Poor recognition vs. recall: Indicator of performance invalidity or legitimate clinical presentation?

Simon Walker
Associate Professor Lynette Tippett
Dr Jamie MacNiven (ADHB).
Background

Neuropsychology

- Assessment, diagnosis, treatment, and rehabilitation of patients across the lifespan with neurological, medical, neurodevelopmental and psychiatric conditions.

- E.g. Stroke, traumatic brain injury, dementia.
  - Co-morbid issues such as anxiety, depression, drug and alcohol use etc.

- Patients/clients undergo interviews/neuropsychological testing with inferences drawn about underlying cognitive strengths or deficits.
Background

Performance validity testing

- Accurate assessment relies upon a patient/client performance that accurately reflects current functioning and ability.

- Research indicates that a considerable minority of cases (5-40%) may involve probable instances of performance invalidity (Mittenberg et al., 2002).
  - Prevalence of invalidity differs considerably cross clinical settings, likelihood of gain and so on (Kirkwood and Kirk, 2010).

- Implications for:
  - Clinical assessment.
  - Provision of rehabilitative resources.
  - Accurate diagnosis.
  - Resourcing.
  - Research findings.
  - Courts, justice, forensic psychology.
Background

Performance validity testing

• Currently viewed as a fundamental and necessary feature of neuropsychological assessment (Bigler, 2012; Larrabee, 2012).

• Contentious and challenging area of clinical work & research with considerable ethical implications.
  • Literature not well defined.
  • Literature is adversarial.
  • Extremely hard to validate objective measures.
  • Reliance upon inference and intuition.
  • Potentially complicated by third party interests (insurance litigation, ACC etc.).
  • Rapid and recent cultural shift within the field.
What does the term ‘invalidity’ capture?

Invalidity

Munchausen’s & Facetious Disorder
- Malingering.
- Financial gain.
- Forensic gain.

- Anxiety about sensitivity of tests.


Intentional.

Unintentional and perhaps normal/expected.
Background

How is invalidity currently determined?

1. **Objective tests (PVT’s/SVT’s)**
   - Test of Memory Malingering (Tombaugh, 1996).
   - Word Memory Test (Green, 2004).
Background

How is invalidity currently determined?

1. **Objective tests (PVT’s/SVT’s)**
   - Test of Memory Malingering (Tombaugh, 1996).
   - Word Memory Test (Green, 2004).

   • **Hard to validate**
     • Inference of invalidity based on litigiousness.
     • Asking non-clinical populations to feign feigning.
Background

How is invalidity currently determined?

1. **Objective tests (PVT’s/SVT’s)**
   - Test of Memory Malingering (Tombaugh, 1996).
   - Word Memory Test (Green, 2004).

   - **Hard to validate**
     - Inference of invalidity based on litigiousness
     - Asking non-clinical populations to feign feigning.

   - **Cut off scores are problematic.**
     - Validity is a continuum, not a dichotomy.

-Taken from Bigler, 2012.
Background

How is invalidity currently determined?

2. Subjective clinical appraisal or qualitative and quantitative information.
   • Test performance vs. injury.
   • Test performance vs. presentation.
   • Test performance between similar cognitive domains.
   • Contextual issues (gain etc.).
Background

How is invalidity currently determined?

3. Patterns of performance on neuropsychological tests that are considered to be ‘unusual’ or ‘atypical’.
   - Impaired recognition vs. recall performance on episodic memory tests (Bernard, 1990, Lezak et al., 2012).
   - Indicative of ‘intentional wrong responding’ (Beetar & Williams, 1995)

“Recognition should almost always be better than encoding and recall unless motivational factors are responsible”.
- Schoenberg & Scott (2011)
Doors and People Test
(Baddeley, Emslie & Nimmo-Smith)

Recall and recognition stimuli.
Visual and verbal domains of episodic memory.

• Normative data tables indicated that recall - recognition discrepancies were approximately normally distributed, suggesting poorer recognition than recall was no less common than the reverse pattern (1994).

• Normative data tables were updated with scores corresponding to poorer performance on recognition than on recall subtests removed (1996).

• No reason was given for this change.
Doors and People (1994) and amendment (1996)

Table 11 – Recall recognition discrepancy
Norms for recall-recognition discrepancy in scaled scores

<table>
<thead>
<tr>
<th>Recall-recognition discrepancy</th>
<th>Scaled recall-recognition scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>-14</td>
<td>2</td>
</tr>
<tr>
<td>-13</td>
<td>2</td>
</tr>
<tr>
<td>-12</td>
<td>3</td>
</tr>
<tr>
<td>-11</td>
<td>4</td>
</tr>
<tr>
<td>-10</td>
<td>4</td>
</tr>
<tr>
<td>-9</td>
<td>5</td>
</tr>
<tr>
<td>-8</td>
<td>5</td>
</tr>
<tr>
<td>-7</td>
<td>6</td>
</tr>
<tr>
<td>-6</td>
<td>6</td>
</tr>
<tr>
<td>-5</td>
<td>7</td>
</tr>
<tr>
<td>-4</td>
<td>8</td>
</tr>
<tr>
<td>-3</td>
<td>8</td>
</tr>
<tr>
<td>-2</td>
<td>9</td>
</tr>
<tr>
<td>-1</td>
<td>9</td>
</tr>
<tr>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>8</td>
<td>15</td>
</tr>
<tr>
<td>9</td>
<td>15</td>
</tr>
<tr>
<td>10</td>
<td>16</td>
</tr>
<tr>
<td>11</td>
<td>16</td>
</tr>
<tr>
<td>12</td>
<td>17</td>
</tr>
<tr>
<td>13</td>
<td>18</td>
</tr>
<tr>
<td>14</td>
<td>18</td>
</tr>
</tbody>
</table>

Scores in **bold** indicate the scores that were excluded in the 1996 amendment.

“It is not unusual for patients to perform better on recognition than recall. The reverse pattern is less common. If observed it should be further investigated, establishing first that the discrepancy is reliably obtained”.
Why is this belief widely held?

• Not actually sure....

• Recognition trials involve re-presenting stimuli to examinee’s at retrieval phase.

• Intuitively easier?
This project...

• Test the intuitive assumption of invalidity.
• Predicts that recall-recognition performance would be approximately normally distributed.
• Congruent with original Doors and People manual.
This project...

**Measures**

- **Doors and People**
- **Executive Functioning Measures.**
  - COWA
  - Stroop
  - Ravens Progressive Matrix
  - Brixton Spatial Anticipation Test
  - Modified Nelson Card Sorting Test
This project...

Participants

• 615 patients
  • Neuropsychiatric Assessment.
  • Clinically heterogeneous.

• 281 healthy controls
  • Community volunteers.
## Clinical cohort

<table>
<thead>
<tr>
<th>Aetiological Classification</th>
<th>Total Patient Group N=615</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBI % (n)</td>
<td>3.3% (20)</td>
</tr>
<tr>
<td>Multiple Sclerosis % (n)</td>
<td>7.2% (44)</td>
</tr>
<tr>
<td>Epilepsy % (n)</td>
<td>13.8% (85)</td>
</tr>
<tr>
<td>Memory Impairment % (n)</td>
<td>14.1% (87)</td>
</tr>
<tr>
<td>Dementia % (n)</td>
<td>1.6% (10)</td>
</tr>
<tr>
<td>Psychiatric % (n)</td>
<td>23.9% (147)</td>
</tr>
<tr>
<td>RTA % (n)</td>
<td>19.8% (122)</td>
</tr>
<tr>
<td>Workplace Accidents % (n)</td>
<td>2.8% (17)</td>
</tr>
<tr>
<td>Other % (n)</td>
<td>13.5% (83)</td>
</tr>
</tbody>
</table>
Results

• 38% of patients and 30% of controls performed ‘atypically’ with a poorer recognition than recall performance.

• In both cases, frequencies of discrepancy scores met ‘rule of thumb’ guidelines for an assumption of normality (skewness and kurtosis values between -1 and 1)
Results

Recall-recognition discrepancy scores for patient group. 
($M=-0.38, SD=3.40$)

Recall – recognition discrepancy scores for the control group. 
($M=-1.11, SD=2.99$)
Results

• Discrepancy scores indicate the *difference* in recall vs. recognition scaled scores.
Results

• Discrepancy scores indicate the difference in recall vs. recognition scaled scores.

• Discrepancy scores say nothing about the actual scaled recall and recognition scores.
Results

Recall-recognition discrepancy scores for patient group. 
\((M=-0.38, SD=3.40)\)

Recall – recognition discrepancy scores for the control group. 
\((M=-1.11, SD=2.99)\)
Patient mean age-corrected scale scores for the ‘typical’ (n=89) and ‘atypical’ (n=76) subgroups.

Recognition: (t\textsubscript{163} = 15.81, p< .001).
Recall: (t\textsubscript{163} = 8.77, p< .001).
Results

Patient mean age-corrected scale scores for the ‘typical’ (n=89) and ‘atypical’ (n=76) subgroups.

Recognition: \( t_{163} = 15.81, \ p < .001 \).
Recall: \( t_{163} = 8.77, \ p < .001 \).

Control group mean age-corrected scale scores for ‘typical’ (n=86) and the ‘atypical’ (n=30) subgroups.

Recognition: \( t_{114} = 7.69, \ p < .001 \).
Recall: \( t_{76.79} = 7.47, \ p < .001 \).
Aetiology and Litigiousness

- Research indicates that risk of invalidity increases with litigiousness of injury (Mittenberg et al., 2002; Suhr et al., 1997).
  - Workplace accidents.
  - Road traffic accidents.
  - Mild Traumatic Brain Injury.
Aetiology and Litigiousness

<table>
<thead>
<tr>
<th>Aetiological Classification</th>
<th>Total Patient Group N=615</th>
<th>Atypical patient group n = 76</th>
<th>Remaining patients n=539</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBI % (n)</td>
<td>3.3% (20)</td>
<td>1.3% (1)</td>
<td>4.5% (19)</td>
</tr>
<tr>
<td>Multiple Sclerosis % (n)</td>
<td>7.2% (44)</td>
<td>9.2% (7)</td>
<td>6.7% (37)</td>
</tr>
<tr>
<td>Epilepsy % (n)</td>
<td>13.8% (85)</td>
<td>14.5% (11)</td>
<td>14.6% (74)</td>
</tr>
<tr>
<td>Memory Impairment % (n)</td>
<td>14.1% (87)</td>
<td>9.2% (7)</td>
<td>13.5% (80)</td>
</tr>
<tr>
<td>Dementia % (n)</td>
<td>1.6% (10)</td>
<td>0% (0)</td>
<td>4.5% (10)</td>
</tr>
<tr>
<td>Psychiatric % (n)</td>
<td>23.9% (147)</td>
<td>19.7% (15)</td>
<td>20.2% (132)</td>
</tr>
<tr>
<td>RTA % (n)</td>
<td>19.8% (122)</td>
<td>26.3% (20)</td>
<td>14.6% (102)</td>
</tr>
<tr>
<td>Workplace Accidents % (n)</td>
<td>2.8% (17)</td>
<td>3.9% (3)</td>
<td>2.2% (14)</td>
</tr>
<tr>
<td>Other % (n)</td>
<td>13.5% (83)</td>
<td>15.8 % (12)</td>
<td>19.1% (71)</td>
</tr>
</tbody>
</table>

\[ \chi^2 (1, n= 615)=1.48, p=0.223 \]
Executive Functioning

• Due to distractor items in recognition tasks, relatively poor performance may be attributable to an executive functioning deficit.
Executive Functioning

Atypical and typical patient subgroups.

<table>
<thead>
<tr>
<th>Measure</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>COWA</td>
<td>$p=0.53$</td>
</tr>
<tr>
<td>Ravens Progressive Matrices</td>
<td>$p=0.14$</td>
</tr>
<tr>
<td>Brixton Spatial Anticipation Test</td>
<td>$p=0.23$</td>
</tr>
<tr>
<td>mNCST (perseverative)</td>
<td>$p=0.52$</td>
</tr>
<tr>
<td>mNCST (categories achieved)</td>
<td>$p=0.14$</td>
</tr>
<tr>
<td>Stroop (RT incongruent – RT neutral)</td>
<td>$p=0.86$</td>
</tr>
<tr>
<td>Stroop (incongruent)</td>
<td>$p=0.81$</td>
</tr>
</tbody>
</table>
Discussion

38% of patients and 30% of controls presented with poor recognition relative to recall and R-R discrepancies are approximately normally distributed.

- Our findings suggest that poor recognition relative to recall is a bona-fide pattern of performance.

- Intuitive inference of invalidity is unfounded.

- Wrongful identification of poor effort givers in clinical practice.

- Implications for The Doors and People Amendment.

- Possible implications for objective validity measures that make use of recognition formats (as the majority do).
Discussion

**Double dissociation in recall and recognition performances**

- Supports contention of validity.
  - Poor effort givers suppress performance across recall and recognition measures (Bernard et al., 1993; Flowers et al., 1996; Sullivan et al., 2002).

- Congruent with current neuro-scientific knowledge.
Discussion

Dual Process Model (Brown & Aggleton, 2001)

Recall: Contextualised *recollective* information.
- e.g. ‘I saw Jim at the bar, it was Sunday, he was wearing a blue shirt’.

Recognition: Either contextualised *recollective* information or a-contextual *familiarity*
- e.g. ‘I feel like I know that person, but I can’t say who they are or where I know them from’.
Discussion

Dual Process Model (Aggleton & Brown, 2001)

**Recognition**

- Recollective process (recall): Mediated by the hippocampus.
- Familiarity process: Mediated by adjacent regions such as perirhinal, entorhinal, parahippocampal cortices.
Discussion

Dual Process Model (Aggleton & Brown, 2001)

• Recall and recognition tests may tap into two qualitatively distinct, dissociable retrieval processes.
  • Familiarity vs. recollection processes.

• Recall and recognition tests do, in some cases engage anatomically distinct neural regions.
  • Hippocampus vs. adjacent cortices.

• In such cases, ‘atypicals’ may be very good at recollection, but poor at familiarity.
Discussion

• Causes of such a pattern remain to be clarified.
  • Function of tests?
  • Developmental contingency?
    • Older people become increasingly reliant upon familiarity (Bastin et al., 2003; Howard et al., 2006).
    • Hippocampus highly susceptible to aging (Rosenzweig & Barnes, 2003)
  • Normal variation?
Discussion

Aetiology/Litigation

• Findings support contention of validity.
Discussion

Executive Functioning

• Findings support contention of validity.
• Poor recognition vs. recall not caused by impairment in executive functioning.
Thanks!