Virtual Environment Lab (V-LAB)
Southern Illinois University

Information Trails

In Progress Assessment of Game-Based Learning

C. Sebastian Loh, Ph.D.
csloh@siu.edu

Horizon Report

- Games-based learning (GBL) — the next new emerging technology to impact Higher Education

- Targetted as learning platform for all kind of learning/ training/ online interaction (military, corporate, higher education).

LEARNING & DOING

- **Brown, Collins and Duguid (1989)**
  - **Different instructional goals of "knowing what" and "know how" → different structures and practices of our education system**
- **DECONTEXTUALIZED: Learning is separated from Doing**
  - **"Activity and situations are integral to cognition and learning"**
  - **Cognitive apprenticeship can provide "the authentic practice through activity and social interaction in a way similar to that evident – and evidently successful – in craft apprenticeship."**
- **Game-Based Learning is a great example of Learning by Doing!**
  - **Authentic "ACTIVITIES" and situations – story-based, social interactions with (non-)player characters**

Situated Cognition

- Knowledge, as a product of a meaning-making (decision-making) process, cannot be separated from the context of its use.

- Learning is a continuous, lifelong process from acting in situations.

- Tools (Game-based learning) and their use "reflect the particular accumulated insights of communities."

- "Given the chance to observe and practice *in situ* the behavior of members of a culture (or environment), people pick up relevant jargon, imitate behavior, and gradually start to act in accordance with its norms."

COGNITIVE APPRENTICESHIP

○ Four Cognitive Apprenticeship Procedures:

1. By beginning with a task embedded in a familiar activity it shows the students the legitimacy of their implicit knowledge and its availability as scaffolding in apparently unfamiliar tasks.

2. By pointing to different decompositions it stresses that heuristics are not absolute but assessed with respect to a particular task – and that even algorithms can he assessed in this way.

3. By allowing students to generate their own solution paths, it helps make them conscious creative members of the culture of problem-solving mathematicians.

4. It helps to enculturate students through this activity, acquire some of the culture's tools – a shared vocabulary and the means to discuss, reflect upon, evaluate and validate community procedures in a collaborative process.
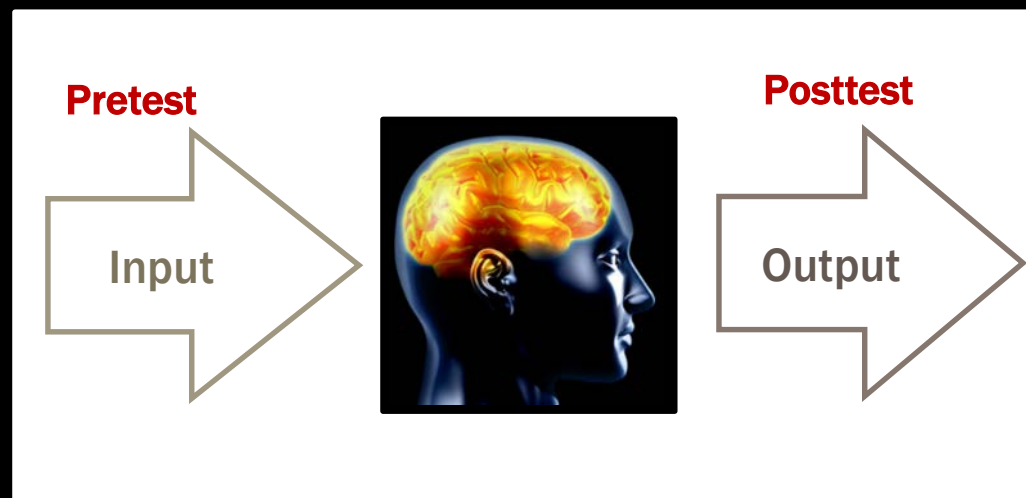
Learning & Assessment

BLACK OR WHITE BOX

- **Inclusion of an <u>assessment</u> components is the main difference distinguishing the more 'serious' games from the ones made for entertainment**
  - **Chen and Michael (2005)**
- **Black Box vs. White Box**
  - Pretest > Intervention > Posttest (Best method?)

DEMONSTRABLE ABILITIES

○ **Learners should be able to demonstrate the 'abilities' they have acquired from the course of instruction (Joosten-ten Brinke, Gorrisen, Latour, 2005)**

• **San the demonstration of the abilities learned, there are no means of knowing if the learners have learned anything**

• **Apart from testing (pretest/posttest), what else can you assess?**

  • Israeli Air Force Study (Tobias, S.)
    – Students who played *Space Fortress* had better rankings in their pilot training than students who did not.

  • Performance ranking...

PNTS	CNTRL	VLCTY	VLNER	IFF	INTRVL	SPEED	SHOTS
200	100	119	0	W		90	70

Case Study

- Serious game improves performance!!
  - U.K. Royal Navy's Maritime Warfare School (MWS)
  - Military recruits who failed to meet required BUTS standard must receive additional training
  - This demands a considerable investment in additional training resources, cost and time off the job.
- Study Design:
  - 10 classes played a new serious game aimed to teach the same skills
  - The results was compared to a control group (taught by traditional method: instructor using book & powerpoint)
  - Pre-existing course curriculum was severely ineffectual.

NEWS FLASH!

- **First Study to prove effectiveness of Serious Games!**
  - BUTS measurement was taken over 10 classes, throughout the term (to determine the effectiveness of the serious game lesson in comparison to the traditional classroom teaching method).
  - One more Media Comparison Studies
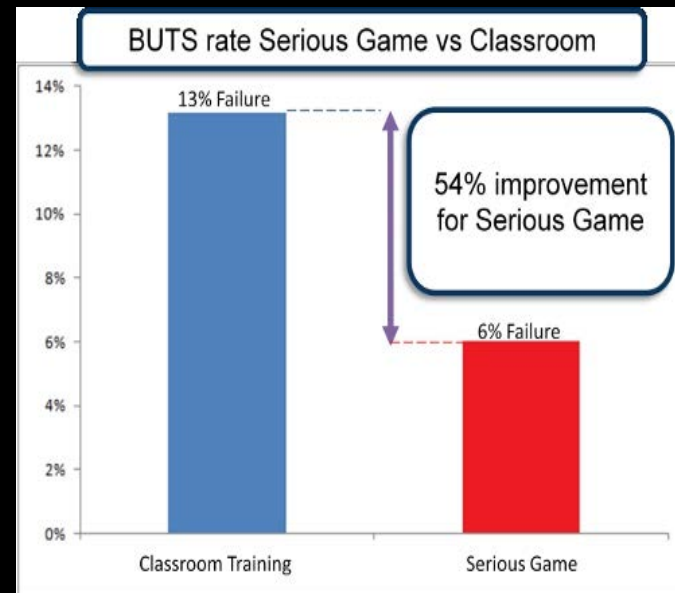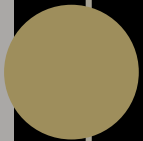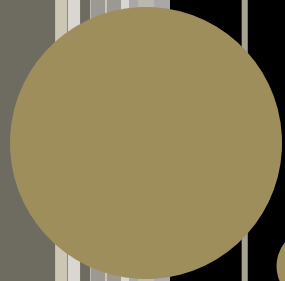
RECONSIDERING ASSESSMENTS

Assessment in Physical Environments

○ **Physical classrooms (physical = face-to-face)**

○ **"Facial expressions and tell-tale physical/emotional behaviors" can be directly interpreted as evidence of learning and participation"**

- Harrington, Meisels, McMahon, Dichtelmiller, & Jablon, 1997
- Body language, eye contact, facial expression, tone of voice, gesture signaling (dis)approval

# Shift from Physical to Virtual Environmt.

Assessment in Virtual Environments

- Virtual, non-physical classroom:
  Facial expressions, body language (Gone!)
- Emote? (Second Life)
  - Will student fake their emote?
- How else can teachers assess learning?
  - Assessment = Mental Dipstick

"IT'LL REVOLUTIONIZE I. Q. TESTING! — I'VE INVENTED THE MENTAL DIPSTICK!"

© Original Artist

Open Mouth   Repulsed   Sad   Shrug
Smile   Suprise   Tongue Out
Toothsmile   Wink   Worry
Afraid   Anger   Bored
Cry   Disdain   Embarrassed
Frown   Kiss   Laugh

DO YOU KNOW WHERE IS YOUR "TRUCK"?

○ **Media is 'a grocer truck' (mere vehicle)**

○ **It *just drives* the learners to the destination....**

○ **Even though they arrived.... Did they take a detour, or stop somewhere where they are not supposed to?**

○ **We did not have the technology then, but we do now.**

Real time tracking
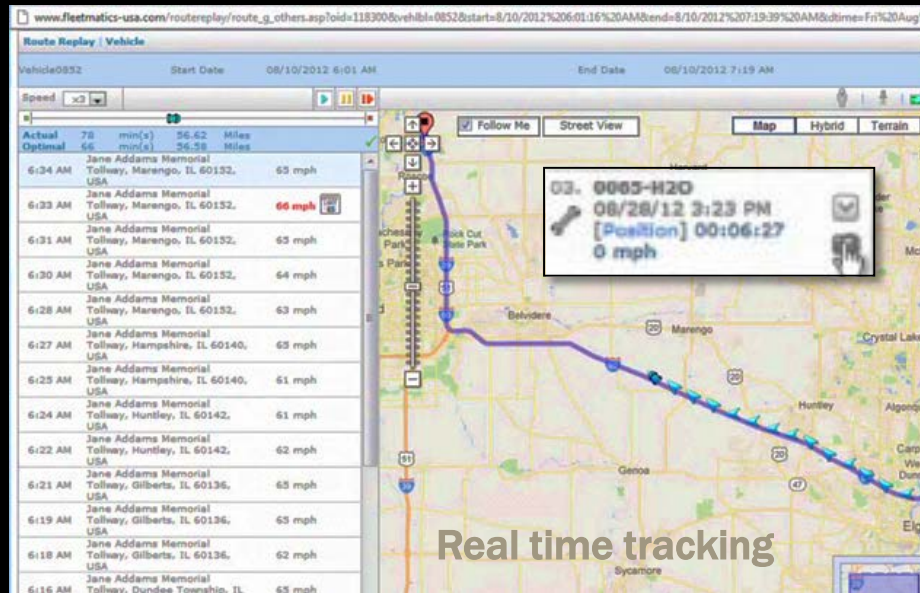
What can be learned?
Driver 1 on schedule, lunch break: 2 eggs, 24 chips, 7 mushrooms and fried bread
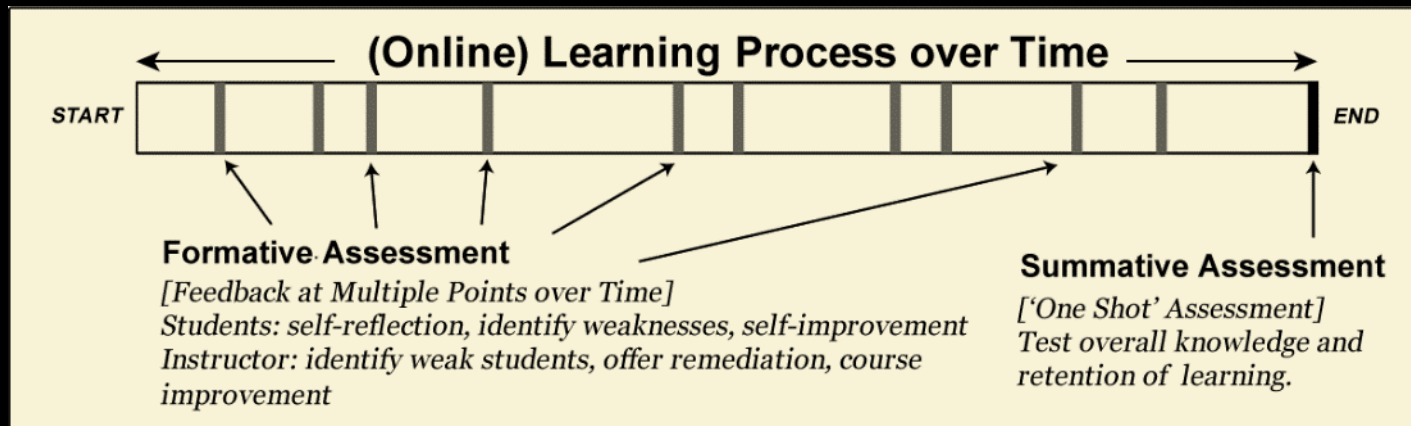Driver 3 helped friend move house

FORMATIVE & SUMMATIVE

- **Formative (along the way)**
  - Quizzes, Assignment throughout a semester
  - *Take small 'tests' as they are doing GBL*
- **Summative (at the end of learning period)**
  - Finals at the end of a semester
  - *Big 'Final' after the GBL period is over (can be 20-40 hours later) – undesirable learning habits*
- Both?

(Online) Learning Process over Time

START   END

Formative Assessment
[Feedback at Multiple Points over Time]
Students: self-reflection, identify weaknesses, self-improvement
Instructor: identify weak students, offer remediation, course improvement

Summative Assessment
['One Shot' Assessment]
Test overall knowledge and retention of learning.

DATA COLLECTION METHODS

ASSESSMENT METHODS

- There are other data collection methods for software assessment

- Friday, 2:15-3:15pm (East- Suite Tower / Wilson)

- Most educational research collect data and measuring performance (before and) after game-play (in a different environment)
  - Pen-and-paper Test → Online test
  - Dependent on short-term memory
  - Interview, Talk-aloud, Self-reports: can be unreliable
  - Game log analysis (D.I.Y. approach, not for the uninitiated)

Game Log Analysis... anyone?

- Item answered, % correct, final score...
- We need better metrics!

Game Data by Mission

SITE NAME	STUDENT NAME	USERNAME	DATE LOGGED IN	TIME ON TASK (MM:SS)	MISSION	LEVEL	TOPIC	SKILL	ITEMS ANSWERED	% CORRECT	FINAL SCORE
Washington MS	Armstrong, Louis	LArmstrong	8/25/2009 02:12 PM	05:05	Tower Storm	G - Middle	Measurement	Solve Proportions	2	0	0
Washington MS	Armstrong, Louis	LArmstrong	8/25/2009 02:18 PM	05:05	Tower Storm	G - Middle	Geometry	Classify Figures	5	60	500
Washington MS	Ellington, Edward	EEllington	8/25/2009 04:27 PM	05:05	Meltdown III	Algebra	Number and Operations	Simplify Radical Terms	3	67	150
Washington MS	Fitzgerald, Ella	EFitzgerald	8/19/2009 04:15 PM	05:04	Meltdown III	G - Middle	Number and Operations	Fractions/Decimals/Per	1	100	50
Washington MS	Fitzgerald, Ella	EFitzgerald	8/19/2009 04:44 PM	15:05	Tower Storm	G - Middle	Number and Operations	Random Mix	13	85	1500
Washington MS	Fitzgerald, Ella	EFitzgerald	8/20/2009 03:57 PM	05:05	Meltdown III	G - Middle	Number and Operations	Random Mix	3	100	450
Washington MS	Fitzgerald, Ella	EFitzgerald	8/20/2009 04:14 PM	16:52	Obstacle Course III	G - Middle	Measurement	Random Mix	16	88	1000
Washington MS	Fitzgerald, Ella	EFitzgerald	8/19/2009 04:39 PM	09:28	Tower Storm	G - Middle	Number and Operations	Compare Numbers	7	71	500
Washington MS	Fitzgerald, Ella	EFitzgerald	8/19/2009 04:15 PM	05:04	Meltdown III	G - Middle	Number and Operations	Fractions/Decimals/Per	2	100	300
Washington MS	Monk, Thelonious	TMonk	8/25/2009 03:19 PM	05:05	Swarm III	G - Middle	Algebra	Translate to Alg.	1	100	100
Washington MS	Monk, Thelonious	TMonk	8/22/2009 03:41 PM	05:04	Meltdown III	G - Middle	Number and Operations	Properties of Addition	1	100	100
Washington MS	Monk, Thelonious	TMonk	8/19/2009 04:42 PM	12:13	Tower Storm	G - Middle	Number and Operations	Random Mix	6	83	500
Washington MS	Monk, Thelonious	TMonk	8/25/2009 03:12 PM	05:05	Meltdown III	G - Middle	Algebra	Random Mix	5	60	200
Washington MS	Monk, Thelonious	TMonk	8/25/2009 03:31 PM	05:04	Swarm III	D	Number and Operations	Random Mix	5	0	0
Washington MS	Monk, Thelonious	TMonk	8/25/2009 03:37 PM	03:03	Meltdown III	D	Number and Operations	Random Mix	5	40	150
Washington MS	Vaughn, Sarah	SVaughn	8/25/2009 02:20 PM	05:05	Meltdown III	G - Middle	Measurement	Equiv. Metric Units	2	0	0
Washington MS	Vaughn, Sarah	SVaughn	8/25/2009 03:12 PM	05:05	Meltdown III	G - Middle	Number and Operations	Estimate to Check	1	100	200
Washington MS	Vaughn, Sarah	SVaughn	8/25/2009 02:14 PM	02:12	Obstacle Course III	G - Middle	Measurement	Find Area/Volume	1	100	0
Washington MS	Vaughn, Sarah	SVaughn	8/25/2009 02:58 PM	04:25	Meltdown III	G - Middle	Number and Operations	Properties of Addition	1	100	200

IN SITU (IN PROCESS)

- How about collecting data during the gameplay?
- *In situ* data collection (can be qualitative or quantitative)
  - No-tech - Face-to-face classroom
- High tech, "In the field" data collection
  - During the Game-Based Learning process

- Game-based learning as a White-Box
  - Can see the variables, tinker with them, change variables to modify outputs
  - Understanding of programming process
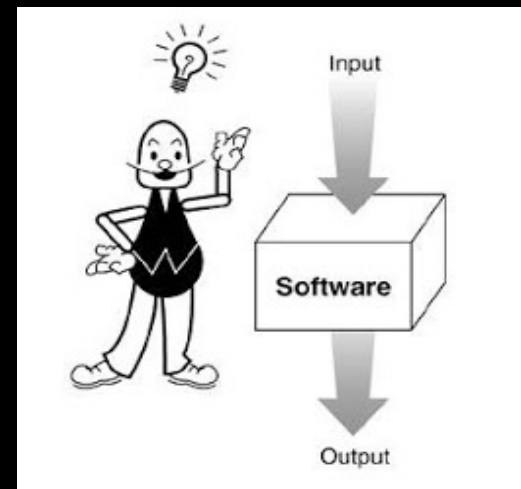
FOUR COMPONENTS

1: User-Generated Data

- **In-Process Data:**
  - Tracking device placed/attached on participants
  - As play-users play games, they generate data on-the-fly (while the game is still on-going)
- **GBL primarily 'train' through repetition**
  - Repetition is important for training muscle memory
  - Matching of learning/training goals to outcomes (Eye-hand coordination, Choosing the right answer?)
  - Verify that people indeed learned what they set out to learn
- **Detection of anomaly: 'unexpected' user behaviors**
  - Gaming the system? Bug? Truly innovative solution?

2: TELEMETRY (INSTRUMENTATION)

- *Telemetry:* **Remotely collecting data, and transmitting the data back to a remote (central) server for analysis**
  - **Transmitter > Receiver**
  - **Tracking the moving agent (participant) in its habitat**
  - **Found in car race, air plane, bio-habitat, truck tracking…**

3/4: Data Visualization

- Visualization of data for easy understanding (don't have to know statistics or read log files)

Heatmap of player death statistics.

3/4: Reporting Tool

- **Display raw data in a user-friendly report *in real-time* From telemetry to reporting**
  - A report is need for stakeholders: students, trainers, administrator
- **Learning Analytics?**

INFORMATION TRAILS & PETRA

(Performance Tracing Report Assistant)

Information Trails® Systems
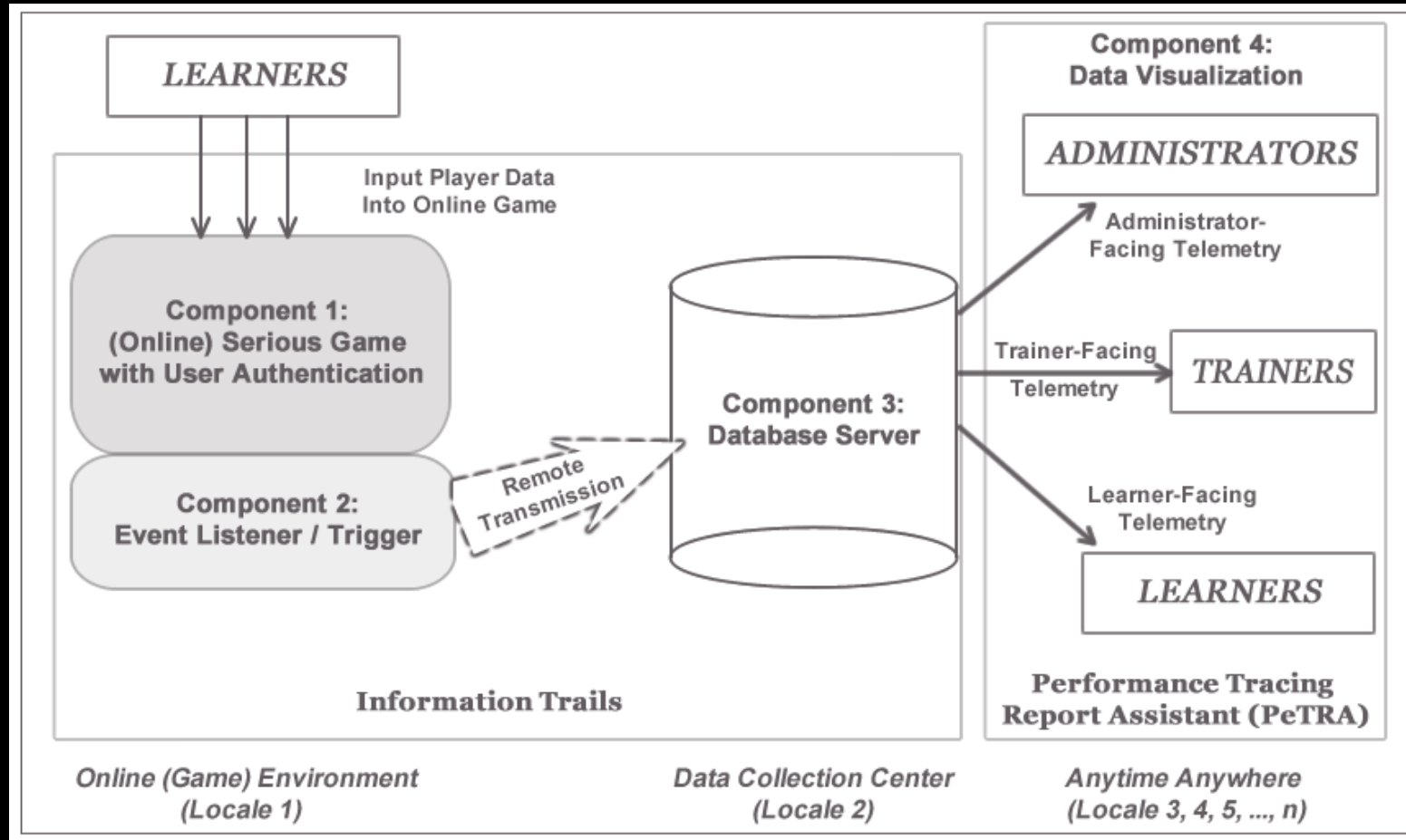
Formative Assessment

- During the game, user actions may be traced remotely or locally in real-time using a browser (desktops/mobile devices)
  - Trainers can monitor progress of trainees from a remote location
  - Play-users can use the report as self-evaluation
  - Administrators can check on performance of individual or group for ROI analysis
  - Anytime anywhere access

Summative Assessment

- After the game, we can compare play-users to his/her peers
  - Identification of 'top' performers
  - SD from expert performance level
- Compare play-users to his/her group
  - Identification of novice vs. expert
  - Predict category of play-user by how they react/do in certain GBL/training scenario

REAL-TIME PERFORMANCE TRACING

- Location 1
  - Game + Telemetry
  - Collection of user-generated data
  - *in situ* assessment

- Location 2
  - Performance Tracing
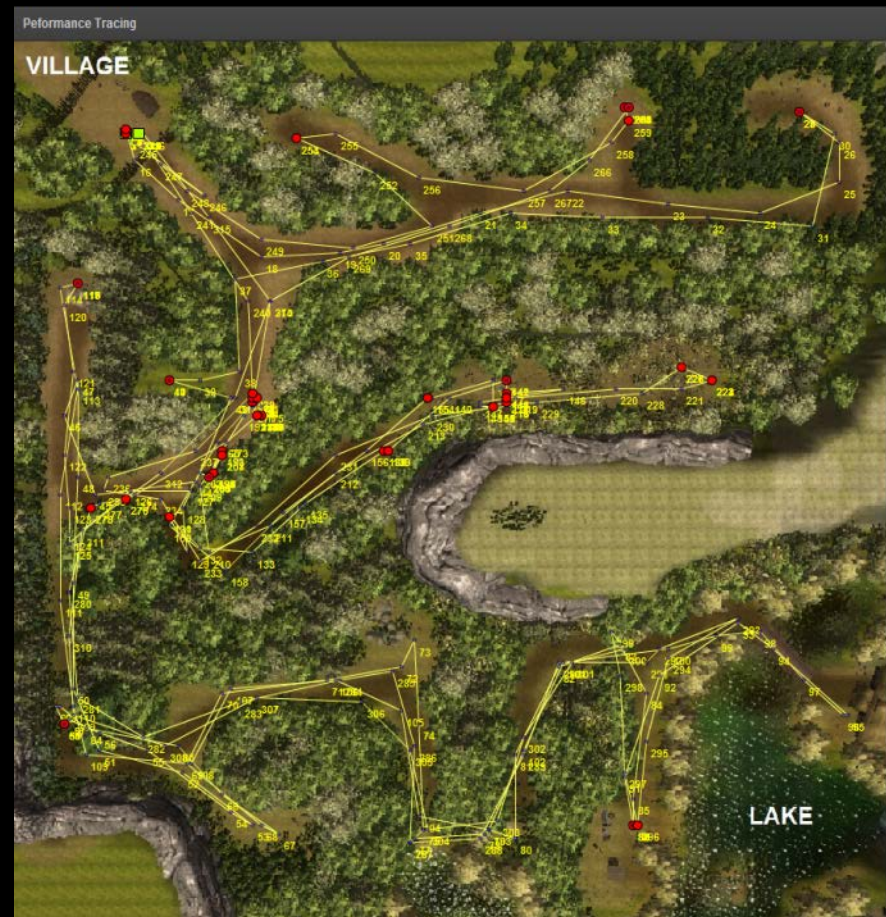 - Visualization of user-generated data
  - Real-time Report

Expert vs Novice Performance

Novice Behavior / Performance

- Confused
- Don't know what to do
- Back tracking
- A lot of wastage (time)
- Highly inefficient behaviors
- Aimless wandering

Expert Behavior / Performance

- Focused
- Know what to do
- Do what is required
- Very little wastage (time)
- Little/no back tracking
- Highly efficient behavior
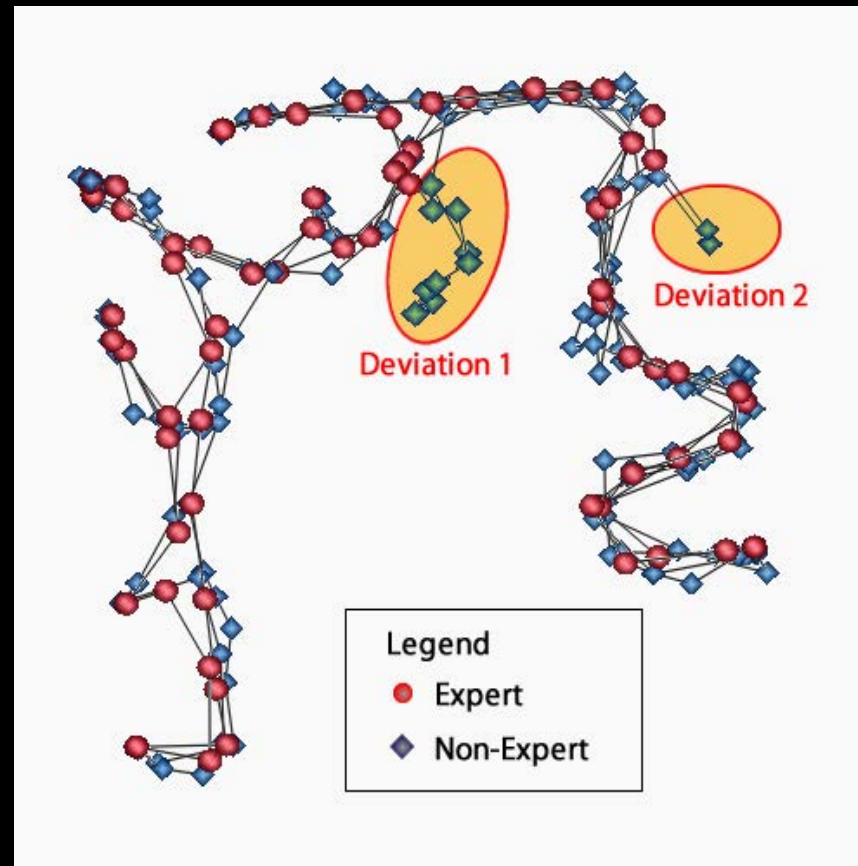- Purposeful actions

Novices vs. Experts (Profile Analysis)

- Expert vs. Novice performance 'graph'
- Performance efficiency
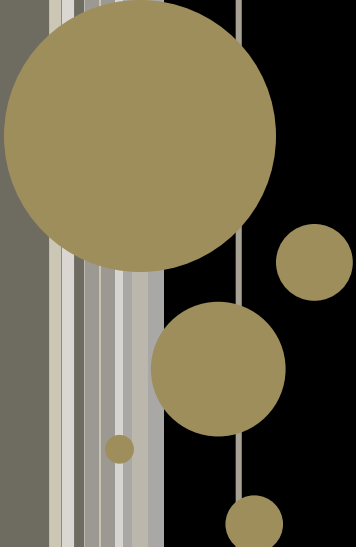- Fewer steps taken
- Shorter time of completion
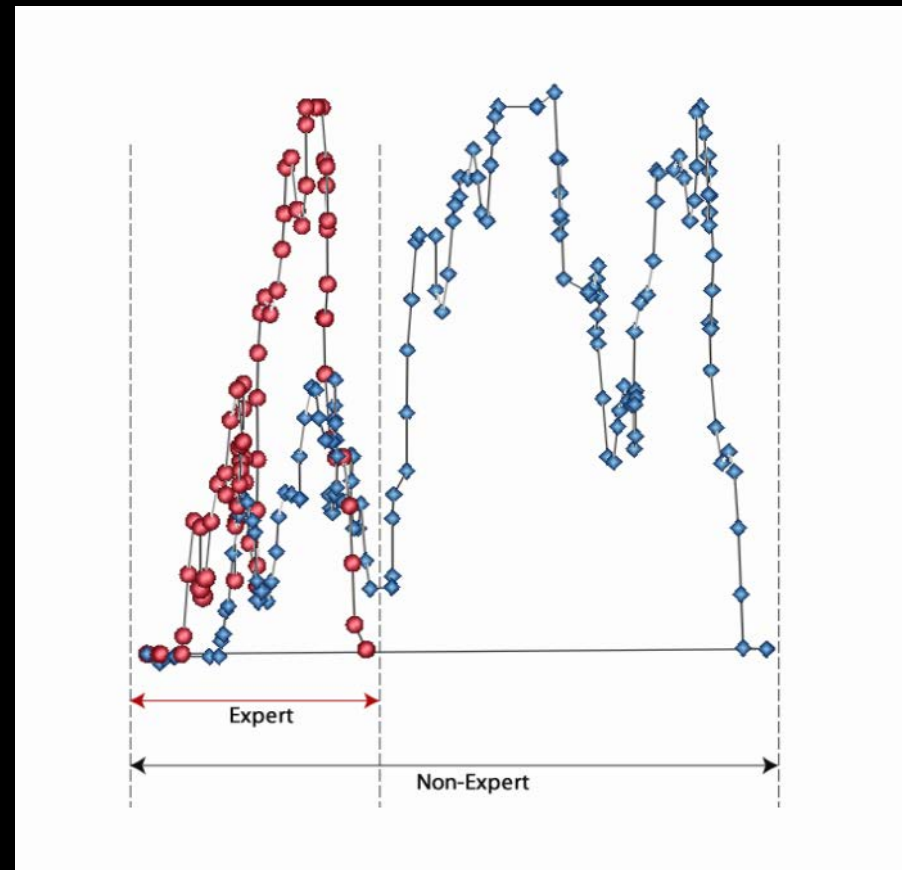
Novices vs. Experts (Profile Analysis)

- **Anomalies (deviations) are extraneous actions not seen in expert behaviors**

- **May constitute man-made "mistake" – worth investigating**

Virtual Environment Lab (V-LAB)
Southern Illinois University

C. SEBASTIAN LOH, PH.D.

DIRECTOR, VIRTUAL ENVIRONMENT LAB (V-LAB)

LEARNING SYSTEMS DESIGN & TECHNOLOGY

SOUTHERN ILLINOIS UNIVERSITY

CARBONDALE, IL

CSLOH@SIU.EDU

INFORMATIONTRAILS.COM, CSLOH.COM

LINKEDIN.COM/IN/CSLOH