An Introduction To Neuropsychological Testing

Overview

- Introduction to neuropsychological testing
- Complications & problems
- History of neuropsychological testing
- Example 1: Bender Visual-Gestalt Test
- Example 2: The Wisconsin Card Sort Test
- Example 3: The Chicago Word Fluency Test
- Example 4: The Wechsler Memory Scale (Revised)
- Example 5: Rey (1941)-Osterrieth (1944) Complex Figure Test
- Conclusion

What is neuropsychological testing?

- Neuropsychological testing looks at two aspects:
  - i.) Functional integrity: Whether or not any particular specifiable function is intact
    - Examples: short-term/long-term memory, lexical access, attention, sensory discrimination, motor strength
  - ii.) Localization: Whether or not any specific neuroanatomical region of the brain is functionally intact

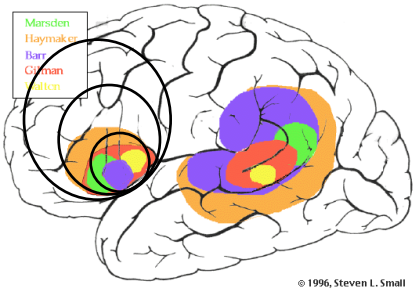
Some complications

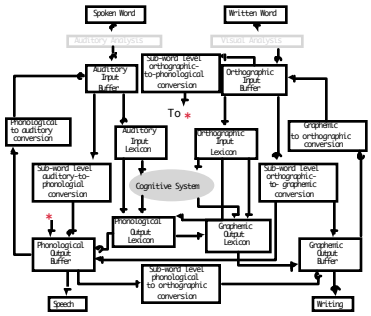
- Function and region do not have a one-to-one mapping
  - Many functions can be affected by lesions at many multiple disparate locations
  - Many brain regions subserve multiple functions
- The brain's functions do not map cleanly onto easily-definable functional categories
  - Neither attention, nor lexical access, nor memory (etc.) are really unitary functions: each can be decomposed into many (sometimes non-intuitive) subfunctions

Some complications

- Functional simplifications and partial relations between function and region- and the relations between these two- can be reified
  - Partial correlations start masquerading as certain facts
  - confirming evidence is piled up without weighting disconfirming evidence, making things seem more certain than they are under close scrutiny
- **Statistically-significant group differences do not guarantee interpretability of individual differences**
  - High overlap = low probability of meaningful interpretation of individual scores

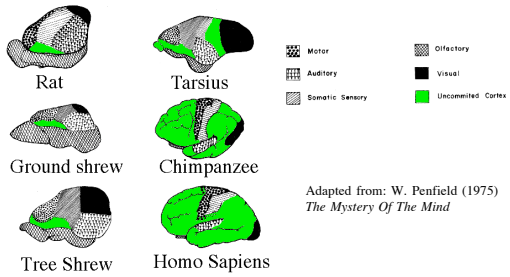
Example: Broca's & Wernicke's areas





A model of single word processing: 14 'nodes'
Redrawn from Howard & Franklin (1988)
(after Morton, 1980)

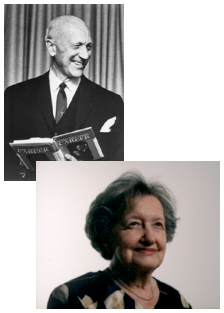
Humans have relatively huge amounts of association cortex



Adapted from: W. Penfield (1975)
The Mystery Of The Mind

History of neuropsych testing

- Neuropsychology began at the MNI in the 1950s, where Wilder Penfield was doing epilepsy surgery
- Brenda Milner did much of the early neuropsychological work



History of neuropsych testing

- Small amounts of information were very valuable in a situation where there was almost no hard data
- Many tests were developed from experimental work with minimal attention to psychometric rigor
- However, EEG, PET, and MRI have all changed the equation in modern times, leaving much modern neuropsychological testing 'orphaned', especially with respect to localization

Functional assessment

- When neuropsychological tests are used (as they often are) for purely functional assessment, they can escape from the constraining demands of validity simply by having face validity (or even just historical precedent)
  - A standard battery can have utility simply be virtue of being standard, and/or by allowing for pre-post testing
  - When inferences are to be made to prior functioning, more psychometric rigor is required, but not always available

The 10 most commonly used tests

1.) Wechsler Intelligence Scale for Children (WISC)
2.) Bender Visual-Motor Gestalt Test
3.) Wechsler Adult Intelligence Scale (WAIS)
4.) Minnesota Multiphasic Personality Inventory (MMPI)
5.) Rorschach Ink Blot Test
6.) Thematic Apperception Test (TAT)
7.) Sentence Completion
8.) Goodenough Draw-A-Person Test
9.) House-Tree-Person Test
10.) Stanford- Binet Intelligence Scale

From Brown & McGuire, 1976

Example 1: Bender Visual-Gestalt Test

- The Bender Visual Motor Gestalt Test (1946) is a widely-used test to assess visual motor processing.
- It is often referred to as the Bender Gestalt.
  - 'Bender' is the person who designed it.
  - 'Gestalt' comes from a German word meaning 'form'.
- The test simply asks you to copy a set of abstract designs (we will see them in the next lab)

Example 1: Bender Visual-Gestalt Test

- The Bender Gestalt is sensitive at identifying organic brain damage, distinguishing it from purely psychiatric diagnoses.
  - Visuographic productive abilities are associated with the parietal lobe, especially in the right hemisphere.
  - A good result cannot rule out brain damage in other regions of the brain.
  - This test is also sometimes used for assessing mental retardation and regression in the psychoanalytic sense (functioning beneath ones actual developmental level).

Example 1: Bender Visual-Gestalt Test

- The original scoring was very unspecified, requiring an expert qualitative judgment
- Many objective scoring systems have since been developed
- Some have inter-judge reliabilities above 0.95.
- Bender scores correlate around 0.5 with all WAIS subtests except Digit Span and Object Assembly, with which they correlate a little lower, around 0.4.

Example 1: Bender Visual-Gestalt Test

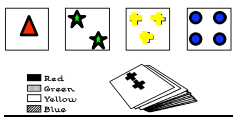
- Using one system, 59% of brain-damaged subjects, but only 8% of non-brain-damaged (normal and psychiatric) subjects, score above the cut-off. What is the chance that a person has brain-damage [P(D)], given that they score above the cut-off [P(S)]? Assume that 5% of patients are brain-damaged.

Example 1: Bender Visual-Gestalt Test

- Using one system, 59% of brain-damaged subjects, but only 8% of non-brain-damaged (normal and psychiatric) subjects, score above the cut-off. What is the chance that a person has brain-damage [P(D)], given that they score above the cut-off [P(S)]? Assume that 5% of patients are brain-damaged.
- $P(D|S) = P(S \& D)/P(S) = P(S|D)P(D)/P(S)$
- $P(S|D) = 0.59$
- $P(D) = 0.05$
- $P(S) = (0.05 \times 0.59) + (0.08 \times 0.95) = 0.0295 + 0.076 = 0.1055$
- $P(D|S) = (0.59 \times 0.05) / 0.1055 = 0.279$, or 28%

Example 2: Wisconsin Card Sorting Task

The Wisconsin Card Sorting Test (WCST)



Aim: To test an individuals ability to repeatedly form, maintain, and switch categories (color, number, form).

Procedure: Based on E's feedback, S must sort cards according to the correct category which switches after 10 appropriate responses.

Score for:
- Cards Used (max. 128)
- ✿ Categories Attained (max. 6)
- ✓ Perseverative Errors
- Non-Perseverative Errors
- "Correct" Responses
- Unique Errors

Wisconsin Slides are courtesy of Aki Caramanos

Effects of Different Brain Lesions on Card Sorting (Milner, 1963)

Pre-Op Findings

Subjects:
18 dorsolateral frontals (DLF)
53 non-DLF controls:
33 temporals,
8 parietals,
5 parieto-occipitals, and
7 orbito-fronto-temporals.

	Cats	Pers.	Non-Pers.
DLF (18)	3.3	39.5	15.4
Controls (53)	4.6	20	17.7

Categories: no sig. diff.
Perseveration Errors: p < 0.01, DLF made more errors but considerable overlap between the groups
Non-Persverative Error: no sig. diff.

Post -Op Findings

Group 1: same patients 2 weeks post-op
Group 2: additional 23 patients tested 2-3 wk. (n=10) or 1-yr. (n=14) post-op

Group 1	Cats	Pers.	Non-Pers.
DLF (18)	1.4	51.5	21.7
Controls (53)	4.7	12.8	17.8

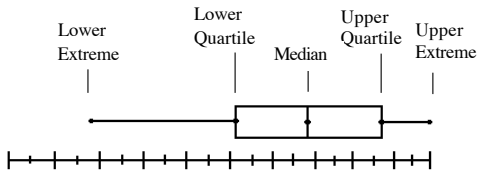
Group 2	Cats	Pers.	Non-Pers.
DLF (7)	1.3	68.1	10.1
Controls (16)	4.6	26	12.3

Cats: p < 0.001, DLF fewer cats
Pers: p < 0.001, DLF more pers. errs
N-Pers: no sig. diff.

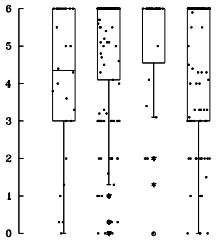
Groups 1 & 2	_ 3 cats	> 3 cats.
DLF (25)	100%	0%
Controls (69)	26%	74%

$\chi^2 = 40.4$, $p < 0.001$

- Milner (1971): "the critical lesion is dorsolateral , not orbital, and a small left frontal excision involving this area causes lasting impairment, whereas larger lesions on the right sometimes produce only transient defects, or even none at all"
- Milner (1975): "even in cases where early severe injury to the LH has caused speech to be mediated by the RH, frontal removals for the non-speaking LH cause lasting deficits in WCST which are not seen after removals from a comparably damaged RH."
- Based mainly on the strength of these, and other similar findings the WCST is widely accepted as an indicator of frontal lobe function, left more than right.

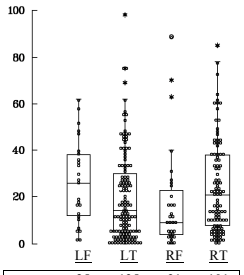


How to read a 'box and whisker' plot



	LF	LT	RF	RT
n:	26	137	31	102
mean:	4.0	4.9	5.0	4.6
st. dev.:	2.1	1.8	1.8	1.9

Categories attained: LF < LT, RF ($p < 0.04$)



	LF	LT	RF	RT
n:	25	135	31	101
mean:	26.4	20.1	17.8	24.3
st. dev.:	18.0	19.0	21.4	19.6

Perseverative errors: LF > RF ($p < 0.1$)

Categories

$\chi^2 = 2.1$, n.s.	LF	LT	RF	RT
impaired (< 3)	8	30	5	26
normal (> 3)	18	107	26	76

Perseverative Errors

$\chi^2 = 5.5$, n.s.	LF	LT	RF	RT
impaired (> 19)	13	56	9	51
normal (< 19)	12	79	22	50

By chi-square, WCST cut-off scores are not useful in pre-operative discrimination of focal epilepsy patients.

Discriminant analyses (by side, location, or location & side) were not successful either

Summary of findings

- While the LF group was statistically impaired on some of the WCST measures relative to the other groups of patients tested, there was almost complete between-group overlap on all measures at all stages of testing = **classification of individual patients based on any one measure is impossible.**
- Individual patient 's pre-operative pattern of performance across the WCST variables could not predict their locus of neural disturbance.
- neither early post-operative, nor late follow-up performance could predict site of cortical excision.
- The WCST may be an adequate measure of an individual 's ability to repeatedly form, maintain, and switch categories, but it is not an effective tool for localising neural dysfunction

Example 3: Chicago Word Fluency Test

- The Chicago Word Fluency Test is used to measure an individual 's symbolic verbal fluency.
- Subjects are required to write as many different words beginning with S as possible in 5 minutes and, after this, as many singular four-letter words beginning with C as possible in 4 minutes.
- The total number of 'S' and 'C words produced, minus the number of rule-breaking and perseverative responses, yield the patients ' measure of verbal fluency.
- Spelling mistakes and socially inappropriate words are noted, but not subtracted from this measure.

Chicago Word Fluency Test: History

- In 1964, Milner found that patients that had undergone discrete cortical excision from the left prefrontal cortex (LF, n=7) for the treatment of focal epilepsy were severely impaired on this task relative to similar patients with excisions from the right-frontal (RF, n=4) or the left-temporal (LT, n=7) lobes.
- In 1974, Perret tested a variety of patients pre-operatively on an oral version of the CWFT and found that patients with frontal lesions performed worse than those with non-frontal lesions (n=68). Moreover, the LF patients (n=23) were more impaired than the RF patients (n=27).
- Based on these and other similar findings, the CWFT has been widely accepted as a measure of frontal lobe function.
- a recent survey of epilepsy centers found it to be the most widely used measure of verbal fluency

Chicago Word Fluency Test

Valid 'S' Words:

	LF	LT	RF	RT
n:	28	120	33	97
mean:	15.9	23.8	24.2	26.4
st. dev.	11.6	11.4	9.6	13.0

- main effect of side, L<R ($p<0.01$)
- main effect of lobe, F<T ($p<0.01$)
- no side*lobe interaction

Chicago Word Fluency Test

Valid 'C' Words:

	LF	LT	RF	RT
n:	28	120	33	97
mean:	5.3	10.2	8.5	10.6
st. dev.:	6.2	5.9	5.8	6.4

- no effect of side
- main effect of lobe, F<T ($p<0.001$)
- no side*lobe interaction

- No significant main effects or interactions of # of spelling errors, perseverations , rule breaks, or swear words

Discriminant Ability: Summary

- Frontal patients, as a group, produced statistically fewer words on the CWFT.
- Nevertheless, there was almost complete between-group overlap on all measures at all stages of testing
- Discriminant analyses not successful at predicting locus of excision in these focal epilepsy patients.
- Classification of individual patients based on any one measure was therefore impossible.

The Structure of Memory

- Memory is a complex construct composed on many differentiable subfunctions

Process	Duration	Associated Concept	Neuroanatomy	Deficit
Registration	Msecs	Awareness	Reticular Activating System	Stupor, coma
Short term memory	0.5-60 mins.	Working memory	Limbic System +	Low memory span
Consolidation	Seconds to years	Learning & Recent memory	Hippocampus +	Defective information storage and retrieval
Long-term storage	Seconds to life	Remote memory	Cortex	Lost skills or memories

Memory testing

- The WAIS is a starting point
  - Digit Span tests retention
  - Information tests remote memory
- Other common memory tests are:
  - The Wechsler Memory Scale (1945)
  - Rey- Osterrieth Complex Figure Recall
  - Corsi Blocks

Example 4: The Wechsler Memory Scale (Revised)

- Consists of 7 subtests:
  1.) Personal & current information: Age, date of birth, current head of state etc.
  2.) Orientation: Time and place
  3.) Mental control: Automatisms such as alphabet recitation; Conceptual tracking: "Count by 4 from 1 to 53"
  4.) Logical Memory: Immediate recall of two paragraphs

The Wechsler Memory Scale (Revised)

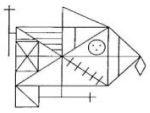
- Consists of 7 subtests:
  5.) Digit Span: Like the WAIS-R, but shorter: no 3-forward/2-back, or 9 forward/8-back
  6.) Visual Reproduction: An immediate visual memory drawing task
  7.) Associate learning: 10 words pairs; 6 easy associations (eg. baby-cries) and 4 hard associations (eg. cabbage-pen).
    - 3 presentations with test after each
    - Score = 0.5 easy + hard

The Wechsler Memory Scale (Revised)

- Issues:
  - MQ assumes memory is a unidimensional function
  - Has been criticized both for an overly-inclusive concept of memory (includes orientation, drawing competency, mental tracking) and for its limitations of functions tested (6/7 tests are verbal; the 7th- Visual recall- has verbal loading)
  - Subtest intercorrelations are low, so one cannot assume that intact subjects will perform well on all well enough to identify deviation
  - Positive correlations with tests of intellectual ability raise questions
  - Not well tuned for differential diagnostic purposes

Example 5: Rey (1941)-Osterrieth (1944) Complex Figure Test

- Investigates both perceptual organization & visual memory



- Copy, sometimes with different colored pens after elements
- Time to completion is recorded
- One or two tests or recall follow

Rey-Osterrieth Complex Figure Test

- Frontal lobe patients perseverate in copies
- LH damage patients tend to break drawing into smaller units than normals (less so at recall) and simplify (eg. by rounding angles such as those on the diamond; drawing dashes instead of each dot; turning the cross into a T)
- RH patients tend to make more omissions
- Parietal patients have difficulty with spatial organization

- Scoring systems exist
- Inter-rater R is very high

Corsi Blocks

- Non-verbal analogue to digit span
- Nine 1.4 inch cubes attached to a black background
- E taps each one in sequence, adding one after each successful copy by the patient
- One pattern is repeated ever third trial (as in Hebb's Digits)

- R temporal lobe damage shows little long-term learning and show deficits of short-term recall as well
- Other RH damage can also affect performance

Special factors in neuropsychological testing

- Normal age-related changes
- Handedness
- Sex
- Premorbid psychological status
- Medication
- Epilepsy
- Psychosis, perhaps secondary
- Malingering

Conclusions

- Tests may (and many do) discriminate groups without succeeding in discriminating individuals
- Functional localization claims are fraught with difficulty and can often be resolved with technological rather than inferential tools
- Functional claims may be made on the basis of tests being their own validation, since it is not always obvious what else could validate the test more appropriately