

Uses models:	No
Uses documents:	Yes
Uses topic statements:	No
Emulates:	Pyramid Responsiveness

	Initial summaries						Update summaries							
	<i>r</i>	p (95% CI)		<i>rho</i>	p	<i>tau</i>	p	<i>r</i>	p (95% CI)		<i>rho</i>	p	<i>tau</i>	p
Pyramid	NaN	NaN (NaN - NaN)		NaN	NaN	NaN	1.000	NaN	NaN (NaN - NaN)		NaN	NaN	NaN	1.000
Responsiveness	NaN	NaN (NaN - NaN)		NaN	NaN	NaN	1.000	NaN	NaN (NaN - NaN)		NaN	NaN	NaN	1.000
Readability	NaN	NaN (NaN - NaN)		NaN	NaN	NaN	1.000	NaN	NaN (NaN - NaN)		NaN	NaN	NaN	1.000

Table 1: System-level correlations (Pearson’s *r*, Spearman’s *rho*, Kendall’s *tau*) for uOttawa1, initial and update summaries, all peers.

	Initial summaries						Update summaries							
	<i>r</i>	p (95% CI)		<i>rho</i>	p	<i>tau</i>	p	<i>r</i>	p (95% CI)		<i>rho</i>	p	<i>tau</i>	p
Pyramid	0.586	0.000 (0.370 - 0.742)		0.713	0.000	0.559	0.000	0.569	0.000 (0.349 - 0.730)		0.631	0.000	0.470	0.000
Responsiveness	0.576	0.000 (0.357 - 0.735)		0.621	0.000	0.437	0.000	0.535	0.000 (0.304 - 0.706)		0.568	0.000	0.428	0.000
Readability	0.435	0.001 (0.182 - 0.635)		0.365	0.008	0.247	0.011	0.380	0.006 (0.117 - 0.594)		0.213	0.133	0.127	0.193

Table 2: System-level correlations (Pearson’s *r*, Spearman’s *rho*, Kendall’s *tau*) for uOttawa1, initial and update summaries, no models.

	Initial summaries			Update summaries		
	all	avg per topic	avg per assessor	all	avg per topic	avg per assessor
Pyramid	0.264	0.343	0.285	0.190	0.287	0.228
Responsiveness	0.235	0.287	0.306	0.205	0.213	0.218
Readability	0.164	0.185	0.192	0.182	0.182	0.193

Table 3: Summary-level correlations (Pearson’s *r*) for uOttawa1, initial and update summaries, no models.

		A>B	A=B	A<B
		uOttawa1	A>B	0
	A=B	408	0	
	A>B			0

Table 4: Discriminative power for uOttawa1 compared with Pyramid. Initial summaries, models vs. non-models.

		A>B	A=B	A<B
		uOttawa1	A>B	0
	A=B	408	0	
	A>B			0

Table 5: Discriminative power for uOttawa1 compared with Pyramid. Update summaries, models vs. non-models.

		A>B	A=B	A<B
		uOttawa1	A>B	0
	A=B	408	0	
	A>B			0

Table 6: Discriminative power for uOttawa1 compared with Responsiveness. Initial summaries, models vs. non-models.

		A>B	A=B	A<B
		uOttawa1	A>B	0
	A=B	408	0	
	A>B			0

Table 7: Discriminative power for uOttawa1 compared with Responsiveness. Update summaries, models vs. non-models.

		A>B	A=B	A<B
		uOttawa1	A>B	0
	A=B	407	1	
	A>B			0

Table 8: Discriminative power for uOttawa1 compared with Readability. Initial summaries, models vs. non-models.

		A>B	A=B	A<B
		uOttawa1	A>B	0
	A=B	408	0	
	A>B			0

Table 9: Discriminative power for uOttawa1 compared with Readability. Update summaries, models vs. non-models.

		A>B	A=B	A<B
uOttawa1	A>B	157	435	
	A=B	80	601	
	A>B			2

Table 10: Discriminative power for uOttawa1 compared with Pyramid. Initial summaries, non-model summaries.

		A>B	A=B	A<B
uOttawa1	A>B	146	465	
	A=B	41	623	
	A>B			0

Table 11: Discriminative power for uOttawa1 compared with Pyramid. Update summaries, non-model summaries.

		A>B	A=B	A<B
uOttawa1	A>B	141	453	
	A=B	80	601	
	A>B			0

Table 12: Discriminative power for uOttawa1 compared with Responsiveness. Initial summaries, non-model summaries.

		A>B	A=B	A<B
uOttawa1	A>B	90	521	
	A=B	38	626	
	A>B			0

Table 13: Discriminative power for uOttawa1 compared with Responsiveness. Update summaries, non-model summaries.

		A>B	A=B	A<B
uOttawa1	A>B	198	372	
	A=B	192	489	
	A>B			24

Table 14: Discriminative power for uOttawa1 compared with Readability. Initial summaries, non-model summaries.

		A>B	A=B	A<B
uOttawa1	A>B	175	425	
	A=B	139	525	
	A>B			11

Table 15: Discriminative power for uOttawa1 compared with Readability. Update summaries, non-model summaries.

Uses models:	Yes
Uses documents:	No
Uses topic statements:	No
Emulates:	Pyramid Responsiveness

	Initial summaries						Update summaries							
	<i>r</i>	p (95% CI)		<i>rho</i>	<i>p</i>	<i>tau</i>	<i>p</i>	<i>r</i>	p (95% CI)		<i>rho</i>	<i>p</i>	<i>tau</i>	<i>p</i>
Pyramid	NaN	NaN (NaN - NaN)		NaN	NaN	NaN	1.000	NaN	NaN (NaN - NaN)		NaN	NaN	NaN	1.000
Responsiveness	NaN	NaN (NaN - NaN)		NaN	NaN	NaN	1.000	NaN	NaN (NaN - NaN)		NaN	NaN	NaN	1.000
Readability	NaN	NaN (NaN - NaN)		NaN	NaN	NaN	1.000	NaN	NaN (NaN - NaN)		NaN	NaN	NaN	1.000

Table 1: System-level correlations (Pearson’s *r*, Spearman’s *rho*, Kendall’s *tau*) for uOttawa2, initial and update summaries, all peers.

	Initial summaries						Update summaries							
	<i>r</i>	p (95% CI)		<i>rho</i>	<i>p</i>	<i>tau</i>	<i>p</i>	<i>r</i>	p (95% CI)		<i>rho</i>	<i>p</i>	<i>tau</i>	<i>p</i>
Pyramid	0.940	0.000 (0.897 - 0.966)		0.866	0.000	0.717	0.000	0.842	0.000 (0.738 - 0.907)		0.731	0.000	0.559	0.000
Responsiveness	0.906	0.000 (0.841 - 0.946)		0.756	0.000	0.603	0.000	0.837	0.000 (0.730 - 0.904)		0.717	0.000	0.546	0.000
Readability	0.730	0.000 (0.569 - 0.837)		0.292	0.038	0.213	0.029	0.588	0.000 (0.373 - 0.743)		0.277	0.049	0.191	0.051

Table 2: System-level correlations (Pearson’s *r*, Spearman’s *rho*, Kendall’s *tau*) for uOttawa2, initial and update summaries, no models.

	Initial summaries			Update summaries		
	all	avg per topic	avg per assessor	all	avg per topic	avg per assessor
Pyramid	0.487	0.609	0.541	0.326	0.492	0.389
Responsiveness	0.448	0.521	0.522	0.318	0.385	0.365
Readability	0.357	0.421	0.391	0.216	0.270	0.287

Table 3: Summary-level correlations (Pearson’s *r*) for uOttawa2, initial and update summaries, no models.

		A>B	A=B	A<B
		uOttawa2	A>B	0
	A=B	408	0	
	A>B			0

Table 4: Discriminative power for uOttawa2 compared with Pyramid. Initial summaries, models vs. non-models.

		A>B	A=B	A<B
		uOttawa2	A>B	0
	A=B	408	0	
	A>B			0

Table 5: Discriminative power for uOttawa2 compared with Pyramid. Update summaries, models vs. non-models.

		A>B	A=B	A<B
		uOttawa2	A>B	0
	A=B	408	0	
	A>B			0

Table 6: Discriminative power for uOttawa2 compared with Responsiveness. Initial summaries, models vs. non-models.

		A>B	A=B	A<B
		uOttawa2	A>B	0
	A=B	408	0	
	A>B			0

Table 7: Discriminative power for uOttawa2 compared with Responsiveness. Update summaries, models vs. non-models.

		A>B	A=B	A<B
		uOttawa2	A>B	0
	A=B	407	1	
	A>B			0

Table 8: Discriminative power for uOttawa2 compared with Readability. Initial summaries, models vs. non-models.

		A>B	A=B	A<B
		uOttawa2	A>B	0
	A=B	408	0	
	A>B			0

Table 9: Discriminative power for uOttawa2 compared with Readability. Update summaries, models vs. non-models.

		A>B	A=B	A<B
uOttawa2	A>B	216	55	
	A=B	23	981	
	A>B			0

Table 10: Discriminative power for uOttawa2 compared with Pyramid. Initial summaries, non-model summaries.

		A>B	A=B	A<B
uOttawa2	A>B	86	34	
	A=B	101	1054	
	A>B			0

Table 11: Discriminative power for uOttawa2 compared with Pyramid. Update summaries, non-model summaries.

		A>B	A=B	A<B
uOttawa2	A>B	199	72	
	A=B	22	982	
	A>B			0

Table 12: Discriminative power for uOttawa2 compared with Responsiveness. Initial summaries, non-model summaries.

		A>B	A=B	A<B
uOttawa2	A>B	75	45	
	A=B	53	1102	
	A>B			0

Table 13: Discriminative power for uOttawa2 compared with Responsiveness. Update summaries, non-model summaries.

		A>B	A=B	A<B
uOttawa2	A>B	225	45	
	A=B	188	816	
	A>B			1

Table 14: Discriminative power for uOttawa2 compared with Readability. Initial summaries, non-model summaries.

		A>B	A=B	A<B
uOttawa2	A>B	93	27	
	A=B	232	923	
	A>B			0

Table 15: Discriminative power for uOttawa2 compared with Readability. Update summaries, non-model summaries.

Uses models:	Yes
Uses documents:	Yes
Uses topic statements:	No
Emulates:	Pyramid Responsiveness

	Initial summaries						Update summaries							
	<i>r</i>	p (95% CI)		<i>rho</i>	p	<i>tau</i>	p	<i>r</i>	p (95% CI)		<i>rho</i>	p	<i>tau</i>	p
Pyramid	NaN	NaN (NaN - NaN)		NaN	NaN	NaN	1.000	NaN	NaN (NaN - NaN)		NaN	NaN	NaN	1.000
Responsiveness	NaN	NaN (NaN - NaN)		NaN	NaN	NaN	1.000	NaN	NaN (NaN - NaN)		NaN	NaN	NaN	1.000
Readability	NaN	NaN (NaN - NaN)		NaN	NaN	NaN	1.000	NaN	NaN (NaN - NaN)		NaN	NaN	NaN	1.000

Table 1: System-level correlations (Pearson’s *r*, Spearman’s *rho*, Kendall’s *tau*) for uOttawa3, initial and update summaries, all peers.

	Initial summaries						Update summaries							
	<i>r</i>	p (95% CI)		<i>rho</i>	p	<i>tau</i>	p	<i>r</i>	p (95% CI)		<i>rho</i>	p	<i>tau</i>	p
Pyramid	0.895	0.000 (0.822 - 0.939)		0.831	0.000	0.671	0.000	0.781	0.000 (0.644 - 0.869)		0.684	0.000	0.518	0.000
Responsiveness	0.868	0.000 (0.778 - 0.923)		0.704	0.000	0.538	0.000	0.761	0.000 (0.614 - 0.857)		0.637	0.000	0.487	0.000
Readability	0.688	0.000 (0.508 - 0.810)		0.366	0.008	0.269	0.006	0.543	0.000 (0.314 - 0.712)		0.293	0.037	0.194	0.047

Table 2: System-level correlations (Pearson’s *r*, Spearman’s *rho*, Kendall’s *tau*) for uOttawa3, initial and update summaries, no models.

	Initial summaries			Update summaries		
	all	avg per topic	avg per assessor	all	avg per topic	avg per assessor
Pyramid	0.465	0.591	0.528	0.332	0.487	0.381
Responsiveness	0.394	0.505	0.497	0.318	0.368	0.358
Readability	0.308	0.387	0.371	0.232	0.277	0.280

Table 3: Summary-level correlations (Pearson’s *r*) for uOttawa3, initial and update summaries, no models.

		A>B	A=B	A<B
		uOttawa3	A>B	0
	A=B	408	0	
	A>B			0

Table 4: Discriminative power for uOttawa3 compared with Pyramid. Initial summaries, models vs. non-models.

		A>B	A=B	A<B
		uOttawa3	A>B	0
	A=B	408	0	
	A>B			0

Table 5: Discriminative power for uOttawa3 compared with Pyramid. Update summaries, models vs. non-models.

		A>B	A=B	A<B
		uOttawa3	A>B	0
	A=B	408	0	
	A>B			0

Table 6: Discriminative power for uOttawa3 compared with Responsiveness. Initial summaries, models vs. non-models.

		A>B	A=B	A<B
		uOttawa3	A>B	0
	A=B	408	0	
	A>B			0

Table 7: Discriminative power for uOttawa3 compared with Responsiveness. Update summaries, models vs. non-models.

		A>B	A=B	A<B
		uOttawa3	A>B	0
	A=B	407	1	
	A>B			0

Table 8: Discriminative power for uOttawa3 compared with Readability. Initial summaries, models vs. non-models.

		A>B	A=B	A<B
		uOttawa3	A>B	0
	A=B	408	0	
	A>B			0

Table 9: Discriminative power for uOttawa3 compared with Readability. Update summaries, models vs. non-models.

		A>B	A=B	A<B
uOttawa3	A>B	226	228	
	A=B	13	808	
	A>B			0

Table 10: Discriminative power for uOttawa3 compared with Pyramid. Initial summaries, non-model summaries.

		A>B	A=B	A<B
uOttawa3	A>B	154	193	
	A=B	33	895	
	A>B			0

Table 11: Discriminative power for uOttawa3 compared with Pyramid. Update summaries, non-model summaries.

		A>B	A=B	A<B
uOttawa3	A>B	208	246	
	A=B	13	808	
	A>B			0

Table 12: Discriminative power for uOttawa3 compared with Responsiveness. Initial summaries, non-model summaries.

		A>B	A=B	A<B
uOttawa3	A>B	107	240	
	A=B	21	907	
	A>B			0

Table 13: Discriminative power for uOttawa3 compared with Responsiveness. Update summaries, non-model summaries.

		A>B	A=B	A<B
uOttawa3	A>B	246	202	
	A=B	162	659	
	A>B			6

Table 14: Discriminative power for uOttawa3 compared with Readability. Initial summaries, non-model summaries.

		A>B	A=B	A<B
uOttawa3	A>B	157	187	
	A=B	165	763	
	A>B			3

Table 15: Discriminative power for uOttawa3 compared with Readability. Update summaries, non-model summaries.