

ISSN: 2711-175X (en línea)
No. 15 de 2022 (III trimestre)
Boletín SIEA
Santiago de Cali



SIEA SISTEMA INSTITUCIONAL DE
EVALUACIÓN DE LOS
APRENDIZAJES



Editora

Sonia Cadena Castillo
Universidad Autónoma de Occidente
Colombia

Comité Editorial

Claudia Alexandra Roldán Morales
Universidad Autónoma de Occidente
Colombia

Oswaldo Rodríguez Díaz
Universidad Autónoma de Occidente
Colombia

Dulfay Astrid González Jiménez
Universidad Autónoma de Occidente
Colombia

Fabio Jurado Valencia
Universidad Nacional de Colombia
Colombia

Daniel Bogoya Maldonado
Consultor Nacional e Internacional
Colombia

Autores: Luz Estela Muñoz, Orlando Arboleda, Jesús Antonio Lemos

Apoyaron la realización del Boletín Nro. 15:
Dulfay Astrid González y Sonia Cadena

Diseño: Pablo Andrés Sánchez

Página web:

<https://sitios.uao.edu.co/docentes/boletines-siea>

BOLETÍN DEL SISTEMA INSTITUCIONAL DE EVALUACIÓN DE LOS APRENDIZAJES

El Boletín del SIEA es un medio de divulgación de resultados de investigación, con periodicidad trimestral y está incluido en bases bibliográficas.

La importancia del pilotaje en la prueba de Algoritmia y Programación. ¿Qué aporta?

Presentación

En el proceso de diseño, construcción y validación de las pruebas estandarizadas del SIEA, el pilotaje juega un papel fundamental. Permite monitorear tanto el funcionamiento general del instrumento en una muestra de participantes, cuyas características son similares a las de la población objeto de interés, como estimar el grado de habilidad y de verosimilitud, entre otros. Tal es el caso de la prueba de Algoritmia y Programación, que después de 3 años de no aplicarse, se retomó en el año 2022, gracias a un sostenido trabajo en equipo por parte de los profesores del área, que focalizaron su atención en la actualización de la tabla de especificaciones, en la construcción de nuevos bloques y en la adaptación de la prueba a algunos cambios curriculares que han devenido recientemente en los programas de Ingeniería.

Así entonces, el lector encuentra en este Boletín algunos de los análisis de la información obtenida mediante el pilotaje en mención, mostrando aspectos tanto del proceso de construcción del instrumentos como de los

análisis estadísticos de los resultados, para derivar en el plan de mejoramiento del instrumento, en la perspectiva de ganar en validez y confiabilidad de la prueba y en el despliegue de la misma para los estudiantes de los 11 programas de ingeniería, a quienes se ofrece el curso de Algoritmia y Programación.

¿Cómo se concibió el pilotaje?

El pilotaje se concibió como una oportunidad de lectura de proceso sustentada en que "Una prueba piloto es aquella experimentación que se realiza por primera vez persiguiendo comprobar ciertas cuestiones. Se trata de un ensayo experimental, cuyas conclusiones pueden resultar interesantes para avanzar con el desarrollo de un programa, recurso, protocolo... Su objetivo es detectar tanto los posibles fallos o problemas como sus elementos positivos, funcionando como un primer paso para conseguir información pertinente" (Socialasturias.es, 2022).

Tras la aplicación del pilotaje, en el curso de Algoritmia y programación, se analizó la eficacia y pertinencia del instrumento y la articulación con los niveles de desempeño en el contexto de la taxonomía SOLO (Structure of Observed Learning Outcome).

Esta prueba piloto aplicada en el 2022-01, suministró datos estadísticos obtenidos mediante el procesamiento con Winsteps y utilizando la Teoría de Respuesta al Ítem. Estos datos son proporcionados por el equipo de profesores de Matemática y también fueron analizados por el profesor Luis Carlos Castillo, experto en lenguaje, quien, en calidad de asesor, ofreció algunas recomendaciones trabajadas con el equipo, que acogió lo necesario para mejorar el instrumento.

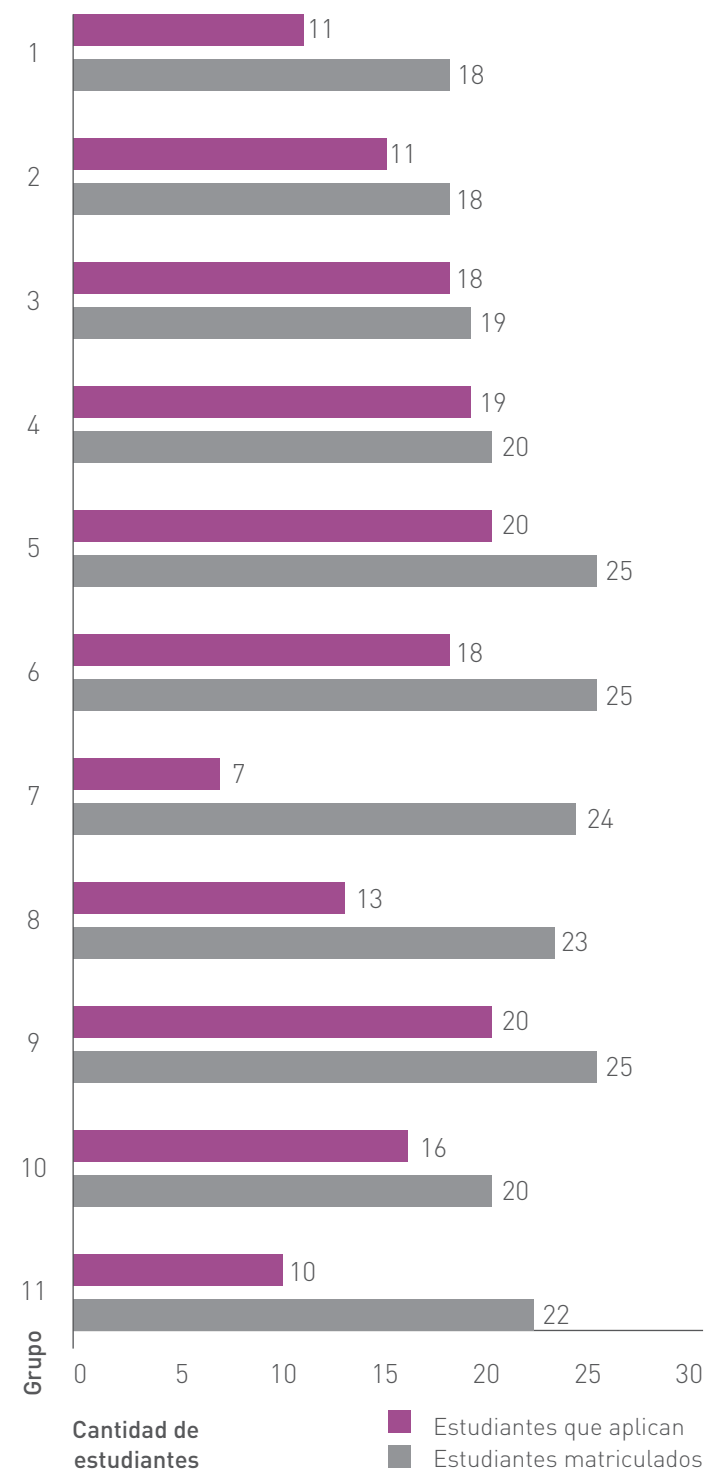
La prueba y sus características

Población evaluada

El pilotaje, de dos bloques, 30 ítems, de la nueva prueba con un grupo de 167 estudiantes que se distribuyen en los 11 grupos a los que se les oferta la asignatura.

Si bien el número de estudiantes aptos para presentar la prueba era de 239, el 70% lo presentó de manera voluntaria. A continuación, la distribución por grupos.

Figura 1. Dsitribución de la población de acuerdo con el curso



Estructura de los cuadernillos

La prueba aplicada en el pilotaje está constituida por un cuadernillo, que agrupa dos bloques, con un total de treinta ítems. Fueron diseñados teniendo en consideración tres contextos (uno personal, uno social y otro universal) y a cada contexto se asoció un total de 5 ítems, construidos de acuerdo con la

distribución en los diferentes dominios cognitivos (niveles de desempeño) y campos conceptuales (Muñoz-Ceballos, Peña-Paz, Lemos-Benavides & Arboleda-Molina, 2020).

Cada bloque responde a la siguiente tabla de especificaciones:

Constructo	Campo conceptual	Dominio cognitivo				Total
		Uniestructural	Multiestructural	Relacional	Abstracto ampliado	
Pensamiento computacional	Análisis	1(P)	1(U) 1(S)	1(P) 1(S) 1(U)	1(P)	7
	Algoritmo	1(S)	1(P) 1(S) 1(U)	1(P) 1(S) 1(U)	1(U)	8
TOTAL		2	5	6	2	15

Elaboración propia.



Los niveles de desempeño de acuerdo con la taxonomía SOLO son:

1. Uniestructural: El estudiante puede identificar, reconocer la existencia de un elemento constitutivo de los conceptos incluidos en los campos conceptuales de análisis y algoritmo.
2. Multiestructural: El estudiante puede enumerar, describir o listar la existencia de dos o más elementos constitutivos de los conceptos incluidos en los campos conceptuales de análisis y algoritmo.
3. Relacional: El estudiante puede relacionar o aplicar conceptos constitutivos de los campos conceptuales de análisis y algoritmo.
4. Abstracto Ampliado: El estudiante puede evaluar, adaptar o extender en torno a los campos conceptuales de análisis y algoritmo asociados con el constructo.

Los cuadernillos se integraron de la siguiente forma:

- Cuadernillo 1: Bloque 1 + Bloque 2
- Cuadernillo 2: Bloque 2 + Bloque 3
- Cuadernillo 3: Bloque 1 + Bloque 3

Resultados de la prueba piloto

Los resultados contienen indicadores estadísticos sobre los ítems y su comportamiento en la prueba, de tal manera, que se ofrezca información ajustada a la cohorte o grupo evaluado sobre los niveles de desempeño definidos en la taxonomía utilizada en la construcción del instrumento.

Sobre los ítems

Para la evaluación del ítem se tienen en cuenta la distribución por cuartiles en los niveles de desempeño (mapa de Wright¹):

- Un ítem es uniestructural si el porcentaje de respuestas correctas está en el rango de 76% al 99%.
- Un ítem es multiestructural si el porcentaje de respuestas correctas está en el rango de 51% al 75%.
- Un ítem es relacional si el porcentaje de respuestas correctas está en el rango de 26% al 50%.
- Un ítem es abstracto ampliado si el porcentaje de respuestas correctas está en el rango del 1% al 25%.

Sobre los estudiantes

Apoyado en la teoría de respuesta al ítem (TRI), se obtiene el valor numérico de la habilidad del estudiante y se aplica la Expresión 1, para obtener su nota en la escala de 0.0 a 5.0:

$$\text{nota} = ((\text{habilidad} - \text{habilidad promedio}) / \text{desviación estándar de la habilidad}) * 0.45 + 3 \text{ (Expresión 1)}$$

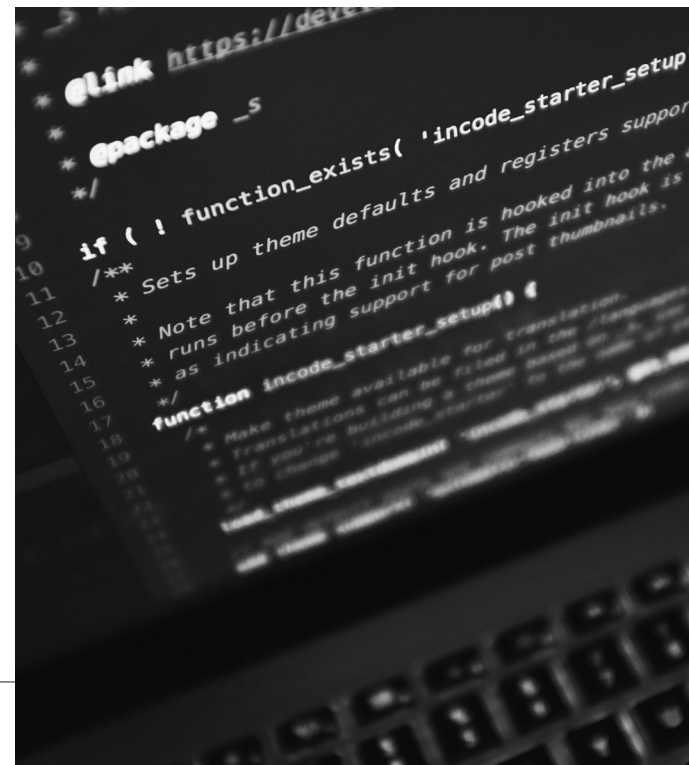
Donde:

- Habilidad corresponde con la habilidad del estudiante
- Habilidad promedio es el promedio de las habilidades de los estudiantes que presentaron la prueba.

- Desviación estándar de la habilidad mide la dispersión estándar de los estudiantes que presentaron la prueba.

Una vez aplicada la prueba y realizado el procesamiento se obtiene la data con las habilidades individuales. Estos insumos son necesarios para comprender el proceso de cálculo de la habilidad promedio, la desviación estándar, la nota del estudiante, y el nivel de desempeño según la taxonomía SOLO.

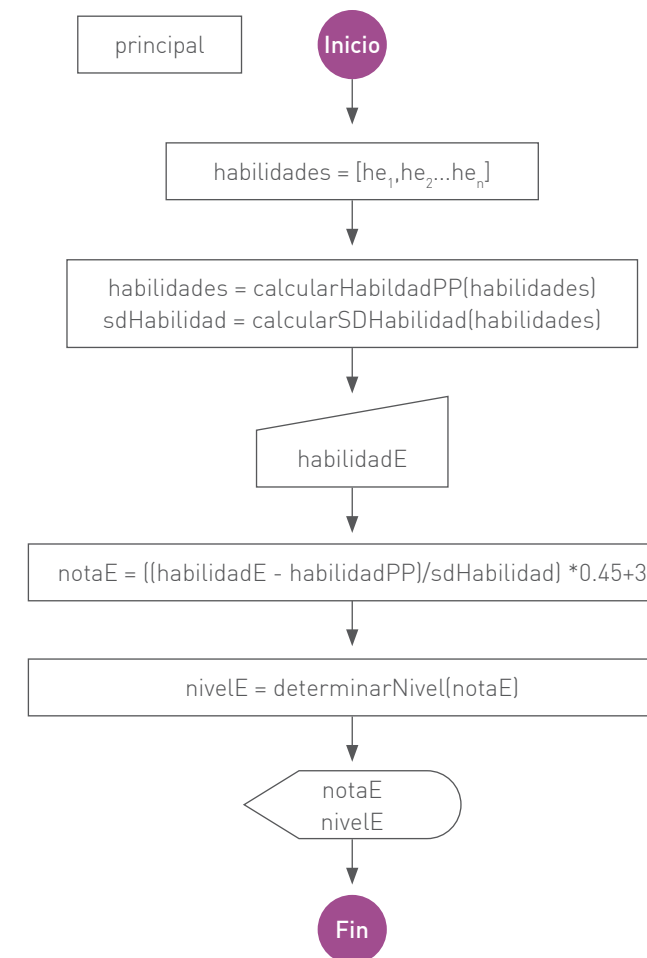
En la Figura 2 se ilustra el diagrama de flujo que permite calcular la nota de un estudiante, determinar su nivel de desempeño y mostrar ambos valores. El proceso inicia tomando de la data el conjunto de habilidades de los estudiantes que presentaron la prueba (habilidades). Con este conjunto, se procede a calcular la habilidad promedio de la prueba (habilidadPP) y su desviación estándar (sdHabilidad). Posteriormente, el algoritmo le solicita al usuario ingresar la habilidad del estudiante (habilidadE), la cual es tomada de la data suministrada. Y, a continuación, se aplica la Expresión 1 para obtener la nota del estudiante (notaE) en la escala de 0.0 a 5.0. Con este valor (notaE), se determina su nivel de desempeño (nivelE). Este proceso es ilustrado en la Figura 3. Finalmente, el algoritmo muestra la nota del estudiante (notaE) y su correspondiente nivel de desempeño (nivelE).



1 Que ofrece datos relevantes como la media de dificultad, la desviación estándar, discriminación, direccionalidad de los ítems, nivel de dificultad, tendencia, entre otros.

En el siguiente diagrama de flujo el proceso detallado:

Figura 2. Flujograma para el cálculo de la nota de un estudiante



habilidades es el conjunto de las habilidades, he_i , de los estudiantes en la prueba. Valores de tipo real. Este valor corresponde a los valores de la columna Habilidad en la hoja **Nota Estudiantes** del libro **Repo** en la data correspondiente.

habilidadPP es la habilidad promedio de la prueba. Valor de tipo real. Corresponde al promedio aritmético de las habilidades de los estudiantes que aplicaron la prueba.

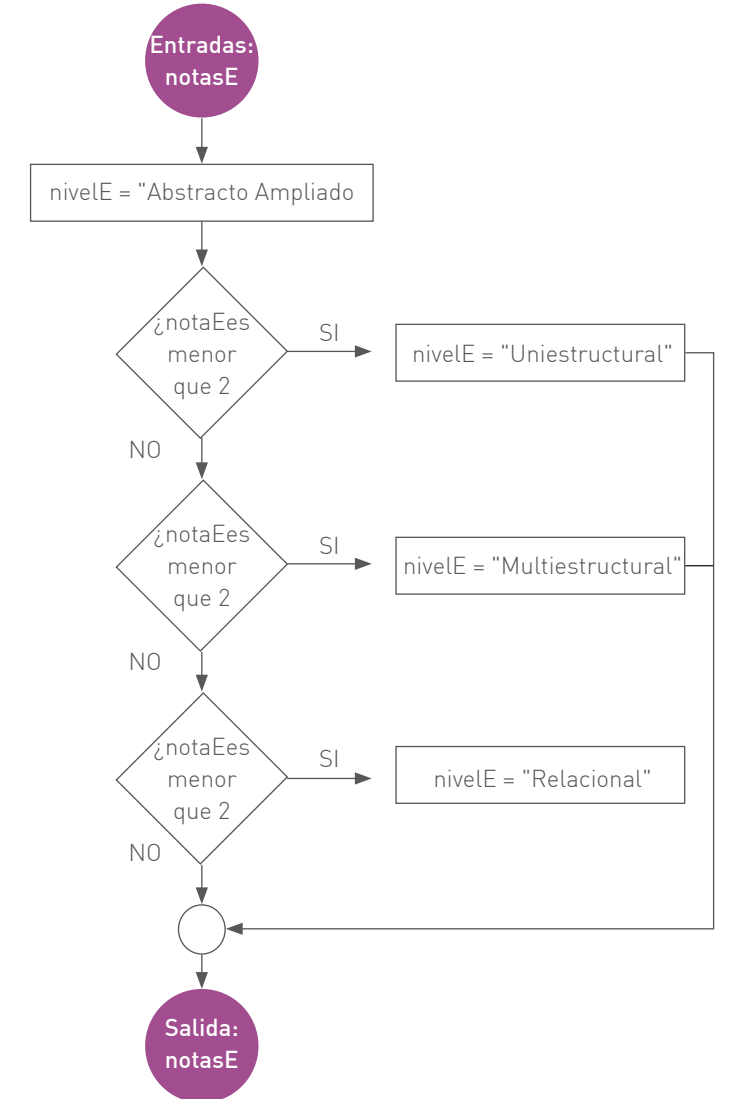
sdHabilidad es la desviación estándar de las habilidades de los estudiantes que aplicaron la prueba.

habilidadE es la habilidad de un estudiantes. Valor de tipo real.

notaE es la nota del estudiante. Valor de tipo real en la escala de 0.0 a 5.0

nivelE es el nivel de desempeño del estudiante según la taxonomía SOLO

Figura 3. Flujograma para determinar el nivel de desempeño de un estudiante



determinarNivel

Flujograma para determinar el nivel de desempeño de un estudiante, según la nota del estudiante y los siguientes rangos:

Nota del estudiante	Nivel de desempeño
(0.0, 2.0)	Uniestructural
(2.0, 3.0)	Multiestructural
(3.0, 4.0)	Relacional
(4.0, 5.0)	Abstracto ampliado



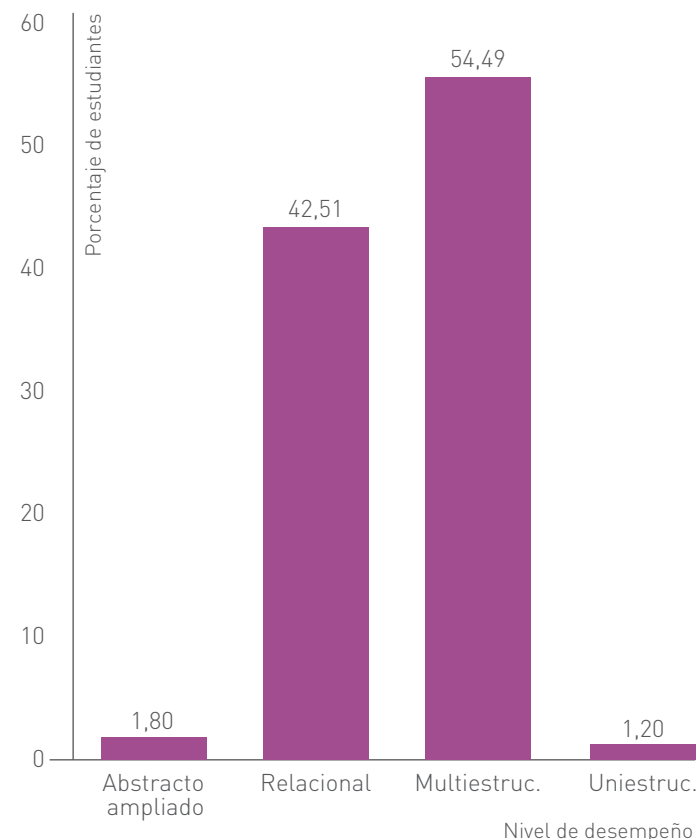
La Figura 3 recoge la representación gráfica de un algoritmo que recibe la nota de un estudiante (notaE) para determinar su correspondiente nivel de desempeño (nivelE). En el algoritmo se inicia suponiendo que el nivel de desempeño es "Abstracto Ampliado". A continuación, se evalúan las diferentes opciones para determinar el nivel de desempeño de un estudiante a partir de su nota.

En estas esquematizaciones gráficas del proceso se muestran los pasos secuenciales para la obtención de resultados y su representación visual permite un sistema de tratamiento adecuado para los resultados.

Análisis de los resultados del pilotaje

Respecto de la distribución de los evaluados en los niveles de desempeño de la taxonomía SOLO, se observa en la figura 4 el comportamiento:

Figura 4. Porcentajes por nivel de desempeño de los estudiantes en la prueba



Fuente: Elaboración propia.

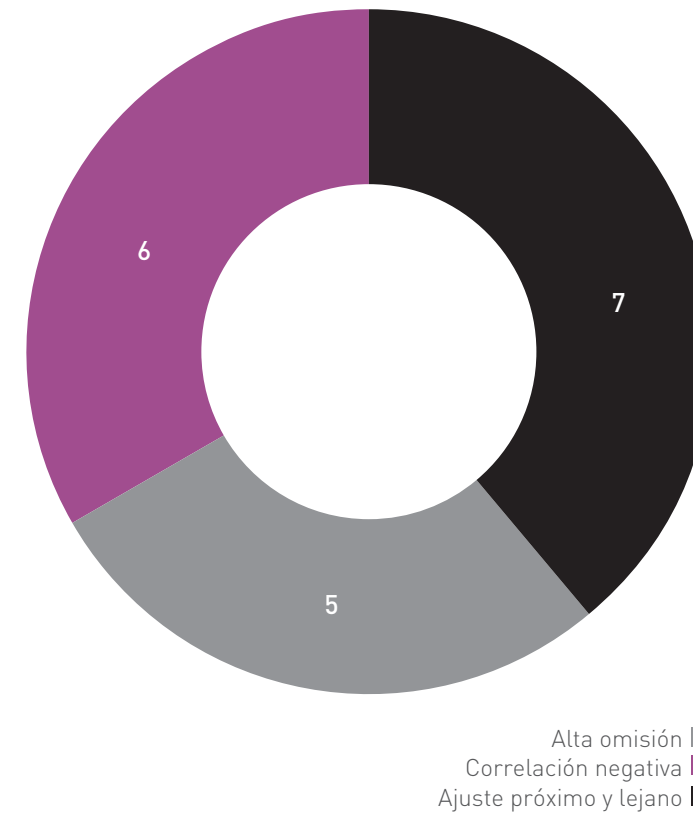


Como se aprecia, una alta proporción de estudiantes quedó ubicado en el nivel de desempeño relacional, lo cual indica el logro de los resultados de aprendizaje propuestos en la asignatura.

A su vez, se obtuvo la siguiente información:

- Alta omisión: estudiantes que omiten responder la mayoría de las preguntas.
- Correlación negativa: un valor negativo de este coeficiente (asociación inversa con la prueba), el ítem está alejado del constructo de la prueba, no aporta a la consistencia interna del test.
- Ajuste próximo y lejano: son indicadores que aportan información sobre la subestimación y la sobreestimación de la dificultad de los ítems y la habilidad de las personas.

Figura 5. Incidencias en la prueba



Fuente: Elaboración propia.

Evaluación de los ítems

En cuanto al análisis de ítems, a partir de la data resultante, en la etapa de calibración, se eliminaron del análisis los siguientes ítems:

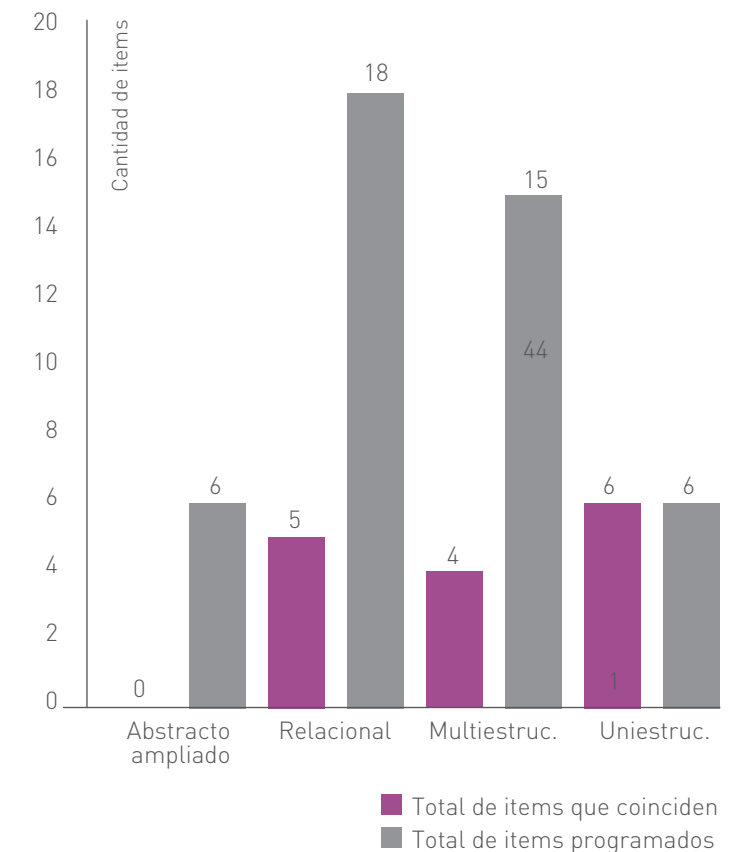
- Correlación negativa: ítem 24.
- Discriminación: ítem 18.
- Azar: ítems 2 y 13.
- Alta correlación entre ítems: ítem 44. (La alta correlación se presenta con el ítem 45 que presenta alcanza mejores indicadores).

Una vez aplicada la prueba piloto se procesan los resultados, se revisan y se produce un informe técnico cuyos detalles son los siguientes:

El presente informe se realiza con base en un análisis de los bloques/ítems y la tabla de especificaciones que da cuenta de la estructura de la prueba. Se describen los comportamientos de los ítems según la

nueva data de la aplicación de esta prueba. Se aconseja al equipo revisar la predicción con el comportamiento de los ítems para extraer de ello aprendizajes que puedan servirles de orientación al ajuste de los bloques/ítems o su completo rediseño, ya que esta vez contamos con un pilotaje que sirve de referencia².

Figura 6. Comportamiento de los ítems programados vs ítems coincidentes, en la prueba.



Fuente: Elaboración propia.

Como se puede observar, el mayor número de coincidencias se da entre preguntas uniestructurales (6 coincidencias), le siguen las preguntas relacionales (5 coincidencias) y las multiestructurales (4 coincidencias). No se presentaron coincidencias en las preguntas relacionadas con el nivel de dominio cognitivo abstracto ampliado. El porcentaje de coincidencia de los ítems corresponde al 33.3%.

² Castillo, L. C. (2022). Informe técnico: pilotaje nuevos bloques prueba de Algoritmia y Programación. Pág. 4.

Plan de mejoramiento y ajustes a los ítems

Con base en las recomendaciones del informe técnico³ se revisaron y aplicaron los ajustes que se consideraron pertinentes teniendo en cuenta las siguientes situaciones:

Recomendación del asesor:

Cuidar el diseño del ítem según el nivel SOLO a evaluar. Recordar que el diseño por parejas de opciones homólogas es conveniente para preguntas relacionales y, que las preguntas abstractas implican que el diseño plantee opciones que se diferencien por cifras, valores mínimos, tal vez de un mismo rango. Es fundamental recordar que la proporcionalidad de plausibilidad en las abstractas es idéntica para todas las opciones y que ninguna debería tener más de un 25% de posibilidades de ser marcada como la respuesta. (Pág. 3)

Plan de mejora:

- En el caso de los ítems catalogados como de nivel abstracto ampliado, se tuvo la precaución de que las 4 opciones fuesen igualmente probables, es decir, cada una con un 25% de probabilidad, lo cual exige por parte del estudiante una reflexión cuidadosa a la hora de seleccionar la opción correcta.
- En el caso de los ítems catalogados como relacionales, se tuvo la precaución de que cada par de opciones fuesen igualmente probables, es decir, cada par de opciones con una probabilidad del 50%.

Recomendación del asesor:

Resulta conveniente revisar las opciones de respuesta diseñadas, pues la clave se resuelve por procesos de comprensión de lectura simples o con una lectura detallada, de elementos de información como inferencias, deducciones, conjeturas, hipótesis o proposiciones. El estudiante no tendría necesidad de un saber experto o específico, por comprensión de lectura, llega a la alternativa de respuesta correcta. (Pag. 5)

3 Castillo. L. C. (2022). Informe técnico: pilotaje nuevos bloques prueba de algoritmia.

Plan de mejora:

Al revisar el ítem se evidencia que muchas de las opciones eran descartables con una lectura y se procedió a modificar, en algunos casos, las opciones, y en otros casos, el enunciado del ítem.

Recomendación del asesor:

Conviene revisar la tendencia que subyace en el caso de preguntas que se proponen como relacionales o abstractas y tienden a comportarse como uniestructurales o multiestructurales. En unos casos se observa que se debe a que lo que el ítem exige no son procesos cognitivos complejos sino procesos más simples, como identificar y cotejar o comparar elementos. (Pag. 7)

Plan de mejora:

Se reformula el enunciado del ítem, de tal manera que le apueste al nivel de desempeño que se espera medir.

Conclusiones

- El análisis de la prueba piloto y su informe técnico muestra un terreno próspero para mejorar la idoneidad, validez y confiabilidad del instrumento.
- A partir del informe técnico se logró identificar algunas falencias de los ítems, para realizar ajustes con oportunidad y pertinencia. Este proceso se realiza con el propósito de mejorar el instrumento utilizado para las pruebas de entrada y salida del periodo 2022-03, dejándolo estable y logrando trazabilidad.
- Las mejoras en el instrumento tienen impacto en la validez y confiabilidad de la prueba, logrando con mayor precisión una ubicación del estudiante en el nivel de desempeño cognitivo de acuerdo con las habilidades y conocimientos detectados.
- Aunque se esperaba un porcentaje de coincidencia mayor entre lo planificado al construir los ítems, y lo que se obtiene en el pilotaje, la participación del 70% de los estudiantes es significativo y permite favorecer los ajustes esperados. Así entonces, el pilotaje, de acuerdo con los parámetros de cali-

dad establecidos por el SIEA, es de gran utilidad, porque además de verificar la coherencia entre el ítem y el indicador de evaluación, la concordancia entre las opciones de respuesta y el enunciado y la verificación de la plausibilidad de los indicadores, entre otras, proporciona información estadística sobre la capacidad discriminativa, la tasa y patrón de omisión, grado de dificultad y distribución de opciones de respuesta en la pregunta. Todo ello permite tomar decisiones informadas para ajustar, descartar, rediseñar. En este proceso, el equipo se fortalece, un riguroso análisis del comportamiento métrico del ítem y un cumplimiento riguroso de las fases del proceso de construcción de pruebas de acuerdo con los criterios de calidad de pruebas estandarizadas.

- La prueba piloto permitió mejorar el instrumento, fue aplicada al final del semestre 2022-1, por lo que no se tuvieron datos comparativos que permitieran establecer rutas de aprendizaje para mejorar los niveles de desempeño de los estudiantes, lo que derivó en una prueba más compacta y con mayor poder de predictibilidad aplicada en el 2022-3.

Referencias

- Biggs, J. B., & Collis, K. F. (1982). Evaluation the quality of learning: The SOLO taxonomy (structure of the observed learning outcome). Academic Press.
- Gutiérrez R., & Cañas A. (2019). Boletín Del Sistema Institucional de Evaluación de los Aprendizajes. Universidad Autónoma de Occidente. Santiago de Cali, Valle del Cauca.
- UNIVERSIDAD AUTÓNOMA DE OCCIDENTE-FACULTAD DE INGENIERÍA (2020). Marco teórico prueba SIEA Algoritmia y Programación. Documento de trabajo. Santiago de Cali, Valle del Cauca.
- Muñoz-Ceballos, L.E., & Peña-Paz L., & Lemos-Benavides J.A., & Arboleda-Molina O. (2020). Boletín 6 Del Sistema Institucional de Evaluación de los Aprendi-

zajes. Universidad Autónoma de Occidente. Santiago de Cali, Valle del Cauca. Boletín 6. ISSN: 2711-175X [En línea]. Descargable desde https://issuu.com/catello-uao.edu/docs/boleti_n_siea__6.

Gobierno del Principado de Asturias. (2022). La importancia del pilotaje como fase de experimentación y evaluación. https://www.socialasurias.es/pilotate/zoom/la-importancia-del-pilotate-como-fase-de-experimentacion-y-evaluacion_1341_596_2265_0_1_in.html



Consulte en línea nuestros tres últimos boletines: <https://sitios.uao.edu.co/docentes/boletines-siea>