

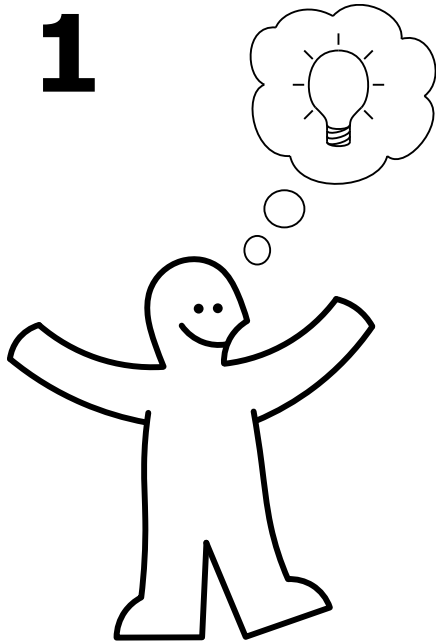
Benchmarking Flaws in Systems Security

Erik van der Kouwe, Gernot Heiser, Dennis Andriessse,
Herbert Bos, and Cristiano Giuffrida

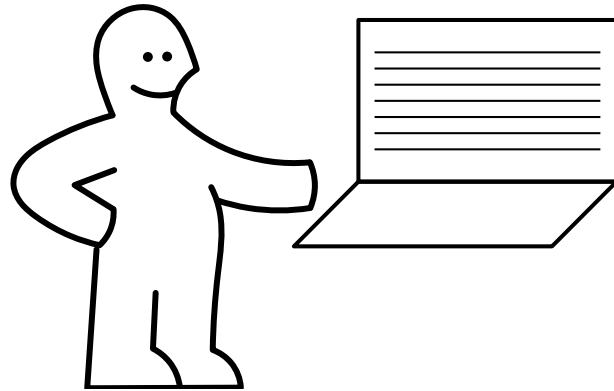
EuroS&P 2019, 18 June 2019, Stockholm, Sweden

PUBLIKATION

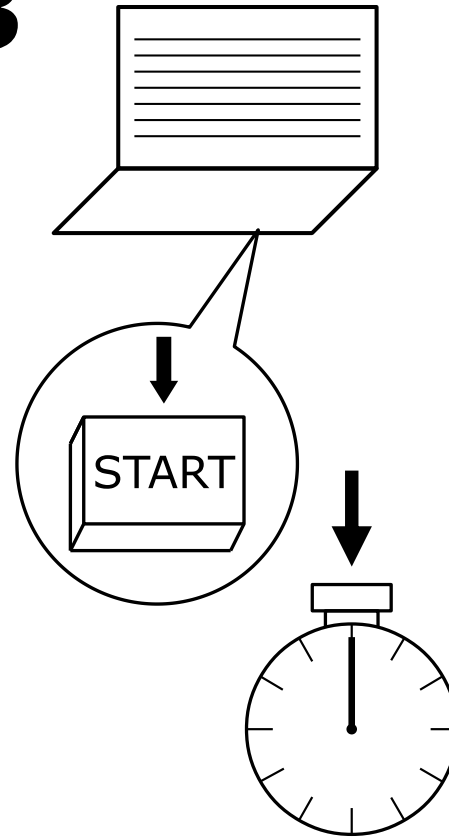
1



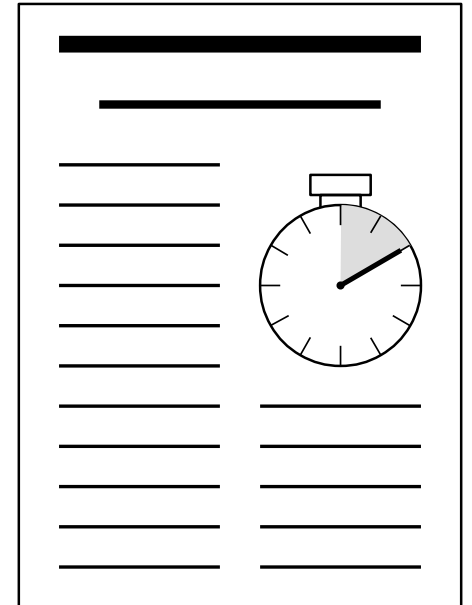
2



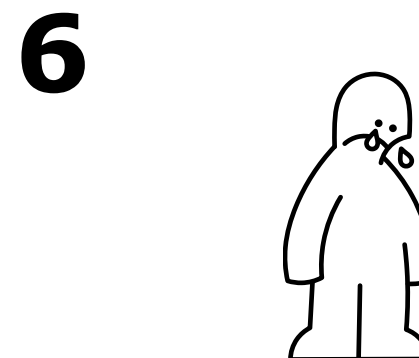
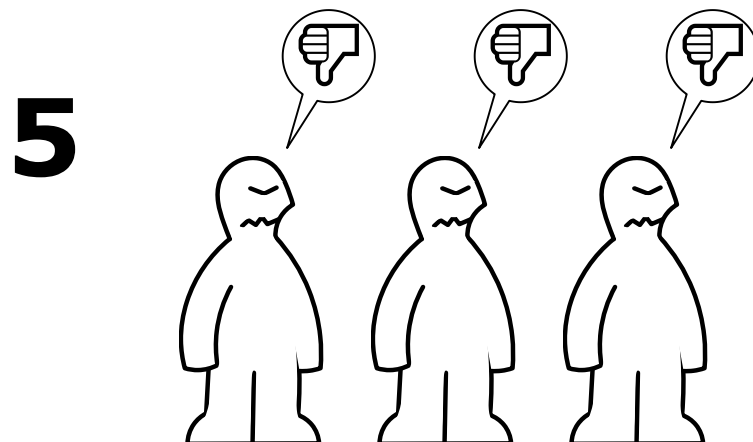
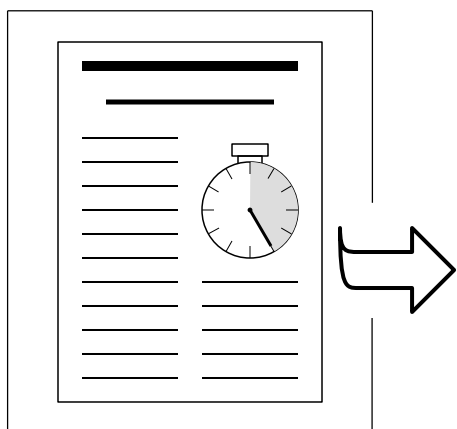
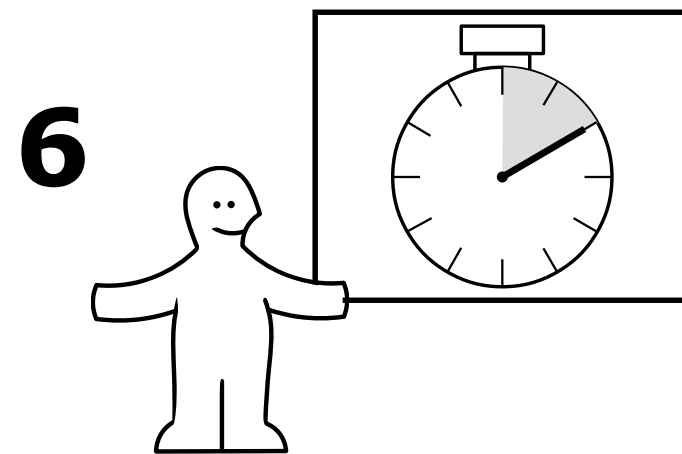
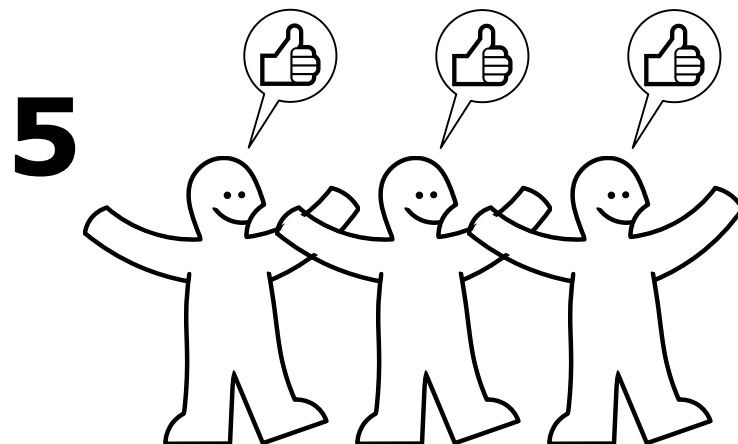
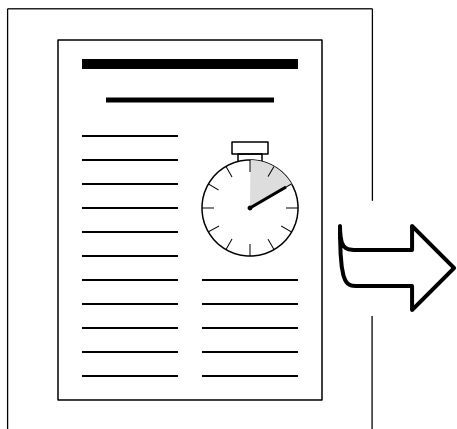
3



4

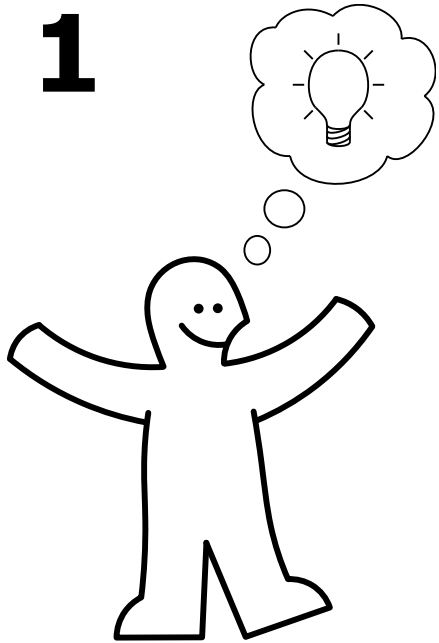


PUBLIKATION

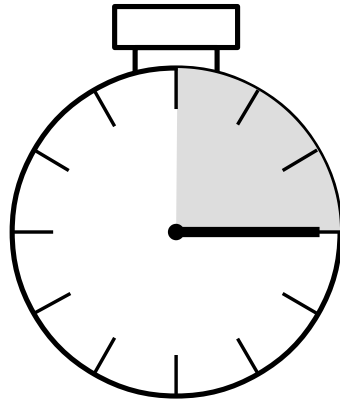


LATER WÖRK

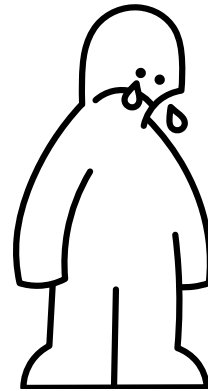
1



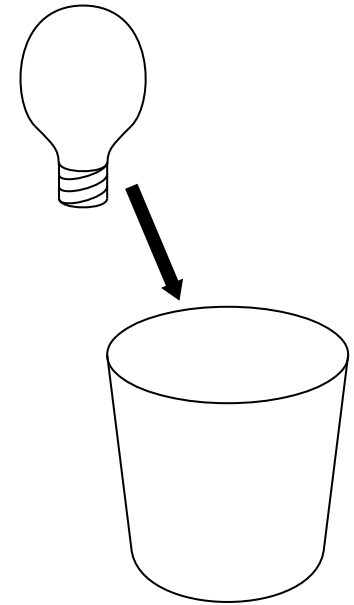
2



3

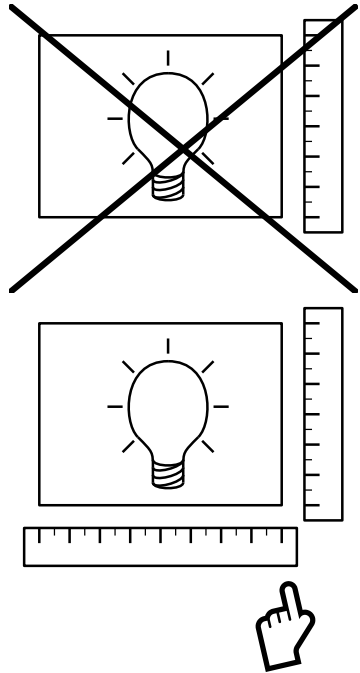


4



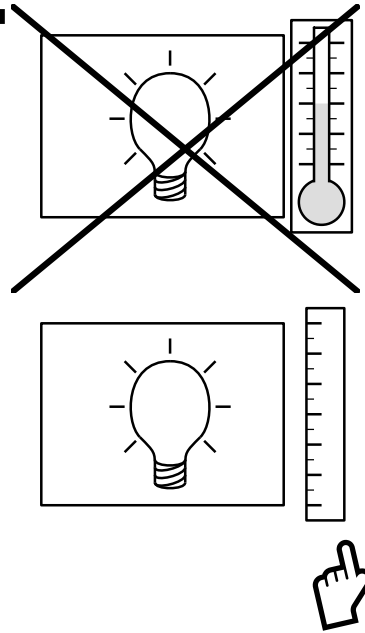
BENCHMARKÄRK

1



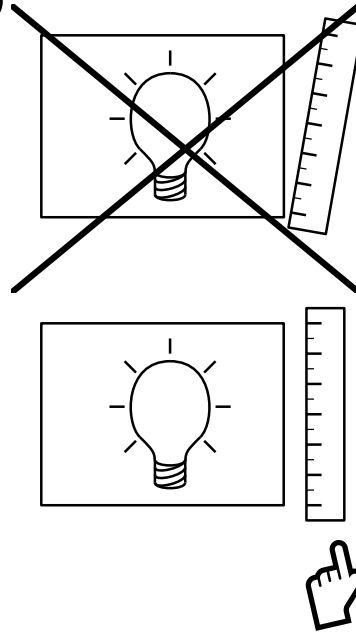
Complete

2



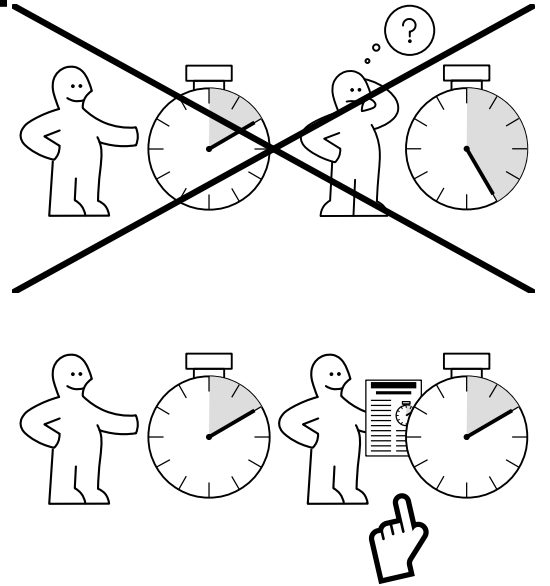
Relevant

3



Sound

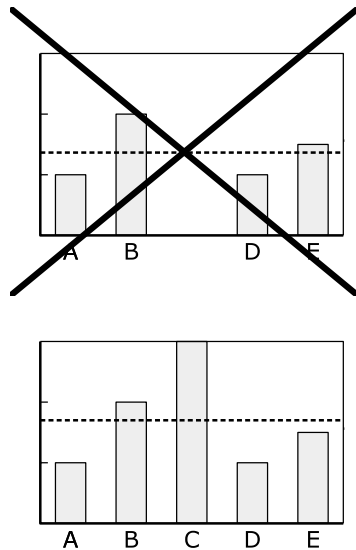
4



Reproducible

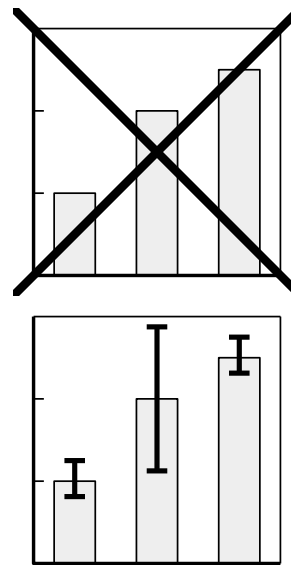
FLÅW KLASSIFIKATION

A



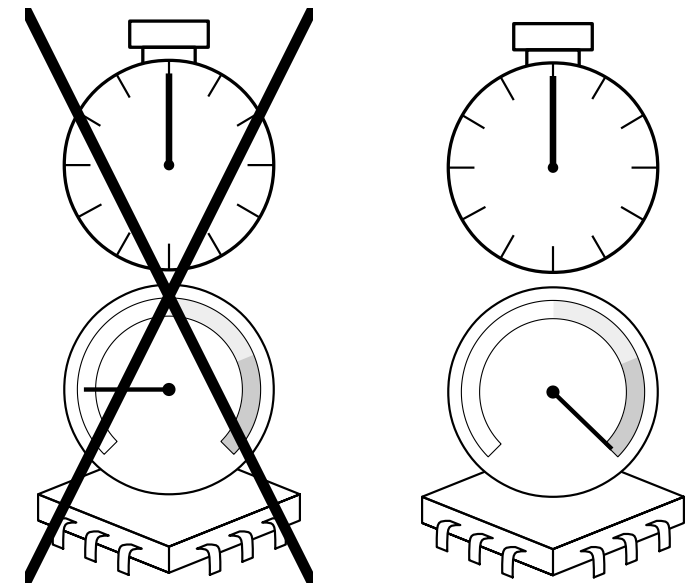
Selective benchmarking

B



Improper handling of results

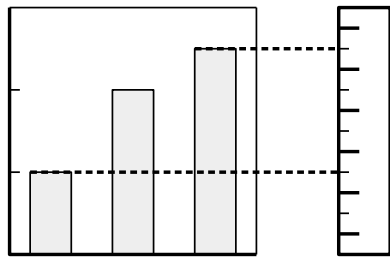
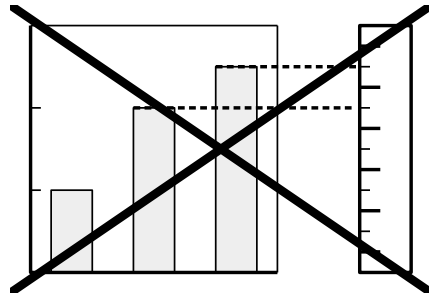
C



Wrong benchmarks

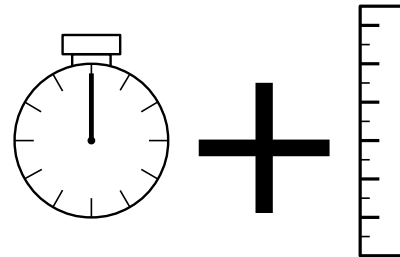
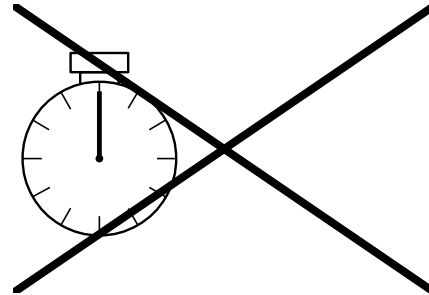
FLÅW KLASSIFIKATION

D



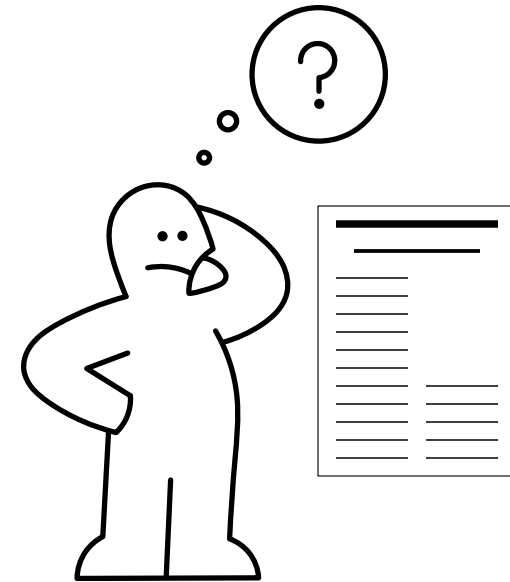
Improper
comparison

E



Benchmarking
omissions

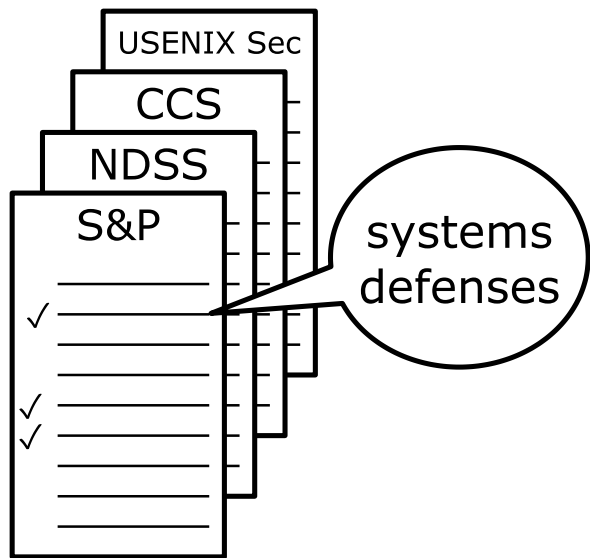
F



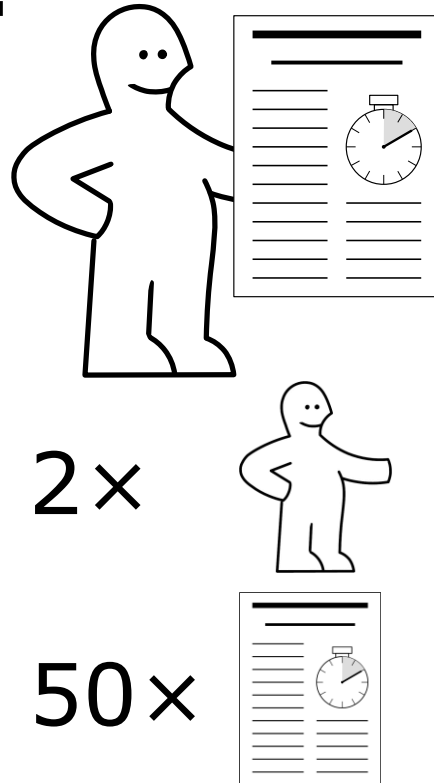
Missing
information

SURVEJ

1



2

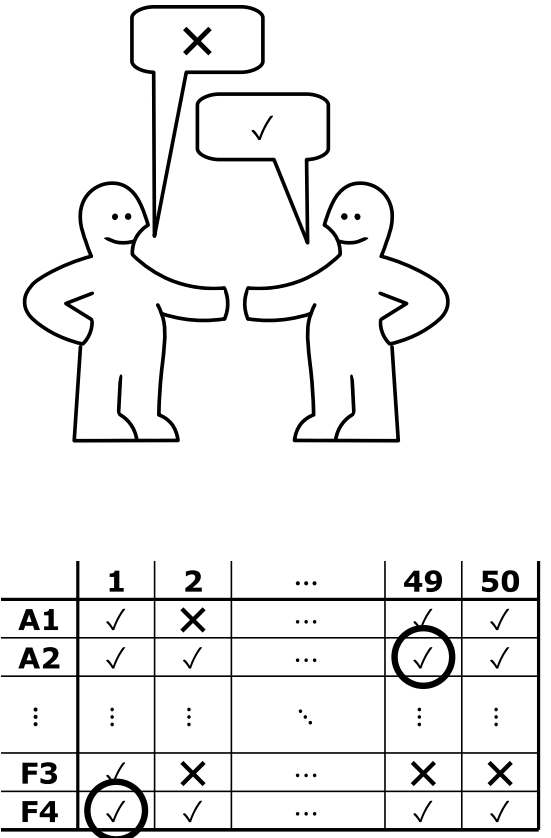


3

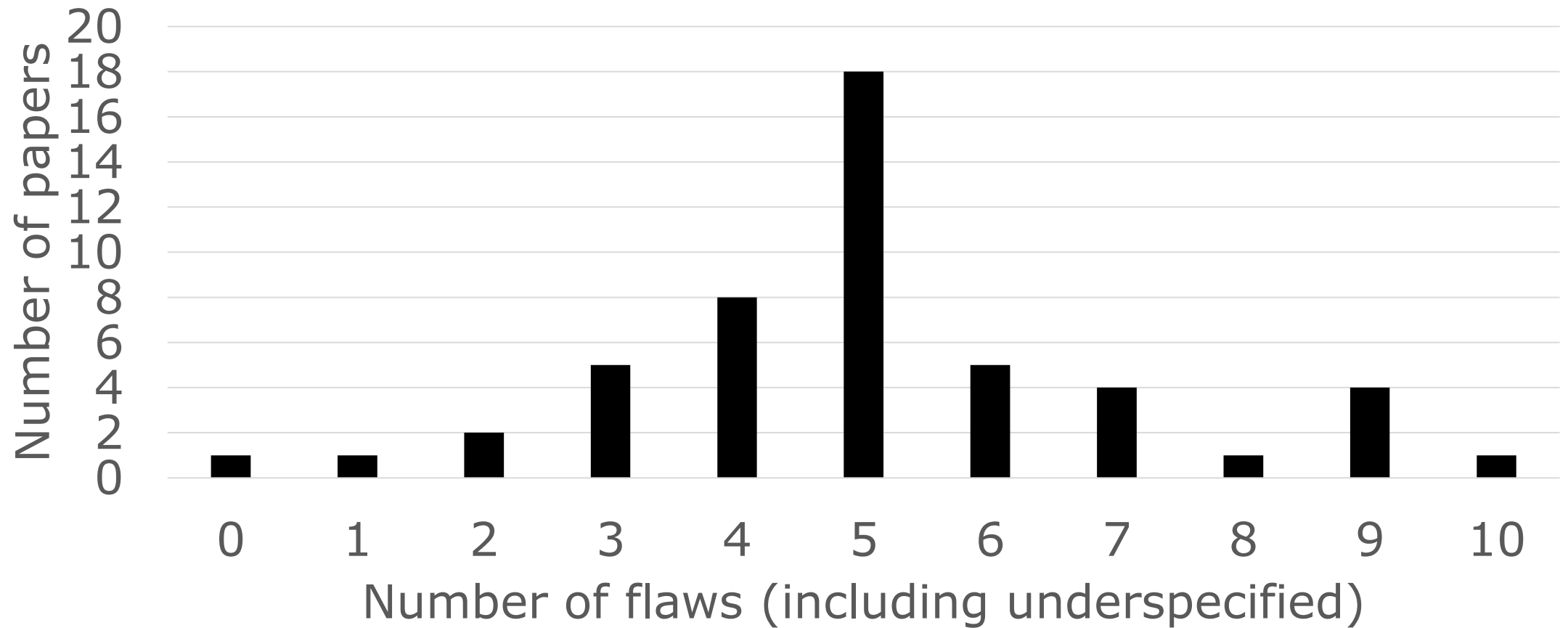
	1	2	...	49	50
A1	✓	X	...	✓	✓
A2	✓	✓	...	X	✓
⋮	⋮	⋮	⋮	⋮	⋮
F3	✓	X	...	X	X
F4	✓	✓	...	✓	✓

	1	2	...	49	50
A1	✓	X	...	✓	✓
A2	✓	✓	...	✓	✓
⋮	⋮	⋮	⋮	⋮	⋮
F3	✓	X	...	X	X
F4	X	✓	...	✓	✓

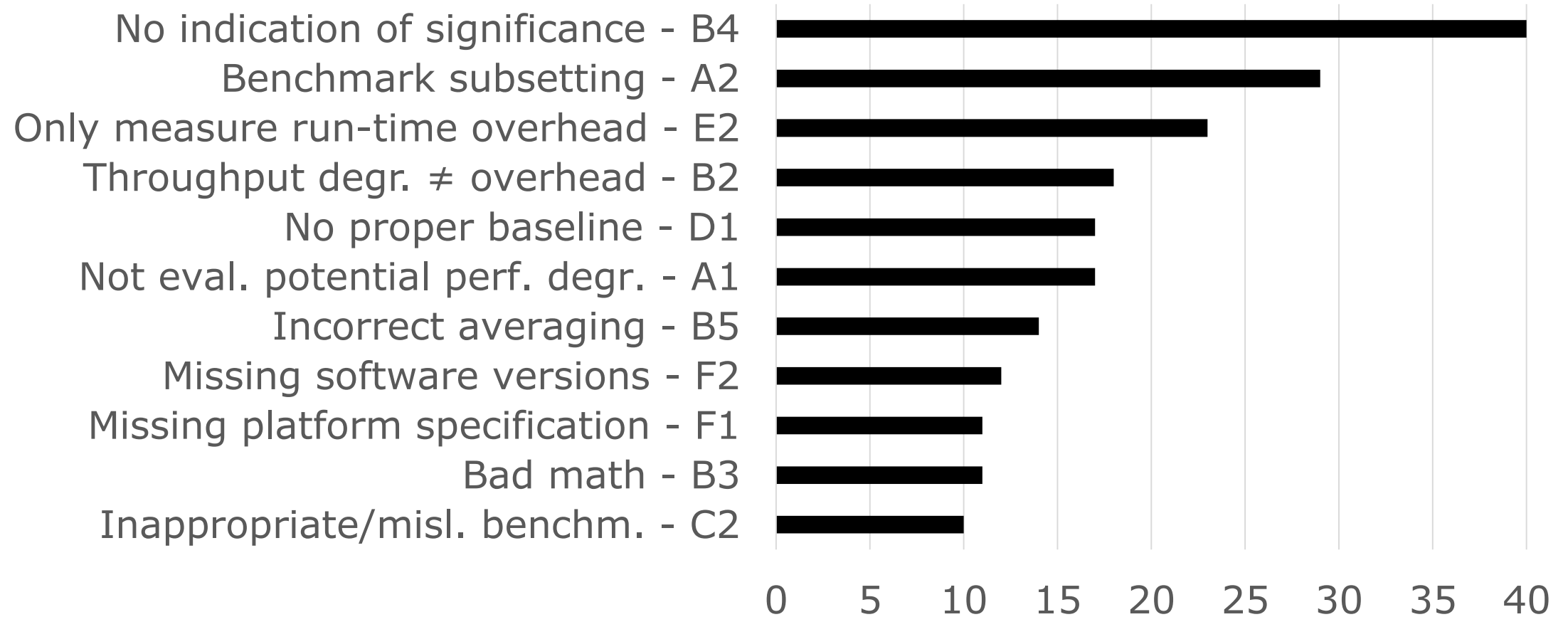
4



RESYLT

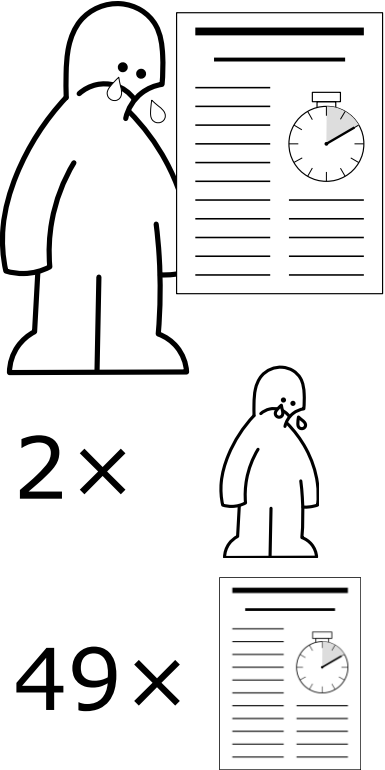


KOMMON FLÅWS

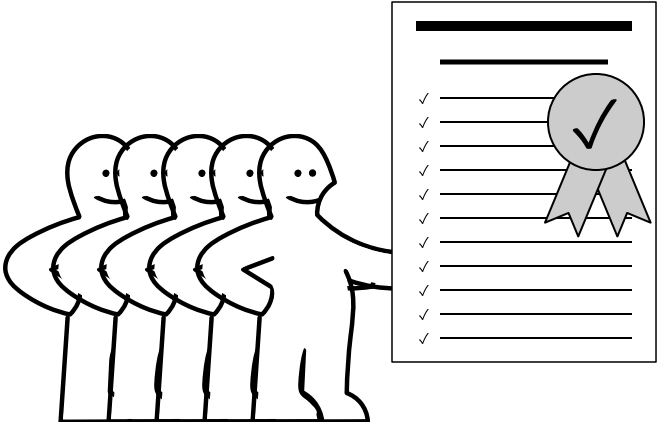


DISKUSSION

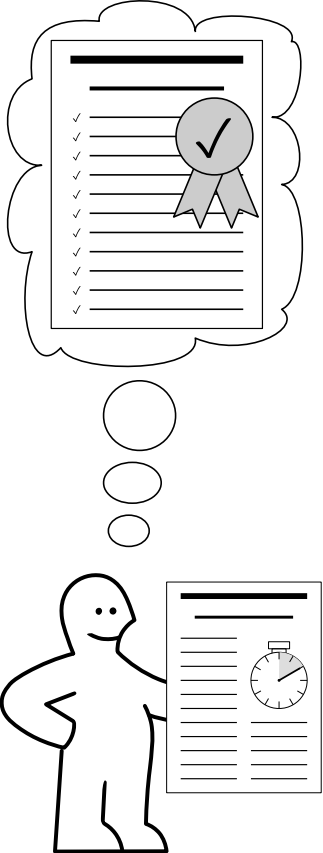
1



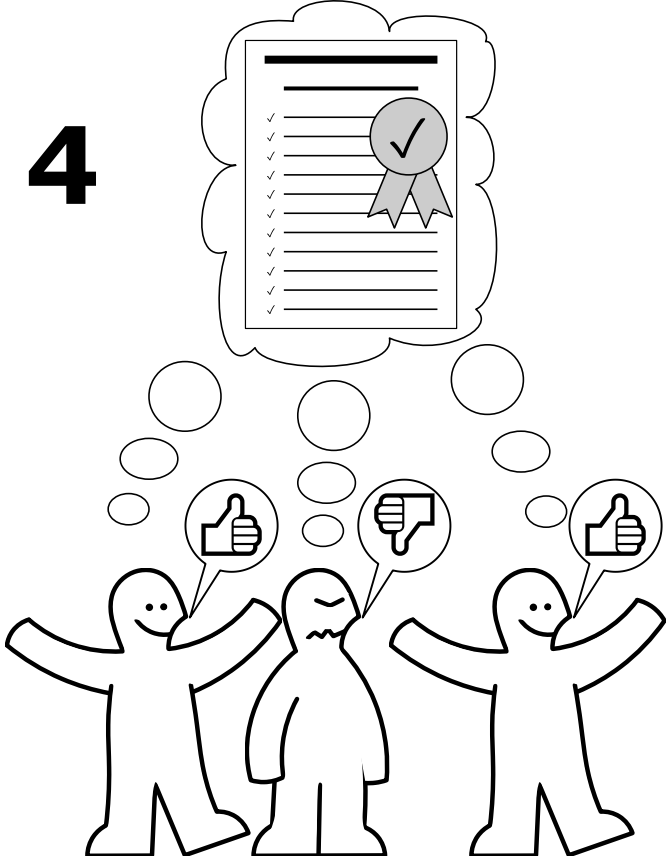
2



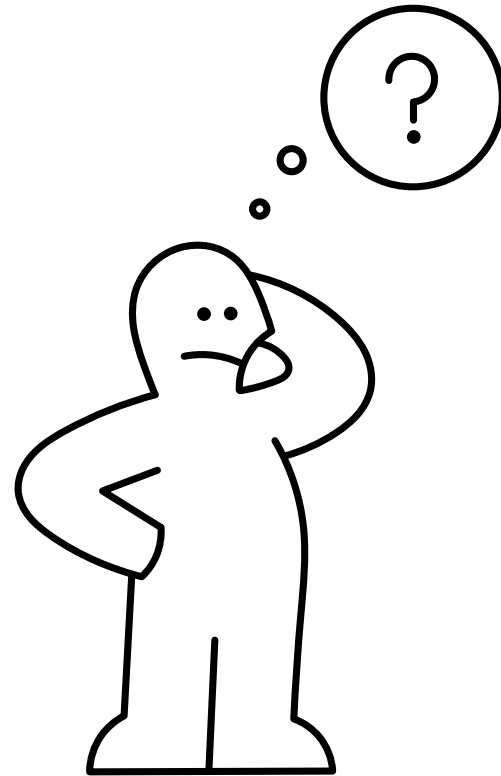
3



4



KVESTIONS?



Acknowledgement: human figures from <https://idea-instructions.com/> (CC by-nc-sa 4.0)