SATE IV Synthetic

Method and Results

Aurelien Delaitre
SATE IV Organizer

March 29, 2012

Method

Was a test case correctly analyzed?

Challenges:

- Too large for manual analysis
- Non-trivial input data

Input

Tool Warning

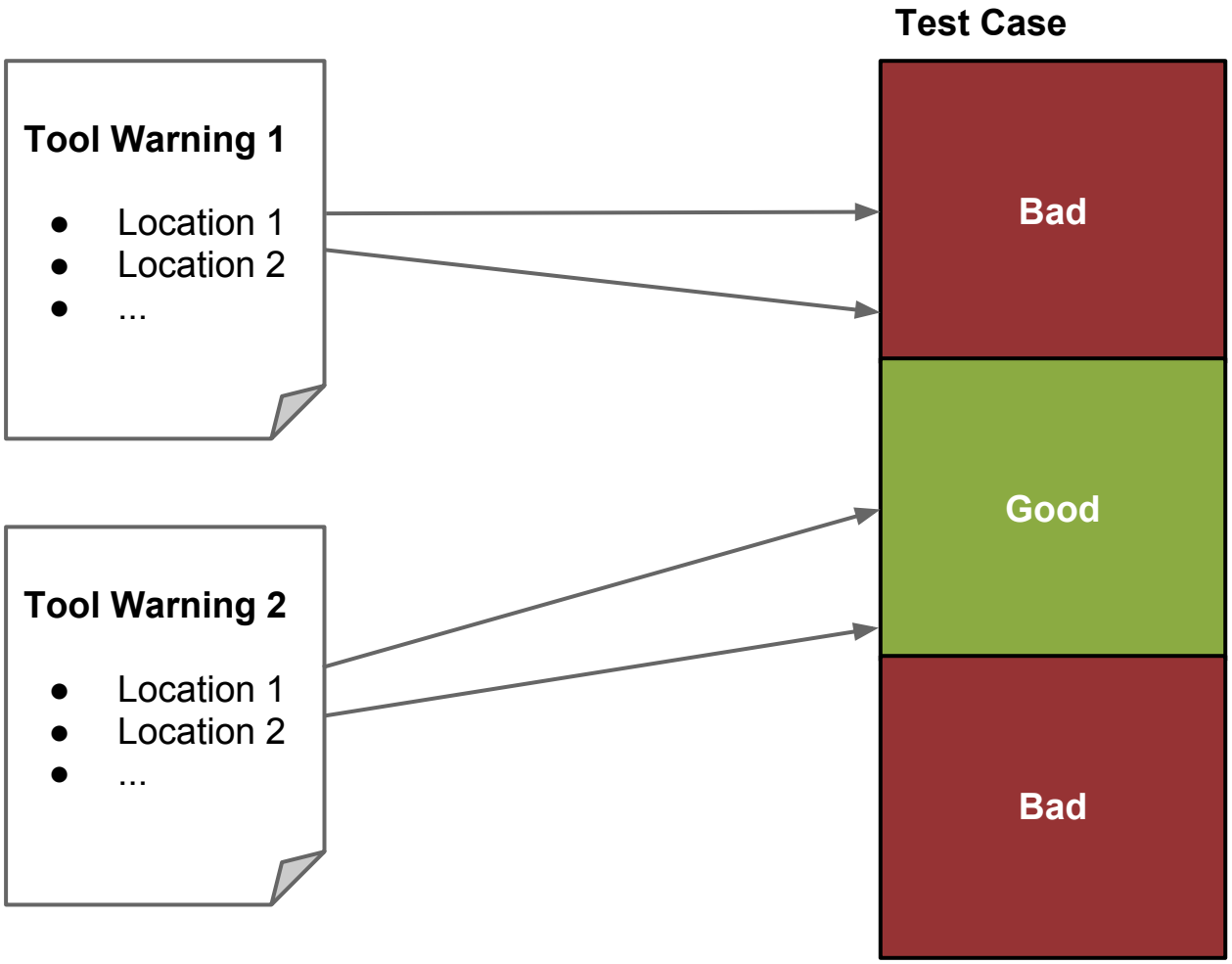
- Assigned CWE ID
- Path to sink

Test Case

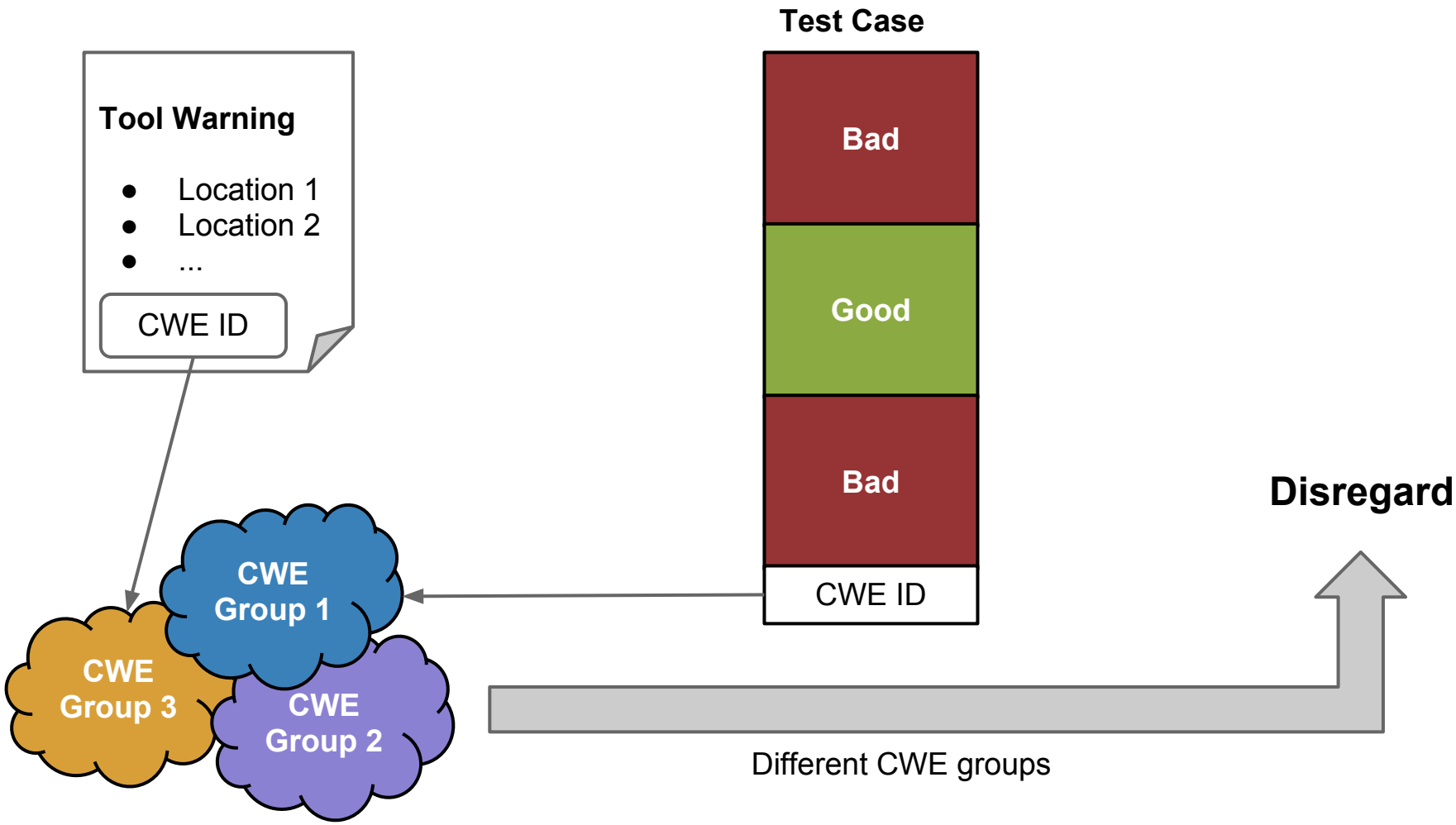
- CWE ID
- Blocks of code

CWE Matching

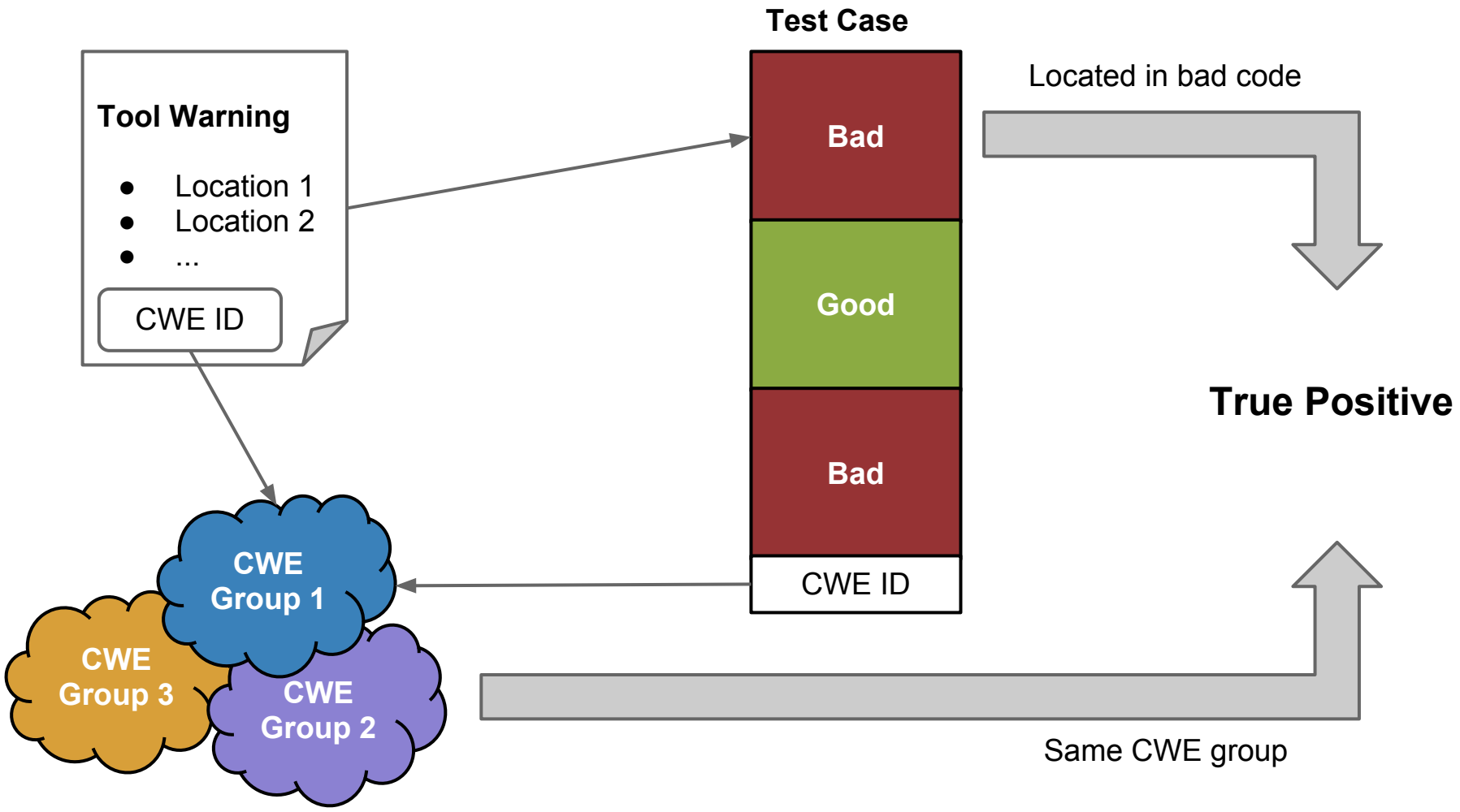
Location Matching



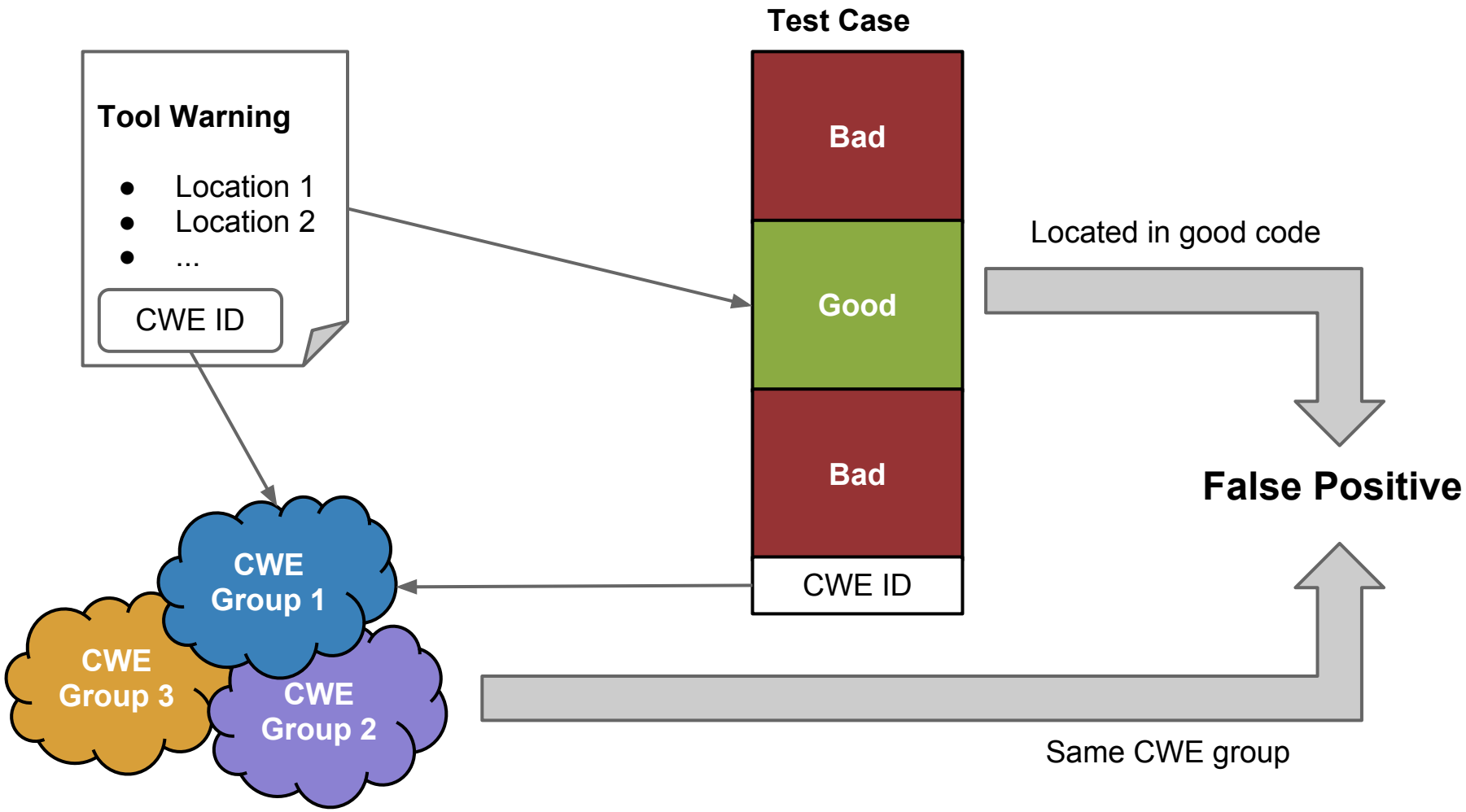
CWE + Location Supercombo



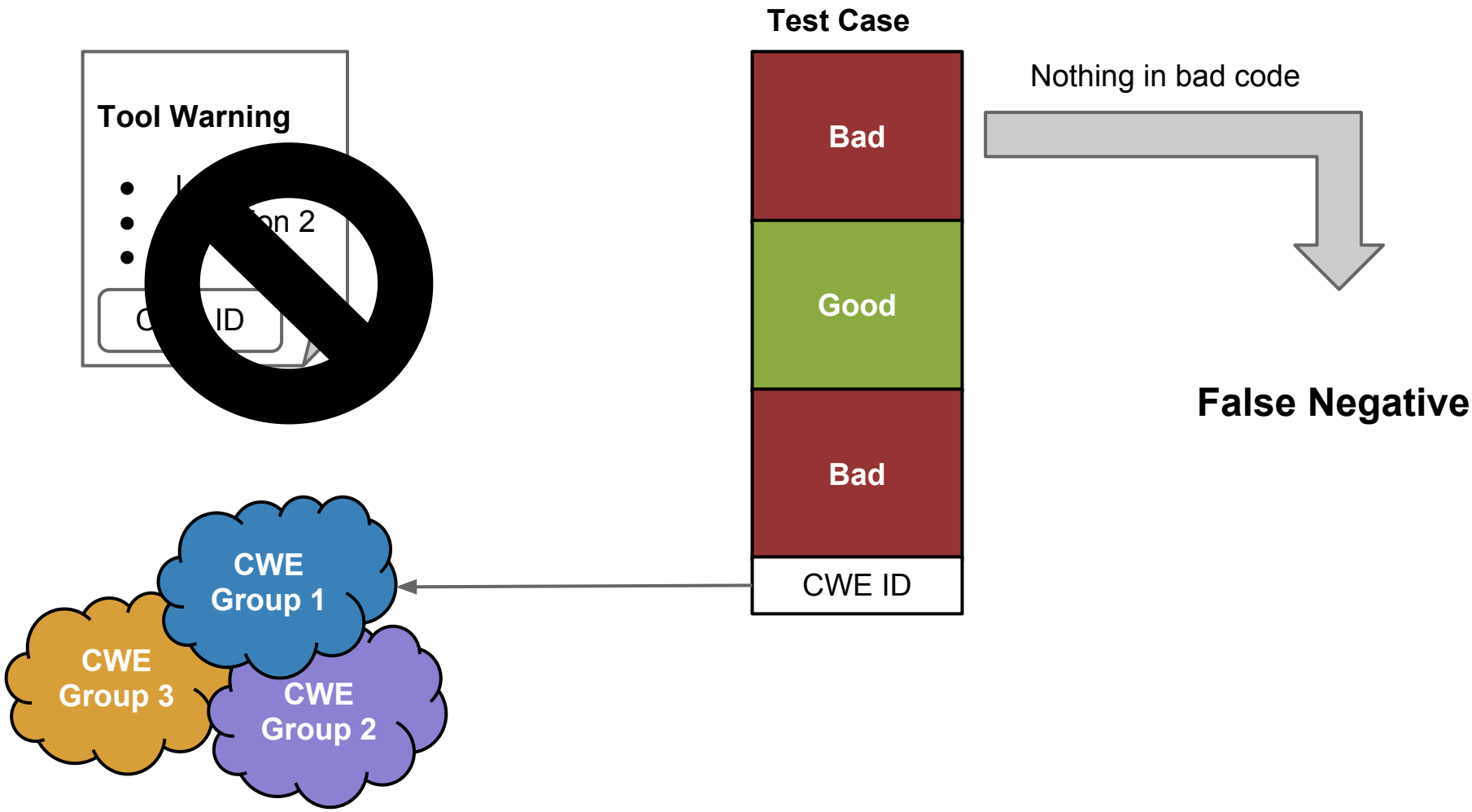
CWE + Location Supercombo



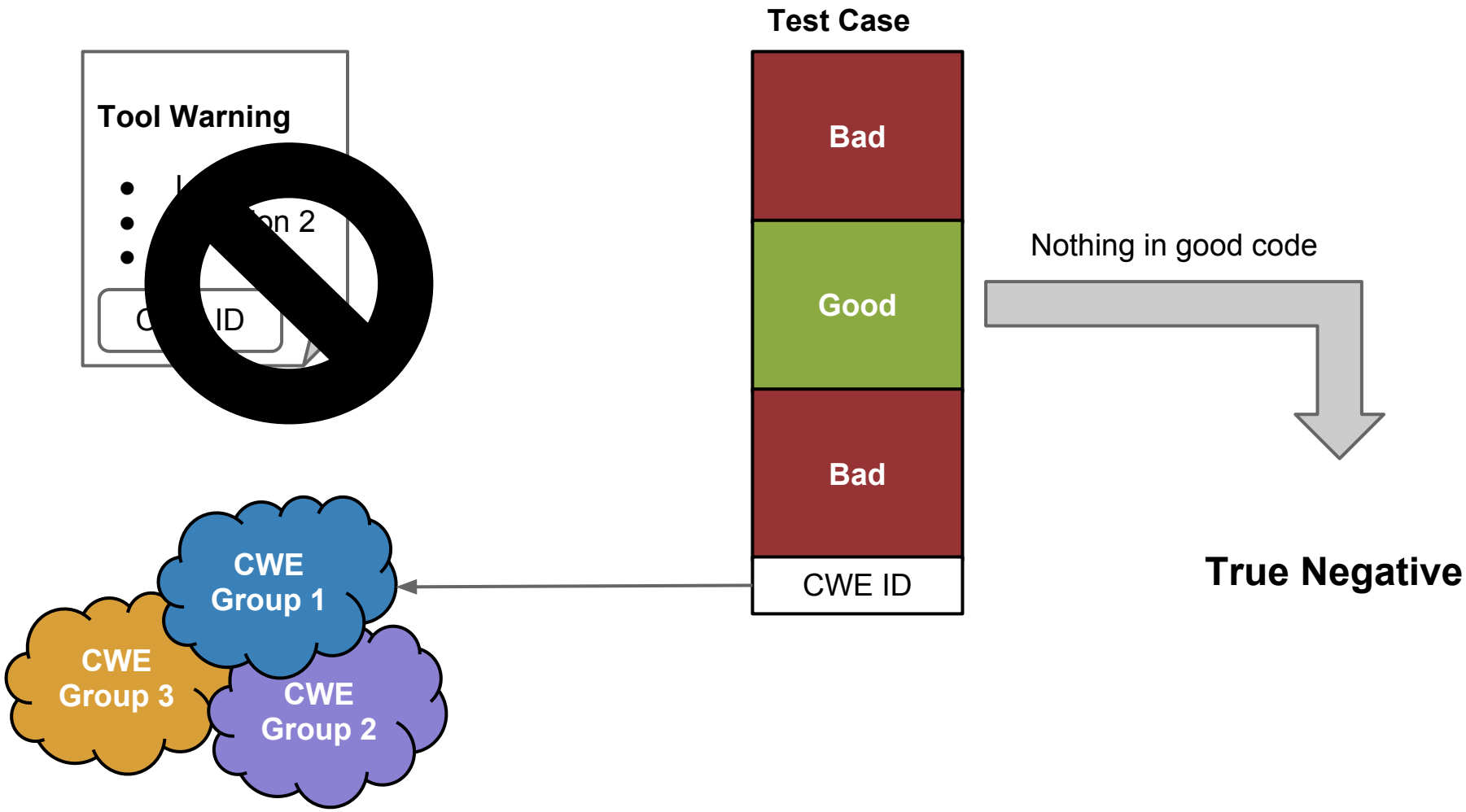
CWE + Location Supercombo



CWE + Location Supercombo



CWE + Location Supercombo



Dead Code Issue

```
if(true)
{
    // good code
}
else
{
    // bad code
}
```

Always true

Dead code

Dead Code Issue

Reporting weaknesses in dead code?

Decision:

- Rate as false-positive
- No viable alternative

CWE Assignment Issue

Based on weakness name

No perfect match between tool warnings and CWE

Remediation (in progress)

- Manual analysis of a sample
- Determination of the error ratio
- Correction if sensible

Results

What can be determined using synthetic test cases?

Discrimination and Coverage

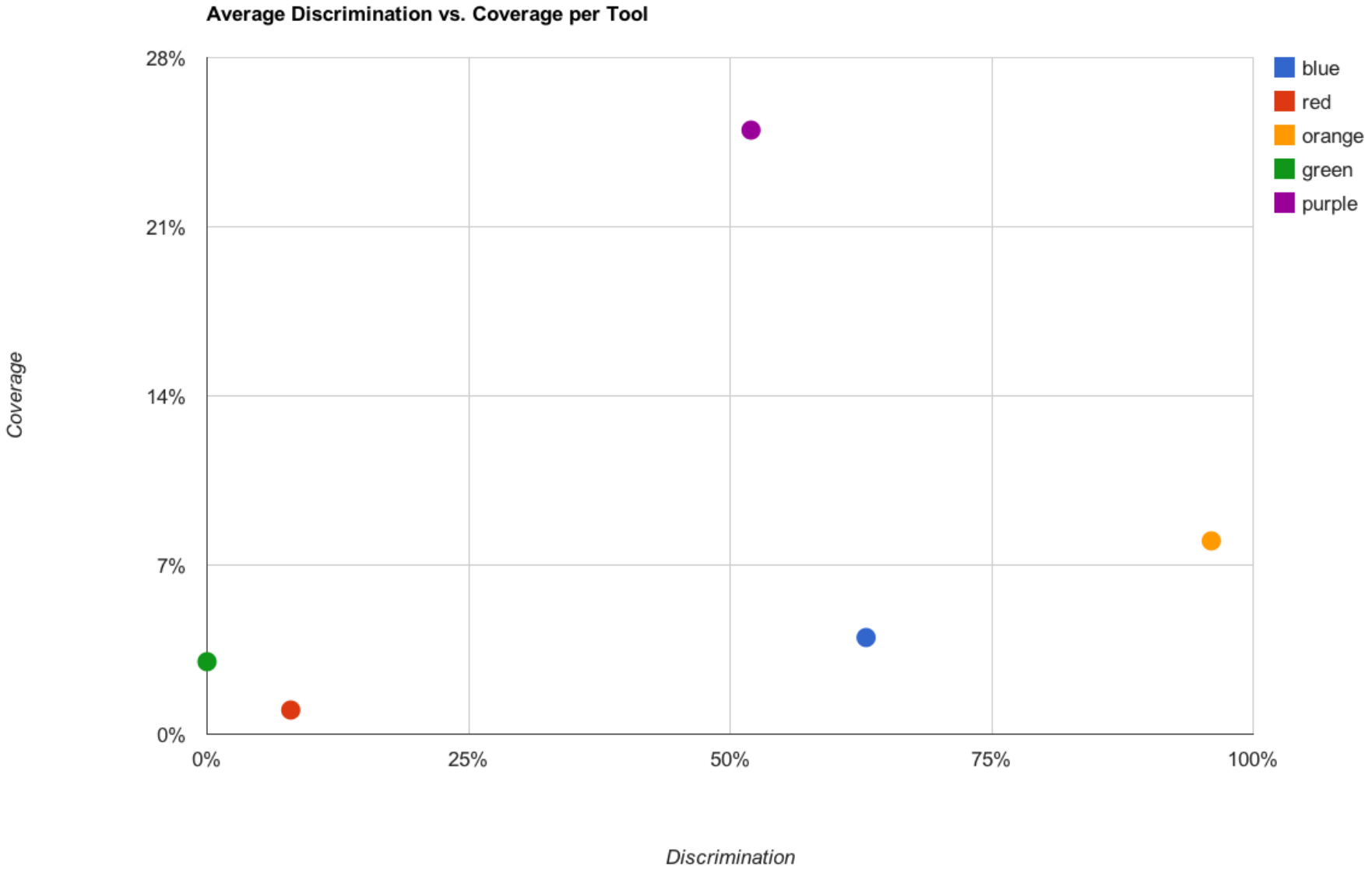
Discrimination:

- Reporting a weakness when there is one
- Keeping quiet when there is none

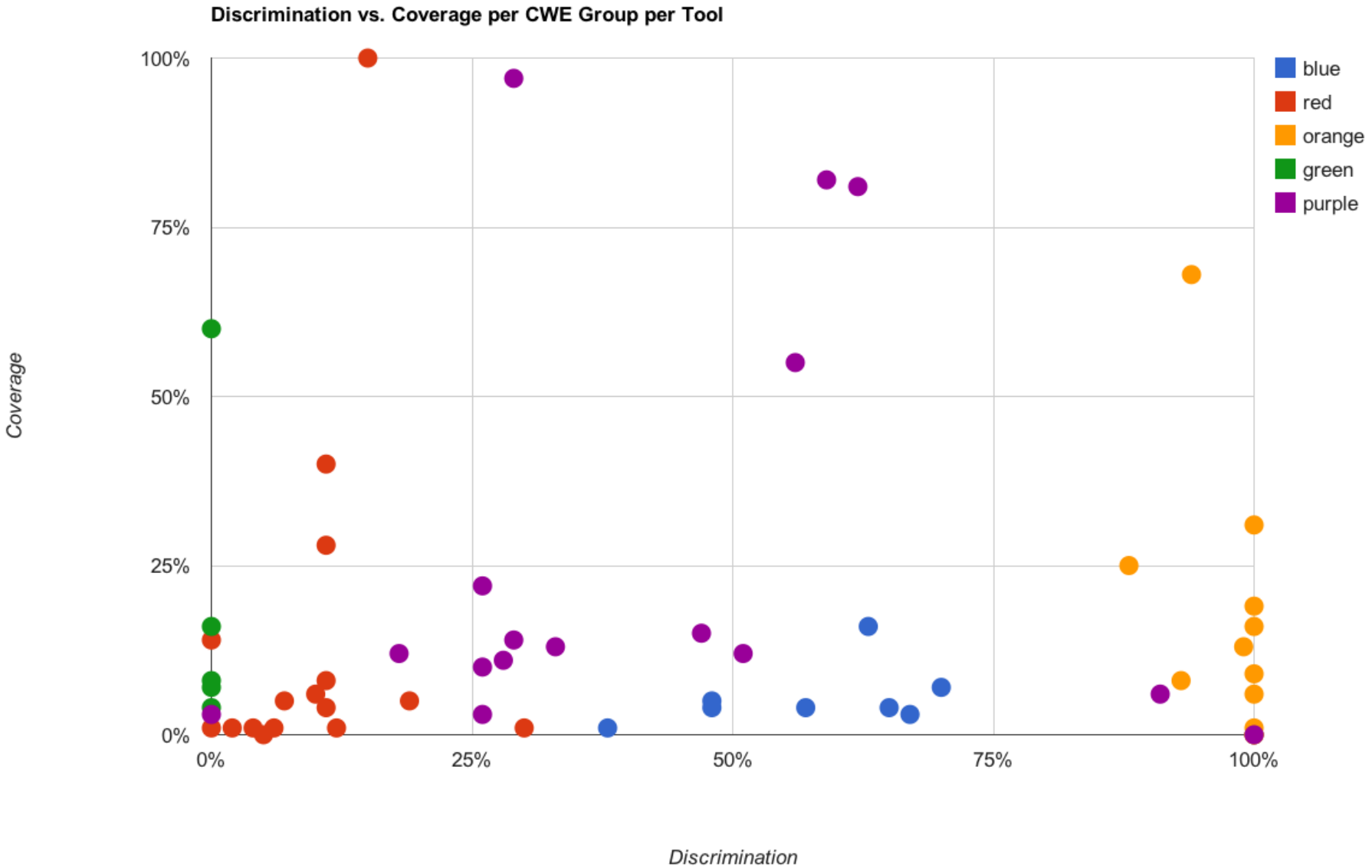
Coverage:

- What proportion of warnings are found?

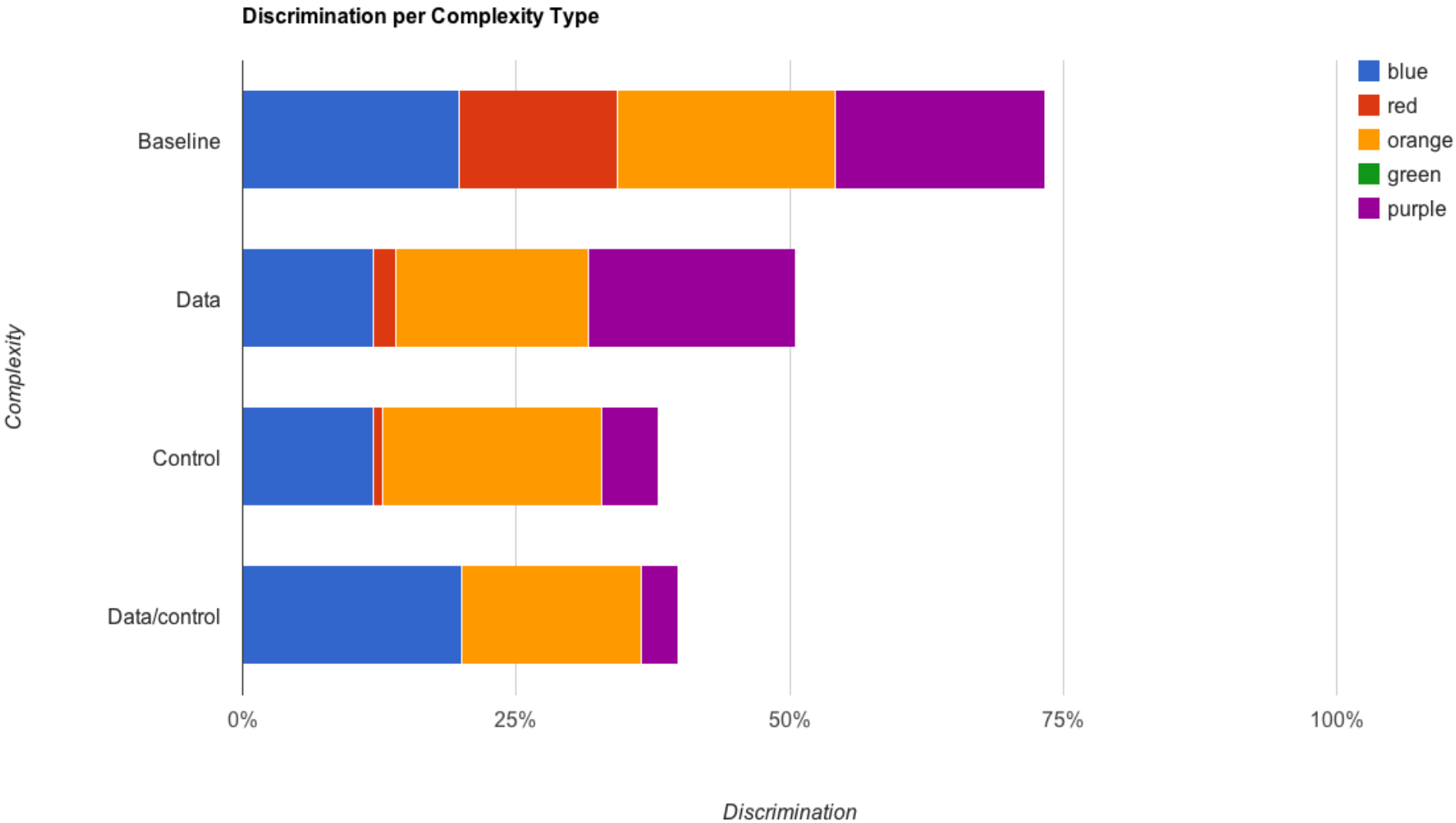
Discrimination vs. Coverage



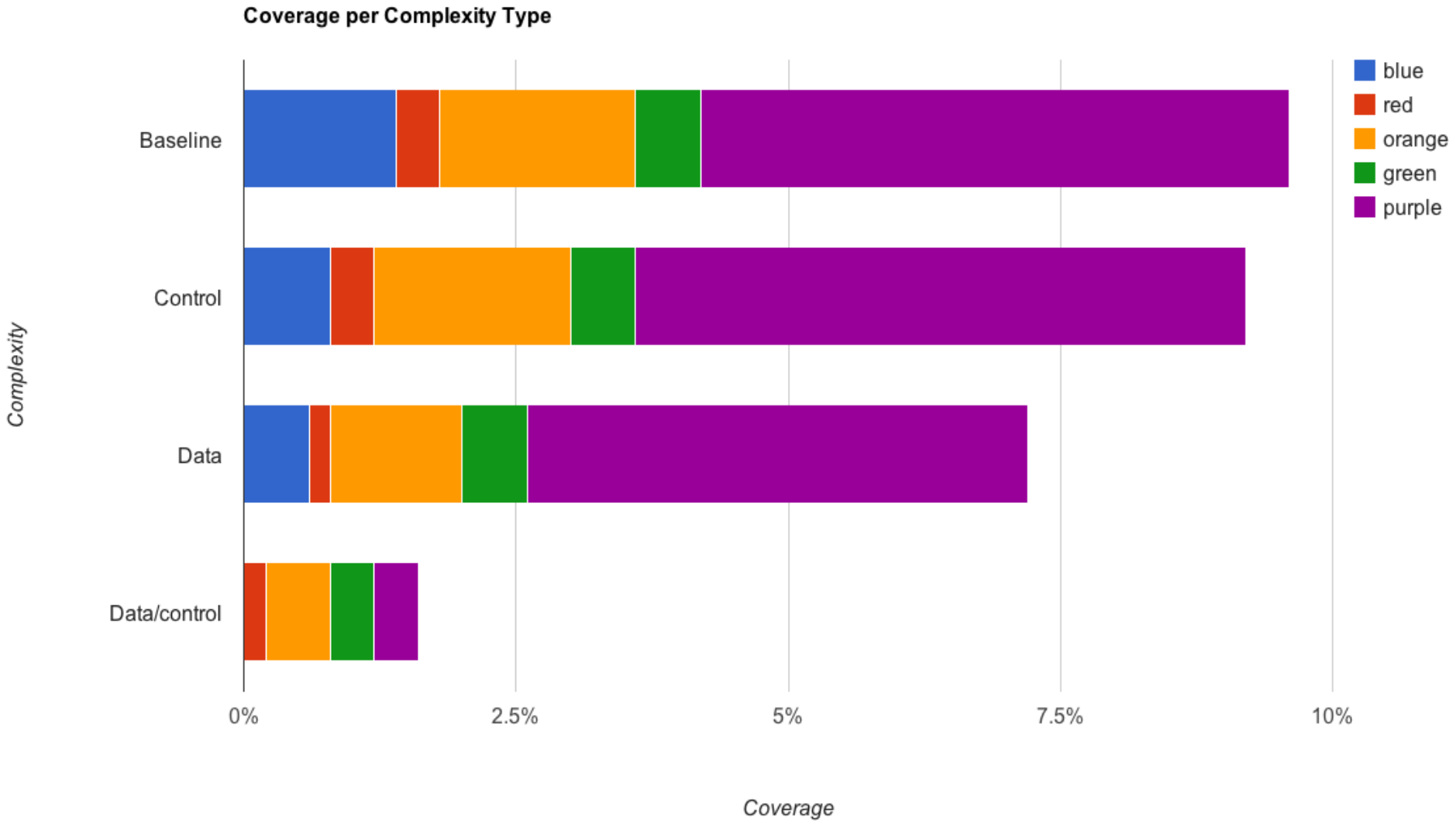
Discrimination vs. Coverage / CWE



Complexity's Toll on Discrimination



Complexity's Toll on Coverage



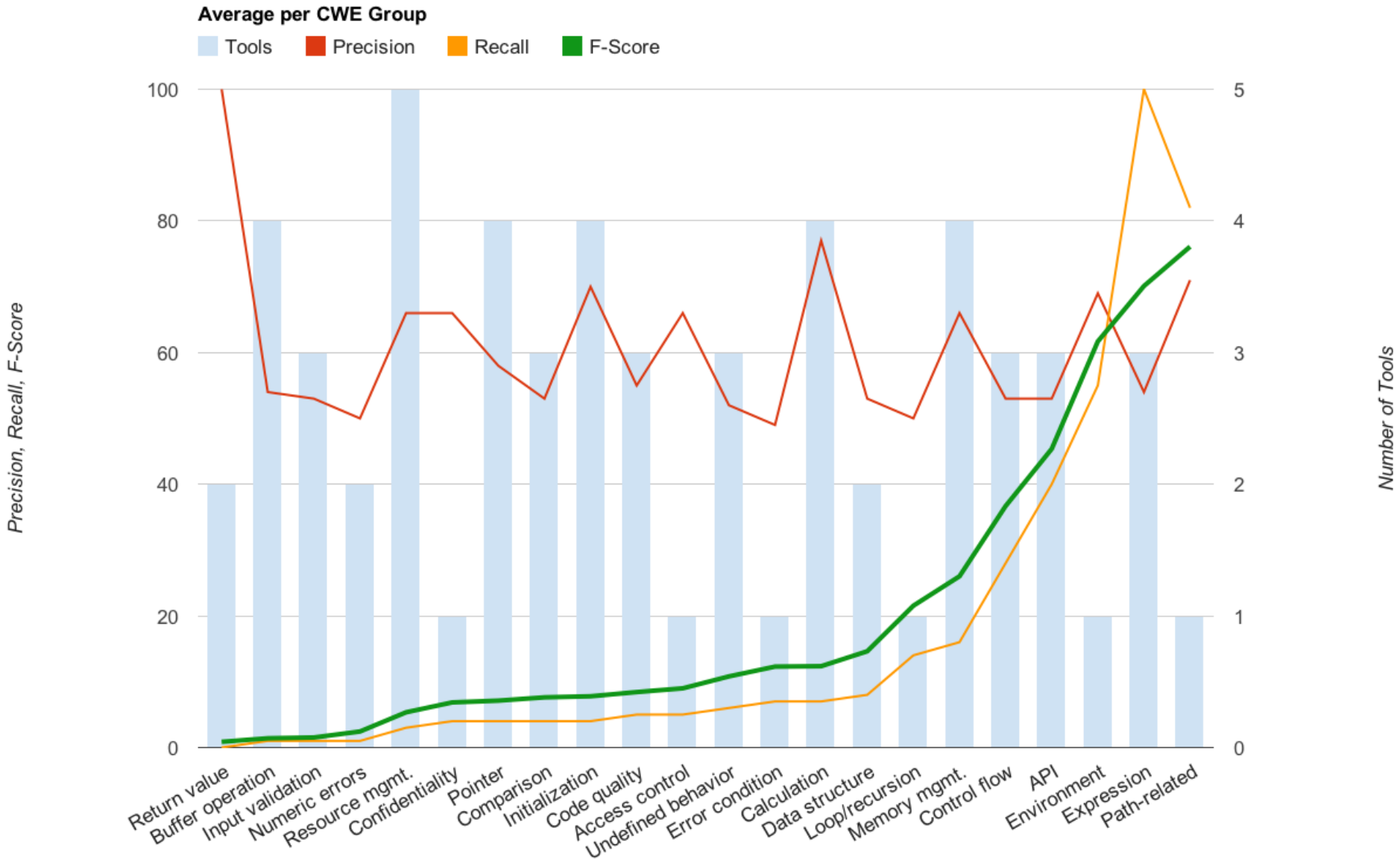
Precision, Recall and F-Score

Precision:
Fraction of retrieved instances that are relevant

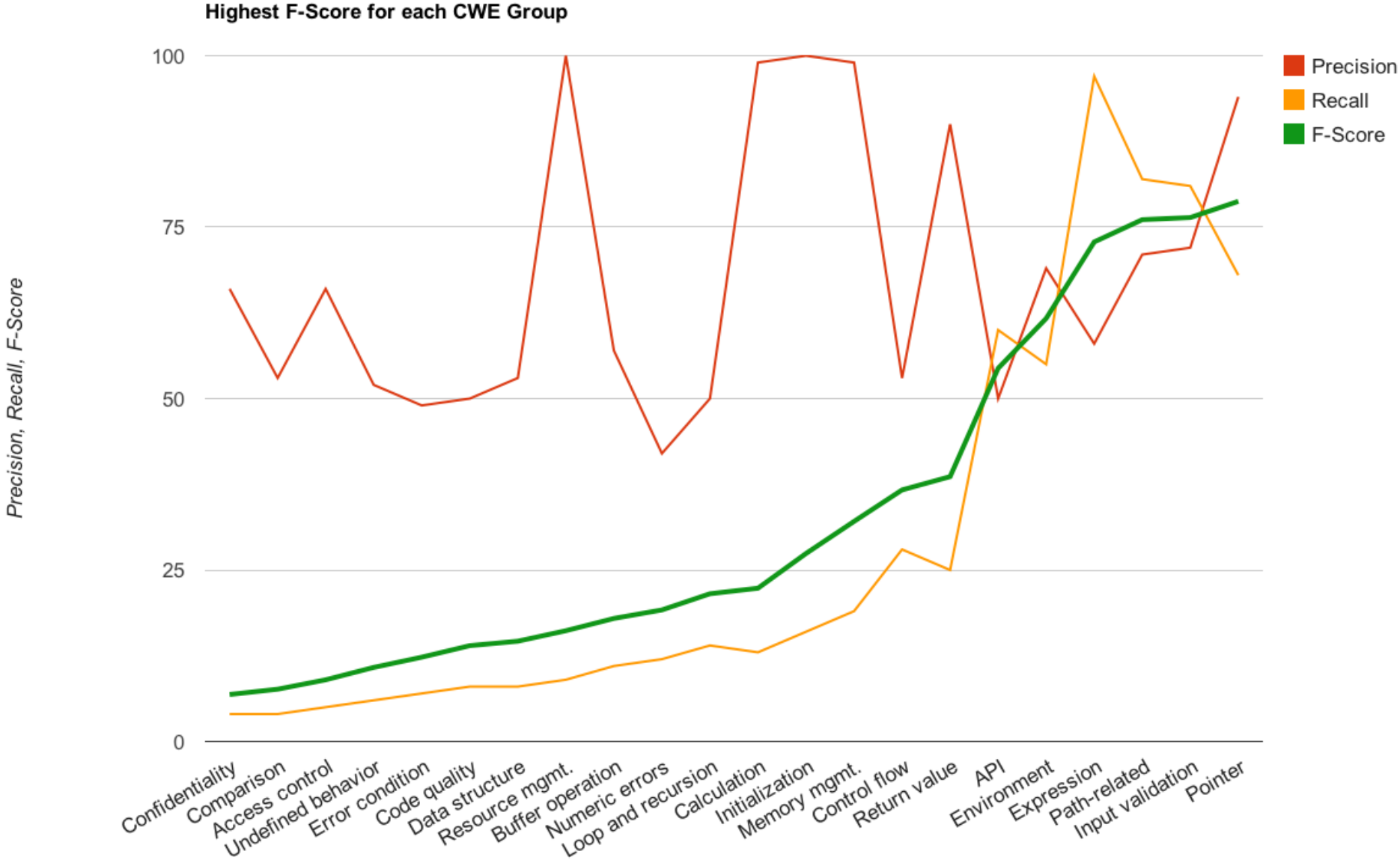
Recall:
Fraction of relevant instances that are retrieved

F-Score:
Harmonic mean of precision and recall

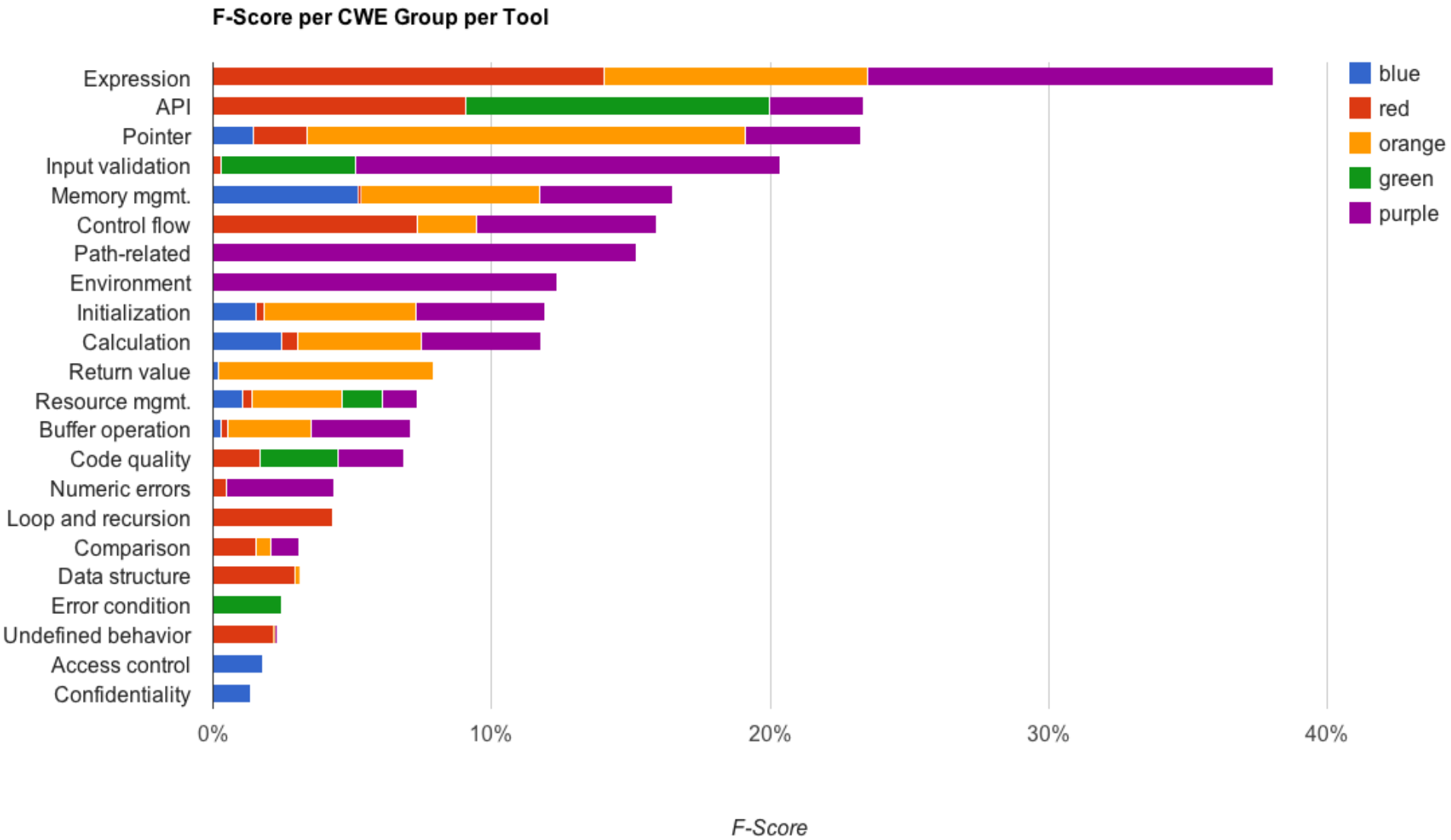
Average per CWE Group



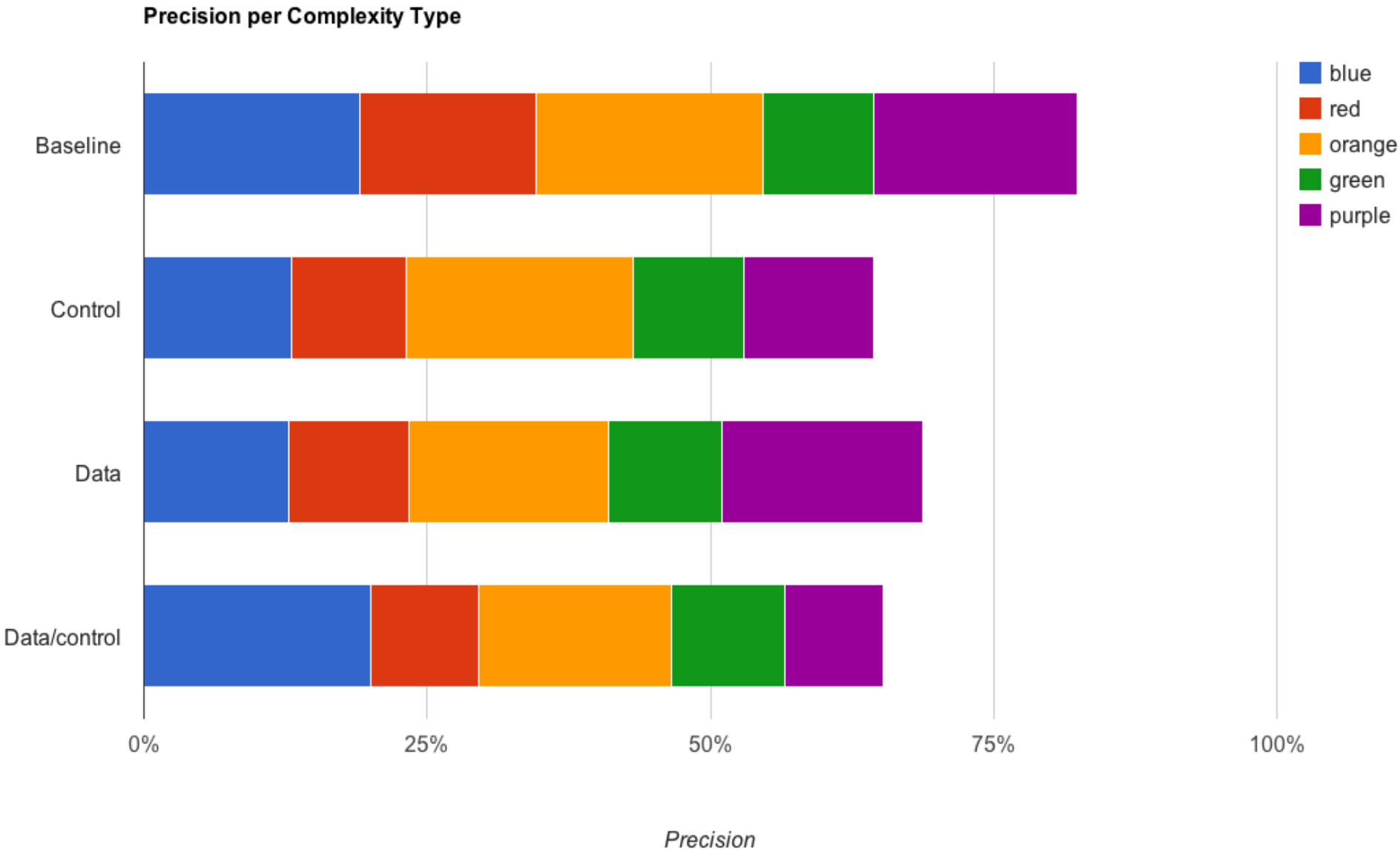
Highest F-Score per CWE Group



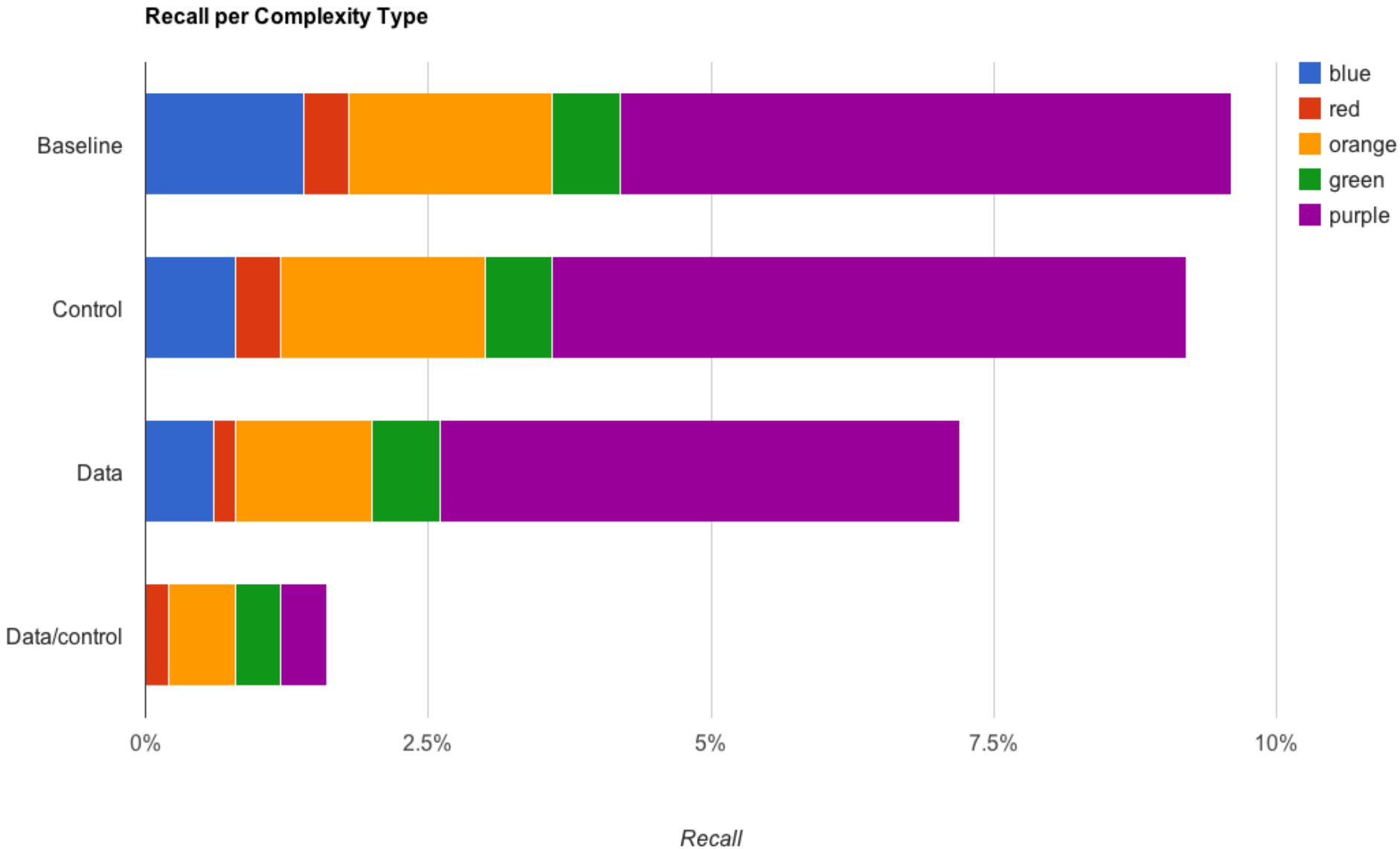
Best-handled CWE Groups



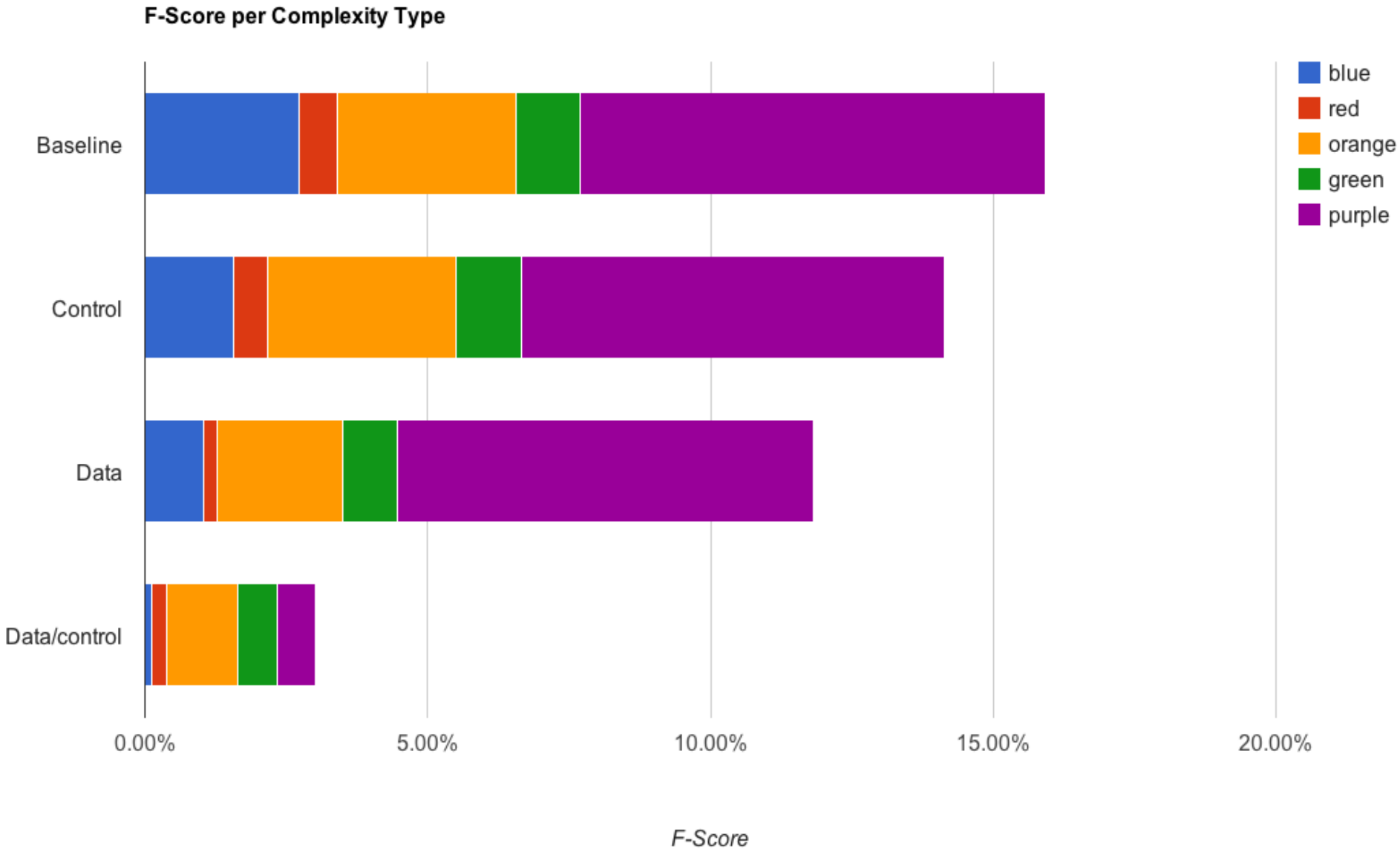
Complexity's Toll on Precision



Complexity's Toll on Recall



Complexity's Toll on the F-Score



Overlap and O-Score

Overlap(N):

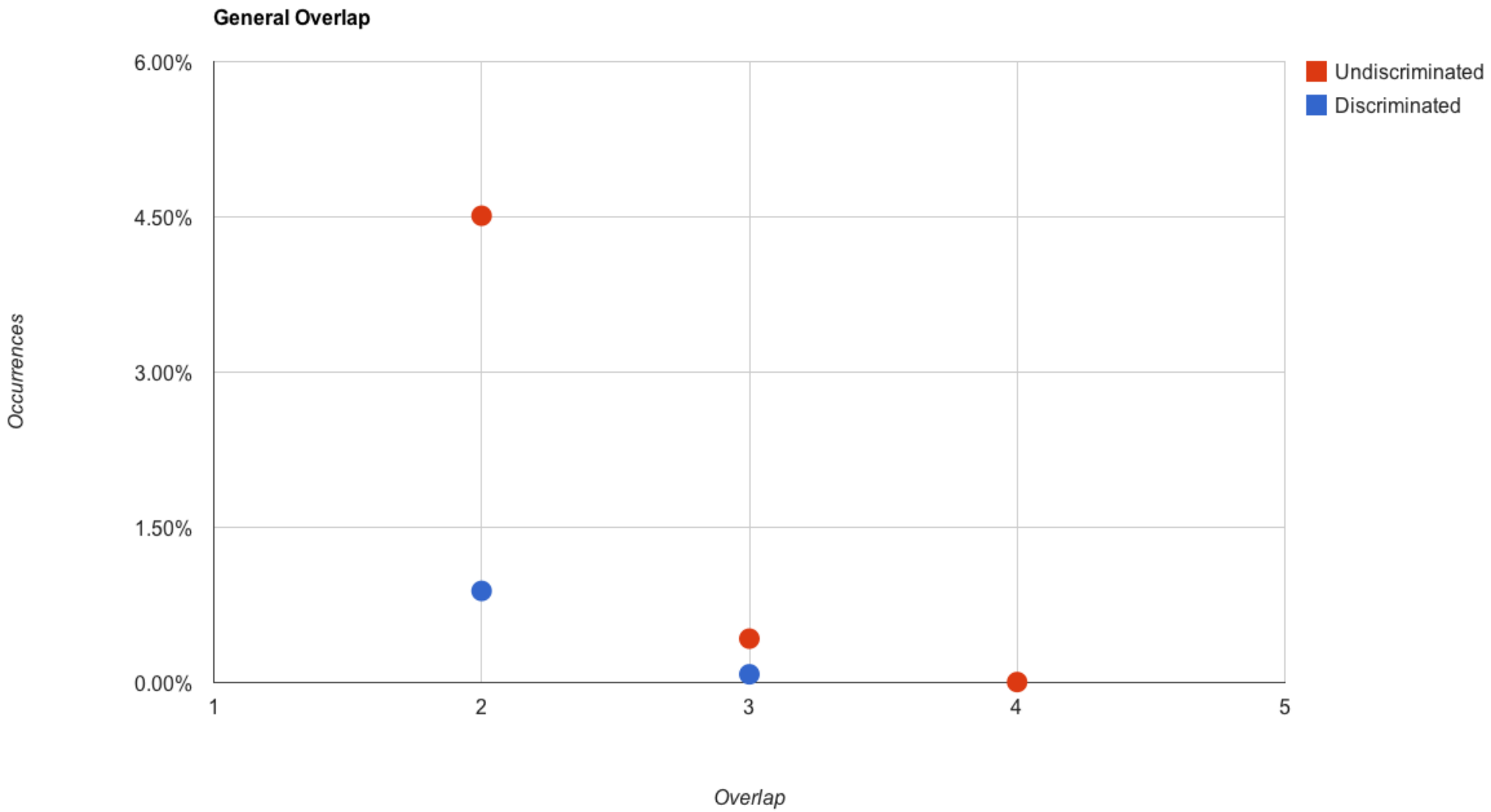
N tools identify the same weakness

O-Score:

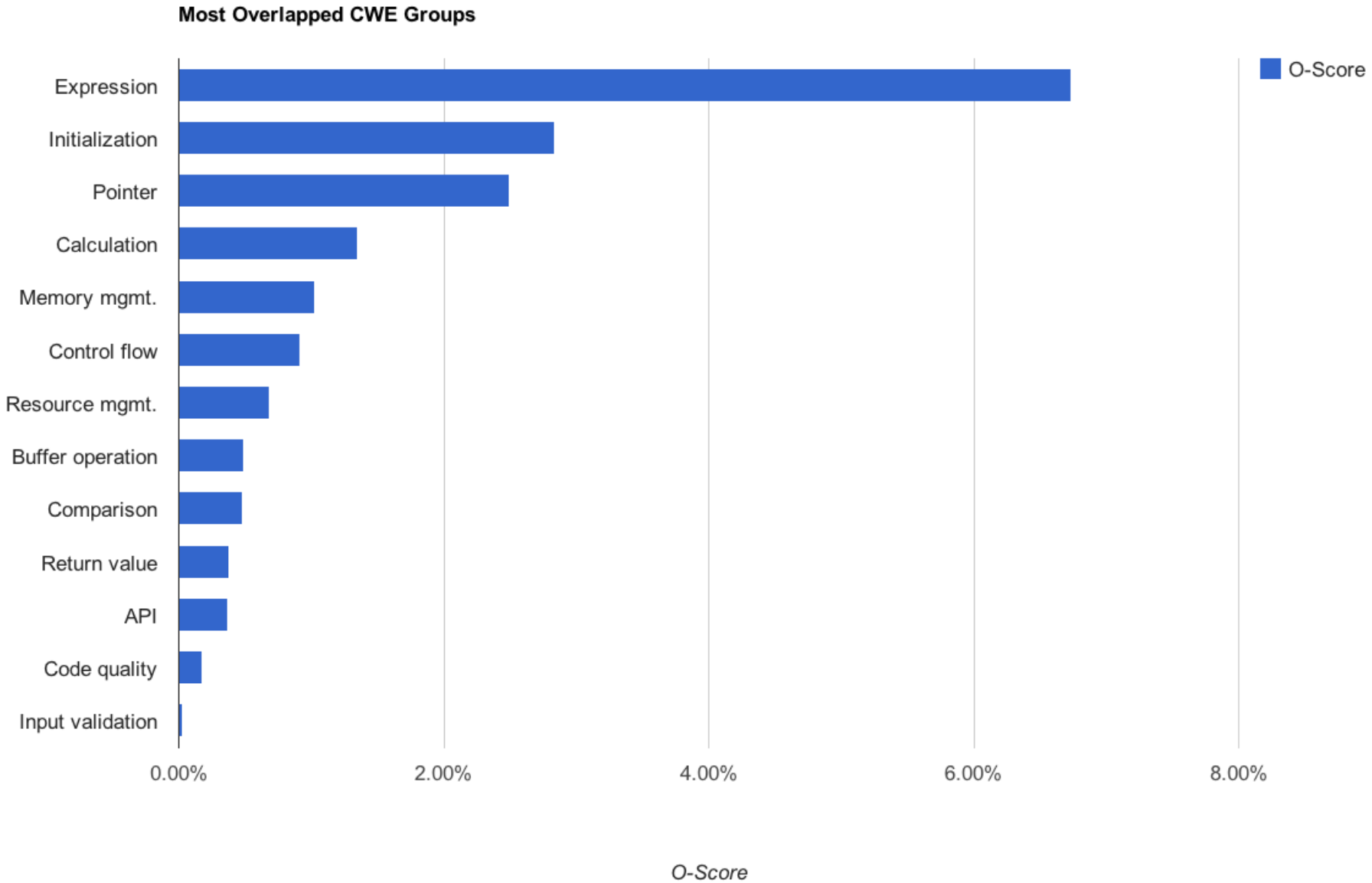
Overlap-based ranking score

O-Score(Overlap(N)) < O-Score(Overlap(N+1))

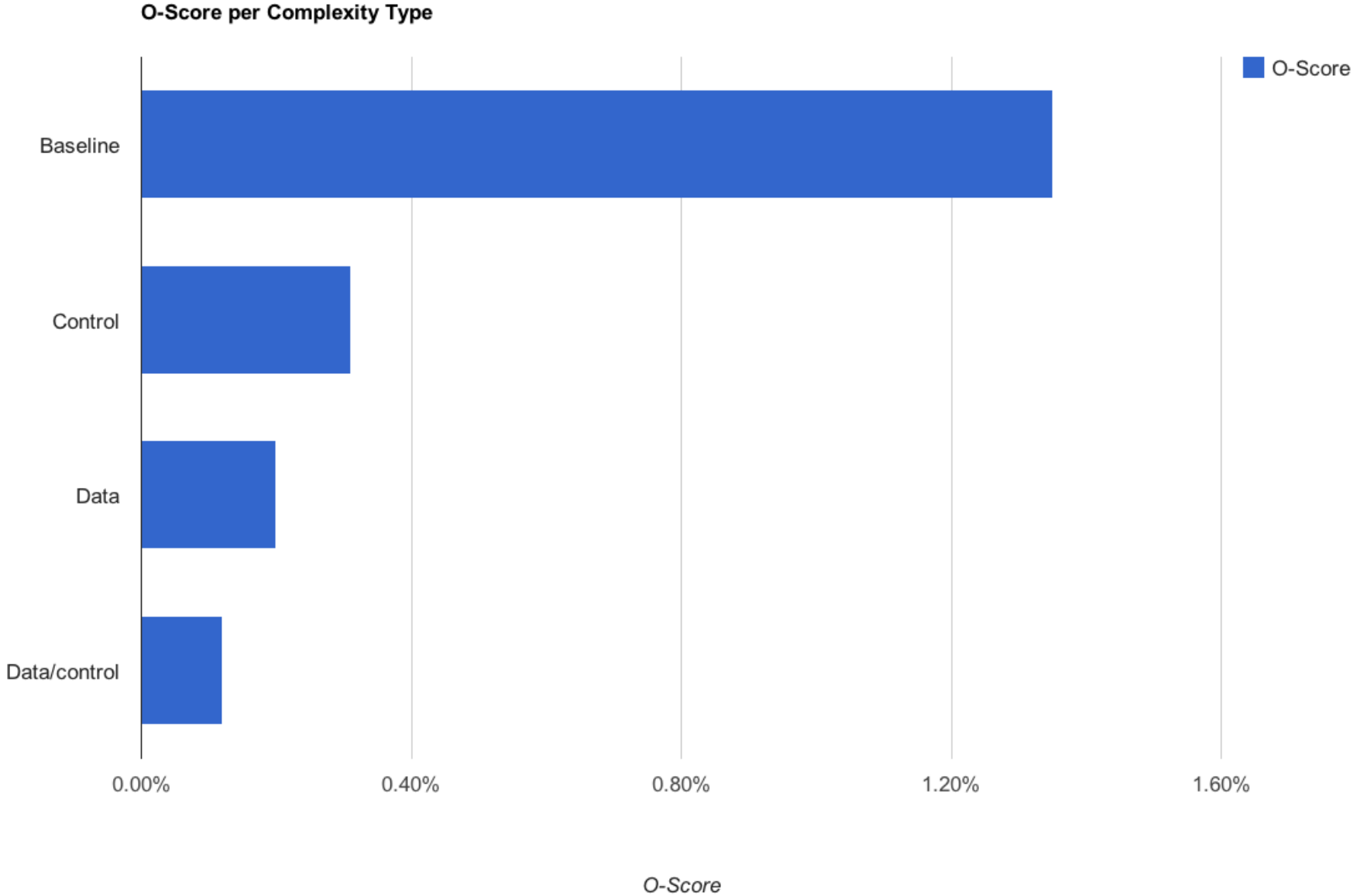
General Overlap



Most Overlapped CWE Groups



Complexity's Toll on Overlap



Conclusion

Tools differ in several dimensions

Using a cleverly picked set of tools significantly increases performance

Weaknesses are not born equal

Complexity is a key