© Paul Shambroom

When Things Go Wrong

Authenticating Nuclear Warheads with High Confidence

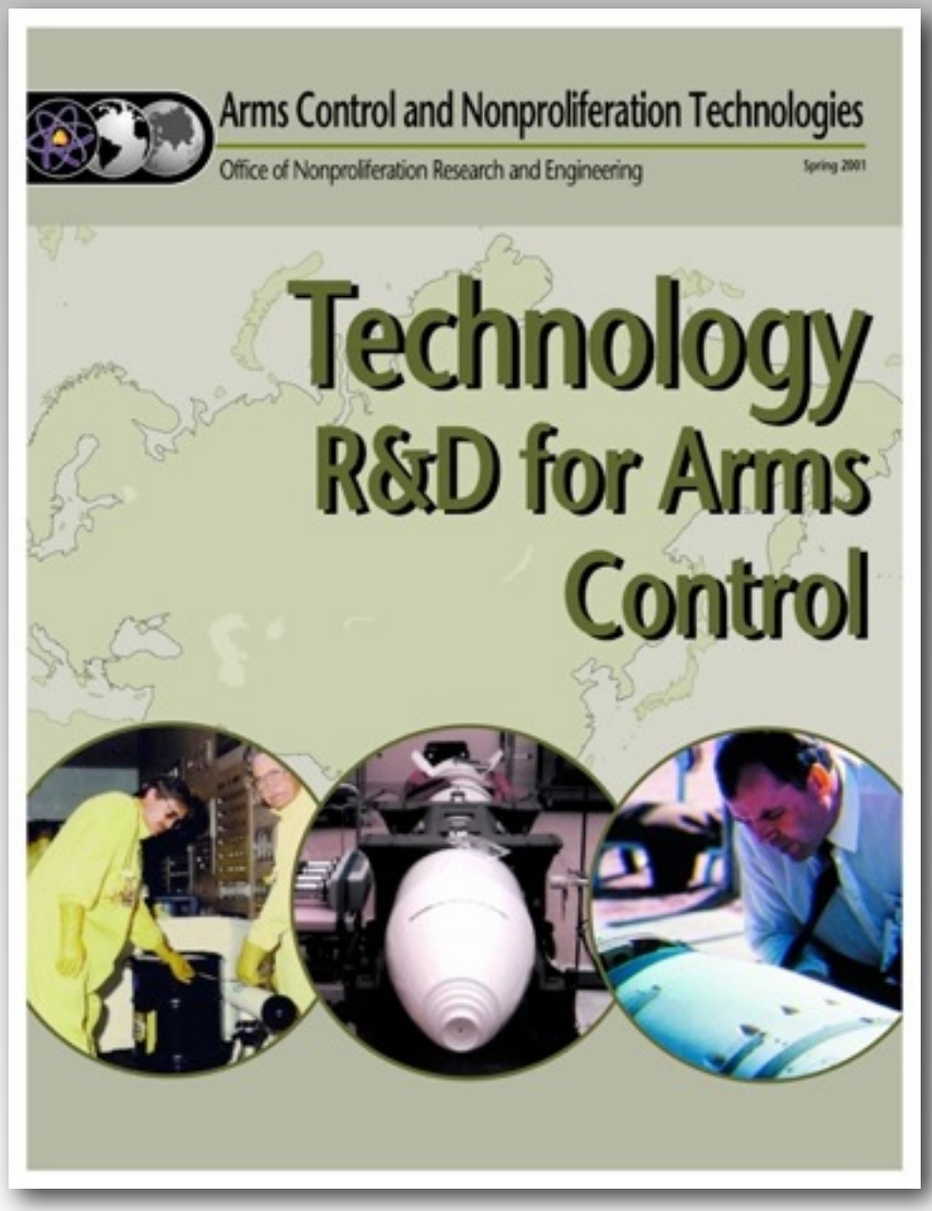
Moritz Kütt, Sebastien Philippe,
Boaz Barak, Alexander Glaser, and Robert J. Goldston

Princeton University, Technische Universität Darmstadt
Microsoft Research New England, Princeton Plasma Physics Laboratory

55th Annual Meeting, Institute of Nuclear Materials Management
Georgia, Atlanta, July 2014

Revision 5

Inspection Systems for Nuclear Warhead Verification Have Been Under Development Since the 1990s



edited by David Spears, 2001

Attribute Approach

Confirming selected characteristics of an object
in classified form
(for example, the presence/mass of plutonium)

Template Approach

Comparing the radiation signature
from the inspected item with a reference item
("golden warhead") of the same type

Information Barrier

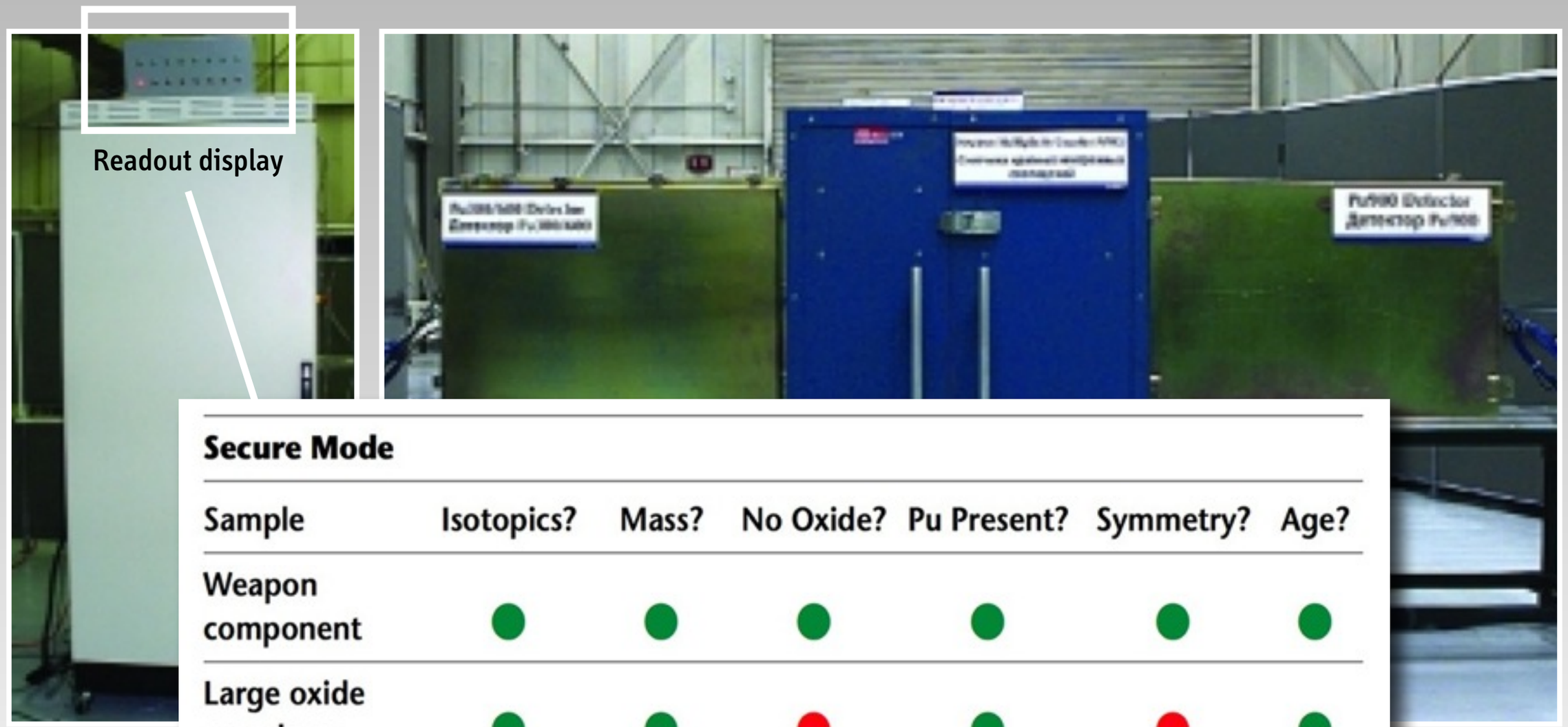
Technologies and procedures that prevent the
release of sensitive nuclear information
(generally needed for both approaches)

What To Do When “Red Lights” Come On?

Fissile Material Transparency Technology Demonstration (FMTTD), Los Alamos, August 2000

David Spears (ed.), *Technology R&D for Arms Control*, U.S. Department of Energy, Office of Nonproliferation Research and Engineering, Washington, DC, 2001

What To Do When “Red Lights” Come On?



Secure Mode

Sample	Isotopics?	Mass?	No Oxide?	Pu Present?	Symmetry?	Age?
Weapon component	●	●	●	●	●	●
Large oxide sample on its side	●	●	●	●	●	●

Fissile Material Transparency Technology Demonstration (FMTTD), Los Alamos, August 2000

David Spears (ed.), *Technology R&D for Arms Control*, U.S. Department of Energy, Office of Nonproliferation Research and Engineering, Washington, DC, 2001

Methodology

(applicable to both attribute and template approaches)

Terminology and Definitions

Assume basic test (single measurement) is “not strong enough”

(i.e., too many valid items can fail, too many invalid items can pass)

We repeatedly apply basic test until we declare item “good” or “bad”

Terminology and Definitions

Assume basic test (single measurement) is "not strong enough"

(i.e., too many valid items can fail, too many invalid items can pass)

We repeatedly apply basic test until we declare item "good" or "bad"

4 Key Parameters

α : probability that valid item passes basic test (e.g. 0.95)

β : probability that invalid item passes basic test (0.05–0.10, but could be higher)

α^* : probability that valid item is declared "good" (host wants high, e.g. 0.99)

β^* : probability that invalid item is declared "good" (inspector wants low, e.g. 0.01)

Re-Testing Protocol

Assume we have observed a sequence of n basic tests $X_1, X_2, \ldots, X_n$
(Bernoulli trials or "biased coin flips") that includes P "passes" and F "fails"

We define two hypotheses that we will test

A. Item is valid and passes each basic test with probability α

B. Item is invalid and fails each basic test with probability $(1 - \beta)$

For any given n, we can check the following inequalities to accept one of the hypotheses:

A. $$\sum_{i=F}^{n} \binom{n}{i} (1-\alpha)^i \, \alpha^{n-i} \leq 1 - \alpha^*$$

B. $$\sum_{i=P}^{n} \binom{n}{i} \beta^i \, (1-\beta)^{n-i} \leq \beta^*$$

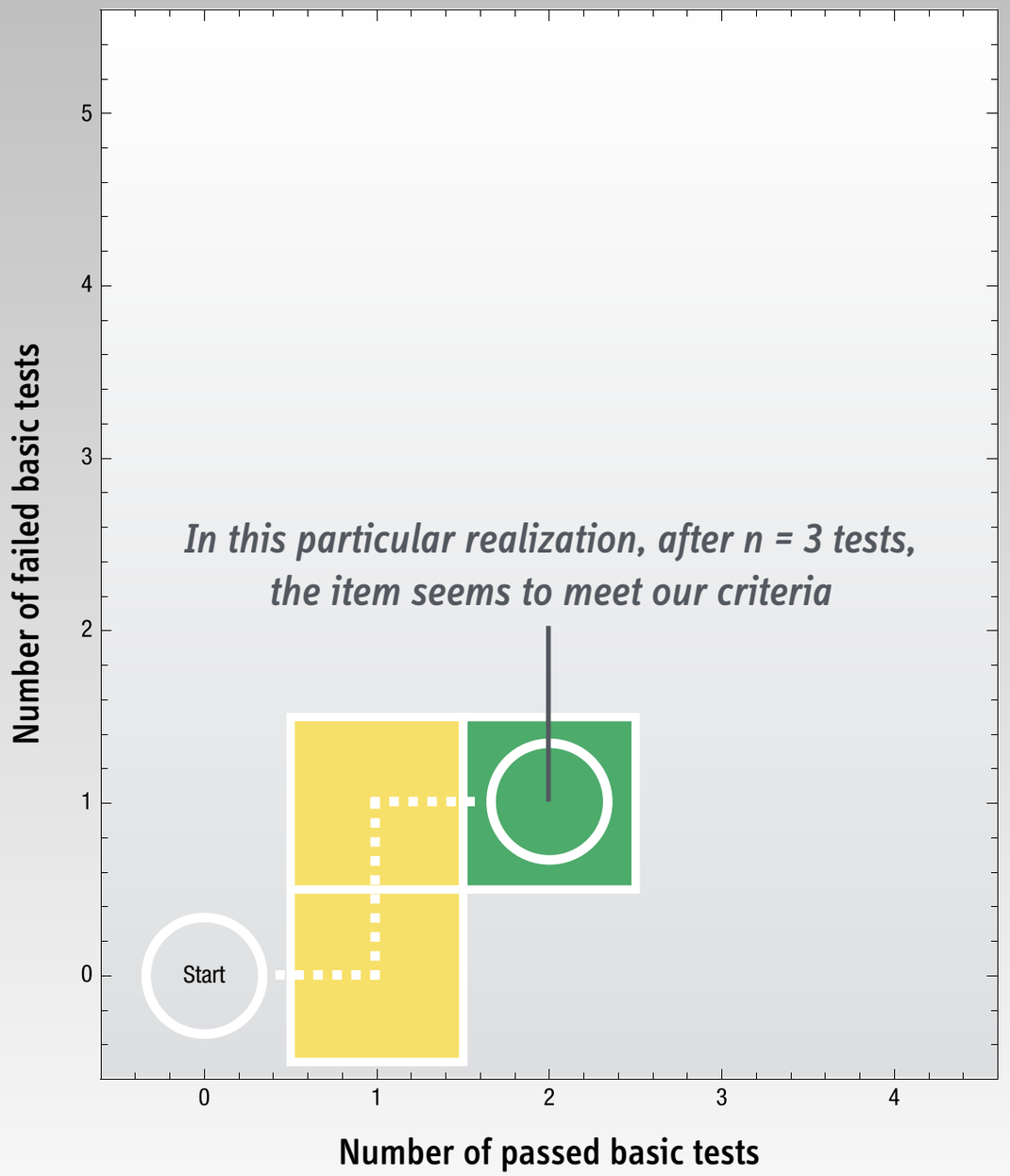
(both can't be true at the same time if $\beta < \alpha$)

We want to find the smallest n_{opt} such that one of the inequalities is true
for any combination of passed tests (k_p) and failed tests (k_f) with $k_p + k_f = n_{opt}$

Scorecards

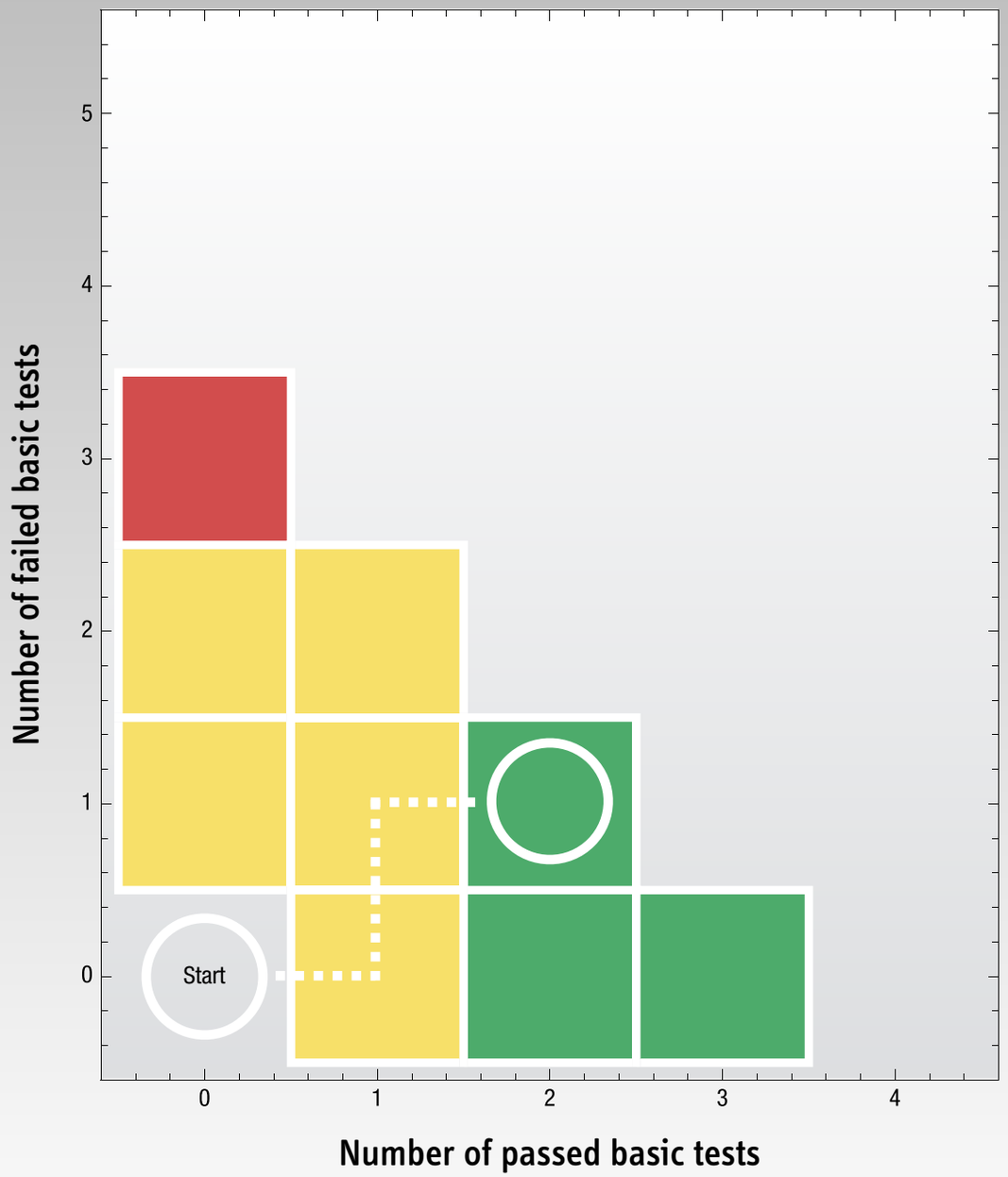
Making Scorecards

Default parameter set: $(\alpha = 0.95, \beta = 0.05, \alpha^* = 0.999, \beta^* = 0.01)$



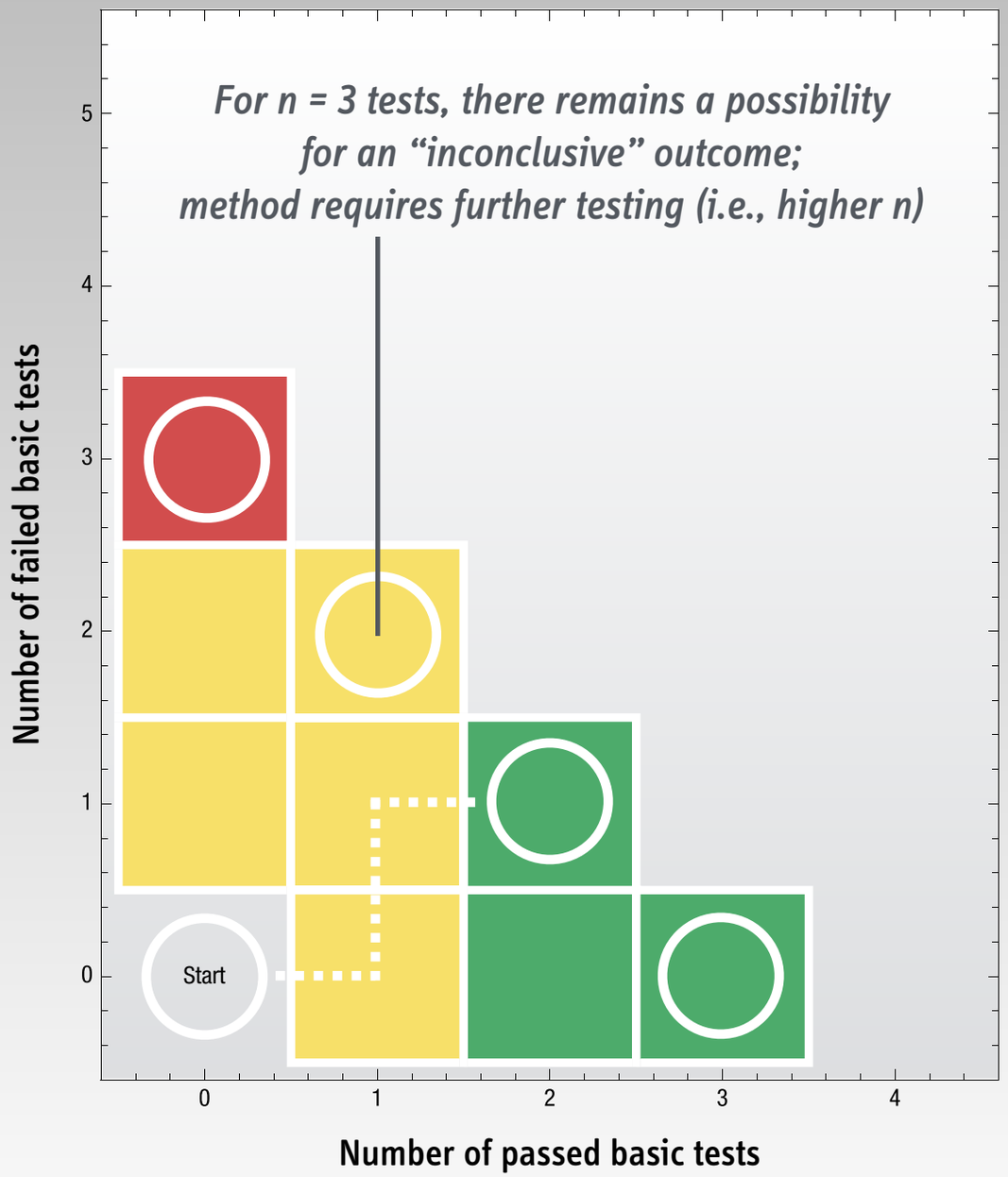
Making Scorecards

Default parameter set: $(\alpha = 0.95, \beta = 0.05, \alpha^* = 0.999, \beta^* = 0.01)$



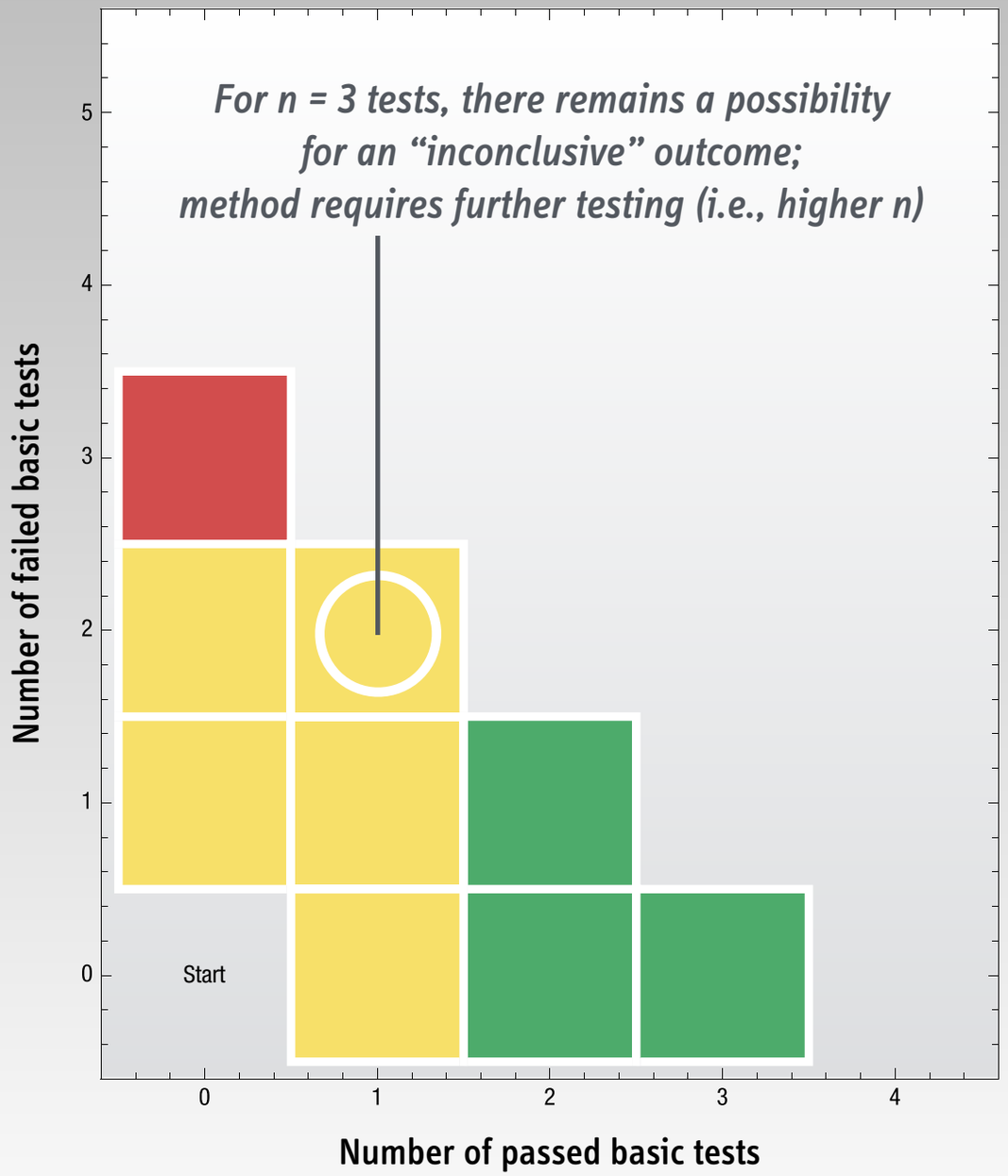
Making Scorecards

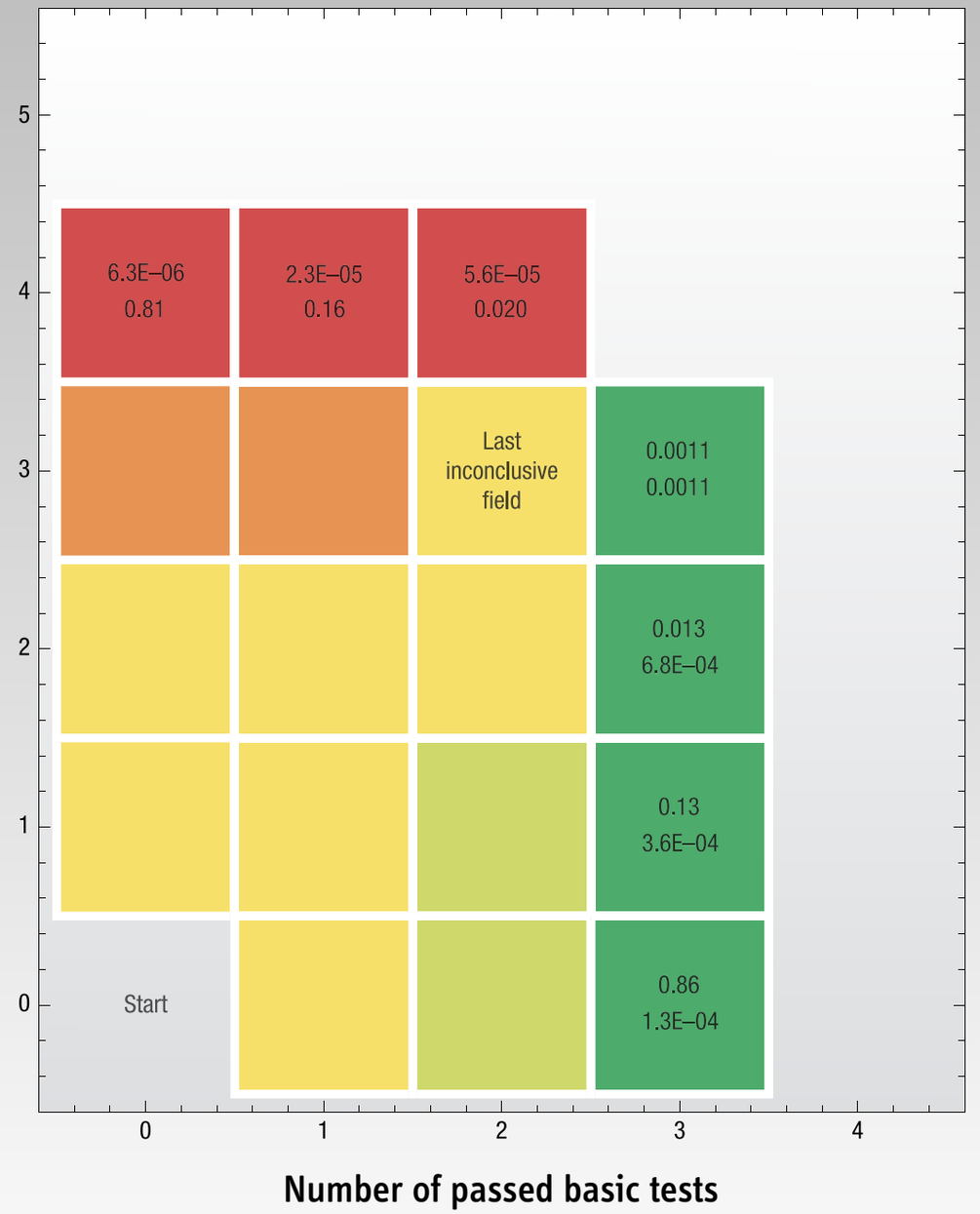
Default parameter set: $(\alpha = 0.95, \beta = 0.05, \alpha^* = 0.999, \beta^* = 0.01)$



Making Scorecards

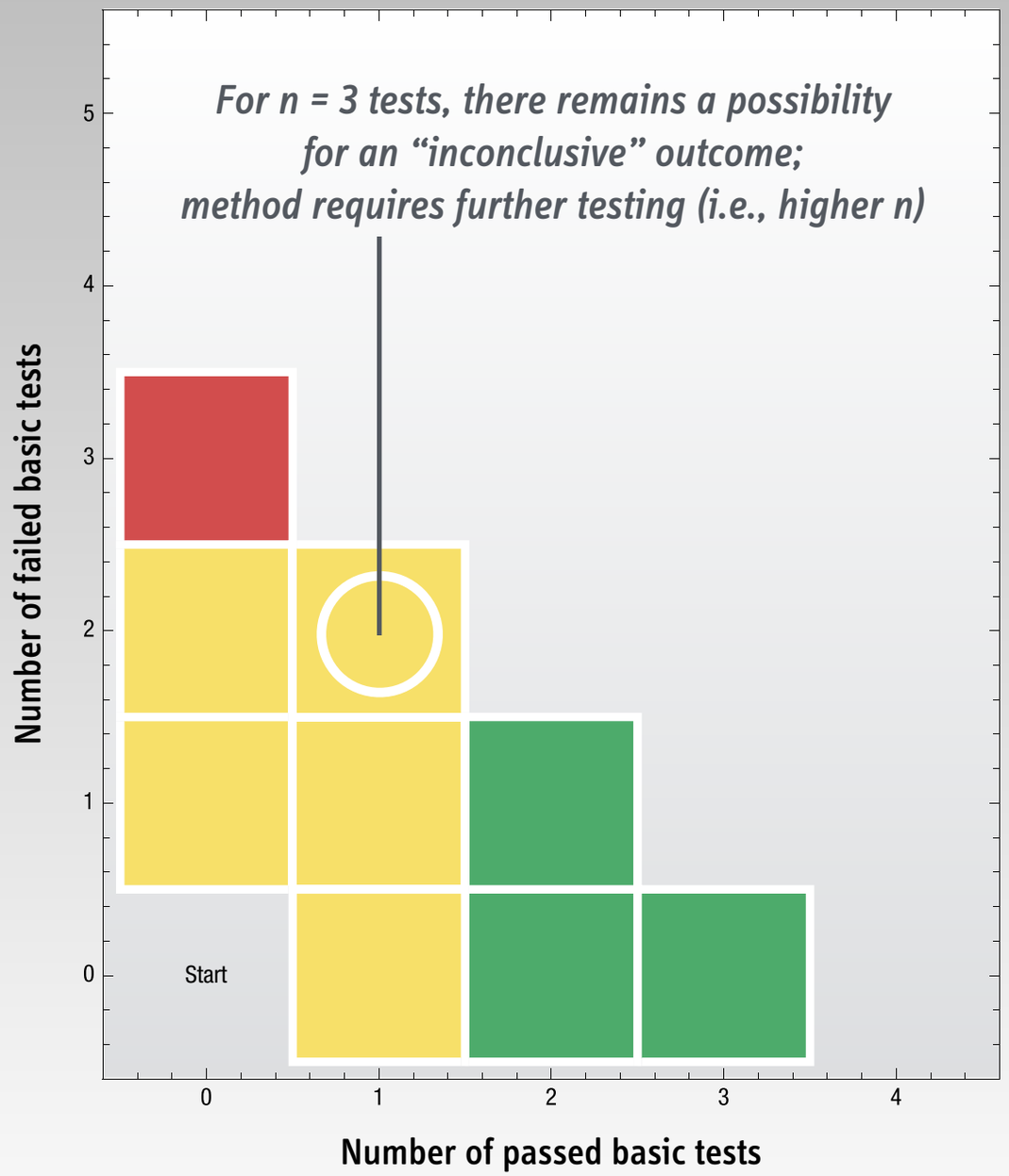
Default parameter set: $(\alpha = 0.95, \beta = 0.05, \alpha^* = 0.999, \beta^* = 0.01)$

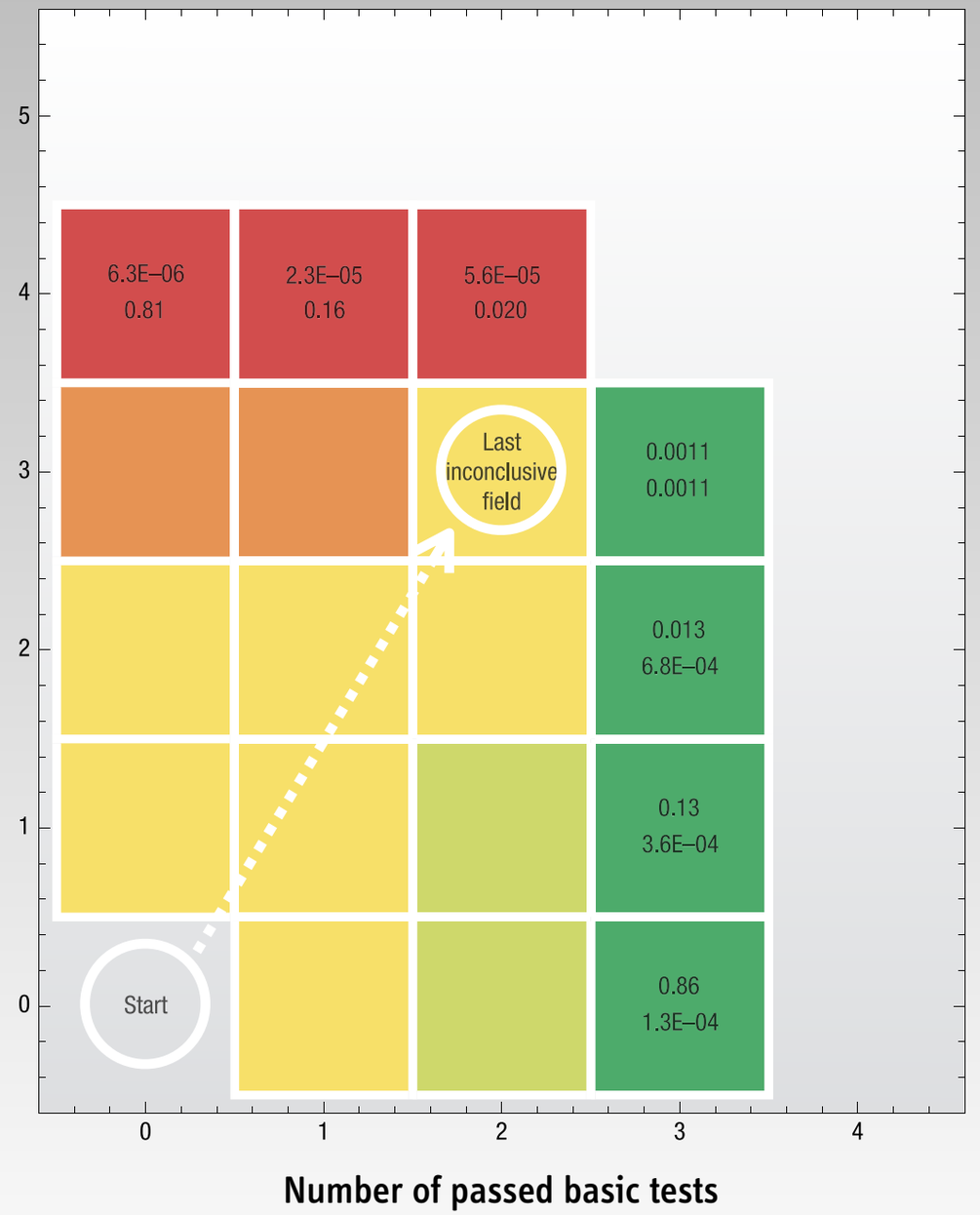




Making Scorecards

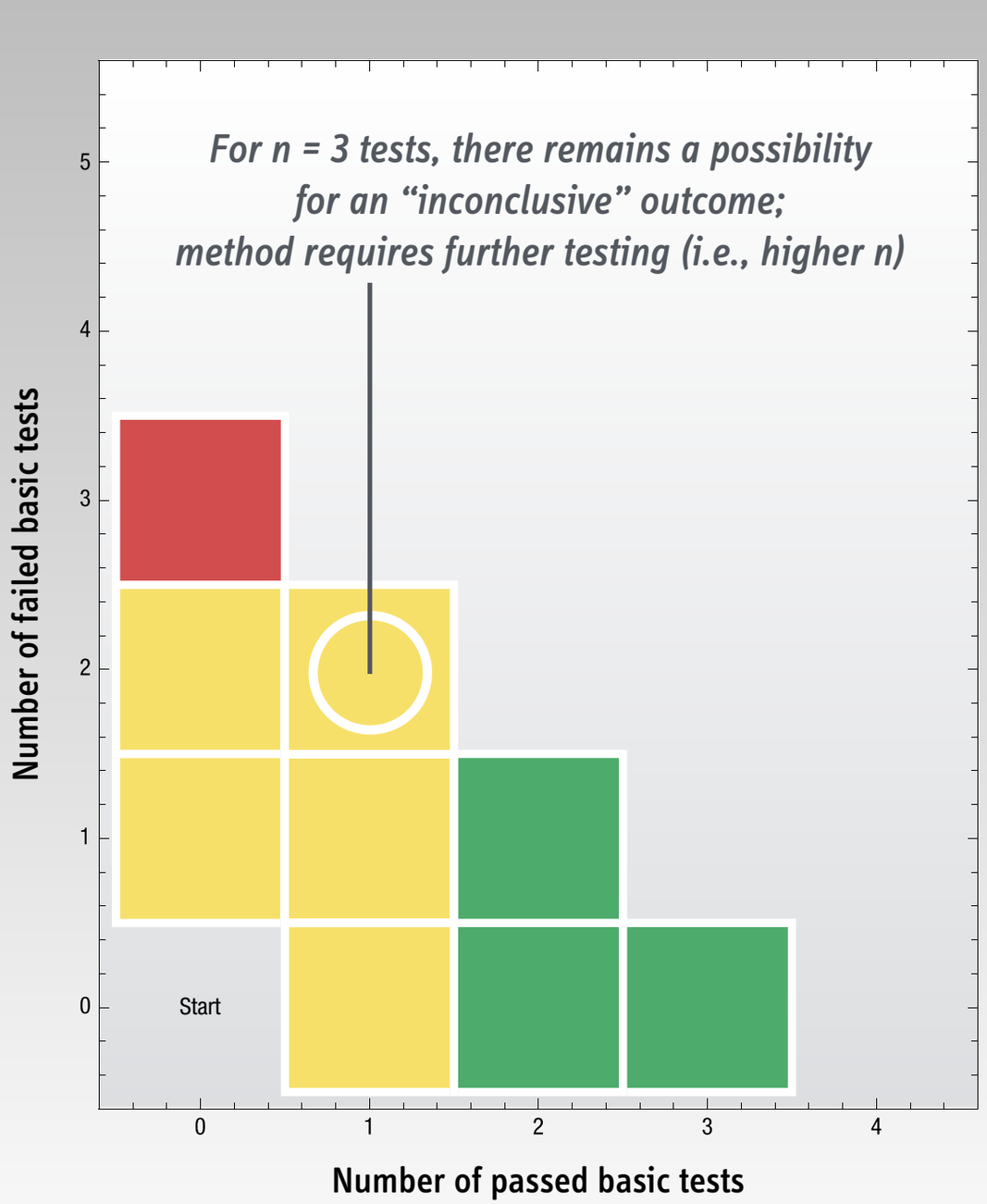
Default parameter set: $(\alpha = 0.95, \beta = 0.05, \alpha^* = 0.999, \beta^* = 0.01)$

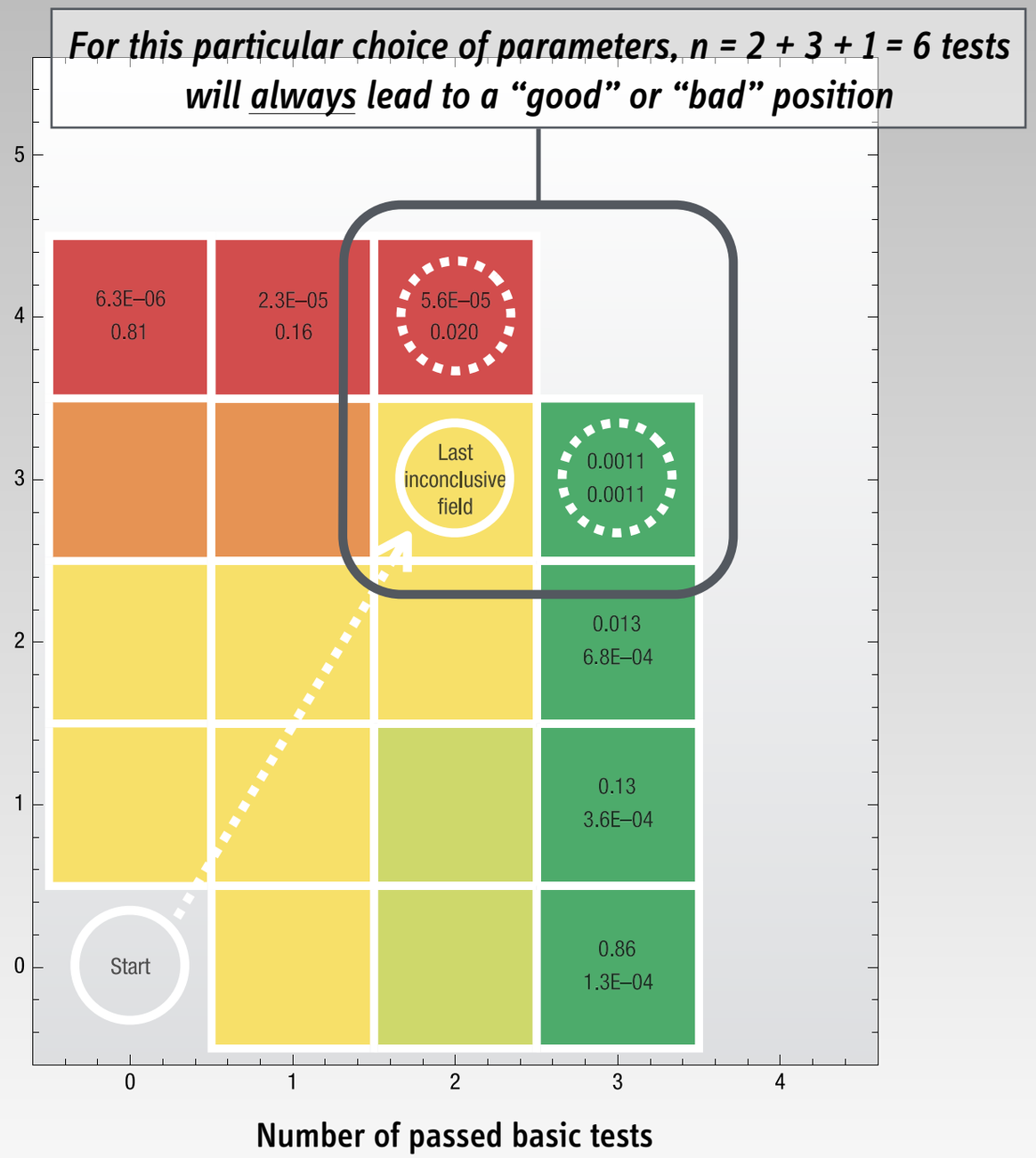




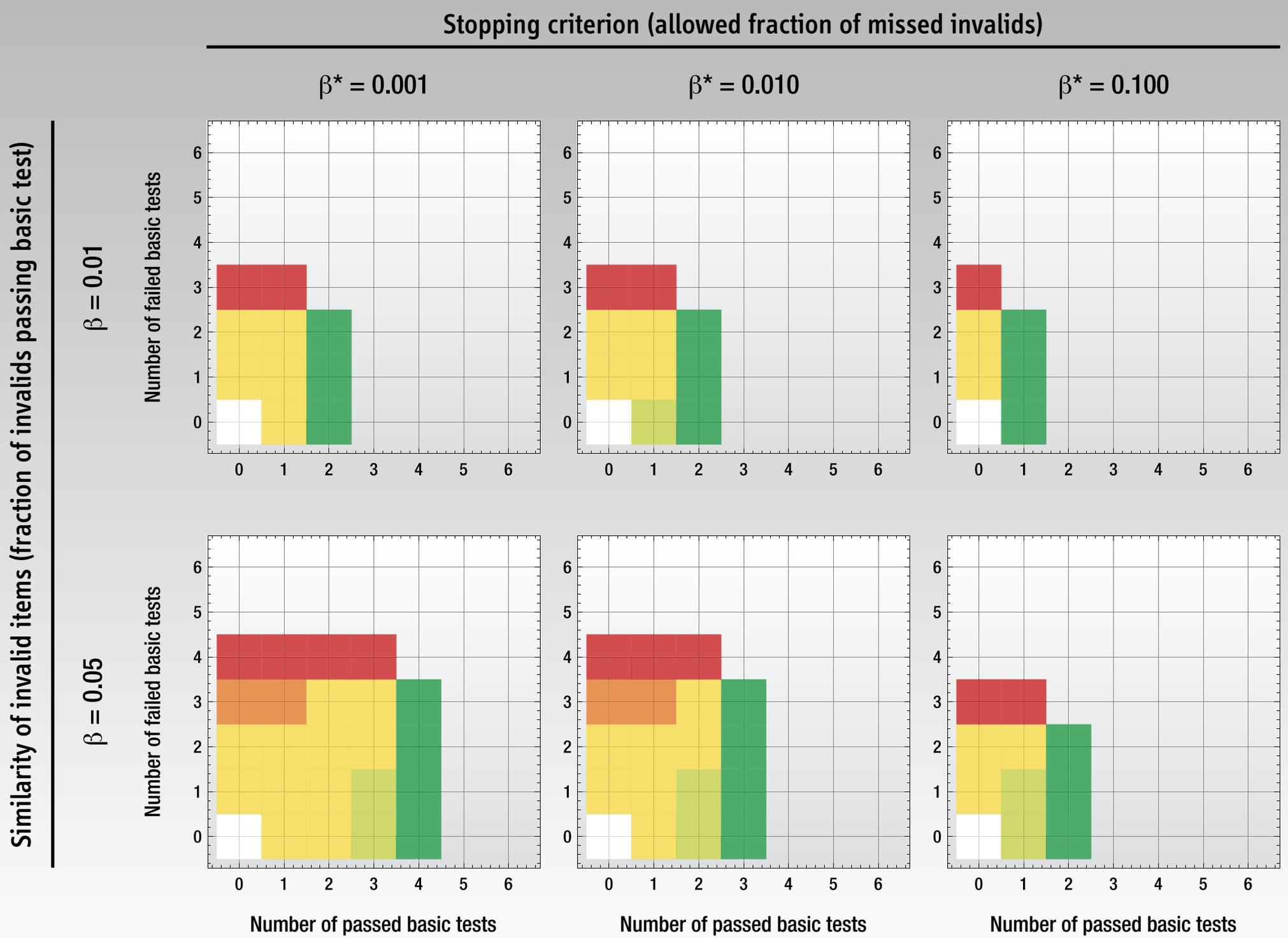
Making Scorecards

Default parameter set: $(\alpha = 0.95, \beta = 0.05, \alpha^* = 0.999, \beta^* = 0.01)$





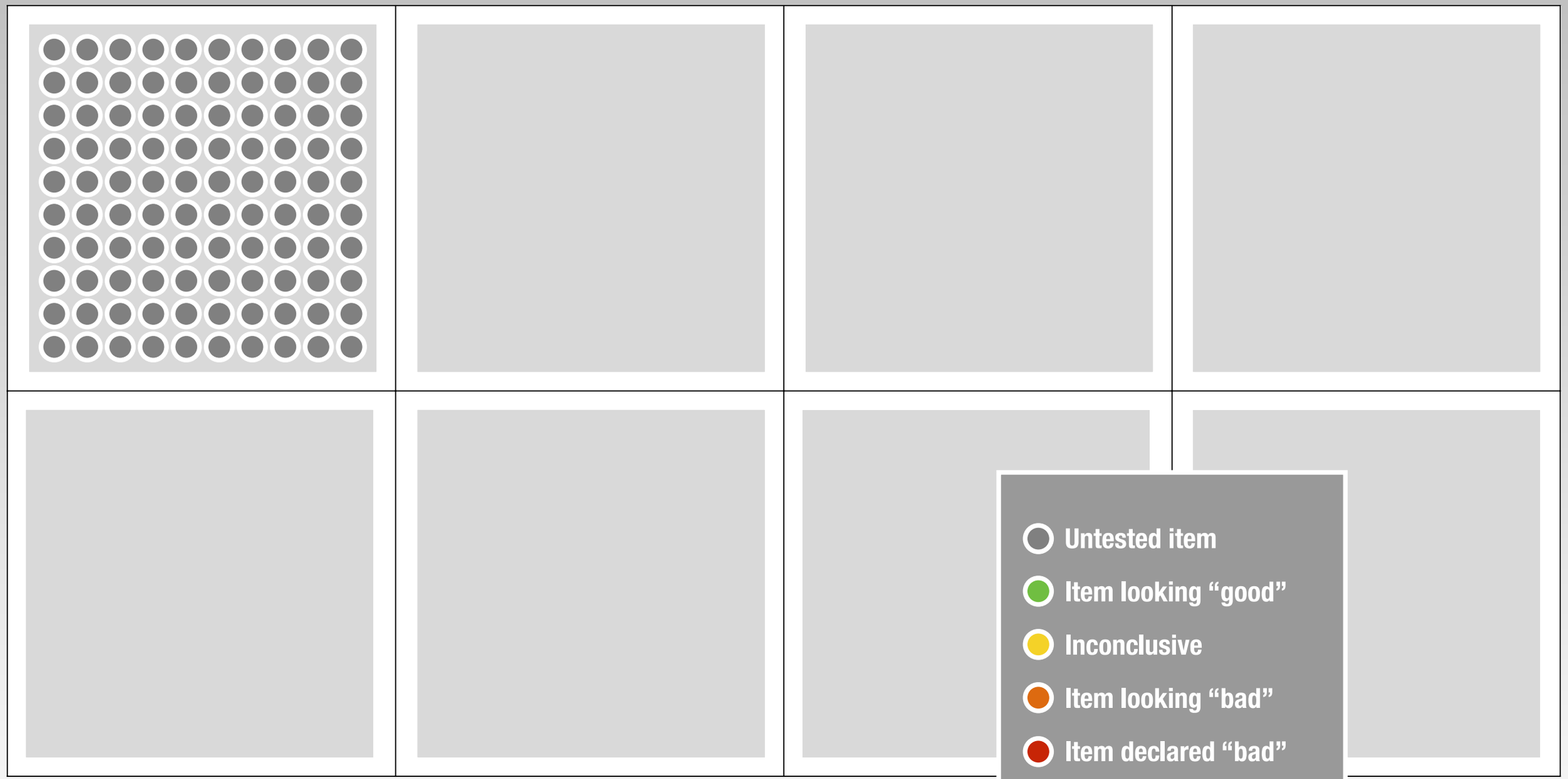
“Pick Your Values … Make Your Scorecard”



Simulated Results

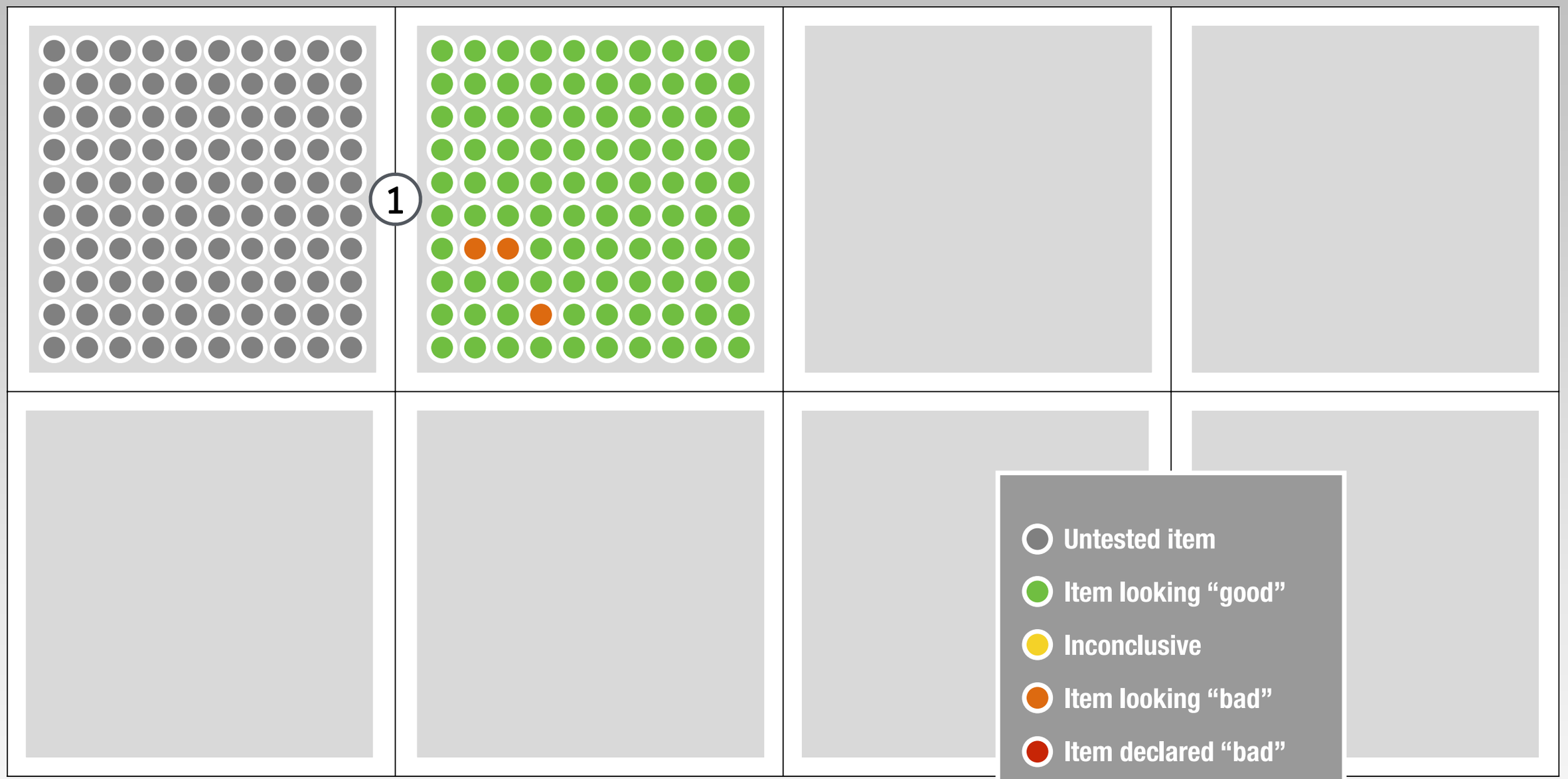
Honest Host

Default parameter set: $(\alpha = 0.95, \beta = 0.05, \alpha^* = 0.999, \beta^* = 0.01)$



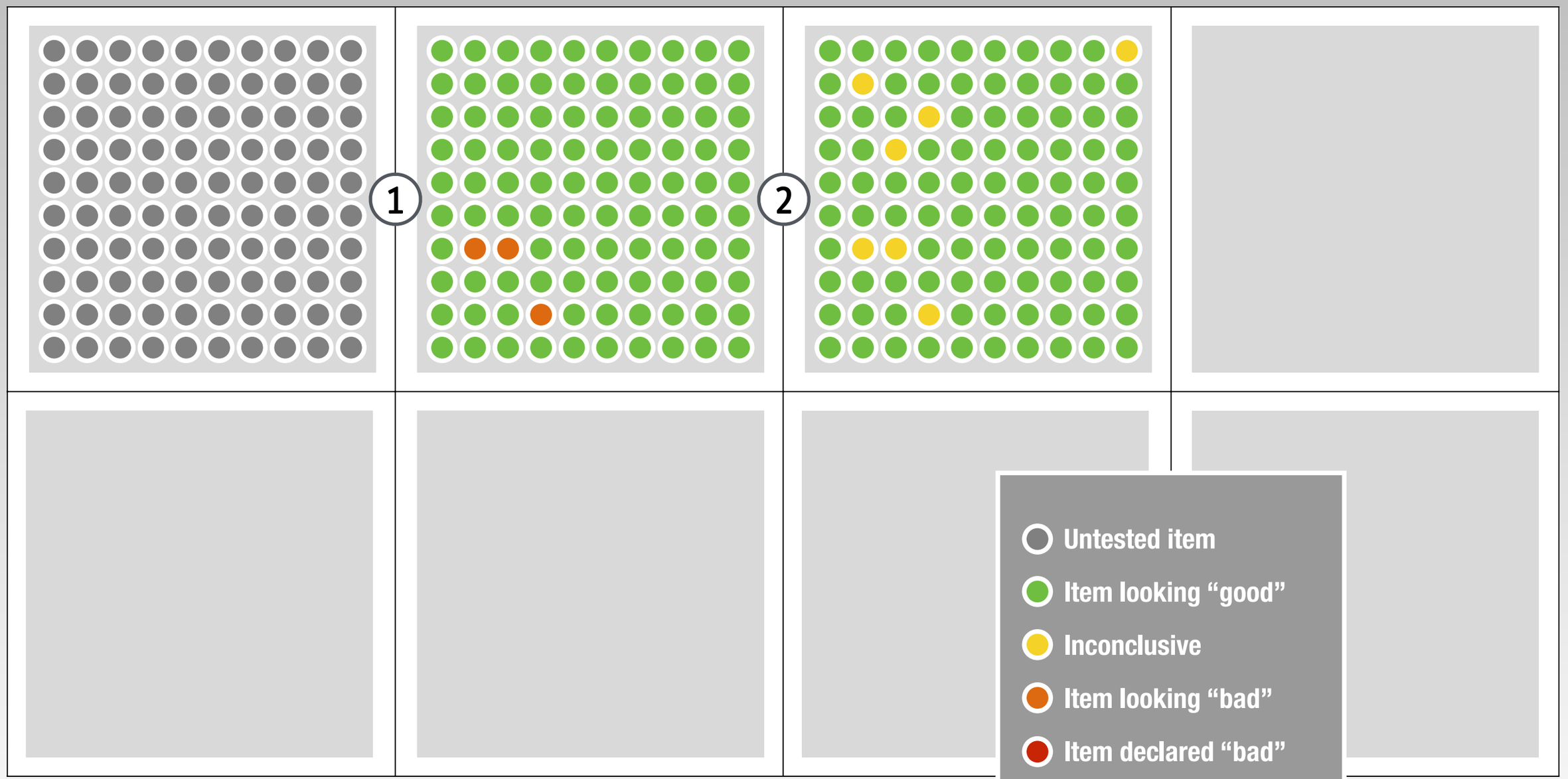
Honest Host

Default parameter set: $(\alpha = 0.95, \beta = 0.05, \alpha^* = 0.999, \beta^* = 0.01)$



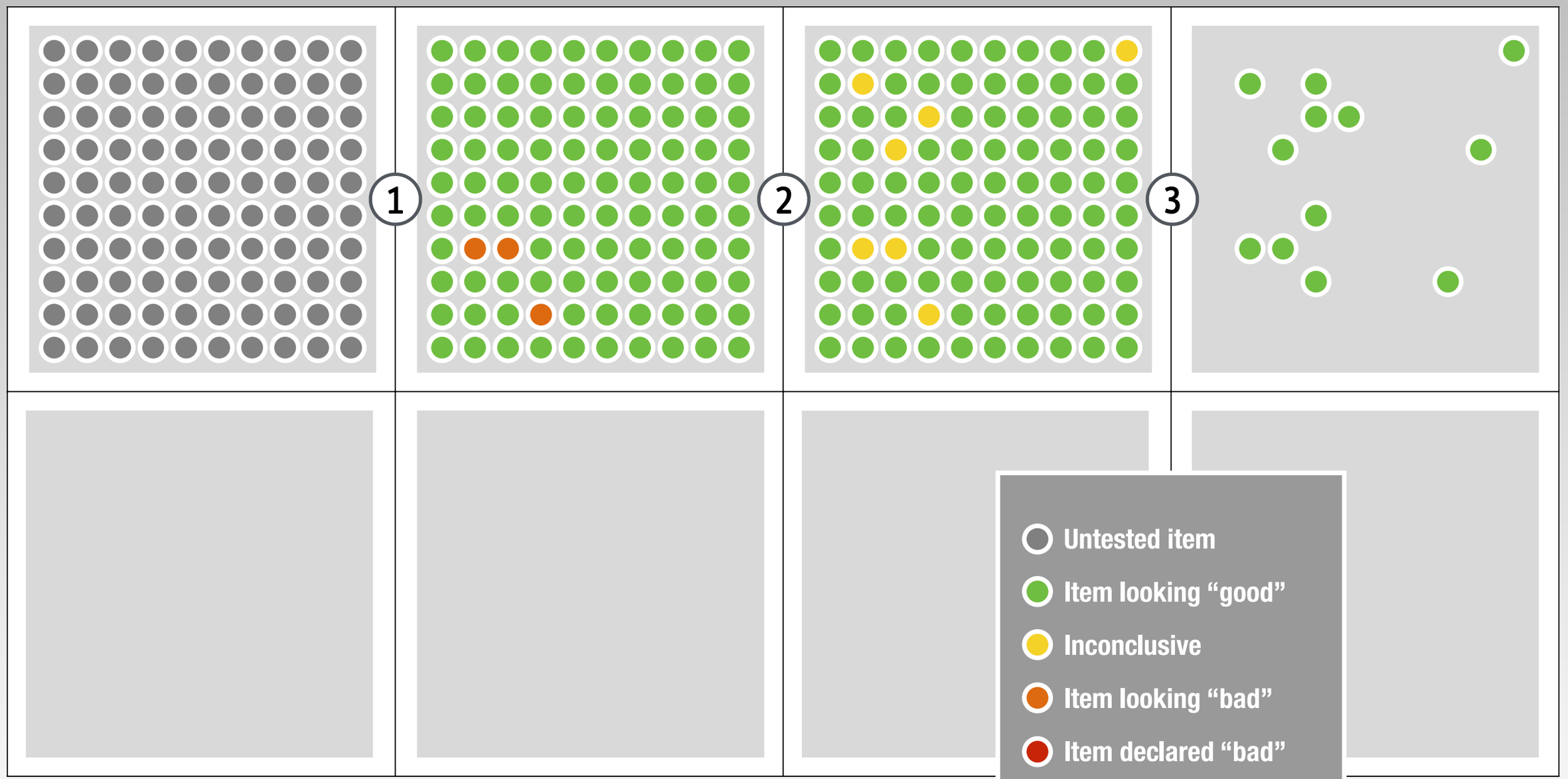
Honest Host

Default parameter set: $(\alpha = 0.95, \beta = 0.05, \alpha^* = 0.999, \beta^* = 0.01)$



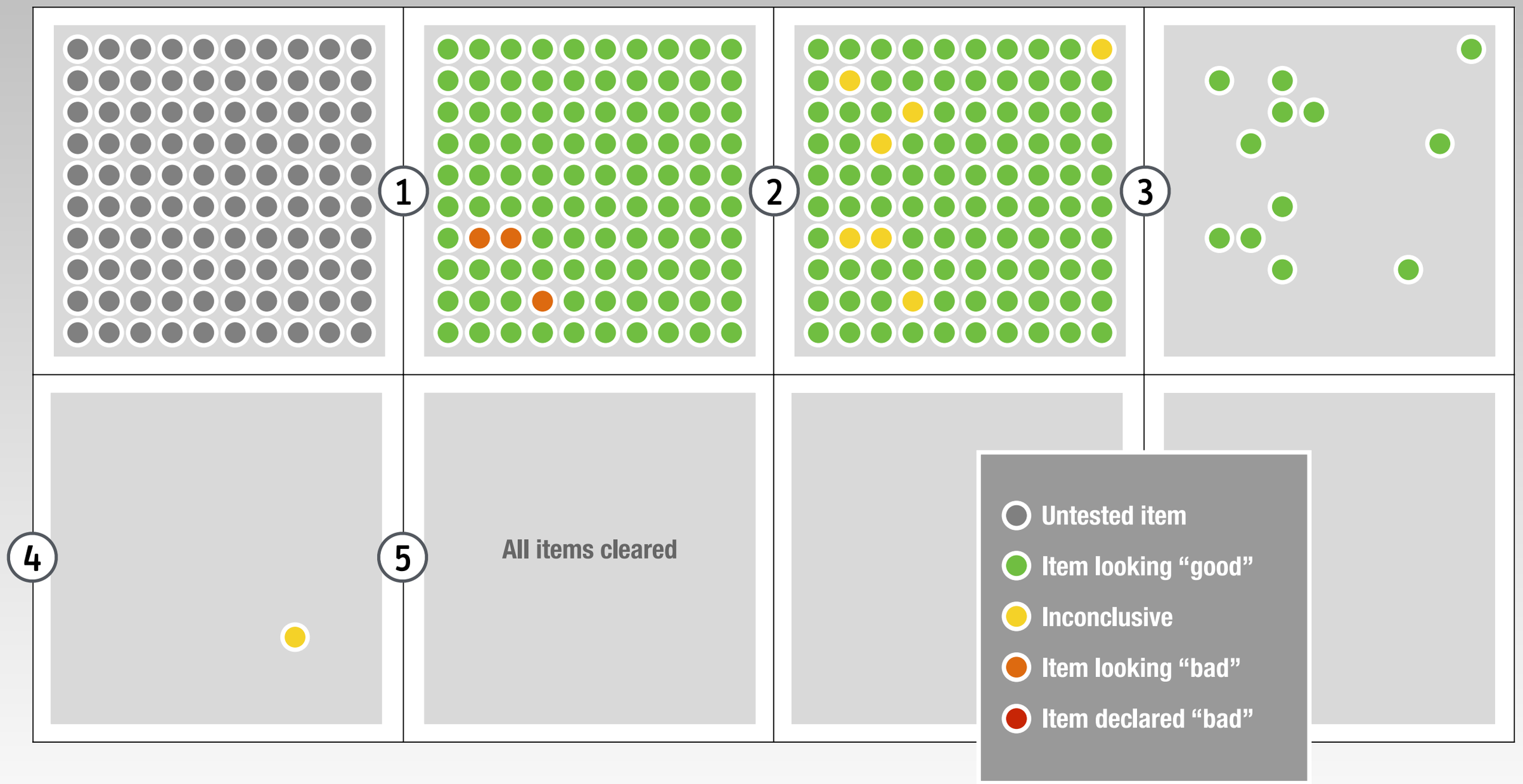
Honest Host

Default parameter set: $(\alpha = 0.95, \beta = 0.05, \alpha^* = 0.999, \beta^* = 0.01)$



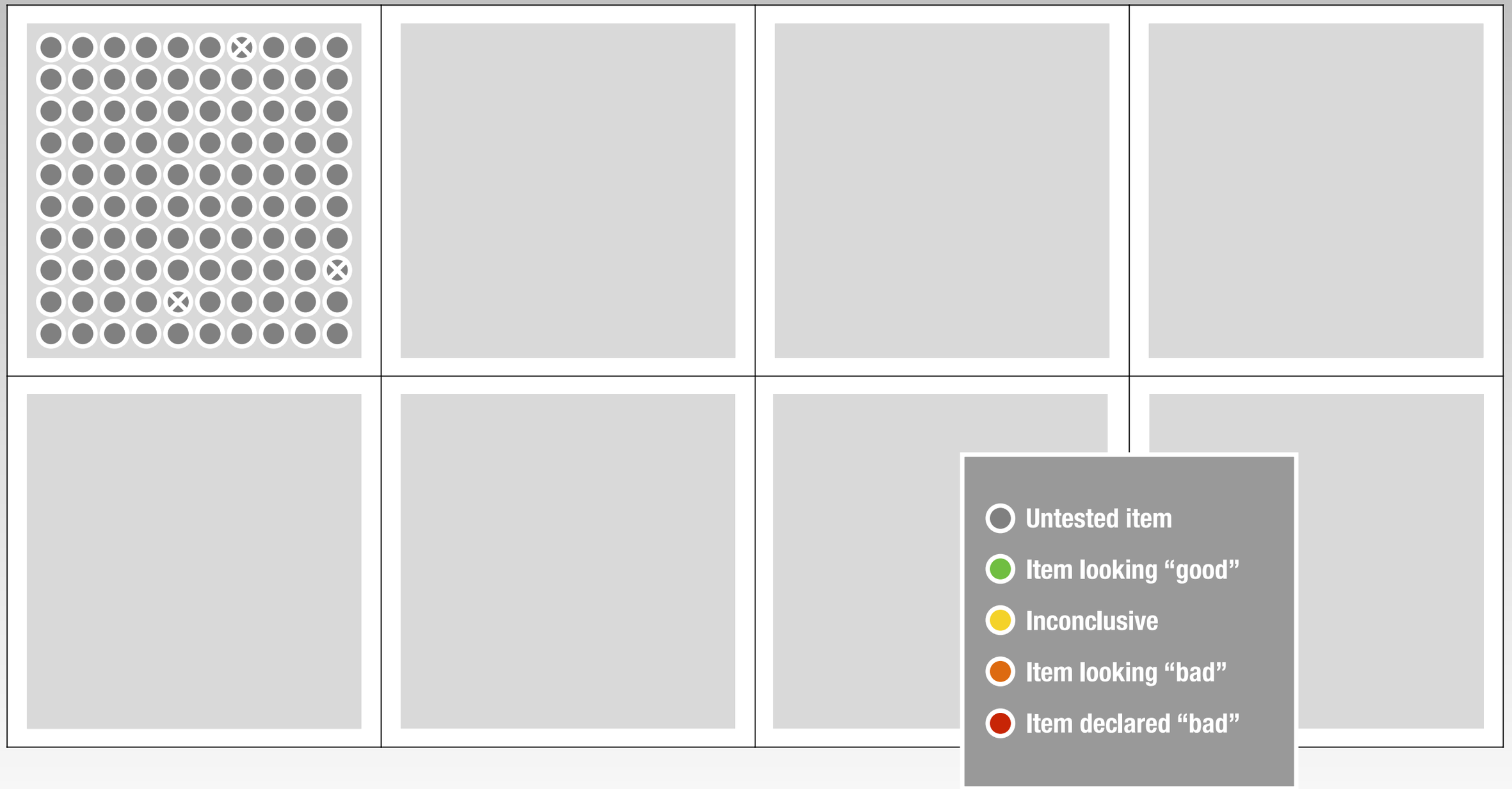
Honest Host

Default parameter set: $(\alpha = 0.95, \beta = 0.05, \alpha^* = 0.999, \beta^* = 0.01)$



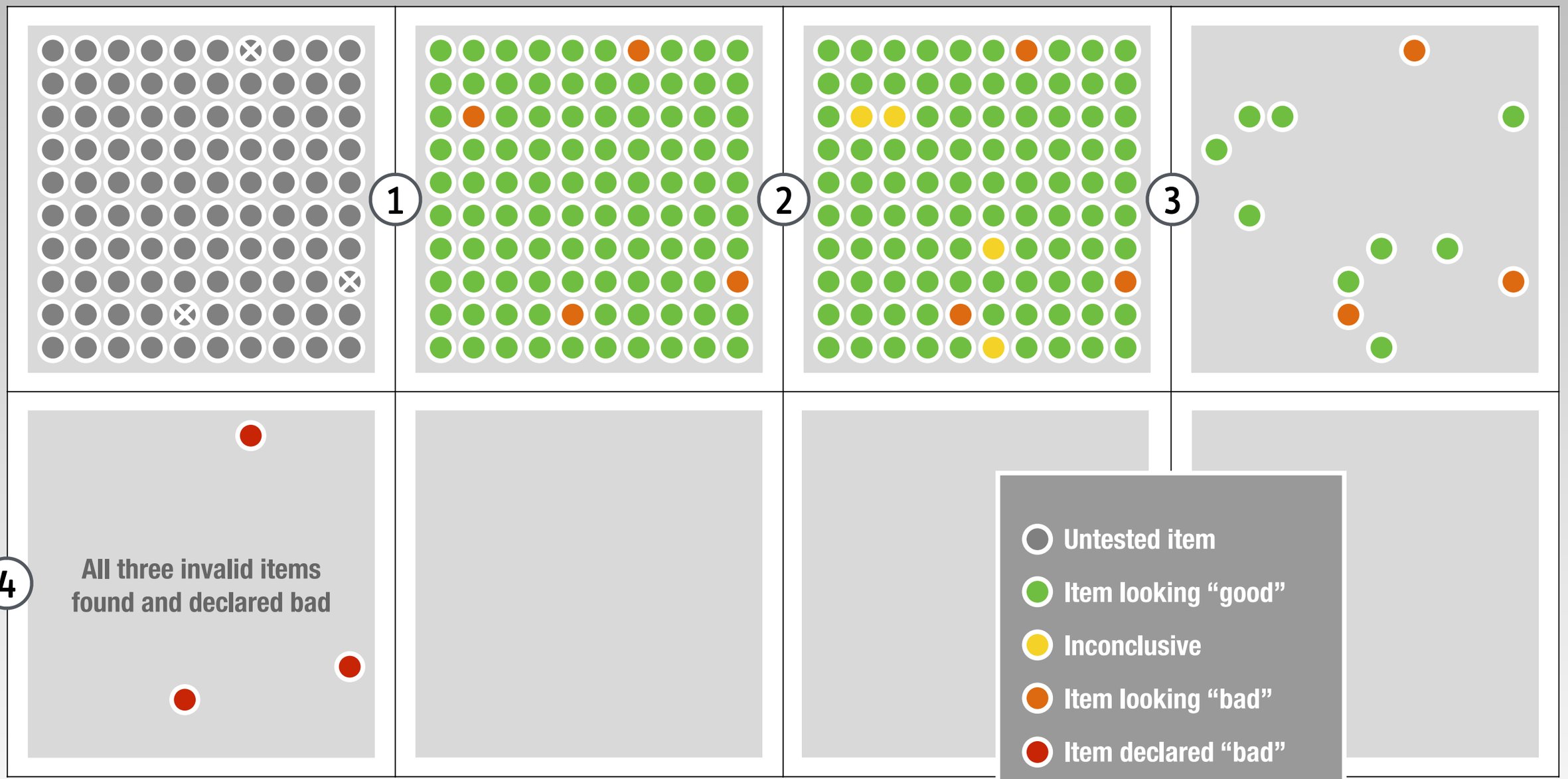
Dishonest Host

Default parameter set: $(\alpha = 0.95, \beta = 0.05, \alpha^* = 0.999, \beta^* = 0.01)$



Dishonest Host

Default parameter set: $(\alpha = 0.95, \beta = 0.05, \alpha^* = 0.999, \beta^* = 0.01)$



Lookup Table for Inspector

(who has to assume dishonest host and worries about invalids declared “good”)

Probability that invalid item passes basic test (β)	0.01	0.05	0.10	0.20	0.30	0.40	0.50
Average number of tests per item (99% detection probability for invalid items)	2.1	3.2	4.2	6.3	8.4	13.7	17.9

Similar lookup table(s) can be constructed for the host

who worries about valid items being declared “bad”

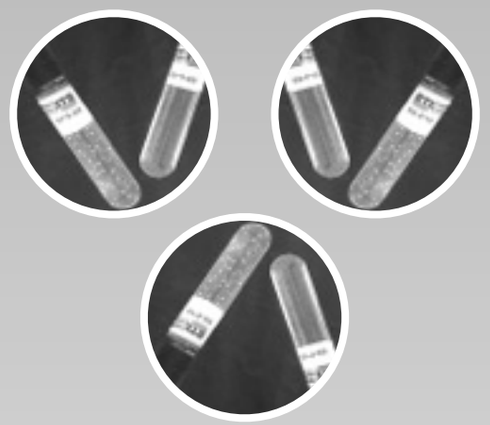
Zero-knowledge Template Approach

(Glaser, Barak, Goldston, *Nature*, 510, June 2014)

Zero-knowledge Template Approach

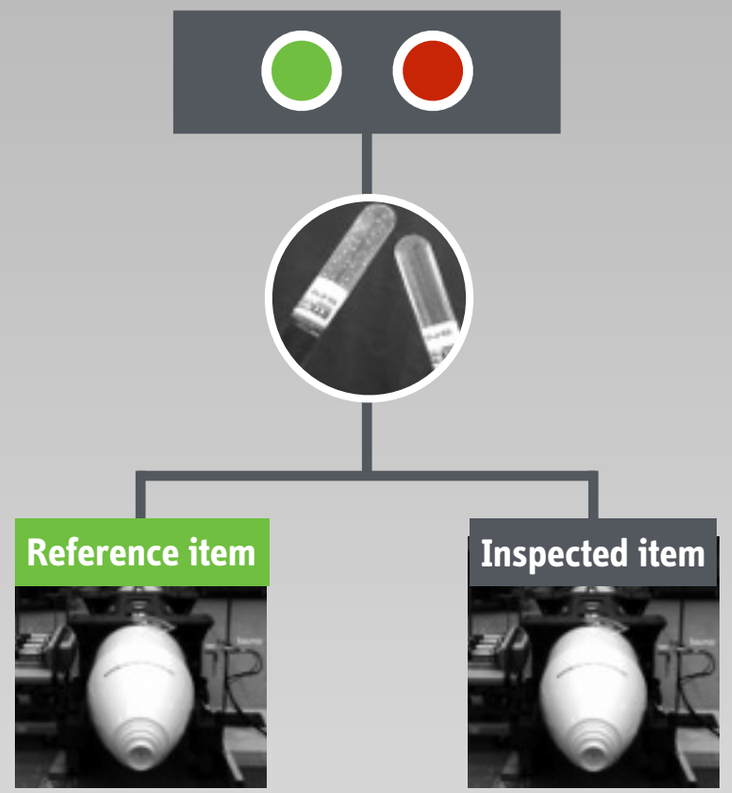
(with preloaded non-electronic detectors, simplified)

1



Preloading detector pairs
(host only)

2



Inspecting item together with reference item
(inspector chooses which preloaded detectors are used with which item)

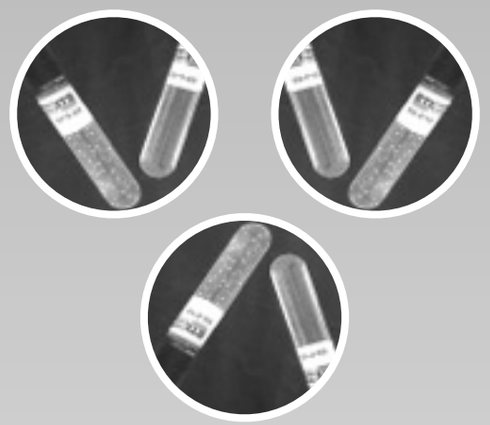
In the basic one-on-one approach, reference item needs to be present for every inspection

(Can we authenticate N items without comparing them to the template each and every time?)

Zero-knowledge Template Approach

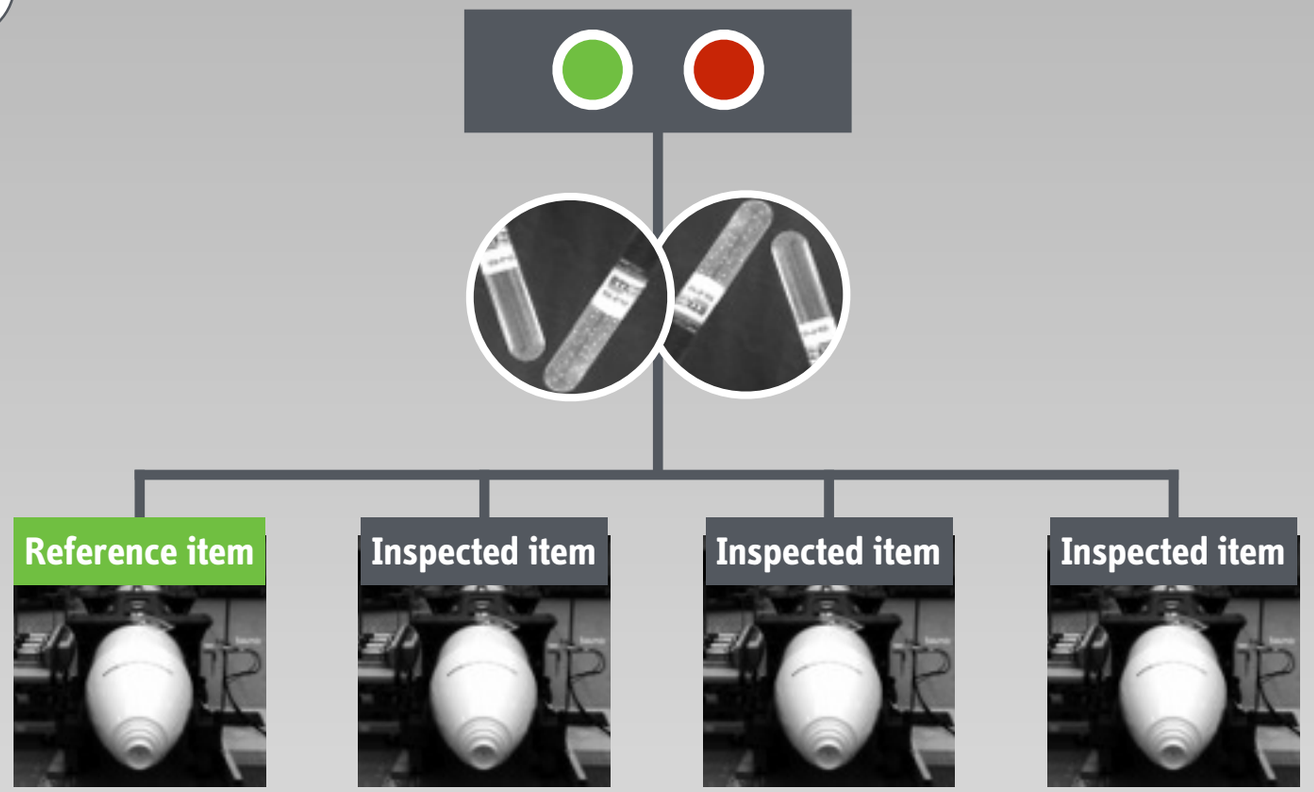
(with preloaded non-electronic detectors, simplified)

1



Preloading detector pairs
(host only)

2



Inspecting item together with reference item
(inspector chooses which preloaded detectors are used with which item)

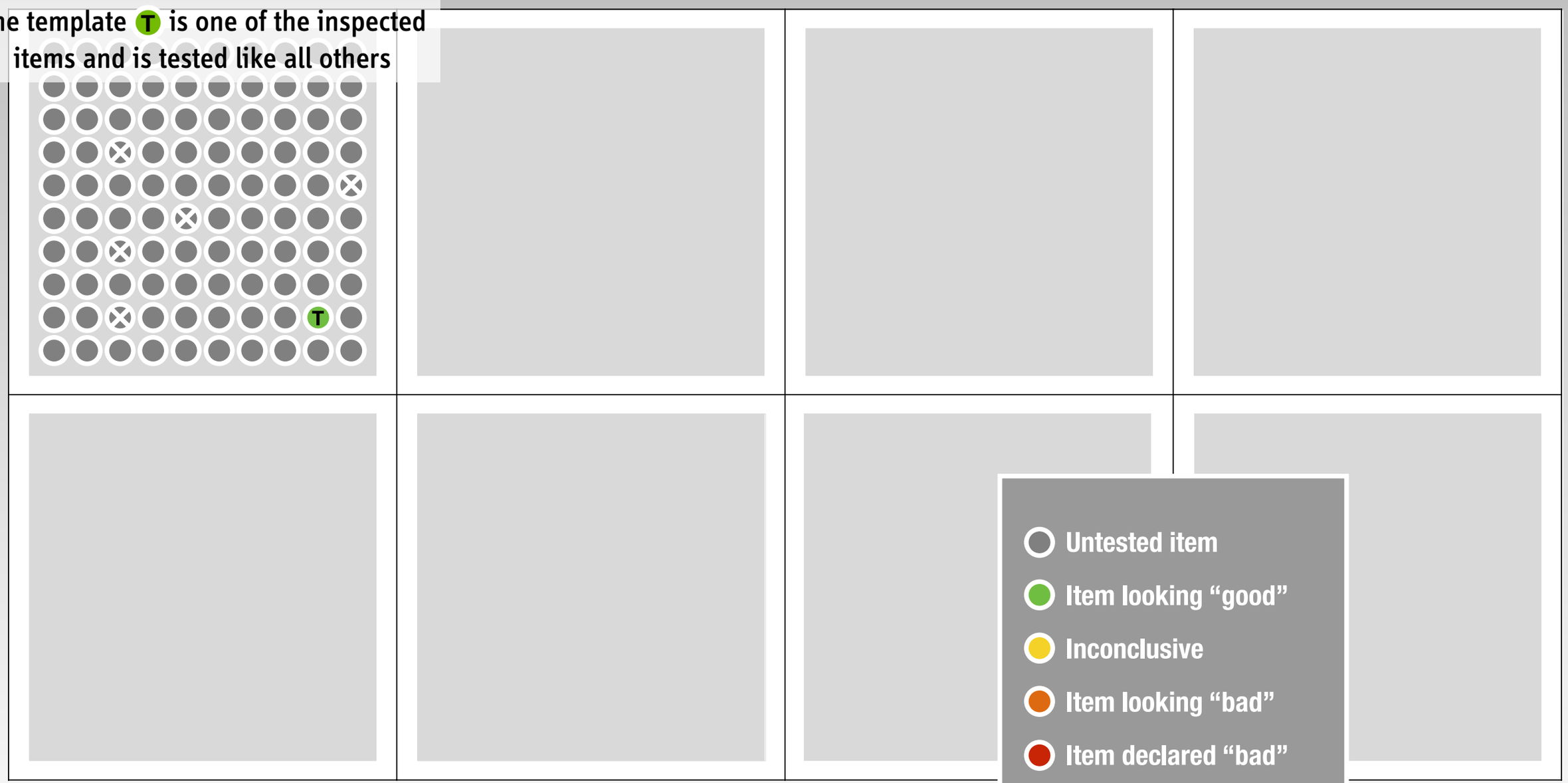
In the basic one-on-one approach, reference item needs to be present for every inspection

(Can we authenticate N items without comparing them to the template each and every time?)

Zero-Knowledge Template Approach

Default parameter set: $(\alpha = 0.95, \beta = 0.05, \alpha^* = 0.999, \beta^* = 0.01)$

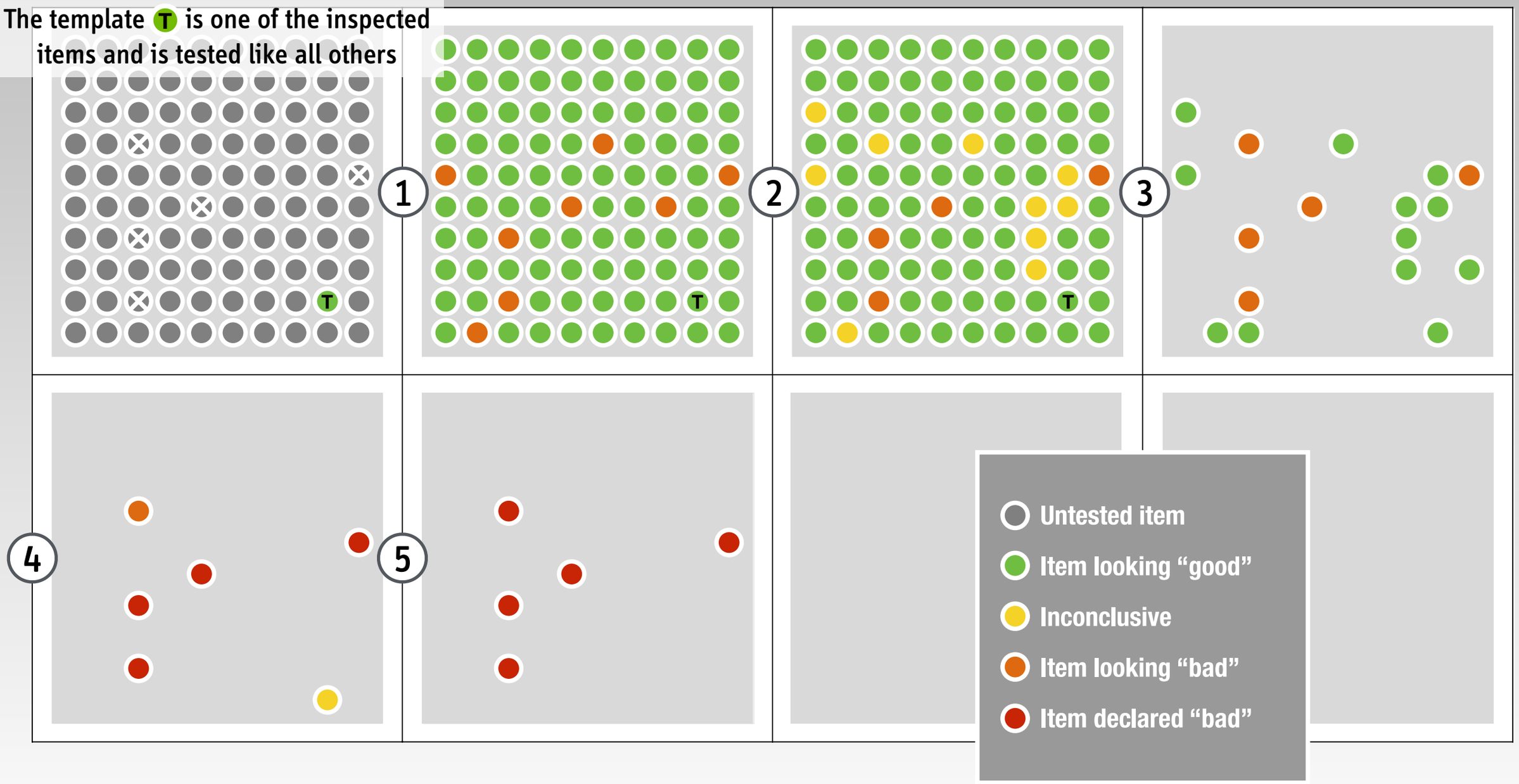
The template (T) is one of the inspected items and is tested like all others



Zero-Knowledge Template Approach

Default parameter set: $(\alpha = 0.95, \beta = 0.05, \alpha^* = 0.999, \beta^* = 0.01)$

The template (T) is one of the inspected items and is tested like all others



Conclusion

Summary and Conclusion

Proposed Testing Protocol

Protocol can be used to control the probabilities of inspection failures
for successful nuclear warhead authentication

Protocol allows the host and the inspector to adjust testing parameters such that
both achieve their desired performance requirements with a minimum inspection effort

Protocol can be used for both attribute-type and template-type measurements

Findings

Critical parameter is the fraction of invalid items passing a basic test (single measurement)

This parameter depends on diversion scenario, cannot be "defined" independently
up to 20–30% may be acceptable; above this range, number of repetitions increases sharply

Future work needs to determine tradeoffs between repeated tests and longer individual inspections