

PROTOCOLO DE INTERVENCIÓN PARA LA CONSTRUCCIÓN Y VALIDACIÓN DE
ÍTEMS DE PRUEBAS FUNCIONALES ESCRITAS EN UN CONCURSO DE MÉRITOS.



ESTUDIANTE
ALANIS CUEVAS MAYORGA

ASESOR UNIVERSITARIO
JOHN SOLORZANO RESTREPO

ASESOR DEL SITIO
JONATAN TOVAR RODRÍGUEZ

COMITÉ TRABAJO DE GRADO

UNIVERSIDAD DE CUNDINAMARCA

FACULTAD DE CIENCIAS SOCIALES, HUMANIDADES Y CIENCIAS POLÍTICAS

PSICOLOGÍA

CUNDINAMARCA

2020

Tabla de Contenido

Introducción	3
Planteamiento del problema	5
Generalidades de la Convocatoria del Sector Defensa.....	7
Levantamiento del Proceso General de Construcción y Validación.....	8
Objetivos	13
Síntesis Teórica	14
Definición del Objeto y Modelo de evaluación por competencias	19
Formato de Pruebas de Juicio Situacional (PJS)	21
Errores Frecuentes en la Construcción de Ítems en Pruebas de Desempeño Óptimo con Ítems de Opción Múltiple con Única Respuesta	24
Etapa de Validación de Ítems	32
Plan de Acción	33
Resultados esperados	39
Referencias	42

Lista de Tablas y Figuras

Tabla 1, <i>Proceso general de Construcción y validación</i>	11
Tabla 2, <i>Ficha técnica pruebas funcionales escritas</i>	16
Tabla 3, <i>Primer momento plan de acción protocolo de intervención</i>	35
Tabla 4, <i>Segundo momento plan de acción protocolo de intervención</i>	37
Tabla 5, <i>Tercer momento plan de acción protocolo de intervención.</i>	38
Figura 1, <i>Diagrama de Flujo Proceso de Construcción y Validación de ítems</i>	12

Introducción

Este trabajo se desarrolla en el marco de la realización de la primera Convocatoria pública para asignación de empleos de carrera administrativa del Sector Defensa. Al ser la primera Convocatoria de su tipo, surgen una serie de necesidades especiales que no tienen un referente directo de satisfacción. Por lo tanto, en el ejercicio de las labores derivadas de la Pasantía, resultó necesario iniciar con una fase diagnóstica, durante esta etapa se determinaron el estado en el que se encontraba el proceso de construcción y validación de pruebas funcionales escritas en la Universidad Libre y, las necesidades a ser atendidas.

La presente Pasantía se realiza dentro de un convenio entre la Universidad de Cundinamarca y la Universidad Libre de Colombia, siendo esta última la entidad encargada de cumplir funciones de operador en el actual concurso de meritocracia para el Sector Defensa, bajo la supervisión de la Comisión Nacional del Servicio Civil (CNSC) y de acuerdo con las generalidades del Concurso y la normatividad legal vigente.

La estandarización en los procesos asociados con este tipo de convocatorias resulta ser importante, ya que, estas pruebas son clasificadas como pruebas de altas consecuencias; es decir, que los resultados de los puntajes obtenidos afectan directamente la vida de los evaluados, ya que, determinan la posibilidad de ocupar un cargo de carrera administrativa dentro de la función pública. Por lo tanto, resulta importante tener los instrumentos de evaluación pertinentes para identificar a los aspirantes más idóneos para ocupar las Ofertas Públicas de Empleos de Carrera (OPEC) que, para este caso particular, se encuentran disponibles en el Sector Defensa.

Teniendo en cuenta la relevancia social de este tipo de procesos, resulta pertinente para la Psicología y puntualmente para la Psicometría, como rama de esta disciplina que se encarga del desarrollo e implementación de modelos que posibilitan procesos de evaluación objetiva, brindar soporte teórico y metodológico para asegurar que la asignación de los cargos

de carrera administrativa se dé verdaderamente por los méritos de los participantes. Es por esto, que este Protocolo, tiene como finalidad generar estrategias para mitigar las fuentes de error de medición que provienen de problemas derivados de la construcción y validación de los Ítems en las pruebas funcionales escritas, ya que, se debe garantizar que los instrumentos objetivos desarrollados midan las competencias laborales de los aspirantes bajo procesos estandarizados, de manera tal que, independientemente de quien aplique la prueba llegue a las mismas conclusiones sobre el evaluado; es decir, que las puntuaciones permitan hacer inferencias acertadas sobre los mismos.

En los procesos supervisados por la CNSC se hace uso del Modelo de Evaluación por Competencias bajo el formato de Pruebas de Juicio Situacional para evidenciar las competencias de los aspirantes para desempeñar un cargo dentro de la función pública, por lo tanto, se asume que las respuestas de los evaluados en estas pruebas objetivas son evidencias de la puesta en práctica de sus conocimientos, habilidades, destrezas y valores para desenvolverse en un contexto laboral específico del Sector Defensa.

Como referentes, se tuvieron en cuenta los supuestos de la Teoría clásica de los Test (TCT), ya que, en la TCT o modelo clásico, las puntuaciones observadas son el resultado de la sumatoria entre la puntuación verdadera y el error de medición, por lo que, al implementar estrategias para disminuir el error de medición se asegura que la puntuación verdadera se asemeje a la puntuación observada. Estos errores de construcción que afectan a toda la población pueden ser controlados mediante la implementación de procesos estandarizados debidamente sustentados.

A lo largo de este documento, se desarrolla el plan de acción para la concreción de un Protocolo de intervención que sirva como guía metodológica para el análisis de los Ítems construidos en el marco de la Convocatoria, ya que, en esta guía se establecen algunos criterios para mitigar los errores en los ítems de las pruebas escritas funcionales. En síntesis,

en el desarrollo de esta propuesta, se presentan referentes normativos, teóricos y metodológicos que sustentan la puesta en marcha de acciones que estandaricen y optimicen las labores cotidianas de los diferentes perfiles que participan en el proceso de construcción y validación (profesionales de apoyo, correctores de estilo, constructores, pares validadores y validadores doble ciego).

Se espera que, con la implementación de este Protocolo, se eviten omisiones de los elementos mínimos de los ítems, errores comunes de forma o estilo (lenguaje y redacción del Ítem) y errores de contenido como falta de claridad, pertinencia, suficiencia, coherencia y relevancia. Para que, en suma, estos factores aporten a un funcionamiento psicométrico adecuado de los ítems y tengan una repercusión directa en la disminución de reclamaciones, durante la etapa de reclamaciones, por parte de por parte de los evaluados, garantizando así un proceso de evaluación objetivo que permita seleccionar el personal idóneo para ocupar cargos de la función pública en el Sector Defensa.

Planteamiento del problema

Dado el marco normativo existente para la asignación de Ofertas Públicas de Empleos de Carrera (OPEC), resulta necesario garantizar el adecuado desarrollo de mecanismos e instrumentos de evaluación objetivos que, permitan identificar dentro del grupo de aspirantes a aquellos que resultan ser idóneos para desenvolverse en los diferentes cargos de la función pública. Muestra de esto es que, como lo señala La Constitución Política de Colombia (1991) en su Artículo 130 “habrá una Comisión Nacional del Servicio Civil responsable de la administración y vigilancia de las carreras de los servidores públicos, excepción hecha de las que tengan carácter especial” (p.65).

Así, la Corte Constitucional, a través de la Sentencia 824 de 2013, define el proceso de carrera administrativa como: “El sistema técnico de administración de personal, que garantiza

la eficiencia de la administración pública y ofrece estabilidad e igualdad de oportunidades para el acceso y ascenso al servicio público. Para tal fin, tanto el ingreso como la permanencia de los empleos, debe llevarse a cabo exclusivamente, a partir del mérito, mediante procesos de selección que garanticen la transparencia, la objetividad y la igualdad”.

Por lo tanto, la Comisión Nacional del Servicio Civil (CNSC) es el ente encargado de llevar a cabo los procesos necesarios para la asignación de empleos públicos de carrera mediante la participación abierta en concursos públicos de méritos. Así mismo, la Ley determina que los empleos públicos deben ser de carrera, salvo algunas excepciones, tal y como se describe en la Constitución Política de Colombia (1991) en el Artículo 125, en donde se establece que “los empleos en los órganos y entidades del Estado son de carrera, se exceptúan los de elección popular, los de libre nombramiento y remoción, los de trabajadores oficiales y los demás que determine la ley. Los funcionarios, cuyo sistema de nombramiento no haya sido determinado por la Constitución o la ley, serán nombrados por concurso público” (p.62).

En consecuencia, se reglamentan la Ley 489 de 1998 en donde se establecen normas sobre la organización y funcionamiento de las entidades del orden nacional, además en esta norma se expiden las disposiciones, principios y reglas generales; en la misma línea, la Ley 909 de 2004 expide normas que regulan el empleo público, la carrera administrativa y la gerencia pública.

Para efectos de la presente convocatoria y dadas las características especiales del sector al que se encuentra dirigida, se toma como normatividad específica de referencia a la Ley 1033 de 2006, mediante la cual crea el Sistema Especial de Carrera Administrativa para los empleados públicos no uniformados al servicio del Ministerio de Defensa Nacional, de las Fuerzas Militares, de la Policía Nacional y de sus entidades descentralizadas, adscritas y vinculadas al Sector Defensa.

De la misma forma, se encuentran los Decretos con Fuerza de Ley, 091 y 092 de 2007; los cuales regulan y establecen las reglas de juego para este primer Concurso Público del Sector Defensa; y el Decreto Ordinario 1070 de 2015; el cual recopila la normatividad para el Concurso.

Generalidades de la Convocatoria del Sector Defensa

Para esta Convocatoria se cuentan con un total de 3.427 Ofertas Públicas de Empleos de Carrera (OPEC), 5.473 vacantes, 41.369 inscritos y la participación de 17 entidades con sus respectivos actos administrativos. En ese orden de ideas, los Manuales de Funciones y Competencias Laborales de las entidades participantes reflejan el contenido de la OPEC; determinando los requisitos mínimos para que cualquier ciudadano que cumpla con ellos pueda concursar. De esta manera, la CNSC garantiza los procesos necesarios para la asignación de empleos, mediante la participación abierta en concursos públicos de méritos, dentro del marco normativo correspondiente.

Dadas las consecuencias que este tipo de pruebas tienen sobre la vida de los sujetos, es necesario garantizar las calidades técnicas de los procesos desarrollados para la Convocatoria, como lo exponen Campos, Morales y Guerrero (2016) este tipo de pruebas pueden ser clasificadas como de “alta consecuencia”, dado que, como lo definen “las pruebas estandarizadas producen altas consecuencias cuando sus resultados impactan directamente la vida de los evaluados y las comunidades” (p. 356).

Por otro lado, es necesario considerar que el impacto de las decisiones que se toman con base en las puntuaciones de los evaluados va más allá de consecuencias para los mismos, trascendiendo a un nivel social, ya que, las puntuaciones determinan quienes ejercerán como funcionarios públicos, por lo que los datos derivados de las pruebas objetivas resultan ser una herramienta idónea para la selección de funcionarios.

Por lo tanto, la Universidad Libre como operador de la Convocatoria tiene una gran responsabilidad social y debe ser garante de las calidades técnicas del proceso, teniendo en cuenta que la aplicación de la prueba tiene como objetivo estimar la capacidad, idoneidad y acondicionamiento de los concursantes a los diferentes empleos que se convocan. Entonces, es necesaria la implementación de estrategias para evitar errores de medición, que afecten la precisión de las calificaciones y puedan generar consecuencias negativas derivadas de las decisiones tomadas con base en las puntuaciones.

Para desarrollar estas estrategias, es necesario reconocer el estado actual de los procesos de construcción y validación de las pruebas de la Convocatoria. Para que, por medio del desarrollo de un protocolo de intervención, se estandaricen las acciones asociadas con estos procesos, realizando una revisión a los procesos que se estaban llevando a cabo en busca de optimizarlos de tal manera que, mediante el cumplimiento de estos procedimientos, se logre mitigar la presencia de factores asociados con errores de medición que tienen una incidencia directa en la interpretación de las puntuaciones de los evaluados y las decisiones que se toman en consecuencia, dado que tal y como se establece en el Artículo 19 de la Ley 909 de 2004, el cual deja de manifiesto que la gestión de personal al interior de las entidades del Estado toma un curso que demanda una alta precisión en la evaluación de competencias.

Levantamiento del Proceso General de Construcción y Validación

Como resultado de las entrevistas realizadas a los integrantes del equipo, a continuación, se presenta el estado del Proceso general de construcción y validación de ítems para el momento en el que se inicia la intervención. El proceso general de construcción y validación comienza con la *Validación de ejes temáticos*, esta cuenta con un total de tres procedimientos, se inicia con un análisis realizado por la CNSC en conjunto con las entidades participantes, en este análisis se revisa el contenido funcional de la Ofertas Públicas de Empleo OPEC para identificar las competencias que requieren los participantes para aspirar a

ocupar cada uno de los cargos ofertados, el resultado de este análisis es el insumo para la propuesta inicial de indicadores a medir en las pruebas; el segundo procedimiento, la Universidad Libre realiza una verificación de las agrupaciones e indicadores (aplicación de conocimientos, capacidades y habilidades) a la luz de las OPEC y los manuales de funciones de las entidades, de esta verificación se deriva una contrapropuesta a la CNSC; Finalmente, se recibe la retroalimentación de la CNSC sobre la propuesta de la Universidad, si no existe acuerdo entre las dos entidades, es necesario que se ajuste la matriz hasta la aprobación de esta, ya que, esta aprobación es requisito para iniciar el proceso de construcción y validación.

Con la aprobación en el proceso de Validación de ejes temáticos, se inicia el proceso de *Reclutamiento*, este consiste en reclutar constructores y validadores con experiencia y conocimientos en los indicadores definidos, la idoneidad de los mismos se verifica mediante los criterios que la CNSC dispone para tal fin, como resultado se debe contar con un consolidado de los perfiles a presentar a la CNSC y las hojas de vida del equipo con los respectivos soportes.

En paralelo al proceso de reclutamiento, se desarrolló la *Propuesta de implementación del formato y modelo de evaluación*, en el primer procedimiento de este proceso, el Psicómetra encargado por la Universidad desarrollo una propuesta inicial para la implementación del modelo de evaluación por competencias bajo el formato de Pruebas de Juicio Situacional, en esta propuesta se definen los referentes teóricos metodológicos a utilizar en la construcción de las pruebas (definición conceptual de los atributos a evaluar, definiciones operacionales, modelo de evaluación, formato de ítem a utilizar) ; posteriormente esta propuesta se socializó con los integrantes del Equipo de pruebas para recibir sugerencias de ajuste sobre la misma, previo a la reunión de revisión y aprobación de la Propuesta por parte de la CNSC.

Una vez aprobada la Propuesta de implementación del formato y modelo de evaluación, se da inicio al proceso de *Capacitaciones y asignación de construcción*. Durante este proceso se

realizan las capacitaciones de constructores y validadores para que se familiaricen con el objetivo de la convocatoria, modelo y formato de evaluación. Una vez realizadas las capacitaciones, se procede a realizar la asignación de las responsabilidades de construcción para los constructores y validadores.

La construcción de ítems inicia con la solicitud del profesional de apoyo a pruebas al constructor, para que este construya un paquete inicial de ítems, esta asignación oscila en un rango de 5 a 10 ítems. Luego, los ítems se suben al aplicativo y son revisados por el profesional de apoyo a pruebas, de ser aprobados se programa el taller de validación y se autoriza al constructor para que continúe con la construcción de los ítems que le fueron asignados.

El primer paso en la *Validación* es programar una sesión de taller sincrónico para validar los ítems cargados en el aplicativo, en este espacio se realiza una revisión de estos reactivos, de acuerdo con los criterios técnicos y metodológicos definidos para tal fin, en este espacio participan el profesional de apoyo a pruebas, un constructor, dos validadores y un corrector de estilo. Posteriormente, el validador doble ciego revisa los ítems bajo los criterios de validación para su rol y de ser aprobados pasan a ser almacenados en el Banco de ítems. El proceso de *Validación* finaliza con La auditoría de la CNSC al banco de ítems para verificar las calidades técnicas de los mismos.

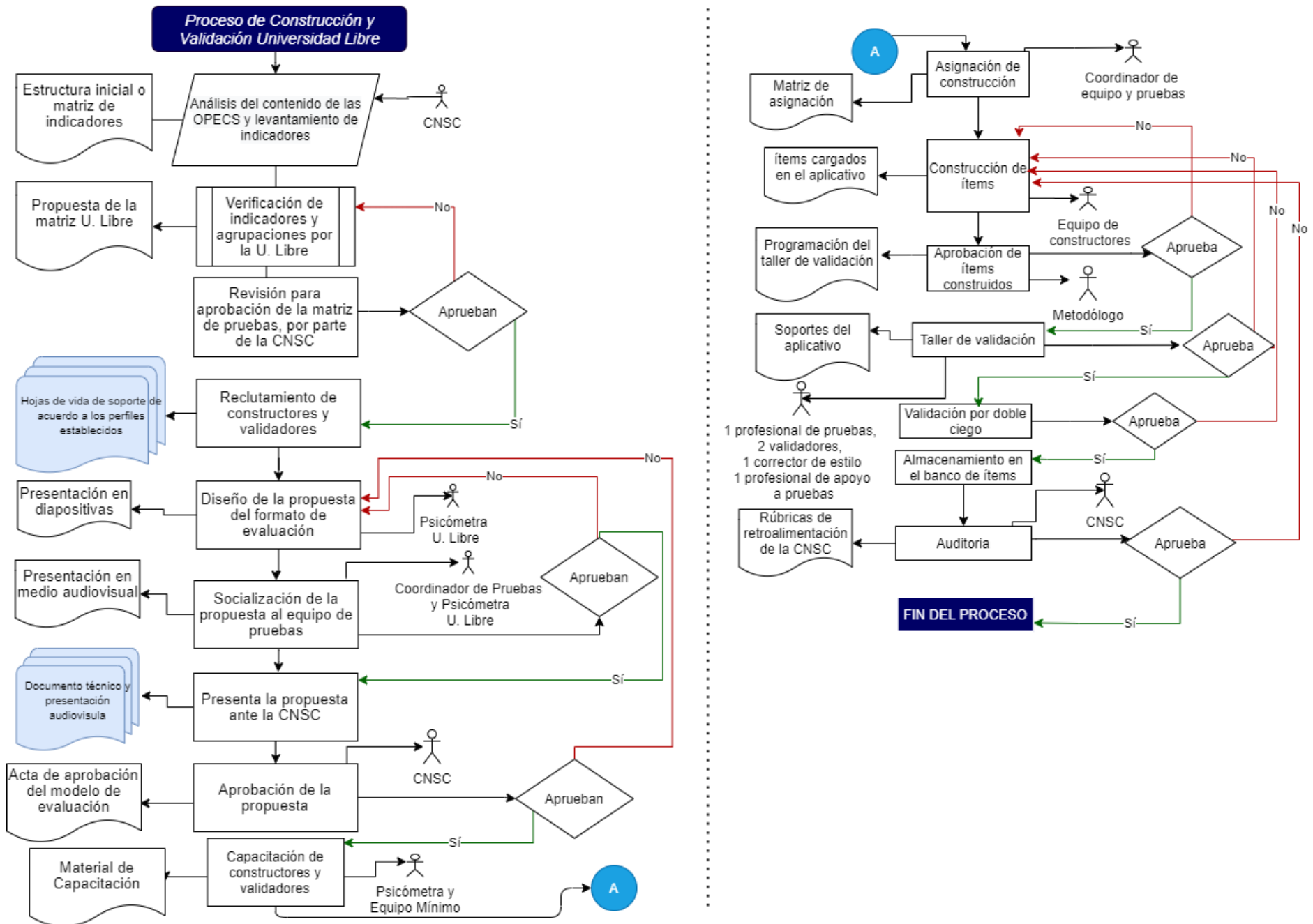
En la Tabla 1, se presentan los diferentes procesos asociados con la construcción y validación de ítems para las pruebas funcionales escritas en la primera Convocatoria del Sector Defensa. En esta Tabla se detallan los procedimientos asociados, su directo responsable y los productos asociados para hacer seguimiento mediante la formulación de indicadores de gestión y desarrollo de métricas, tal y como se sugiere en la Norma ISO 9001.

Tabla 1*Proceso general de Construcción y validación*

PROCESO	PROCEDIMIENTOS	RESPONSABLE	DOCUMENTO
Validación de ejes temáticos	Análisis del contenido funcional de las OPEC y levantamiento de indicadores.	Comisión Nacional del Servicio Civil	Estructura inicial o matriz de indicadores
	Verificación de indicadores y agrupaciones por la U. Libre	Coordinador y equipo de pruebas.	Propuesta de matriz U. Libre.
	Revisión para aprobación de la matriz de pruebas	Comisión Nacional del Servicio Civil	Matriz definitiva y acta de aprobación de la misma.
Reclutamiento	Reclutamiento de constructores y validadores	Coordinador y equipo de pruebas.	Perfiles para contratar y Hojas de vida de soporte
Propuesta de implementación del formato y modelo de evaluación.	Diseño de propuesta	Psicómetra	Presentación inicial en diapositivas.
	Socialización de la propuesta al equipo de pruebas.	Coordinador de prueba y psicómetra	Presentación en medio audiovisual
	Revisión de la propuesta del formato de evaluación	Equipo de Pruebas	
	Presentar la propuesta ante la CNSC.	Equipo de Pruebas	Documento técnico y Presentación audiovisual.
	Aprobación de la propuesta.	Comisión Nacional del Servicio Civil	Acta de aprobación del modelo de evaluación.
Capacitaciones y asignación de construcción	Capacitaciones de constructores y validadores.	Psicómetra y equipo mínimo	Material de capacitación
	Asignación de construcción	Coordinador y equipo de pruebas.	Matriz de asignación
Construcción	Construcción de ítems.	Equipo de constructores	Ítems cargados en el aplicativo.
	Aprobación del auxiliar de pruebas para pasar a validación de los ítems.	Profesional de apoyo a pruebas/Metodólogo	Programación del taller de validación
Validación	Taller de validación.	Un profesional de apoyo a pruebas, un constructor, dos validadores y un corrector de estilo	Soportes del aplicativo.
	Validación doble ciego.	Doble ciego	Soportes del aplicativo.
	Almacenamiento del ítem en el banco de ítems.	Equipo de Pruebas	Soportes del aplicativo.
	Auditoria de la CNSC	CNSC	Rubricas de retroalimentación de la CNSC.

Nota: Elaboración propia con base en la información recopilada en entrevistas con el equipo de la Universidad Libre y acorde con los criterios de la Norma ISO9001 para el levantamiento de procesos.

Figura 1
 Diagrama de Flujo Proceso de Construcción y Validación de ítems



Fuente: elaboración propia con información recopilada en entrevistas con los miembros del equipo y acorde con los lineamientos de la Norma ISO 9001.

Como se puede ver la Figura 1 es una representación gráfica de las secuencia lógica y estructura bajo la cual se debe desarrollar el proceso de construcción y validación de ítems. La implementación del Protocolo de intervención que se desarrolla en esta propuesta recopila los acuerdos del equipo de prueba entorno al formato y modelo de evaluación a implementar en las pruebas escritas y funcionales e, impacta todas las actividades de construcción, validación, auditoria y consolidación del banco de ítems, ya que, en el protocolo para la construcción y validación de ítems, se definen las condiciones y criterios mínimos para asegurar la calidad de los ítems.

Objetivos

Objetivo General

Implementar estrategias que ayuden a minimizar las fuentes de error de medición derivadas de problemas en la construcción de los ítems.

Objetivos específicos

Determinar el estado actual en que se encuentra el proceso de construcción y validación previo a la intervención.

Establecer criterios teóricos y metodológicos para la validación de preguntas en las pruebas escritas funcionales.

Implementar protocolo de intervención con los respectivos equipos de trabajo.

Síntesis Teórica

A modo de contexto general, a continuación, se realizará una descripción de las características de las pruebas funcionales escritas, tomando como referente los criterios de clasificación propuestos por Martínez, Hernández M.V. y Hernández M.J. (2016). En primer lugar, al diseñar una prueba objetiva, se debe considerar cuál es el propósito de esta y si la estrategia utilizada para hacer la medición permite o no, tomar una muestra representativa de lo que se quiere observar, garantizando de esta manera que el instrumento mida lo que pretende medir y que las decisiones tomadas con base en las puntuaciones de los evaluados se encuentren adecuadamente sustentadas.

Según los autores, una prueba objetiva es un instrumento que permite medir variables bajo un proceso estandarizado, estas variables de interés suelen ser conceptos o características latentes para las que no existen herramientas de medición directa como un metro, una balanza o un termómetro; en otras palabras, estas variables de medición indirecta o variables latentes son el resultado de construcciones teóricas, razón por la cual, reciben el nombre de constructos. Entonces, el carácter objetivo de un proceso de evaluación reside en que, en este tipo de procesos se recogen un conjunto de modelos formales que establecen las bases para que la medición de los fenómenos psicológicos, organizacionales y educativos se realicen de forma adecuada. De manera tal que, las decisiones tomadas con base en las puntuaciones de los evaluados deben atender a estándares, a los hechos y a la lógica, minimizando así, los efectos de los sentimientos o creencias del evaluador; es decir, que independientemente de quién realice la evaluación, llegue a conclusiones similares sobre el evaluado.

La intención de este protocolo de intervención es fortalecer las competencias del equipo de trabajo de la Universidad Libre para desarrollar instrumentos de evaluación objetiva; es decir, instrumentos de medición que permiten obtener una muestra de la conducta de los

evaluados en un dominio específico, haciendo uso de un procedimiento estandarizado para la asignación de sus puntuaciones.

El objeto de evaluación dentro de los procesos desarrollados por la Comisión Nacional del Servicio Civil (CNSC) y la Universidad Libre se remite a las competencias laborales, la definición de competencia laboral puede variar de un concurso a otro, dependiendo de la normatividad que aplique para cada concurso, más adelante, se detallará la definición de competencia empleada para esta Convocatoria.

Teniendo en cuenta que, en el contexto de esta Convocatoria, el proceso de evaluación por competencias tiene un propósito clasificatorio o eliminatorio, las pruebas escritas funcionales están orientadas a identificar a aquellos aspirantes que tienen una mayor idoneidad para desempeñarse en el contexto de la función pública dentro del Sector Defensa, la CNSC demanda que los Ítems a construir para la Convocatoria se mantengan en un rango de dificultad media, siendo esta una característica representativa de este tipo de convocatorias. A modo de ejemplo, para explicar porque se utiliza el rango de dificultad medio, se presentan dos pruebas que utilizan un modelo de evaluación por competencias, pero difieren en la dificultad de sus Ítems, ya que responden a diferentes propósitos: el Examen de admisión a la Universidad Nacional y la Prueba Saber 11 ° evalúan competencias básicas de aprendizaje, no obstante, para el primer caso, dado que, el objetivo es identificar dentro de un grupo de aspirantes cuáles son los que muestran mejor desempeño para otorgarles cupos a la Universidad Nacional, el nivel de dificultad de todas preguntas de este examen tiende a ser alto; en contraste, la Prueba Saber 11 ° tiene un propósito diagnóstico, siendo su objetivo proporcionar información sobre el estado de la educación a nivel nacional, por lo que, incluye preguntas de diferentes niveles de dificultad que permiten identificar tanto a aquellos sectores de la población que muestran fortalezas, como a los que muestran un desempeño promedio o dificultades de aprendizaje. En síntesis, las pruebas a utilizar en esta Convocatoria no tienen

un carácter diagnóstico, sino clasificatorio y eliminatorio, por lo tanto, la CNSC determina que las preguntas a incluir deben ubicarse en un rango de dificultad medio.

Partiendo de la propuesta de clasificación de pruebas objetivas realizada por Martínez et al.(2016) y teniendo en cuenta que la definición de competencia tiene un carácter multidimensional, se entiende que en la ejecución de la competencia se involucran diferentes atributos o constructos; es decir, al hacer uso de sus competencias, el evaluado integra los conocimientos, habilidades y capacidades en un saber hacer, ya que, aplica sus conocimientos a la resolución o atención de problemas derivados de un contexto específico, que en este caso está constituido por situaciones derivadas del ejercicio cotidiano de la función pública en el Sector Defensa.

Haciendo uso de la taxonomía propuesta por los autores, en función del área de comportamiento acotado y el tipo de interpretación que se realiza de las puntuaciones, las pruebas funcionales escritas, al evaluar competencias y ser interpretadas a partir de criterios establecidos en la normatividad vigente, son pruebas de dominio o desempeño óptimo que están referidas al criterio, ya que, mediante la normatividad que regula la Convocatoria, se define previamente el desempeño esperado o umbral aprobatorio. En otras palabras, las competencias laborales evaluadas, hacen referencia a un deber ser dentro del contexto de la función pública, sobre el cual se pueden hacer interpretaciones como aprobado o reprobado, acorde con lo que se ha establecido previamente como el desempeño óptimo.

En la misma línea, de acuerdo con la modalidad de aplicación, para esta Convocatoria, las pruebas a las que se refiere este protocolo pueden ser clasificadas como pruebas de aplicación colectiva en formato de lápiz y papel. Acorde con las directrices de la CNSC, estas pruebas deben ser construidas bajo el formato de Pruebas de Juicio Situacional (PJS), de forma tal que, se garantice el realismo de las situaciones planteadas; es decir, que las situaciones

hipotéticas a incluir en las pruebas deben reflejar la cotidianidad del Sector Defensa. En las pruebas funcionales escritas se utilizarán Ítems de opción múltiple con única respuesta que, considerando las demandas temporales de la prueba, buscan evidenciar la potencia de los evaluados y no su rapidez; es decir, que el interés de la evaluación se centra en la calidad de las respuestas de los aspirantes más que en la rapidez con la que responden. No obstante, dada la contingencia generada por la pandemia COVID-19, al momento de la realización de este informe, el mecanismo y fecha en la que se va a llevar a cabo la aplicación de las pruebas es incierto.

Tabla 2

Ficha técnica pruebas funcionales escritas

Nombre de la prueba o dominio evaluado	Pruebas escritas funcionales.
Propósito de la prueba-Alcance de los resultados (Certificación, diagnóstico, clasificación...)	Seleccionar a los aspirantes idóneos para ocupar los cargos de Carrera Administrativa para la Convocatoria del Sector Defensa.
Población objetivo	Aspirantes que completaron la fase de Valoración de Requisitos Mínimos.
Dimensionalidad	Multidimensional
Consecuencias para el sujeto	Altas
Modalidad de aplicación	Colectiva de Lápiz y papel
Tipo de interpretación de las puntuaciones	Criterio definido por la normatividad vigente.
Planteamiento del problema y tipo de respuesta	Opción múltiple con única respuesta
Área del comportamiento acotada	Desempeño óptimo
Demandas temporales (continuo)	Potencia
Modelo estadístico de calificación	TCT

Creación propia con base en la Taxonomía de clasificación de pruebas propuesta por Martínez et al. (2016)

En el marco normativo derivado de la Ley 909 de 2004, la Universidad Libre determina los referentes teóricos y metodológicos para desarrollar la Convocatoria, ajustándose a las

características especiales del Sector Defensa, como población objetivo de este proceso. Dado que el objeto de evaluación de este proceso es determinar las competencias laborales de los evaluados, resulta necesario definir el término competencia, el cual implica el conocimiento y las habilidades concretas, por lo que se entiende que la competencia es inseparable de la acción y exige la aplicación de conocimientos. Por lo tanto, como lo expone Irigoín y Vargas (2002), las competencias están más allá de los conocimientos, porque están ligadas a un dominio específico de la tarea e implican la puesta en juego de esquemas operativos que se han desarrollado en la experiencia de la persona en su labor. Por lo que, al momento de hablar de evaluación de competencias es indispensable que la medición de estas se dé en la práctica de un contexto específico.

Con la intención de garantizar el realismo de las situaciones planteadas, la Universidad Libre adopta la propuesta metodológica que la CNSC utiliza en los diferentes concursos, por lo tanto, desde el formato de Pruebas de Juicio Situacional, se busca que el concursante se exponga a situaciones cotidianas o incidentes críticos del contexto funcional en el que aspira desempeñarse el evaluado. Para la identificación de estas situaciones, se solicita a trabajadores modelo o expertos que identifiquen y describan en detalle, aquellas situaciones que consideren importantes en relación con el logro de objetivos en sus trabajos, destacando tanto los resultados positivos como los negativos, estas descripciones posibilitan que se levante un perfil del cargo, siendo este el insumo para delimitar los comportamientos observables o conductas esperadas en el evaluado, que a la vez, funcionan como predictores confiables y válidos del desempeño futuro del mismo en el empleo al que aspira (Pereda y Berrocal 1999). En síntesis, la evaluación de competencias en contextos realistas debe ser considerada como la evaluación de características individuales que están directamente relacionadas con el desempeño eficiente de las personas en sus puestos de trabajo.

Definición del Objeto y Modelo de evaluación por competencias

En el marco de la presente Convocatoria, la Universidad Libre adopta para el desarrollo de las pruebas, la definición de competencia, tal y como se presenta en el Decreto 1070 (2015), en donde las competencias laborales

Se definen como la capacidad de una persona para desempeñar, en diferentes contextos y con base en los requerimientos y resultados esperados en el sector público y en especial en el sector defensa, las funciones inherentes a un empleo; capacidad que está determinada por los conocimientos, destrezas, valores, habilidades, actitudes y aptitudes que debe poseer y demostrar el empleado público del Sector Defensa. (p.9)

Esta definición es acorde con la perspectiva de competencia como un atributo integral que, empezó a consolidarse y a desarrollarse de manera ininterrumpida desde que McClelland en 1989 afirmó que, para el éxito en la contratación de una persona, no era suficiente con el título que aportaba, ni el resultado de las pruebas psicológicas a los que se les sometía. Este tipo de aproximación ha repercutido en diferentes países de América y Europa, siendo los Estados Unidos, Inglaterra, Francia y, en cierta medida, México, los principales centros de elaboración conceptual y aplicación práctica del modelo de evaluación por competencias (Ruiz, Barrios y Romero, 2005).

Es evidente que, en la actualidad, en los procesos evaluativos las competencias se conciben como actuaciones idóneas, que emergen de la comprensión y el entendimiento de situaciones contextualizadas en un marco cultural (Gonazi & Althanasou, 1996). Por lo tanto, la evaluación por competencias requiere garantizar que el evaluado ponga en juego de manera integral, tanto su dominio de los conceptos del área de conocimiento que se quiere evaluar, como sus capacidades y habilidades para desempeñarse eficazmente en situaciones específicas de trabajo (Ducci, 1997 y Agudelo, 1998); es decir, que en realidad se evalúe la puesta en juego de los

atributos necesarios para que el evaluado se desenvuelva de forma autónoma y flexible en la resolución de problemas ocupacionales (Bunk, 1994). Por lo que, en la evaluación de competencias, estas deben ser entendidas como un atributo teórico práctico, algo que integra el conocimiento técnico referido al *saber hacer*, pero está más allá del mismo (Posada, 2004).

Desde esta aproximación teórica, se sobreentiende que, en la manifestación del repertorio comportamental del evaluado, es posible identificar qué personas tienen mayor dominio para desenvolverse en una situación determinada. Por lo tanto, el desempeño diferencial entre los evaluados puede ser explicado por la interacción de las demandas del medio y las características individuales (Leboyer, 1997); en otras palabras, las características individuales guardan una relación causal con el desempeño efectivo (Ruiz, Barrios y Romero, 2005). Como lo señala Le Boterf (1998) las competencias son una construcción, a partir de una combinación de los recursos individuales y los recursos del ambiente, que se integran para apuntar a un desempeño esperado.

En síntesis, partiendo de la conceptualización presentada sobre evaluación de competencias, la Universidad Libre fija como objetivo, recopilar evidencias suficientes que permitan identificar a los candidatos más adecuados para desempeñarse en los empleos ofertados. Por lo tanto, se asume que las competencias al expresarse en la acción funcionan de manera análoga a la estructura de un iceberg; es decir, que la parte invisible del comportamiento es mucho más densa y voluminosa que el comportamiento visible, de igual forma la parte invisible sostiene la parte visible. Dicho de otra manera, cada acción observable está sustentada en un sustrato integrado por los valores adquiridos en la vida “saber ser”, los conocimientos aprendidos “saber-saber”, las destrezas y habilidades desarrolladas “saber hacer” y la empatía, entendida como un saber relacional e inteligencia emocional “saber convivir” (UNESCO, 1997 como se cita en Delors et. al, 1997). De tal modo, el desempeño observado en los evaluados se puede entender como una expresión integral de comportamientos, que como se ejemplifico con el

iceberg, a pesar de que algunos de ellos no son observables directamente, a partir de las inferencias de evaluación se les puede objetivar, medir, mejorar, acompañar y desarrollar.

En conclusión, tanto en la definición de competencia que de la normatividad se deriva, como en las definiciones tomadas del contexto académico, existe acuerdo en que los conocimientos sobre un dominio específico son insuficientes al momento de explicar los atributos necesarios para desenvolverse eficientemente en un contexto determinado. Es por estas razones, que la CNSC determina el uso del modelo de evaluación por competencias; es decir que, al evaluar la aplicación de conocimientos, estos conocimientos deben apalancarse en una serie de habilidades y capacidades que le permiten al sujeto desenvolverse acorde a las necesidades propias del Sector Defensa. Por lo tanto, las pruebas de la Convocatoria deben reflejar esta relación entre conocimientos, habilidades y capacidades. En otras palabras, las pruebas funcionales miden la competencia del evaluado para aplicar sus conocimientos, en conjunto con las habilidades y capacidades que son requeridas para ejecutar una función laboral determinada por el contenido funcional de las OPEC; es decir, un Saber y un Saber Hacer.

Por otra parte, la prueba de Valores en Defensa y Seguridad apunta a la evaluación de un Saber Ser; es decir, las competencias del evaluado para mantener un comportamiento íntegro en la ejecución de sus labores como funcionario público y sus competencias para comportarse en el marco de la legalidad.

Formato de Pruebas de Juicio Situacional (PJS)

En cuanto a los métodos y herramientas para la evaluación de competencias, como se mencionó anteriormente, la Universidad Libre adopta el modelo de evaluación de competencias utilizado por la CNSC en los diferentes concursos que desarrolla, por lo tanto, se implementó el uso de Pruebas de Juicio Situacional (PJS), las cuales han sido utilizadas por diversas organizaciones, demostrando una efectividad clara para la evaluación y selección de aspirantes, en la medida en que el test se centra principalmente en la interpretación que da el

candidato al simular un contexto situacional específico para cada cargo; la elaboración de un contexto realista permite que el evaluado muestre sus habilidades y capacidades para aplicar su conocimiento, sentido común y experiencia en la interpretación de un contexto particular.

Dado que las PJS dan una muestra de la conducta característica en un puesto de trabajo, para este tipo de pruebas se persigue la consistencia conductual; es decir, que el rendimiento del aspirante durante la prueba de juicio situacional debe ser coherente y predictivo respecto a su desempeño en dicho puesto. Para garantizar la consistencia conductual Motiwidlo (1990, como se citó en Lievens, 2007) sugiere que el desarrollo de las PJS se dé bajo las siguientes etapas:

En la primera etapa, se realiza el análisis del cargo, es decir, se recopilan eventos críticos y situaciones cotidianas, que son proporcionados por Expertos en la Materia en Cuestión (EMC) o tomadas de fuentes de archivo. Luego, estas situaciones son agrupadas por áreas de contenido similar, se eligen guiones representativos para cada área y se editan para que tengan una longitud y un formato similar. En el segundo paso, los EMC generan una o más respuestas a cada situación; ellos son los que indican qué harían con más probabilidad y qué actuación creen que sería la mejor. Los EMC son útiles porque deberían ser capaces de identificar las mejores respuestas y también de generar algunas alternativas menos adecuadas. Una vez recogidas las respuestas alternativas, el creador edita todas las alternativas de respuestas de tal modo que cada ítem (reactivo) disponga de un amplio rango de efectividad laboral. (Lievens, 2007). En el último paso se desarrolla la matriz de soluciones. Aunque en la literatura que concierne a las PJS se discuten varios métodos de puntuación, el enfoque más utilizado es el basado en las opiniones de los expertos y el empírico.

Diferentes investigaciones sustentan el uso de la metodología de PJS en el contexto de procesos de selección para evaluar diferentes atributos individuales y predecir el desempeño futuro de los evaluados. Con relación a la predicción del desempeño futuro, autores como,

Rodríguez y López (2018) en un proceso de selección de auxiliares administrativos del Gobierno Vasco, concluyen que las PJS tienen una alta validez predictiva y generan una adecuada percepción de justicia en cuanto a las decisiones subsecuentes del proceso de selección y el desempeño mostrado por las personas seleccionadas. De manera similar, McDaniel (2001, como se citó en Lievens, 2007) al realizar lo que puede ser denominado el primer metaanálisis en el ámbito laboral (basándose en 95 estudios) concluye que, las PJS basadas en el análisis de puesto mostraron mayor poder predictivo del nivel de desempeño mostrado por los evaluados, en relación con aquellas que no estaban basadas en el análisis de puesto.

Para explicar el poder predictivo de las PJS Motowidlo (1990, como se citó en Lievens, 2007) propone tres supuestos generales, como primer supuesto se plantea que: “el mejor predictor del comportamiento futuro es el comportamiento pasado” (es decir, el principio de consistencia conductual). El segundo supuesto, altamente relacionado, es que las PJS miden las intenciones y metas de los aspirantes (la teoría del establecimiento de metas). Y el tercer supuesto, es que estas evalúan *constructos* tales como conocimiento tácito, inteligencia práctica y capacidad cognitiva general, los cuales han mostrado ser predictores del desempeño laboral importantes, robustos o útiles.

En la misma línea, diferentes autores sostienen que las PJS permiten evaluar una amplia gama de atributos individuales, por ejemplo, Salter & Highhouse (2009) identificaron que las PJS además de ser un método eficiente para evaluar la inteligencia práctica o sentido común, predicen el desempeño futuro en el trabajo de mejor manera que otro tipo de predictores. Así mismo, Bledow & Frese (2009) como se citó en Ambiel et al. (2015) sostienen que además de sus virtudes predictivas del desempeño laboral futuro, las PJS permiten medir iniciativa personal en el trabajo. Para demostrarlo, utilizaron simulaciones mentales de escenario de trabajo como indicadores de su nivel general de iniciativa personal, sus resultados evidenciaron

que la Prueba de Juicio Situacional es un método promisorio para medir iniciativa personal y que su uso puede ser un medio para aumentar la validez de las medidas organizacionales. Finalmente, Motowidlo y Beier (2010, como se citó en Ambiel, 2015) concluyeron que las PJS son una buena herramienta para recopilar tanto información de dominio de conocimientos específicos como de características de personalidad.

Para garantizar que la construcción de las pruebas se dé bajo los parámetros del modelo de evaluación por competencias y las PJS, es necesario focalizar los esfuerzos en la identificación de fuentes de error de medición y así minimizar las consecuencias negativas.

Errores Frecuentes en la Construcción de Ítems en Pruebas de Desempeño Óptimo con Ítems de Opción Múltiple con Única Respuesta

Partiendo de los supuestos de la Teoría clásica de los test (TCT) existe una relación lineal que como según Muñiz (2018) puede expresarse formalmente así: $X = V + e$; es decir que, la puntuación observada en el evaluado representada por X, es el resultado de la sumatoria de la Puntuación verdadera (V) y el error de medición (e), entendiéndose que el error puede ser ocasionado por factores aleatorios que no son controlables o errores de construcción que afectan a toda la población evaluada. En esta línea, el error representa la información irrelevante que distorsiona la calificación de la persona. Por lo tanto, en la medida en que se reduce el error de medición, la puntuación será más precisa, por esta razón, es necesario implementar controles para minimizar el error de medición.

En pro de garantizar la equidad y validez de la medición en el proceso de construcción y validación de pruebas, es necesario, identificar fuentes de información irrelevante; es decir, aquella información que no es objeto de interés y puede distorsionar las interpretaciones de las puntuaciones de los evaluados. Tal y como lo mencionan Clauser & Mazor (1998) citados por Cuevas (2013) y Ackerman (1992); Mellenbergh (1989); Shealy & Stout (1993) citados por Jauregui, Álvarez y Oviden (2000) el error de medición, es una característica no

intencional del ítem que interfiere en la estimación de las puntuaciones, ya que está causado por la intervención de atributos espurios que son medidos junto con el atributo de interés, generando así, distorsiones en la precisión de la medición.

Como lo señala Fidalgo (1996), la probabilidad de acertar o no en un ítem debe estar en función del atributo de interés; es decir, a una mayor competencia del evaluado en un dominio específico, se espera que este tenga una mayor probabilidad de responder correctamente a los ítems que evalúan dicha competencia, esta relación causal entre la magnitud de atributo y la probabilidad de acertar el ítem recibe el nombre de impacto. En caso contrario; se observan sesgos en la medición; es decir, que el error de medición lleva a que la probabilidad de acertar o no al ítem, sea independiente a la competencia del evaluado, siendo atribuible no a diferencias reales entre lo evaluados, sino a diferencias espurias. En otras palabras, como lo exponen Huang, Wilson, & Wang (2016) podemos hablar de errores en la medición por sesgo, cuando evaluados con un mismo nivel de competencia, tienen probabilidades significativamente diferentes de responder correctamente a un ítem determinado.

Para garantizar la precisión de la medición y la validez de las decisiones que se toman con base en las puntuaciones de los evaluados, resulta necesario implementar controles para minimizar las fuentes de errores de construcción que afecten la evaluación. Por lo tanto, a partir de la revisión literaria, a continuación, se presentan algunas fuentes de información irrelevante de frecuente aparición y que deben ser consideradas tanto en el proceso de construcción como en la etapa de validación, la clasificación de estas es arbitraria y solo se hace con fines ilustrativos, puesto que, se entiende que los diferentes tipos de errores son interdependientes y un mismo error podría ser clasificado en más de una categoría.

En su investigación Urdanibia (2009) propone dos grandes categorías que agrupan los errores de medición, la primera categoría se refiere a variables asociadas con la persona evaluada y la segunda hace referencia a errores en el ítem. Esta clasificación es coherente con la propuesta por, Mellenbergh y Kok (1991), citados por Jauregui, Álvarez y Oviden (2000) quienes también proponen dos categorías globales para agrupar las fuentes de error de medición, la primera de ellas recoge elementos que actúan sobre el sujeto (rasgos y variables) y la segunda condensa variables asociados con los ítems (relación de los ítems con el objeto de medición).

La primera categoría propuesta por Urdanibia (2009) abarca elementos propios de la persona, como ansiedad, memoria, aspectos fisiológicos, motivación y algunas características de aprendizaje y desarrollo. Sin embargo, como lo señalan Eslava y Escobar (2012) dichos componentes individuales se enmarcan en el error aleatorio de medición, ya que, las mediciones de estas variables varían de manera no predecible y por lo tanto escapan al control del constructor de la prueba.

En contraste, Urdanibia (2009) señala que la segunda categoría agrupa los errores derivados del instrumento que pueda interferir en el desempeño del evaluado, ya que, no corresponden con el atributo objetivo de medición. Ahora bien, según Eslava y Escobar (2012) este tipo de componentes varían de manera predecible por lo que afectan la medición de todos los casos por igual, por lo que, tienen un carácter sistemático y al ser identificables, permiten que el constructor desarrolle un control sobre los mismos.

Entonces, el primer tipo de fuentes de información irrelevante para el objeto de evaluación, son los errores en la construcción de los ítems. A este respecto, Banks (2006) resalta la importancia de que, en los análisis de fuentes de información irrelevante en el ítem, se indague por características de los diferentes elementos que componen el ítem, por lo que

resulta igualmente importante analizar el caso, los enunciados y las alternativas de respuesta; prestando igual atención tanto a la clave como a los distractores. En la misma línea, Costa & Araújo (s.f.) citados por Cuevas (2013) señalan que características del formato del ítem, el tipo el tipo de texto, el aspecto evaluado y la situación pueden interferir en el proceso de medición. De manera similar, Sireci & Allouf (2003) señalan que variables como la dificultad de las palabras y las oraciones, el tipo de contenido y su formato pueden afectar la precisión de la medición.

Para referirse a los errores presentes en el ítem se pueden establecer dos vías, la primera referida a la forma del ítem, en la cual Haladyna, Downing y Rodríguez (2002) identifican errores orientados al lenguaje y la redacción del ítem; es decir, que el ítem ideal debe contener una gramática adecuada y un vocabulario que se ajuste a los evaluados. Según Soubirón y Camarano (2006) aspectos a considerar en la construcción y validación de ítems son la claridad, brevedad e inteligibilidad; es decir que, las pruebas estén escritas en lenguaje directo y comprensible para el evaluado; de manera tal que, se refieran a un solo contenido.

Para efectos de esta Convocatoria, los constructores en conjunto con los profesionales de apoyo a pruebas, correctores de estilo y validadores deben garantizar que exista:

Uso de lenguaje que se ajusta al nivel: El lenguaje se ajusta al nivel del empleo de los aspirantes; es decir, no presenta precisión en los conceptos empleados, es difícil e innecesario, por lo que, dificulta la comprensión del caso, enunciado y/u opciones de respuesta.

Redacción: Que no se evidencien problemas de redacción en el caso, enunciado y/u opciones de respuesta lo que dificulta su comprensión, como, por ejemplo:

- Presentar información innecesaria o insuficiente,
- Exponer la información en forma de listado,

-Presentar errores de coherencia de género, número y tiempo verbal.

Repetición: Evitar que ocurra repetición, esta se puede presentar cuando se emplean palabras en el caso o enunciado que se repiten en las alternativas de respuesta, sesgando la elección del evaluado o, cuando la respuesta correcta está contenida en el caso o enunciado.

Negación: No se hagan uso de negaciones en el caso, enunciado y/u opciones de respuesta, que interfieren con su comprensión.

Ambigüedad o Vaguedad: Evitar el uso tecnicismos, siglas o conceptos, que no tienen un único significado, haciendo que la posibilidad de respuesta sea amplia y/o ambigua.

Convenciones, símbolos o siglas: Que no se realice un uso inadecuado de convenciones, símbolos o siglas en el caso, enunciado y/u opciones de respuesta que hacen que estos sean inciertos afectando la elección del evaluado.

Desbalance en las alternativas: Se emplean opciones de respuesta que se encuentran balanceadas en términos de: extensión, tipo de lenguaje, tipo de contenido, número de variables, para evitar el sesgo en la elección del evaluado. De manera tal que, los distractores cumplen su función, porque no son fáciles de descartar y no resultan ser más atractivos que la alternativa correcta.

Extensión: Se suministra en el caso, enunciado y/u opciones de respuesta información que es relevante y no genera confusión; en ese sentido, la extensión sugerida para el caso es de 100 palabras y para el enunciado es de 30 palabras.

Interferencia sensorial: El texto, imágenes, tablas o diagramas se presentan en un formato que facilita su comprensión. (disposición, tipo de fuente, tamaño, color...)

La segunda vía propuesta por Haladyna, Downing y Rodríguez (2002) es la referida con el contenido, para asegurar que el ítem no tenga errores de contenido debe ser, claro, pertinente,

relevante y contener la información suficiente para medir la competencia evaluada. Por lo tanto, el equipo de la Universidad Libre debe garantizar:

Realismo: El ítem se ajusta a una situación o necesidad de la cotidianidad laboral y brinda los elementos para tomar cursos de acción.

Atributo: Se observa el objeto de evaluación en alguno de los elementos que conforman el ítem (la situación, enunciado, la opción de respuesta o el formato utilizado).

Claridad: El ítem se comprende fácilmente, es decir, su sintáctica y semántica no son adecuadas.

Pertinencia: El ítem refleja la dimensión, el indicador o el eje temático que está midiendo; es decir, está relacionado con lo que pretende medir.

Agrupación por contexto (independencia local): Evitar que la información contenida en un enunciado o las opciones de respuesta de este, puedan dar luces respecto de la respuesta de otro enunciado, sesgando la elección del evaluado.

Multiclave: Evitar el uso de distractores que sean correctos o parcialmente correctos.

Parafraseo de las alternativas: Evitar que una de las alternativas sea un parafraseo de la otra alternativa, es decir, que dos alternativas de respuesta signifiquen lo mismo.

Alineación: El enunciado esté alineado con el objeto de evaluación (subeje temático, funciones del empleo, propósito del empleo, criterios de verdad)

Articulación: La clave es una solución que se deriva del caso o del enunciado.

Sesgo por formación: Evitar que el contenido del caso o del enunciado pueda estar asociado con funciones que no son propias de la OPEC convocada o de las formaciones académicas solicitadas en esta.

Interferencia emocional: Evitar que el contenido del caso, enunciado y/u opciones de respuesta sea sensible para ciertos grupos de personas y pueda generar respuestas emocionales (ofensa, molestia, controversias, temas políticos o religiosos) que pueden

afectar la habilidad del candidato para responder el enunciado adecuadamente. Por lo tanto, se sugiere evitar contenidos con un alto nivel de crudeza o que haga referencia a la naturaleza política, religiosa, sexual o regional, cuando estos contenidos no son directamente el objeto de evaluación.

Evaluación de conocimientos: Evitar llevar al evaluado a que, cite, defina, describa, enumere, indique, nombre, reconozca o señale; dado que, estos procesos cognitivos suelen estar asociados con la evaluación de memoria o conocimiento.

Incoherencia entre elementos: Garantizar que la situación, enunciado y opciones de respuesta se relacionan entre sí y funcionan como un todo.

En la misma línea el Educational Testing Service (2009) ETS señala que variables físicas de la prueba son fuente de error de medición, esto sucede cuando características del ítem dificultan que el evaluado lo perciba (ayudas inadecuadas de accesibilidad o información innecesaria). De manera similar, este Instituto, identifica como fuentes de error de medición a las variables cognitivas, estas pueden ser consideradas como una fuente de información irrelevante cuando no están relacionados con el propósito de la evaluación y su distribución es inequitativa entre los evaluados (el uso de lenguaje difícil e innecesario, regionalismos y tópicos específicos como aspectos militares, herramientas especializadas, deportes, religión...).

El ETS (2009) también señala que las variables afectivas pueden ser consideradas como fuentes de información irrelevante cuando el contenido del ítem se torna sensible para algunos evaluados y les genera molestias o emociones negativas que pueden interferir con el desempeño del evaluado (la crudeza del material, la presencia de aspectos amenazadores, el apoyo a ideologías o causas particulares). En conclusión, variables que tienen un efecto directo sobre el evaluado pueden afectar la precisión de la medición.

Dentro de la categoría de errores de medición relacionados con el sujeto, Cuevas (2013) sugiere que los ítems pueden estar permeados por sesgos culturales por las siguientes fuentes de información irrelevante: en primer lugar, se encuentra la frecuencia de ocurrencia de ciertas experiencias en cada grupo o la llamada familiaridad; en segundo lugar, se encuentran las variaciones en la epistemología de las culturas o la manera en la que se relacionan con el mundo y; finalmente se encuentran las dificultades en la construcción de los ítems. De manera similar, esta autora identifica cuatro fuentes emergentes de sesgo cultural, los tecnicismos, el colonialismo- la religión-el Estado, los juicios de valor hacia teorías sociales o de mercado y la escuela tradicional.

Sin embargo, como se mencionó previamente, una de las características del modelo de juicio situacional es que las situaciones empleadas en la evaluación deben tener un carácter realista; es decir que sean situaciones cotidianas que se deriven del contexto particular de la función pública en el Sector Defensa. Esta condición de realismo debe entrar en equilibrio para evitar que la familiaridad con ciertos contenidos dentro de una OPEC arroje ventajas de unos evaluados sobre otros. Tal y como lo exponen Tellegen & Laros (2004), una fuente informativa del funcionamiento diferencial de un ítem puede ser la familiaridad que uno de los grupos tenga con el contenido o formato del ítem. Según Luykx, Lee, Mahotiere, Lester, Hart & Deaktor (2007), en situaciones de incertidumbre en las que el evaluado percibe no tener la habilidad necesaria para responder una pregunta, es frecuente que acuda a sus experiencias cotidianas y los elementos del lenguaje que de ellas se derivan (interferencia fonológica, ortográfica o semántica, estructura pragmática, convenciones textuales y género de los textos); estas variables median las respuestas y desde una lógica funcional sirven, entre otros, para dar significado a términos técnicos desde lo cotidiano (nivel semántico).

En síntesis, en el Anexo 1, Protocolo-Guía metodológica para la construcción y validación de ítems se agrupan los elementos mínimos con los que debe contar el ítem y una serie de tips

para el control y seguimientos de errores frecuentes de forma y contenido que fueron identificados en esta síntesis teórica.

Etapa de Validación de Ítems

El método utilizado para la validación de los ítems por parte de la Universidad Libre es el juicio de expertos. Por medio de este método, se busca la obtención de evidencias de validez de contenido; es decir, que mediante la revisión del cumplimiento de los criterios mínimos establecidos para los ítems, se busca garantizar que los ítems construidos respondan al objetivo de evaluación propuesto (Escobar y Cuervo, 2008; Galicia et al., 2017; Lievens, 2007). En concordancia, la Universidad realiza talleres de validación conjunta, en los que están presentes el constructor junto a un equipo de expertos compuesto por un metodólogo o auxiliar de apoyo a pruebas, dos validadores, y un corrector de estilo. Posterior a esto pasan los ítems por la validación de un tercero bajo el rol de doble ciego.

Los criterios utilizados por el equipo de trabajo para la validación de los ítems están sustentados en la propuesta del juicio de expertos de Escobar y Cuervo (2008) y se evalúa la claridad, coherencia y relevancia que cada ítem posee con referencia a los indicadores de aplicación de conocimientos, habilidades y capacidades que se pretenden evaluar, así como la suficiencia del conjunto de ítems construidos. Estos criterios son complementarios a los presentados en el Anexo 1, Protocolo-Guía metodológica para la construcción y validación de ítems, en donde se especifican los elementos mínimos con los que debe contar cada ítem para responder a la construcción de reactivos bajo este modelo, y una descripción de los errores de forma y contenido más comunes en la construcción. A continuación, se realiza una descripción de los participantes del proceso de validación y la descripción del rol que desempeñan en este ejercicio.

Metodólogo / Auxiliar. de apoyo a pruebas: Revisar que los ítems sean construidos acorde al formato establecido para pruebas de juicio situacional y al modelo por competencias; para

esto utilizan el Anexo 1, Protocolo-Guía metodológica para la construcción y validación de ítems, formato para la construcción de ítems de PJS bajo el Modelo de Evaluación por Competencias.

Validadores: Su labor se centra en la validación de contenido de cada ítem para esto se hace uso de los criterios propuestos por Escobar y Cuervo (2008). Además, cumplen con la función de revisar que no se presenten errores frecuentes de contenido descritos en el Anexo 1, Protocolo-Guía metodológica para la construcción y validación de ítems.

Corrector de Estilo: Verifica que los ítems construidos no presenten errores de forma relacionados en el Anexo 1, Protocolo-Guía metodológica para la construcción y validación de ítems.

Doble Ciego: Su labor está centrada en la revisión de validez aparente en el ítem, la función principal de este perfil es verificar que el diseño del ítem establezca una coherencia con la habilidad o competencia a evaluar, siendo este un indicador de validez aparente; es decir, que el ítem aparentemente mida lo que pretende medir.

Plan de Acción

En el siguiente Plan de Acción se relacionan las actividades, responsables y productos desarrolladas como parte del rol de Pasante durante el semestre II-2020 en la Universidad Libre. Las actividades se dividen en tres momentos, cada uno de ellos correspondiente a uno de los objetivos específicos establecidos en la propuesta. Para detallar las fechas en las que se desarrollaron las actividades ver Anexo 2, Cronograma de Actividades Pasantía.

En la Tabla 3, se detallan las actividades relacionadas con el levantamiento de información para la contextualización general del Proceso de Construcción y Validación que es llevado a cabo por la Universidad Libre bajo la supervisión de la CNSC. En este marco, se realizaron reuniones periódicas a lo largo de aproximadamente ocho semanas, de las cuales fueron participes el psicómetra asesor de la U. Libre, la pasante y la practicante de

la Universidad de Cundinamarca. Además, se realizó una revisión de los documentos técnicos utilizados por la Universidad Libre, para determinar así el estado en que se encontraba el Proceso de Construcción y Validación al inicio de las actividades de la Pasantía e, identificar las necesidades de intervención manifestadas por el equipo de trabajo de la Universidad. Como objetivo de este momento se proyectó determinar el estado del proceso de interés y priorizar las necesidades sobre las cuales debería recaer la intervención. El resultado de estas actividades satisface las expectativas fijadas, ya que, con la participación activa de los involucrados, se logró realizar el levantamiento del proceso de construcción y validación con base en los estándares establecidos por la Norma ISO 9001 (acorde con la propuesta del Equipo Vértice, 2010) tal y como se detalla en la Tabla 1 y Figura 1, en donde se realiza una descripción del Proceso de Construcción y Validación con base en la identificación de actividades, encargados y productos asociados a las mismas.

Con relación a las necesidades identificadas, las fuentes consultadas coincidieron en que en el desarrollo del Proceso de Construcción y Validación, algunos de los referentes teóricos y metodológicos utilizados por la Universidad Libre eran de difícil comprensión y manejo o, se prestaban a interpretaciones subjetivas o ambigüedades que, afectaban la objetividad y estandarización del Proceso; en concordancia, se determinó que el protocolo de intervención propuesto debería priorizar la estandarización de los siguientes tópicos: a) las definiciones de cuáles son los elementos mínimos con los que debe contar la estructura de un ítem que evalúa competencias laborales desde el modelo de evaluación por competencias y el formato de PJS, b) el diseño de estrategias para el monitoreo a errores frecuentes de forma y contenido en los ítems y, c) la delimitación de los criterios de validación a monitorear por los diferentes roles participantes en el Proceso.

Tabla 3*Primer momento plan de acción protocolo de intervención.***PLAN DE ACCIÓN PROTOCOLO DE INTERVENCIÓN PARA LA CONSTRUCCIÓN Y VALIDACIÓN DE ÍTEMS****OBJETIVO:** Implementar estrategias que ayuden a minimizar las fuentes de error de medición derivadas de problemas en la construcción de los ítems

NRO	OBJETIVO ESPECIFICO	ENTREGABLE	INDICADOR	META	ACTIVIDAD	RESPONSABLE
1	Determinar el estado en que se encuentra el proceso de construcción y validación previo a la propuesta de intervención.	Protocolo del proceso de Construcción y validación.	Existencia de 1 tabla, 1 diagrama de flujo del Protocolo del proceso de construcción y validación.	100%	1.1 Reunión de inducción con el Psicómetra encargado de las diferentes convocatorias desarrolladas por la Universidad Libre.	Psicómetra y Pasante
					1.2 Reunión periódica de seguimiento semanal para concretar el levantamiento del proceso general de construcción y validación.	Psicómetra y Pasante
					1.3 Reunión periódica de seguimiento semanal para concretar el levantamiento del proceso específico de construcción y validación de pruebas funcionales escritas.	Psicómetra y Pasante
					1.4 Revisión de los estándares de la Norma ISO 9001 para el levantamiento de protocolos.	Pasante
					1.6 Realización de la matriz de actividades del Proceso de construcción y validación de pruebas y realización del diagrama de flujo.	Pasante

En la Tabla 4, se describe el segundo momento desarrollado en el plan de acción, en el cual, con base en las necesidades identificadas en el primer momento de la intervención, en un lapso de aproximadamente ocho semanas de duración, se concretó el protocolo de intervención al que se apuntaba como objetivo central del trabajo de Pasantía. La finalidad de este momento fue atender las necesidades identificadas, partiendo de la experiencia del equipo de la Universidad Libre y el sustento proporcionado por los referentes teóricos y metodológicos sustraídos de la revisión de la literatura. Por lo que, en este segundo momento, se enfatizó en desarrollar herramientas para asegurar las calidades técnicas de los Ítems de prueba, ya que, garantizar su calidad, aporta a lograr la finalidad general del Concurso; es decir, proveer las vacantes disponibles en cargos de carrera administrativa para la función pública del Sector Defensa, tomando como referencia para esta selección, información válida y confiable.

Las expectativas relacionadas con las actividades de este momento fueron alcanzadas en plenitud, dado que, las actividades planeadas se ejecutaron al 100 % y se concretó el Anexo 1, Protocolo-Guía metodológica para la construcción y validación de ítems; es decir, bajo la supervisión del psicómetra encargado y la búsqueda de referentes, en este Protocolo y documentos asociados, se detallan los elementos mínimos a incluir en la estructura general de Ítem, se hace un recuento de los errores de forma y contenido más comunes que se pueden presentar y se facilita el trabajo de seguimiento y monitoreo de las tareas ejecutadas por los participantes en la construcción y validación de ítems.

Al mismo tiempo, la información de soporte de este Protocolo repercutió en un resultado no esperado, ya que, la síntesis teórica fue utilizada para que el equipo de pruebas pudiera ajustar el Manual Técnico de Prueba a algunos requerimientos de la CNSC, aportando al soporte técnico y metodológico de las actividades desarrolladas en esta Convocatoria.

Tabla 4

Segundo momento plan de acción protocolo de intervención.

PLAN DE ACCIÓN PROTOCOLO DE INTERVENCIÓN PARA LA CONSTRUCCIÓN Y VALIDACIÓN DE ÍTEMS						
OBJETIVO: Implementar estrategias que ayuden a minimizar las fuentes de error de medición derivadas de problemas en la construcción de los ítems						
NRO	OBJETIVO ESPECIFICO	ENTREGABLE	INDICADOR	META	ACTIVIDAD	RESPONSABLE
2	Establecer criterios teóricos y metodológicos para la validación de preguntas en las pruebas funcionales escritas.	Síntesis teórica y Protocolo guía metodológica para la construcción y validación de ítems y, Manual Técnico de pruebas ajustado para revisión de la CNSC. .	Existencia de 1 documento de Síntesis teórica y 1 Protocolo guía metodológica para la construcción y validación de ítems y, 1 Manual Técnico de pruebas ajustado para revisión de la CNSC.	100%	2.1. Revisión de la literatura para generar la síntesis teórica.	Pasante
					2.2. Generación de protocolo metodológico a emplear para la construcción y validación de ítems	Psicómetra y Pasante
					2.3 Revisión y ajuste del Manual Técnico de pruebas	Psicómetra y Pasante
					2.4 Reunión periódica con el asesor en sitio para seguimiento de documentos técnicos.	Pasante

Finalmente, en la Tabla 5 se describe el tercer momento, cuyo objetivo principal fue iniciar la implementación del Anexo 1, Protocolo-Guía metodológica para la construcción y validación de ítems, para ello, durante las cinco semanas previas a la presentación de este informe, se realizó la preparación del material para los ejercicios de capacitación teórico prácticos y las sesiones iniciales de talleres de validación. Cabe aclarar que algunas de las actividades planeadas para este momento, a la fecha de entrega del informe aún se encuentran en ejecución, dado el cronograma de Construcción y Validación, no obstante, la implementación del Protocolo se viene desarrollando bajo diferentes modalidades, ya que, además de ser entregado a cada uno de los participantes del Proceso, los referentes teóricos y metodológicos del mismo fueron incorporados en el material de capacitación y la

Universidad Libre tomó la decisión de incorporar los criterios de revisión en la interfaz del aplicativo utilizado para la construcción de las pruebas.

Tabla 5

Tercer momento plan de acción protocolo de intervención.

PLAN DE ACCIÓN PROTOCOLO DE INTERVENCIÓN PARA LA CONSTRUCCIÓN Y VALIDACIÓN DE ÍTEMS						
OBJETIVO: Implementar estrategias que ayuden a minimizar las fuentes de error de medición derivadas de problemas en la construcción de los ítems						
NRO	OBJETIVO ESPECIFICO	ENTREGABLE	INDICADOR	META	ACTIVIDAD	RESPONSABLE
3	Implementación de protocolo de intervención con los diferentes equipos de trabajo	Material de capacitación	Insumos para la preparación del material de presentaciones, y listado de asistencia de los convocados a las capacitaciones realizadas (ver Anexo 3).		3.1. Preparación del material para las capacitaciones teórico-prácticas al equipo de profesionales de apoyo.	Psicómetra y Pasante
					3.2. Socialización del material desarrollado en las capacitaciones teórico-prácticas al equipo de profesionales de apoyo.	Capacitadores y equipo mínimo
					3.3. Preparación del material para las capacitaciones teórico-prácticas a constructores.	Psicómetra y Pasante
					3.4. Socialización del material desarrollado en las capacitaciones teórico-prácticas a constructores.	Capacitadores y equipo mínimo
					3.5. Preparación del material para las capacitaciones teórico-prácticas a validadores.	Psicómetra y Pasante
					3.6. Socialización del material desarrollado en las capacitaciones teórico-prácticas a validadores.	Capacitadores y equipo mínimo
					3.7 Reunión periódica con el asesor en sitio para seguimientos de la implementación y capacidad instaurada	Psicómetra y Pasante

Con relación a los objetivos alcanzados en este momento del plan de acción, desde una mirada cualitativa, en las actividades de socialización, se recibió una retroalimentación positiva por parte del equipo mínimo de pruebas, constructores y validadores, enfatizando en la practicidad del insumo para dar a conocer la estructura del ítem y las características que debe incluir para asegurar sus calidades técnicas. Finalmente, tal y como se desarrolla en la siguiente sección, se espera que, en el corto plazo, este Protocolo se utilice como insumo a la hora de validar y verificar los Ítems en el aplicativo.

Resultados esperados

Durante las actividades de socialización del Anexo 1, Protocolo-Guía metodológica para la construcción y validación de ítems y, herramientas asociadas, varios integrantes del equipo de trabajo manifestaron de manera espontánea que el uso de este insumo ha facilitado sus labores en el desarrollo de las actividades de construcción y validación, dado que, mediante esta herramienta y la estandarización de procesos, se unificaron criterios y conceptos minimizando las ambigüedades e interpretaciones subjetivas de cómo se deben desarrollar las tareas asignadas a cada rol, para conocer el listado de participantes en las actividades de socialización, ver Anexo 3, Listados de asistencia Capacitaciones.

En primer lugar, se espera que se realice una revisión metodológica de todos los ítems construidos para las pruebas funcionales escritas de la Convocatoria, se espera que esta revisión se haga con base en los parámetros definidos en Anexo 1, Protocolo-Guía metodológica para la construcción y validación de ítems, esto se garantiza por medio de la inclusión de estos criterios en el aplicativo que se utiliza para el cargue de los Ítems. De esta manera, se asegura que, dentro del sistema utilizado para guardar la cadena de seguridad, aparezcan los criterios establecidos para cada uno de los roles participantes en el proceso y que su uso sea obligatorio; es decir, que se dé una estandarización del proceso.

En la misma línea, se espera que con el uso del Protocolo y el material guía, se facilite el trabajo de cada uno de los roles participantes, dado que, gracias a su estructura práctica y fácil acceso, proporciona una ventaja en la dinámica misma de los talleres de validación, teniendo en cuenta que dicho proceso no se encontraba totalmente estandarizado. Por lo tanto, con la estandarización de este, se espera que el material proporcionado sirva como guía de criterios que agilicen tanto la construcción, como la revisión de las calidades técnicas de los Ítems.

En consecuencia, se pretende que dichas calidades técnicas sirvan como pilar fundamental para la aprobación de los ítems en las auditorías realizadas por parte de la Comisión Nacional del Servicio Civil; es decir, que durante el proceso de auditoría muestral el comportamiento de los ítems aporte a que no sea necesario que la CNSC realice una auditoría censal del banco de ítems. Estas auditorías inician cuando se completa el 10 % del total de la construcción de los ítems y de su resultado depende la aprobación del ensamblaje. Por lo tanto, se espera que con la implementación del Anexo 1, Protocolo-Guía metodológica para la construcción y validación de ítems, se minimice la cantidad de Ítems rechazados por la CNSC.

De manera similar, se espera que, en la fase de reclamaciones, se evidencie una disminución en la tasa de reclamaciones relacionadas con errores de forma o de contenido, ya que, los evaluados pueden solicitar acceso para conocer las preguntas y respuestas de la prueba que ejecutaron, razón por la que se espera que se presente un mínimo número de reclamaciones que correspondan con errores de construcción. Así mismo, en relación con el análisis psicométrico, se espera que, en la etapa de calificación, la tasa de Ítems eliminados de la calificación no supere el umbral establecido por la Comisión Nacional del Servicio Civil que, representa un máximo del 30 % de los Ítems construidos para cada una de las pruebas funcionales escritas.

Finalmente, se espera que este Protocolo sea insumo inicial de trabajo para futuros procesos adelantados por la Universidad Libre y la CNSC; es decir, de acuerdo con necesidades particulares y con los ajustes necesarios, este Protocolo y herramientas relacionadas pueden ser adaptados para la estandarización de futuros procesos e, incluso utilizados en el desarrollo de planes de acción, métricas e indicadores de gestión.

Referencias

- Agudelo S. (1998) *Certificación de competencias laborales*. Aplicación en Gastronomía. Montevideo: Cinterfor/ OIT.
- Ambiel, R., Campos, M., Alves, B., y Silva, C. (2015). Fundamentos de aplicabilidad de dos Testes de Juzgamiento Situacional contexto da Psicología Organizacional. *Revista Psicologia*. Organizações e Trabalho, 15(3), 298-308.
- Banks, K. (2006). A comprehensive framework for evaluating hypotheses about cultural bias in educational testing. *Applied Measurement in Education*, 19(2), 115-132.
- Bunk G.P. (1994) La transmisión de las competencias en la formación y perfeccionamiento profesionales en la RFA. *Revista CEDEFOP*; 1 (Nd): 29-54.
- Campos, M, Morales, J y Guerrero, P. (2016). *Efectos indeseados de la medición de la calidad educativa en Chile*. La respuesta de la sociedad civil. Cuadernos Cedes, 36(100), 355-374.
- Colombia, *Constitución Política*. (1991). Bogotá, Colombia.
- Colombia, Presidencia de la República, *Decreto 091* (2007).
- Colombia, Presidencia de la República, *Decreto 092* (2007).
- Colombia, Presidencia de la República, *Decreto 1070* (2015).
- Congreso de la República de Colombia, *Ley 489* (1998). D.O 43.464.
- Congreso de la República de Colombia, *Ley 909* (2004). D.O 45.680.
- Congreso de la República de Colombia, *Ley 1033* (2006).
- Corte Constitucional República de Colombia, *Sentencia 824* (2013).
- Cuevas, M. L. (2013). *Sesgo cultural en los ítems de las pruebas del examen Saber 11 en Colombia*. Departamento de Psicología.
- Delors, J., Amagi, I., Carneiro, R., Chung, F., Geremek, B., Gorham, W., & Stavenhagen, R. (1997). *La educación encierra un tesoro: informe para la UNESCO de la Comisión Internacional sobre la Educación para el Siglo Veintiuno*. UNESCO.
- Ducci M.A. (1997) El enfoque de competencia laboral en la perspectiva internacional. *Formación basada en competencia laboral*. Montevideo: Cinterfor/OIT.
- Escobar, J. y Cuervo, M. (2008). Validez de contenido y juicio de expertos: una aproximación a su utilización. *Revista Avances en Medición* 10.
- Eslava, J., y Escobar, F. (2012). Error aleatorio, sesgo y fraude en las publicaciones científicas. *Revista Colombiana de Anestesiología*, 40(2), 91-94.
- Educational Testing Service. (2009). *ETS guidelines for fairness review of assessments*.
- Equipo Vértice. (2010). *Gestión de la calidad (ISO 9001/2008)*. Editorial Vértice.
- Fidalgo, A.M. (1996). *Funcionamiento diferencial de los ítems*. En J. Muñiz (Coord.), *Psicometría*, (pp. 371-455), Madrid: Universitas

- Galicia, L., Balderrama J., Edel R. (2017) *Validez de contenido por juicio de expertos: Propuesta de una herramienta virtual*. *Apertura* (Guadalajara, Jal.), 9(2), 42-53. <https://doi.org/10.32870/ap.v9n2.993>
- Haladyna, T., Downing, S. & Rodriguez, M. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied measurement in education*, 15(3), 309-333.
- Huang, X., Wilson, M., & Wang, L. (2016). Exploring plausible causes of differential item functioning in the PISA science assessment: language, curriculum or culture. *Educational Psychology*, 36(2), 378-390.
- Irigoin, M., & Vargas, F. (2002). *Competencia laboral: manual de conceptos, métodos y aplicaciones en el sector salud* (pp. 252-252). Oficina Internacional del Trabajo, CINTERFOR.
- Jauregui, A., Álvarez, E., & Oliden, P. (2000). Desarrollos didácticos y funcionamiento diferencial de los ítems. Problemas inherentes a toda investigación empírica sobre sesgo. *Psicothema*, 12(2), 198-202.
- Lievens, F. (2007). Tests de Juicio Situacional: Introducción y revisión de la investigación. *Revista de Psicología del Trabajo y de las Organizaciones*, 23(1), 93-110.
- Le Boterf, G. (1998). Évaluer les compétences. Quels jugements? Quels critères? Quelles instances. *Education permanente*, 135(2), 143-151.
- Leboyer L. (1997) *Gestión de las competencias*. Barcelona: Gestión 2000.
- Luykx, A., Lee, O., Mahotiere, M., Lester, B., Hart, J., & Deaktor, R. (2007). *Cultural and home language influences on children's responses to science assessments*. *Teachers College Record*, 109(4), 897-926.
- Martínez, A., Hernández M.J. y Hernández M.V. (2016). *Psicometría*. Alianza Editorial.
- McClelland, D. C. (1989). *Estudio de la motivación humana* (Vol. 52). Narcea Ediciones.
- Muñiz, J. (2018). *Introducción a la psicometría*. Difusora Larousse-Ediciones Pirámide.
- Pereda, S. y Berrocal, F (1999). *Gestión de recursos humanos por competencias*. Madrid: Centro de Estudios Ramón Areces.
- Posada, Á. (2004) *Formación superior basada en competencias, interdisciplinariedad y trabajo autónomo del estudiante*. Facultad de Educación, Universidad del Atlántico, Colombia. Disponible en: https://www.ugr.es/~ugr_unt/Material%20M%F3dulo%205/competencias_univ.pdf
- Ruiz de Vargas M, Barrios B y Romero L. Competencias laborales y la formación universitaria. *Psicología desde el Caribe. Revista Universidad del Norte* 2005; 16 (n.d): 64-91.
- Rodríguez, A., y López, J. (2018). Predictores selectivos en el sector público: validez predictiva y reacciones de los aspirantes. *Journal of Work and Organizational Psychology*, 34, 16-2.
- Salter, N., y Highhouse, S. (2009). Assessing managers common sense using situational judgment tests. *Decision*, 47(3), 392-398.

- Sireci, S. G., & Allalouf, A. (2003). Appraising item equivalence across multiple languages and cultures. *Language testing*, 20(2), 148-166.
- Soubirón, E., & Camarano, S. (2006). Diseño de pruebas objetivas. Montevideo: *Unidad académica de Educación Química*.
- Tellegen, P. J., & Laros, J. A. (2004). Cultural bias in the SON-R test: Comparative study of Brazilian and Dutch children. *Psicologia: Teoria e Pesquisa*, 20(2), 103-111.
- Urdanibia, A. (2009). Error de medición. *Revista Psicología*, 28(1), 37-55.



CONVOCATORIA SECTOR DEFENSA
Protocolo-Guía metodológica para la construcción y validación de ítems



Objetivo: Identificar los errores de forma y contenido más comunes que se presentan en el proceso del diseño de ítems construídas bajo el modelo de juicio situacional, facilitando la construcción y validación de los ítems.

Instrucciones generales: A continuación, se relacionan los elementos mínimos del ítem y una serie de errores comunes que se presentan en los casos, enunciados y/u opciones de respuesta, diseñados bajo el modelo de juicio situacional; para evitar que se presenten este tipo de errores, deberá emplear la siguiente guía, de manera que, se garantice la calidad de los ítems durante su proceso de construcción y validación.

ELEMENTOS MÍNIMOS DEL ÍTEM

Situación	Tenga en cuenta que el caso debe estar escrito en tercera persona e incluir el rol; es decir, el sujeto dentro del caso y las acciones a realizar; las cuales deben ser acordes al nivel jerárquico evaluado (asistencial, técnico, profesional o asesor), la agrupación de las OPEC y el eje temático.
Problema o necesidad	Eventos crítico(s): es una situación problemática o necesidad que deriva del contexto de la función pública, esta situación hipotética debe tener una posibilidad de solución o mejora que no resulte ser obvia. El evento crítico debe estar acompañado de una serie de herramientas o restricciones que lleven al aspirante a valorar y tomar decisiones en torno al proceso o acción idónea para atender la situación. -El caso, idealmente, debe contener un máximo de 100 palabras. -Dentro de un mismo caso pueden existir varios incidentes críticos, tenga en cuenta que estos deben estar relacionados entre sí. -Los eventos críticos deben contener información de referencia que no esté dispuesta en forma de listado. -De un incidente crítico se pueden construir varios enunciados. -El caso debe iniciar con un contexto general. -Refleja un problema típico relacionado con la OPEC para la que participa el evaluado.
Enunciado	El enunciado es una afirmación que se hace sobre el desempeño esperado en el evaluado, esta afirmación se hará en tercera persona y se debe evitar el uso de verbos que sugieran preferencias o intereses (desea, elige, selecciona, considera, piensa, cree, entre otros); tenga en cuenta que, esta afirmación (enunciado) se encuentra parcialmente completa y conforma una unidad gramatical con cada una de las alternativas de respuesta (clave y respuestas incorrectas). Esta afirmación es un deber ser, por lo tanto, en su elaboración debe incluir expresiones tales como: -Por lo tanto, él debe -Teniendo en cuenta lo anterior, él tiene que -Con base en lo anterior, lo pertinente es, -Según la situación expuesta, lo adecuado es. NOTA: El enunciado, idealmente, debe contener un máximo de 30 palabras.
Criterio de verdad	El criterio de verdad justifica que la clave sea la única de las alternativas que es correcta, sustentándola en la normatividad vigente, documentos técnicos, normas técnicas, manuales, cartillas u otros tipos de información de acceso público que validen la respuesta correcta. En lo referente al criterio de verdad, tenga en cuenta que: -Pueden ser criterios de verdad el más eficiente, económico, oportuno, efectivo, eficaz entre otros. -El criterio de verdad puede estar ubicado en el caso o en el enunciado.
Proceso cognitivo (Habilidad- capacidad)	Procure que la respuesta induzca a procesos cognoscitivos complejos de manera que se busque en las opciones de respuesta en las que se realice un proceso de aplicación, explicación, justificación, demostración, clasificación, análisis, estimación, formulación, conclusión, síntesis o similares. Recuerde que estas pruebas buscan evaluar competencias y no conocimientos. Por lo tanto, los ejes de aplicación de conocimientos deben estar apalancados en una de las capacidades o habilidades del catálogo de competencias de la CNSC.
Alternativa	La alternativa es el complemento gramatical del enunciado, tenga en cuenta que: -Las tres alternativas deben apuntar en la misma dirección o todas en diferentes direcciones; se debe evitar que dos alternativas sean muy similares y la otra sea claramente distinta. -No se utilicen expresiones como, todas las anteriores o ninguna de las anteriores. -Evite utilizar negaciones (no, ninguno, nunca) o generalizaciones (todos, siempre) -Las opciones de respuesta se deben encontrar balanceadas en términos de extensión, tipo de lenguaje, tipo de contenido y número de variables,
Justificación	La Justificación debe estar ligada tanto al criterio de verdad, como a la competencia evaluada y ser un argumento suficiente que permita sustentar por qué la alternativa es correcta o no. Esta debe estar acompañada de la respectiva referencia en formato APA.

LOS ERRORES DE FORMA MÁS COMUNES

Uso de lenguaje que no se ajusta al nivel	El lenguaje no se ajusta al nivel del empleo de los aspirantes; es decir, no presenta precisión en los conceptos empleados, es difícil e innecesario, por lo que, dificulta la comprensión del caso, enunciado y/u opciones de respuesta.
Redacción	Se evidencian problemas de redacción en el caso, enunciado y/u opciones de respuesta lo que dificulta su comprensión, como por ejemplo: -Presentar información innecesaria o insuficiente. -Exponer la información en forma de listado. -Presentar errores de coherencia de género, número y tiempo verbal.
Repetición	Ocurre repetición cuando: -Se emplean palabras en el caso o enunciado que se repiten en las alternativas de respuesta, siguiendo la elección del evaluado. -La respuesta correcta está contenida en el caso o enunciado.
Negación	Se hace uso de negaciones en el caso, enunciado y/u opciones de respuesta, que interfieren con su comprensión.
Ambigüedad o Vaguedad	Se emplean tecnicismos, siglas o conceptos, que no tienen un único significado, haciendo que la posibilidad de respuesta sea amplia y/o ambigua.
Convenciones, símbolos o siglas	Se realiza un uso inadecuado de convenciones, símbolos o siglas en el caso, enunciado y/u opciones de respuesta que hacen que estos sean inciertos afectando la elección del evaluado.
Desbalance en las alternativas	Se emplean opciones de respuesta que se encuentran desbalanceadas en términos de: extensión, tipo de lenguaje, tipo de contenido, número de variables, siguiendo la elección del evaluado. De manera tal que, los distractores no cumplen su función, por que son fáciles de descartar o resultan ser más atractivos que la alternativa correcta.
Extensión	Se suministra en el caso, enunciado y/u opciones de respuesta información que no es relevante y puede generar confusión; en ese sentido, la extensión sugerida para el caso es de 100 palabras y para el enunciado es de 30 palabras.
Interferencia sensorial	El texto, imágenes, tablas o diagramas se presentan en un formato que obstaculiza su comprensión, (disposición, tipo de fuente, tamaño, color...)

LOS ERRORES DE CONTENIDO MAS COMUNES

Realismo	El ítem no se ajusta a una situación o necesidad de la cotidianidad laboral y no brinda los elementos para tomar cursos de acción.
Atributo	No se observa el objeto de evaluación en alguno de los elementos que conforman el ítem (la situación, enunciado, la opción de respuesta o el formato utilizado).
Claridad	El ítem no se comprende fácilmente, es decir, su sintáctica y semántica no son adecuadas.
Pertinencia	El ítem no refleja la dimensión, el indicador o el eje temático que está midiendo; es decir, no está relacionado con lo que pretende medir.
Agrupación por contexto (independencia local)	La información contenida en un enunciado o las opciones de respuesta de este, pueden dar luces respecto de la respuesta de otro enunciado, siguiendo la elección del evaluado.
Multiclave	Los distractores son correctos o parcialmente correctos.
Parfraseo de las alternativas	Una de las alternativa es un parfraseo de la otra alternativa, es decir, que dos alternativas de respuesta significan lo mismo.
Alineación	El enunciado no está alineado con el objeto de evaluación (subje temático, funciones del empleo, propósito del empleo, criterios de verdad)
Articulación	La clave no es una solución que se deriva del caso o del enunciado.
Sesgo por formación	El contenido del caso o del enunciado puede estar asociado con funciones que no son propias de la OPEC convocada o de las formaciones académicas solicitadas en esta.
Interferencia emocional	El contenido del caso, enunciado y/u opciones de respuesta es sensible para ciertos grupos de personas y puede generar respuestas emocionales (ofensa, molestia, controversias, temas políticos o religiosos) que pueden afectar la habilidad del candidato para responder el enunciado adecuadamente. Por lo tanto, se sugiere: -Evitar contenidos con un alto nivel de crudeza o que haga referencia a la naturaleza política, religiosa, sexual o regional, cuando estos contenidos no son directamente el objeto de evaluación.
Evaluación de conocimientos	Evite llevar al evaluado a que, cite, defina, describa, enumere, indique, nombre, reconozca o señale; dado que, estos procesos cognitivos suelen estar asociados con la evaluación de memoria o conocimiento.
Incoherencia entre elementos	Situación, enunciado y opciones de respuesta no se relacionan entre sí y no funcionan como un todo.