Game Metrics & Biometrics

The Future of Player Experience Research

Nacke, L., Ambinder, M., Canossa, A., Mandryk, R., Stach, T. (2009). "Game Metrics and Biometrics: The Future of Player Experience Research" Panel at Future Play 2009.

Overview

1. Introduction of Panelists
2. Methods Overview
3. Discussion and Questions

Introduction of Panelists

1. Mike Ambinder
2. Alessandro Canossa
3. Regan Mandryk
4. Tad Stach
5. Lennart Nacke

Mike Ambinder

- User Experience Designer
  - Valve Corporation
- PhD in Experimental Psychology
- Application of knowledge and methodologies from psychology to game design

Alessandro Canossa

3 years experience at EIDOS

- Play pattern modelling techniques (Hitman Blood Money, Kane & Lynch and Tomb Raider Underworld)

Play-Persona framework

- Tool used in design phase to integrate different players' needs and motivations
- Backed by game metrics
- Tool used to evaluate experience

Speaking at

- Nordic Game (Sweden)
- NLGD (Holland)
- DGExpo (North Carolina, USA)
- Future Play (Canada)
- DIGRA (Japan)

Regan Mandryk

- User engagement in games
  - Sensing and modeling
- Interaction techniques
  - Emerging devices
- Assistant Professor
  - Computer Science
- University of Saskatchewan
  - Canada

www.reganmandryk.com

Tad Stach

- PhD student
- Computer science
  - Queen's University
- Exercise video games
- Heuristics and Usability

Lennart Nacke

- Blekinge Institute of Technology
  - PhD Candidate
    - Digital Game Development Degree
- EU FUGA ("Fun of Gaming") project
- Fun & player experience research
- Biometrics consulting

Methods Overview

1. Mike: *Direct Observation, Q&A, Verbal Reports, Surveys*
2. Alessandro: *In-game metrics, GIS/Heatmaps, Play-Personas*
3. Regan: *EMG, Skin Conductance, Heart-Rate*
4. Tad: *Heuristics, Usability evaluations*
5. Lennart: *EEG, Eye Tracking*

Direct Observation

- "Typical" playtest
  - Watch people play the game
  - Observe their gameplay/behavior
  - Simulate at-home experience
- Have a design goal
- Is it fun?

Direct Observation

PRO

+ Get a feel for player interaction with game
+ Importance of what people do—not what they say

CON

– Presence of observers can bias results
– Salient event can slant interpretation
– Behavior requires interpretation


Bill
50 480
15
Louis
Bill
Zoey
+100


QUARANTI
CONTAGIOUS DIS
NO ONE MAY ENTER OR LE
THIS BUILDING BY ORDER O
THE CIVIL EMERGENCY AND
DEFENSE AGENCY
TRESPASSERS WILL BE
PROSECUTED
CEDA

Verbal Reports

- Think-aloud protocol:
  - People describe their actions as they play
  - Unprompted and uncorrected
- In conjunction with direct observation

Verbal Reports

PRO	CON
+ Enables realtime glimpse into player thoughts, feelings, and motivations	– Interferes with gameplay
+ Bring up unnoticed details	– Creates an artificial experience
+ Effective for 'why' questions	– Inaccurate and biased


Louis
32 180
7
Don't shoot teammates!

Q&As

- Structured (usually) querying of playtesters
- Validate playtest goals
- Source of supplemental information

Q&As

PRO

+ Answer specific design questions
+ Determine specific player intent

CON

– Group biases (anchoring, social pressure, saliency, etc.)
– People don't know why they do what they do
– Potential for biased questions


10 78
1
Zoey
Don't shoot teammates!
Zoey
Friends
Louis
+04

Surveys

- Set of standardized questions
- Forced choice responses
- Quantify feedback/opinions
- Player categorization

Surveys

PRO	CON
+ Less biased responses	– Eliminate nuance
+ Response validation	– Difficulty in converting ratings to meaningful decisions
+ Forced choice helpful for revealing preference	– Limited solution space
+ Time-based comparisons	

How challenging were the following enemies (**1** = very easy; **7** = very hard)?

Boomer:	1	2	3	4	5	6	7
Common Infected:	1	2	3	4	5	6	7
Hunter:	1	2	3	4	5	6	7
Smoker:	1	2	3	4	5	6	7
Tank:	1	2	3	4	5	6	7
Witch:	1	2	3	4	5	6	7

Please rank order your preference for the following weapons from **1** (most liked) to **12** (least liked)

Assault Rifle _____
Auto Shotgun _____
Dual Pistols _____
Gas Can _____
Hunting Rifle _____
Molotov Cocktail _____
Mounted Turret _____
Pipe Bomb _____
Pistol _____
Propane Tank _____
Pump Shotgun _____
SMG _____

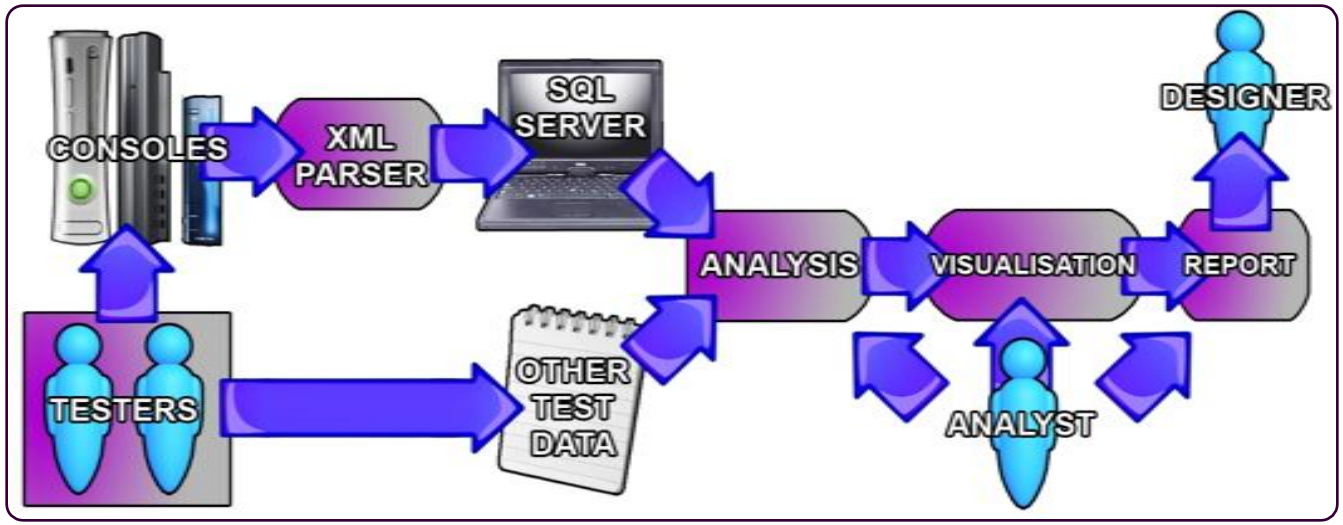
Gameplay Metrics

Gameplay metrics = Player behavior

- *Numerical data from game software about player behavior*

Types:

- Continuous / Frequency / Triggered
- Spatial / Non-spatial



Gameplay Metrics

PRO

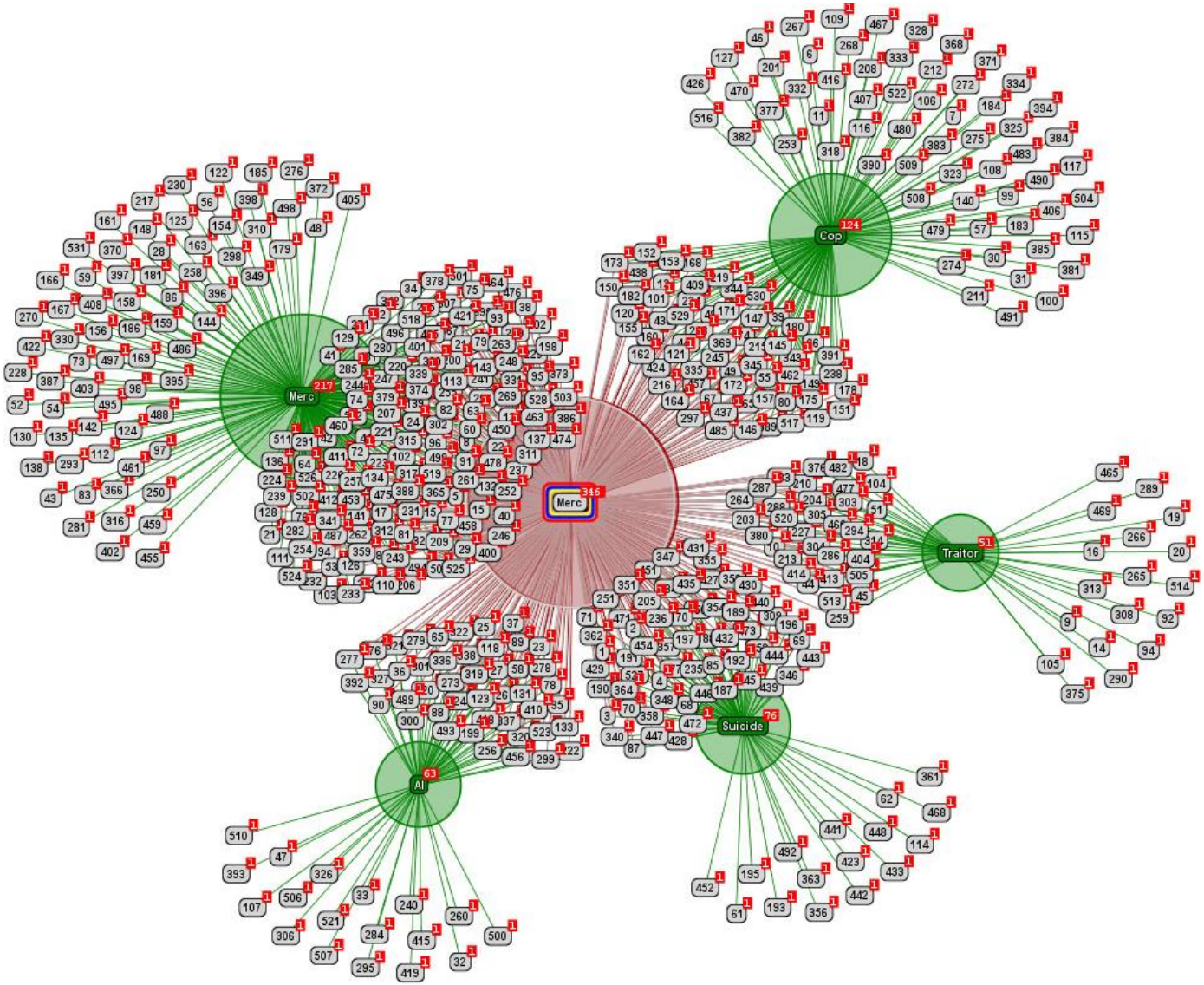
Answers to

- What?
- Where?
- When?

CON

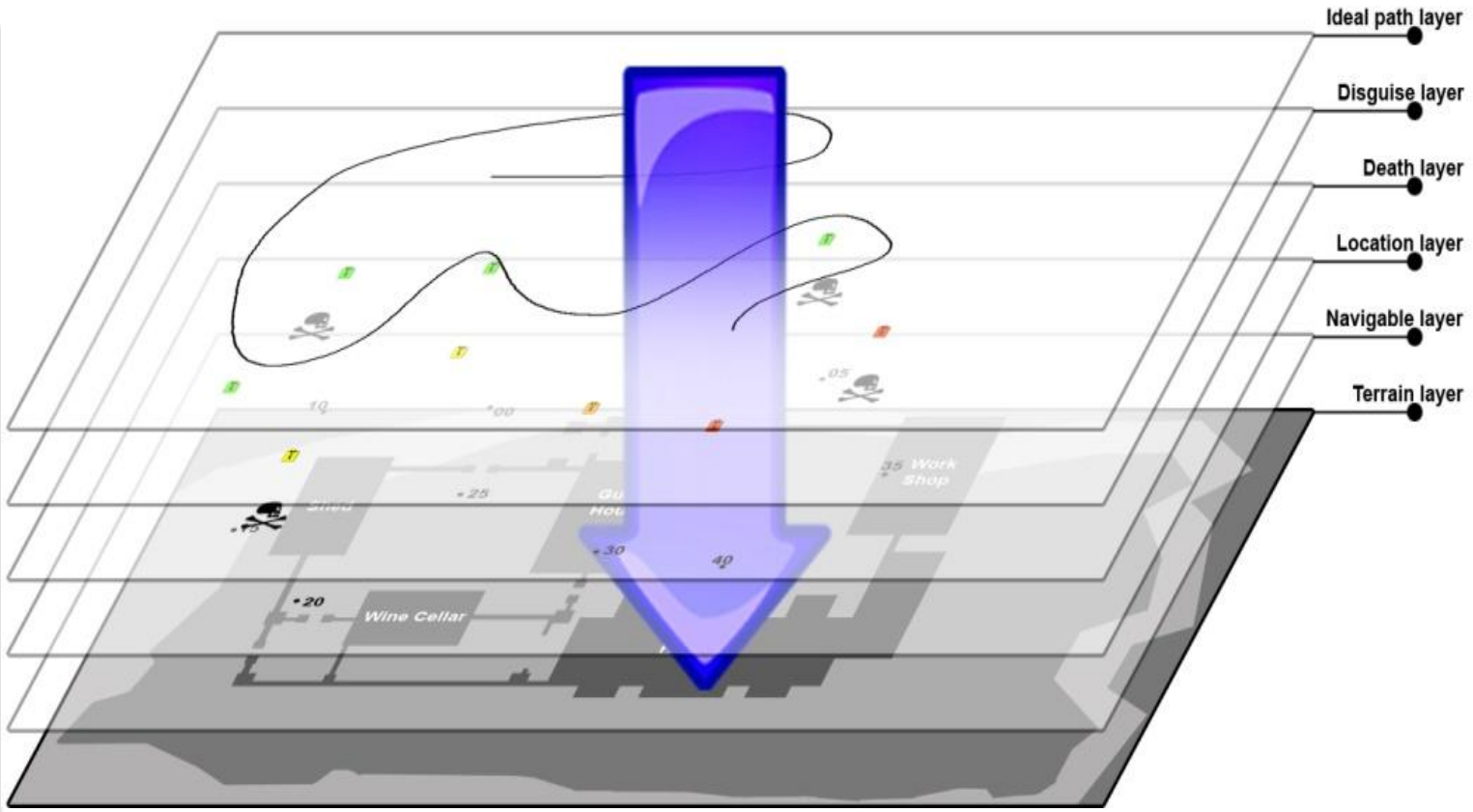
No answers to

- Why?
- How?



Flower of Death

Generated by a cluster visualization tool (shows data from *Fragile Alliance*, it relates ***role*** *at death* with ***cause*** *of death*)



Geographical Information System (GIS)

GIS are computerized data management systems used to capture, store, manage, retrieve, analyze, and display information with spatial dimension.

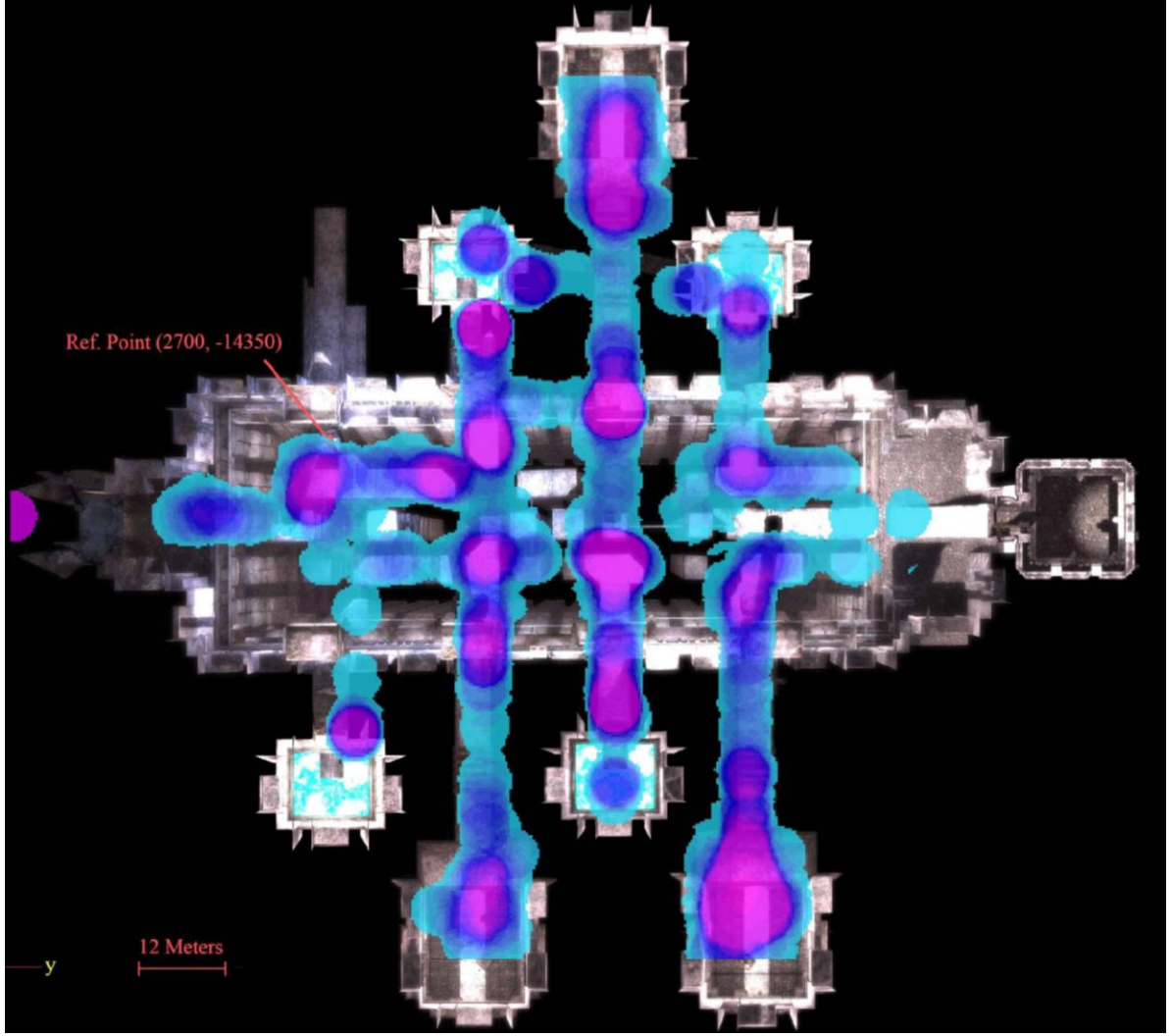
Geographical Information Systems (GIS)

PRO

- Flexible
- Off-the-shelf
- Cheaper
- Minimal customization needed

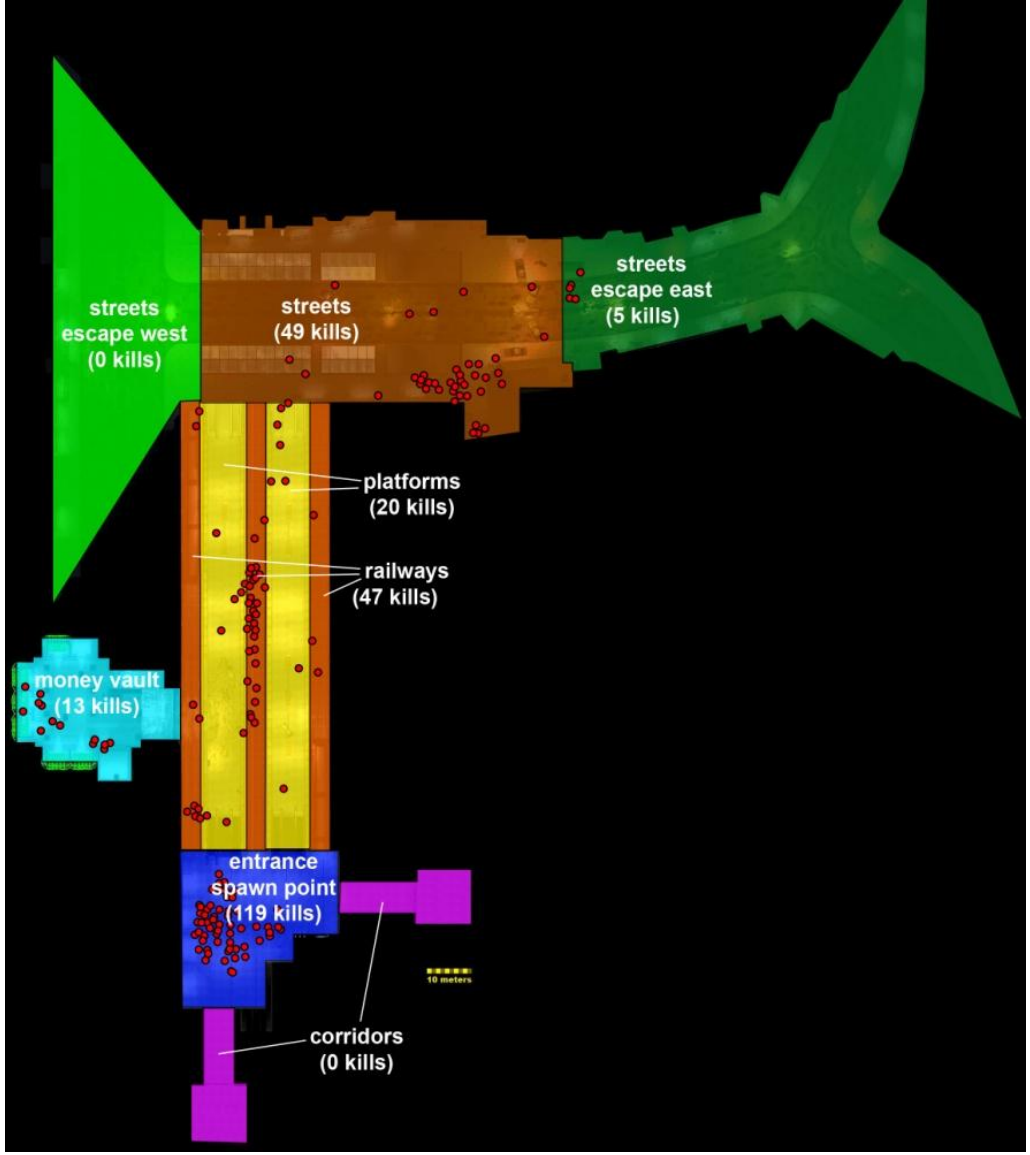
CON

- Overkill
  - for simple, non-spatial analyses
- Not integrated with game engine
- Limited 3D representation



Heatmap

HoD requests have been plotted and a density kernel calculated into a heatmap to visualize the distribution of areas with high and low intensities of requests



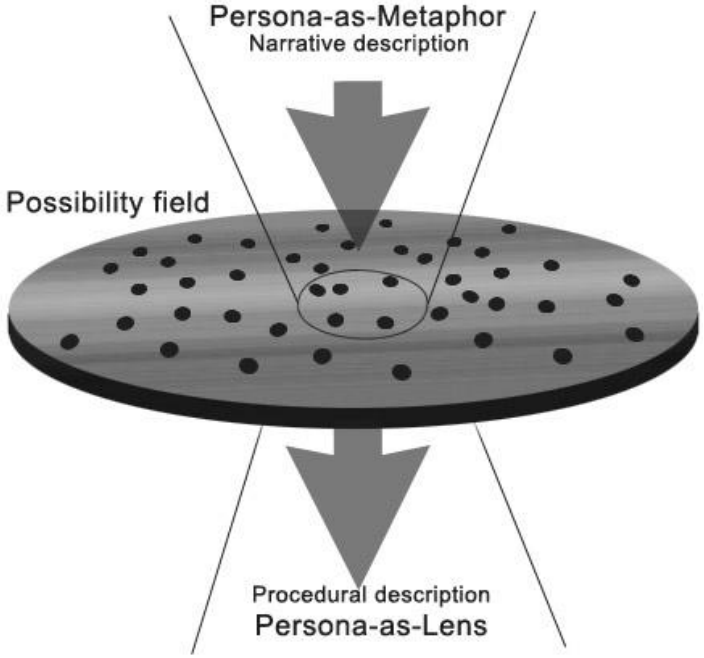
Deaths in Sectors

Plotted deaths divided per sub-sector

Play-Personas

aggregate descriptions of possible player behaviour:

- a-priori ***metaphor*** (hypotheses): theoretical models of ideal users, expectations of the designer



- a-posteriori ***lens*** (analyses): data-driven representations of player behaviors, descriptions of what actual, real players do during play

Play-Personas

- Pre/Production
  - Envision different play experiences
- After Launch
  - Evaluation of experiences

Hypothesising and analysing what players repeatedly do, sheds light on what their **goals**, **intentions** and **desires** are at a precise moment in time and in a precise context: the game.

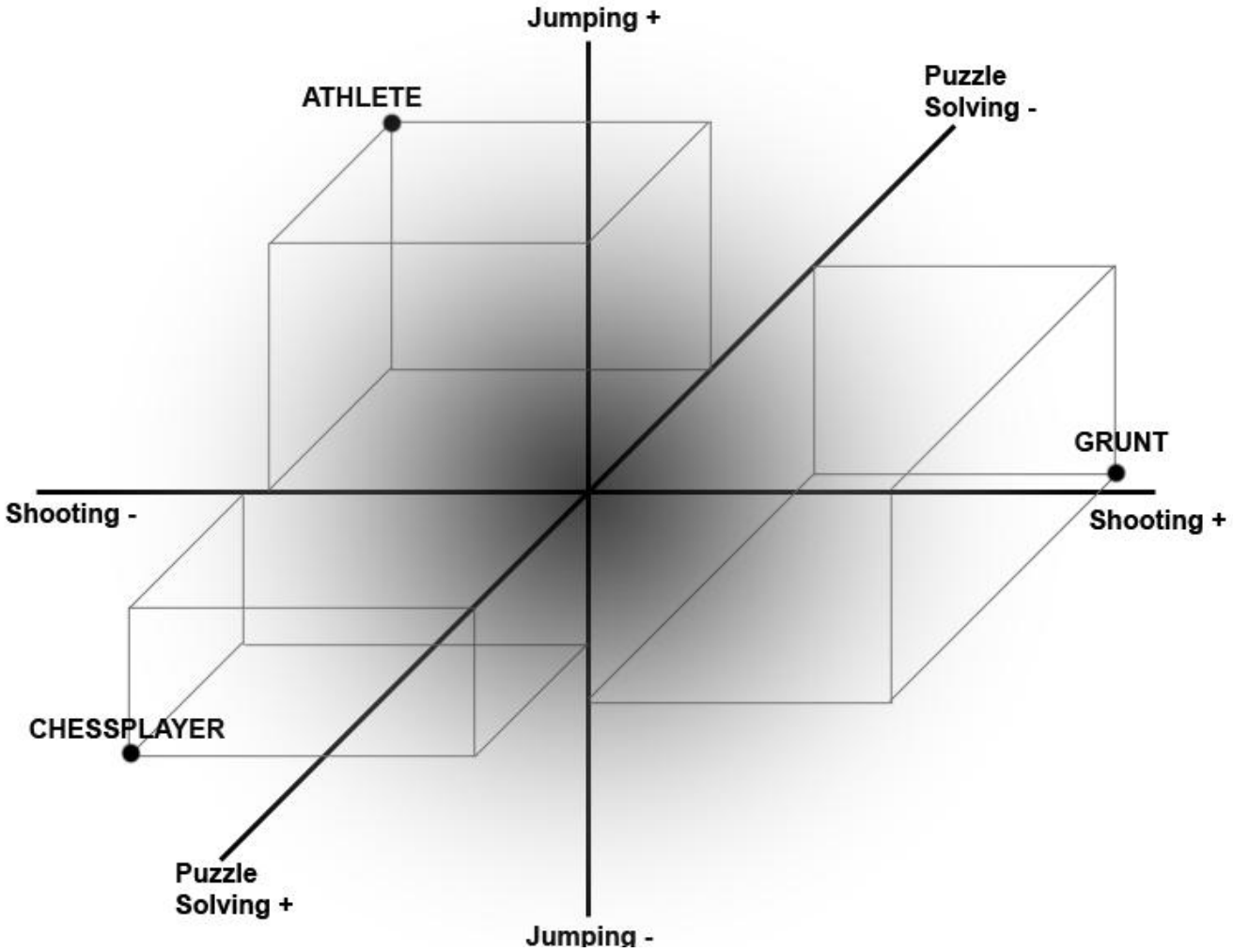
Play-Personas

PRO

- Experiences easier to
  - Design
  - Analyse
- Focus
  - Play experience
  - Player behavior
- Map playing landscape
- Provide varied experience

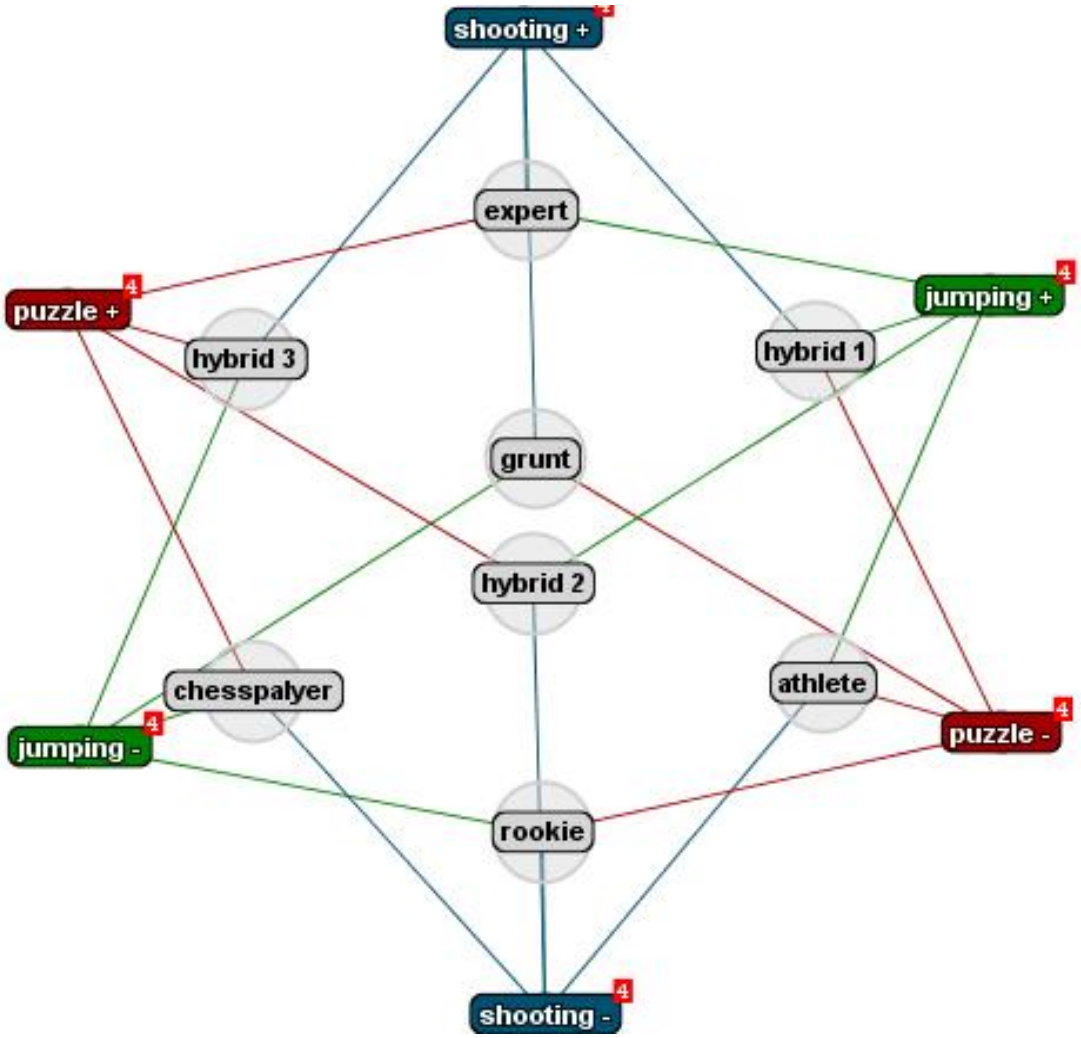
CON

- Risk of truisms
- No detection of
  - problems unrelated to patterns of play
- Not useful for usability issues



Play persona possibility space

Mapping the possibility space with play-personas



Gameplay parameter relations

Persona hypotheses emerge as relations between parameters that have been derived from gameplay mechanics.

Some references

Tychsen, A. and Canossa, A., Defining personas in games using metrics. In 2008 Conference on Future Play: Research, Play, Share, (Toronto, Ontario, Canada, 2008), ACM, 73-80.

Canossa, A. and Drachen, A., Play-Personas: Behaviours and Belief systems in User-Centred Game Design. Interact Conference 2009. Uppsala, Sweden.

Tychsen, A. and Canossa, A., Analyzing User Behavior via Gameplay Metrics. Future Play 2009.

Galvanic Skin Response

- Measured on palms and soles of feet- Eccrine sweat glands
- Measures electrical resistance (or conductance) between two electrodes
- Correlate to psychological arousal

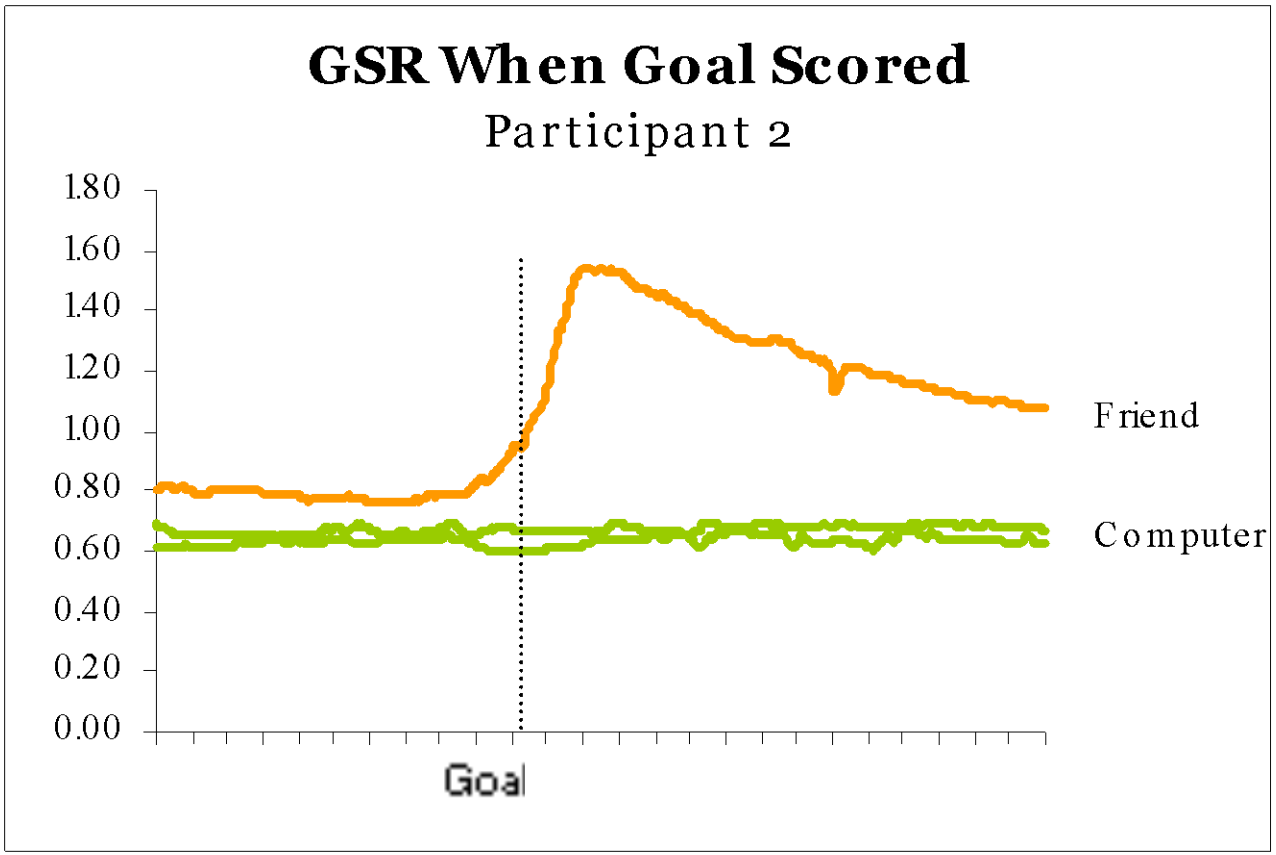
Galvanic Skin Response

PRO

- Easy to measure
- Inexpensive hardware
- Easy to interpret
- Non-intrusive (could be built into a device)

CON

- Noisy signal
- Large individual variations in baseline and responsivity
- Slow decay (signals add together)

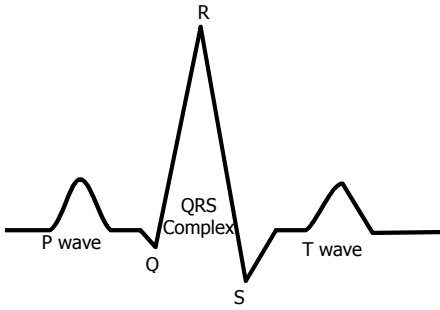


Example Usage

Three instances of GSR when a goal was scored in NHL 2003

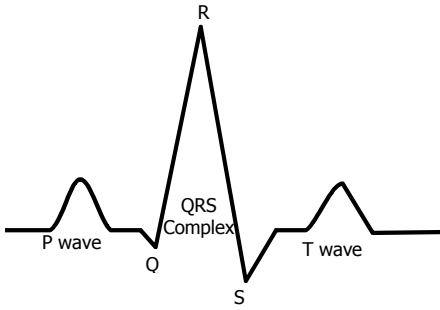
- twice against the computer and once against a friend

Cardiovascular Measures



- Electrocardiography (EKG)
- Heart Rate (HR)
- Interbeat Interval (IBI)
- Heart Rate Variability (HRV)
  - Spectral analysis of sinus arrhythmia
  - Indicative of mental effort, cognitive load
- Blood Volume Pulse (BVP) (periodic)
- Blood Pressure (BP)

Cardiovascular Measures



PRO

- Easy to measure some signals (HR)
- Inexpensive hardware (HR)
- Salient and established measures

CON

- Intrusive to measure accurately
- Affected by many things (e.g., physical activity)
- Complex analysis (HRV)

Electromyography

- Isometric tension, or detection of motion
- Needles or surface electrodes
- Tension in the jaw
- Forehead (smiling vs. frowning)
- Can be used on any muscles

Electromyography

PRO

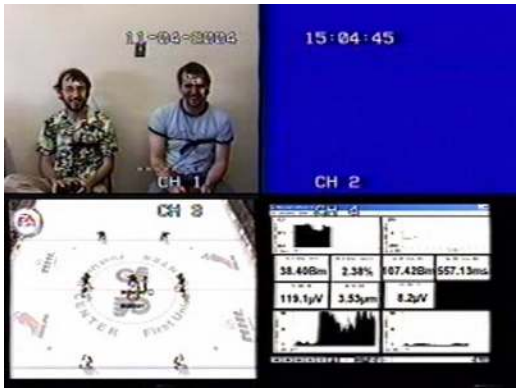
- Analysis of signals easy
- More sensitive than image processing for facial expressions
- Easy to interpret

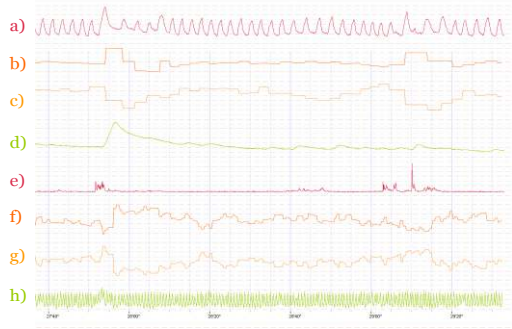
CON

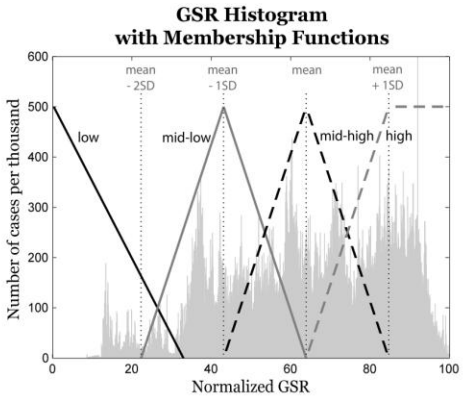
- Intrusive to measure
- Difficult to get natural measures
- Hardware is expensive
- Interference of muscle groups

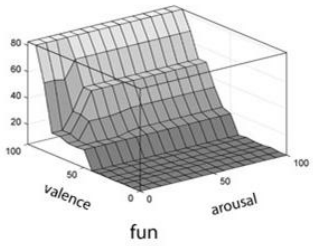
EMG, HR, and GSR (and respiration)

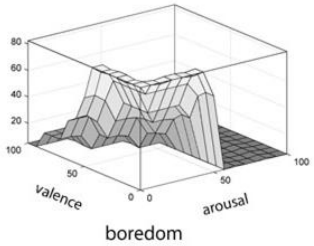
Intrusiveness of sensors is clear, but participants forgot about them after a short time

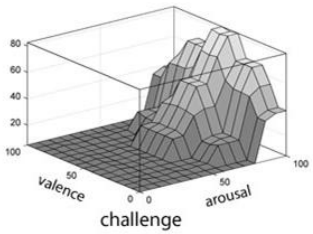

15:04:45
CH 1
CH 2
CH 3

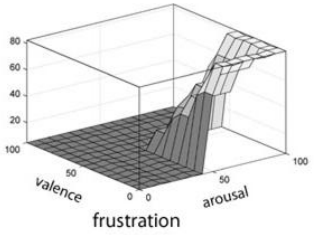

a)
b)
c)
d)
e)
f)
g)
h)

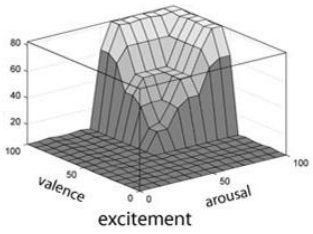

GSR Histogram
with Membership Functions
mean - 2SD
mean - 1SD
mean
mean + 1SD
low
mid-low
mid-high / high
Number of cases per thousand
Normalized GSR


valence
arousal
fun


valence
arousal
boredom


valence
arousal
challenge


valence
arousal
frustration


valence
arousal
excitement

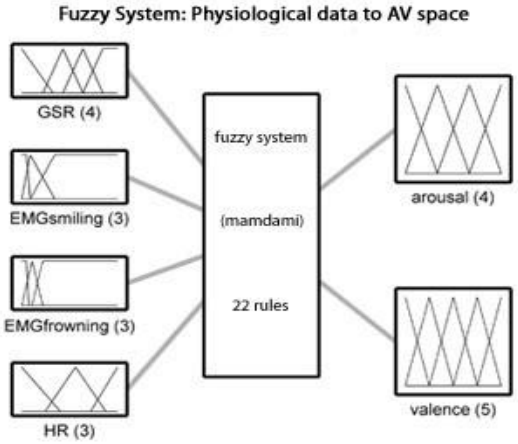

Fuzzy System: Physiological data to AV space
GSR (4)
EMGsmiling (3)
EMGfrowning (3)
HR (3)
fuzzy system
(mamdami)
22 rules
arousal (4)
valence (5)

TABLE OF CONTENTS

For more information:

R.L. Mandryk (2008). Physiological Measures for Game Evaluation. in *Game Usability: Advice from the Experts for Advancing the Player Experience*. (K. Isbister and N. Shaffer, Eds.), Morgan Kaufmann.

Some References

R.L. Mandryk, and M.S. Atkins (2007). A Fuzzy Physiological Approach for Continuously Modeling Emotion During Interaction with Play Environments. International Journal of Human-Computer Studies, 6(4), pg. 329-347. The original publication is available at Elsevier Online.

R.L. Mandryk, K.M. Inkpen, and T.W. Calvert (2006). Using Psychophysiological Techniques to Measure User Experience with Entertainment Technologies. Behaviour and Information Technology (Special Issue on User Experience), Vol. 25, No.2, March-April 2006, pg. 141-158.

R.L. Mandryk, M.S. Atkins, and K.M. Inkpen (2006). A Continuous and Objective Evaluation of Emotional Experience with Interactive Play Environments. in Proceedings of the Conference on Human Factors in Computing Systems (CHI 2006). Montreal, Canada, April 2006, pg. 1027-1036.

R.L. Mandryk (2008). A physiological approach for continuously modeling user emotion in interactive play environments. in Proc of Measuring Behavior 2008, Maastricht, NE, August 2008, pg. 93-94.

Heuristic Evaluation for Games

- Few formal methods exist for evaluating the usability of game interfaces
- Developed usability principles for video game design
- Heuristics can be used to carry out usability inspections of video games

Developing Game Usability Heuristics

- Step 1: identify problems from game reviews
  - 108 reviews from *GameSpot*
  - 6 major PC game genres
- Step 2: develop problems categories
  - 12 common categories found
- Step 3: develop game heuristics
  - 10 heuristics created from problem categories

Usability Heuristics

1. Provide consistent responses to user's actions
2. Allow users to customize video and audio settings, difficulty and game speed
3. Provide predictable and reasonable behaviour for computer controlled units
4. Provide unobstructed views that are appropriate for the user's current actions
5. Allow users to skip non-playable and frequently repeated content
6. Provide intuitive and customizable input mappings
7. Provide controls that are easy to manage, and that have an appropriate level of sensitivity and responsiveness
8. Provide users with information on game status
9. Provide instructions, training, and help
10. Provide visual representations that are easy to interpret and that minimize the need for micromanagement

Usability Heuristics

PRO

- help identifying game-specific usability problems
- applicable to mockups and prototypes
- can be used to evaluate most games

CON

- does not address engagement and “playability”
- limitations in the development of heuristics

Some References

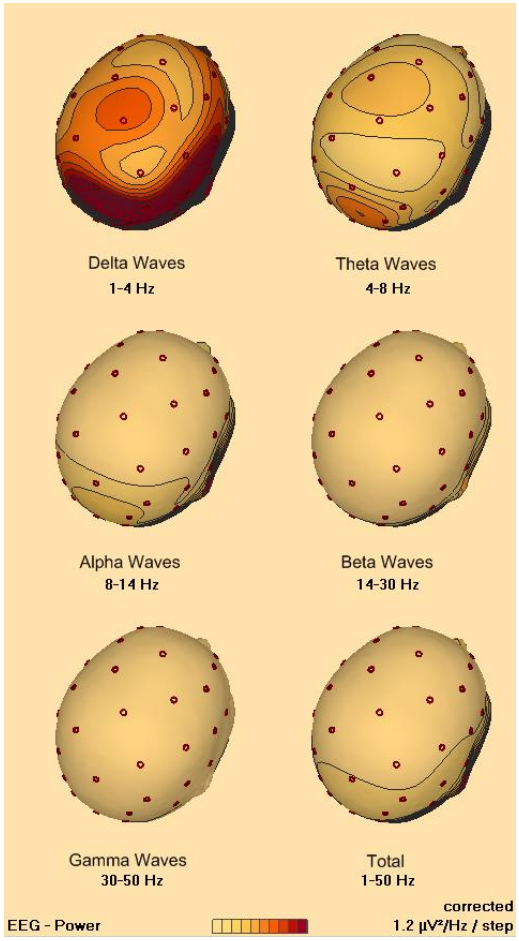
- Pinelle, D., Wong, N., Stach, T., Gutwin, C. (2009) Usability Heuristics for Networked Multiplayer Games. To appear in *GROUP 2009*.
- Pinelle, D., Wong, N., Stach, T. (2008) Using Genres to Customize Usability Evaluations of Video Games. *Future Play 2008*, 129-136.
- Pinelle, D., Wong, N., Stach, T. (2008) Heuristic Evaluation for Games: Usability Principles for Video Game Design. *CHI 2008*, 1453-1462.

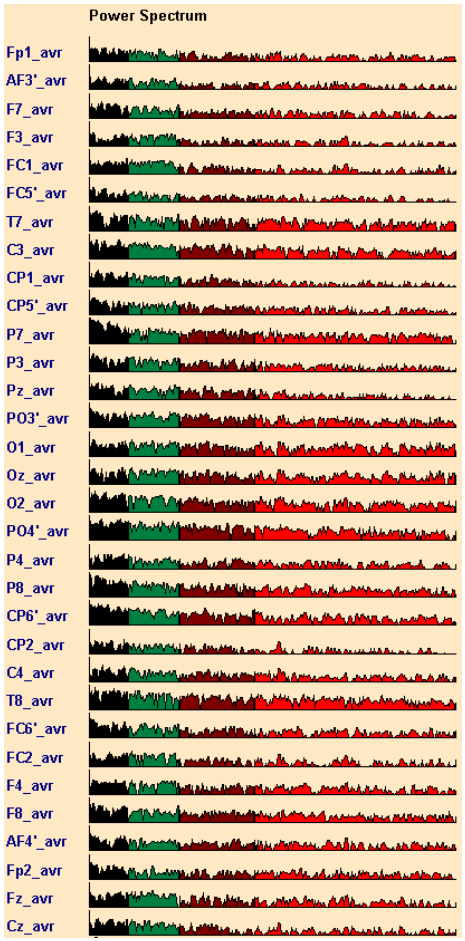
EEG

- Electrodes placed on scalp (from 20 to 256)
- Measures electric potentials
- Brainwaves are described in frequency bands
  - Delta (trance, sleep)
  - Theta (emotions, sensations)
  - Alpha (calm, mental work)
  - Low beta (focus, relaxed)
  - Mid beta (thinking, alert)
  - High beta (alert, agitated)
  - Gamma, seldom (information processing)

Game experiment Setup

EEG and EMG electrodes are being attached. The Biosemi electrode cap consists of 32 electrodes in the areas: frontal (F), parietal (P), temporal (T), occipital (O), central (C).





EEG Frequencies and Spectrum

EEG Analysis is difficult. After artifact scoring, values have to be transformed for spectral analysis.

EEG

PRO

- Objective
- Covert & continuous recording
- Quantifiable
- Reliable
- Replicable
- Empirical power

CON

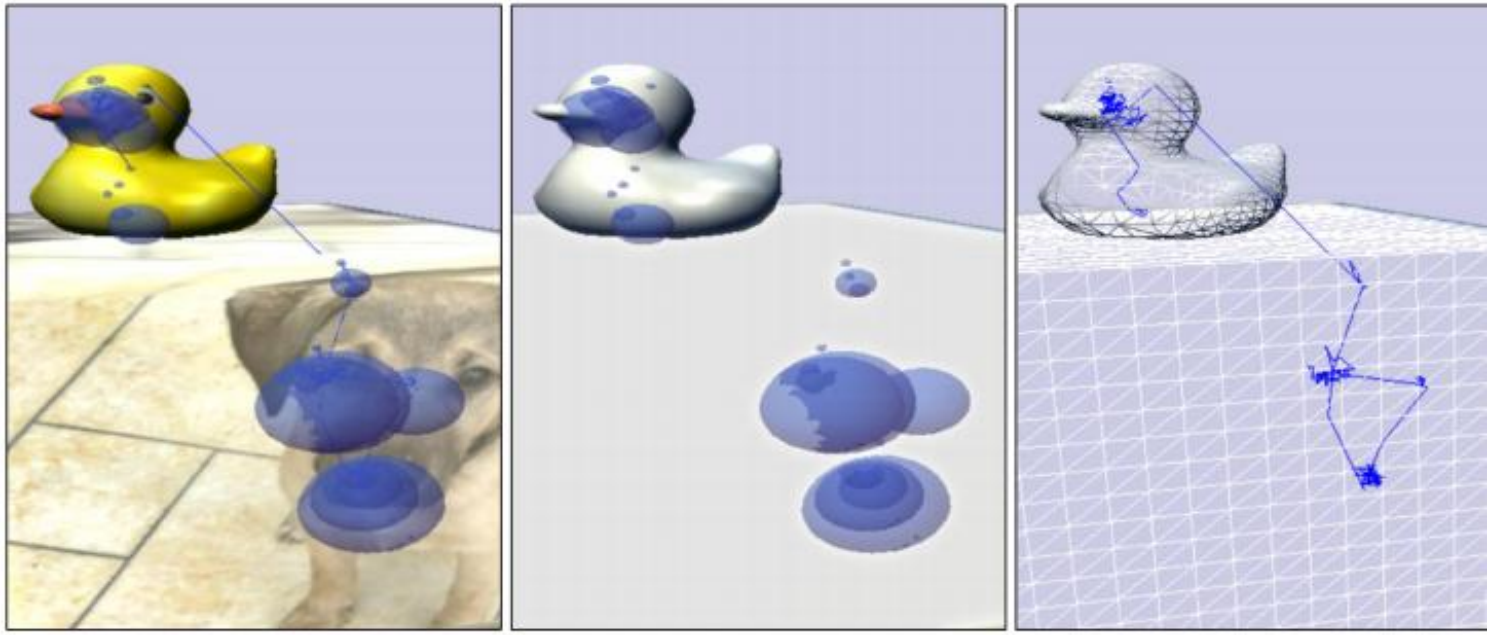
- Intrusive
- Expensive
- Artifact scoring
- Time-consuming
- Sometimes hard to interpret

Eye Tracking

- Measures what eyes look at
  - Saccades (fast movement)
    - Gaze path
  - Fixations (dwell times)
    - Attention focus
  - Pupil dilation/blink rate
- Attention precedes gaze (200ms)
- Used mainly to improve interface
  - Lack of 3D analysis tools

Experimental playing session

Experimental gaming session with all logging equipment in place.



Example of 3D Eye Tracking Visualization

Viewed game world objects can be displayed together with their gazepaths in 3D (see also Stellmach, 2009)

Eye Tracking

PRO

- Easy to use
- Objective
- Covert
- Continuous
- Quantifiable
- Replicable
- Empirical power

CON

- Can be expensive
- Lack of good tools
- Time-consuming

Some References

Nacke, L. and Lindley, C.A., Flow and Immersion in First-Person Shooters: Measuring the player's gameplay experience. In *Proceedings of the 2008 Conference on Future Play: Research, Play, Share*, (Toronto, Canada, 2008), ACM, 81-88.

Grimshaw, M., Lindley, C. A., & Nacke, L. (2008). Sound and Immersion in the First-Person Shooter: Mixed Measurement of the Player's Sonic Experience. Audio Mostly Conference 2008, Piteå, Sweden.

Nacke, L., Lindley, C., and Stellmach, S. (2008) Log Who's Playing: Psychophysiological Game Analysis Made Easy through Event Logging. In Proceedings of the 2nd international Conference on Fun and Games (Eindhoven, The Netherlands, October 20 - 21, 2008). P. Markopoulos, B. Ruyter, W. Ijsselsteijn, and D. Rowland, Eds. Lecture Notes In Computer Science, vol. 5294. Springer-Verlag, Berlin, Heidelberg, 150-157.

Stellmach (2009). Visual Analysis of Eye Gaze Data in Virtual Environments. Master's Thesis. University of Magdeburg.

Discussion

- Is an integration of the presented methods feasible?
- Can they be integrated in a cost-efficient way?
- Which methods are suitable for evaluating which parts of game development?
- Can empirical data be applied to game design? How?

Discussion

- Should there be a discussion about separating quantitative from qualitative or do we agree on integrated measures?
- What can these methods be used for beyond evaluation? Exergames? Biofeedback?
- Are those methods improving games? If yes, how can (or should) they be adopted by the majority of the game industry?

More questions?

- Audience

Find out more…

- Mike: www.valvesoftware.com
- Alessandro: www.dkds.dk
- Regan: www.reganmandryk.com
- Tad: equis.cs.queensu.ca
- Lennart: www.acagamic.com
  - project.hkkk.fi/fuga/

Thanks a lot!