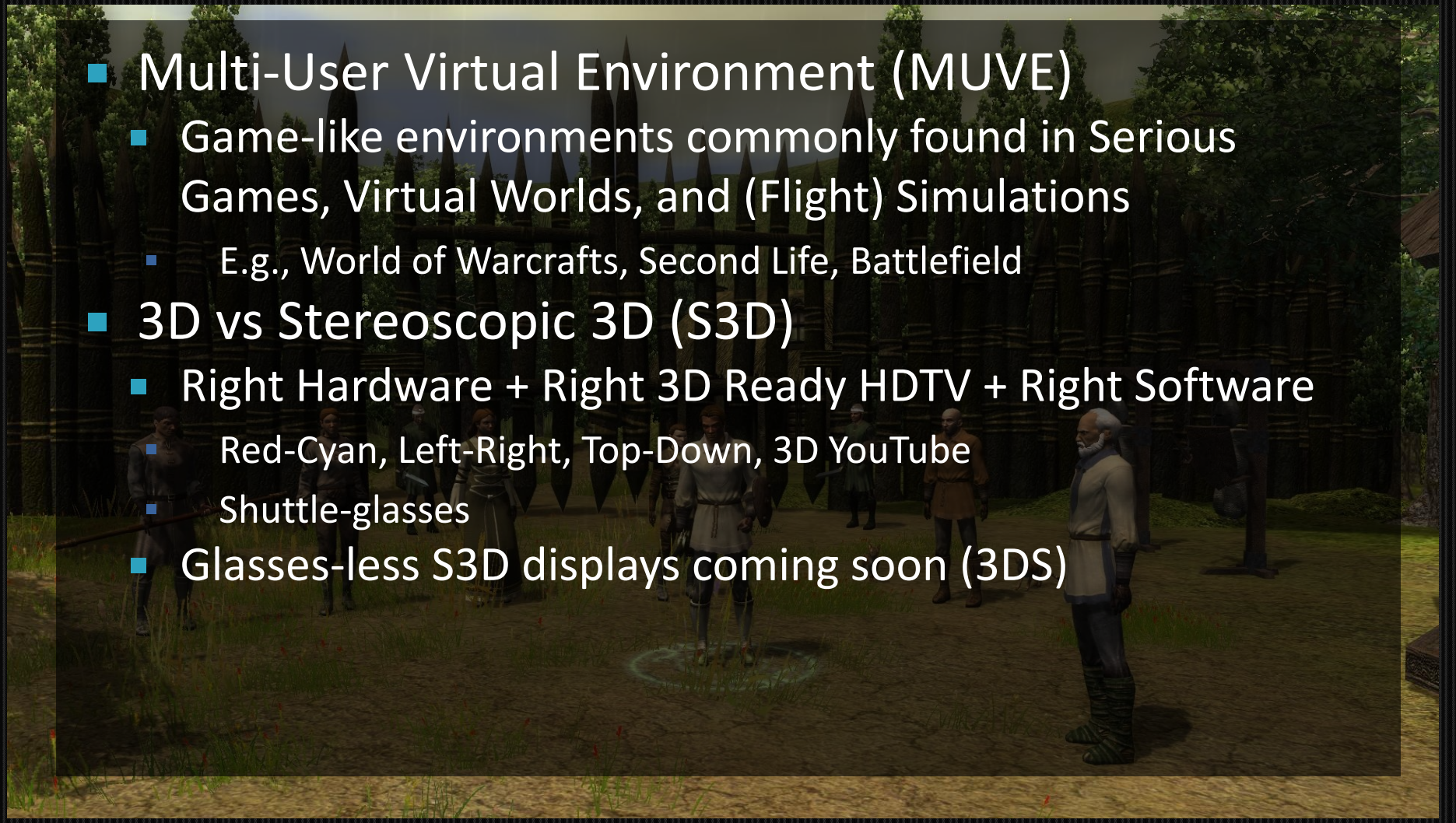
Assessment for Game-Based Learning using *in situ* Data Collection & Telemetry

Christian Sebastian Loh, Ph.D.
Virtual Environment Lab (V-Lab)
Southern Illinois University Carbondale, IL

MUVE



- Multi-User Virtual Environment (MUVE)
  - Game-like environments commonly found in Serious Games, Virtual Worlds, and (Flight) Simulations
    - E.g., World of Warcrafts, Second Life, Battlefield
- 3D vs Stereoscopic 3D (S3D)
  - Right Hardware + Right 3D Ready HDTV + Right Software
    - Red-Cyan, Left-Right, Top-Down, 3D YouTube
    - Shuttle-glasses
  - Glasses-less S3D displays coming soon (3DS)

Why 3D MUVE

- Advantages:
  - (Photo)Realism – new graphic engines
  - Spatial placement – 3D placement (depth)
  - Co-location – different locations (trainer/trainees)
  - Mitigate costs – save on transporting resources
  - = Interactive Learning (e-/m-Learning)
    - Anytime Anywhere
  - *Game* changing…

The 4 'P's of Training

- 3D MUVE affects several 'P's in training:
  - *Process* of Training
    - From boring, repetitive CBT-like 'drill and kill' to fun and engaging
    - E.g., Learn to speak foreign languages while having fun by playing 3D video games that teach what to say, how to say it, and when to say it

Immersive simulations of real life social communication

Interactive 3D video games involving spoken dialogs and cultural protocols with "socially intelligent virtual humans"

- *Proximity* of Training
  - We are no longer limited by physical locations or proximity to training resources
  - Social interaction in a global village (Virtual Worlds)
- *Place* of Training
  - Simulations can help (re-)create authentic but rare occurring situations (injection of events into real-life)
  - Final frontier, inner space, fantastic settings
- *Performance* of Training

Watch Out for Pitfalls

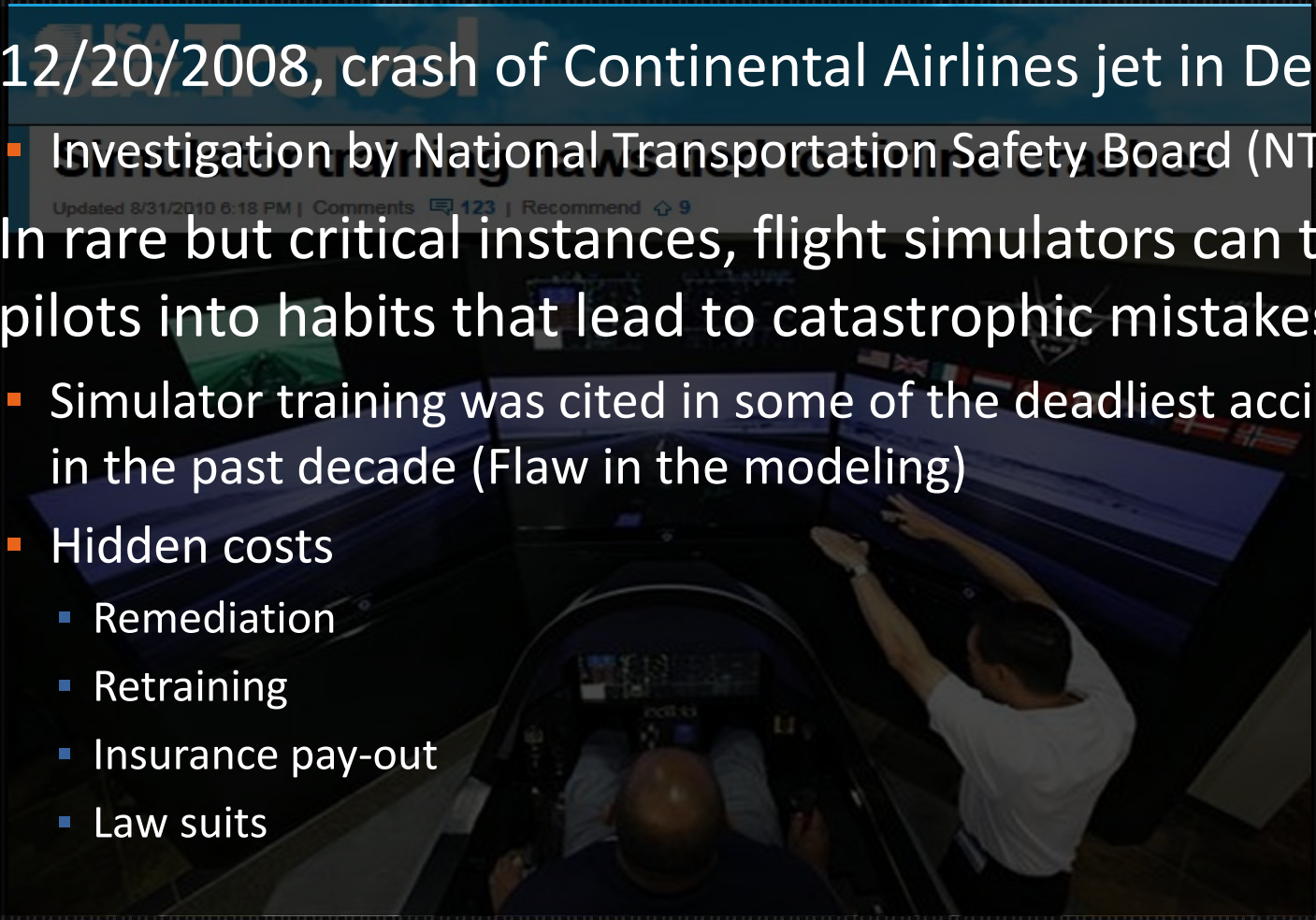
- Development costs
  - Game engine is exorbitantly expensive ($ 350-700 k)
    - E.g., Military model (American Army)
    - Development team expensive… time ($) consuming
  - L.U.B.I.F.Y.
    - E.g., Grant funded (Project Serene – NSF funded)
    - Locked-In to one technology
    - Maintenance Fees (monthly subscriptions)
    - Lesson learned from SecondLife (vs. OpenSimulator)
  - Free GDK
    - E.g., Low-cost approach, Education Arcade, NWN
    - Game *mod*ification (modding)
    - Scenario mismatch (e.g., medieval setting)

Validation,...

Verification, …

- Verification: Is the thing built right?
  - Assumption: A ‘readied’ game is ready to train
    - Especially after spending $$$
  - Creation process – faulty model/scenario (G.I.G.O.)
  - Flight Simulators
    - Used by aviation companies to train pilot
    - Since 1970, over 40-year ‘success’ story
    - Entirely on-the-ground training

- 12/20/2008, crash of Continental Airlines jet in Denver
  - Investigation by National Transportation Safety Board (NTSB)
- In rare but critical instances, flight simulators can trick pilots into habits that lead to catastrophic mistakes
  - Simulator training was cited in some of the deadliest accidents in the past decade (Flaw in the modeling)
  - Hidden costs
    - Remediation
    - Retraining
    - Insurance pay-out
    - Law suits



Accreditation (Testing)

- How do you ‘test’ that the 3D MUVE built fulfill its intended purpose (i.e. training)?
  - Findings will affect adoption
  - A lot of people are asking for this data
- Repetition is important for training muscle memory
  - Games still ‘train’ through repetition
    - learning scenario/tasks
  - Undetected mistakes become entrenched
  - More time and money needed to un-train/un-learn and then to re-train/re-learn
  - Just-in-time assessment is needed to ‘catch’ mistakes made

Accreditation (Testing)…

- No built-in Evaluation & Performance Assessment (EPA) process at the moment
  - 'External' data collection process – human errors introduced during data entry
  - *Post hoc* EPA (After Action Report) – undetected mistakes became entrenched (prolonged training)
- Game logs
  - Most commonly available, cheap and easy, not 'standardized'
  - Available at "End of game" (*post hoc* analysis)
    - Number of kills, time taken, etc
  - In plain text, XML

Accreditation (Testing)…

- Video Analysis
  - Qualitative Approach like (usability testing)
    - Very details (too much details?)
    - Time consuming ($$)
    - 2-3 hours of transcription time per hour of game play
    - Imagine COTS/GOTS game that req. 20-40 hours
  - Instant replay
    - Commonly found in sports games
    - More entertainment than testing
    - Self-gratifying YouTube moments

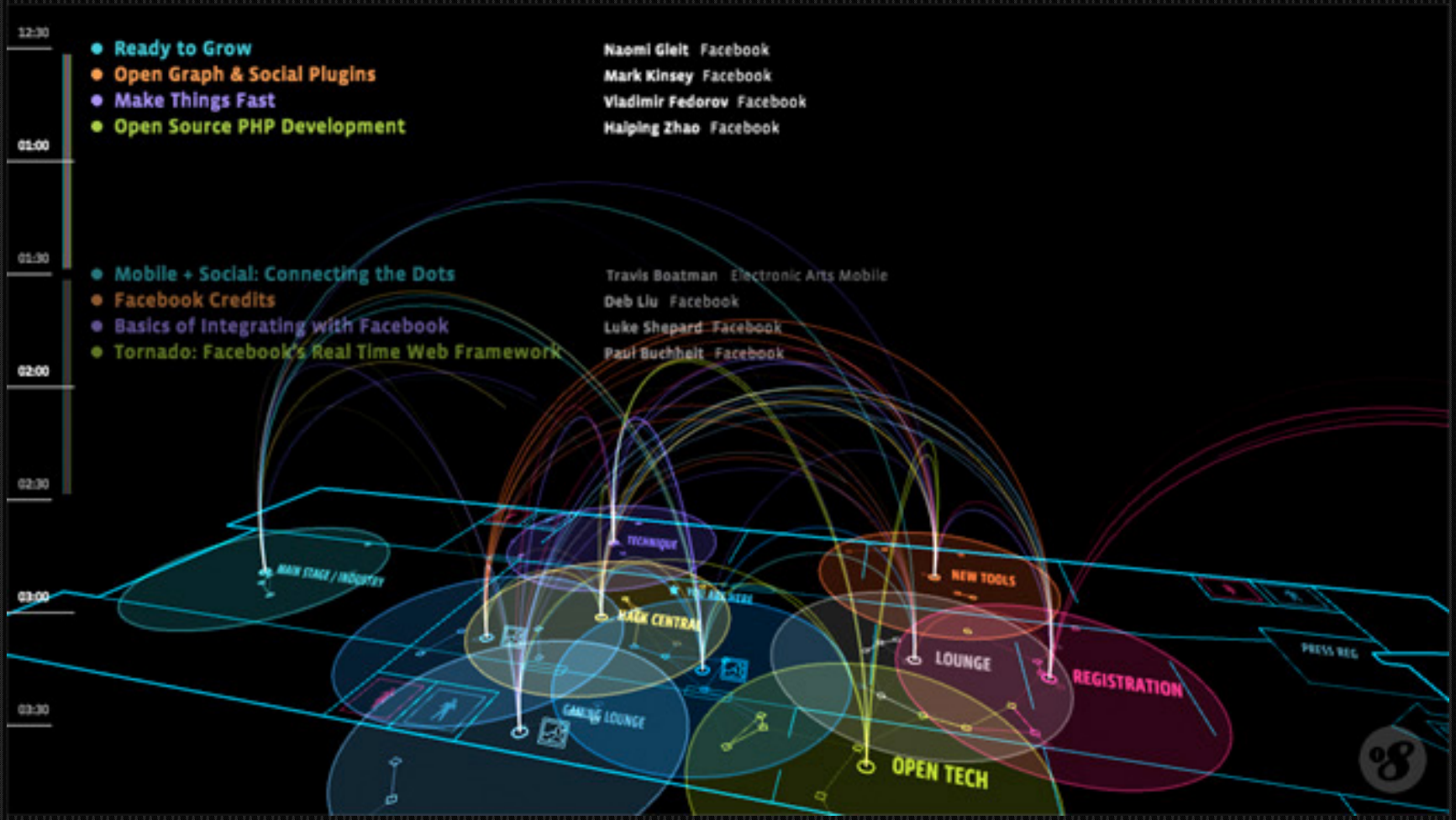
Accreditation (Testing)…

- Pre-test/Post-test
  - Quantitative
- Compare achievement scores of a ‘game’ class against a ‘traditional’ class
  - When student group 1 used a (named game/simulation), Testers noticed a significant improvement over those students (group 2) who didn't use the simulation.
  - Media Comparison studies
    - BAD design, yield no significant difference result (R.E. Clark)
  - Move away from comparing different product/ media
    - Compare product design

Why 3D MUVE

- Advantages:
  - (Photo)Realism – new graphic engines
  - Spatial placement – 3D placement (depth)
  - Co-location – different locations (trainer/trainees)
  - Mitigate costs – save on transporting resources
  - = Interactive Learning (e-/m-Learning)
    - Anytime Anywhere
  - Game Changing
  - Traceable learning objectives

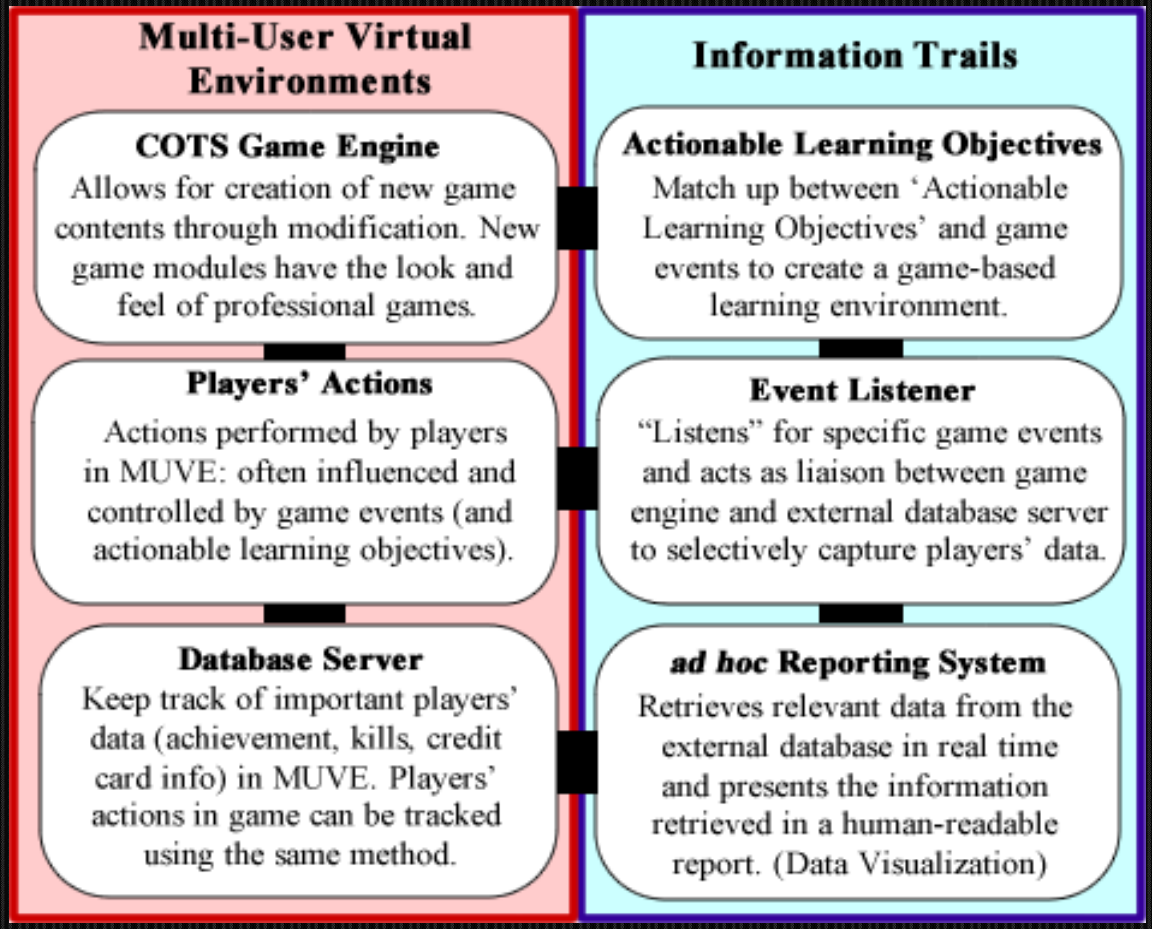
Traceable Objectives



Assessment Component for MUVE

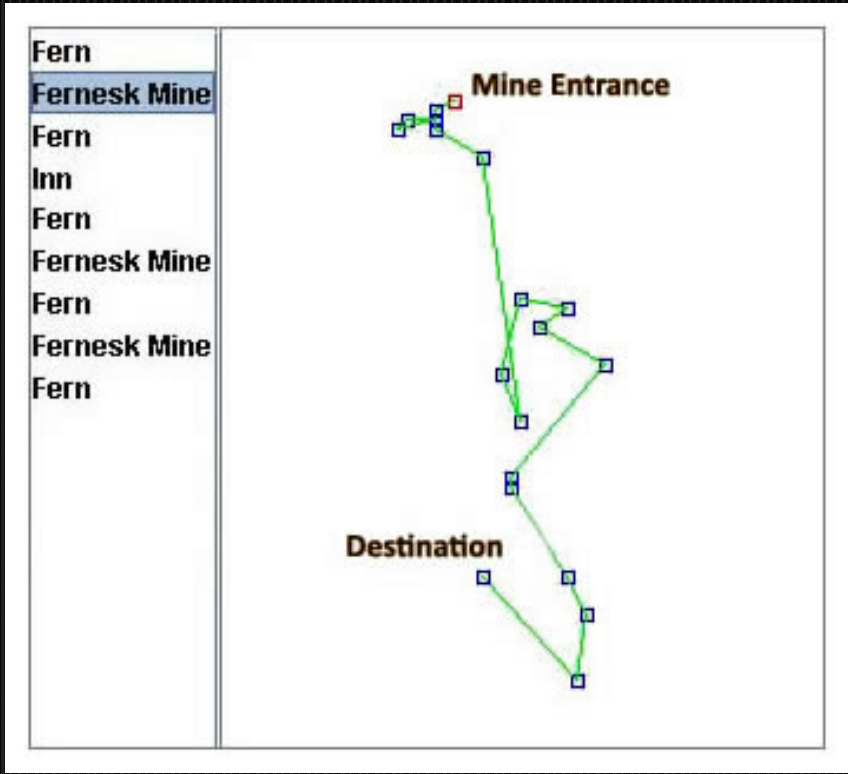
- *Assessment is the key difference between entertainment games and serious games* – Michael & Chen (2004)
  - Performance assessment is important for serious games (and 3D MUVE) for training
  - We need an on-demand (*ad hoc*) assessment system
  - *Post hoc* After Action Reporting can incur more costs!
- Making good MUVE for training:
  - Don't do 'multiple choice' selection (= edutainment)
  - Action should have more 'open choices', not (feels like) scripted action, or preprogrammed actions

Information Trails (Loh, 2007)

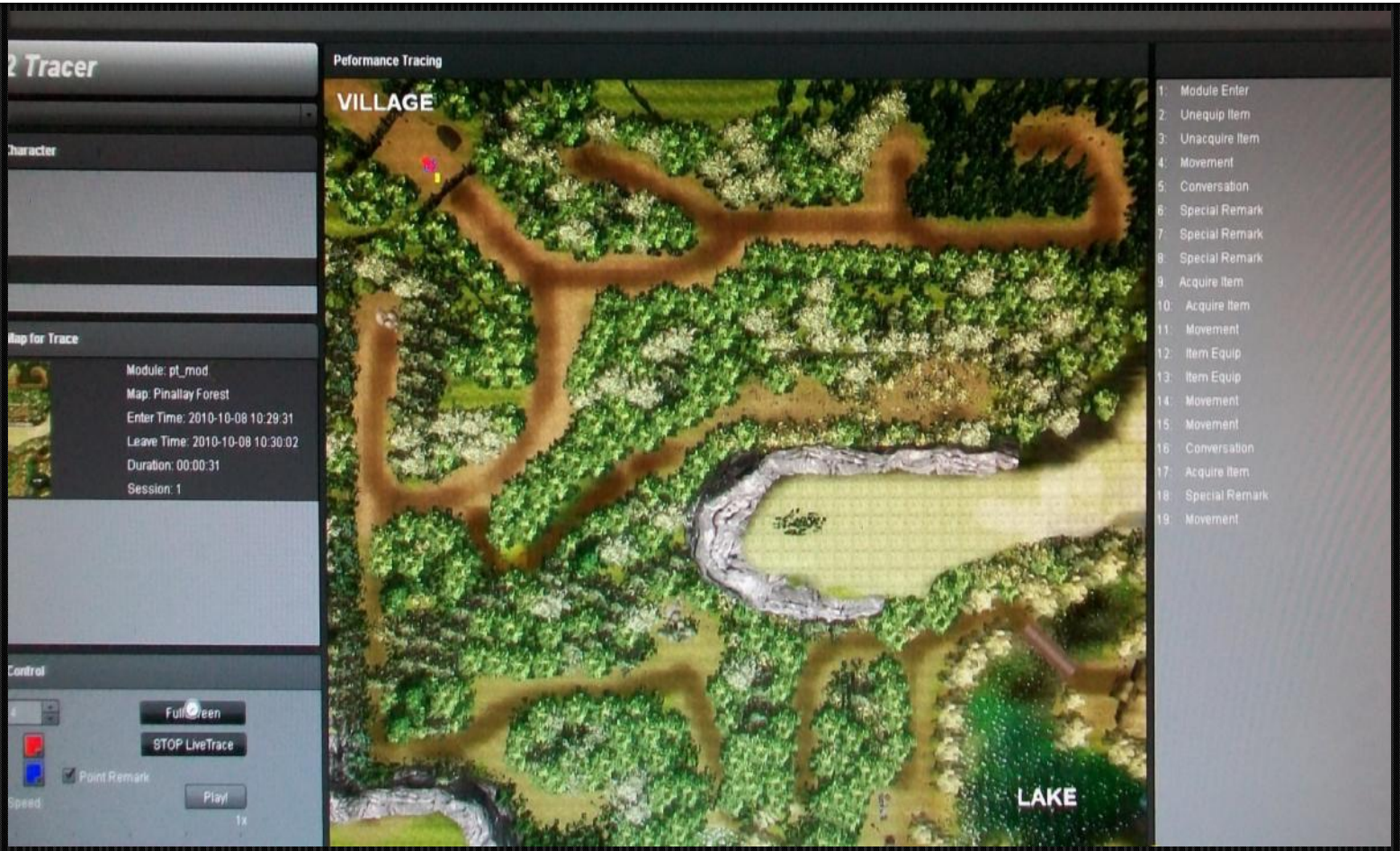


- Dynamic Data-Driven Assessment
- *Renamed* **Performance Trails** (2010)

From Concept…

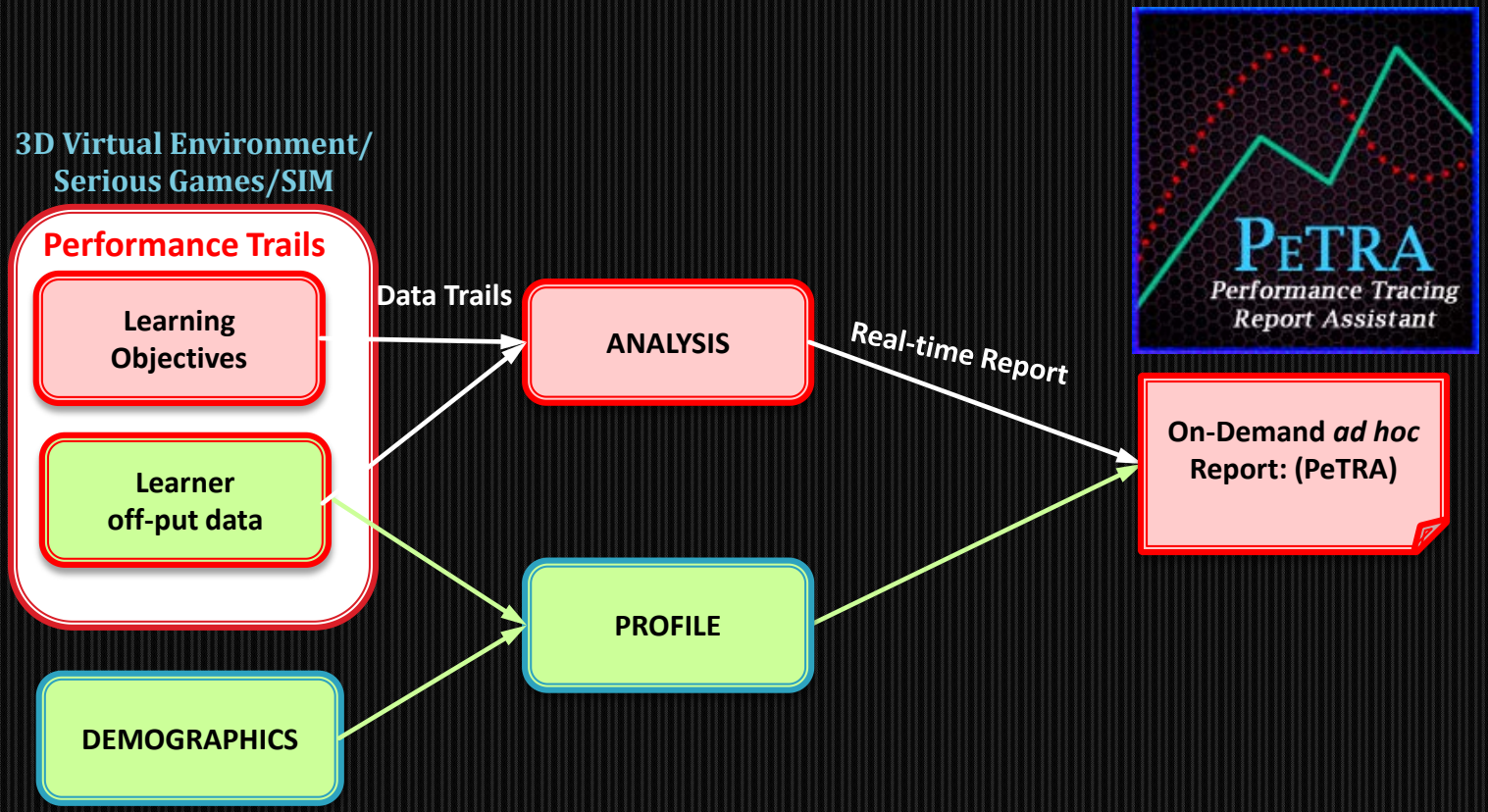


…to Finished Product



- *Performance Tracing Report Assistant* (PeTRA) displays learner's performance data in a human readable report.

Performance Trails + Petra (2010)



- Intelligent Data (Curriculum, learning objectives, etc.)
- Learner off-put data (actions, behaviors, game decisions)

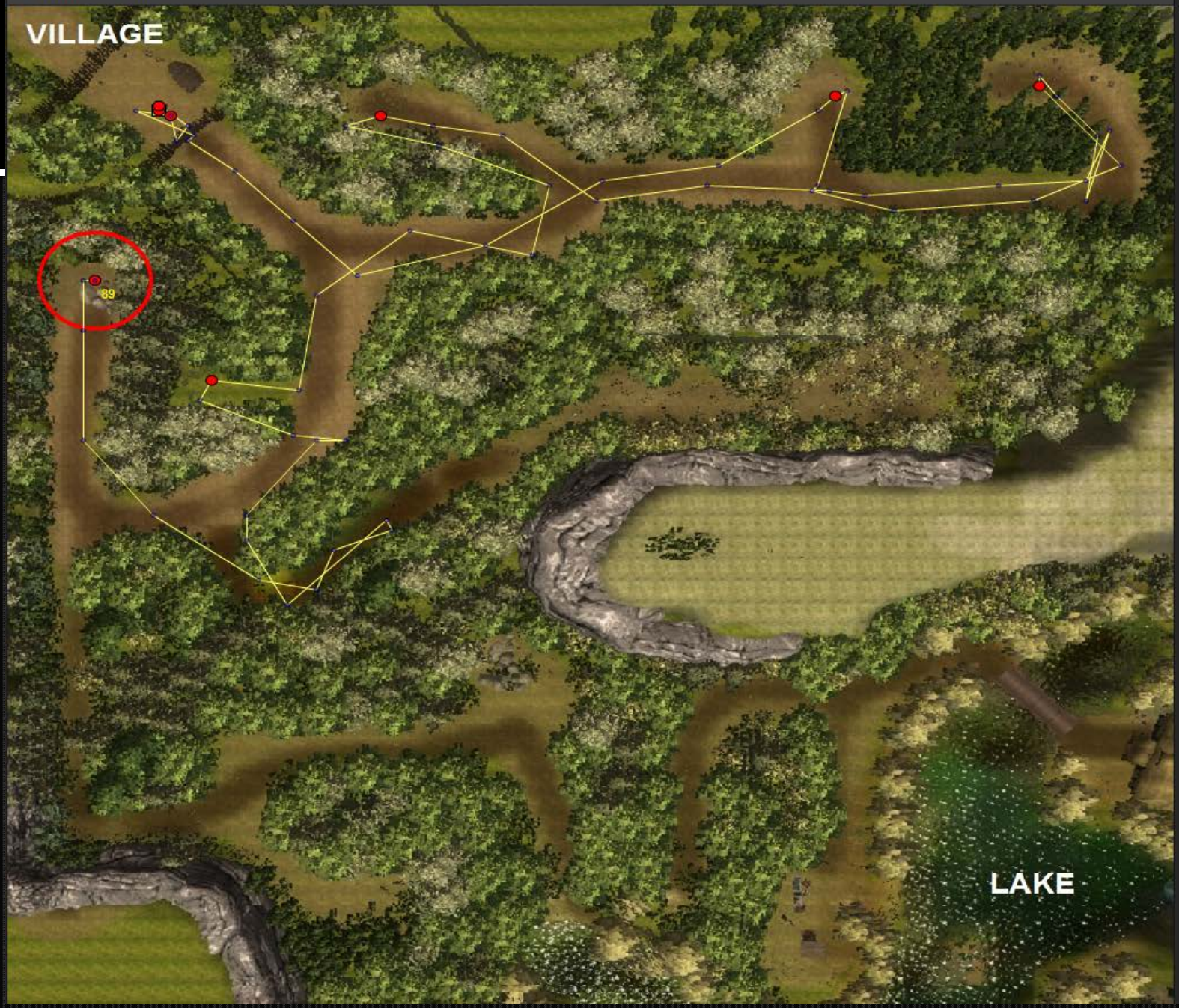
3D MUVE Training

- Player off-put data into Performance Trails enabled 3D MUVE system

3D MUVE & PETRA (side by side)

- On-demand *ad hoc* (PeTRA) Report allows trainer/administrator to monitor trainee's in real-time (to provide intervention/remediation).

Peformance Tracing



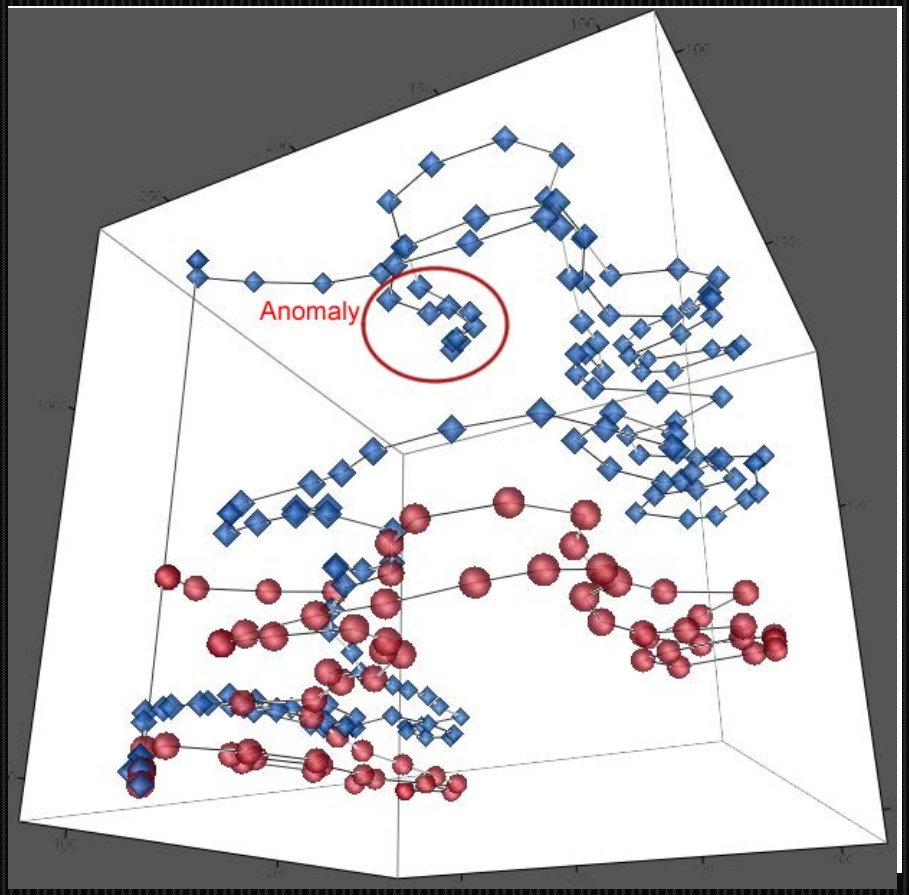
Action List

28: Movement
29: Movement
30: Conversation
31: Special Remark
32: Movement
33: Conversation
34: Conversation
35: Movement
36: Movement
37: Movement
38: Movement
39: Movement
40: Movement
41: Movement
42: Conversation
43: Special Remark
44: Movement
45: Movement
46: Movement
47: Movement
48: Movement
49: Movement
50: Movement
51: Movement
52: Movement
53: Movement
54: Movement
55: Movement
56: Movement
57: Conversation
58: Special Remark
59: Movement
60: Movement
61: Movement
62: Movement
63: Movement
64: Movement
65: Movement
66: Conversation
67: Special Remark
68: Movement

Novice Data

Expert Data

Novices vs. Experts (Profile Analysis)



Anomalies are extraneous actions not found in expert profiles. Possibly produced during man-made “mistake” – worth investigating

Question/Comments?

Contact:

Christian 'Sebastian' Loh, Ph.D.
Director, *Virtual Environment Lab (V-Lab)*
Assoc. Prof., *Instructional Design & Technology*
Curriculum & Instruction
Southern Illinois University Carbondale
csloh@siu.edu
618. 453.4206