How are the Civil Service examinations developed?

The purpose of this document is to inform the reader of the process generally followed in making entry-level or promotional tests. Like most things, there can be slight deviations from a standard process when necessary, and this document will provide an outline of the steps *usually* followed. **Please note that the time estimates can be affected by the competing priorities of the department that we are working with to make the test.**

Before embarking on a labor-intensive journey of developing a promotional or entry-level exam for any position within the City, we – the Test Development and Validation Division – takes specific steps. Here is more information on this process (along with the time it generally takes to complete each step in this process).

1. **Background research (2-3 Weeks):** We conduct thorough research on the position for which we are tasked to develop an examination. This phase helps our team develop an initial list[1] of tasks, knowledge, skills, and abilities (KSAs) for the position. The background-research phase involves the following activities:
    a. Referring to the current job description. Here's an example of an Accountant's job description.
    b. Obtaining past (and the most recent) job analysis information (i.e., task and KSAs)
    c. Studying the departmental structure and the role of the position in that department
    d. Gathering additional information from the Occupational Information Network[2] (O*Net) database
    e. Speaking to the in-house Civil Service personnel who have worked on recruiting and testing for this position in the past

[1] If the initial list is long, we conduct a preliminary review. This helps us remove any redundancies, clean up the task statements/KSAs, etc.
[2] This database contains information on several occupations; it was developed under the sponsorship of the US Department of Labor/Employment and Training Administration and is often used by several public and private organizations.

f. Conducting site visits, interviews with individuals occupying this position (i.e. incumbents), and their supervisors

2. **Job Analysis (4-6 Weeks):** We develop an online survey that includes the tasks and KSAs. For this, we currently use the survey-engine Qualtrics. Then, we administer the job-analysis survey to a certain percentage of incumbents and their supervisors.[3] This phase helps in narrowing down the list of tasks and KSAs for further consideration in the test-development process.
   a. The departmental SMEs rate every task and KSA on specific rating scales. Some examples of ratings scales are: needed-at-entry, frequency, and importance. Generally, KSAs are categorized into broad work behaviors. They are also rated on different rating scales. Here's an example of a list of tasks, KSAs, work behaviors for one of the NOFD positions, along with the job-analysis scales used.
   b. Our goal is to narrow down the list to specific tasks and KSAs that can be used to develop the test content.

3. **Data Analysis (1-2 Weeks):** Once we have gathered SME ratings on every task and KSA, we statistically analyze those ratings to calculate some indices like mean, standard deviation, etc. This phase helps us get closer to the test-specific/job-related tasks and KSAs, based on which the examination content will be designed.
   a. Along with some general statistics, we also establish a cut-off[4] for screening out those tasks and KSAs that were generally rated lower by the SMEs.

[3] They are considered (and will be referred) as the departmental subject-matter experts (throughout this document). We generally require a certain percentage of incumbents and supervisors to serve as our SMEs. Typically, we require them to have more than 1-3 years of experience in the current positions, to have served in different roles within their departments earlier, and they should be considered high performers by the higher management of their departments. Based on these stipulations, the department provides us a list of their employees, and we select a certain percentage of them for different phases in the test-development process. Our selection is based on their demographic representation (i.e., gender and race) and areas of specialty.

[4] The cut-off for shortlisting tasks and KSAs is generally based on a combination of factors. We generally want to shortlist tasks and KSAs that are frequently performed/used and are very important for the job in question, and that have been "endorsed" by the majority of departmental SMEs during the job-analysis.

b. The tasks and KSAs that make our cut-off are then reviewed by another (smaller) panel of Civil Service personnel and departmental SMEs.

c. As a result of this data analysis, the job analysis weights are calculated for each of the shortlisted KSAs and for the broad categories in which the KSAs fit.

4. **Test Plan and Reading List (1-2 Weeks):** Once the shortlisted KSAs and tasks are available to us, we develop a tentative[5] test plan. This phase is generally considered a blueprint of the overall exam. The test plan includes ideas on the method of testing (e.g., multiple-choice test, interview, etc.) which may make up different components of the exam. It also includes the percentage weight for each component in the exam. There are various methods of testing, some of which are listed below along with their advantages and disadvantages.

• **Multiple-Choice Written Test** o Advantage: Ease of administration/scoring large numbers; objective; high validity and utility. o Disadvantage: Reflects real life less well (esp. if not content valid); not worth for few applicants; time-consuming; adverse impact.
• **Structured/Situational Interview** o Advantage: Reasonably valid; good for small numbers of applicants. o Disadvantage: Difficult to administer on large numbers; time-consuming to develop.
• **Assessment Center/Simulation/In-Basket/Work-Sample Test** o Advantage: Good for managerial/supervisory types of jobs; moderate predictive validity; high content validity. Potentially less adverse impact. o Disadvantage: Expensive, difficult and time-consuming to administer to large numbers.

a. Based on the various factors[6], an exam can be designed as either a multiple-hurdle or a compensatory-style exam.

i. A multiple-hurdle exam involves testing candidates in two or three phases. Those who pass the first phase are allowed to appear for

[5] The plan is not final until the Civil Service and the City Department requesting the exam agree upon its feasibility and quality.

[6] These factors generally include consideration for the number of candidates being tested, the location of the position in the organizational hierarchy, availability of departmental funding for the exam, and number of vacancies, to name a few.

the next phase, and so on. The eligible list includes only those candidates that have passed the final phase of the exam.

ii. A compensatory-style exam involves weighted combination of exam components (e.g., 30% multiple-choice test and 70% in-basket). A candidate's total score is based on a total of component scores. Therefore, the eligibility list includes only candidates that have appeared for all components and have scored above an overall cut-off score.

b. During the same time, we develop a list of reading resources for prospective candidates. The list includes books and other departmental documents that candidates are required to study. This reading list and the test plan (once finalized) are sent out in an official information bulletin to all the candidates.

5. **Test Development and Administration Protocol (16-20 Weeks):**[7] Once the tentative test plan is established, we begin developing test items/questions and exercises with the help of departmental SMEs. <u>This is the most time-consuming and crucial phase in the overall test-development process as this involves developing the material that the candidate is exposed to.</u>

   a. Generally, the test-development meetings are held once a week. Every week, roughly five to six departmental SMEs[8] gather to discuss test items. The departmental SMEs are highly involved in developing the rating scale, scoring rubric, and check list of correct responses.

   b. Based on what happens in the meetings, we draft the instructions, test-questions, response options, and other elements of the exam. All of the

[7] It takes longer if the examination includes multiple components, especially a multiple-choice test. Most of the time spent developing a multiple-choice test requires Civil Service personnel to read through department-specific resources to develop the initial pool of questions that SMEs can review.

[8] These SMEs are not allowed to discuss the test content with anyone outside of the meeting or help any potential candidates with the upcoming test. They have been explained the importance of maintaining confidentiality and asked to sign the "Civil Service Confidentiality Agreement" before the test-development process commences.

drafted material is again reviewed by the departmental SMEs. Based on their feedback, the material is revised.

c. At the end of these meetings, we are ready with the exam, instructions for candidates, and the scoring rubric.
d. Once the entire exam is developed, we also create an administration protocol[9] which includes how the exam will be administered.

6. **Pilot Testing and Revision:** After the exam is developed, we normally administer it to a small group of departmental SMEs (generally some incumbents) to find out (a) whether the test instructions and questions are clear and understandable; (b) how long the candidates will take to complete the exam; (c) whether the scoring rubric for the exam needs more information; and (d) if the exam can be administered in a better manner than already planned. <u>This phase serves as a check on what has already been done by us with the help of the departmental SMEs.</u>

7. **Test Administration:** The work is not done after the test has been created. The administration of the test to qualified applicants is done by the members of the Recruitment and Selection Division of Civil Service and additional steps (described below) are required of the Test Development and Validation Division. In some cases, candidates are required to attend a *mandatory* orientation session before the exam is administered. Generally, such orientation sessions are conducted roughly 2-3 weeks prior to the exam. During the administration of exams, the Civil Service staff makes sure that:
   a. the standard instructions are read out to all candidates in the same manner;
   b. their questions about the administration logistics are answered appropriately before the exam begins; and,
   c. candidates are given the same time to complete the examination.

[9] This includes specific instructions to the Civil Service proctors, their assignments, and a plan for how the candidates will be rotated/moved from one phase to the other (if the exam involves more than one component).

In short, the administration protocol is strictly followed to maintain standardization across all candidates.

8. **Post-Examination Activities:** Once the examination is successfully administered, candidates' responses are saved for scoring at a later date (unless they are scored live, on the actual day of the exam). This phase involves activities geared towards getting the scores/ranks out to the candidates and the eligible list out to the concerned department.
    a. For our high-stakes tests (e.g., Fire and Police promotional tests), a panel of external assessors is invited to score their responses.[10] These assessors are from other cities and counties (most likely, from other states) similar in size and scope as the City of New Orleans.
    b. Generally, assessors are grouped in a panel of 2 or 3. A single candidate response is, therefore, scored by such panel of 2 (or possibly 3) assessors.
    c. The assessors are required to be within 1 point of each other on their scores for the same response. In case of disagreements (i.e., difference of greater than a point), assessors are required to come within a point. If they are unable to do so, they are required to provide a thorough explanation of why their scores are more than a point apart.
    d. These scores are later entered into an MS Excel spreadsheet[11] for further data analysis. The total scores of candidates are calculated based on the weights established during Steps# 3 and 4 listed in this document.
    e. Additional important analyses are conducted to determine: (a) a cut-off and overall pass rate, (b) any gender and racial differences, and (c) statistics for each assessor and panel.

[10] In the case of a multiple-choice test, this step is not followed. Such tests are generally scored using a scantron machine.
[11] The data entry is done by one Civil Service employee, and it is checked by another one for accuracy. Sometimes, we enter data twice separately and compare the spreadsheets for any errors.

9. **Civil Service Eligible List and Candidate Feedback:** After the additional analysis, a final eligible list is sent to the concerned department and the candidates are notified of the results.
    a. During this time, all candidates are given an opportunity to sign up for a feedback session.[12] These session are NOT for protesting exam items; they are an opportunity for candidates to learn how they can improve their performance on the areas measured by the exam. We review candidates' areas of strengths and improvement. We also offer candidates general tips on test-taking and study habits during this session.

Note: if you have any questions about the aforementioned process or this document, please contact the Personnel Administrator of the Test Development and Validation Division – Bharati Belwalkar at (504) 658 3508 or bbbelwalkar@nola.gov.

[12] Traditionally, these feedback sessions are heavily attended by failing candidates. However, they are open to passing candidates as well. The reason for offering these sessions is to explore opportunities for further improvement.