Table 1: Task completion times and efficiency for each of the tasks. Highlighted cells contain values that are computed by us.

Task	Kind of specification	Time	Efficiency
B	Plain	1639	0.139
	Structured	1614	0.349
	W	0.670	-3.891
	p value	0.503	¡ 0.001
C	Plain	1220	0.109
	Structured	1648	0.239
	W/t	-1.282	3.884
	p value	0.207	0.0001
D	Plain	1911	0.139
	Structured	2109	0.146
	t	-0.582	-0.003
	p value	0.564	0.998
E	Plain	1373	0.185
	Structured	1701	0.111
	W	-1.252	1.673
	p value	0.21	0.094
Aggregated	Plain	1763.53	0.21
	Structured	1483.81	0.14
	W	-1.15	-2.85
	p	0.25	0.004

Different Result: Does Organizing Security Patterns Focus Architectural Choices?

Summary: Yskout and others conducted a large-scale study to analyze the effect of structuring security requirements on the time and efficiency to improve a given software architecture to meet a security requirement [?][1]

- Independent variable: Type of requirement specification (2 levels, operationalized as structured and plain)

- Tasks: 4 tasks to improve a given software architecture to meet a security requirement (plus a warm-up task that was not analyzed).

- Dependent variables:

  1. Task completion time (referred to as "effort") [metric scale]
  2. Efficiency (operationalized as the number of security patterns implemented, divided by the total number of available security patterns) [metric scale]

- Null hypotheses (for each of the four tasks):

  $H_0 time$: The kind of requirement specification does not affect task completion time

  $H_0 efficiency$: The kind of requirement specification does not affect efficiency

- Results:

  $H_0 time$: No significant difference for any of the tasks regarding completion time.

  $H_0 efficiency$: Significant difference for 2 of the 4 tasks.

As aggregation function, we applied the arithmetic mean to both dependent variables (i.e., task completion time, efficiency). We aggregated over all tasks (excluding the warm-up task). Before that, we removed outliers as denoted by Yskout and others [?]. We applied the Wilcoxon test to both dependent variables, as they were not normally distributed. We summarize our reanalysis in Table 1. For the task completion time, we could replicate the result of no significant difference, and come to the same conclusion as the authors. For efficiency, the aggregation yielded a significant difference, so we also reject the null hypothesis.

---

[1]Link to data: `https://distrinet.cs.kuleuven.be/software/securitypatterns/`

In general, we can confirm the results of the authors. However, we like to note that in the introduction and discussion of results, the authors were not very precise when discussing the benefit for efficiency. That is, although they defined one null hypothesis for each task, they did not restrict their discussion to the tasks for which there was a difference in efficiency, but kept the discussion very general, saying that they could find a benefit for efficiency. This can easily be interpreted that, independent of the task, a structured requirement specification leads to a higher efficiency, which would be a too strong statement. Instead, the authors should have discussed the results task-wise, allowing them a more nuanced discussion, especially since there are two tasks that seem to be better solvable with the structured requirement specification.