

Supplemental Digital Content B: Organizational Context Scores at Baseline (n=201) and Post (n=140) in 15 BMW Practices

	Office Culture (%)		Office Structure / Processes (%)		Office Climate (%)		Office Technology (%)		IRA ($r_{WG(J)}$)*		CVF: Blau Index**	CVF: Prop. Grp/Dev**
	Baseline (n=15)	Post (n=14)	Baseline (n=15)	Post (n=14)	Baseline (n=15)	Post (n=14)	Baseline (n=15)	Post (n=14)	Baseline (n=15)	Post (n=14)	Baseline (n=15)	Baseline (n=15)
1	63.6	84.0	58.0	87.9	68.8	77.2	44.0	85.0	.574	.865	0.62	0.70
2	72.7	86.6	62.0	83.4	60.0	67.1	55.0	76.0	.412	.745	0.55	0.76
3	59.5	78.0	44.0	72.5	61.6	65.0	56.0	73.3	.506	.635	0.70	0.60
4	61.1	78.2	57.6	78.8	55.2	61.8	52.5	90.0	.682	.883	0.65	0.41
5	68.5	78.4	70.5	77.9	65.4	77.1	89.1	83.3	.899	.917	0.68	0.53
6	84.8	75.1	74.2	81.9	76.9	82.5	65.7	88.0	.884	.915	0.67	0.60
7	70.5	82.8	68.9	84.5	66.7	72.1	72.5	80.0	.615	.775	0.66	0.53
8	77.8	N/A	46.7	N/A	74.1	N/A	26.7	N/A	.878	N/A	0.48	0.83
9	83.1	86.3	72.6	87.3	73.3	77.8	72.0	100	.940	.958	0.63	0.62
10	67.3	75.2	62.2	73.4	66.1	71.0	62.8	80.0	.850	.889	0.66	0.65
11	65.4	72.8	64.6	71.8	60.6	64.9	56.0	82.8	.666	.695	0.69	0.53
12	51.4	63.3	53.4	71.3	57.6	60.4	30.0	60.0	.878	.802	0.61	0.52
13	65.0	67.4	66.4	71.8	60.9	62.2	75.7	84.4	.896	.768	0.61	0.54
14	46.9	76.5	42.9	77.6	57.5	77.4	30.0	80.0	.204	.901	0.59	0.55
15	85.8	93.7	76.6	92.2	67.3	63.3	92.0	95.0	.583	.915	0.74	0.56
Mean	68.2	78.4	61.4	79.4	64.8	70.0	58.7	82.7	.698	.833		

* For each practice, inter-rater agreement ($r_{WG(J)}$) based on the uniform theoretical null distribution was calculated for each domain using the $r_{WG(J)}$ index. It ranges in value from 0.00 (no agreement) to 1.00 (perfect agreement). Scores presented here represent the average across domains and are based on the uniform theoretical null distribution (i.e. assuming no systematic bias in rating given participants' anonymity).

** All staff rated their practice culture using competing values framework (CVF). Uniformity of distribution of points across the four values ("Blau Index") and proportion of points assigned to developmental and group types ("Prop. Grp/Dev") were calculated for each respondent. Practice scores represent the average across individual respondent scores for each practice.