423***	0.485	0.752	0.097**
	(0.090)	(0.398)	(0.531)	(0.032)
Question order effect	0.084	0.044	-0.673*	0.001
	(0.060)	(0.333)	(0.273)	(0.022)
Component I is an accurate measure of practice	-0.115	-0.275	0.571	-0.015
	(0.083)	(0.476)	(0.378)	(0.030)
Component II is an accurate measure of practice	0.239**	0.566	0.116	0.017
	(0.081)	(0.493)	(0.336)	(0.030)
Component III is an accurate measure of practice	0.013	-0.141	0.406	-0.002
	(0.077)	(0.434)	(0.342)	(0.027)
Component IV is an accurate measure of practice	0.016	-0.010	-0.236	0.034
	(0.073)	(0.393)	(0.332)	(0.026)
Component V is an accurate measure of practice	0.211*	0.478	-0.526	0.029
	(0.084)	(0.412)	(0.391)	(0.031)
Hours of Instruction	-0.002	0.000	0.005	0.001
	(0.003)	(0.014)	(0.012)	(0.001)
Specialist's school is a good place to work – "Strongly Agree"/"Agree"	0.154	1.299	0.306	-0.012
	(0.079)	(0.721)	(0.306)	(0.031)
Student's family background determines achievement – "Strongly Agree"/"Agree"	-0.079	0.284	-0.171	0.023
	(0.064)	(0.356)	(0.277)	(0.023)
Teacher/Spec/Admin can help all students achieve academic growth – "Strongly Agree"/"Agree"	0.252***	1.095**	0.722*	0.060*
	(0.068)	(0.356)	(0.343)	(0.024)

Table 9C (Continued): Factors that Influence Specialists Views of DPAS-II, Fairness of DPAS-II, Understanding of DPAS-II, and Fidelity of Implementation, Multilevel Linear and Logistic Regressions

	Dependent Variable: Views of GPA	Dependent Variable: Views of Fairness	Dependent Variable: Understanding of DPAS-II	Dependent Variable: Fidelity of Implementation
	ModelA coef/se	ModelB coef/se	ModelC coef/se	ModelD coef/se
DPAS-II is a positive influence on school culture	0.623***	2.434**	0.174	0.092
	(0.173)	(0.805)	(1.135)	(0.058)
DPAS-II is a mixed influence on school culture	0.550***	1.118*	0.729*	0.032
	(0.073)	(0.472)	(0.348)	(0.027)
Don't know type of influence DPAS-II has on school culture	0.406***	0.233	-0.190	0.018
	(0.087)	(0.616)	(0.338)	(0.033)
Sex of Respondent: Female	0.111	-0.247	0.395	-0.042
	(0.097)	(0.521)	(0.415)	(0.036)
Total Years of Experience	-0.010***	-0.003	0.018	-0.002
	(0.003)	(0.016)	(0.013)	(0.001)
School Poverty Measure: Students eligible for Free and Reduced-Price Lunch	0.000	0.005	-0.010*	0.000
	(0.001)	(0.006)	(0.005)	(0.000)
_cons	0.334	-6.833***	0.909	0.608***
	(0.176)	(1.584)	(0.681)	(0.067)
Number of observations	497	469	500	327
Log-Likelihood	-497.40	-131.94	-186.82	72.40
bic	1,137.606	393.048	497.929	-11.639

note: *** p<.001, ** p<.01, * p<.05

note: Standard Error in Parentheses

Table 10C: Factors that Influence Changes in Practice, Views of the Workplace, and Views of Continuing to Work as an Educator for Specialists DPAS-II for, A Multilevel Logistic Regression

	Dependent Variable: Change in Practice	Dependent Variable: Views of the Workplace	Dependent Variable: Desire to Continue Working as an Administrator
	model1 coef/se	model2 coef/se	model3 coef/se
Grade given to DPAS-II	0.455*	0.711*	0.293
	(0.211)	(0.306)	(0.247)
Fidelity of Implementation Index	3.956***	-0.421	0.807
	(0.787)	(1.029)	(0.814)
Specialists understands DPAS-II evaluation system – “Strongly Agree”/”Agree”	-0.103	0.870	-0.207
	(0.508)	(0.619)	(0.510)
DPAS-II improves practice – “Often”/”Somewhat”	1.146**	-0.647	-0.362
	(0.409)	(0.793)	(0.600)
DPAS-II informs Professional Development – “Often”/”Somewhat”	0.213	1.529	0.560
	(0.374)	(0.850)	(0.628)
Question order effect	0.233	-0.147	-0.326
	(0.276)	(0.426)	(0.348)
Component I is an accurate measure of practice	-0.231	0.113	-0.622
	(0.361)	(0.589)	(0.491)
Component II is an accurate measure of practice	0.880*	0.644	0.232
	(0.400)	(0.609)	(0.486)
Component III is an accurate measure of practice	0.513	-0.266	0.347
	(0.338)	(0.526)	(0.424)
Component IV is an accurate measure of practice	-0.316	-0.608	0.227
	(0.320)	(0.525)	(0.425)
Component V is an accurate measure of practice	-0.131	-0.147	-0.573
	(0.378)	(0.630)	(0.553)
Specialist's school is a good place to work – “Strongly Agree”/”Agree”	0.120		2.055***
	(0.394)		(0.379)

Table 10C (Continued): Factors that Influence Changes in Practice, Views of the Workplace, and Views of Continuing to Work as an Educator for Specialists DPAS-II for, A Multilevel Logistic Regression

	Dependent Variable: Change in Practice	Dependent Variable: Views of the Workplace	Dependent Variable:Desire to Continue Working as an Administrator
	model1	model2	model3
	coef/se	coef/se	coef/se
Specialist can help all students achieve academic growth – "Strongly Agree"/"Agree"	0.185	1.273*	2.176***
	(0.302)	(0.573)	(0.578)
Student's family background determines achievement – "Strongly Agree"/"Agree"	0.188	0.078	-0.341
	(0.290)	(0.457)	(0.348)
DPAS-II is a positive influence on school culture	-1.791*		-0.232
	(0.741)		(1.219)
DPAS-II is a mixed influence on school culture	-0.390		0.421
	(0.358)		(0.434)
Don't know type of influence DPAS-II has on school culture	-0.264		-0.018
	(0.419)		(0.478)
Sex of Respondent: Female	0.518	1.102	0.258
	(0.466)	(0.692)	(0.618)
Total Years of Experience	-0.015	0.030	-0.019
	(0.013)	(0.022)	(0.017)
School Poverty Measure: Students eligible for Free and Reduced-Price Lunch	-0.005	-0.008	-0.006
	(0.005)	(0.009)	(0.006)
_cons	-4.936***	-0.181	-0.629
	(1.123)	(1.317)	(1.067)
/lnsig2u	-1.450	1.272*	-9.658
	(1.633)	(0.605)	(17.125)
Number of observations	356	358	353
Log-Likelihood	-184.44	-133.96	-118.76
bic	498.123	373.773	366.590

note: *** p<.001, ** p<.01, * p<.05

note: Standard Error in Parentheses

Table 11C: Descriptive Statistics for Specialists

Variable	Obs	Mean	Std. Dev	Min	Max
Grade given to DPAS-II	724	1.36326	0.938984	0	4
DPAS-II is fair and equitable – "Strongly Agree"/"Agree"	705	0.141844	0.349138	0	1
Specialist understands DPAS-II evaluation system – "Very much"/"Somewhat"	748	0.846257	0.360944	0	1
Specialist's school is a good place to work – "Strongly Agree"/"Agree"	666	0.833333	0.372958	0	1
Specialist changed practice based on feedback from DPAS-II	669	0.337818	0.47332	0	1
Would like to continue working as a Specialist – "Strongly Agree"/"Agree"	665	0.806015	0.395715	0	1
Fidelity of Implementation Index	468	0.713408	0.208307	0.13	1
DPAS-II improves practice – "Often"/"Somewhat"	745	0.218792	0.413705	0	1
DPAS-II informs Professional Development – "Often"/"Somewhat"	747	0.24498	0.430364	0	1
Question order effect	1,560	0.498077	0.500157	0	1
Component I is an accurate measure of practice	689	0.449927	0.497848	0	1
Component II is an accurate measure of practice	689	0.676343	0.468211	0	1
Component III is an accurate measure of practice	689	0.555878	0.497229	0	1
Component IV is an accurate measure of practice	689	0.554427	0.49739	0	1
Component V is an accurate measure of practice	689	0.182874	0.386844	0	1
Hours of Instruction	580	10.51724	12.47317	0	40
Student's family background determines achievement – "Strongly Agree"/"Agree"	666	0.33033	0.470686	0	1
Teacher/Spec/Admin can help all students achieve academic growth – "Strongly Agree"/"Agree"	658	0.346505	0.476218	0	1
DPAS-II is a positive influence on school culture	663	0.031674	0.175264	0	1
DPAS-II is a mixed influence on school culture	663	0.374057	0.484244	0	1
Don't know type of influence DPAS-II has on school culture	663	0.197587	0.398479	0	1
Sex of Respondent: Female	1,559	0.868506	0.338049	0	1
Total Years of Experience	1,559	16.70815	10.27773	0	43
School Poverty Measure: Students eligible for Free and Reduced-Price Lunch	561	65.40463	28.6877	0	100

Table 12C: Factors that Influence views of DPAS-II for Administrators (on a 4.0 GPA scale), A Multilevel Regression

	Model0G coef/se	Model1G coef/se	Model2G coef/se	Model3G coef/se	Model4G coef/se
Administrator: Principal versus All Others	0.055	-0.010	0.050	0.002	-0.001
	(0.114)	(0.089)	(0.117)	(0.108)	(0.089)
DPAS-II improves practice – "Often"/"Somewhat"		1.003***			0.787***
		(0.113)			(0.121)
DPAS-II informs Professional Development – "Often"/"Somewhat"		0.459***			0.375***
		(0.113)			(0.116)
Administrator understands DPAS-II evaluation system – "Strongly Agree"/"Agree"					0.427**
					(0.167)
Component I is an accurate measure of practice			0.228		0.135
			(0.133)		(0.099)
Component II is an accurate measure of practice			-0.117		-0.147
			(0.147)		(0.117)
Component III is an accurate measure of practice			0.159		-0.015
			(0.132)		(0.101)
Component IV is an accurate measure of practice			-0.047		0.057
			(0.123)		(0.093)
Component V is an accurate measure of practice			0.438***		0.034
			(0.119)		(0.094)
Administrator's school is a good place to work – "Strongly Agree"/"Agree"				0.194	-0.049
				(0.152)	(0.127)
DPAS-II is a positive influence on school culture				1.588***	1.026***
				(0.203)	(0.178)
DPAS-II is a mixed influence on school culture				0.936***	0.657***
				(0.172)	(0.147)
Don't know type of influence DPAS-II has on school culture				0.665***	0.560***
				(0.178)	(0.149)
Student's family background determines achievement – "Strongly Agree"/"Agree"					0.070
					(0.116)

Table 12C (Continued): Factors that Influence views of DPAS-II for Administrators (on a 4.0 GPA scale),
A Multilevel Regression

	Model0G	Model1G	Model2G	Model3G	Model4G
	coef/se	coef/se	coef/se	coef/se	coef/se
Total Years of Experience		-0.007	-0.002	-0.005	-0.006
		(0.005)	(0.007)	(0.006)	(0.005)
Sex of Respondent: Female		0.036	-0.041	-0.055	0.014
		(0.087)	(0.115)	(0.106)	(0.088)
Question order effect					0.010
					(0.086)
_cons	1.805***	1.368***	1.522***	0.911***	0.486**
	(0.073)	(0.113)	(0.177)	(0.213)	(0.236)
Number of observations	316	302	291	279	265
Log-Likelihood	-441.75	-334.61	-396.40	-353.39	-272.37
bic	906.525	714.910	855.215	763.091	656.342

note: *** p<.01, * p<.01, ** p<.05
note: Standard Error in Parentheses

Table 13C: Factors that Influence Administrators Views of DPAS-II as Fair and Equitable, A Multilevel Logistic Regression

	Model0 coef/se	Model1 coef/se	Model2 coef/se	Model3 coef/se	Model4 coef/se
Administrator: Principal versus All Others	0.443**	0.485	0.376	0.344	0.198
	(0.238)	(0.309)	(0.269)	(0.283)	(0.370)
DPAS-II improves practice – "Often"/"Somewhat"		2.812***			2.743***
		(0.405)			(0.508)
DPAS-II informs Professional Development – "Often"/"Somewhat"		0.033			-0.337
		(0.389)			(0.491)
Administrator understands DPAS-II evaluation system – "Strongly Agree"/"Agree"					0.447
					(0.874)
Component I is an accurate measure of practice			0.840***		0.994**
			(0.318)		(0.435)
Component II is an accurate measure of practice			0.324		0.325
			(0.364)		(0.503)
Component III is an accurate measure of practice			0.139		-0.027
			(0.315)		(0.427)
Component IV is an accurate measure of practice			-0.158		0.016
			(0.286)		(0.378)
Component V is an accurate measure of practice			0.733***		0.298
			(0.272)		(0.377)
Administrator's school is a good place to work – "Strongly Agree"/"Agree"				0.560	-0.053
				(0.454)	(0.569)
DPAS-II is a positive influence on school culture				2.816***	2.181***
				(0.618)	(0.789)
DPAS-II is a mixed influence on school culture				1.309**	1.064
				(0.535)	(0.673)
Don't know type of influence DPAS-II has on school culture				0.832	0.613
				(0.560)	(0.697)
Student's family background determines achievement – "Strongly Agree"/"Agree"					0.491

Table 13C (Continued): Factors that Influence Administrators Views of DPAS-II as Fair and Equitable, A Multilevel Logistic Regression

	Model0 coef/se	Model1 coef/se	Model2 coef/se	Model3 coef/se	Model4 coef/se
Total Years of Experience		-0.013	0.002	0.002	-0.015
		(0.020)	(0.017)	(0.017)	(0.483)
Sex of Respondent: Female		-0.405	-0.345	-0.391	-0.521
		(0.306)	(0.268)	(0.282)	(0.367)
Question order effect					0.218
					(0.361)
_cons	-0.548***	-1.370***	-1.565***	-2.099***	-3.620***
	(0.146)	(0.421)	(0.450)	(0.676)	(1.203)
Number of observations	316	301	279	267	253
Log-Likelihood	-211.57	-148.08	-172.76	-156.90	-110.81
bic	440.408	336.111	401.834	364.084	326.755

note: *** p<.01, ** p<.05, * p<.1

note: Standard Error in Parentheses

Table 14C: Factors that Influence Administrators Understanding of DPAS-II, A Multilevel Logistic Regression

	Model0U coef/se	Model1U coef/se	Model2U coef/se	Model3U coef/se	Model4F coef/se
Administrator: Principal versus All Others	0.835**	0.638	0.482	0.734	0.198
	(0.476)	(0.509)	(0.544)	(0.556)	(0.370)
DPAS-II improves practice – "Often"/"Somewhat"		2.732**			2.743***
		(1.096)			(0.508)
DPAS-II informs Professional Development – "Often"/"Somewhat"		0.365			-0.337
		(0.640)			(0.491)
Administrator understands DPAS-II evaluation system – "Strongly Agree"/"Agree"					0.447
					(0.874)
Component I is an accurate measure of practice			0.002		0.994**
			(0.516)		(0.435)
Component II is an accurate measure of practice			1.007**		0.325
			(0.547)		(0.503)
Component III is an accurate measure of practice			0.733		-0.027
			(0.508)		(0.427)
Component IV is an accurate measure of practice			-0.038		0.016
			(0.518)		(0.378)
Component V is an accurate measure of practice			0.006		0.298
			(0.512)		(0.377)
Administrator's school is a good place to work – "Strongly Agree"/"Agree"				1.441***	-0.053
				(0.526)	(0.569)
DPAS-II is a positive influence on school culture					2.181***
					(0.789)
DPAS-II is a mixed influence on school culture				0.585	1.064
				(0.679)	(0.673)
Don't know type of influence DPAS-II has on school culture				-0.128	0.613
				(0.640)	(0.697)
Student's family background determines achievement – "Strongly Agree"/"Agree"					0.491
					(0.483)

Table 14C (Continued): Factors that Influence Administrators Understanding of DPAS-II, A Multilevel Logistic Regression

	Model0U coef/se	Model1U coef/se	Model2U coef/se	Model3U coef/se	Model4F coef/se
Total Years of Experience		0.016	0.017	0.014	0.015
		(0.027)	(0.027)	(0.029)	(0.023)
Sex of Respondent: Female		-0.063	-0.250	0.056	-0.521
		(0.449)	(0.484)	(0.497)	(0.367)
Question order effect					0.218
					(0.361)
_cons	2.097***	1.458***	1.234**	0.622	-3.620***
	(0.253)	(0.560)	(0.646)	(0.792)	(1.203)
Number of observations	335	319	293	233	253
Log-Likelihood	-99.04	-80.80	-71.87	-64.66	-110.81
bic	215.514	201.950	200.550	172.937	326.755

note: *** p<.01, ** p<.05, * p<.1

note: Standard Error in Parentheses

Table 15C: Factors that Influence Fidelity of Implementation of DPAS-II for Administrators (on a 0 to 1 FOI scale), A Multilevel Regression

	Model0I coef/se	Model1I coef/se	Model2I coef/se	Model3I coef/se	Model4I coef/se
Administrator: Principal versus All Others	0.079***	0.080***	0.084***	0.065**	0.061**
	(0.026)	(0.027)	(0.026)	(0.026)	(0.027)
DPAS-II improves practice – "Often"/"Somewhat"		0.069**			0.042
		(0.035)			(0.039)
DPAS-II informs Professional Development – "Often"/"Somewhat"		-0.017			-0.039
		(0.035)			(0.036)
Grade given to DPAS-II			0.032**		0.019
			(0.013)		(0.019)
Administrator understands DPAS-II evaluation system – "Strongly Agree"/"Agree"					0.163**
					(0.075)
Component I is an accurate measure of practice					0.038
					(0.031)
Component II is an accurate measure of practice					-0.035
					(0.036)
Component III is an accurate measure of practice					0.018
					(0.031)
Component IV is an accurate measure of practice					-0.006
					(0.029)
Component V is an accurate measure of practice					0.020
					(0.028)
Administrator's school is a good place to work – "Strongly Agree"/"Agree"				0.083**	0.065
				(0.040)	(0.041)
DPAS-II is a positive influence on school culture				0.019	-0.034
				(0.050)	(0.059)
DPAS-II is a mixed influence on school culture				-0.024	-0.055
				(0.043)	(0.049)
Don't know type of influence DPAS-II has on school culture				-0.097**	-0.123**
				(0.046)	(0.051)

Table 15C (Continued): Factors that Influence Fidelity of Implementation of DPAS-II for Administrators (on a 0 to 1 FOI scale), A Multilevel Regression

	Model0I coef/se	Model1I coef/se	Model2I coef/se	Model3I coef/se	Model4I coef/se
Student's family background determines achievement – "Strongly Agree"/"Agree"					0.043
					(0.037)
Total Years of Experience		-0.001	-0.000	-0.001	-0.001
		(0.002)	(0.002)	(0.002)	(0.002)
Sex of Respondent: Female		-0.046	-0.047	-0.032	-0.032
		(0.027)	(0.026)	(0.026)	(0.027)
Question order effect					-0.003
					(0.027)
_cons	0.721***	0.741***	0.693***	0.729***	0.556***
	(0.018)	(0.037)	(0.045)	(0.058)	(0.097)
Number of observations	251	231	236	223	211
Log-Likelihood	44.56	43.60	47.71	53.02	53.65
bic	-67.015	-43.669	-57.164	-51.975	5.092

note: *** p<.01, * p<.01, ** p<.05

note: Standard Error in Parentheses

Table 16C: Factors that Influence Administrators' Change in Practice Due to DPAS-II Feedback, A Multilevel Logistic Regression

	CHANGE1	CHANGE2	CHANGE3	CHANGE4
	coef/se	coef/se	coef/se	coef/se
Administrator: Principal versus All Others	0.283	-0.136	-0.209	-0.466
	(0.261)	(0.294)	(0.299)	(0.347)
Fidelity of Implementation Index		3.818***	3.766***	4.413***
		(0.846)	(0.865)	(1.040)
DPAS-II improves practice – "Often"/"Somewhat"			0.523**	0.479
			(0.294)	(0.476)
Grade given to DPAS-II				0.258
				(0.240)
Administrator understands DPAS-II evaluation system – "Strongly Agree"/"Agree"				0.014
				(1.053)
DPAS-II informs Professional Development – "Often"/"Somewhat"				-0.174
				(0.438)
Component I is an accurate measure of practice				0.300
				(0.398)
Component II is an accurate measure of practice				0.118
				(0.459)
Component III is an accurate measure of practice				-0.139
				(0.385)
Component IV is an accurate measure of practice				0.309
				(0.354)
Component V is an accurate measure of practice				0.176
				(0.353)
Administrator's school is a good place to work – "Strongly Agree"/"Agree"				-0.957**
				(0.555)
Student's family background determines achievement – "Strongly Agree"/"Agree"				-0.509
				(0.476)
DPAS-II is a positive influence on school culture				-0.386
				(0.764)
DPAS-II is a mixed influence on school culture				-0.049
				(0.645)
Don't know type of influence DPAS-II has on school culture				-0.469
				(0.681)
Sex of Respondent: Female				0.285
				(0.345)
Total Years of Experience				0.017
				(0.021)
Question order effect				-0.079
				(0.345)

Table 16C (Continued): Factors that Influence Administrators' Change in Practice Due to DPAS-II Feedback, A Multilevel Logistic Regression

	CHANGE1	CHANGE2	CHANGE3	CHANGE4
	coef/se	coef/se	coef/se	coef/se
_cons	-0.514***	-3.223***	-3.351***	-3.921***
	(0.193)	(0.681)	(0.697)	(1.398)
Number of observations	259	222	219	206
Log-Likelihood	-174.59	-138.68	-134.96	-118.97
bic	365.846	298.980	296.867	349.828

note: *** p<.01, ** p<.05, * p<.1

note: Standard Error in Parentheses

Table 17C: Factors that Influence Administrators' Views District/School as a Good Place to Work, A Multilevel Logistic Regression

	SCHOOL1	SCHOOL2	SCHOOL3	SCHOOL4
	coef/se	coef/se	coef/se	coef/se
Administrator: Principal versus All Others	-0.233	-0.577	-0.543	-0.202
	(0.349)	(0.455)	(0.472)	(0.540)
Fidelity of Implementation Index		3.110***	2.772**	1.970
		(1.055)	(1.100)	(1.206)
DPAS-II improves practice – "Often"/"Somewhat"			1.439***	1.148
			(0.556)	(0.816)
Grade given to DPAS-II				0.042
				(0.338)
Administrator understands DPAS-II evaluation system – "Strongly Agree"/"Agree"				0.864
				(1.054)
DPAS-II informs Professional Development – "Often"/"Somewhat"				0.542
				(0.757)
Component I is an accurate measure of practice				0.069
				(0.597)
Component II is an accurate measure of practice				-0.369
				(0.666)
Component III is an accurate measure of practice				0.880
				(0.593)
Component IV is an accurate measure of practice				-0.519
				(0.581)
Component V is an accurate measure of practice				0.744
				(0.556)
Student's family background determines achievement – "Strongly Agree"/"Agree"				0.207
				(0.747)
Sex of Respondent: Female				-0.647
				(0.549)
Total Years of Experience				-0.031
				(0.030)
Question order effect				0.211
				(0.530)
_cons	2.202***	0.424	0.259	0.239
	(0.363)	(0.738)	(0.756)	(1.409)
Number of observations	288	226	223	213
Log-Likelihood	-118.34	-79.67	-75.33	-66.92
bic	253.668	181.017	177.694	224.976

note: *** p<.01, * p<.01, ** p<.05
note: Standard Error in Parentheses

Table 18C: Factors that Influence Administrators Interest in Continuing to Work as Administrators, A Multilevel Logistic Regression

	TEACH1 coef/se	TEACH2 coef/se	TEACH3 coef/se	TEACH4 coef/se
Administrator: Principal versus All Others	-0.167	0.236	0.184	1.114
	(0.373)	(0.440)	(0.450)	(0.750)
Fidelity of Implementation Index		-0.214	-0.374	-1.202
		(1.051)	(1.086)	(1.750)
DPAS-II improves practice – "Often"/"Somewhat"			0.492	-1.808
			(0.464)	(1.035)
Grade given to DPAS-II				1.280**
				(0.508)
Administrator understands DPAS-II evaluation system – "Strongly Agree"/"Agree"				(dropped)
DPAS-II informs Professional Development – "Often"/"Somewhat"				-0.258
				(0.926)
Component I is an accurate measure of practice				-0.216
				(0.838)
Component II is an accurate measure of practice				-0.621
				(1.041)
Component III is an accurate measure of practice				0.404
				(0.843)
Component IV is an accurate measure of practice				-0.019
				(0.777)
Component V is an accurate measure of practice				0.172
				(0.713)
Administrator's school is a good place to work – "Strongly Agree"/"Agree"				3.920***
				(0.914)
Student's family background determines achievement – "Strongly Agree"/"Agree"				1.306
				(1.285)

Table 18C (Continued): Factors that Influence Administrators Interest in Continuing to Work as Administrators, A Multilevel Logistic Regression

	TEACH1 coef/se	TEACH2 coef/se	TEACH3 coef/se	TEACH4 coef/se
DPAS-II is a positive influence on school culture				2.117
				(1.403)
DPAS-II is a mixed influence on school culture				1.517
				(0.964)
Don't know type of influence DPAS-II has on school culture				2.378**
				(1.084)
Sex of Respondent: Female				-0.334
				(0.708)
Total Years of Experience				-0.012
				(0.043)
Question order effect				0.443
				(0.745)
_cons	2.093***	2.210***	2.238***	-2.586
	(0.282)	(0.821)	(0.840)	(2.062)
Number of observations	289	227	224	204
Log-Likelihood	-104.58	-81.84	-78.81	-40.94
Bic	226.154	185.371	184.684	188.249

note: *** p<.01, * p<.01, ** p<.05

note: Standard Error in Parentheses

Table 19C: Descriptive Statistics from the Administrator Survey

Variable	Obs	Mean	Std. Dev	Min	Max
Grade given to DPAS-II	316	1.825949	0.981459	0	4
DPAS-II is fair and equitable – "Strongly Agree"/"Agree"	316	0.405063	0.491683	0	1
Administrator changed practice based on feedback from DPAS-II	286	0.395105	0.48973	0	1
Administrator's school is a good place to work – "Strongly Agree"/"Agree"	288	0.84375	0.363724	0	1
Administrator understands DPAS-II evaluation system – "Very much"/"Somewhat"	335	0.910448	0.285966	0	1
Fidelity of Implementation Index	251	0.756474	0.20681	0.125	1
Wishes to continue working as Administrator – "Strongly Agree"/"Agree"	289	0.882353	0.322749	0	1
Administrator: Principal versus All Others	691	0.296672	0.457121	0	1
DPAS-II improves practice – "Often"/"Somewhat"	329	0.379939	0.486111	0	1
DPAS-II informs Professional Development – "Often"/"Somewhat"	327	0.388379	0.488129	0	1
Question order effect	691	0.500724	0.500362	0	1
Component I is an accurate measure of practice	304	0.648026	0.478373	0	1
Component II is an accurate measure of practice	304	0.733553	0.44283	0	1
Component III is an accurate measure of practice	304	0.657895	0.475197	0	1
Component IV is an accurate measure of practice	304	0.460526	0.499261	0	1
Component V is an accurate measure of practice	304	0.450658	0.49838	0	1
Student's family background determines achievement – "Strongly Agree"/"Agree"	292	0.174658	0.380326	0	1
Administrator's school is a good place to work – "Strongly Agree"/"Agree"	288	0.84375	0.363724	0	1
DPAS-II is a positive influence on school culture	288	0.166667	0.373327	0	1
DPAS-II is a mixed influence on school culture	288	0.385417	0.487541	0	1
DPAS-II is a negative influence on school culture	288	0.125	0.331295	0	1
Don't know type of influence DPAS-II has on school culture	288	0.322917	0.468405	0	1
Total Years of Experience	689	16.22642	8.663332	0	47
Sex of Respondent: Female	689	0.561684	0.496541	0	1

Appendix D: Teacher Artifacts as Evidence of Validity and Implementation Fidelity of DPAS-II

Pete Goldschmidt, Ph.D.
National Center for the Improvement of Educational Assessment

We use artifacts collected from teachers teaching in four schools to develop validity evidence and to evaluate the fidelity of implementation of specific components of DPAS-II. The analyses are based on artifacts from teachers who were recruited and volunteered to provide the data. Given this is not a simple random sample of teachers, results and inferences derived from results should be viewed as speculative. This exploratory and descriptive analysis is based on a small sample (N=42 teachers) and we are therefore limited in the analyses we can conduct. Generally, we collected artifacts from teachers who are directly related to their evaluation as well as an independent criterion demonstrated to be related to teacher's instructional practices – teacher assignments (Data Matsumura, Garnier, Slater, and Boston, 2008; Matsumura and Pascal, 2003; Clare Valdes, Pascal, and Steinberg, 2001). We are interested in identifying exemplars for each of the three artifacts as well as examining the relationships among them and the teacher's official observation rating. We develop several scoring rubrics that we use to identify exemplary practices as well as to examine the relationship among the measures. We focus on three instruments: one, an indicator of assignment quality, based on previous work by Matsumura, et. al., (2003); two, we develop an indicator of feedback quality based on Nicol, and Macfarlane-Dick (2004) and Thurlings et., al., (2012); and three, we develop an indicator of quality for teacher growth plans based on Wallings, et., al., (2013). We used the relevant literature to develop scoring rubrics for each construct. We also collect teacher's official principal observation ratings.

While teacher evaluation has existed in some form in most states for some time, the results often suffered from lack of face validity, were considered superficial (Stiggens and Duke, 1988), and of insufficient quality to yield reliable indicators of teacher performance – in that 99% of teachers were deemed effective in virtually every state, despite results in many states indicating that 50% of the students are not on grade level and significant gaps in achievement continued to exist among various subgroups.

Limited evidence suggests that evaluation systems can have a positive impact on student outcomes (Taylor and Tyler, 2012; Kimball, et., al., 2008 Milonowski, 2004). The literature indicates that a key component is the fidelity with which principals can manage an evaluation system and provide meaningful guidance to teachers (Milonowski, 2004). Despite rhetoric to the contrary, most state systems rely heavily on subjective indicators of quality, rather than on indicators of student learning. This further increases the importance of implementation fidelity – not only in accurately identifying and classifying teachers into performance categories, but also in guiding teachers that need improvement towards meaningful mechanisms to improve. Overall, early results from state-wide systems have returned results similar to those under the old systems - with 97% to 99% of teachers being rated as effective – including 99% in Delaware. Few states have attributed increased student performance to their teacher evaluation system (although Tennessee is a notable exception).

The contribution of teachers to student learning (i.e. value added models –VAM) has received considerable scrutiny over the past 10 years (Ladd, 2008; Lockwood and McCaffrey. 2007; McCaffrey, 2004). Observations have received some attention in terms of the rubric itself, or the

amount of time it may take to adequately observe a teacher, but the fidelity with which principals and teachers implement specific portions of the evaluation system has not been fully examined. The fact that 99% of teachers are classified as effective does not diminish the need for concrete, actionable, feedback for continued improvement.

While teacher evaluation systems consist of multiple measures, some, such as VAM results can merely provide corroborating evidence as to the effectiveness of instruction, while others will be more closely linked to instructional practices. Two important elements to consider are the feedback principals provide and the specific professional growth goals teachers ascribe to. These two elements directly impact what and how teachers will adjust their practices. Professional growth plans have received little attention – although elements that make plans meaningful have been proposed (Walling, Shapiro, and Ast, 2013). Formative feedback has received a significant amount of attention; although substantively less is related to changes in performance[3] (Bjorn, Wurth, and Hergovich, 2013; Kluger and DeNisi, 1996). If a teacher evaluation system has potential to improve instructional practices, an important mechanism for improvement will be principal feedback.

Methodology

We applied both quantitative and qualitative strategies to evaluating teacher artifacts related to the DPAS-II. We detail the rubrics and coding strategies within the presentation of results for each type of artifact collected.

We examine artifacts and surveys for their technical amenability (e.g. score properties, reliability) for use in analytical models before utilizing results in subsequent evaluation models. Analyses of artifacts consist of three connected elements. It is important to note that artifacts and observations are attempting to provide evidence of teacher pedagogical content knowledge (PCK) (Schulman, 1986; Phelps and Schilling, 2004), which is a combination of knowing what to teach and how to teach it. PCK should be aligned with the evidence sought in Components I – IV (Heneman and Milanowski, 2004; Kimball and Milanowski, 2009). Teacher created artifacts can provide evidence of instructional fidelity (Matsumura, Garnier, Slater, and Boston, 2008; Matsumura and Pascal, 2003; Clare Valdes, Pascal, and Steinberg, 2001) which is related is to evidence associated with teachers' instructional quality and PCK. Artifacts might consist of many representations; however, we focus on those that have previously been demonstrated to impact instruction and student academic outcomes. Given that a primary interest is evidence of implementation when DPAS-II is done well and identification of best practices, it is important to establish what "well" implies – in this case well takes on two dimensions: DPAS-II results can be validated with external criteria; and two, exemplars are associated with high quality instruction.

We utilize an abbreviated form of artifact analysis rubric developed by Clare, Valdes, Pascal, and Steinberg (2001) for some artifacts. This includes asking teachers to provide a brief cover sheet to submitted artifacts. We present the details of the elements of the assessment quality indicator in the results section pertaining to teacher assessments. We collected multiple types of artifacts and attempt to ascertain whether assessments are suitable for large scale use in validating teacher performance indicators. Teacher assignments have been demonstrated to be reliable indicators from which valid inferences about instruction can be drawn (Matsumura and Pascal, 2003) and are amenable for use across a wide range of content (Silver, 2009, Borko, Stecher, 2007). Inferences

[3] Performance as differentiated specifically from learning – as in Black and William, 1998).

based on assignments are related to observations of teachers (Matsumura and Pascal, 2003), which is important in this context because it helps address the seemingly skewed principal observation results. Although, given the small sample size, results are less skewed, but are highly centered without much variation. The assignment scores allow us to address important goals of the evaluation. It will afford us an opportunity to examine the extent to which principal observation results are related to an external criteria, and whether this relationship varies among districts – which potentially provides evidence of where DPAS-II was done well. Analyses using assignment scores are described in detail below.

We develop additional indicators for artifacts. This includes an indicator of feedback quality based on qualities that good feedback ought to exhibit (Thurlings, Vermeulen, Kreijins, Bastiaens, and Stijnen, 2012; Meyer, 1991; Heneman and Milanowski, 2004; Kimball and Milanowski, 2009). Details of the Feedback Quality Indicator (FQI) we develop are presented in the section results section pertaining to principal[4] feedback.

We also develop a professional growth goal indicator. Although reflections and growth plans are common components of educator effectiveness systems, little research has examined what might constitute a "good" plan. We develop an indicator based on the relevant recent literature (Walling, Shapiro, and Ast, 2013; Nicol and MacFarlane-Dick, 2004). Specific criteria for this indicator are detailed in the results section pertaining to Component IV.

Results

We received artifacts from 42 teachers[5]. These 42 provided artifacts related to the written feedback they received from principals, their Component IV, part 5, professional growth plans, and class assignments (including instructions, grading rubrics, and student work samples). Not all teachers provided all information.

Table 1:

Response Rates for Artifacts

Artifact	Responses
Assignment Information	23
Component IV PGG	41
Principal Feedback	37
Complete Data	18

[4] Although administrators (as well as teachers) can conduct observations, we refer to observers as principals given the majority of raters are principals.

[5] The scope of work indicated that we would conduct the artifact analysis on 25 teachers; however, given the uneven distribution of responses we felt it beneficial to evaluate the data at hand. This provided a larger sample than planned for two of three analyses and a smaller sample for one analysis. Also, the combined complete sample size is less than 25.

The results in Table one indicate that, although 42 teachers responded, there was a considerable amount of missing data. This is an important caveat to bear in mind because we have insufficient responses to examine the pattern of missingness. Hence, we are limited by both a small sample and the potential of both non-response and reponse bias. Given that the emphasis on developing a proof of concept, this is less problematic than it might be if high stakes were attached to the evaluations of artifacts. Despite the limitations, several interesting patterns emerge and warrant continued examination or monitoring. These findings will be addressed in turn.

Principal Feedback

As noted in Table 1, 37 teachers provided their written feedback from principals. We rated principal feedback using an index consisting of 10 items. Table 2 presents the 10 items used. Given the prospective nature of the analysis, we scored each item solely on the presence or absence of the construct identified in the item. Additional iterations might develop a scoring rubric that better differentiates feedback on each construct. Overall the Feedback Quality Instrument (FQI) appeared to work well. The sample size was insufficient to fully examine the psychometric properties of the instrument; however casual analysis indicates that the full 10 item FQI has a Cronbach's of .6, while a nine item version (FQI2) has a reliability of .77.

Table 2:

Elements of the Feedback Quality Instrument

Question	Feedback
Q1	Directed at the task/practice/goal and not the teacher
Q2	Language aligns with rating
Q3	Based solely on observation
Q4	Refers to specific events in classroom
Q5	Focuses on actions not justifications for actions
Q6	Addresses actionable behavior and provide concrete recommendation for improvement
Q7	Separates good from bad.
Q8	Compares/connects between actual and desired outcome
Q9	Is corrective - issues are brought up with corrections
Q10	Provides constructive criticism rather than affirmations and encouragement

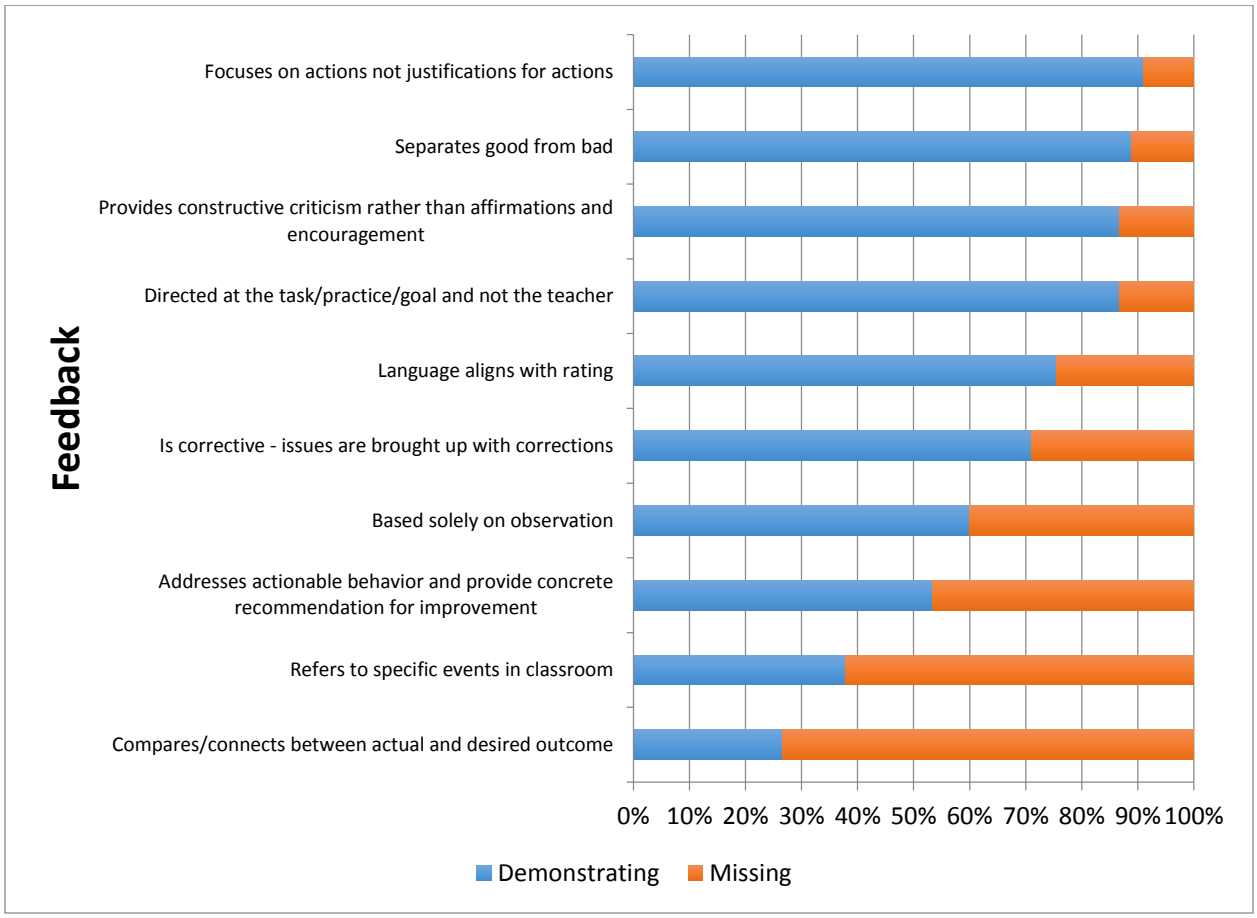
The difference between FQI and FQI2 is whether feedback made reference to post-observation meeting conversations. It is unclear whether the feedback should or should not include discussion. On the one hand the feedback should be the motivator for discussion, while on the other hand, feedback including discussion can provide a record of guidance that a teacher can refer back to. Another consideration is whether the construct of the FFT instrument, including feedback, is altered if post classroom observation information is taken into account in forming ratings and

feedback. This report is agnostic and presents findings related to both constructs. However, the DDOE may want to solidify, if it has not already done so, its expectations regarding this practice[6].

Figure 1 presents the distribution of performance on each of the FQI items. The items are ordered from "easiest" to "hardest;" that is, elements of the FQI that were most readily observed in the feedback are at the top of the figure. For example, we find that most principals were able to provide feedback that focused on actions and not the justification for actions.

On the other hand, principals had a much more difficult time comparing/connecting actual and desired behavior. The results in Figure 1 clearly indicate that there are attributes to high quality feedback that vary in the propensity of their appearance on written teacher feedback.

Figure 1: Performance on Principal Feedback Components



The individual items provide an interesting summary of principal feedback, but do not allow for generalizations that might be helpful in identifying how principals might benefit from professional development. One approach would be to collapse the items presented above into meaningful sub-domains of principal feedback.

Although the sample size would generally be considered too small to conduct factor analysis, guidance in the literature tends to vary with recommendations focusing on either an absolute

[6] From a generalizability standpoint, including any post-observation information introduces another source variation into the resulting scores and feedback.

minimum N or a subject to variable ratio. Minimum N's as low as 40 and subject to variable ratios as low as 2:1 have been utilized in the literature. While there is no specific cut-off, the robustness of results depends to a large extent on the empirical results (Zhao, 2009).

We applied principal component exploratory factor analysis and found that the 10 items presented in Table 2 behaved quite well in forming a two factor solution. We consider these results preliminary[7]. The variance explained for FQI is consistent with it reliability, about .57. The two factor solution is quite informative. The two factor solution represents two domains: one, instructional coach and; and two, human resource manager. The instructional coach factor focuses on specific observed classroom practices, areas for improvement, and specific recommendations for improvement. The human resources manager factor focuses on communication – the feedback language is aligned to the rating, comments are directed at teachers, and feedback clearly delineates strengths from weaknesses. Figure 2 summarizes how well principals provide feedback related to providing guidance as an instructional coach and a human resource manager.

Figure 2: Principal Feedback by Domain

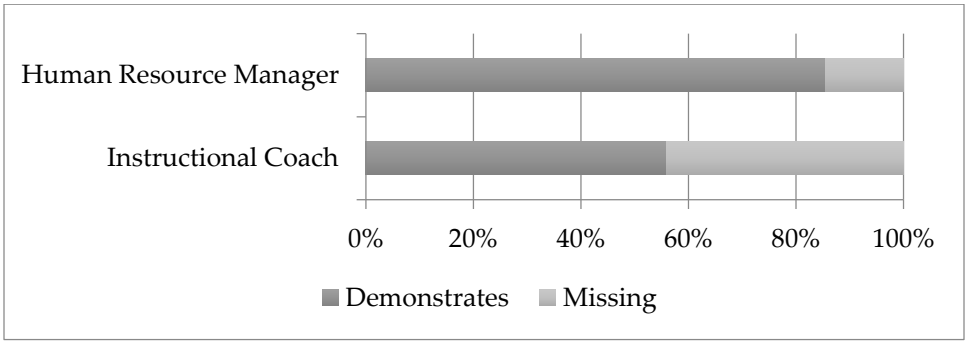


Figure 2 highlights that principals are able to, for the most part, provide feedback that is communicated well, in terms of aligning to the rating and focusing on actions as opposed to the person. However, principals seem to be less able to consistently provide specific feedback that provides concrete examples from the classroom linked to areas of improvement that are aligned to desired outcomes and specific recommendations as how to achieve the desired outcomes.

In order to solidify the concepts presented above, we provide specific examples from the feedback forms. Although principals generally do well in communication, and despite language aligning with ratings about 75%, there are several examples of misalignment.

> "[T] did not get to cover what she wanted in the lesson because time ran short. It has been recommended that she use a visual timer..." Teacher scored Distinguished in 2a. maximize learning time.
>
> "T has shown the capacity to meet each challenge with grace, poise, and the ability to continually perfect her teaching practice. Well done!" Teacher scored straight Proficient.
>
> "It is recommended that you use formative assessment to gauge student progress...you did not directly assess understanding of the text prior...additionally,

[7] The empirical evidence suggests that the results appear well behaved, particularly with no cross-loaded factors and each of the two factors having at least 2 items with loadings over .8.

> you did not present a summarizing task...recommended that you devote the majority of your instructional time to content-related learning tasks..." Teacher scored straight Proficient.

In terms of feedback related to being an instructional coach, principals had a significantly more difficult time providing concrete guidance. For example in terms of constructive criticism rather than affirmations and encouragement:

> "It is recommended that you continue this program with fidelity." (Provides no constructive criticism and focuses on affirmation.)

While some principals do provide a straight-forward example of meeting this criterion:

> "…in order to move to a distinguished level have students plan to ask 1-2 questions after they compared their markings…"

Although occurring less than 40% of the time, some feedback did refer to specific events in classroom, such as:

> "As we discussed in the post-conference, you not only gave students recall questions to answer as they read, but you told them exactly where to find the answers. When giving students an important text to read, determine your purpose first and then provide an appropriate graphic organizer and/or require the use of an effective reading strategy that promotes deeper understanding of the text."

Principals can also present concrete issues with corrective actions, as in:

> "Prepare to move the lesson along when/if students are able to grasp concepts more quickly than anticipated. Students appeared to quickly understand the significance of a PSA and the components of an effective PSA. More time can be spent on student production of their PSA ..."

It is also possible for feedback to addresses actionable behavior and provide concrete recommendations for improvement, for example:

> "…make the lesson more relevant to students and their lives. We talked about providing them choices when picking foods…"

While occurring about a quarter of the time in written feedback, comparison/connection between actual and desired outcome was present in some feedback, for example:

> "…have the students to share their data for finding right angles instead of her sharing that information. That would have given students who did not finish the activity [opportunity] to complete the task as well."

We examined both characteristics of the teachers and of the observation to determine whether there were any systematic relationships with FQI scores. Overall, the average teacher rating on the

observation protocol does not vary by any of the characteristics for which we have indicators. Hence, we focus on the quality of the feedback, and whether either of the domain scores vary. For example, principal feedback was not qualitatively different for whether the observation was announced or unannounced. Overall feedback quality does not differ between novice and experienced teachers; however, specific feedback related to instructional practice is of significantly lower quality for novice teachers ($p < .05$).

Importantly there is evidence that the number of elements of the FFT[8] scored relates to both teacher overall ratings[9] and the quality of feedback they receive. Teachers scored on fewer elements of the FFT tend to have lower overall ratings ($p< .05$). Also, the instructional coach subdomains is positively related to the number of elements scored ($p < .01$) as is the overall indicator of feedback quality ($p < .01$). There is suggestive evidence that the human resource manager subdomain is positively related to the number of elements scored ($p < .10$). Consistent with expectations, there is suggestive evidence that feedback quality is inversely related to teacher overall ratings ($p < .10$)[10].

Table 3 summarizes the variability of the quality of feedback. In other words, table three provides some evidence as to whether there are statistically significant differences in feedback among evaluators and schools[11]. There is suggestive evidence that the quality of feedback, particularly instructional feedback quality varies by evaluator. The results suggest that there tend to be systematic differences among schools in teacher ratings as well as the quality of the feedback. It is important to note that differences among schools represent not only differences among the principals in providing feedback, but also among other staff that provide feedback a within a school.[12]

Table 3:

Variation in Principal Feedback

	Evaluator	School
Overall Teacher Rating	no	$p < .10$
Instructional Coach	$p < .10$	$p < .05$
Human Resource Mgmt.	no	no
FQI	no	$P < .01$
FQI2	$p < .10$	$P < .01$

Overall, the evaluation of written principal feedback provides meaningful results despite the limited sample size. Much of the potential benefit lies in teachers receiving meaningful constructive feedback that can concretely direct them towards improvement. The results imply that principals are having difficult time providing specific recommendations linked to desired outcomes. The evaluation of feedback also indicates that feedback appears to be of lower quality specifically for those teachers who likely need the most direction. Novice teachers are receiving lower quality instructional coaching. The evidence also suggests that feedback is linked to how much principals observe.

[8] Framework for Teaching.
[9] We use the average score on the 4 point rubric and not the classifications used on the DPAS-II summative reports.
[10] This relationship is significant when using FQI2 ($p < .05$).
[11] Results are based on 11 evaluators and 4 schools.
[12] There is insufficient data to fully examine the structure of the relationships (e.g. the between rater, within school variability).

Component IV, Part 5 Professional Growth Goals Results

We next examine teachers' self-described professional growth goals. We note that, overall, this form is not being completed nor evaluated with fidelity. For example, there is no place for evaluator comments on the form. We recommend that teachers are provided more concrete guidance for this form and that the form focus on part 5, with goals addressing the criteria we present below.

We evaluated teachers' written plan using a set of items considered to meaningfully describe aspects of quality related to goal setting. Each teacher's goal was scored on

Table 4:

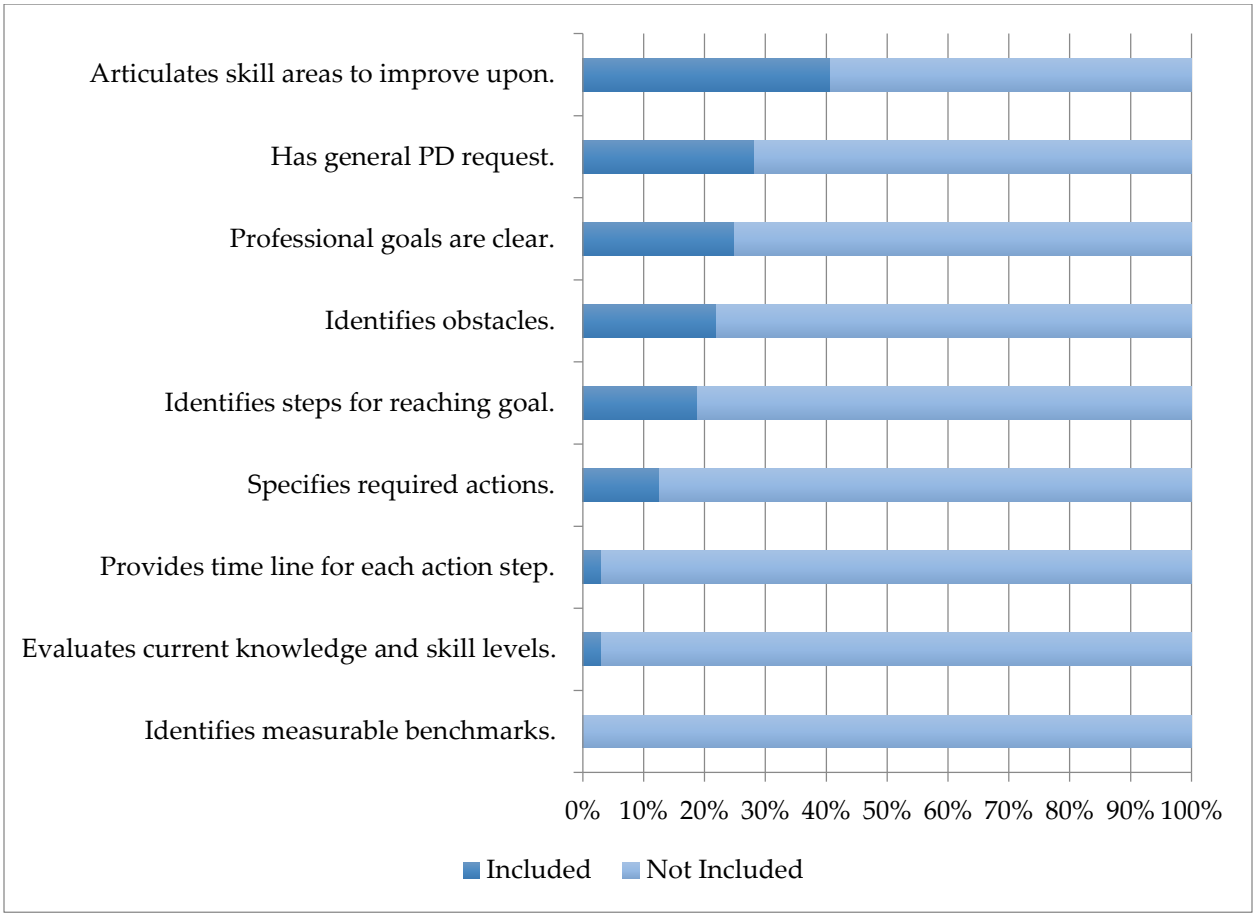
Elements of the Growth Goal Evaluation Instrument

Item	Dimension
Q1	Articulates skill areas to improve upon.
Q2	Professional goals are clear.
Q3	Evaluates current knowledge and skill levels.
Q4	Identifies steps for reaching goal.
Q5	Specifies required actions.
Q6	Identifies obstacles.
Q7	Identifies measurable benchmarks.
Q8	Provides time line for each action step.
Q9	Has general PD request.

the nine items listed in table four. While teachers could demonstrate different degrees to which stated plans might address each of the quality elements, given the limited sample size and the exploratory nature of this construct[13] we simply coded for the presence of absence of the element. In this way we are able to determine the extent to which teachers, without guidance, are able to develop a plan that minimally addresses elements that a high quality growth plan should include. Figure 3 summarizes teacher goals on each of the criteria identified in table four.

[13] We also note that teachers were not provided guidance related to the criteria listed in table 4, so the degree to which a teacher addressed these with fidelity is less amenable to analysis than simply identifying the presence or absence of the element.

Figure 3: Growth Goal Criteria Included in Plans



The results indicate that teachers' plans only loosely develop professional plans. Less than 50% of the plans identify a specific skill area to improve upon. In terms of specific actions required to meet goals, less than 20% of plans specify actions and no plan identifies a measurable benchmark that would provide evidence that the goal has been met. Again, it is important to note that the criteria were not provided to guide teachers in developing a plan; hence, these results provide evidence more for the need to solidify component with more concrete direction than it is evidence of teachers' inability to develop coherent growth goals.

Overall, the average growth plan scored about a 1.2 out of a possible of 9 points. These low scores impact the reliability of the instrument because the modal score is 0. Additional research would be required to determine whether the instrument is incapable of identifying the distribution of quality in growth plans, or whether growth plans are not developed with fidelity across the state.

While unequivocal claims about plans would be unjustified, substantive evidence does indicate that the plans are not completed with fidelity. For example, one articulated goal was to "be the best teacher I can…" while another plan indicated that the goals are to …"continue to learn as a teacher… and attend workshops…".

On the other hand there are examples of plans meeting specific criteria. Table 5 presents representative samples from growth plans. Table 5 also indicates what proportion of plans met the criteria. In some instances there are few exemplars from which to choose. Also, providing

measurable benchmarks was not indicated on any plan consistently--although in some instances teachers would indicate an objective such as a degree, which to some extent provides a benchmark automatically; however, given that this goal was generally part of a set of goals it was not explicitly scored as meeting this criteria.[14]

[14] This is an instance where a more sophisticated scoring rubric would be beneficial (e.g. no benchmark =0, some =1, and all =2).

Table 5:

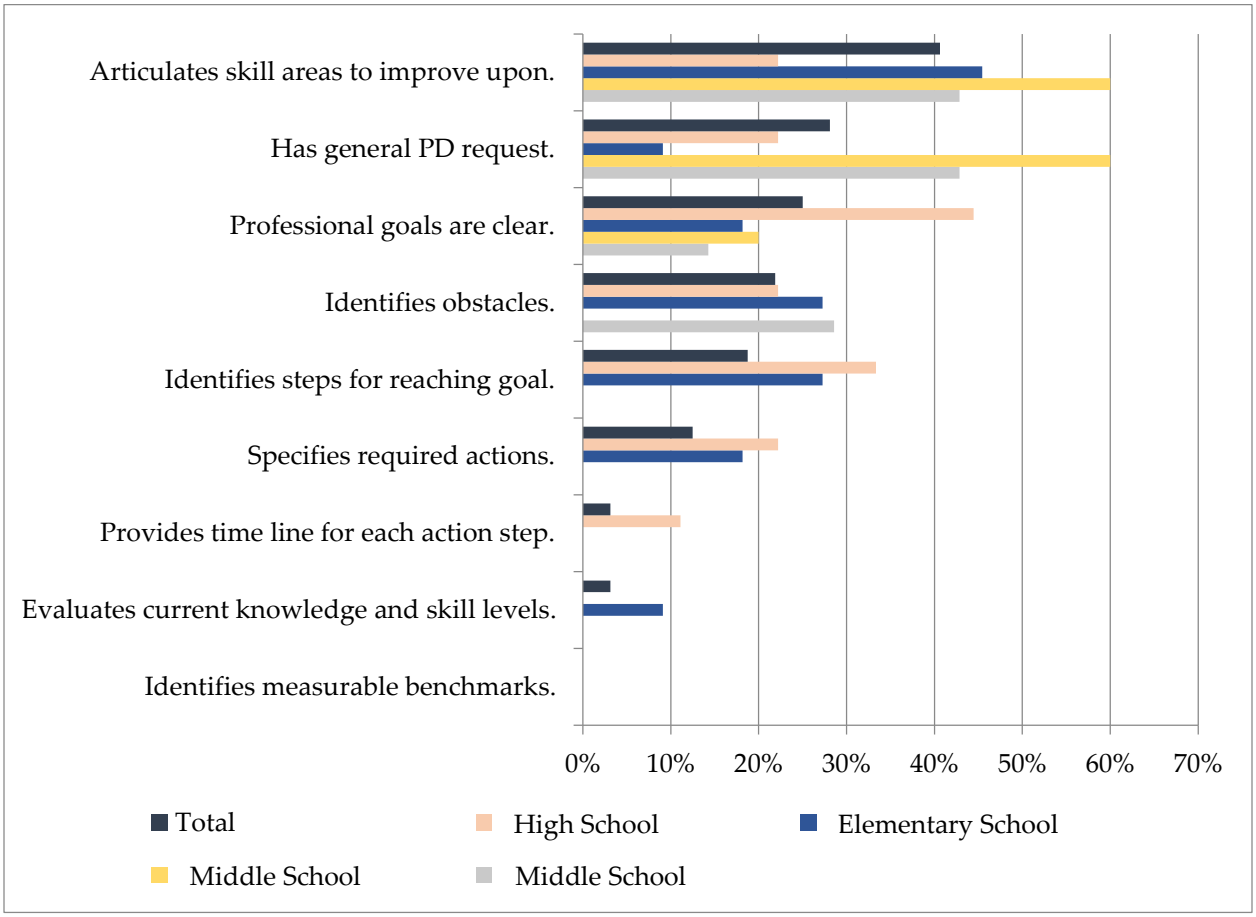
Examples of Goals Meeting Criteria

Criteria	Percent Demonstrating
Articulates skill areas to improve upon.	41%
"I would love to participate in professional development which will help me improve upon differentiating assignments within a SAM (single approach to mastery) classroom."	
Has general PD request.	28%
"I would like to attend more content specific professional development workshops to shape my instruction to fit the needs of my students."	
Professional goals are clear.	25%
"Go back to school and receive a certificate in Educational Technology."	
Identifies obstacles.	22%
"My ELL period 8 class is my main struggle this year to help them overcome the language barrier and be successful with 9th grade math concepts."	
Identifies steps for reaching goal.	19%
"My goal this year is to focus on the RTQ Problem solving process. I will be working with my plc teams to analyze student work and identify areas of concern. We will de designing a plan for each student and implementing interventions."	
Specifies required actions.	13%
"I want to master small group instruction and differentiation. I am attending small group instruction trainings to help me towards this goal."	
Provides time line for each action step.	3%
"… complete my requirements for my Masters before December. The paper I am writing involves investigating the fairness of school funding in Delaware. Help with school funding data would be helpful"	
Evaluates current knowledge and skill levels.	3%
"I can personally see myself struggling when it comes to instructing reading. I am intimidated by reading due to the fact that kindergarten students have little or no phonemic awareness/phonics skills before entering kindergarten."	
Identifies measurable benchmarks.	0%

Despite the limited range in overall teacher ratings and despite the limited range of scores on the growth goal instrument, there is evidence that more effective teachers write stronger growth goals

as evidenced by a positive correlation between overall teacher ratings and growth goal scores ($p < .05$). Likely consistent with expectations is that principal feedback is inversely related to growth goal scores ($p < .05$). This indicates that teachers writing better plans (who tend to be more highly rated teachers) receive lower quality feedback. It is also interesting to note that the strength of goals does not vary systematically among the schools in the sample, but that two schools in the sample have means that are two to three times higher than the other two schools in the sample. The difference between the two pairs of schools is significant ($p < .05$). Again we note that these results are not based on simple random samples and that inferences based on statistical tests should be considered with caution.

Figure 4: Growth Goals Compared Across Schools



The results in Figure 4 appear to support the notion that, although there is insufficient evidence to support the notion that growth goals vary systematically among schools, in general there are suggestive patterns to the criteria. For example, while the majority of plans at one school articulate a skill area to improve upon, no plans at this school meet six of the criteria. A similar pattern exists for another school in a different district. Additional investigation can examine whether teachers in these schools were provided specific direction or guidance (that coincidentally met some of the criteria applied in this evaluation. The other two schools seem to demonstrate more variability among the plans—with less concentration on particular aspects, but broader coverage. Together, the results suggests that there may be differences among the schools in how they approach growth plans and that developing guidance and policy can impact how teachers address this task. While

there is a positive relationship between overall teacher rating and the quality of plans, there are no significant correlations within schools. The point estimates of the correlations vary from .23 to .45, but there is insufficient sample size within school detect relationships. Again, the results provide suggestive evidence for school-wide differences in approaching Component IV.

Overall, the evaluation of component for growth goals reveals that the exercise would benefit from additional guidance to teachers. A particular feature that is in line with student growth goals is for teachers to provide measureable benchmarks that provide evidence of meeting goals.

Teacher Assignment Results

Student Assignments were collected from 25 teachers[15], which limits results to general findings and likely does not fully point to the potential for assignments to provide a legitimate validation criterion. One benefit of pursing assignments for evaluation purposes is the low stakes nature of the assignments.

Consistent with the aforementioned literature, we examined student assignments on three related domains: cognitive challenge; clarity; and, learning goals. We rate the overall quality of the assignment based on the elements in table six. Using a composite of the indicators in table six results in a reliability estimate of .79 for the overall assignment quality measure. Although each indicator is equally weighted, given that the first indicator ranges from 1 to 4 (because it assess Web's Depth of Knowledge (DOK), it is naturally weighted more heavily.

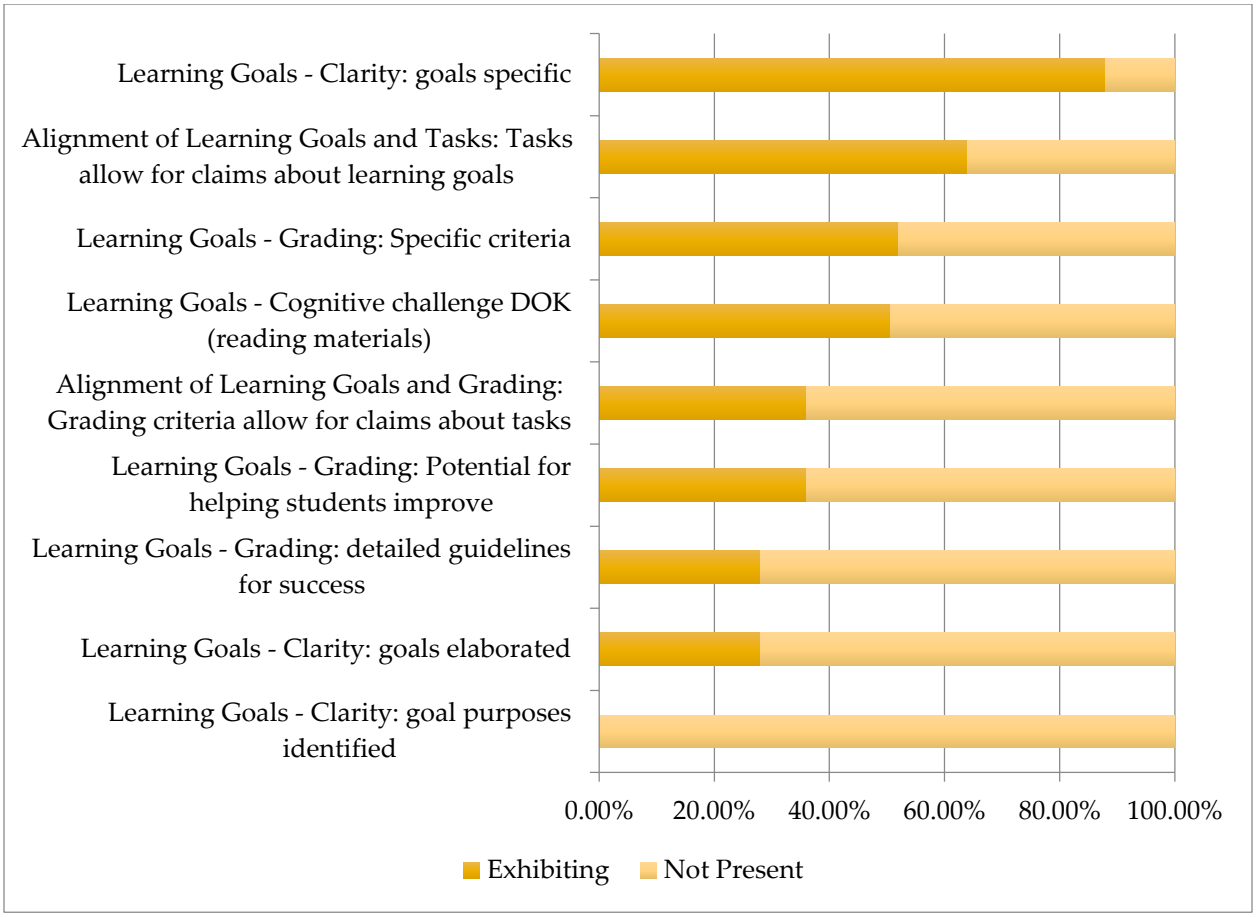
Table 6:

Indicators of Assignment Quality

Learning Goals - Cognitive challenge DOK (reading materials).
Learning Goals - Clarity: goals specific.
Learning Goals - Clarity: goals elaborated.
Learning Goals - Clarity: goal purposes identified.
Learning Goals - Grading: Specific criteria.
Learning Goals - Grading: detailed guidelines for success.
Learning Goals - Grading: Potential for helping students improve.
Alignment of Learning Goals and Tasks: Tasks allow for claims about learning goals.
Alignment of Learning Goals and Grading: Grading criteria allow for claims about tasks.

[15] Not all teachers had complete responses.

Figure 5: Indicators of Assessment Quality



Results in Figure 5 indicate that assignments varied in the degree to which indicators were met. For example, over 80% of assignments clearly articulated specific learning goals for assignments. Figure 5 indicates that the mean DOK was about 50% (out of a total of 4 – indicating that the mean DOK of assignments was about 2). We measured DOK in three ways on the assignment rubric – teacher constructed response, teacher selected responses of tasks, and teacher indicators of preparatory reading materials and associated goals.

Specifically we asked teachers: Learning Goals – What were the specific goals you expected the students to demonstrate through this assignment? We evaluated scored each response for DOK.We also asked teachers: What were the students asked to do in this assignment? Check up to 4 primary tasks ordered from primary (1) to ancillary (4)? We provided teachers with a list of tasks (see appendix D1).Finally, we asked teachers: How were the students to use the reading material for this assignment? We provided teachers with a list of uses (see appendix B).

The correlations among the three indicators are modest (r ~ .34 to .46) and the means are relatively close – ranging from 2.5 to 2.0. We elected to use the latter indicator as it both the most conservative and most aligned with the holistic indicator of cognitive challenge (described below). Needless to say, with such a small sample size, additional analyses are certainly warranted. The assessment quality index ranges from a potential score of 1 to 12, with a mean of 5.3 and a standard

deviation of 2.4. The actual assignment quality index scores range from a low of 2 to a high of 10. There is suggestive evidence that assessment quality varies among schools ($p < .10$)[16].

We independently and holistically scored the assignments on cognitive challenge, clarity, and grading. We calculated an overall quality score based on the three holistic scores. Three scores are developed – an equally weighted composite, a regression-based score, and a score based only on cognitive challenge and grading[17]. The correlations between the assignment quality index based on teacher responses and the assignment quality based on the various holistic scores range from .73 to .80, providing some criterion related evidence with respect to inferences about assignment quality based on the assignment quality index.

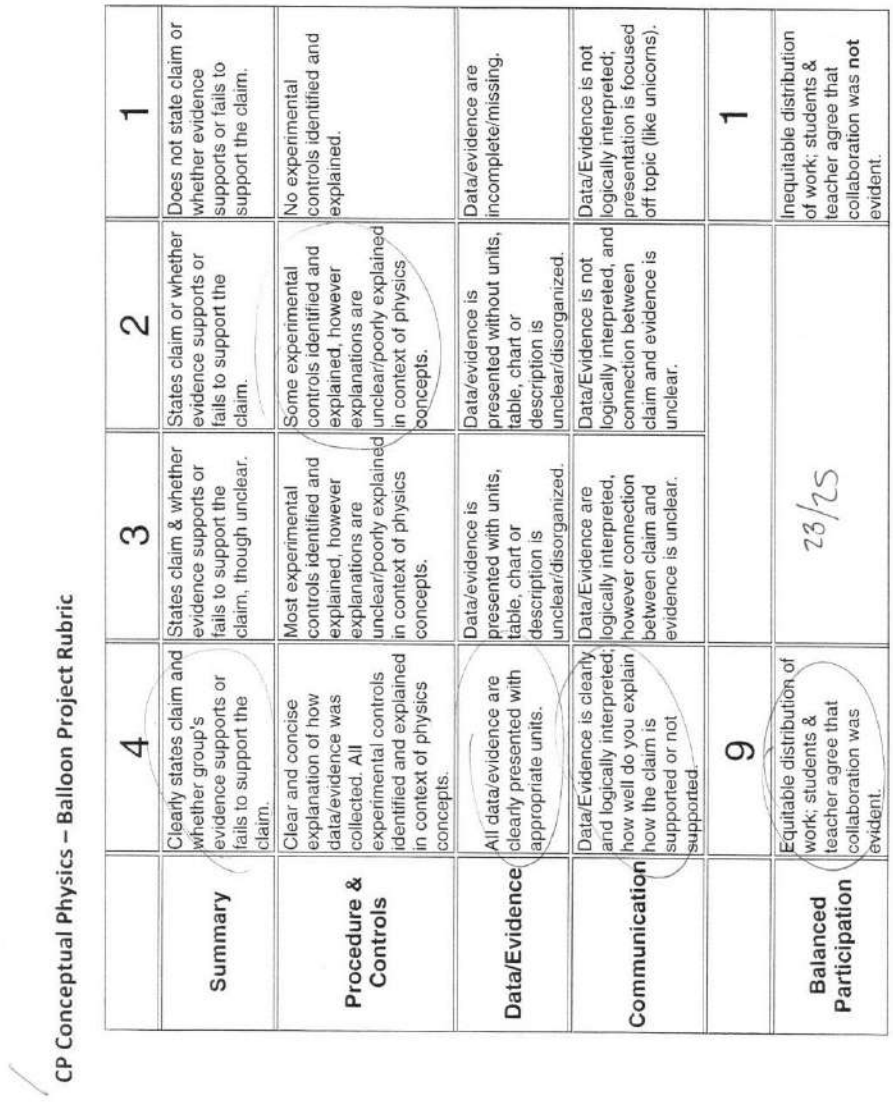
An example of an assignment scoring high on overall quality includes the assignment presented in Figures 6a and 6b from a high school physics class:

[16] The potential systematic variation in assignment quality may be attributable to systematic differences in teachers among schools (PLCs, for example), and/or differences among student preparedness. This can be further examined by utilizing prior student performance. Similar analyses were completed by the MET project in examining the variability of observation and student survey results.

[17] Examination of the correlations among the scores reveals that clarity of goals is least correlated with the other constructs and eliminating increases the reliability of the indicator.

Figure 6a: A High Quality Assignment (part 1)

Summary	The manufacturer claims that their water balloon launcher can launch a balloon 200 yards. Our data supported/failed to support the manufacturer's claim. Our group's data revealed that the average launch range of the launcher was _____ yards.	
Procedure & Controls	**Control 1:**	**Physics Concept:**
	Control 2:	**Physics Concept:**
	Control 3:	**Physics Concept:**
	Control 4:	**Physics Concept:**
Data/Evidence	**Trial 1** _____ m	Claimed Range: 200 yards. Average Range: _____ yards.
	Trial 2 _____ m	
	Trial 3 _____ m	
	Trial 4 _____ m	
	Trial 5 _____ m	
	Trial 6 _____ m	
	Trial 7 _____ m	
	Trial 8 _____ m	
	Trial 9 _____ m	
	Trial 10 _____ m	



CP Conceptual Physics – Balloon Project Rubric

	4	3	2	1
Summary	Clearly states claim and whether group's evidence supports or fails to support the claim.	States claim & whether evidence supports or fails to support the claim, though unclear.	States claim or whether evidence supports or fails to support the claim.	Does not state claim or whether evidence supports or fails to support the claim.
Procedure & Controls	Clear and concise explanation of how data/evidence was collected. All experimental controls identified and explained in context of physics concepts.	Most experimental controls identified and explained, however explanations are unclear/poorly explained in context of physics concepts.	Some experimental controls identified and explained, however explanations are unclear/poorly explained in context of physics concepts.	No experimental controls identified and explained.
Data/Evidence	All data/evidence are clearly presented with appropriate units.	Data/evidence is presented with units, table, chart or description is unclear/disorganized.	Data/evidence is presented without units, table, chart or description is unclear/disorganized.	Data/evidence are incomplete/missing.
Communication	Data/Evidence is clearly and logically interpreted; how well do you explain how the claim is supported or not supported.	Data/Evidence are logically interpreted, however connection between claim and evidence is unclear.	Data/Evidence is not logically interpreted, and connection between claim and evidence is unclear.	Data/Evidence is not logically interpreted; presentation is focused off topic (like unicorns).
	9	23/25		1
Balanced Participation	Equitable distribution of work; students & teacher agree that collaboration was evident.			Inequitable distribution of work; students & teacher agree that collaboration was **not** evident.

We scored this assignment as being of high quality because:

Task Summary: Required students to work in a group to test an advertising claim and present their findings to the class.

Cognitive Challenge: The task provided an opportunity to use physics in a real life situation, thus creating relevance for the students. The rigor lay in the exploration of the procedures/controls and the presentation of the data/evidence[18]. The work samples show that depth of knowledge of the work accepted was less than the

[18] We note, however, the work samples show that depth of knowledge of the work accepted was less than the assignment allowed for, but we did not score for this misalignment.

assignment allowed for. The task also required students to participate in peer discussions regarding findings.

Clarity: The teacher provided detailed guidelines for success: assignment template with summary of activity, thesis statement, graphic organizer for procedures/controls and data/evidence.

Grading: Grading criteria provided detailed guidelines for success and contains potential for student improvement: four-point rubric correlated to instruction template, two categories, project itself and balanced [group] participation.

Another example of a high quality assignment is presented in figure 7. We scored this assignment high because:

Task Summary: Students chose from four prompts and wrote a RACE (**R**estate the question, **A**nswer the question, **C**ite evidence, **E**xplain your answer) response.

Cognitive Challenge: The differentiated prompts allowed for student meta-cognition. Each prompt requires the students to extend their thinking from what is presented directly in the text (Romeo and Juliet) to what it possibly represents and why.

Clarity: The teacher provided detailed guidelines for success: written assignment directions were provided to each student, directions included key factors for success in bold, the prompts and rubric were on that same page allowing students to compare directions, prompts, and rubric and then determine which prompt would provide the maximum opportunity for success.

Grading: Grading criteria provided detailed guidelines for success and contains potential for student improvement: each element of the RACE response was a category, with three levels of proficiency.

Figure 7: A High Quality Assignment in Language Arts

Short Response

Please choose ONE short response prompt. Write a RACE response to answer the prompt that you choose. You MUST include specific details from the text to earn full credit.

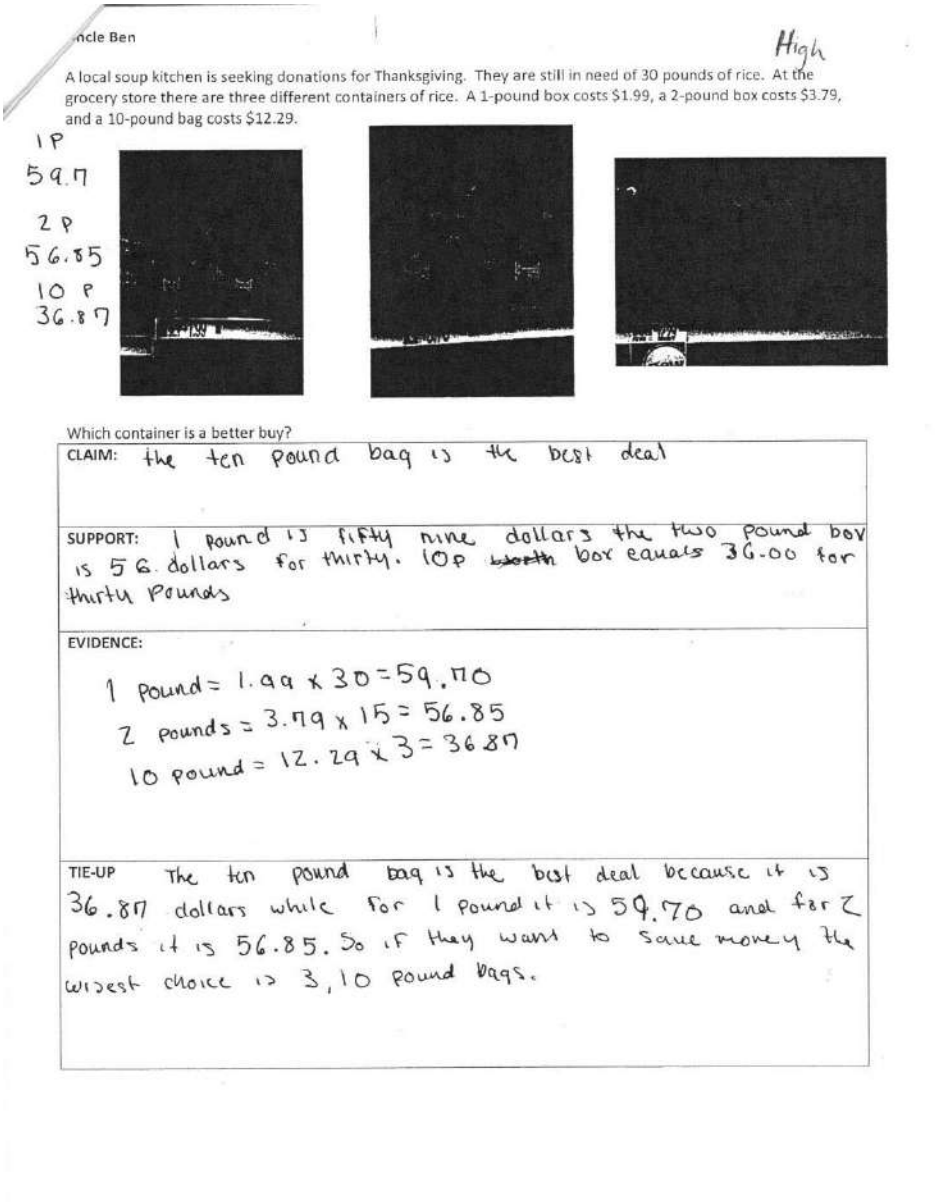
- Compare and contrast Romeo's view on love with Friar Lawrence's view on love. How do Romeo and Friar Lawrence view love differently? Use specific details from the text to support your answer.
- At the beginning of the play, Romeo is upset because he loves Roseline but she does not love him back. How does including Roseline in the play help the audience understand Romeo's personality? How would the play be different if Roseline was not included? Use specific details from the text to support your answer.
- Who do you think is to blame for Romeo and Juliet's deaths? Explain which character you think is most guilty and why, using specific details from the text to support your answer.
- One of the play's main themes is that fate is unavoidable, meaning certain things will happen no matter what you try to do. Explain how this theme is conveyed in the play, using specific details from the text to support your answer.

RACE Rubric	10-9	8-7	6
Restate the question	Student restates question accurately. A reader would be able to tell what the original prompt was.	Student attempts to restate but is not accurate	Student does not restate the question accurately
Answer the question	Answer was accurate, specific, and complete. All parts of the question are answered.	Answer was accurate but incomplete; may be vague or lacking details	Student makes an attempt but the answer is unclear, incomplete, or inaccurate.
Cite Evidence and/or examples	Student provides specific evidence, ideally a direct quote, that clearly supports the answer.	Student provides an example but it may be irrelevant or too general.	Student provides no evidence or examples to support the answer.
Explain Your Answer	Student gives a clear explanation of how the example supports the answer. The reader does not need to infer how the example and answer are related.	The explanation is not detailed enough and does not clearly explain how the evidence supports the answer. There may be too much focus on plot summary.	Student provides no explanation or explanation does not help answer the question.

Figure 8 provides an example of an assignment of lower quality.

Grading: There is no information provided regarding scoring.

Figure 8: Lower Quality 8th Grade Math Assignment



ncle Ben

High

A local soup kitchen is seeking donations for Thanksgiving. They are still in need of 30 pounds of rice. At the grocery store there are three different containers of rice. A 1-pound box costs $1.99, a 2-pound box costs $3.79, and a 10-pound bag costs $12.29.

1P
59.7

2P
56.85

10P
36.87

Which container is a better buy?

CLAIM: the ten pound bag is the best deal

SUPPORT: 1 pound is fifty nine dollars the two pound box is 56 dollars for thirty. 10P ~~box~~ box equals 36.00 for thirty pounds

EVIDENCE:

1 pound = 1.99 x 30 = 59.70

2 pounds = 3.79 x 15 = 56.85

10 pound = 12.29 x 3 = 36.87

TIE-UP The ten pound bag is the best deal because it is 36.87 dollars while for 1 pound it is 59.70 and for 2 pounds it is 56.85. So if they want to save money the wisest choice is 3, 10 pound bags.

We scored this assignment as being of low quality because:

Task Summary: In this Middle School, 8th Grade math assignment students were to consider the better buy of two shopping choices.

Cognitive Challenge: While students did need to support their answers, the two questions were both DOK One requiring very basic multiplication/division calculations. (Note: 77% of students are on grade-level; 1% ELL.)

Clarity: The assignment is clear with written instructions and a graphic organizer for students to use.

We asked teachers to indicate what the primary (i.e. more than 50% of the instruction) instructional strategy was in preparing students to be able to complete the assignment. Six teachers indicated that they did not use any of the instructional strategies presented (see Appendix D3) for more 50% of instruction. Of the remaining teachers (19) 13 of them indicated that they used at least three of the instructional strategies more than 50% of the time.

As part of the validation process we are interested in whether assignment quality is related to teacher effectiveness and teacher observed effectiveness. As part of this analysis we are also interested in whether there are assignment quality differences associated with instructional strategies and whether there are teacher effectiveness differences associated with instructional strategies. Given the exploratory nature of the analysis we test whether there are differences among the specific assignment quality criteria for each instructional strategy. Overall, the results indicate that there is no difference in teacher effectiveness across the instructional strategies based on observations ratings. In general, assignment quality or its criteria are about equally likely to be endorsed across the various instructional strategies. There is some suggestive evidence that instructional strategies aligned with direct instruction tended to score more highly than more inquiry-based methods; specifically, on assignments providing detailed guidelines for success and the potential of assignments to enhance student learning opportunities[19].

The quality of a teacher's assignment is intended to provide an independent criterion to evaluate the quality of instructional practices – or specifically in this instance teacher effectiveness. We find that assignment quality is not related to teacher effectiveness as measured by observation scores (in fact the results suggest a slightly negative relationship ($p < .10$). This is inconsistent with the literature, which finds low to moderate correlations between assignment quality and both observation and student achievement results (Matsumura, Garnier, Slater, and Boston, 2008) but may be the result of several limitations associated with this set of exploratory analyses. The literature indicates that results tend to be most stable and reliable when using three assignments. This analysis is limited to a single assignment. Also, we utilized an abbreviated instrument with which to evaluate assignment quality. The abbreviated form was selected for two reasons: one, to reduce the burden on teachers and increase the likelihood of responses; and two, to determine whether assignment quality might be a viable multiple measure for an educator effectiveness system. Also, as noted previously, the sample size is not only small, but lacks much variability in teacher rating (a coefficient of variation of less than 0.06).

Summary of Artifact Analyses

This evaluation examined three sets of artifacts from teachers. Teachers provided written feedback, their component IV growth goals, and student assignments. Overall we find the feedback tends to fall along two dimensions. One dimension focuses on communication and human resource related issues. This dimension of feedback provides results related to how well principals feedback is written – whether it is clear and objective. Principals generally do a god job in this dimension. The other dimension focuses more critically on instructional practices. This dimension focuses on feedback using specific classroom practices to highlight strengths and weaknesses and develop concrete recommendations for improvement as well as strategies that engender that improvement. Principals were less successful at providing this sort of feedback. Overall, there are differences in

[19] Whether this is related to actual student performance can be addressed with student assessment results.

feedback, with some evidence suggesting that the quality of feedback varies by school. Also, more effective teachers tended to get poorer feedback while novice teachers also received poorer instructional feedback. Evidence also suggests that the fewer FFT elements that were scored, the poorer the quality of the feedback.

Tangentially related to the number of FFT elements scored, we note that elements not scored appeared to be somewhat ad hoc, and seemingly particularly germane to either feedback or teaching practices.

Component IV appears to be completed with less fidelity than other elements of the system, both in terms of the form provided to complete this task, and in terms of the guidance provided. Both teachers and evaluators (although we did not specifically evaluate feedback related to growth goals—the form lacked sufficient space for evaluators to comment). The results indicate that, based on the criteria applied, teachers generally did not develop succinct growth goal plans that incorporated concrete steps as well as measureable benchmarks for success. Component IV guidance might provide teachers with a better understanding of what the purpose of writing goals is, and may facilitate more critical reflection by teachers. Consistent with expectations more effective teacher (as measured by observation ratings) tended to write better growth goal plans.

We also examined teacher assignment quality in hopes of using the assignment quality instrument as a tool to help provide validity evidence for inferences made about teachers based on DPAS-II results. The instrument (while fairly reliable) was did not perform well in differentiating teacher effectiveness and did not align well with inferences about instructional quality as measured by the assignments and observation ratings Whether this hold true for student learning is unknown. It is recommended to conduct such as analysis as it may shed additional light on inferences about teachers.

References

Ballou, Dale, William Sanders and Paul Wright. (2004). "Controlling for Student Background in Value-Added Assessment of Teachers." *Journal of Educational and Behavioral Statistics*, 29 (1), pp. 37-66.

Bjorn, k., S. Wurth,.and A. Hergovich (2013). The Impact of Feedback on Goal Setting and Task Performance, *Swiss Journal of Psychology*, 72(2): 79-89.

Bryk, A., Thum, Y.M., Easton, J. & Luppescu, S.(1998). Assessing school academic productivity: the case of Chicago school reform. *Social Psychology of Education*, 2, 103-142.

Daly, Timothy, and Arun Ramanathan. 2010. "California's Quality-blind Layoffs Law Harms Teachers and Students." *The Los Angeles Times*, March 24, 2010.

Goldhaber, D., & Hansen, M. (2009). *Assessing the potential of using value-added estimates of teacher job performance for making tenure decisions*. (CRPE Working Paper # 2009_2). Retrieved from Center on Reinventing Public Education website: http://www.crpe.org/cs/crpe/download/csr_files/wp_crpe_vamtenure_dec09_.pdf

Goldhaber, D., & Hansen, M. (2010). *Using performance on the job to inform teacher tenure decisions.* (CALDER Brief 10). Retrieved from Urban Institute website: http://www.urban.org/uploadedpdf/1001385-using-performance-calder-brief10.pdf

Goldschmidt, P. and J.F. Martinez-Fernandez (2007). The Relationship among Measures as Empirical Evidence of Validity: Incorporating Multiple Indicators of Achievement and School Context, *Educational Assessment* 12(3&4).

Hahnel, C. and O. Jackson (2012). Learning Denied: The Case for Equitable Access to Effective Teaching in California's Largest School District, The Education Trust – West

Heneman III, H. and A. Milanowski (2004). Alignment of Human Resource Practices and Teacher Performance Competency, *Peabody Journal*, 79(4): 108-125.

Jacob, B.A., Lefgren, L., and Sims, D.P. 2010. "The Persistence of Teacher-Induced Learning." *Journal of Human Resources* 45(4): 915-943.

Kimball, S. and A. Milanowski (2009). Examining Teacher Evaluation Validity and the Leadership Decision Making Within a Standards-Based Evaluation System, *Educational Administrative Quarterly*, 45(1):34-70.

Kluger, A. and A. DeNisi (1996). The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory, Psychological Bulletin, 119(2): 254-284.

Ladd, H. (2008). Teacher Effects: What Do We Know?. Paper prepared for the Teacher Quality Conference, Northwestern University, May, 2008.

Lockwood, J.R. and Daniel F. McCaffrey. 2007. "Controlling for individual heterogeneity in longitudinal models, with applications to student achievement," *Electronic Journal of Statistics.* Vol 1, ppl 223-252.

Matsumura L.C., Garnier, H., Slater, S.C., and M. Boston (2008). Toward Measuring Instructional Interactions "At-Scale", *Educational Assessment*, 13:267-300.

McCafffrey, Daniel, F, J.R. Lockwood, Daniel Koretx, Thomas A. Louis, and Laura Hamilton. 2004. "Models for Value-Added Modeling of Teacher Effects," *Journal of Educational and Behavioral Statistics,* vol. 29, no. 1 (Spring), pp. 67-101.

McCaffrey, D. F., Sass, T. R., Lockwood, J. R., & Mihaly, K. (2009). The intertemporal variability of teacher effect estimates. *Education Finance and Policy*, *4*, 572-606.

Meyer, Herbert (1991). A solution to the performance appraisal feedback enigma, Academy of Management Executive, 5(1):68-76.

Nicol, D. and D. Macfarlane-Dick (2004) Rethinking Formative Assessment in HE: a theoretical model and seven principles of good feedback, Briefing Paper. Retrieved rom: http://www.heacademy.ac.uk/assessment/ASS051D_SENLEF_model.doc.

Raudenbush, S. & Willms, D. (1995). The estimation of school effects. *Journal of Educational and Behavioral Statistics*, 20(4), 307-335.

Sanders W, Saxton A, Horn B. The Tennessee Value-Added Assessment System: A quantitative outcomes-based approach to educational assessment. In: Millman J, editor. Grading teachers, grading schools: Is student achievement a valid evaluational Measure? Thousand Oaks, CA: Corwin Press, Inc; 1997. pp. 137–162.

Thurlings, M, Vermeulen, Kreijins, K., Bastiaens, T., and S. Stijnen (2012). Development of the Teacher Feedback Observation Scheme: evaluating the quality of feedback in peer groups, *Journal of Education for Teaching: International research and pedagogy*, 38(2):193-208.

Walling, A., Shapiro, J., and T. Ast (2013). What Makes a Good Reflective Paper? Family *Medicine*, 34(1):7-12.

Willms, D. & Raudenbush, S. (1989). A longitudinal hierarchical linear model for estimating school effects and their stability, *Journal of Educational Measurement*, 26(3), 209-232.

Zhoa, N. (2009). The Minimum Sample Size in Factor Analysis, Retrieved from: https://www.encorewiki.org/display/~nzhao/The+Minimum+Sample+Size+in+Factor+Analysis

Appendix D1

Arrange		Calculate		Design		Hypothesize	
Assess		Compile		Differentiate		Infer	
Classify		Create		Formulate		Justify	
Compare		Describe		Label		Locate	
Conclude		Develop		Measure		Modify	
Connect		Extend		Predict		Quote	
Critique		Generalize		Prove		Represent	
Define		Graph		Solve		Synthesize	
Summarize		Tabulate					

Appendix D2

Purpose	Not Applicable	Secondarily	Generally	Primarily
Gather background knowledge				
Report on the presented information				
Reference to support student ideas				
Predict sequence of events				
Predict character actions				
Help predict results				
Summarize the presented information				
Summarize the plot				
Compare views/information within the text				
Compare views/information between texts				
Develop own idea using text as launching point				
Collect factual information				
Analyze presented information for new use				

Appendix D3

What types of instructional strategies did you use to prepare students for this assignment prior to distributing the assignment? Please check the appropriate boxes.

Strategy	Amount of time spent			
	less than 10%	11%-25%	26%-50%	More than 50%
Modeling Task				
Explicit Instruction				
Teacher Presentation				
Student Presentations				
Small Work Groups				
Discussion Opportunities				
Detailed Sequences of Work				
Feedback of Prior, Relevant Student Work				
Available Resources				
Jigsaws/Think Pair Share				
Videos				
Socratic Questioning				
Service Learning				
Game Based Learning				
Guided Discovery Learning				
Student Collaboration				
Modeling Thinking Processes				
Inquiry-based activities				

Appendix D4.

Assignment Information Sheet

Teacher Name: _____

School: _____

Grade: _____

Subject: _____

Please attach a copy of a typical assignment, the assignment directions you distributed to students, copies of any supporting materials you distributed to the students (notes, handouts, texts, etc.), and any scoring rubrics you used and/or distributed to students. Also, please provide the examples of student work (low and high quality).

Learning Goals – What were the specific goals you expected the students to demonstrate through this assignment?

_____

_____

_____

_____

_____

_____

How often do you assign an assignment like this? _____ times per school year.

Detailed Assignment Description – What were the students asked to do in this assignment? Check up to 4 primary tasks ordered from primary (1) to ancilary (4).

Arrange		Calculate		Design		Hypothesize	
Assess		Compile		Differentiate		Infer	
Classify		Create		Formulate		Justify	
Compare		Describe		Label		Locate	
Conclude		Develop		Measure		Modify	
Connect		Extend		Predict		Quote	
Critique		Generalize		Prove		Represent	
Define		Graph		Solve		Synthesize	
Summarize		Tabulate					

Material Information – What kinds of reading and/or reference material did you use for this assignment (e.g. textbook, article, graphs, maps, photographs or other primary source documents, etc.)?

Text Type	Sources	Text Title	# of Pages Assigned

How were the students to use the reading material for this assignment? Please check the appropriate boxes.

Purpose	Not Applicable	Secondarily	Generally	Primarily
Gather background knowledge				
Report on the presented information				
Reference to support student ideas				
Predict sequence of events				
Predict character actions				
Help predict results				
Summarize the presented information				
Summarize the plot				
Compare views/information within the text				
Compare views/information between texts				
Develop own idea using text as launching point				
Collect factual information				
Analyze presented information for new use				

What types of instructional strategies did you use to prepare students for this assignment prior to distributing the assignment? Please check the appropriate boxes.

Strategy	Amount of time spent			
	less than 10%	11%-25%	26%-50%	More than 50%
Modeling Task				
Explicit Instruction				
Teacher Presentation				
Student Presentations				
Small Work Groups				
Discussion Opportunities				
Detailed Sequences of Work				
Feedback of Prior, Relevant Student Work				
Available Resources				
Jigsaws/Think Pair Share				
Videos				
Socratic Questioning				
Service Learning				
Game Based Learning				
Guided Discovery Learning				
Student Collaboration				
Modeling Thinking Processes				
Inquiry-based activities				

Monitoring and Assessment – How did you monitor student performance during the assignment? Please check the appropriate boxes.

	Amount of time spent			
	less than 10%	11%-25%	26%-50%	More than 50%
Circulating the classroom				
Work with small groups throughout				
Peer Monitoring				
Students report out progress				

Did this assignment require any reteach or modification? Please check the appropriate boxes.

	Whole Class	Small Group	Individual Student(s)
Of other skills prior to distribution of this assignment?			
While assessing progress of this assignment?			

After evaluating results of this assignment?			

-Class Demographics –

What is the percentage of ELLs in your classroom? _____

What is the percentage of students classified with disabilities in your classroom? _____

What percentage of your students are:
at grade level? _____
below grade level? _____
above grade level? _____

If the assignment was not completed by the end of class, what percentage of the assignment was done in class and what percentage was done at home?
In class:_____ At home:_____

How many class minutes did students take to complete the assignment? _____

Appendix E: Views of Current and Past Changes to DPAS-II

The Delaware Department of Education has instituted—or is considering—various changes to the DPAS-II system based on feedback and statewide goals for educator effectiveness, Table 1.2 summarizes these changes.

Table 1E. Current and Proposed Changes to DPAS-II

Change	Description
Changes Made for 2014-15 School Year	
Changes to Components II and III	Evaluators may use **short** observations, which must be at least 10 minutes, after at least one full observation has occurred; applicable to Components II & III only.
Changes to Component IV	Districts/charters can opt to strengthen Component IV, for example by substituting a collaboratively-developed Component
Credentialed Observers	Districts can credential additional observers to assist with the DPAS-II process
Criterion-level Ratings	All educators are required to receive ratings on each of the criteria in the DPAS-II for Teachers and Specialists rubric
Proposed Future Changes	
Changes in Weighting	Components I through IV would receive greater emphasis, as evaluators would have more discretion in using Component V scores when Components I through IV are strong
Increasing the Number of Rating Categories for Components I through IV	Components I through IV would be scored along a 4-point scale rather than a binary ("Satisfactory"/"Unsatisfactory") scale
Annual Appraisals	Beginning in 2016-2017, Annual Summative Appraisals would be required of all teachers

This appendix provides insight of the views from teachers, specialists, and administrators of recent and proposed changes described in the table above.

Table 2E. Teachers Views on whether the Current and Proposed Changes Enhance DPAS-II

Change	Description					
Changes Made for 2013-15 School Years		Very much	Some-what	Slightly	Not at All	Do Not Know
Changes to Component II and III	Evaluators may use **short** observations, which must be at least 10-minutes, after at least one full observation has occurred for Components II & III only.	10.1	25.8	18.1	20.4	25.7
Changes to Component IV	Districts/charters can opt to strengthen Component IV, for example by substituting a collaboratively developed Component	6.6	19.8	16.3	15.7	41.7
Credentialed Observers	Districts can credential additional observers to assist with the DPAS-II process	6.0	16.5	15.7	27.4	34.5
Criterion-level Ratings	All educators are required to receive ratings on each of the criteria in the DPAS-II for Teachers and Specialists rubric	7.2	26.3	23.0	21.8	21.7
Proposed Changes for 2015-16 School Year						
Changes in Weighting	Components I through IV would receive greater emphasis, as evaluators would have more discretion in using Component V scores when Components I through IV are strong	25.2	31.5	17.9	11.1	14.3
Increasing the Number of Rating Categories for Components I through IV	Each of Components I through IV would be assigned a score along a 4-point scale rather than a binary ("Satisfactory"/"Unsatisfactory") scale	16.9	32.4	20.4	17.3	13.0
Annual Appraisals	Beginning in 2016-2017, Annual Summative Appraisals would be required of all teachers	7.7	21.5	19.2	36.2	15.4

Table 3E. Specialists Views on whether the Current and Proposed Changes Enhance DPAS-II

Change	Description					
Changes Made for 2013-15 School Years		Very much	Some-what	Slightly	Not at All	Do Not Know
Changes to Component II and III	Evaluators may use **short** observations, which must be at least 10-minutes, after at least one full observation has occurred for Components II & III only.	8.7	23.3	10.2	14.9	42.9
Changes to Component IV	Districts/charters can opt to strengthen Component IV, for example by substituting a collaboratively developed Component	7.0	18.0	10.6	12.8	51.6
Credentialed Observers	Districts can credential additional observers to assist with the DPAS-II process	7.4	15.2	10.6	20.5	46.4
Criterion-level Ratings	All educators are required to receive ratings on each of the criteria in the DPAS-II for Teachers and Specialists rubric	5.4	20.6	17.4	22.8	33.8
Proposed Changes for 2015-16 School Year						
Changes in Weighting	Components I through IV would receive greater emphasis, as evaluators would have more discretion in using Component V scores when Components I through IV are strong	18.0	25.9	16.7	13.3	25.9
Increasing the Number of Rating Categories for Components I through IV	Each of Components I through IV would be assigned a score along a 4-point scale rather than a binary ("Satisfactory"/"Unsatisfactory") scale	12.6	26.4	15.8	20.4	24.8
Annual Appraisals	Beginning in 2016-2017, Annual Summative Appraisals would be required of all teachers	5.6	18.3	14.2	35.7	26.3

Table 4E. Administrators Views on whether the Current and Proposed Changes Enhance DPAS-II

Change	Description					
Changes Made for 2013-15 School Years		Very much	Some-what	Slightly	Not at All	Do Not Know
Changes to Component II and III	Evaluators may use **short** observations, which must be at least 10-minutes, after at least one full observation has occurred for Components II & III only.	18.07	33.33	19.68	20.48	8.43
Changes to Component IV	Districts/charters can opt to strengthen Component IV, for example by substituting a collaboratively developed Component	7.23	18.47	16.47	22.09	35.74
Credentialed Observers	Districts can credential additional observers to assist with the DPAS-II process	16.80	22.40	14.00	27.60	19.20
Criterion-level Ratings	All educators are required to receive ratings on each of the criteria in the DPAS-II for Teachers and Specialists rubric	13.65	34.94	22.49	23.29	5.62
Proposed Changes for 2015-16 School Year						
Changes in Weighting	Components I through IV would receive greater emphasis, as evaluators would have more discretion in using Component V scores when Components I through IV are strong	34.00	32.80	18.00	9.20	6.00
Increasing the Number of Rating Categories for Components I through IV	Each of Components I through IV would be assigned a score along a 4-point scale rather than a binary ("Satisfactory"/"Unsatisfactory") scale	17.60	37.60	17.20	17.20	10.40
Annual Appraisals	Beginning in 2016-2017, Annual Summative Appraisals would be required of all teachers	8.00	22.00	11.20	52.40	6.40