Mobile Monitoring of Cognitive Change
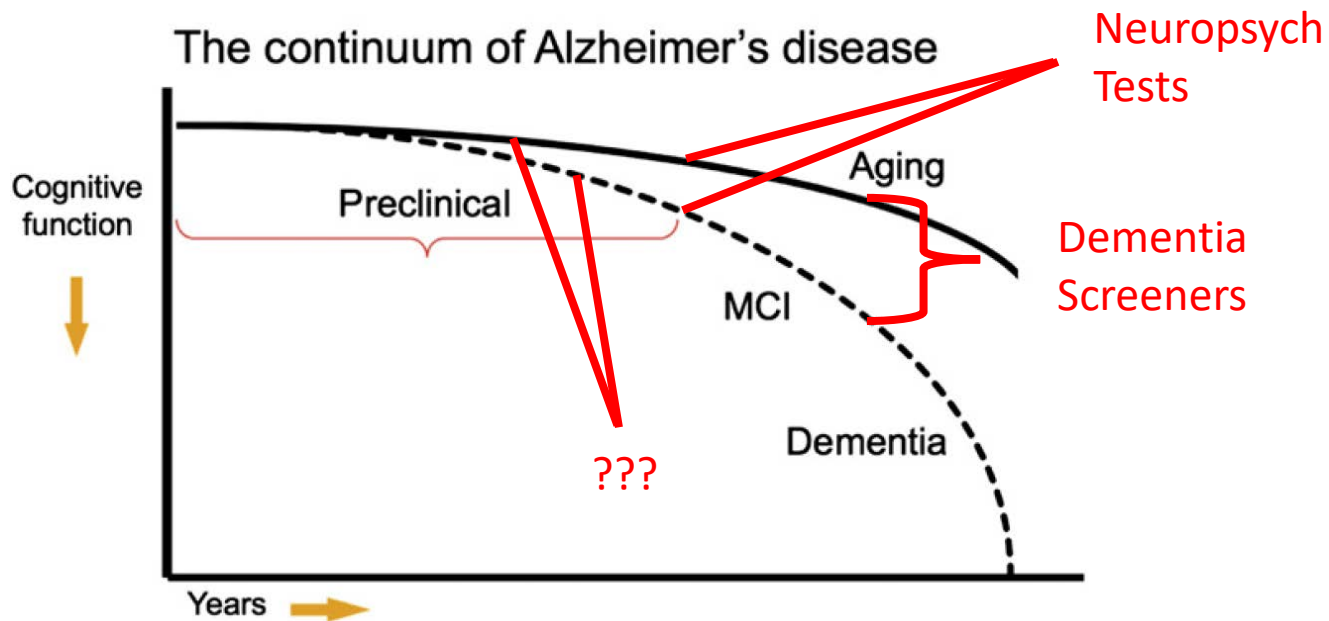
Jonathan W. King

Division of Behavioral and Social Research
National Institute on Aging

The continuum of Alzheimer's disease

Neuropsych Tests

Dementia Screeners

Cognitive function

Preclinical

Aging

MCI

Dementia

???

Adapted from Sperling et al., 2011

Years

Dementia screeners can detect moderate or more severe impairment. More sensitive neurospsych tests can help assess milder impairment. But what can we do to detect earlier differences and changes?

What more is needed to measure age-related changes in cognition?

- Instruments need to be reliable and sensitive to changes in level (slope) **and** acceleration of change.

- Instruments must be valid measures of constructs, that, to help us understand the reasons for change.

- We need to measure a wider variety of cognitive constructs (traditional ones may not be sensitive to age-related change or disease).

- Instruments need to be less burdensome on participants, and more cost-effective for funders.
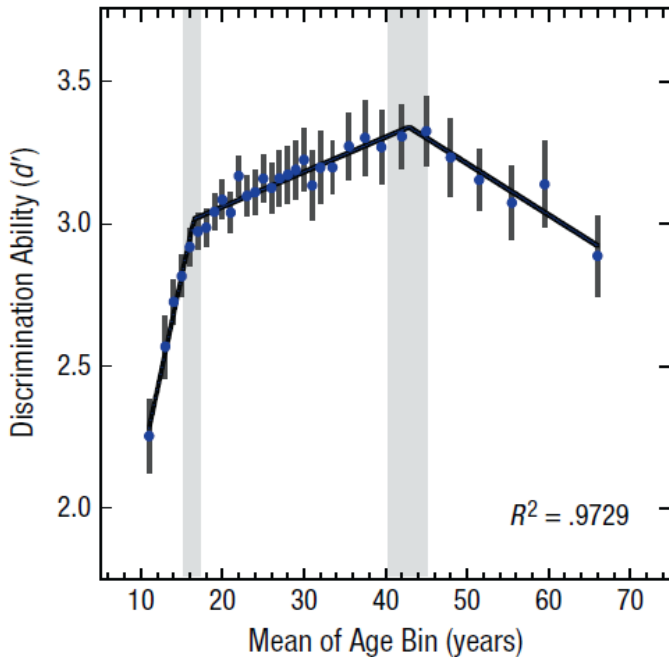
Cross-sectional Measures of Cognition

Age, in years
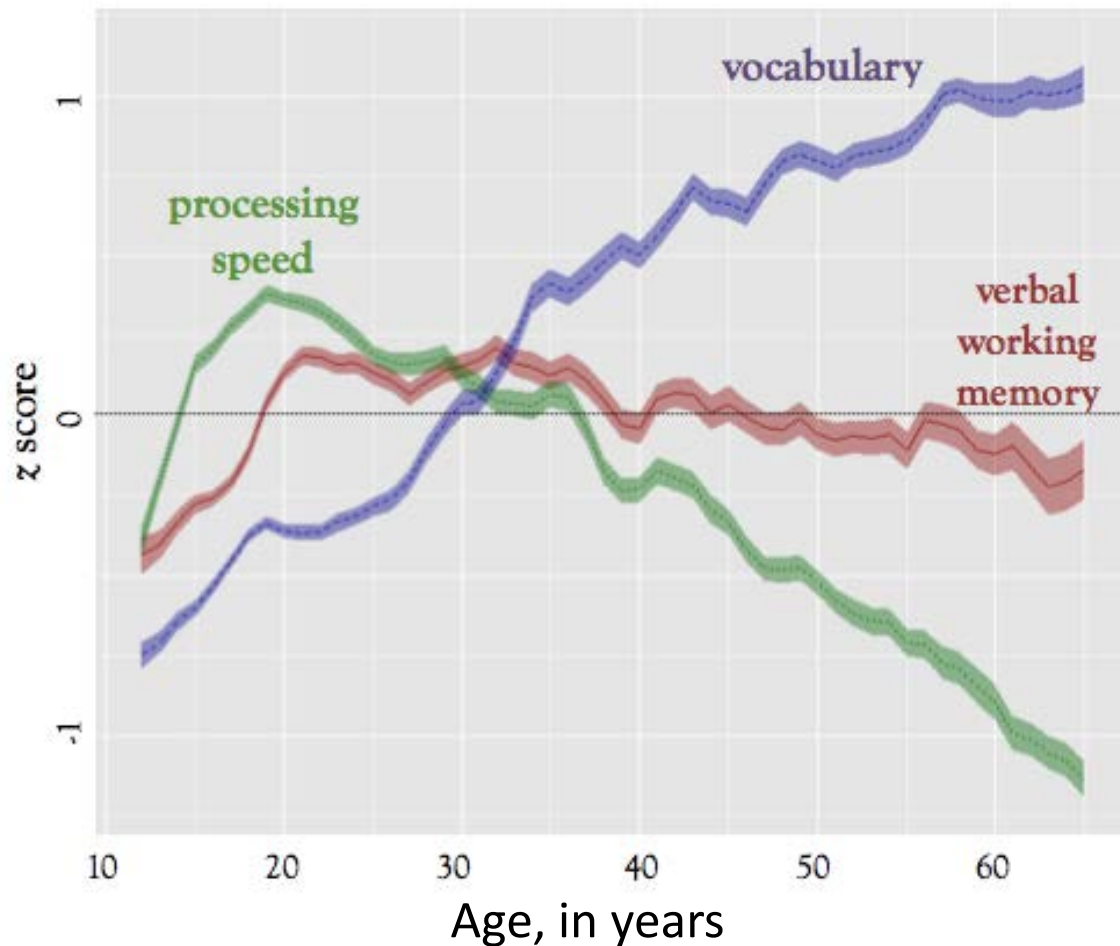
Data from N=10,384 people visiting the **website** testmybrain.org over the course of one year. (Hartshorne & Germine, 2015)

THE MANY BRAINS PROJECT
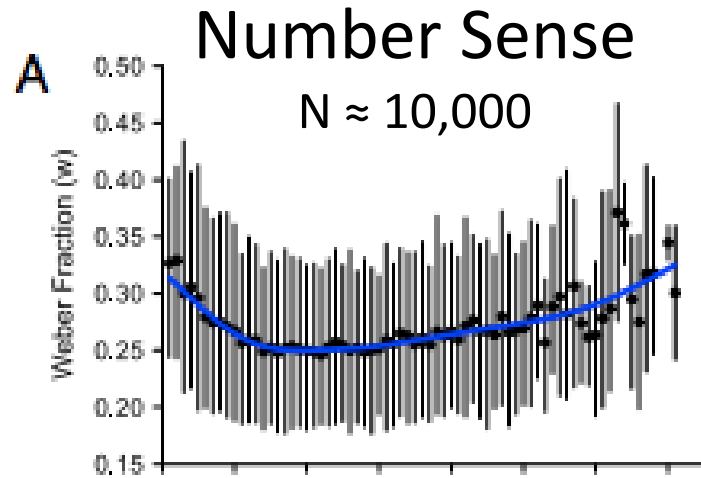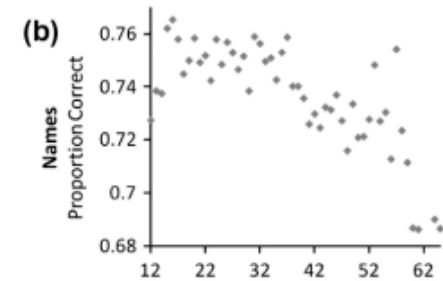
Sustained Attention
N ≈ 10,000

Fortenbaugh et al., (2015)

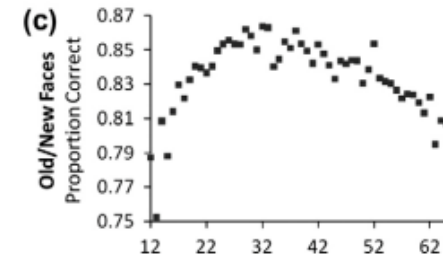## Number Sense
N ≈ 10,000

A

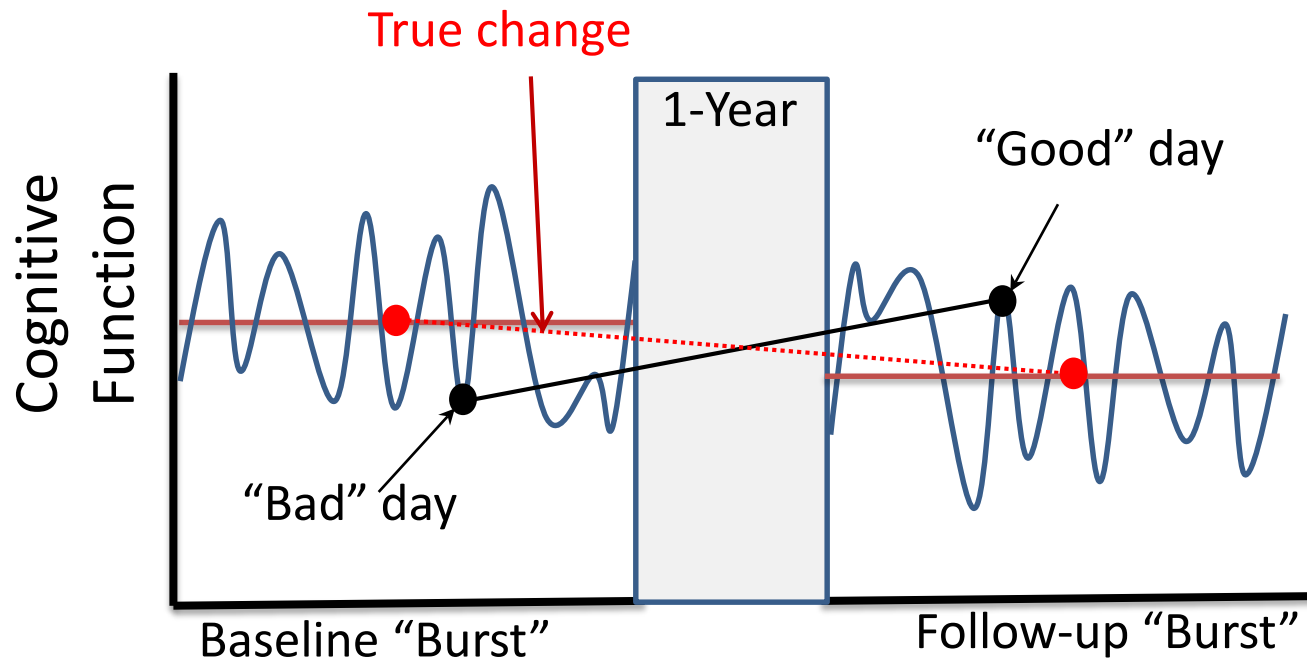Halberda et al., (2012 *PNAS*

Name Recognition

Face Recognition

Germine et al., (2011; N ≈ 65,000)

Burst Measurement Design

Burst measurement designs give you estimates of mean level and variability as well as better measures of true change over time. (Sliwinski, 2015)

This is 2016; let's just phone it in

iPhone 6s   Galaxy S6

Make an App for that, in iOS and Android (market share: 95+%)

- Two smartphone platforms are stable and very popular.
- Distribution model (free online) likely will increase uptake.
- People are willing to spend a **lot** of time on their devices
- Programming and database issues well understood.
- Smartphones give access to many sensor types.
- This is now becoming the most prevalent computing platform.

App Measures are Reliable & Valid
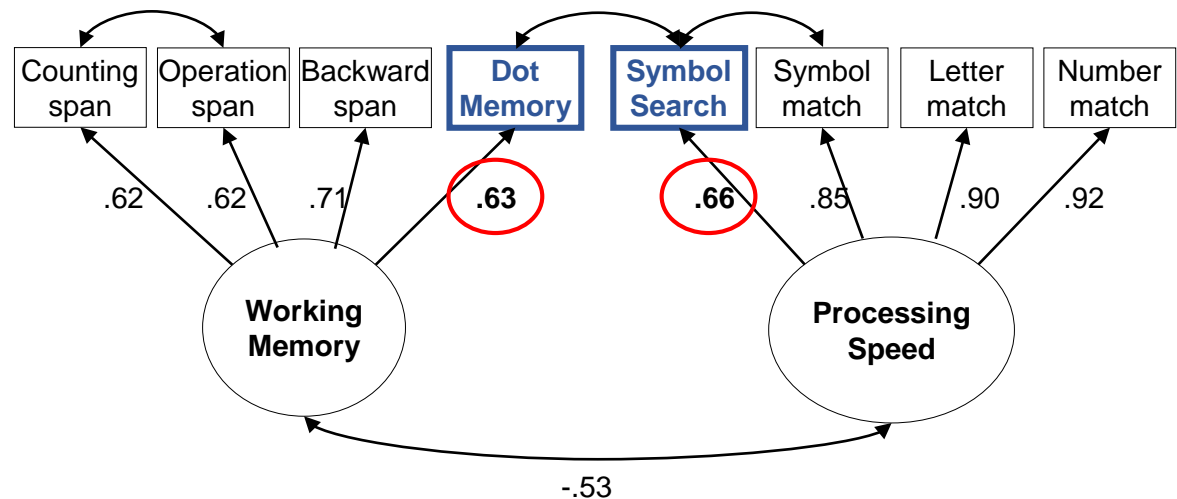
The ESCAPE Project showed that 14-day assessment bursts yield reliable (ICC > 0.95) measures of the underlying constructs measured in the lab. (Scott et al., 2015; Sliwinski, 2015)

Proposed Approach

- Support the development of iOS and Android apps that can be used to administer cognitive tests.

- Develop back-end database facilities that will allow the tracking of individuals over time, the selection of tests and items to administer or prompt, and the ability to provide selected feedback to participants.

- Validate initial tests against the NIH Toolbox® as well as legacy measures and tests in the UDS.

- Support the development and incremental validation of additional cognitive and non-cognitive measures from the scientific community.

Extra Slides (if questions arise)

NIH Efforts to Cooperate With

- OBSSR-led "Mobilizing Research" U2C Effort awarded to UCSF to fund "a reliable infrastructure to conduct mHealth research" that itself does not, however, support the generation or testing of apps.

- PMI Cohort Program Participant Technologies Center (U24) will generate infrastructure for use with **PMI participants**, focusing on **passive** sensor technologies; note that "The PTC is not responsible for generating the content of the surveys and other data collection procedures or protocols."

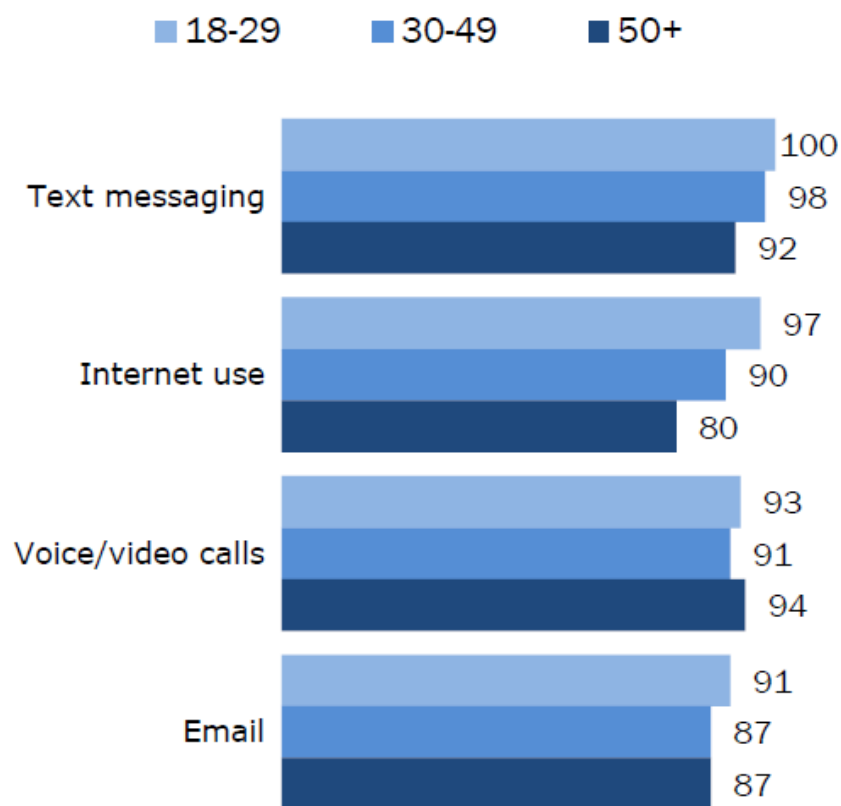# MMSE Measures Cognition Poorly

General cognitive performance
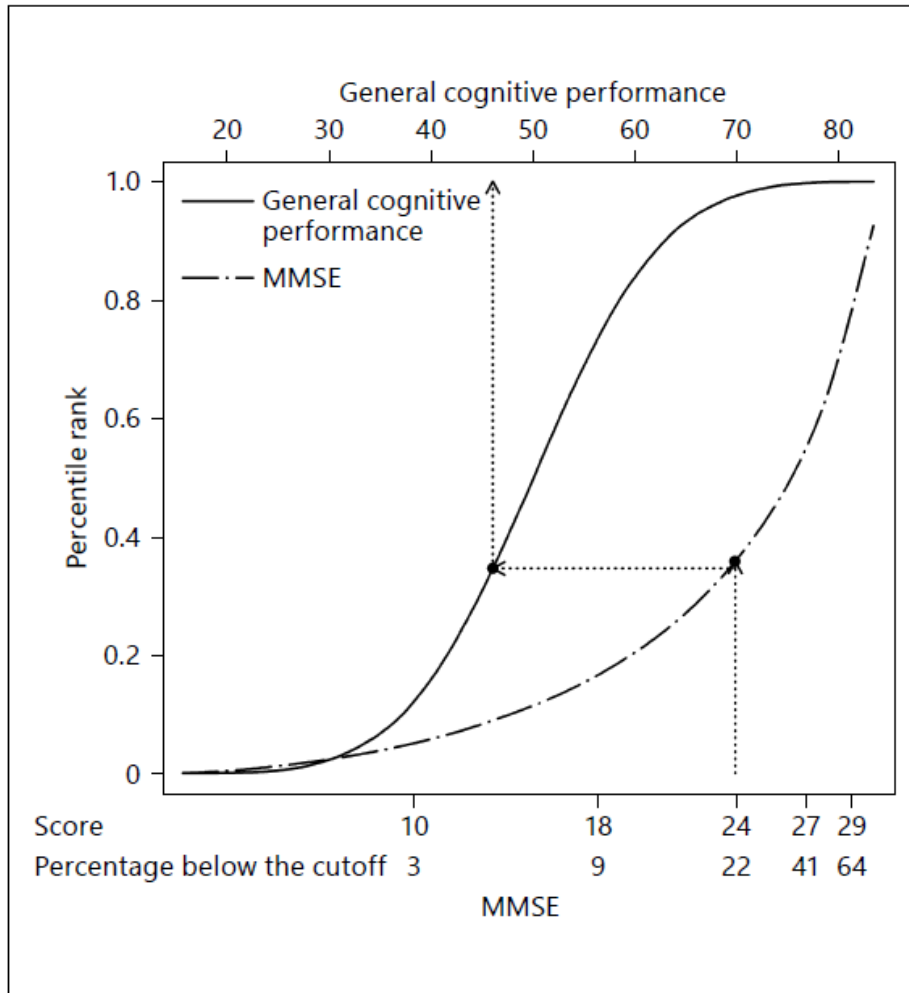
Gross et al. (2014) used IRT methods to derive a General Cognitive Performance scale from ADAMS data calibrated to have a mean of 50 and an SD of 10 in that sample.

The crosswalk to MMSE shows the insensitivity of the latter to higher levels of function.
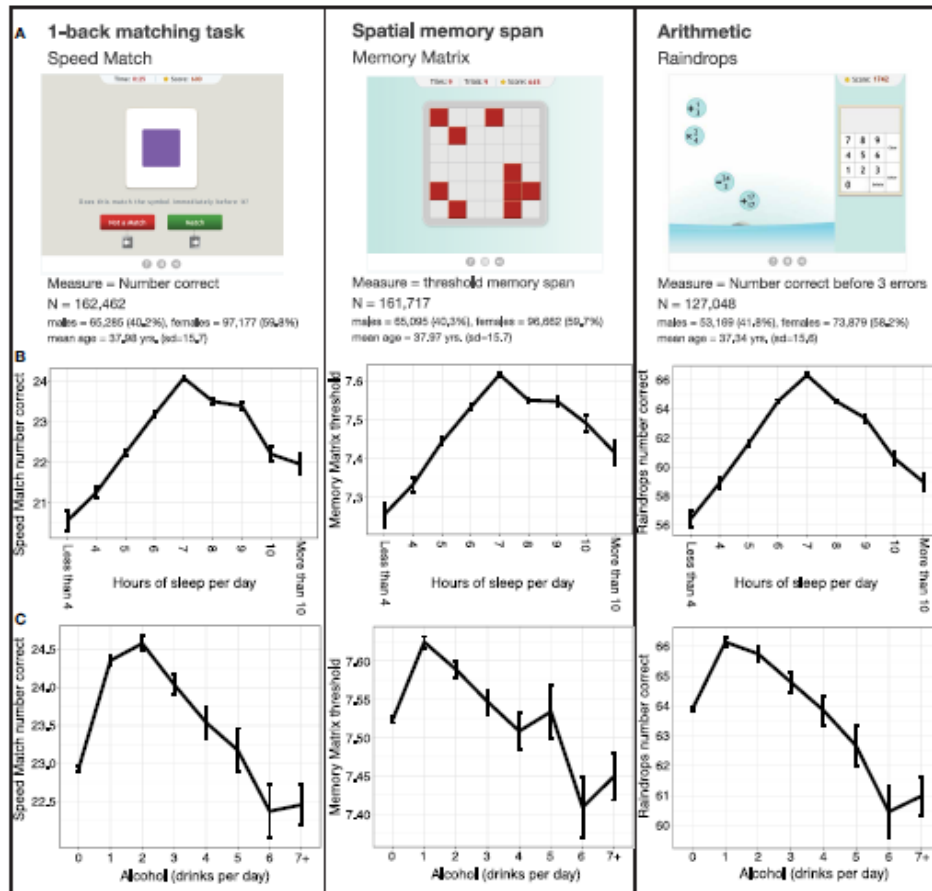
Smartphones are nearly universal

% of U.S. adults in each group who own a smartphone

All adults	64%
Male	66
Female	63
18-29	85
30-49	79
50-64	54
65+	27
White, non-Hispanic	61
Black, non-Hispanic	70
Hispanic	71
HS grad or less	52
Some college	69
College+	78

Legend: ■ 18-29 ■ 30-49 ■ 50+

	18-29	30-49	50+
Text messaging	100	98	92
Internet use	97	90	80
Voice/video calls	93	91	94
Email	91	87	87

Pew Research Center (2015)

Lumosity.com: Big Money…

Sternberg et al. (2013) showed cognitive
effects of sleep duration and drinking

Project Talent:
Big Data before it was Cool

Project TALENT collected cognitive and other data from over 400,000 high school students in 1960.

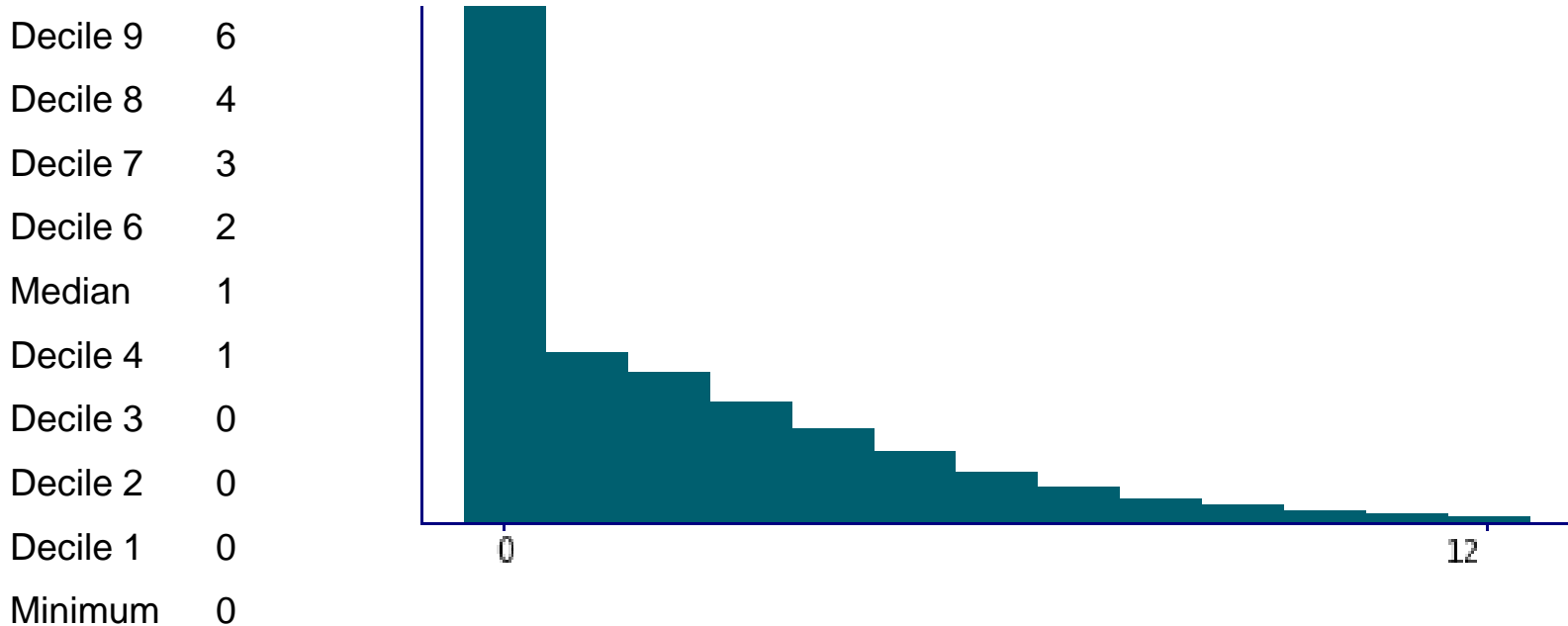Total test battery took 2.5 school days to administer.

Current proposals to re-test Project TALENT members will require ~1 hour of adaptive testing online.

Ultra-brief Batteries are Informative

UK BioBank Fluid Intelligence includes 13 items including a synonym item, an antonym item, number series items, and other brief measures. An IRT version would have had even greater sensitivity. (N = 165,645)
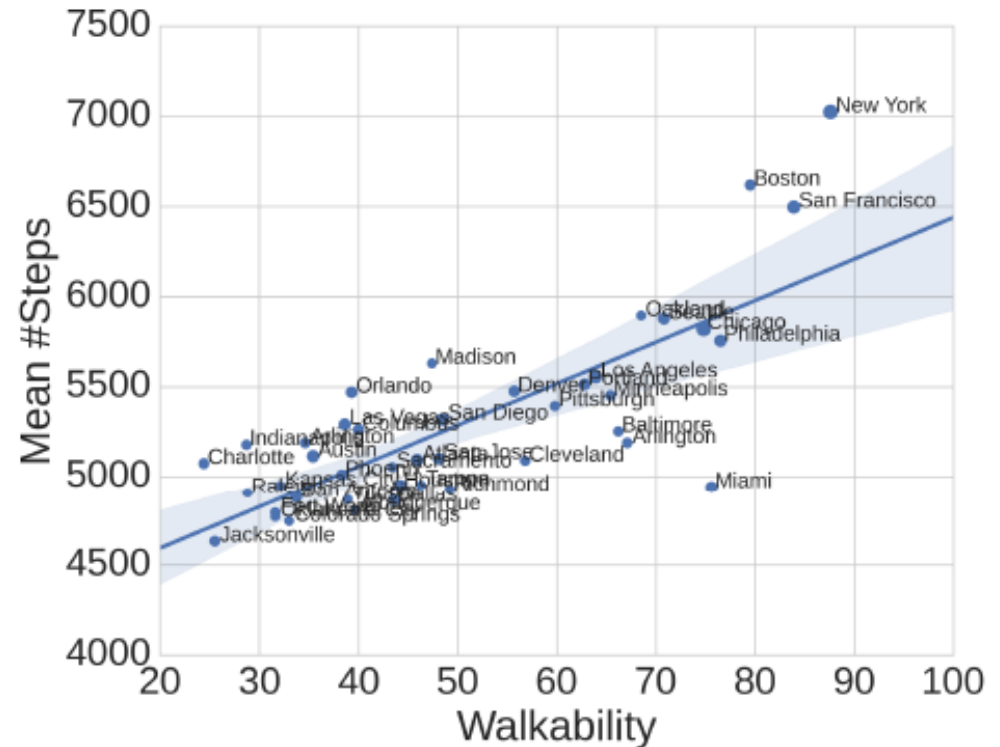
Why Incremental Validation is Key

Decile 9	6
Decile 8	4
Decile 7	3
Decile 6	2
Median	1
Decile 4	1
Decile 3	0
Decile 2	0
Decile 1	0
Minimum	0

UK Biobank "Pairs" memory data was inadequately tested; over 30% of participants received perfect scores, limiting explainable variation. (N=498,709)

Argus App   Walkability & Steps

Althoff et al. (2015; Stanford BD2K Mobilize Center) analyzed Argus app data on ~900,000 people in the US and collected step data on **~30 million person days.**