
UNIVERSITY OF CAPE TOWN
Department of Computer Science
Measuring Presence
Edwin Blake
edwin@cs.uct.ac.za

# Outline



- Measurement Issues
- Introspection
  - Examples in Presence measurement
  - Breaks in Presence
- Measures of experience
  - Sense of Presence Inventory
- Behavioural Measures
- Physiological Measures
- Comparison

# General problems



- Presence is a multi-dimensional concept that involves psychological processes
- Researchers face significant challenges in developing valid and reliable measures of presence
- Two general approaches:
  - subjective and
  - objective

# Measurement primer (quick revision)



- Measures are evaluated on two properties:
  - *Validity*: is the scale measuring the correct construct?
    - request information logically related to what we understand presence to be
  - *Reliability*: How sensitive is the measure to noise?
    - proven test-retest repeatability
- These properties are largely independent
- A “good” measure should have high degrees of each

# Further problems



- Presence is not well defined theoretically
  - What to look for?
  - What does a present subject look like?
- Is presence a continuous or discreet phenomenon?
  - Does “intensity” make any sense?
  - Is there a response curve to estimate?
  - Do we need to worry about ‘false positives’ and ‘false negatives’?

# The measurement equation I



Any measurement can be expressed as:

$$X = T + e$$

Where:

- X is the observed (measured) score
- T is the “true” score (what we want to get)
- e is the error (noise of the scale)

# The measurement equation II



- Things to note:
  - Noise (e) can be reduced but never eliminated
  - The degree of noise (reliability) does not affect what is measured
  - If the incorrect T is chosen, the X will also be wrong (no validity)
  - Reducing noise (e) only approximates T better, but does not ensure measurement of the correct thing

# Bad measure: example



Consider this calculus question:

> “What is the name of the big city in the middle of that little place in the middle of South Africa, a bit up from Joburg?”

# Bad measure: example II



Still the same calculus question, but reduce noise by making it more specific:

"What is the capital of Lesotho?"

Less ambiguous, therefore less noise (improved)

Note it is still not a calculus question!!! (reliability and validity are independent)

# Other Desirable Features for Measures



- Apart from Valid and Reliable:
- Sensitive — to variations in the variables affecting presence
- Non-intrusive — avoid unintentional degradation of performance or sense of presence
- Convenient:
  - portable
  - low cost
  - easy to learn
  - easy to administer
  - easy to analyse

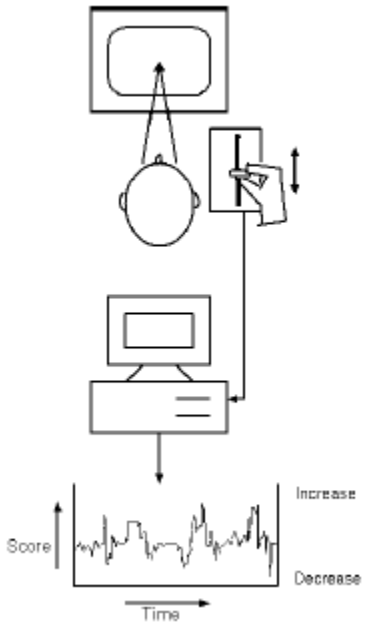
# Outline



- Measurement Issues
- **Introspection (self-report)**
  - Examples in Presence measurement
  - Breaks in Presence
- Measures of experience
  - ITC Sense of Presence Inventory
- Behavioural Measures
- Physiological Measures
- Comparison

# Introspection Measures






- Ask subjects to examine own experiences and report on it
  - Subjective measures: conscious, introspective judgment regarding experience
- Usually after the experience
  - Exception: IJsselsteijn et al (1998) and Freeman et al (1999) — continuous slider method;
    - Nonintrusive?
  - Schloerb (1995) – reality decision

# Personal Presence Measurement



- Early major approaches to measuring Personal Presence:
- The Slater, Usoh & Steed (SUS) questionnaire
- The Presence Questionnaire (PQ) of Witmer & Singer
  - This is accompanied by an Immersive Tendencies Questionnaire (ITQ)

# Introspection example — SUS I



- Good example: Slater, Usoh & Steed (1994)
  - Widely used (most widely used?)
- Six items only
  - 7 point semantic differential scale
  - When you think back about your experience, do you think of the virtual room more as images that you saw, or more as somewhere that you visited?

    *The virtual room seems to me more like ... 1 Images that I saw ... 4 About 50/50 ... 7 Somewhere that I visited*
  - During the course of the experience, which was stronger on the whole, your sense of being in the virtual room, or of being in the real world of the laboratory?

    *I had a stronger sense of being in ... 1 The real world in the lab ... 4 About 50/50 ... 7 The virtual room*

# Introspection example — SUS II



- Not validated by its authors
  - Some evidence exists to show it has construct validity
  - Sensitive to hardware changes and other immersive variables
  - Significant but weak correlations with other scales (e.g., Witmer & Singer's Presence Questionnaire)
- Fairly good levels of reliability

# Introspection — discussion



- Slater (1999) – introspection is too subjective
  - Tied into personal factors – 2 users, same experience, different scores on the same measure
- Subjectivity is only a concern for reliability, not validity
  - Signal:Noise will be high
- Not yet clear if presence is not tied into personal factors or not

# Introspection — Advantages



- Appear to be valid
  - ITC-SOPI, TPI, SUS, PQ, MEC-PQ show correlations with each other and detect changes in immersion
- Easy to use
  - Include instructions, simple Likert format
- Cheap
  - No special equipment, can run electronically also
- Easy to analyse
  - Linear modelling techniques work well

# Introspection — limits I



- Is introspection an inherently poor method?
  - Nisbett & Wilson (1977) investigated this with the “stocking study”
  - Found people always gave reasonable introspections even if the reasoning was flawed
  - Nisbett & Wilson argue that introspection is not a memory of a process, but a constructive process itself which maintains a sense of meaning

# Introspection — limits II



- Introspection has limits
  - OK: Memories, intensity of stimuli, contents of consciousness
  - Not OK: why they did something, abstract comparisons
    - Especially not if you refer to presence concepts themselves
- Many items on presence questionnaires go beyond reasonable limits

# Improved introspection — BIPs I



- BIP: any perceived phenomenon during the VE exposure that launches the participant into awareness of the real world setting of the experience
- Breaks their presence in VE
  - gross events, collisions with equipment,
  - subtle effects, seeing a tree as a pixel map rather than a solid object.
- Equivalent to the Aha! of Gestalt that switches rabbit to duck

# Improved introspection — BIPs II



- Slater & Steed (2000) – count the “breaks in presence” during the experience
  - Validated against SUS (r = 0.8; n = 20)
  - Validated using immersion test



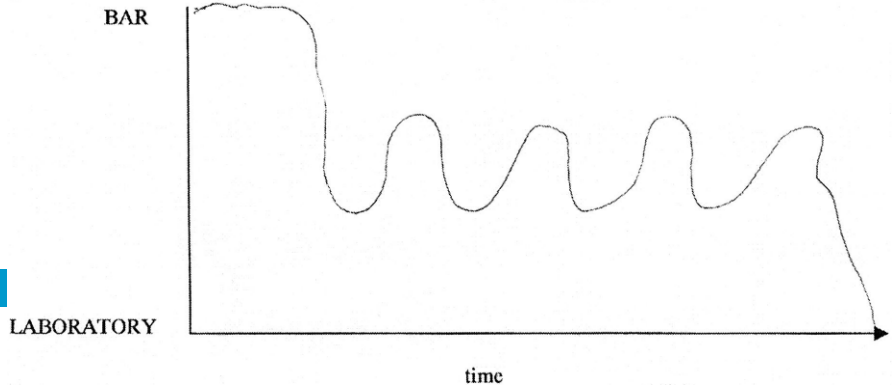
# Improved Introspection — BIPs II



- Simple perceptual task
  - May be prone to underestimation (forget to report a break)
- Criticisms of BIPs method
  - Validation not impressive – should have correlated with multiple scales
  - Almost no evidence that presence is dichotomous (not solved by curve threshold explanation)
  - VE task, emotions, etc can all affect attention and thus BIPs reporting rate

# Measuring BIPs






- Users draw a graph of presence
  - Starts off high, interrupted 4 times by triggered anomalies
- Questionnaire about BIPs (“transitions to real”):

3. Were there any moments when you became suddenly aware of the laboratory?
4. When?
5. How often?
6. What triggered these moments?
7. How did they make you feel?
8. How easy or quick was your recovery (recovering your sense of being in the bar)?
9. Did the intensity of the transitions vary?
   - Garau et al. 2008

# Outline



- Measurement Issues
- Introspection
  - Examples in Presence measurement
  - Breaks in Presence
- **Measures of experience**
  - ITC Sense of Presence Inventory
- Behavioural Measures
- Physiological Measures
- Comparison

# Measures of experience



- Try to get a ‘direct’ hold of the experience
- ‘Divide & conquer’ technique – look for the cluster of small experiences which makes up presence
- Often have a multi-factor structure; they see presence not as monolithic (tied to particular theories)

# Psychometrically Validated: IPQ



- Igroup Presence Questionnaire (IPQ)
  - Three subscales
    - emerged from principal component analyses
    - fairly independent factors
  1. Spatial Presence —sense of being physically present
  2. Involvement — attention devoted to VE & involvement experienced
  3. Experienced Realism —subjective experience of realism

  + general item: “sense of being there”

# Experience measure — ITC SOPI



- The Independent Television Commission's Sense of Presence Inventory
  - Extensive psychometric evaluation
    - Data from more than 1200 subjects
  - Ongoing development
    - 2001 version is 'stable'
  - Cross-media
    - Originally developed to research 'immersive TV' (which never worked)
- J. Lessiter, J. Freeman, E. Keogh & J. Davidoff, "A cross-media presence questionnaire: The ITC-Sense of Presence Inventory", Presence: Teleoperators and Virtual Environments, 10 (2001), 282-297 — on web page.

# Four Factors of ITC-SOPI



**Spatial** — traditional 'being there' presence

- Physical placement within VE and interaction with objects in VE
- 19 items

**Engagement** — enjoyable/compelling experience

- Psychological involvement and a tendency to enjoy the VE experience.
- 13 items

**Naturalness** —matched subject's expectations

- VE is believable, lifelike or realistic (ecological validity)
- 5 items

**Negative effects**

- Negative physiological reactions like dizziness, eyestrain and headaches
- 6 items

# Scoring ITC-SOPI



- No one presence score — totally independent
  - Each of the factors provides a separate score for the experience.
- Satisfy a number of presence concepts at once
  - Slater: simply consider the spatial presence factor
  - IJsselsteijn and colleagues: consider both the engagement and spatial presence factors.

# ITC-SOPI Example Questions



- A AFTER MY EXPERIENCE OF THE DISPLAYED ENVIRONMENT...
  1. I felt sad that my experience was over..... 1 2 3 4 5
  3. I had a sense that I had returned from a journey....
- B DURING MY EXPERIENCE OF THE DISPLAYED ENVIRONMENT...
  1. I felt myself being 'drawn in'.....1 2 3 4 5
  4. I felt I could interact with the displayed environment...
  12. I felt I wasn't just watching something..
  14. I felt dizzy.....
  17. I paid more attention to the displayed environment than I did to my own thoughts (e.g., personal preoccupations, daydreams etc.)....
  18. I had a sense of being in the scenes displayed....
  22. I could almost smell different features of the displayed environment....
  34. I felt as though I was in the same space as the characters and/or objects.
  36. It felt realistic to move things in the displayed environment.....
  37. I felt I had a headache.....

# Using the ITC-SOPI



- Can be used for any medium (books, TV, caves, etc)
  - Useful for cross-media comparisons
- Non-reliance on a particular definition of presence (e.g. like the SUS is) allows more discovery
- Separate factors allow a far more fine grained examination of media experiences
  - E.g., Books might always produce low spatial scores, but what about engagement and naturalness?

# Outline



- Measurement Issues
- Introspection
  - Examples in Presence measurement
  - Breaks in Presence
- Measures of experience
  - Sense of Presence Inventory
- **Behavioural Measures**
- Physiological Measures
- Comparison

# Objective Measures

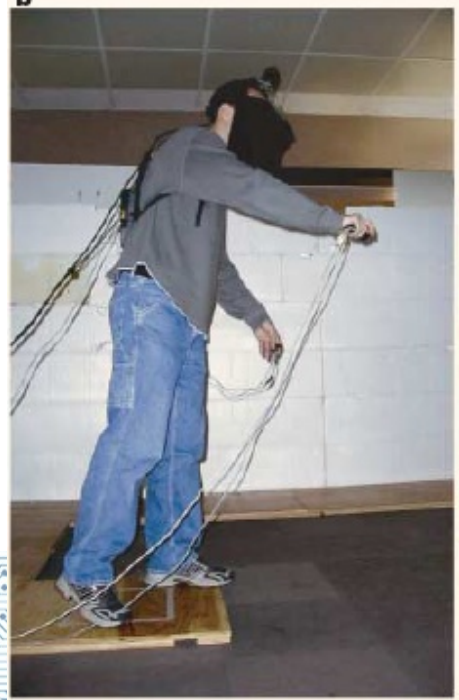


- *Physiological* and/or *Behavioural* responses logically correlated with psychological responses.
- No conscious introspection
- Administered during the experience (not following)
- For example, changes in
  - skin conductance,
  - blood pressure,
  - heart rate,
  - muscle tension,
  - respiration,
  - ocular responses,
  - posture.

# Behavioural Measures



- Check to see if subjects have bodily responses to virtual stimuli
- Usually the VE contains so structure which can be expected to produce response
  - Cliffs
  - Flying missiles
  - Sound sources

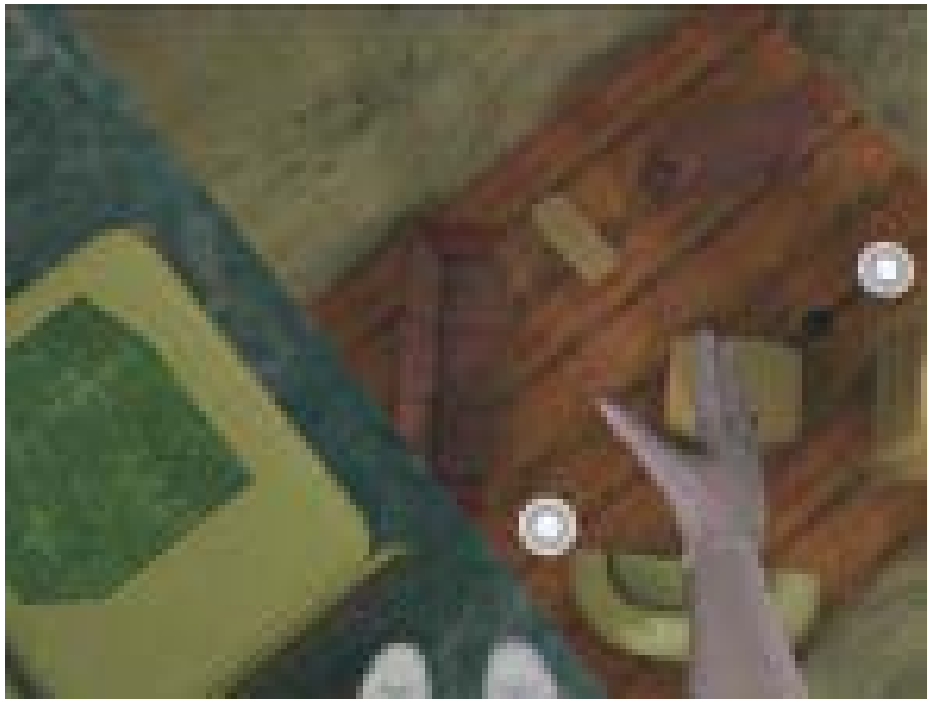



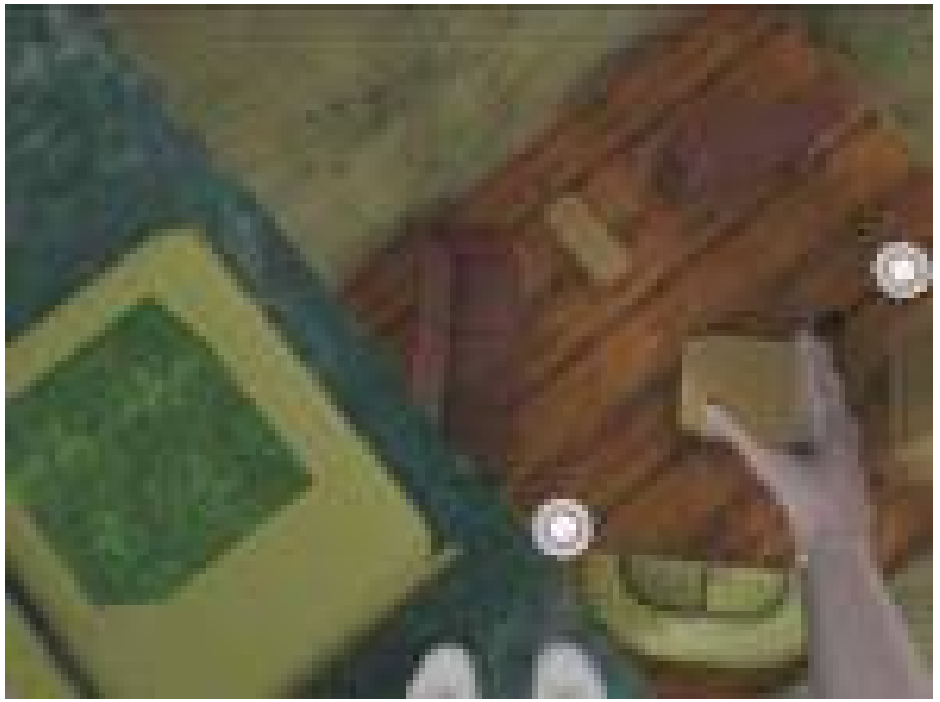


- a Enter left-most room, used for familiarization.
- Task: go into the next room, select object left on the plank and take it to the other side of the room.
- b,c ‘Passive haptics’. Participant positioned at small but real ledges.
- Computer Science, Univ of North Carolina, Chapel Hill

# Stereo Pair







# Behavioural measures example — Loomis



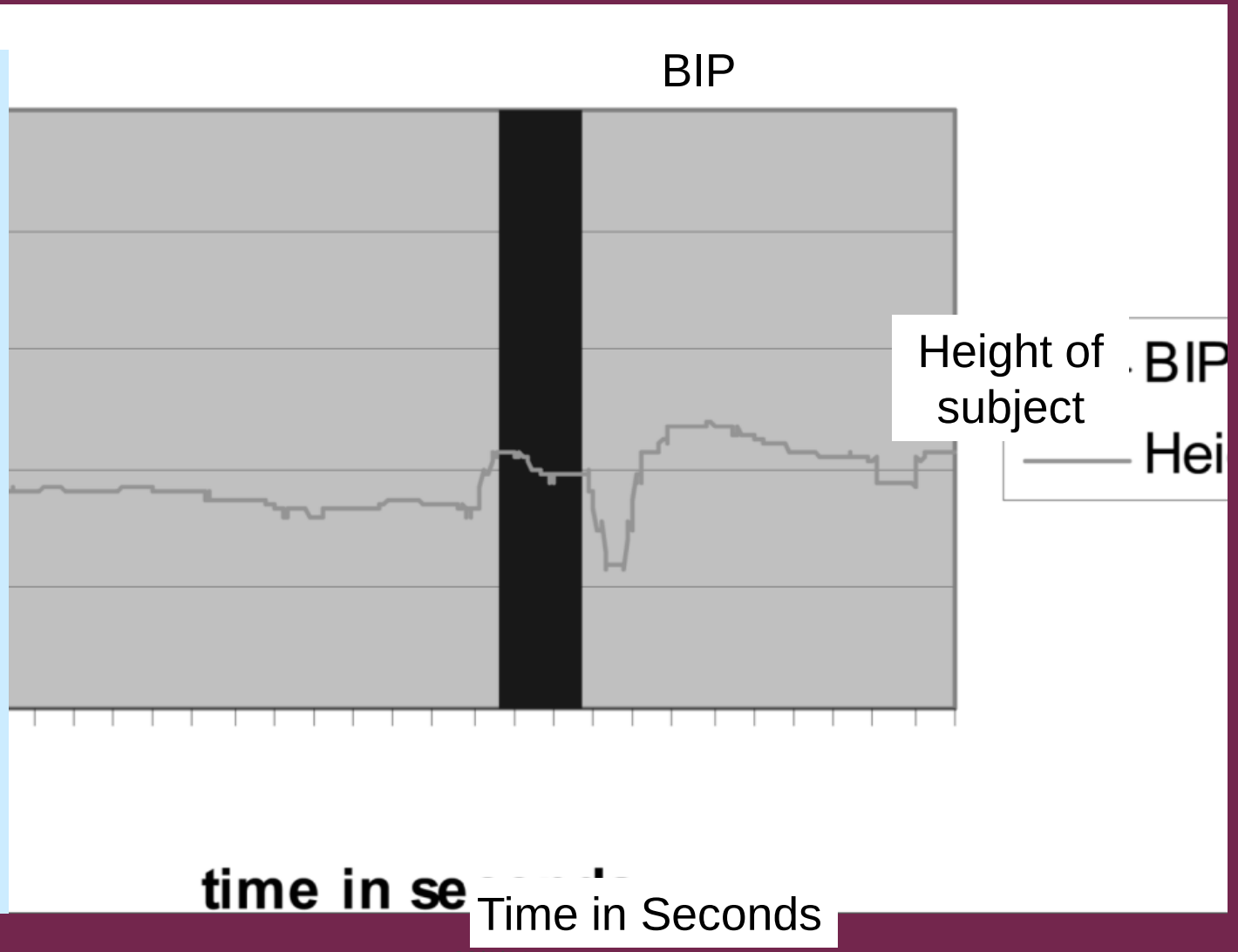
- Loomis (1992)
  - Certain behaviours initiated automatically
  - Includes reflexes, strong social cues
  - Responses may differ, but the initiation of behaviour is guaranteed
  - Loomis argues that this type of response to a virtual object shows belief that the object is “real”

# Subtle Behavioural Measures



A graph of one subject's head height, as measured by the head tracker, following a BIP event.

Friedman et al. Presence 15(5) Oct 2006, 599-610



# Behavioural measures — discussion



- Free from subject interpretation (fast)
- No memory effects (during the experience)
- Problem: is believability an issue?
  - Automatic behaviours occur below level of awareness
  - Subject can still believe the VE to be fake
  - Reflexes can be triggered by extremely low quality stimuli
  - May tell us little about the VE

# Behavioural measures — discussion II



- Not suitable as a general purpose measure
  - Requires changes to VE (cliff, etc)
  - Less intrusive forms exist
    - Social interactions (turn taking, posture, gaze, etc)
    - Posture changes during simulated motion
- No clear method of quantifying behaviour
- Adds problem of reactivity
  - Was that movement presence or not — researcher must decide (and their bias can creep in)

# Outline



- Measurement Issues
- Introspection
  - Examples in Presence measurement
  - Breaks in Presence
- Measures of experience
  - Sense of Presence Inventory
- Behavioural Measures
- **Physiological Measures**
- Comparison

# Physiological Measures



- Measure involuntary physiological reactions
  - Heart rate, galvanic skin response, breathing, etc.
- Physical measurements increase objectivity
  - Reduce memory effects, interpretation, etc

# Physiological measure example — Meehan



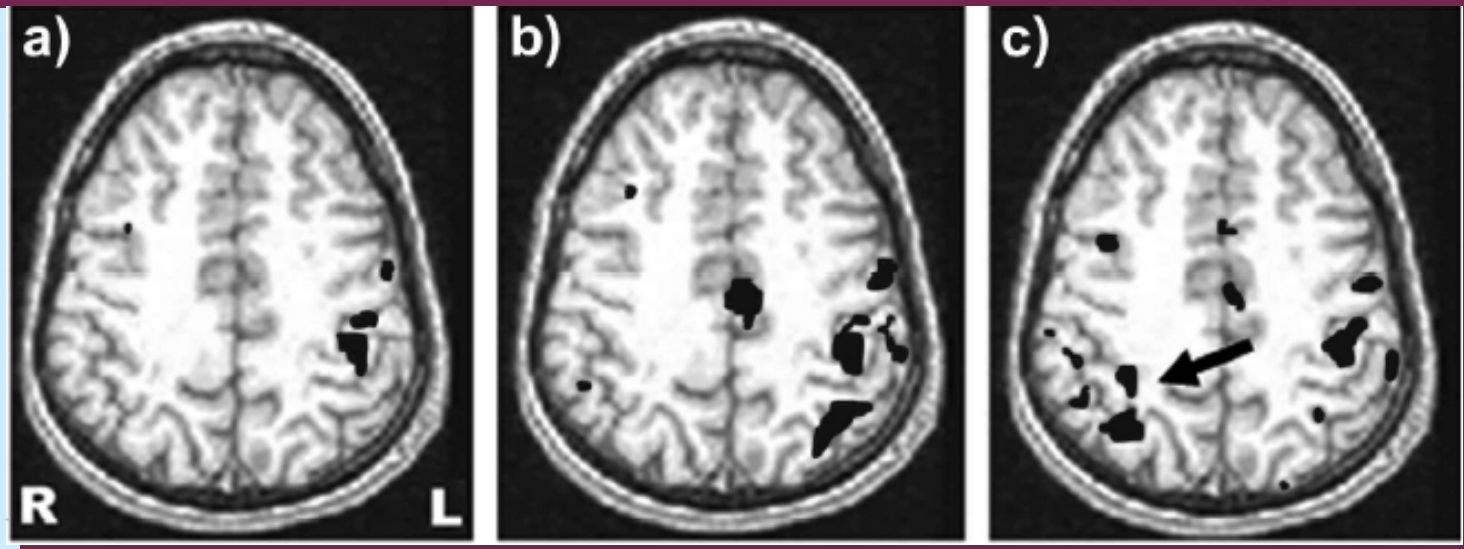
- Meehan et al (2002):
  - Measure presence by heart rate, breathing rate, skin conductance
  - Environment showed an unexpected sudden drop
  - Weak correlations to introspection measures (best r = 0.27; heart rate change to SUS)

# Brain Activity During Finger Tapping



Increased activation of brain regions involved in motor control is shown when visual feedback is provided

Differences in brain activity are also observed for the task conditions involving looking at the real and virtual hands.




a) without looking at the real hand,
b) looking at the real hand,
c) operating the data glove and looking at the virtual hand.

there is a marked increase in ipsilateral parietal lobe activity when operating the data glove (c: black arrow).

# Physiological measures — discussion I



- Objectivity may not be a worthwhile goal
  - Objectivity guarantees an increase in reliability (signal:noise increase)
  - Reliability is already quite good!
  - Validity is the central problem, but objectivity does not guarantee an increase
  - Other subjective scales show better validity then Meehan's method

# Physiological measures — discussion II



- Lack of theoretical basis
  - No particular physiological correlates of presence found
  - If found, likely to be confounded with other mental states
  - In Meehan et al, no way to argue that it was presence rather than anxiety from the cliff
- Physiological measurements of psychological states usually fail
  - E.g., Lie detection

# Outline



- Measurement Issues
- Introspection
  - Examples in Presence measurement
  - Breaks in Presence
- Measures of experience
  - Sense of Presence Inventory
- Behavioural Measures
- Physiological Measures
- **Comparison**

# Comparison to other states



- Measure psychological variables which are related to presence
  - Some other states are easier to measure than presence itself
- Based on the notion that being in an environment has psychological effects
  - E.g., Being in a scary place should produce fear

# Comparison example — simulator sickness



- Experiencing simulator sickness reduces presence
  - Can use it as an inverse measure of presence
- Can show the simulation is incomplete
  - Incomplete simulations may lead to less presence
- Simulator sickness will lead to a reduction in attention focusing
  - Attention is known to be a necessary condition for presence

# Comparison — simulator sickness II



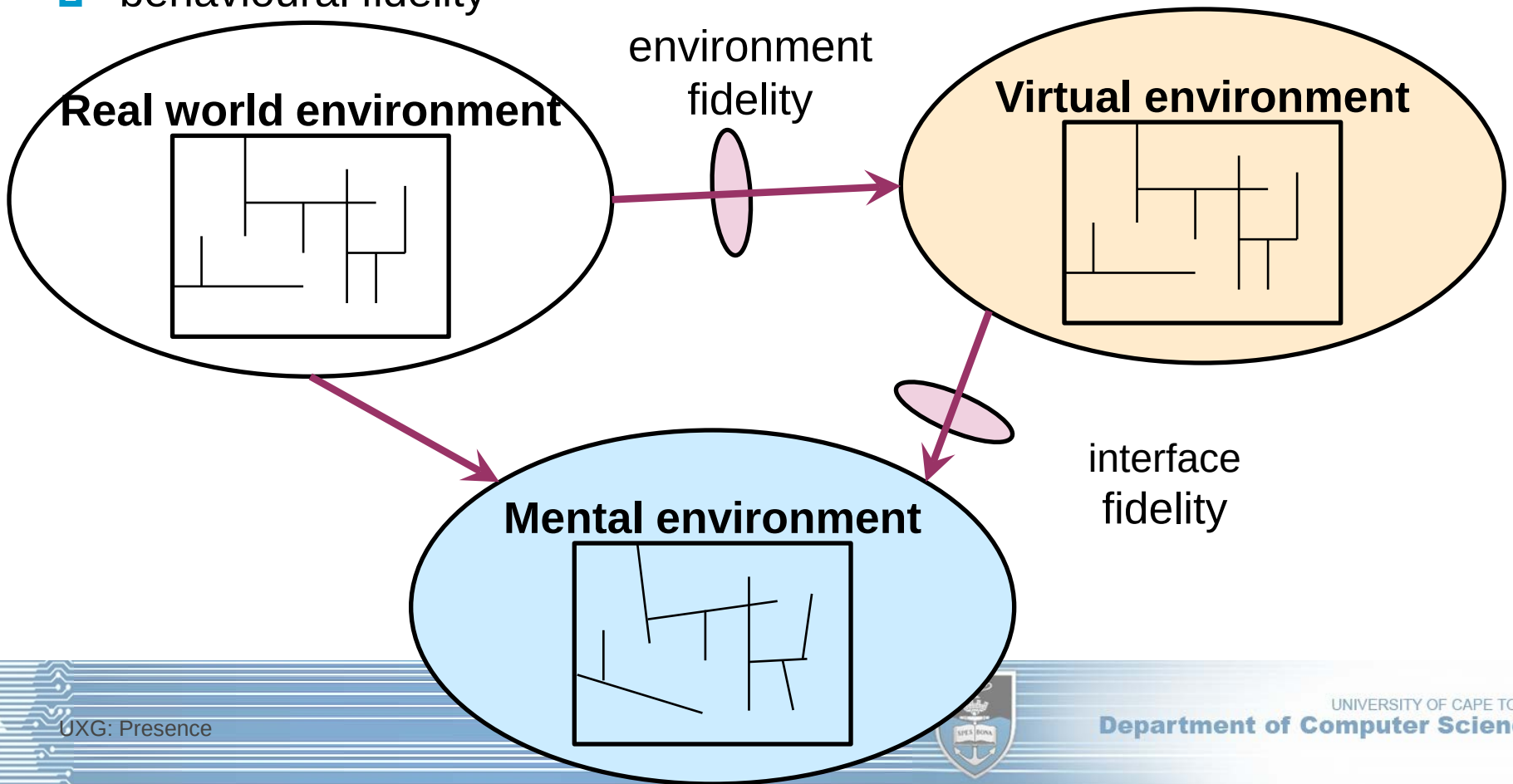
- Simulator sickness is simple to measure
  - Simulator Sickness Questionnaire (Kennedy et al, 1993)
  - Can be measured physiologically
- Problems in using SS:
  - Relationship between SS and presence not clear
  - May only be able to indicate when presence was not possible

# Comparison example — fidelity



Waller, Hunt & Knapp (1998): When a person is present, their reactions should be indistinguishable from those in a similar reality

- "behavioural fidelity"



# Comparison example — fidelity



- Implemented by Mania & Chalmers (2001)
  - Investigated memory performance
  - Found some relationship between immersion & memory for the VE space