Constructor Method

- Tests
  - the purpose of testing is to make inferences about latent constructs.
- Simulations
- Interviews

Test

- A systematic procedure for observing behaviour and describing it with the aid of numerical scales or fixed categories.
- Two broad categories:
  - Speed tests
  - Power tests
- the key distinction between speed and power tests is **the source of variance in test scores**

Speed tests

- consist of a large number of relatively easy items
- speed tests are given under very strict time limits
- Although most test-takers will answer the questions correctly, most test-takers will not finish the test
- variance on speed tests is derived from **the number of items** that are completed

Power tests

- ower testsinclude items of varying difficulty, and test-takers are given sufficient time to answer all of the questions.
- variance in power test scores comes from the number of items that are answered correctly
- university exams are power tests

Computer Adaptive Testing

- a process in which a computer program can be used to administer a test that adapts to the individual test-taker's level of proficiency on the construct being measured
- **Procedure**
  - The test-taker sees an item that is moderately difficult
  - If he or she answers correctly, the next item will be slight more difficult
  - Likewise, if the item is answered incorrectly, the next item will be easier
  - The test continues to adapt, ultimately settling at the test-taker's level of proficiency
- might get the same number of items correct,but answered more difficult questions compared to the other person
- scores are computed as the sum of correct answers, weighted by the difficulty of the questions to which the test-taker responded
- **Advantage**
  - Test-takers in the "tails" of the distribution can be differentiated
  - Far fewer items are needed
- it is quite useful to be able to differentiate people with high GMA from people with VERY high GMA

Testing summary

- traditionally most test items have been written so as to maximize differentiation among people in the moderate range of the distribution
- An item characteristic curve for a moderately difficult item is shown
  - This item simply does not work very well for differentiating these two individuals with high GMA. (as high GMA but the differences between the number of correct questions they answer is quite small)
- How can we differentiate between people in the extreme "tails" of the distribution?
  - add easier(for the lower tail) and harder(for the upper traditionally) item for the test
  - Easy items differentiate between people at the low end of the distribution
    - people with true score below 0, as true score GMA increases, the odds of answering the item correctly increase

    - nearly everyone with a true score above 0 will answer the item correctly
  - difficult items differentiate between people at the high end of the distribution
    - with true score above 0, as true score GMA increases, the odds of answering the item correctly increase
    - everyone with a true score below 0 will answer the item incorrectly

Simulations

- Work simulations are based on the principle of behavioural consistency.
- Simulations tend to have strong validities for predicting overall work performance
- Three important considerations when using simulations:
  - There must be a **strong match** between the behaviours measured during the simulation and the actual behaviours required on the job
  - Simulations are only appropriate in situations in which the applicant **already has the requisite KSAs**
  - Costs - simulations can be more expensive to develop and administer relative to other selection methods

Work samples

- work sample simulations are tests in which the applicant performs an actual sample of the behaviours required on the job
  - both physical tasks or cognitive problem-solving tasks
- Work samples may also be used to assess trainability
- Work samples must be developed in accordance with the results of the job analysis.
  - job analysis is used to identify the most frequent and important tasks/behaviours performed on the job.
- systematic scoring schemes must be developed in order to standardize scores across applicants
- work samples exhibit strong criterion-related validity
- work samples consistently demonstrate incremental validity
- work samples can provide applicants with a realistic job preview
  - result in less voluntary turnover

Assessment centers

- collections of simulation exercisesused to measure KSAs that are critical for managerial performance
- ACs are used for both selection and professional development
- ACs are completed over multiple days, and thus, are quite expensive and intensive
- ACs are typically used only for individuals in high level positions
- Assessment centers are used to measure multiple dimensions using several exercises
- Common exercises include:
  - In-basket
  - Leaderless groups
  - Role playing
  - Case analysis
- **In-basket exercises** simulate the process of dealing with **paperwork and other common tasks** that arrive in a manager's "in basket" throughout the day
- **eaderless group discussions** involve a group of participants working to solve a common problem in a set period of time
  - assess the degree to which participants demonstrate teamwork and leadership characteristics when working in an unstructured environment with others
- **Role playing exercises** involve one-on-one discussions between an assessee and another individual
  - to simulate common scenarios, such as delivering negative feedback or dealing with an unsatisfied customer.
  - evaluated based on the degree to which their behaviours are appropriate for resolving the issue at hand
- **Case analysis exercises** involve reading a packet containing a summary of a problem or dilemma being faced by a hypothetical organization
  - required to prepare an oral presentation in which they deliver a set of recommendations for management
  - are evaluated based on the specificity and quality of their recommendations, and the degree to which their recommendations are based on supporting facts and data.

- These scores are averaged across exercises to generate both dimension ratingsand an overall assessment rating
  - Dimension ratings would be computed by taking the average score for each dimension across exercises
  - Overall assessment rating is the average of these dimension ratings
- both dimension ratings and overall assessment ratings are valid predictors of overall work performance
- AC scores exhibit incremental validity over GMA
- **Dimension ratings** are more specific than overall ratings, and thus can be useful for predicting more specific criteria
  - given their specificity, dimension ratings are more useful for providing feedback

Situational judgement tests

- verbal simulations – written descriptions of work situations to which test-takers respond by selecting the most appropriate course of action.
- SJTs typically do not have one "correct" response with the others being "incorrect."
  - Rational: Subject matter experts provide a judgment of the degree to which each response is appropriate
  - Empirical: Each response is correlated with a criterion (e.g., performance) and weights are assigned based on the strengths of the correlations.
  - inappropriate actions are typically given negative weights
- Drawbacks: SJT scores tend to be highly correlated with GMA, thereby reducing their incremental validity
  - reducing the reading demands associated with SJTs – using video-based SJTs

Interview

- a selection method that involve answering a series of oral questions
- Different purposes
  - Prediction of future performance
  - Recruitment of applicants
  - Providing applicants with information about the job/organization

Structured / unstructrued Interviews

- Adding structure to the interview process greatly improves criterion-related validity
- Unstructured interviews are fairly free-flowing conversations
  - there is not necessarily consistency across applicants (i.e. diffeerent applicants are likely to be asked different questions)
  - questions can vary in the degree to which they are actually tied to the job
  - scoring is generally holistic, with interviewers forming an overall impression of the applicant
- Highly Structured interviews tend to have the following
  - The same questions are asked of all applicants
  - Questions are job-relevant; derived from careful job analysis
  - Consistent scoring rules are used for all applicants
  - Each question is scored
  - Multiple interviewersmay be used to increase reliability
  - Interviewers are often given extensive training
- Two stratgies to improve interview quality
  - Limit probing/follow-up questions
    - ensure the same set of questions are asked across applicants
  - Avoid viewing information about the applicant before the interview
    - Don't check applicant's profile; doing so will bias perceptions and influence interview scores
- Structured interview validities: $r_{xy} = 0.44 - 0.62$
- Unstructured interview validities: $r_{xy} = 0.20 - 0.33$
- adding information from an unstructured interview can actually decrease the overall validity of the selection system
- Unstructured interviews are not as useful as structured interviews, and can actually hurt prediction and introduce bias
- Why unstructrued interview still use?
  - having a free-flowing conversation they are able to gather more relevant information and make better predictions

  - structured interviews are far more labor-intensive to develop
  - unstructured interviews may be more useful for recruitment purposes
- Interviewers seeking to give a good impression of the organization, but the highly structured interviews run counter to these goals
  - Solution: Allocate some time at the beginning and end of the interview for conversation; the bulk of the interview should consist of highly structured questions

Interview key concepts

- what do interviews measure?
  - measure a wide range of constructs.(constructs can be difficult to assess via a written tests)
    - Job knowledge and skills
    - An applicant's fit with the organization
    - Personality and habitual behaviour patterns
    - Social and interpersonal skills
  - measure concepts that cannot easily be distilled down to a multiple choice format
  - Interviews should be designed to measure only 2 or 3 constructs
    - identify the constructs that are most easily assessed via interviews rather than other means
  - each construct should be assessed with multiple items
    - should be scored individually and averaged to form construct scores
- Adding structure to an interview results in higher reliability, less bias, and more accurate predictions
- Unstructured interviews should be avoided.

Type of Interviews

- **Situational Interviews**: applicants are given a description of a hypothetical situation and asked "how would you behave in this situation?"
  - The applicant's answer would be scored according to a pre-determined rubric based on the results of a job analysis.
  - situational interviews are more useful for entry-level positions
- **Behavioural interviews**: applicants are asked to describe their actual past behaviour

  - Behavioural interviews tend to require the person being interviewed to have experience with the job for which they are applying
  - behavioural interviews tend to be more useful for complex jobs

- both situational and behavioural interviews are useful for predicting performance across both low and high complexity jobs