Mobile Cognitive Assessment:

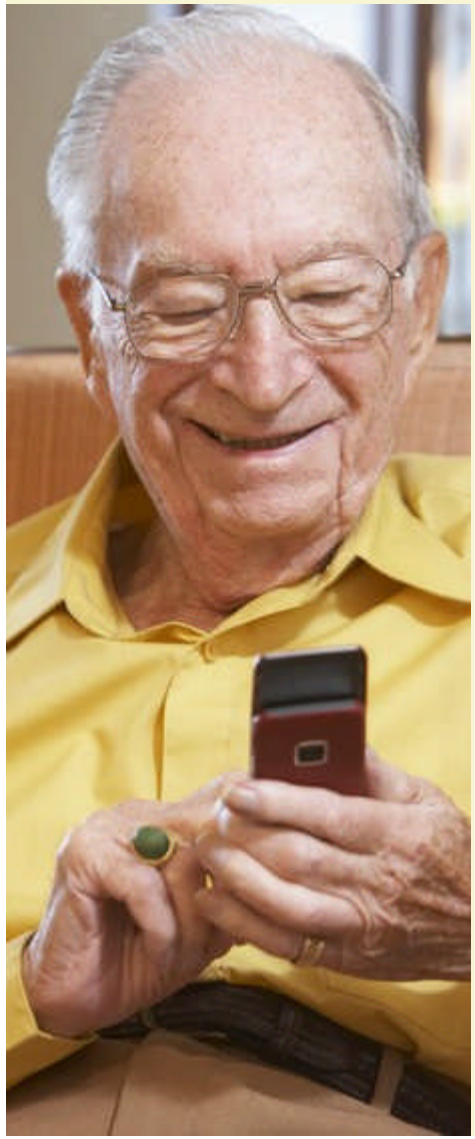
Validation of Neuropsychological Assessment Administered on a Mobile Phone

Brian Tiplady

Windermere,

September 2010

Mobile phone as cognitive test platform



- Many people already have a mobile phone and are comfortable using it
- Mobile phone has ample computing power for test presentation
- Small screen can be a limitation, but a wide variety of tests can be set up
- Tests can be run outside the laboratory
  - Field locations, such as workplace, pub, hospital ward
  - “Everyday life”, unsupervised testing

Validation questions with mobile phones

- Is the device capable of accurate timing?
- Is the software working as intended?
- Are tests with a small screen equivalent to other methods?
- Can unsupervised users cope with the tests?
- Are they using the system properly
- Are the tests capable of detecting the effects of interest?

Validation Model

Device Level Validation	Includes correct operation of phone and software, e.g. timing of responses, randomisation algorithms
Intrinsic Validation	Aspects of test data such as differences between responses to stimuli of differing difficulty
Extrinsic Validation	Ability of test to detect changes within individuals or between individuals due to well-established influences on cognition

Memory Scanning

A set of five digits is shown on the phone screen

Memory Scanning

A set of five digits is shown on the phone screen

Single digits appear. The user presses YES or NO as quickly as possible

Speed and accuracy of responses are recorded

Memory Scanning: Device Level Issues

- Randomisation of target sets
- Randomisation of stimuli in blocks
- Correct recording and classification of responses
- Response timing

Device Level Validation: Timing

- Video recording of screen display allows measurement of time taken for image to be fully displayed
- Transducer to give external record of response timing for button press
- Tap of screen button can be timed directly from video

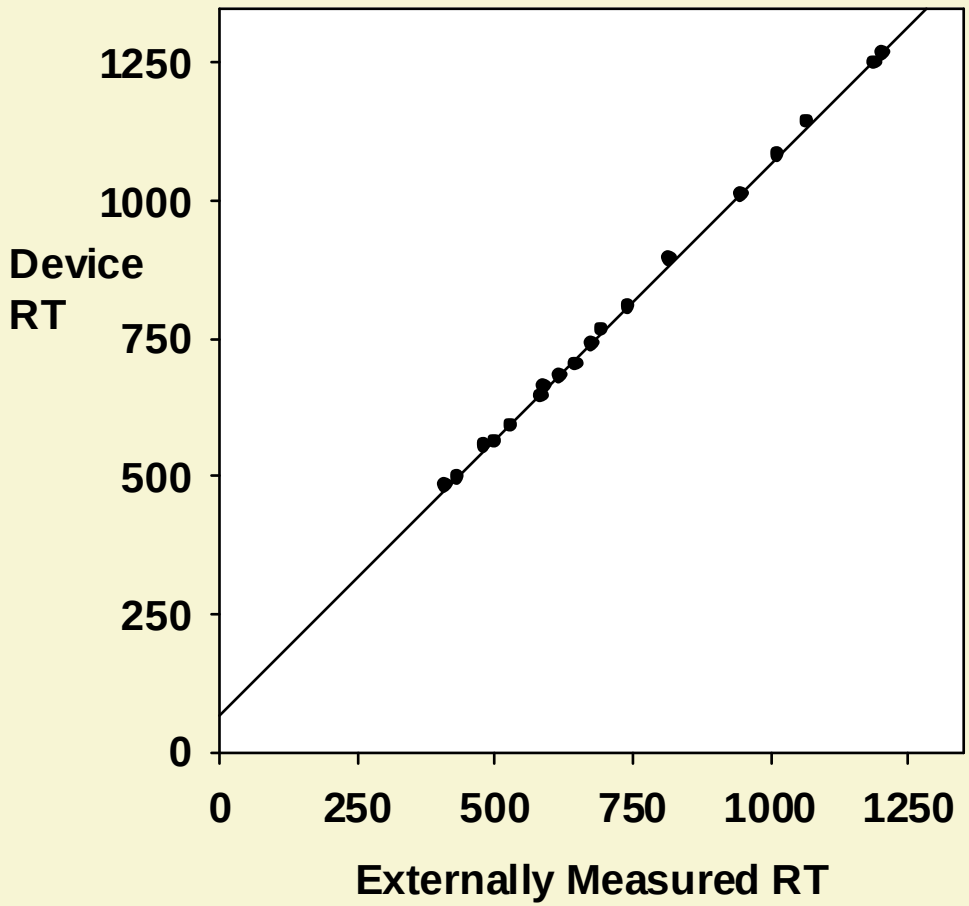
Device Level Validation: Timing

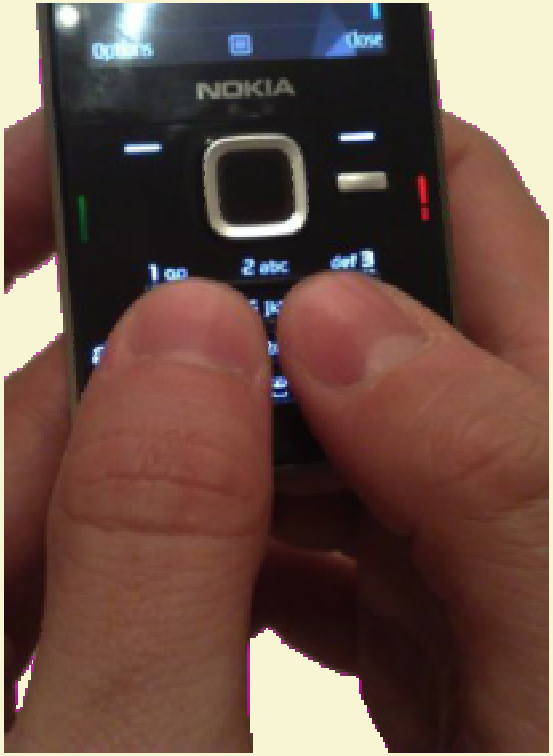
Successive screen shots from mobile phone display at 210 fps:

← 9.5 msec →

Is this compatible with adequate timing accuracy for summary measures (usually means of 20+ responses in a given category)?

Device Level Validation: Timing with key press (msec)





Device Level Validation: Timing with key press

- Correlations in all cases ≥ 0.9995
  - Less than 0.1% of error variance due to timing inaccuracy in device
- Device RT consistently greater than external recording
  - Probably due to delay in appearance of stimulus.
  - Not relevant when differences in scores are used, so long as delay is consistent
  - Would need to be allowed for in norm generation

Device Level Validation: Timing with stylus tap

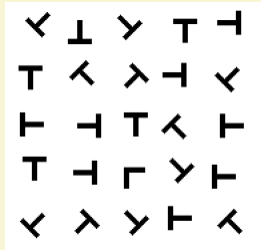
- Correlations lower, but still ≥ 0.995
- Time for stylus to leave screen not timed accurately
  - Makes measurements of response latency impossible
  - Needs investigation whether this is a problem with all screen types.

Intrinsic Issues

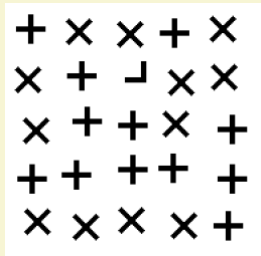
- Visual Search
  - Response time increases as number of distractors increases, for targets which do not “pop out”
  - Used in Newton validation, not yet used for current PDAs
- Memory scanning
  - Responses to digits in the memory set are shorter, typically about 100 msec less, than for digits not in the set.
- Establishment of these differences allows comparison of mobile phones to other systems, e.g. PC
- Differences also provide an indication of test integrity in an unsupervised setting

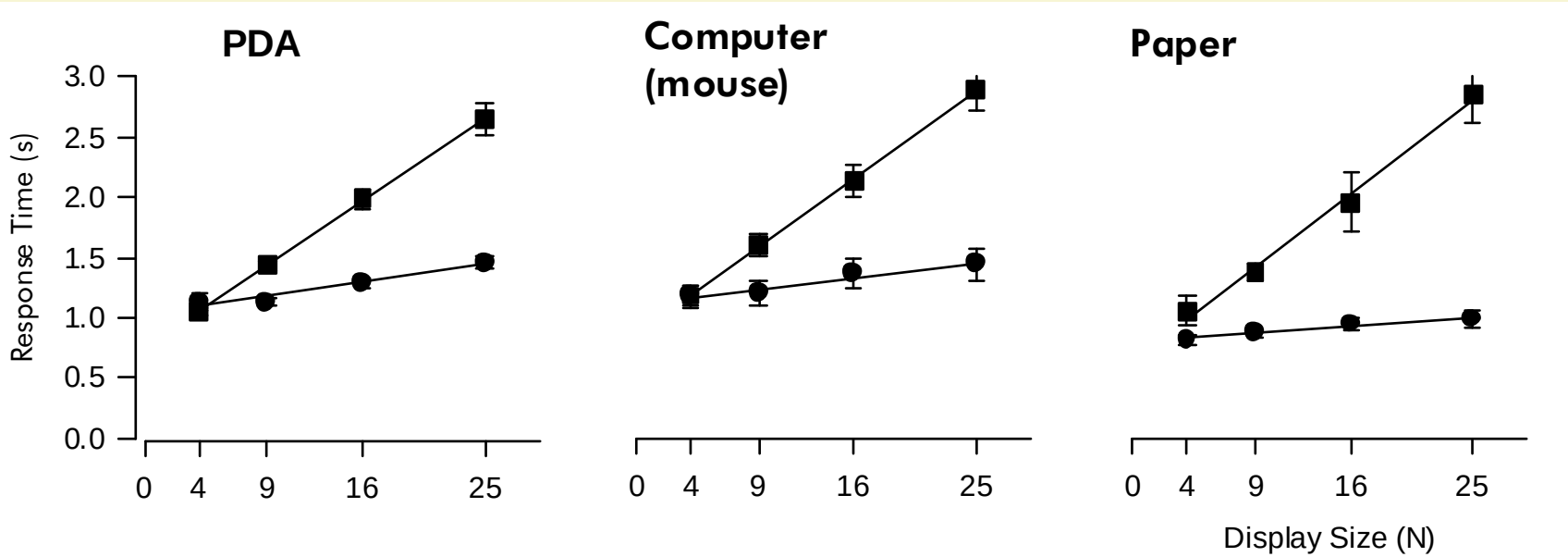
Visual Search Intrinsic validation

No Pop-Out



Pop-Out

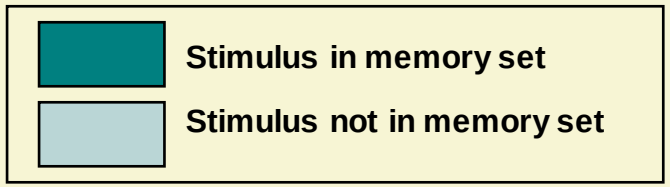




Everyday Life (“Free-Range”) Study

- 38 healthy volunteers (20 male) aged 18-54 years (mean 22.8) took part.
- Text (SMS) messages were sent twice a day at different times to the phones over 14 days.
- Volunteers completed assessments as soon as possible after receiving each text.
- Assessments included recording alcohol consumption, cognitive tests, and questions about disturbance during completion

Memory Scanning: Intrinsic Validation



Assessments included if no alcohol consumption reported in last 24h

Memory Scanning: Intrinsic Validation

Stability of differences in response times to different stimulus types indicates:

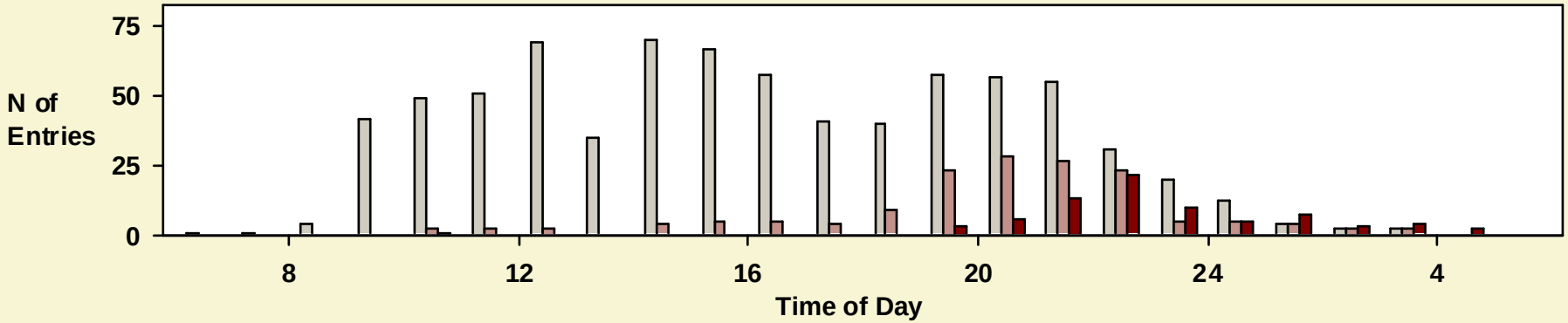
- Response time measurements are detecting differences as expected
- Task characteristics as expected on small-screen system.
- Volunteers are performing task as intended in unsupervised setting

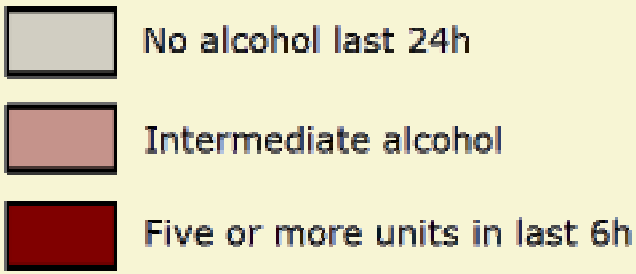
Memory Scanning: Extrinsic Issues

Ability of the mobile phone system to detect effects of factors known to affect (usually impair) cognition

- Fatigue or sleep deprivation
- Drugs, e.g. alcohol, benzodiazepines
- Disease states

Distribution of Free-range Entries





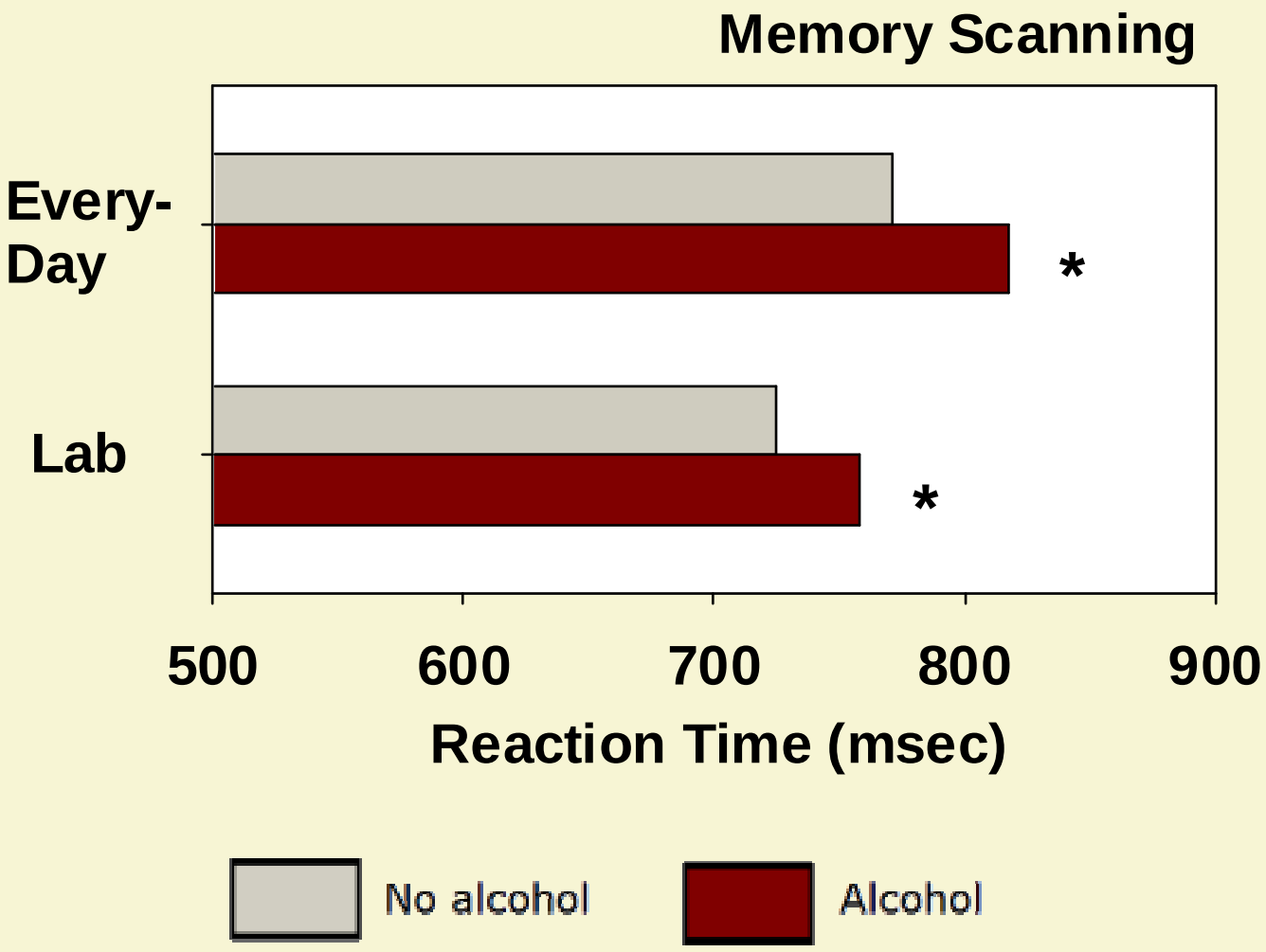
Entries with at least 5 units:

- 30/38 volunteers had at least one entry ≥ 5
- Maximum was 20 units (median 7)
- Previous work suggests that 7 units (reported) corresponds to a BAC of about 95 mg/100 ml.

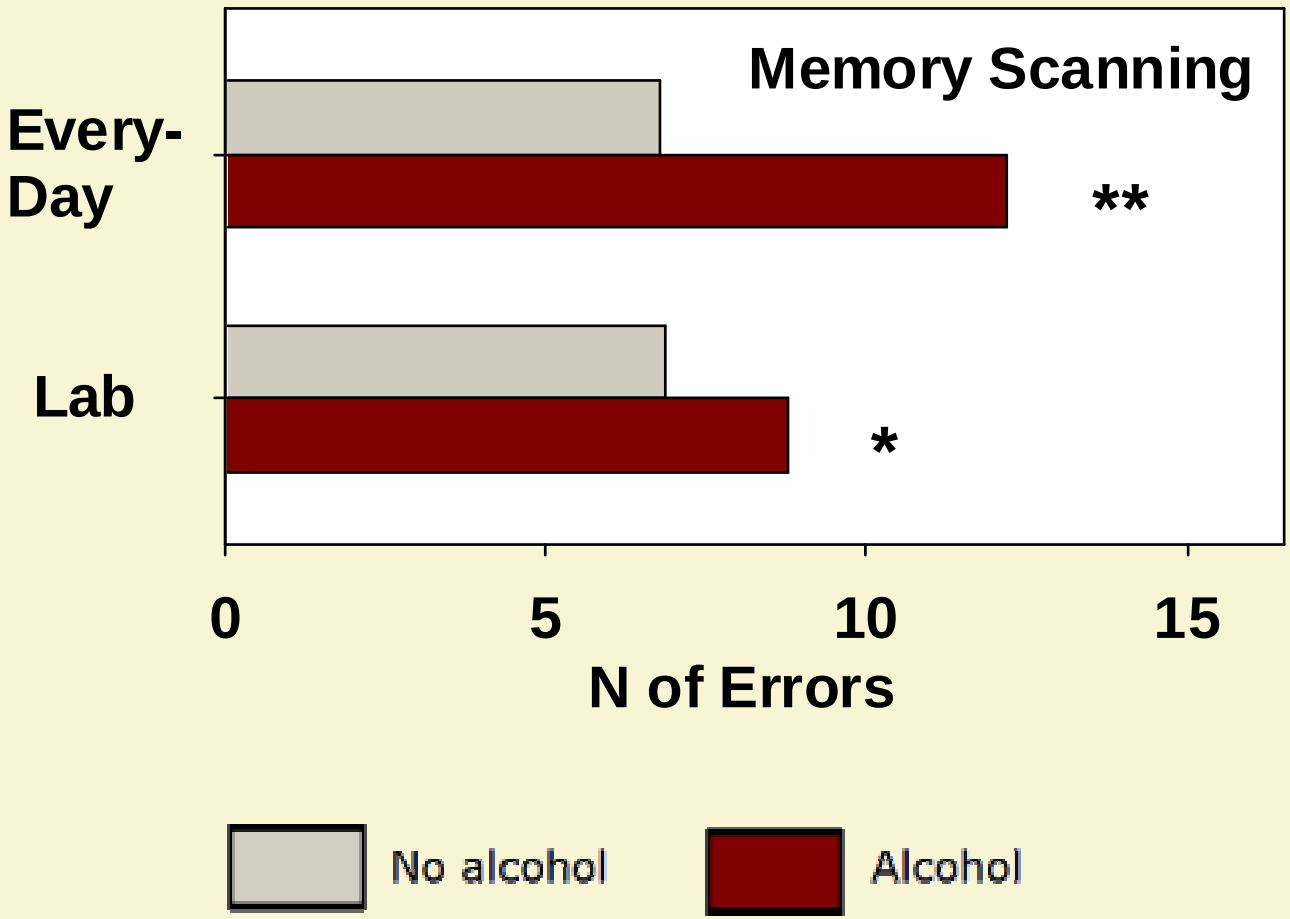
Laboratory (Battery)

- 26 of the volunteers took part in the lab study.
  - They received ethanol and placebo on separate days in random order
  - They completed the same assessments at intervals up to 2h after the drink.
- Mean blood alcohol concentrations were 124 mg/100 ml

Working Memory - Speed

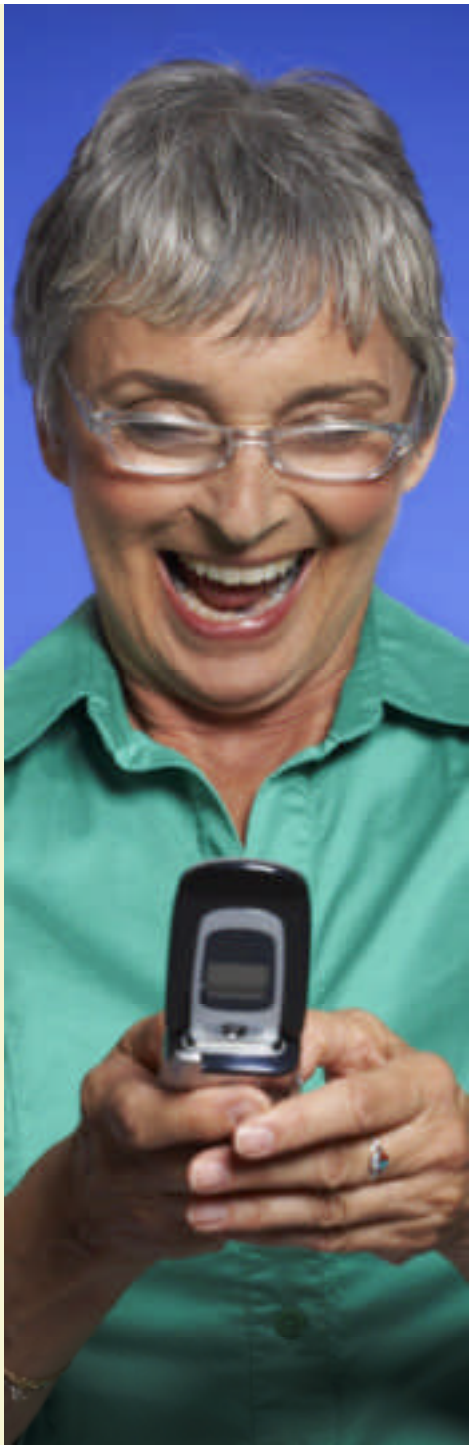


Working Memory - Accuracy



Conclusions

- All three levels of the model give important information for assessing system validity.
  - Device level validation shows consistent bias in RT scores, but excellent agreement otherwise
  - Intrinsic validation shows expected patterns with stimuli of greater difficulty taking longer.
  - Test battery reliably shows effects of extrinsic factors such as alcohol impairment, both in everyday setting and laboratory
- These data support the validity of using mobile phones to collect cognitive performance data in both supervised and unsupervised settings



Acknowledgements

Gordon Drummond,
Federica Amati,
Catriona Dunn

Clinical Research Facility,
Royal Infirmary of Edinburgh

All our volunteers

Contact: Brian Tiplady
brian@penscreen.com
www.penscreen.com