

Table 3 Statistical of the  $F$ -test and the Wilcoxon test Results

Semantic :  $\alpha = \beta = \gamma = 0$ ,  $\eta = 1.0$  and  $Th_1 = SSW_{med}$ ;

Structural:  $\alpha = 0.7$ ,  $\beta = 0.2$ ,  $\gamma = 0.1$ ,  $\eta = 0$  and  $Th_1 = SSW_{med}$ ;

Combined  $\alpha = 0.5$ ,  $\beta = 0.2$ ,  $\gamma = 0.2$ ,  $\eta = 0.1$  and  $Th_1 = SSW_{med}$ ;

System	method	mean	median	Std.dev	Compared method	p-Value	Effect-size
Elasticsearch	Semantic	0.4135	0.4219	0.1003	Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
	Structural	0.8499	0.8510	0.0146	Combined vs Semantic	<b>&lt;0.01</b>	0.9(High)
	Combined	0.9060	0.8709	0.0718	Semantic vs Structural	<b>&lt;0.01</b>	0.9(High)
Android-async-http	Semantic	0.4530	0.4406	0.1216	Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
	Structural	0.8710	0.8938	0.1150	Combined vs Semantic	<b>&lt;0.01</b>	0.9(High)
	Combined	0.8976	0.8911	0.1544	Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Iosched	Semantic	0.4391	0.4620	0.1107	Combined vs Structural	0.0129	0.7(Medium)
	Structural	0.8312	0.8533	0.1322	Combined vs Semantic	<b>&lt;0.01</b>	0.9(High)
	Combined	0.9156	0.9218	0.1200	Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Libgdx	Semantic	0.4518	0.4262	0.0097	Combined vs Structural	<b>&lt;0.01</b>	0.6(Medium)
	Structural	0.8921	0.8654	0.1249	Combined vs Semantic	<b>&lt;0.01</b>	0.9(High)
	Combined	0.9077	0.9178	0.1093	Semantic vs Structural	<b>&lt;0.01</b>	0.9(High)
Android-annotations	Semantic	0.4382	0.4456	0.1410	Combined vs Structural	0.0152	0.7(Medium)
	Structural	0.8508	0.8621	0.1574	Combined vs Semantic	<b>&lt;0.01</b>	0.9(High)
	Combined	0.9123	0.9315	0.1308	Semantic vs Structural	<b>&lt;0.01</b>	0.9(High)
Okhttp	Semantic	0.4581	0.4758	0.1210	Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
	Structural	0.8798	0.8787	0.1205	Combined vs Semantic	<b>&lt;0.01</b>	0.9(High)
	Combined	0.8808	0.8787	0.1218	Semantic vs Structural	<b>&lt;0.01</b>	0.9(High)
RxJava	Semantic	0.4819	0.4787	0.1206	Combined vs Structural	<b>0.0209</b>	0.7(Medium)
	Structural	0.8849	0.8787	0.1259	Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
	Combined	0.9489	0.9478	0.1248	Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)

System	method	mean	median	Std.dev
Spring-framework	Semantic	0.4838	0.4787	0.1270
	Structural	0.8797	0.8788	0.1255
	Combined	0.9287	0.9445	0.1208
Zxing	Semantic	0.4452	0.4505	0.1212
	Structural	0.8385	0.8223	0.0919
	Combined	0.8945	0.8861	0.1167
MPAndroidChart	Semantic	0.4326	0.4484	0.1189
	Structural	0.8207	0.8489	0.1238
	Combined	0.8938	0.8868	0.1225
Netty	Semantic	0.4332	0.4407	0.1242
	Structural	0.8482	0.8907	0.1293
	Combined	0.9002	0.8896	0.1341
Guava	Semantic	0.4384	0.4474	0.1366
	Structural	0.8521	0.8236	0.1349
	Combined	0.8958	0.8667	0.1451
Junit	Semantic	0.4372	0.4593	0.1196
	Structural	0.8883	0.8893	0.1236
	Combined	0.9488	0.8993	0.1208
Presto	Semantic	0.4501	0.4593	0.1331
	Structural	0.8017	0.8483	0.1333
	Combined	0.9049	0.9095	0.1381
Glide	Semantic	0.4483	0.4533	0.1371
	Structural	0.8202	0.8176	0.1371
	Combined	0.8883	0.8862	0.1434

Compared method	p-Value	Effect-size
Combined vs Structural	<b>&lt;0.01</b>	0.6(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.9(High)
Combined vs Structural	<b>0.0100</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>&lt;0.01</b>	0.6(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>0.0180</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)

System	method	mean	median	Std.dev
Disruptor	Semantic	0.4379	0.4593	0.1202
	Structural	0.8933	0.8893	0.1256
	Combined	0.8975	0.8893	0.1310
Stetho	Semantic	0.4500	0.4496	0.1325
	Structural	0.8058	0.8435	0.2257
	Combined	0.9073	0.8856	0.1999
Titan	Semantic	0.4314	0.4159	0.1674
	Structural	0.8801	0.8702	0.0932
	Combined	0.8919	0.8893	0.2253
Fastjson	Semantic	0.4369	0.4465	0.207
	Structural	0.8694	0.8496	0.1724
	Combined	0.9504	0.9566	0.0774
AndroidAsync	Semantic	0.4172	0.4432	0.2006
	Structural	0.8515	0.8453	0.1739
	Combined	0.8958	0.9053	0.0836
gradle	Semantic	0.4181	0.4489	0.1662
	Structural	0.8478	0.8499	0.0839
	Combined	0.9059	0.8932	0.0801
HoloEverywhere	Semantic	0.4209	0.4429	0.2231
	Structural	0.8406	0.8353	0.1975
	Combined	0.8931	0.8912	0.1535
Storm	Semantic	0.4486	0.4496	0.0935
	Structural	0.8143	0.8432	0.1997
	Combined	0.9174	0.8968	0.1555

Compared method	p-Value	Effect-size
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>0.0180</b>	0.6(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>&lt;0.01</b>	0.6(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>&lt;0.01</b>	0.6(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)

System	method	mean	median	Std.dev
Guice	Semantic	0.4258	0.4491	0.0777
	Structural	0.8992	0.8712	0.1575
	Combined	0.9115	0.8929	0.0855
Dubbo	Semantic	0.4209	0.4468	0.0809
	Structural	0.8643	0.8529	0.2023
	Combined	0.9028	0.9088	0.1477
Orientdb	Semantic	0.4204	0.4492	0.0928
	Structural	0.8631	0.8543	0.1492
	Combined	0.8964	0.9083	0.084
Neo4j	Semantic	0.4124	0.4516	0.0805
	Structural	0.8918	0.8705	0.1538
	Combined	0.8912	0.9069	0.0886
jna	Semantic	0.4112	0.4428	0.0766
	Structural	0.8958	0.8898	0.1001
	Combined	0.8988	0.8898	0.2257
Smile	Semantic	0.4035	0.4424	0.1999
	Structural	0.8037	0.8494	0.1674
	Combined	0.8908	0.8964	0.0932
Tachyon	Semantic	0.4092	0.4054	0.2253
	Structural	0.8197	0.8496	0.207
	Combined	0.8991	0.8962	0.1724
ExoPlayer	Semantic	0.4164	0.4268	0.0774
	Structural	0.8943	0.8907	0.2006
	Combined	0.8983	0.8924	0.1739

Compared method	p-Value	Effect-size
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	0.0101	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	0.0122	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.9(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.9(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>&lt;0.01</b>	0.6(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)

System	method	mean	median	Std.dev
Grails	Semantic	0.4332	0.4424	0.1902
	Structural	0.8079	0.8492	0.1068
	Combined	0.9125	0.8964	0.1342
Mockito	Semantic	0.4139	0.4511	0.1329
	Structural	0.8266	0.8496	0.1498
	Combined	0.9084	0.9271	0.1569
Mongo-java-driver	Semantic	0.4411	0.4769	0.1914
	Structural	0.8933	0.8929	0.1851
	Combined	0.8992	0.8924	0.1021
Hibernate-orm	Semantic	0.4067	0.4468	0.1409
	Structural	0.8142	0.8032	0.1499
	Combined	0.9162	0.8953	0.1685
Deeplearning4j	Semantic	0.4299	0.4508	0.2012
	Structural	0.8618	0.8278	0.18
	Combined	0.8947	0.9075	0.1055
Hazelcast	Semantic	0.4003	0.4568	0.1529
	Structural	0.8503	0.8964	0.1764
	Combined	0.9113	0.9032	0.2132
Jadx	Semantic	0.4195	0.4099	0.2016
	Structural	0.8396	0.8076	0.0943
	Combined	0.9253	0.9245	0.1923
Blueprints	Semantic	0.4207	0.4319	0.2076
	Structural	0.8053	0.8496	0.1868
	Combined	0.9098	0.9032	0.1053

Compared method	p-Value	Effect-size
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	0.0142	0.6(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.9(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.9(High)
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.9(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.9(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>&lt;0.01</b>	0.6(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)

System	method	mean	median	Std.dev
Quasar	Semantic	0.4229	0.4512	0.2035
	Structural	0.8549	0.8507	0.1841
	Combined	0.9224	0.9754	0.1102
Cucumber-jvm	Semantic	0.4014	0.4135	0.1898
	Structural	0.8123	0.8544	0.1091
	Combined	0.9294	0.9555	0.1271
Webmagic	Semantic	0.4547	0.4506	0.134
	Structural	0.8418	0.8199	0.1402
	Combined	0.8979	0.9268	0.1468
Zookeeper	Semantic	0.4331	0.4568	0.1736
	Structural	0.8557	0.8496	0.1931
	Combined	0.9648	0.9648	0.176
ninja	Semantic	0.4587	0.4284	0.0938
	Structural	0.8217	0.8256	0.1393
	Combined	0.8773	0.9335	0.1566
Rundeck	Semantic	0.4081	0.4089	0.1836
	Structural	0.8416	0.8495	0.1839
	Combined	0.8969	0.9077	0.1696
Moco	Semantic	0.4284	0.4626	0.0751
	Structural	0.8736	0.8437	0.1617
	Combined	0.8926	0.8743	0.1888
Flyway	Semantic	0.4463	0.4612	0.1987
	Structural	0.8782	0.8612	0.1730
	Combined	0.8994	0.8719	0.0765

Compared method	p-Value	Effect-size
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.9(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.9(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.9(High)
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.9(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)

System	method	mean	median	Std.dev
Flume	Semantic	0.4008	0.4119	0.1946
	Structural	0.8437	0.8508	0.2015
	Combined	0.9206	0.9562	0.1744
Astyanax	Semantic	0.4266	0.4743	0.0727
	Structural	0.8778	0.8749	0.2011
	Combined	0.8803	0.9474	0.1673
Syncany	Semantic	0.4327	0.4731	0.0688
	Structural	0.8916	0.8737	0.1607
	Combined	0.9294	0.8975	0.0764

Compared method	p-Value	Effect-size
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.9(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.8(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.9(High)
Combined vs Structural	<b>&lt;0.01</b>	0.7(Medium)
Combined vs Semantic	<b>&lt;0.01</b>	0.9(High)
Semantic vs Structural	<b>&lt;0.01</b>	0.8(High)