UNIVERSITY OF TRENTO

THE ROLE OF CATALOGUES OF THREATS AND SECURITY CONTROLS IN SECURITY RISK ASSESSMENT: AN EMPIRICAL STUDY WITH ATM PROFESSIONALS

Joint work with
M. De Gramatica,
K. Labunets,
F. Paci,
A. Tedeschi

Fabio Massacci
University of Trento – Italy
securitylab.disi.unitn.it

THE PROBLEM

EMFASE

UNIVERSITY OF TRENTO

- **ENAV – national large air traffic management authority**
  - New cyber operational concepts (eg Remotely Operated Tower)
  - Must identify cyber security measures
  - Lots of domain specialists but few security experts
- **Use a Risk Based Methodology, but which one?**
- ✓ ISACA's CoBIT, SABSA → focus on Business Goals
- ✓ US NIST 800-53, UK's IAS → focus on Threats
- ✓ Eurocontrol or SESAR's SecRAM → focus on Assets
- **Who should execute the methodology?**
- ✓ Ask security experts → they are expensive and in high demand.
- ✓ Use a threat/controls catalogue →which one? is bigger = better?
  - ISO27002→ general measures
  - German's BSI → general measures with specific details
  - Eurocontrol's Risk Toolkit → specific to Air Traffic Management

WHAT IS A SECURITY RISK ASSESSMENT METHOD?

EMFASE

UNIVERSITY OF TRENTO

- **A SRA method**
  - ✓ examines system's security threats
  - ✓ proposes set of system changes (security measures, controls, requirements)
  - ✓ to bring system within acceptable risk
- **Example statements**
  - X helps enterprises create optimal value from IT by maintaining a balance between realising benefits and optimising risk levels and resource use.
  - Y is a proven methodology for developing business-driven, risk and opportunity focused Security Architectures
  - W collect business requirements from risk owners and budget holders. Abstract them in business-language into business drivers for security then execute and measure value
  - The aim and purpose of Z is to analyse a proposed or existing system to identify risks and estimate the levels of those risks; Select appropriate controls to manage the treatable risks.

SIZE MATTERS…

EMFASE

UNIVERSITY OF TRENTO

❑ German's IT-Grundschutz Catalogue (aka BSI)

- ✓ Intro → 40 pages
- ✓ Assets → 350 pages
- ✓ Threats → 1.000 pages
- ✓ Controls → 3.000 pages

❑ Eurocontrol's ATM Security Risk Management Toolkit

- ✓ Guidance Material → 100 pages
- ✓ ATM specific Threats → 57 pages
- ✓ ATM specific controls → 99 pages (pre 72 + 27 post)

❑ Remotely Operated Tower Scenario

- ✓ Operational Focus Area Description → 100+ pages
- ✓ Short "essential" description → 24 pages

RESEARCH QUESTIONS

EMFASE

UNIVERSITY OF TRENTO

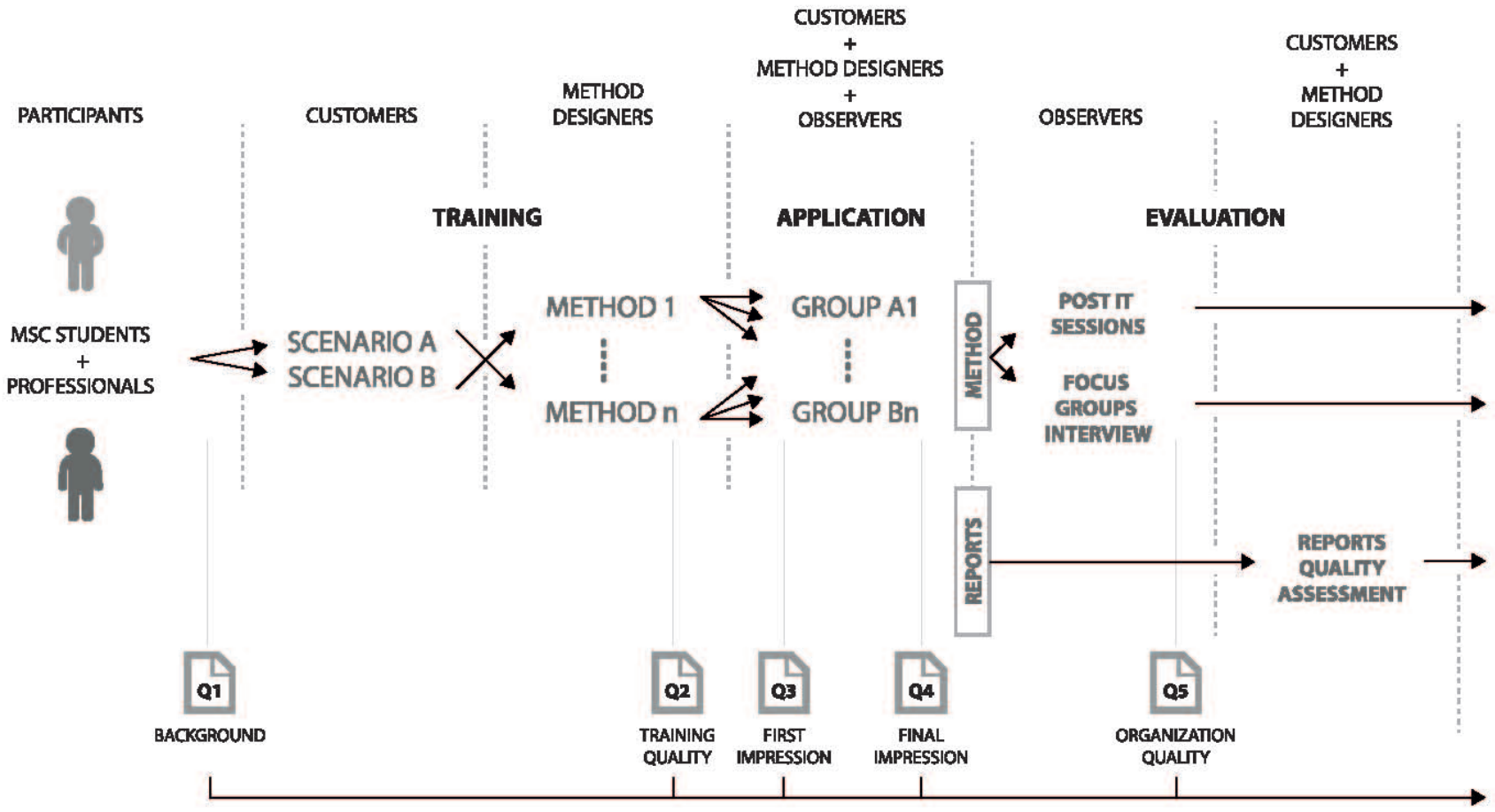
❑ How do we evaluate a method with a catalogue?

- ✓ A clinical procedure with a physician's desk reference (aka drug index) → We know well how to evaluate it →clinical trial protocol
- ✓ A risk assessment procedure with security catalogue → Same idea

❑ Research question: do catalogues make a difference?

- ✓ Evaluate 15 ATM Professionals applying an ATM Risk Method
  - Domain **Security** Experts **without** catalogue
  - Domain Experts **with a generic** catalogue (BSI)
  - Domain Experts **with a specific** catalogue (Eurocontrol)
- ✓ We apply a **trial protocol** to estimate the **efficacy** of a methodology with and without catalogues

OUR EXPERIMENTAL PROTOCOL



PROTOCOL ACTORS

EMFASE

UNIVERSITY OF TRENTO

❑ Participant

✓ Important to have both students (novices to the treatment but unbiased opinion) and practitioners (expert but may have prejudices on what it works).

- Students → preliminary pilot
- ATM Practitioners → THIS Paper at REFSQ 2015

❑ Designer → Expert in the method

✓ Provide the best possible training for the method.

- Security Trainer at Eurocontrol

❑ Customer → Expert in the scenario

✓ Indipendent validation of quality of results (irrespective of treatment!)

- Any method can produce "enough" security requirements if quality doesn't matter.

✓ Expert in method ≠ Expert in domain → former may give good score if method is followed → bias

PROTOCOL STEPS

EMFASE

UNIVERSITY OF TRENTO

❑ Training Participants

- ✓ Designer(s) train on treatment
- ✓ Customer(s) describe scenario

❑ Application

- ✓ Participants apply treatment for a **time span that is sufficiently long (>1day) to be challenging**

❑ Evaluation

- ✓ **Customers evaluate results of participants**
- ✓ Participants tell their opinion on how the experiment went
- ✓ ~~Designers evalute results~~

❑ No initial learning bias

- ✓ Avoid my stuff vs competitor's stuff
- ✓ Not wrong focus

❑ Can test actual efficacy

- ✓ Experiments <1h too short to tell results apart
- ✓ With large catalogues/scenarios 1h not enough even to browse docs

❑ Measures different things

- ✓ Customers → actual efficacy
- ✓ Participants → perceived efficacy
- ✓ Designers should only evaluate compliance

PROTOCOL MEASUREMENTS

❑ Actual Efficacy

✓ Participants Reports

- Quantitative (#threats/controls) → easy to generate huge numbers (of junk)
- Qualitative Analysis →likert scale

❑ Perceived Efficacy

✓ Questionnaires → likert scale

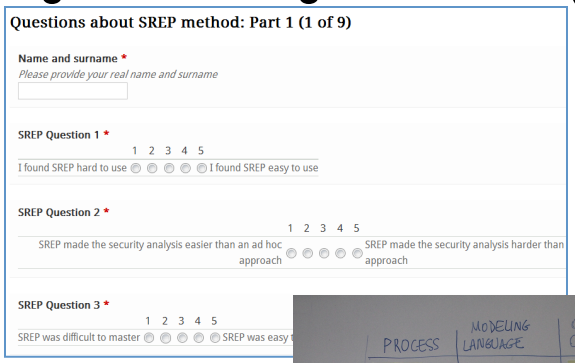
❑ Qualitative analysis

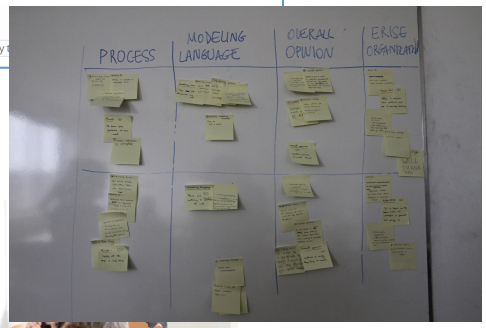
✓ Post-it notes

- Affinity Analysis

✓ Focus Groups Interviews

- Coding, qualitative analysis





PROTOCOL MEASUREMENTS - II

EMFASE

UNIVERSITY OF TRENTO

❑ Actual Efficacy - AE

- ✓ whether the treatment improves performance of the task

❑ Perceived Efficacy – PE

- ✓ Perceived Ease Of Use – PEOU
  - the degree to which a person believes that using a treatment would be free of effort
- ✓ Perceived Usefulness – PU
  - the degree to which a person believes that a treatment will be effective in achieving its intended objectives

❑ Qualitative Feedback

❑ AE Null Hypothesis

- ✓ No difference between the treatements in identified risk/ controls measured as
  - 5 point scale of expert evalutaion

❑ PE Null Hypothesis

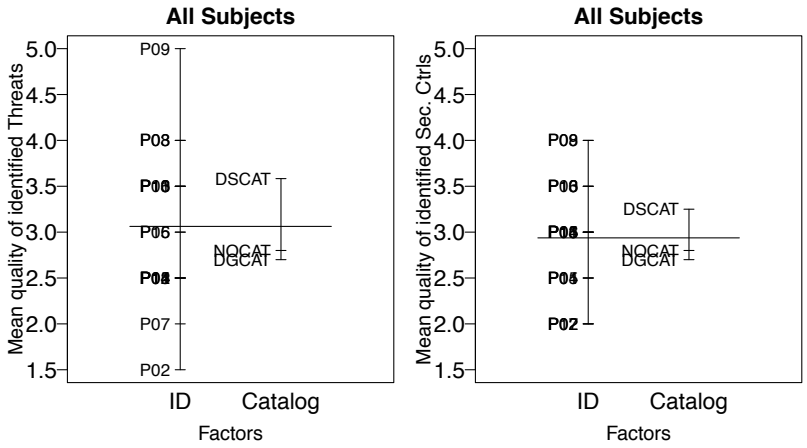
- ✓ No difference between the perceived efficacy (PEOU, PU) by the participants measured as
  - 5-point scale on questionnaire about identifying threats
  - 5-point scale on questionnaire about identifying security measures

ACTUAL EFFECTIVENESS

EMFASE

UNIVERSITY OF TRENTO

❑ Quality of identified threats and controls



❑ Median Scores (Controls): Expert+No Cat NoExpert+ATM Cat

- ✓ Actual Efficacy → 3.0 vs 3.5
- ✓ Perceived Ease of Use → 4.0 vs 3.0

❑ Summary

- ✓ No catalogues slightly worse results + Catalogues slightly worse to use
- ✓ No statistical diff

❑ Domain Experts + Security Catalogues ≈ (Domain + Security) Experts

WHY DO THINGS WORK?

EMFASE

UNIVERSITY OF TRENTO

- **Key features emerging from qualitative analysis**
  - ✓ Structure and Navigation
  - ✓ Coverage and Size (*)
  - ✓ Common Language (*)
  - ✓ Checklist
  - ✓ Quality of knowledge (*)
- **(*) present in qualitative study on which features are important in a ATM risk assessment**
  - ✓ 20 Experts working on Risk Assessement in SESAR
    - Labunets et al. SESAR's Innovation Conference 2014.
- **Big difference between expert and non-experts**
- **Experts**
  - ✓ Common Language
  - ✓ Checklist
    - "*The first step is to use your own experience and then to use the catalogue to cover generic aspects that could be forgotten*"
- **Not Experts**
  - ✓ Navigation is judge, jury and executioner
    - "*I saw people near to me; they were not able to find out stuff in the catalogue, they kept on getting lost in the pages and eventually they came up always with the same two or three items*"
    - "*Once identified the threat, finding out controls was really a mechanical work*"

SUMMARY OF EXPERIMENT

EMFASE

UNIVERSITY OF TRENTO

❑ Do catalogues work?

- ✓ MAYBE YES → not experts perfomed equally to experts without catalogues
- ✓ BUT → people work better with domain specific information
- ✓ AND → experts and non-experts use them in radically different ways

❑ Open Issues

- ✓ What about comprehensibility of results?
  - Risk assessment must be piped down the line for implementation
  - This was a critical issue when we interviewed stakeholders in ATM
- ✓ What about scaling to really large risk assessment?

❑ What is next

- ✓ More Info? → http://securitylab.disi.unitn.it
- ✓ Want to join the effort? → we are looking for replications

ADVERTISING

EMFASE

UNIVERSITY OF TRENTO

❑ We are hiring for a industry-academia lab

- ✓ European Electronic Crime Task force

❑ Positions

- ✓ 2 Phd Students → deadline 20May
- ✓ 2 Post-doctoral positions → open

❑ Further info on Trento

- ✓ Fabio.Massacci@unitn.it
- ✓ http://securitylab.disi.unitn.it
- ✓ http://en.wikipedia.org/wiki/European_Electronic_Crime_Task_Force

DISCUSSION OF TONG'S PAPER

EMFASE

UNIVERSITY OF TRENTO

- **What can we do now that we could not do before?**
  - ✓ Try to model different aspects of a design (from business to physical)
- **How sound is the solution?**
  - ✓ A formal model is behind the graphs → a minimal Tarski semantics exists
  - ✓ A small scale scenario was modelled by the author(s)
- **Whose goals are served or helped by this?**
  - ✓ One can fend off attacks across different layers
- **What is the next step to take?**
  - ✓ Demonstrate that you really capture and fold cross layers attacks
  - ✓ Address scale → goal models quickly evolve into "spaghetti" models
- **Controversial Question: How do you avoid the "beholder" effect?**
  - ✓ "Beauty is in the eye of the beholder"
  - ✓ You are the only one who really used it. How do we know it really works?