

Would I lie to you?

*An introduction to Performance and Symptom Validity Testing
during Psychology Assessments*

Dr Laura Pison-Young
Clinical Psychologist



Learning Points

01	To understand what is meant by performance validity in psychological assessments
02	What can make a psychological test unreliable or invalid?
03	What methods do psychologists use to assess validity?
04	Why does validity matter and why test security is so important
05	An introduction to concepts around denial and self-deception

Key Resources

- **British Psychological Society (2021) Guidance on the assessment of performance validity in neuropsychological assessments.**



Validity in Psychological Assessments

Assumptions

- That the person is giving a genuine account of their symptoms;
- That the person is trying to complete psychological tests to the best of their ability;
- That the person has not practiced tests in advance.

Without the above

- The results of a test/ interview are INVALID: i.e. conclusions that are not an accurate assessment of the person.
- Increases the risk of a false positive or a false negative conclusion

What might affect validity?

Why did the person give the answers they did when I tested them?

- Mental illness
- Physical illness and pain
- Cognitive impairment
- Learning difficulty (e.g. dyslexia)
- Sensory impairment
- Drugs and alcohol
- Tiredness
- Incorrect test conditions
- Incomplete or lack of insight
- Mood
- Hormonal differences
- Trauma
- Lack of motivation
- Inattention & distractibility
- Lack of rapport with examiner
- Cultural & linguistic matters
- Repetition and learning effects
- Diagnostic overlap
- Malingering
- Coaching
- Deliberate falsification & exaggeration
- Confirmation bias

Malingering, Effort and Performance Validity

Malingering: “The dishonest and intentional production or gross exaggeration of physical or psychological symptoms for external gain such as avoiding work, obtaining financial benefits, evading criminal prosecution or obtaining drugs/ treatment”

Different from psychiatric conditions such as factitious somatoform or conversion disorders in which symptoms are expressed for psychological reasons without obvious external gains.

Types of Malingering (Lipman, 1961)

- Invention of symptoms
- Exaggeration of genuine symptoms
- Perseveration: describing symptoms that previously existed but have now resolved.
- Transference: genuine symptoms to a false cause

-

In medicolegal contexts, malingering is usually for the judicial process to decide: psychologist indicates whether the interview data is reliable or not.

“Effort Testing”: Out of date term now, as it takes effort to underperform and pejorative term.

Performance validity: i.e. how likely is it that the results (performance) are credible and accurate? Has the person manipulated the interview or test results in some way in order to gain an intended outcome?

Validity and “malingering” are not the same: can be many reasons for test “failure”.

Why does validity matter?

Serious consequences both for over- and under-diagnosis of problems

- Incorrect treatment (e.g. medication; therapy)
- Financial (e.g. compensation; access to disability benefits)
- Over- or under-referrals to relevant services/ treatment
- Inefficient use of health, social care and charitable resources
- Employment and education (sickness; extra time in exams; suitability for a job e.g. pilots)
- Driving
- Evading responsibilities (e.g. military service)
- Criminal matters (e.g. evading prosecution; reducing sentencing)
- Risk assessment (e.g. child protection; forensic settings)

How does a psychologist assess validity?

- **Formulation driven: integrate a range of information to reach an overall opinion.**
- Weigh up information for and against possible false/ inaccurate reporting.
- Motives: what are the reasons someone might not provide accurate information.
- Alternative explanations: can the invalid presentation be explained by some other reason?
- Performance Validity Tests (PVTs) – credibility of test performance
- Symptom Validity Tests (SVTS) – credibility of symptoms.
- Consistency across time, setting, sources of information
- Observed behaviours
- Plausibility: whether the presentation fits with clinical experience, research and diagnostic criteria

In the Interview

“Magnificent 7” (Barnett, 2021)

1. **Rare symptoms:** over-reporting of symptoms that are usually very infrequent or unusual
2. **Improbable symptoms:** unlikely, bizarre or over-dramatic symptoms.
3. **Symptom combinations:** psychological symptoms that do not usually go together.
4. **Symptom severity:** Symptoms expressed as severe, continuous or extreme.
5. **Indiscriminate symptom endorsement:** claiming to suffer from a vast array of symptoms, whilst being vague and non-specific.
6. **Obvious versus subtle symptoms:** focusing on well known, obvious stereotypical symptoms
7. **Reported versus observed symptoms:** inconsistencies between these, do the symptoms increase after it has been observed or commented on.

Test Security and “trade secrets”

- Psychological testing and detecting performance management relies on our tests not being widely known, practiced or distributed.
- This is particularly important in PVTs and SVTs
- Risk of tests becoming redundant over time and tests no longer being discriminatory.
- We only have a finite number of PVTs – tests take years to validate and develop.
- Internet and digitisation increases need for test security
- Many tests are bought under licence, which prohibits distribution and dissemination.
- Can cause tensions with court process: hot topic!

How to Preserve Test Security

- Not describing test items or procedures in detail in reports
- Administering tests in private, to a minimal number of people (avoiding observers)
- Avoid naming exact tests before the interview
- Don't alert when a test of validity is being used (e.g. Gervais et al., 2001)
- Administer protected tests under direct supervision - don't allow materials to be taken away.
- No videoing or use of phones during testing
- We can be expected to name the tests we've used in reports and the sources we've used to determine scores but we should not be distributing test materials or describing the process in detail.

Motives to manage performance

First question:

What is the context and does the person have a clear and obvious external incentive to provide an inaccurate picture?

In some populations there is a high rationale to manage performance (e.g. personal injury claims; criminal proceedings; family court). Performance validity issues can be high: 25% across all groups (McWhirter et al., 2020;

Common Motives:

- Financial gain
- Perceived benefits to the self or to avoid perceived negative consequences
- Access to some perceived privilege
- To access treatment/ drugs
- To avoid negative outcomes
- Belonging to a particular group or distancing self from a diagnosis
- Stigma or shame

Performance Validity Tests (PVTs)

- PVTs are standardised ways of testing for malingering, exaggeration, or other inaccuracies.
- Mainly linked to cognitive testing such as memory and IQ – multiple examples.
- Neuropsychologists are the experts in this area.
- Can be free standing or embedded.
- Have been standardised against groups
- Specificity: How well does my PVT detect unreliable performance in the group I want it to
- Sensitivity: How many inaccurate performers does my PVT catch?
- Recent research highlights the need for tailored failure cutoffs depending on clinical populations (e.g. people with borderline IQ; different cultural groups)

Symptom Validity Tests (SVTs)

- Different from a PVT: an SVT assesses the credibility of reported symptoms and can be used to assess general mental health symptoms (Newns, 2025)
- Can be stand alone or embedded- often a questionnaire or scale on a wider questionnaire
- Can be generic (e.g. in personality scales) or targeted (e.g. over-exaggeration of psychotic symptoms; tendency to be present positive impression)
- Validated against norms groups: i.e. how do the responses by the interviewee compare to someone who is known to have the condition or is not trying to present well.

Common Methods used in PVTs and SVTs

- **Hold Tests:** Measures performance in areas known to be resistant to effects of the condition but this is not widely known in the general population
- **Forced Choice:** Person's performance/ responses are compared to that expected by chance (i.e. random responding).
- **Floor Effects:** Performance on tasks/ problems using overlearned material or easily accomplished tasks (a naive population will over-estimate impairment).
- **Atypical Performance or Symptoms:** Comparing what the interviewee describes or does against known profiles for genuine cases.
- **Consistency:** Does the person answer similar questions in the same direction (random responding) or if given the same test twice at different times.
- **Magnitude of Error:** How close was the person to "passing" the tests.
- **Performance curve analysis:** Normal performance shows typical patterns such as increased failure rate as difficulty increases, consistent performance on easier items.

How do we know if someone has failed a PVT?

- PVTs often use “cut off” scores: these are often derived from specificity and sensitivity. There will often be different cut off scores for different groups. E.g. dementia patients.
- Important for the psychologist to think about which group you are comparing your scores to
- Cut offs are generally quite conservative for some groups (i.e. it’s quite hard to fail!)
- Generally, the more PVTs are “failed” the more likely it is that there is underperformance (e.g. AACN, 2021 if 3 tests are failed, 99% likelihood of validity issues).
- When PVTs are failed, high likelihood that rest of testing is invalid (McWhirter et al., 2020)
- **Reminder!** Cannot infer intent; can only say test results are unlikely to be accurate.
- Also: No test is perfect! Some degree of error is to be expected. Some tests may “fail” genuine cases (e.g. 90% sensitivity = 10% of genuine cases did not pass)
- Formulation driven: Why might this person not pass this PVT when I administered it?

How many PVTs and SVTs to use

- Depends on context and what you are assessing.
- Psychologist should be prepared to explain and justify their decision making re use or not of PVTs and SVTs
- Too many = test fatigue, time consuming and increases risk of false positive due to chance
- Too few or none = an assessment that lacks validity and cannot rule out performance issues.
- Generally one PVT is harder to interpret in isolation (exception in below chance responding)
- Aim for a mix of embedded and stand alone PVTs.
- Use a variety of methods and triangulate the information.

Collateral Information

- Medical records
- Court papers/ other information (e.g. school reports; previous tests; other expert reports)
- Self-report data – does it match what symptom profiles look like?
- Presentation and clinical observation (formal and informal)
- Interviews with others (e.g. parents; carers; social workers) BUT also need to consider their possible motives.
- Look for exceptions: “bad days” and “good days”
- Ask for specific and individualised examples

Deception and Denial

Lying is universal: can be prosocial, functional, self-deceptive, self-serving etc

Crittenden Self-Deception Scale

Intentional Deception: Conscious, calculated misleading. Speech and behaviour are organised to maintain the deception. As the truth is known to the deceptive person, signs can leak out (e.g. verbal dysfluency, sweating, vague; not answering questions).

Self-deception: To avoid signifiers of dishonesty, uses a variety of psychological approaches such as recalling an imagined rather than a real event, relying on emotional memories not “facts” (von Hippel & Trivers, 2011). Crittenden suggests this can be corrected with factual information – the deceiver may not be aware they are mistaken.

Involving & Self-justified Deception: Deceptive individual uses strategies to ensure they feel justified in deceiving others: rationalises & justifies deception (e.g. to protect the self, relationship or status). Information about complicity is omitted from memory, blames others presenting self as a victim, blanket denial and expressions of emotion (e.g. anger). Uses excessive detail to distract or confuse the listener, becomes confrontational or highly distressed. Examples can be seen in domestic abuse: e.g. “DARVO”

Involving, reciprocal and intentional self-deception: Provides false, distorted information about the self and others, deceived person comes to believe that the person is trustworthy. People may use sophisticated psychological devices/ strategies to omit & hide own contributions to problems: e.g. revealing partial responsibility or faux humility. The deceived person can then confirm the “false” persona is accurate, thus confirming the lie in the person’s mind.

Factitious, Functional & Dissociative Disorders

Factitious disorder - presents as ill, injured or impaired to others through falsification of symptoms, exaggeration or induction of injury/ disease.

Functional disorders (conversion disorders): physical symptoms that cannot be explained by an organic disease or injuries (e.g. non-epileptic seizures, movement problems)

Dissociative disorders: a group of conditions where you feel disconnected from reality: dissociative identity disorder, depersonalization, derealization and dissociative amnesia.

- How to separate these mental health conditions from “malingering”
- Rewards: what are the conscious and unconscious gains and psychological drivers (e.g. sick role; access to care; attention; reduces loneliness or demands; way to express trauma/ psychological distress)
Difficult to separate from “malingering”: not either/ or, both external and internal gains might be present.
- Formulation: what is the clinical history, trauma and abuse, what else is going on, when do symptoms change? Relationship to professional help and medical professionals. How to communicate performance management issues? We must be wary of what we don't know re unexplained or perplexing symptoms. Already marginalised groups tend to be labelled as “faking it” more often (e.g. women; people of colour).

Can we tell when people are not being truthful?

Not really!

No reliable behavioural signs of lying that are 100% accurate: detection relates more to skill of the liar than the detector!

Lie detectors do not work

We do have individual “lie” profiles, which we can learn to recognise in people we know well... (not failsafe)

May reflect what the lie is for and function of deception: e.g. prosocial lying (e.g. Father Christmas? Not hurting someone’s feelings)

BUT: there are signifiers for deception which can be interpreted into general clinical picture (physiological signs; communication differences; behavioural changes; factual discrepancies; an absence of specific information).

Some groups are better at deception:

- Cognitive capacities: higher IQ = better deception
- Personality structure: antisocial, narcissistic, psychopathic and egocentricity (Machiavellianism) traits versus conscientiousness & sociability (Verigin *et al*, 2019; Hilbig *et al.*, 2025)
- Agreeableness?

Case Example 1: Cognitive Functioning

Context: Asylum seeker's cognitive ability and eligibility for support. A previous IQ test had determined a moderate learning disability and recommended supported accommodation. No PVTs used or discussion of validity. Second opinion requested.

Test results: Used a variety of culturally sensitive IQ tests, almost uninterpretable due to "floor effects". Non-verbal PVT was used plus embedded measures. Did not pass the PVT at below chance level. Worse than patients with dementia.

Clinical Interview: Low validity of self-report (e.g. very unusual recall of past events and personal history); poor engagement.

How was performance validity tested: PVTs. Clinical interview. Discussion with support worker, case worker and review of medical records which included social work reports

Presentation: Able to talk in English with solicitor outside of interview, noted to buy items from shop during break, tell the time and use a mobile phone. Contrasted with presentation during interview.

Motive to underperform? Yes.

Collateral information from third party: Fully functioning independent, able to travel independently from one major city to another, drive a car, cook for self, manage own environment and budgets well. Seen to read and write.

Other explanations: Yes, possibly. Cultural issues, lack of education, trauma, low mood, attitude towards people in authority?

Conclusions:

Not possible to give a firm view of IQ but unlikely to be as impaired as first thought given functional abilities. Nevertheless, the person is likely to be vulnerable in some areas. Some targeted support recommended re mood and trauma.

Example 2: Case Discussion - Family Court

Context:

Possible induced illness and medical neglect in the child due to the parent's mismanagement of the child's medical condition. Parent reported history of dyscalculia and learning needs: claiming this was the reason for the child's illness not FII or medical neglect.

Test results: Consistent with dyscalculia and borderline learning disability. Passed all PVTs apart from embedded test linked to working memory (could be explained by dyscalculia).

Process: School reports, medical records, PVTs, clinical information, specific tests, conversation with parents and social worker.

Motive to underperform? Yes – wanted child returned to her care. Wanted to avoid criminal charges.

Outcome: Performance likely to be valid; meant risks better understood and rehabilitation of the child with safety plans in place could be implemented. Allowed for targeted interventions and safety planning (e.g. adapted equipment that did not require knowledge of numbers; outsourcing of some tasks to the grandparents and carers).

Example 3: Over-exaggeration of symptoms

Context:

Queried health anxiety or fabricated or induced illness in the children by mother. Mother believed she may have autism, query whether autism has relevance to the FII allegations.

Procedure:

Specific autism assessment, plus cognitive assessment & full psychological. Autism test blind coded by another clinician, review of medical records, court bundle, third party interview with maternal grandmother. SVTs in personality tests plus stand alone SVT. No PVTs used apart from those embedded. SVTs showed some positive impression management and over-reporting of symptoms.

Motive: Yes, wanted children returned and did not agree with FII, felt symptoms were valid.

Outcome: Some exaggeration and health anxiety was possible but did not explain underlying autism diagnosis. Symptom validity supported by collateral history and presentation. Positive impression management likely artefact of court process. Blind coding was persuasive argument during cross examination. Helped to formulate mother's difficulties as literal, over-interpretation of symptoms and health anxiety with FII less likely.

Brown v Morgan Sindall

Construction and Infrastructure Ltd [2025] EWHC 2204

Cyclist who suffered an orthopaedic injury subsequent PTSD and depression symptoms. Validity testing in relation to mental health rather than neuropsychological symptoms.

The claimant was assessed by two psychiatrists: one who used SVTs and the other who did not.

The psychiatrist using SVTs opined that the claimant reached the criteria for malingering.

This was challenged on the basis that:

- The claimant felt rushed and had dyslexia: not been properly considered by the psychiatrist.
- Over-reliance on test data - lacked a holistic assessment.
- The psychiatrist did not have expertise in validity testing
- Reliance on psychometric data that could not be adequately scrutinised in court: raised questions about transparency and fairness.

The judgment highlighted challenges when experts rely on psychometric data that is not disclosed to the court. The judgement argued that data available for scrutiny should include questions and answers given (This is being challenged!).

Detailed that PVTs are less widely studied in the UK and is an emerging field, limits our conclusions.

Judgement underscores that validity tests in expert evidence are useful but often not sufficient on their own to conclude that the credibility of psychiatric symptoms is undermined.

Some confusion about what the psychiatrist actually used and whether they were or were not appropriate.

For more info: see Dr Newns (2025)

Take Home Points

- Psychologists can and should report on effort, consistency and credibility in their assessment; however, intent and determinations of malingering are for the Court to decide.
- PVTs and SVTs are part of a psychologist's toolkit but caution against over-reliance.
- Aim for a holistic opinion on performance management triangulating full clinical picture: history, collateral information, consistency, comorbidity, presentation and performance on PVTs and SVTs.
- Psychologists should understand the tests they're using and be aware of their limitations.
- Test security is everyone's business
- Lying is difficult to detect accurately: our own biases, attitudes and experiences can impact our judgements.

References

- Dandachi-FitzGerald, B., Merckelbach, H. & Ponds, R.W. (2017). Neuropsychologists' ability to predict distorted symptom presentation. *Journal of Clinical and Experimental Neuropsychology*, 39(3), 257–264.
- Gerhand, S., Jones, C. A., & Hacker, D. (2021). Effort testing, performance validity, and the importance of context and consistency. *In Neuropsychological Aspects of Brain Injury Litigation (pp. 89-115). Routledge.*
- Gervais, R. O., Russell, A. S., Green, P. A. U. L., Allen, L. M., Ferrari, R. O. B. E. R. T., & Pieschl, S. D. (2001). Effort testing in patients with fibromyalgia and disability incentives. *The Journal of Rheumatology*, 28(8), 1892-1899.
- Hilbig et al., 2025), B. E., Thielmann, I., & Heck, D. W. (2025). Filling in the missing pieces: Personality traits (un) related to dishonest behavior. *European Journal of Personality*, 39(5), 732-746.
- Jerry J. Sweet, Robert L. Heilbronner, Joel E. Morgan, Glenn J. Larrabee, Martin L. Rohling, Kyle B. Boone, Michael W. Kirkwood, Ryan W. Schroeder, Julie A. Suhr & Conference Participants (2021) American Academy of Clinical Neuropsychology (AACN) 2021 consensus statement on validity assessment: Update of the 2009 AACN consensus conference statement on neuropsychological assessment of effort, response bias, and malingering, *The Clinical Neuropsychologist*, 35:6, 1053-1106,
- McWhirter, L., Ritchie, C. W., Stone, J., & Carson, A. (2020). Performance validity test failure in clinical populations—A systematic review. *Journal of Neurology, Neurosurgery & Psychiatry*, 91(9), 945-952..
- Newns, K. (2025) Clarifying the role of validity testing in expert evidence: *Case Update*. www.ewi.org.uk/News/Case-Updates/clarifying-the-role-of-validity-testing-in-expert-evidence-2
- Roor, J. J., Peters, M. J., Dandachi-FitzGerald, B., & Ponds, R. W. (2024). Performance validity test failure in the clinical population: A systematic review and meta-analysis of prevalence rates. *Neuropsychology Review*, 34(1), 299-319.
- Tracy, D. K., & Rix, K. J. (2017). Malingering mental disorders: Clinical assessment. *BJPsych Advances*, 23(1), 27-35.
- Von Hippel, W., & Trivers, R. (2011). The evolution and psychology of self-deception. *Behavioral and brain sciences*, 34(1), 1-16.
- Verigin, B. L., Meijer, E. H., Bogaard, G., & Vrij, A. (2019). Lie prevalence, lie characteristics and strategies of self-reported good liars. *PloS one*, 14(12).

Thank you for joining us.

Do you have any questions?

