

SL IB Psychology

Factors Influencing Diagnosis: Validity & Reliability of Diagnosis

Contents

- * How are Validity & Reliability Relevant to Diagnosis?
- * Two Key Studies of Validity & Reliability of Diagnosis: Rosenhan et al. (1973); Nicholls et al. (2000)
- * Summary Table: Key Studies of Factors Influencing Diagnosis

© 2025 Exam Papers Practice. All Rights Reserved

How are Validity & Reliability Relevant to Diagnosis?

Validity and Diagnosis

- For a mental illness **diagnosis** to be **valid** it must accurately reflect the patient's **symptoms**, free from **bias** (**clinical biases** in diagnosis are covered in a separate revision note)
- A valid diagnosis is one which should **classify** and describe a genuine **pattern of symptoms** resulting from a real underlying cause
- A valid diagnosis will result in appropriate **treatment** being prescribed with the expectation of improvement and progress as a result of this treatment
- Due to the complex nature of mental illness the diagnostic process is not always straightforward e.g. is the patient's low mood due to **depression, anxiety, OCD** or could it be part of a potentially more serious disorder such as **schizophrenia**?
- It is arguably more difficult for a clinician to diagnose a specific mental illness than it is for them to diagnose a physical illness e.g. Covid-19 is detectable by testing saliva; a broken bone shows up on an X-ray

Reliability and Diagnosis

- For a mental illness diagnosis to be **reliable** there should be **agreement** and **consistency** across different diagnostic settings i.e. the same diagnosis for the same symptoms presented by the same patient should be made, regardless of who is in charge of the diagnosis
- **Classification systems** such as the **DSM-5** and the **ICD 11** aim to **standardise diagnostic criteria** so as to ensure built-in reliability
- Due to the complex nature of mental illness the likelihood of **symptom overlap** is common, making reliable diagnosis problematic e.g. if a patient reports hearing voices along with the compulsion to wash their hands every 10 minutes this may result in one diagnosis of schizophrenia whereas another clinician may diagnose OCD
- Patients who are **comorbid** may find that the treatment prescribed is based on only one of their disorders rather than taking both of them into consideration (e.g. **SSRIs** for their depression but nothing for their **social phobia**)
- The symptoms of mental illnesses are difficult to measure as they are experienced **subjectively** and may even defy measurement (e.g. how can **delusions** be measured **objectively**?) which is a challenge when it comes to making a reliable diagnosis

Which studies investigate validity & reliability of diagnosis?

- **Rosenhan et al. (1973)** - mental illness diagnosis may not be valid and may result in people being stigmatised
- **Nicholls et al. (2000)** - eating disorders in children are not diagnosed reliably and the process needs to be reviewed

Both Rosenhan et al. (1973) and Nicholls et al. (2000) are available as Two Key Studies of Validity & Reliability of Diagnosis – just navigate the Factors Influencing Diagnosis section of this topic to find them.

Two Key Studies of Validity & Reliability of Diagnosis: Rosenhan et al. (1973); Nicholls et al. (2000)

Key Study One: Rosenhan (1973)

Key study one (validity of diagnosis): Rosenhan (1973)

Aim:

- To investigate the **validity** of mental illness **diagnosis**
- To investigate the consequences of the '**sticky label**' of a mental illness diagnosis

Participants:

- The study used **naive participants** from the following:
- The staff and patients from 12 **mental hospitals** from across the USA
- The hospitals varied in terms of age, location, staff-patient ratios, expertise

Observers:

- Rosenhan recruited eight **confederates** who comprised his sample of **pseudopatients** who infiltrated the mental hospitals and made **covert observations** of the hospital staff and patients
- The pseudopatients consisted of 3 females and 5 males with Rosenhan himself assuming a pseudopatient role as well
- The pseudopatients were from a range of different backgrounds and none of them had a mental illness
- The pseudopatients were told to use fake names and occupations when they presented themselves for diagnosis

Procedure:

- The confederates recruited by Rosenhan (known as 'pseudopatients' as they would be faking their symptoms) were instructed to present themselves at one of the 12 hospitals selected by Rosenhan

- Upon getting an appointment with a doctor they were told to report the following symptoms: *I have been hearing a same-sex voice in my head which repeats the words 'empty', 'hollow' and 'thud'*
- The pseudopatients were told to behave normally during the consultation and not to fake any other symptoms of mental illness
- All but one of the pseudopatients were admitted to hospital with a diagnosis of **schizophrenia** (one of them was admitted with a diagnosis of **bi-polar disorder**)
- Once the pseudopatients had been admitted to hospital Rosenhan's instructions were that they were to never mention their (fake) symptoms again, to behave normally and to persuade the hospital to release them as soon as possible
- Rosenhan also told the pseudopatients to keep notes of what they **observed** during their time in hospital relating to both staff and patients
- The pseudopatients were told not to take any drugs administered to them by hospital staff but to dispose of them discreetly
- The **dependent variable** was the number of days spent in hospital before release
- The overarching **method** of this research is a **covert participant observation**

Results:

- The notes made by the pseudopatients while in hospital detailed the everyday interactions between staff and patients
- Interactions between staff and patients was sparse, with staff often ignoring patients, dismissing their requests (e.g. asking when visiting hours were), making **little eye contact** with the patients
- Normal behaviours were often interpreted by staff as aspects of mental illness e.g. three pseudopatients were told that their writing was evidence of **pathological behaviour**, labelling this is 'writing behaviour' rather than simply 'writing'
- On one occasion a psychiatrist pointed to a group of patients queuing for lunch and labelled this behaviour as '**oral-acquisitive syndrome**' rather than simply accepting that they were just queuing up for lunch
- None of the staff suspected that the pseudopatients were fake, however 35 out of 118 patients approached the pseudopatients and voiced their suspicions that the pseudopatients were not actual patients (some of the patients thought that the pseudopatients might be undercover journalists)
- The pseudopatients spent from 7 to 52 days in hospital (**mean**=19 days)
- All but one of the pseudopatients were released from hospital with a diagnosis of 'schizophrenia in **remission**'

Conclusion:

- There are questions to be asked re: the validity of mental illness diagnosis as the doctors should not have diagnosed any of the pseudopatients with schizophrenia or bi-polar disorder as their (fake) symptoms do not align with either of these diagnoses
- Once someone has been diagnosed with a mental illness this becomes a 'sticky label' through which all subsequent behaviours are viewed and judged
- Patients hospitalised with a mental illness experience **depersonalisation** due to the indifferent, sometimes hostile treatment at the hands of hospital staff

Evaluation of Rosenhan (1973)

Strengths

- The use of research in the field via covert observational methods means that the observed participants are unlikely to have succumbed to the **observer effect**, making the findings high in **ecological validity**
- This was a controversial, ground-breaking study which provoked important discussion about how people suffering from mental disorders are treated by institutions

Weaknesses

- The study does raise some **ethical concerns**: the staff and patients of the hospitals were **deceived**; the hospital participants could not give their **informed consent** or be given the **right to withdraw** plus their **privacy** was compromised
- A sample of only 8 pseudopatients is not enough from which to draw strong and meaningful conclusions plus there is the possibility that the pseudopatients might have succumbed to **confirmation bias** in reporting their observations



One of Rosenhan's key findings was that mental hospitals rob people of their individuality.

Key Study Two: Nicholls et al. (2000)

Aim: To evaluate the reliability of **diagnostic classification systems** for **eating disorders** when applied to children and young adolescents.

Participants: 81 children aged 7–16 who had been selected via **random sampling** from a **population** of 226 child patients attending a clinic specialising in **eating disorders**.

Procedure:

- Each child was assessed by one of six clinicians
- The clinicians were asked to use either the DSM-IV, the ICD 10 or the Great Ormond Street Hospital (GOSH) diagnostic manual in to form their diagnosis of each child
- Each clinician gave their diagnosis as to which specific eating disorder the child was suffering from, using one of the three diagnostic manuals cited in the above bullet point
- Two clinicians assessed each child (each clinician having used a different diagnostic manual to the other) without knowing about each other's diagnosis i.e. they were **blind** to the pre-existing diagnosis

Results:

- **Inter-rater reliability** values were calculated for each of the three diagnostic manuals used to come to reach the diagnosis
- The higher the inter-rater value is, the more reliable the diagnosis is
- The results per diagnostic manual were as follows:
 - GOSH: 0.879
 - DSM-IV: 0.636
 - ICD 10: 0.357
- The GOSH definitions included **anorexia** and **bulimia nervosa**, **food avoidance emotional disorder**, **selective eating** and **pervasive refusal to eat** amongst their classification of eating disorders
- GOSH criteria had been specifically developed to classify child and adolescent eating disorders: they were more **reliable** than the DSM IV and ICD 10 criteria, which showed little consistency, especially the ICD 10, which had the lowest inter-rater reliability of all the classification systems
- The DSM-IV and the ICD 10 focused too much on body shape and weight which are invalid criteria when diagnosing eating disorders in children

Conclusion: The DSM and ICD are not suitable classification systems for the diagnosis of eating disorders in children; a clinician working diagnosing children with eating disorders requires tailor-made criteria such as those supplied by GOSH.

Evaluation of Nicholls et al. (2000)

Strengths

- The study's use of blind clinicians (who did not know the diagnosis given by their counterpart) increases the validity of the findings as it helps to eliminate **bias** from the assessments provided
- The findings are vital in that they pinpoint flaws in the more traditional classification systems and highlight how children with eating disorders should be diagnosed

Weaknesses

- A sample of 81 children from the UK is small and unrepresentative of the wider population, making the results difficult to generalise
- The research only highlights how children with eating disorders should be diagnosed, it does not account for other disorders which may also require a separate and specific classification system



Summary Table: Key Studies of Factors Influencing Diagnosis

Key Studies Summary of Factors Influencing Diagnosis

SUMMARY TABLE: KEY STUDIES OF FACTORS INFLUENCING DIAGNOSIS	
Topic	Two Key Studies
Normality vs Abnormality <ul style="list-style-type: none"> Use both of these studies to answer a question on normality vs abnormality Use Mojtabai (2011) to answer a question on classification systems as well 	Jahoda (1958) Mojtabai (2011)
Classification Systems <ul style="list-style-type: none"> Use both of these studies to answer a question on classification systems Use Mojtabai (2011) to answer a question on normality vs abnormality as well Use Haroz et al. (2017) to answer a question on validity and reliability of diagnosis as well 	Haroz et al. (2017) Mojtabai (2011)
The Role of Clinical Biases in Diagnosis <ul style="list-style-type: none"> Use both of these studies to answer a question on the role of clinical biases in diagnosis 	Longnecker et al. (2010) Jenkins-Hall & Sacco (1991)
Validity & Reliability of Diagnosis <ul style="list-style-type: none"> Use both of these studies to answer a question on the role of validity and reliability of diagnosis Use Nicholls et al. (2000) to answer a question on classification systems as well 	Rosenhan (1973) Nicholls et al. (2017)

How do I use these studies in an exam question on this topic?

- IB students have a lot of content to cover (particularly students taking Psychology at Higher Level) so the purpose of this revision resource is to slim down and streamline the number of studies you need per topic/exam question
- Remember that all Paper 2 questions are ERQs (Extended Response Questions) which are worth 22 marks, take an hour to write and need to be rich in critical thinking
- The exam question command term will be one of the following: 'Evaluate', 'Discuss', 'Contrast' or 'To what extent'
- Each command term requires you to answer the question in slightly different ways, using the content as shown in the summary table above i.e. specific studies per topic/question
- In order to slim down the content you need to revise you can see above how some of the studies can be used for more than one potential exam question

- Mojtabai (2011), Haroz et al. (2017) and Nicholls et al. (2000) can be used to answer more than one potential exam question so you may decide to keep all of these studies and 'throw away' any studies which you find that you don't need to revise

