



Best Practices for Unsupervised Testing

Document prepared by **The Personnel Psychology Centre**



OCTOBER 2015

The purpose of assessment in staffing is to determine whether the job candidate possesses the knowledge, skills, abilities or competencies necessary to perform the job in an effective way. To accomplish that, assessment instruments must be valid, reliable, fair and defensible.

This document presents best practices on the use, development, administration and scoring of unsupervised tests to protect the integrity, merit and fairness of assessment in appointment processes. The primary focus is on unsupervised knowledge tests taken at a location of the candidate's choosing, often referred to as "take-home exams." Much of the content may apply to other types of unsupervised assessment and may also vary by test type.

What is a “take-home exam”?

- It is a test that is given to candidates to complete on their own time, without supervision, over a specified period of time. It is typically assumed to be open-book.
- It is taken in an unsupervised environment, usually with access to reference texts and materials.
- It can take many forms (including computer-scored, subjectively scored, on-line or paper and pencil) and can assess a range of different qualifications.
- It typically consists of essay or constructed response questions. There are two types of questions:
 - Extended response – allows for lengthier responses from candidates and permits candidates lots of flexibility in their responses (e.g., essay); and
 - Restricted response – limits the form and content of candidates’ responses (e.g., short answer, problem sets).
- Depending on the time limit set, candidates may be able to work at their own pace, thus permitting longer and more elaborate responses to questions.

Advantages of take-home exams

- The main benefit is that they provide an opportunity for candidates to demonstrate their ability to reason, organize and synthesize knowledge and convey judgement and higher-order critical thinking skills in writing.
- They reflect realistic job tasks in which employees are required to organize and interpret information from various sources and generate and communicate thoughts and ideas.

- They increase accessibility as candidates can complete the test at a time and location of their choosing and in their own workspace that is set up to best meet their needs. By using their own computers, candidates have access to special hardware, software or adaptive technology, as required, such as ergonomic mice and keyboards, screen readers and speech-to-text software.
- They reduce operational costs as candidates do not need to travel and organizations do not need to reserve testing facilities, ship test booklets or administer the test.

Disadvantages of take-home exams

- For assessments involving subjective scoring (i.e., ratings) such as essay questions or case studies, the nature of the scoring can be unreliable and subject to a number of biases (e.g., context effect, halo effect).
- For extended response formats, it is time-consuming to appropriately read and score candidate responses.
- The unsupervised testing format allows the opportunity for unauthorized collaboration and consultation with others, plagiarism and access to unauthorized materials.

Best use of take-home exams

- To assess analysis, synthesis and evaluation skills by requiring candidates to organize and integrate information, interpret information, construct arguments, give explanations, evaluate ideas and develop original thoughts/ opinions. Take-home exams also provide a way of testing a candidate’s ability to communicate in writing and to present their responses in a logical manner.



- To assess how candidates apply their knowledge to a specific situation, context or problem (e.g., a case study), rather than regurgitating facts. Take-home exams are not recommended to assess knowledge where candidates can simply copy and paste the correct answer from the Internet.
- To assess small groups of candidates because they can take an extended amount of time to score.
- Consider using several restricted response questions rather than one or two extended response questions to obtain greater sampling of content.
- Avoid giving candidates a choice as to which questions they should answer. This decreases the validity and reliability of the test because each candidate is taking a potentially different test.
- Provide clear directions and expectations:
 - Indicate the length of desired responses (e.g., less than 500 words), expected format (i.e., complete sentences) and point values for each question;
 - Indicate how grammar, sentence structure, etc. will be handled; and
 - Provide a time limit to complete the work (e.g., 48 hours, 5 days).

Guidelines for constructing essay questions

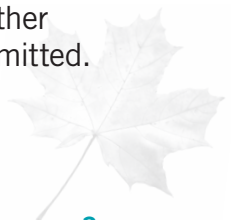
- Clearly define what is to be assessed.
- Formulate the question so that the task is clearly defined for candidates.
- Situate the question within a context or problem/situation.
- Use directive verbs to provide an indication of the types of thinking and content expected in responding to the question. The following directive verbs can help to focus the candidates on the desired approach ([click here for additional directive verbs](#)):
 - Agree/disagree
 - Analyze
 - Compare/contrast
 - Describe
 - Illustrate
 - Justify
 - Summarize
 - Note: Avoid using the term “Discuss” as this can lead to responses that are too broad or vague.
- It may also be helpful to identify the intended audience (e.g., general public, senior management). This provides a context to help guide candidates’ responses.

- For more information, consult the [Checklist for constructing essay questions](#).

Test security guidelines

It is essential to provide explicit instructions about what is and what is not permitted. The following points are provided as examples of information that can be included in the instructions to increase the security and defensibility of a take-home exam.

- Inform candidates of the confidential nature of the test and their responsibility to respect it and to take the test honestly.
- If access to reference materials is deemed appropriate, specify the types of materials that may and may not be consulted (e.g., resources, reports and reference materials) and the expectation that all materials must be cited accordingly.
- If appropriate, inform candidates that consultation or collaboration with other candidates or individuals is not permitted.



- Have candidates confirm and sign that they have read and understood the instructions.
- Include consequences for breach of guidelines (e.g., invalidation of test results, termination of job application, fraud investigation).

The following is an example of take-home exam instructions given to students at Harvard University:

- “You may consult your books or other reference material. You may NOT consult any person other than (the instructor) about any aspect of this exam. The rules of attribution apply to take-home exams: All sources must be cited.”
- Students are also required to sign a compliance affidavit: “I affirm that I have had no conversation regarding this exam with any persons other than (the instructor) and other authorized persons. Further, I certify that the attached work represents my own thinking. Any information, concepts or words that originate from other sources are cited in accordance with (university) guidelines. I am aware of the serious consequences that result from improper discussions with others or from the improper citation of work that is not my own.”

Guidelines for scoring essay questions

- Choose a scoring method with explicit criteria to increase fairness and ensure consistency in scoring. One approach is to prepare an expected response to the questions in advance, where the major components of the response are defined and assigned a point value. Candidates’ responses can then be compared to the expected response and given credit or partial credit when their responses include desired elements from the expected

response. An overall score for a question can then be generated by adding the points received for each required element.

- Decide in advance how elements of written communication will be handled, such as grammar, punctuation, organization and flow, use of vocabulary/terminology and irrelevant/inaccurate information.
- Allow plenty of time for scoring. Often, responses need to be read more than once for accurate scoring.
- Score all responses to one question before moving to the next question.
- Occasionally change the order in which the candidates’ tests are scored to avoid any order effects (e.g., score all question 1 responses, shuffle the exams and then score all of question 2 responses, etc.).
- Score responses without looking at the candidates’ names to avoid possible bias in scoring due to factors such as gender, familiarity with the candidate, etc.
- Have at least two qualified individuals score candidates’ responses independently and discuss/negotiate their ratings.
- For more information, consult the [Checklist for scoring essay questions](#).

Additional recommendations

- Given the potential for unauthorized collaboration and consultation during unsupervised assessment, the Public Service Commission recommends that supervised verification testing be conducted after the unsupervised assessment for those candidates who passed the unsupervised version. Notifying candidates that their scores on the unsupervised exam will be verified through another independent assessment method can be a disincentive to unauthorized test-taking behaviour.



- It is recommended that essay questions only be used once, since they are easily shared and easy to remember.
- A plagiarism check may be performed by reviewing likely documents that candidates would be consulting.

References

Center for Innovation in Teaching & Learning (2009). Improving your test questions. From http://cte.illinois.edu/testing/exam/test_questions.html, accessed 27 August 2014.

Centre for Teaching Excellence, University of Waterloo (n.d.). Preparing tests and exams. From <https://uwaterloo.ca/centre-for-teaching-excellence/teaching-resources/teaching-tips/developing-assignments/exams/exam-preparation>, accessed 27 August 2014.

Criswell, J.R. & Criswell, S.J. (2004). Asking essay questions: Answering contemporary needs. *Education*, 124 (3), 510-516.

Evaluation and Examination Service, University of Iowa (n.d.). Preparing and evaluating essay test questions, Technical bulletin 36. From http://www.uiowa.edu/~examserv/resources_fees/Technical_Bulletins/Preparing_and_Evaluating_Essay_Tests_Questions.html, accessed 27 August 2014.

Harvard Kennedy School of Government (2014). Exam procedures and policies. From <http://www.hks.harvard.edu/degrees/registrar/faculty/exams-and-grading/exams>, accessed 27 August 2014.

Instructional Assessment Resources, University of Texas (2001). Rubrics. From <https://www.utexas.edu/academic/ctl/assessment/iar/students/report/rubrics.php>, accessed 27 August 2014.

Jacobs, L.C. (2004). How to write better tests: A handbook for improving test construction skills. From http://www.indiana.edu/~best/pdf_docs/better_tests.pdf, accessed 27 August 2014.

Piontek, M.E. (2008). Best practices for designing and grading exams. Center for Research on Learning and Teaching, University of Michigan.

Reiner, C.M., Bothell, T.W., Sudweeks, R.R., & Wood, B. (2002). Preparing effective essay questions: A self-directed workbook for educators. New Forums Press.

Salkind, N.J. (2013). Essay items. *Tests & Measurement for People Who (Think They) Hate Tests & Measurement*, 2nd edition. Sage Publications.

University of New South Wales (2013). Open book and take home exams. From <https://student.unsw.edu.au/open-book-and-take-home-exams>, accessed 27 August 2014.



List of directive verbs

(from Reiner, Bothell, Sudweeks & Wood, 2002)

- Analyze
- Apply
- Attribute
- Classify
- Compare
- Compose
- Contrast
- Create
- Critique
- Defend
- Define
- Describe
- Design
- Develop
- Differentiate
- Explain
- Evaluate
- Generate
- Identify
- Illustrate
- Infer
- Interpret
- Justify
- List
- Predict
- Propose
- Recognize
- Recall
- Summarize
- Trace

Checklist for scoring essay questions

(adapted from McMillan, 2001)

- Is the model response outlined prior to testing candidates?
- Is the scoring method appropriate?
- Has the role of elements of written communication been clarified?
- Are questions scored one at a time?
- Is the order in which the tests are scored changed?
- Is the identity of the candidate anonymous?

Checklist for constructing essay questions

(adapted from Reiner, Bothell, Sudweeks & Wood, 2002)

- Could the competency/ability be better assessed with a different kind of assessment?
- Is the essay question aligned with the competency or ability being assessed?
- Is the essay question too long and should it be split up into several relatively short essay questions?
- Does the essay question contain a clear and delimited task and a specific problem/situation?
- Is the question worded and structured in such a way that it will be clear to the candidates what they are expected to do?
- Is the task presented to the candidates reasonable?
- Is the problem/situation included in the essay question a novel situation?
- Do the candidates know the time frame for completing their responses?
- Do the candidates know how many points the essay is worth?
- Has the use of optional questions been avoided?
- Has an expected response or an outline of major points that should be included in the response been written and discussed? Is the expected response aligned with the essay question and the competency or ability being assessed?

