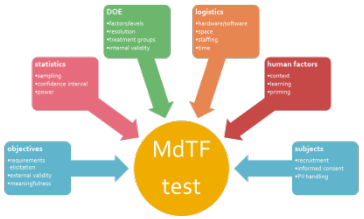
Efficient test design for biometric exit scenarios.

Yevgeniy B. Sirotin, Ph.D.

Outline

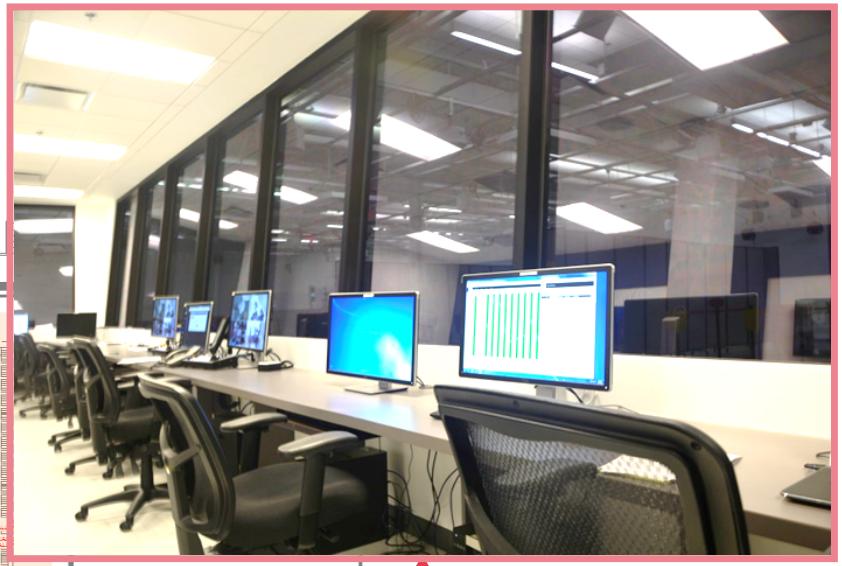
- Maryland Test Facility
  - Facilities and Capabilities

- Test Design Process
  - Balancing multiple requirements and constraints within a single test



The Maryland Test Facility

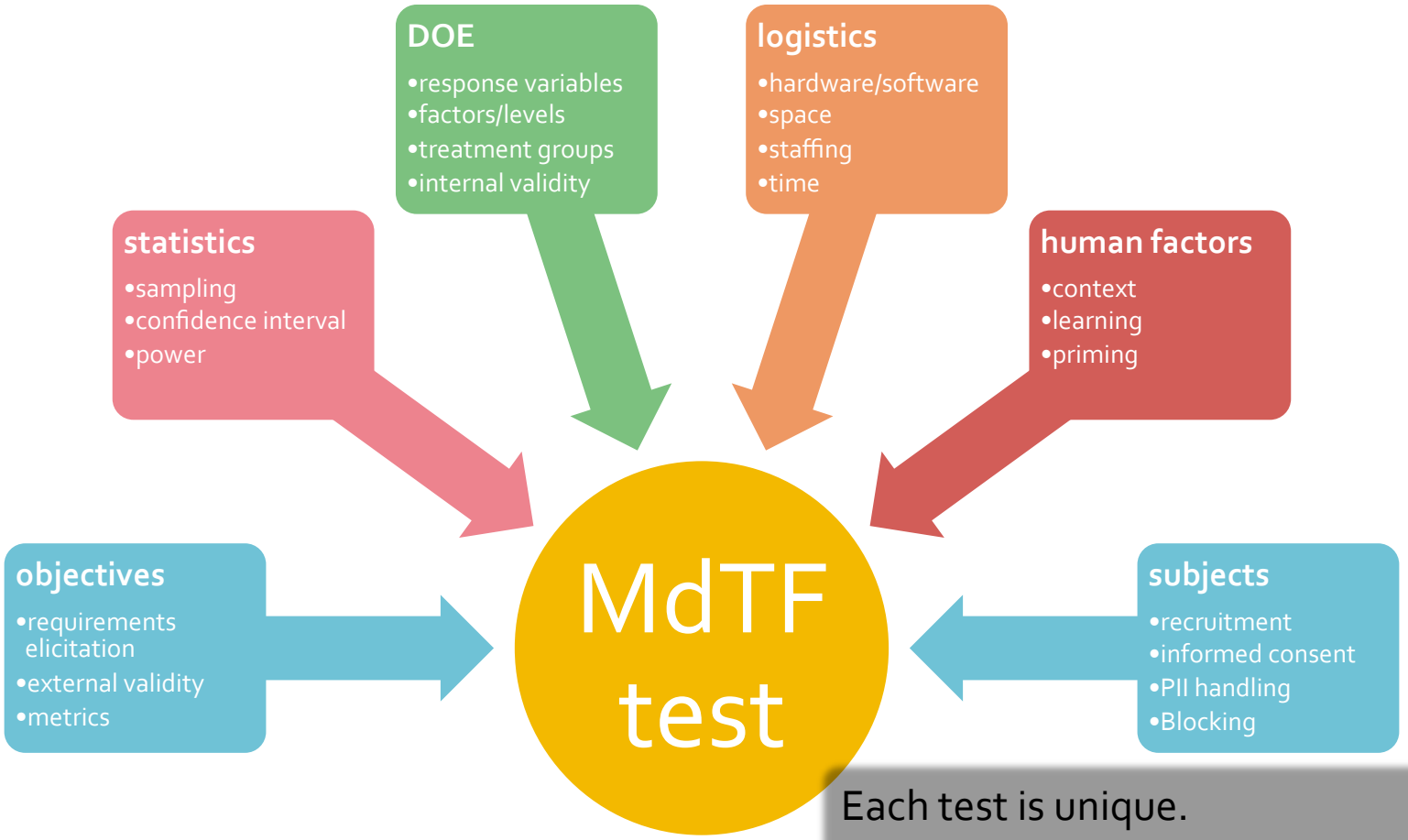
- 20,000+ ft² of office and laboratory space
- Video Recording
  - 16 PTZ and Zoom video cameras
- Eye Tracking
- Environmental sensors
  - ambient light, noise, humidity, …
- Control center
  - real-time monitoring and analysis



The Maryland Test Facility

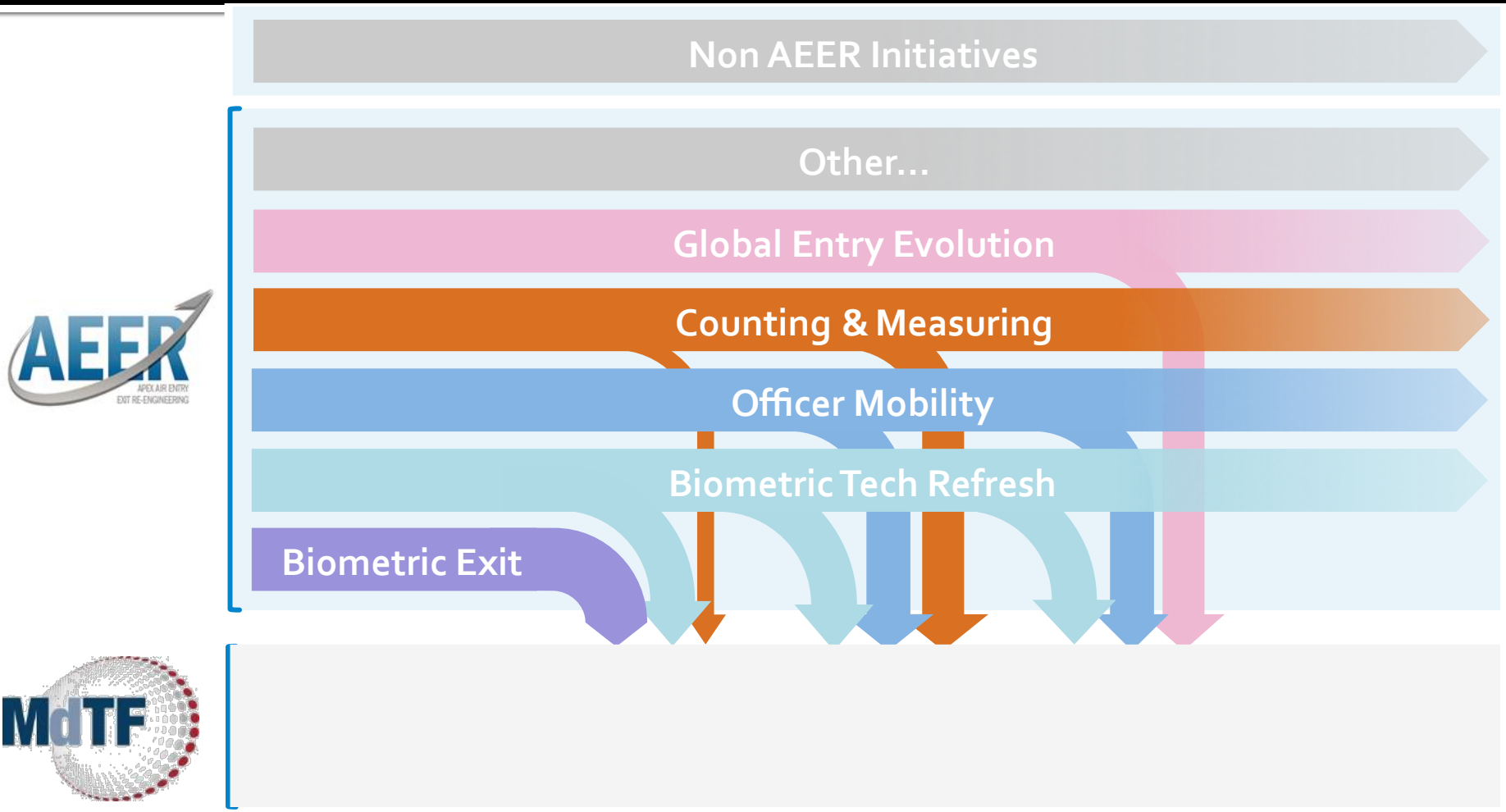
- Private briefing & debriefing rooms
  - Informed consent: closed door, white noise
- Work areas
  - Focus Groups
  - Workshops
  - Training
- Laboratory space

Elements of an MdTF test



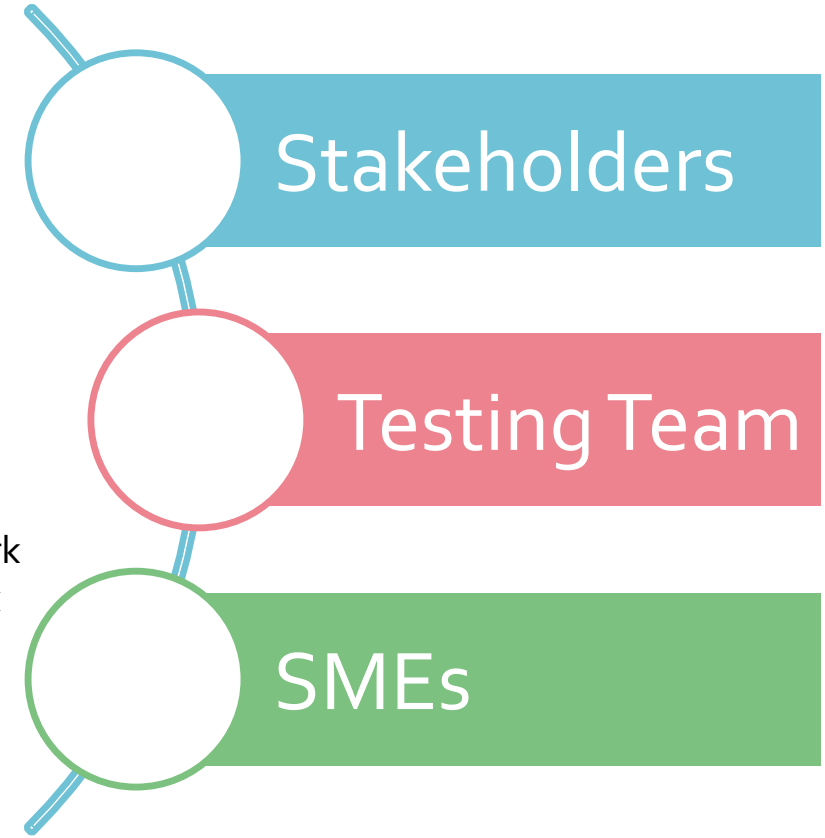
Each test is unique.
There is no common approach that will work for all tests.

MdTF Sequence Testing



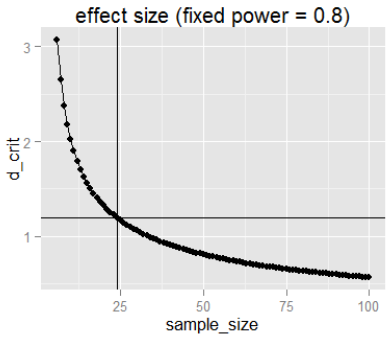
Identifying test objectives

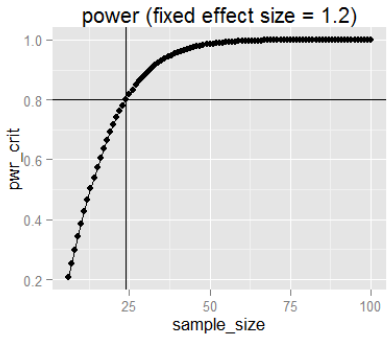
- Why is the test being performed?
  - Single process
    - validation – demonstrate function
    - exploration – identifying important factors
    - benchmarking – compare to specific criteria
    - refinement – identify best factor levels
  - Different processes
    - differentiation – identify differences
    - comparison – compare along specific metrics
- What is the desired outcome?
  - data – obtain specific data
  - knowledge – know why it works / does not work
  - demonstration – can show it can / cannot work
- What will happen as a result of the test?
  - recommendations
  - further analysis
  - down-selection

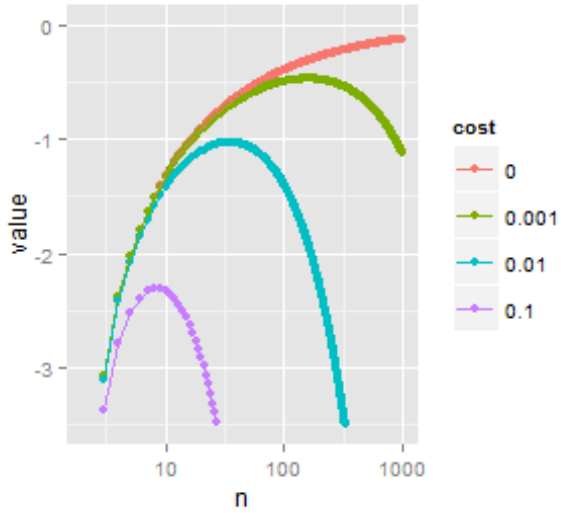


Statistical considerations

- Identify desired effect size (E) or confidence interval
  - Percentage? Benchmark? Critical difference?
- Power
  - Usually ≥ 0.8 (80% chance of detecting effect, when present)
  - Depends on sample size (n) and effect size
- Sample size
  - **Ideal world:** Sample size determined from power and effect size
  - **Reality:** Sample size drives cost (C = **$** & **t**), tradeoff depends on test objectives
- Test Value = -[E(n) + C(n)]





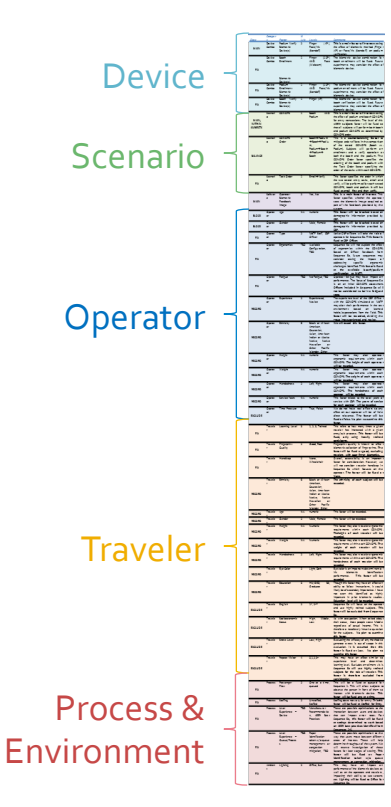


Note: cost function need not be linear and may have discontinuities, but good to be aware of ballpark when scoping the test.

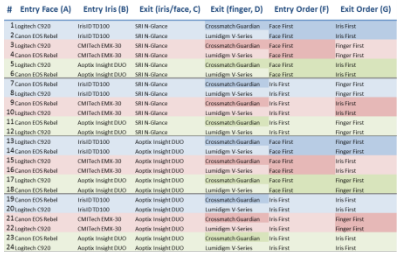
Design of Experiments

- Formalize response variables
  - E.g. Timing, performance, user feedback
- Identify factors and levels.
  - Build factor list (use stakeholders & SMEs)
  - Identify factor categories:
    - Device, Scenario, Operator, Subject, Process, Environment
  - Identify manipulated, fixed, or blocked factors
  - Do once and use many times
- We often include counterbalancing factors.
- Can everything be tested at once?
  - Treatment = tested factor/level combination
  - Use fractional factorial designs to reduce number of treatments needed
  - Use separate sub-experiments to reduce design complexity

Factors/Levels

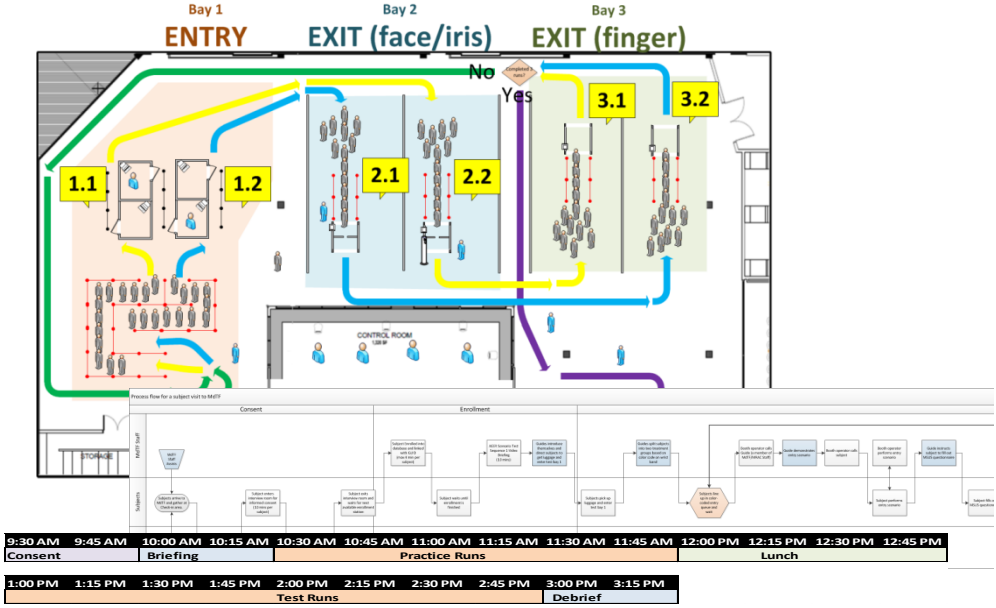


Treatments



Planning Logistics

- Space and movement
  - Facility layout
  - Movement plan
  - Process flowchart
- Time budget
  - Estimate time for each activity
    - Informed Consent, Transactions, Surveys, etc.
  - Generate test day schedule
- Time reduction through parallelization
  - Cost: more complex movements
- Staffing
  - Define test staff roles
  - Create and execute staff training sessions
  - Conduct walkthroughs
- Hardware/software
  - Build required test hardware
  - Design required test software
  - Test hardware + software



Human Factors Considerations

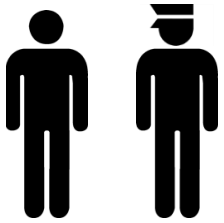
- Context
  - Video orientation / PPT Briefing
  - Context-appropriate props
    - Booths, podiums, stanchions, baggage, etc…
  - Scripted interactions
- Learning
  - Assess learning using repeated transactions
  - Mitigate learning via practice sessions
  - Simulate state of target population
    - E.g. 56% of travelers have not previously been to US
- Priming
  - Avoid pre exposure to test system
  - E.g. acquire iris images for ground-truth using non-tested collection method
  - Avoid testing multiple similar processes in sequence
    - Counterbalancing schemes
    - Make factor between subjects instead of within subjects
      - Better measure less of what is relevant than more of what is irrelevant

Test Subjects

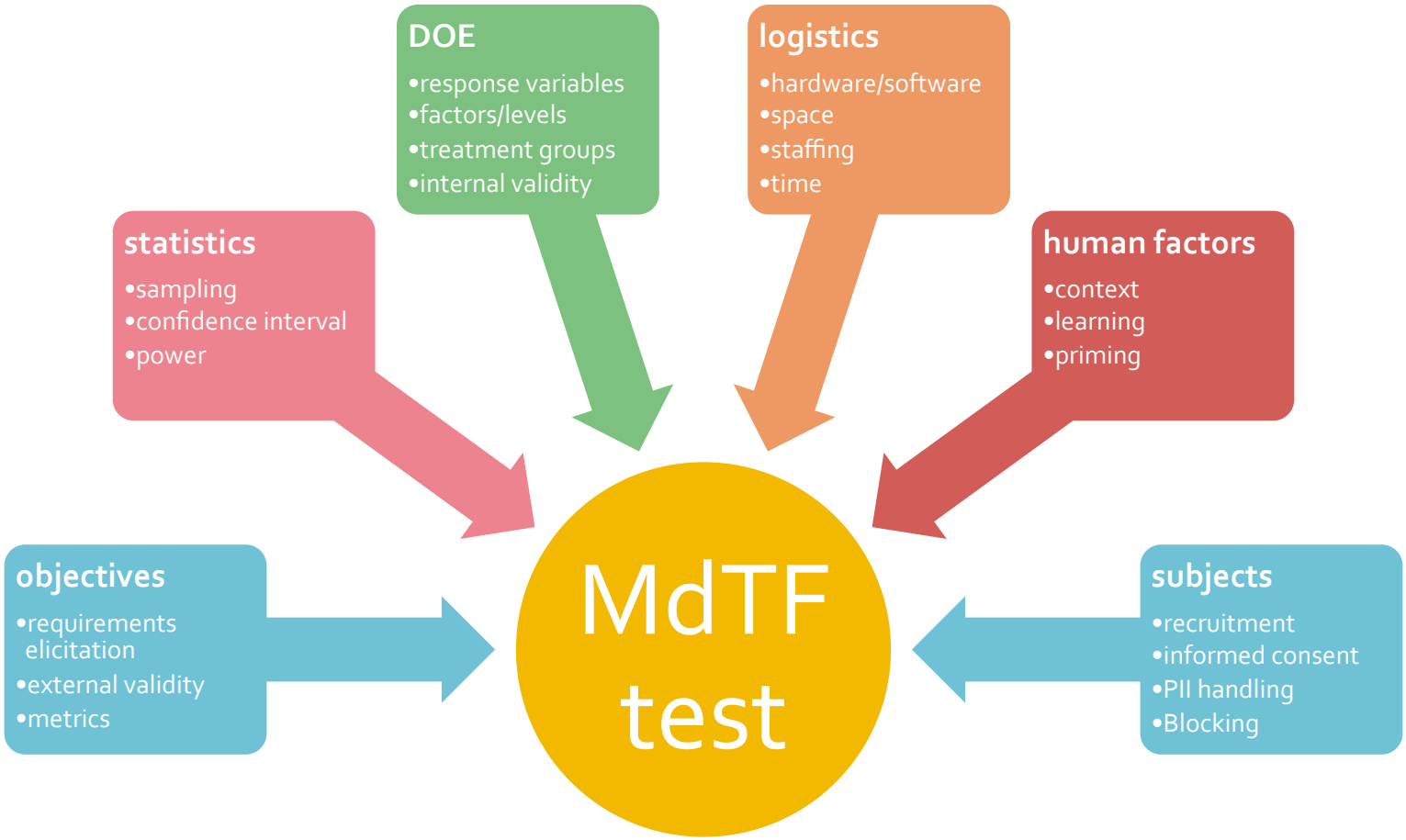
- Identify test population
  - What are the demographics?
  - What do the subjects know about the tested process?

- Develop blocks
  - Amount of stratification depends on block group size
  - Identify factors for blocking
    - May need to down-select based on importance
      - Age, gender, eye color, experience level, etc…
    - We typically block on no more than 3-4 factors for group size ~30

- Recruiting
  - Different populations require different approaches
    - Local population – for public facing devices
    - Professional population – for operator interface design
  - Population depletion?
    - Over time, becomes hard to find naïve subjects



Conclusions



Thank You!

The research for this paper was fully funded by the Department of Homeland Security Science and Technology Directorate on contract number W911NF-13-D-0006-0003. The views presented here are those of the authors and do not represent those of the Department of Homeland Security or of the U.S. Government.