



INFORME TÉCNICO

PILOTO 2022

Índice

Introducción	2
1. Características del Piloto PAES	3
1.1. Pruebas Piloto	3
1.2. Selección de la muestra	5
1.3. Consideraciones especiales	6
1.4. Metodología de análisis	7
1.5. Calibración del banco de preguntas	9
2. Datos de rendición	10
2.1. Rendición por zona	11
2.2. Rendición por sexo	12
2.3. Rendición por dependencia	13
2.4. Rendición por rama educativa	14
3. Resultados	15
3.1. Competencia Lectora	15
3.1.1. Funcionamiento de ítems anclas	16
3.1.2. Estado de ítems	18
3.2. Competencia Matemática 1 (M1)	20
3.2.1. Funcionamiento de ítems anclas	21
3.2.2. Estado de ítems	23
3.3. Competencia Matemática 2 (M2)	25
3.3.1. Funcionamiento de ítems anclas	26
3.3.2. Estado de ítems	28
3.4. Historia y Ciencias Sociales	30
3.4.1. Funcionamiento de ítems anclas	31
3.4.2. Estado de ítems	33
3.5. Ciencias	35
3.5.1. Funcionamiento de ítems anclas	35
3.5.2. Estado de ítems	38
4. Anexo	41
4.1. Anexo A: Modelo de Rasch	41
4.1.1. Método de estimación de las dificultades	41
4.1.2. Método de estimación de las habilidades	42
4.1.3. Supuestos del modelo Rasch	42
4.2. Anexo B: Población que rinde por zona geográfica	43
5. Glosario	44

Introducción

Cada año DEMRE realiza procesos de pilotaje de ítems, llamado Simulación PAES o Piloto, con el propósito de validar el comportamiento de las preguntas que se construyen para cada una de las Pruebas de Acceso a la Educación Superior (PAES). Esto permite contar con un banco de preguntas disponibles que cumplan con las características psicométricas y técnicas adecuadas para las pruebas oficiales que se utilizan en cada proceso de admisión universitaria.

El proceso de pilotaje del año 2022 se realizó en dos periodos. El primer punto se aplicó durante en el mes de abril el cual se realizó un llamado a nivel de colegios. El segundo punto se aplicó en el mes de octubre y tuvo una convocatoria a nivel de colegios, como en el punto 1, y a nivel individual en un llamado a inscritos a la prueba oficial PAES del año en curso.

La aplicación de la Simulación PAES 2022 se llevó a cabo siguiendo estrictamente las medidas de prevención recomendadas por las autoridades sanitarias. Estas medidas se implementaron con el objetivo de prevenir la propagación y contagio del COVID-19 en el país, garantizando así la seguridad sanitaria de todas las personas que participaron en el proceso de aplicación.

Este informe tiene finalidad de describir características, aspectos técnicos y resultados de la aplicación del piloto realizada en el año 2022. El informe se divide en 3 capítulos: en el primer capítulo describen las pruebas y sus características, la selección de la muestra y la metodología de análisis. En el segundo capítulo, se muestran los datos de rendición y en el tercer capítulo se muestran los resultados para cada una de las pruebas.

1. Características del Piloto PAES

La aplicación piloto de la PAES es una tarea crucial en el proceso de construcción de instrumentos de evaluación para el Sistema de Acceso, ya que en esta se realiza la validación estadística de los ítems que luego formarán parte de las pruebas oficiales. Esta validación se realiza en una muestra de la población, bajo estrictos parámetros que aseguran la calidad del producto final.

El Piloto PAES se aplicó dos grupos de participantes: En primer lugar, a estudiantes que actualmente están cursando tercero y cuarto año de Enseñanza Media. En este caso, fue posible organizar el operativo directamente con sus establecimientos educativos. El segundo grupo corresponde a egresados de Enseñanza Media en promociones anteriores o actuales inscritas al proceso de Admisión 2023, para quienes se realizó un llamado a nivel individual.

1.1. Pruebas Piloto

El proceso de pilotaje llevado a cabo el año 2022 se aplicó en dos fechas. En el mes de abril se realizó un pilotaje a nivel colegios, mientras que en el mes de octubre contó con 2 modalidades: a nivel colegios e individual.

Las diversas pruebas pilotos son similares en cuanto a estructura, ajustándose a los cambios y necesidades de cada una de las pruebas, como poblar el banco de ítems y las restricciones a nivel horario (horas lectivas) de los establecimientos en los que se aplica. La batería PAES del piloto tiene la siguiente composición:

- La prueba de Competencia Lectora consta de 65 preguntas, con 4 o 5 opciones de respuestas que fueron probadas en 18 formas.
- La prueba de Competencia Matemática 1 (M1) consta de 65 preguntas, con 4 o 5 opciones de respuestas que fueron probadas en 18 formas.
- La prueba de Competencia Matemática 2 (M2) consta de 65 preguntas en el mes de abril y 55 en el mes de octubre, con 4 o 5 opciones de respuestas que fueron probadas en 18 formas.
- La prueba de Historia y Ciencias Sociales consta de 65 preguntas, con 4 o 5 opciones de respuestas que fueron probadas en 19 formas.

- La prueba de Ciencias consta de 80 preguntas, con 4 o 5 opciones de respuestas que fueron probadas en 35 formas.

A continuación, se detallan el número de formas aplicadas en cada una de las pruebas:

Prueba	Formas Aplicadas
Competencia Lectora	18
Competencia Matemática 1 (M1)	18
Competencia Matemática 2 (M2)	18
Historia y Ciencias Sociales	19
Ciencias	35

Cuadro 1: Número de formas Simulación PAES

A cada uno de los estudiantes convocados para el piloto aplicado en colegios le fueron asignados dos pruebas:

- La primera de Competencia Lectora o Competencia Matemática 1 (M1) o Competencia Matemática 2 (M2)
- La segunda una prueba electiva de Ciencias o Historia y Ciencias Sociales..

La asignación de la prueba electiva se realizó en el mes de abril primordialmente en base a la información que proporcionaron los establecimientos, dado que no se contaba con la información de inscripción. En el mes de octubre la asignación se realizó según su inscripción para la PAES Admisión 2023. En el caso del mes de octubre, para el pilotaje a nivel individual, cada estudiante debía rendir la prueba de Competencia Matemática 2 (M2) y la prueba de Ciencias, como se muestra en el cuadro 2.

Prueba	Mes de Abril		Mes de Octubre
	3ro Medios	4to Medios	Individual
Competencia Lectora	X	X	
Competencia Matemática 1 (M1)	X	X	
Competencia Matemática 2 (M2)	X	X	X
Historia y Ciencias Sociales	X	X	
Ciencias	X	X	X

Cuadro 2: Resumen de modalidad Simulación PAES

1.2. Selección de la muestra

Uno de los pasos que se realiza en el proceso de pilotaje es seleccionar una muestra que nos permita representar lo mejor posible a la población que rinde, ya que esto permite obtener ensamblajes adecuados para cada una de las pruebas.

Para obtener esta muestra se deben considerar los siguientes aspectos:

- **Caracterización de la población:** Para cumplir con el objetivo de representar la población que rinde las pruebas oficiales, esta se caracteriza por medio de la dependencia administrativa y rama educacional del establecimiento de egreso, así como la zona geográfica. Para estos efectos, se consideran tres zonas: zona norte, centro y sur, cuya composición se detalla en el anexo B.
- **Restricciones logísticas:** En la práctica, es necesario considerar la parte de la población disponible en las diversas localidades así como la facilidad de acceso a ellas.

Luego de conocer estos aspectos se procede de la siguiente forma para escoger la muestra:

1. Se realiza la selección de regiones que participarán del proceso de pilotaje. Esta selección se realiza teniendo en cuenta ciertas limitaciones geográficas y operacionales que podrían dejar fuera ciertos lugares del país a los que es complejo acceder. Además, se considera a aquellas regiones que permitan capturar la mayor información posible de la zona geográfica a la que pertenecen, tanto por cantidad de participantes en el piloto como por cobertura de las distintas variables de interés del establecimiento educacional de egreso, como lo es dependencia, sexo y rama educacional.
2. De las regiones seleccionadas previamente, se deben seleccionar algunas comunas. Siguiendo la misma lógica anterior, estas deben capturar la mayor información posible de la región a la que pertenecen.
3. Se distribuyen los folletos asignados a las comunas previamente seleccionadas.
4. Se seleccionan establecimientos educacionales en cada comuna, de manera tal que se cumpla con la cantidad de folletos asignados y que a nivel país se cumpla con las distribuciones asociadas a la caracterización de la población.
5. Se envía la lista de establecimientos a la unidad encargada de los contactos y en caso de ser necesario, se realizan reemplazos, cuidando siempre de mantener las distribuciones.

1.3. Consideraciones especiales

- Para caracterizar a la población que rinde, solo se utiliza la información de los postulantes egresados de establecimientos diurnos y con educación para niños y jóvenes. Es decir, para la selección de la muestra, quedan fuera todos los establecimientos de educación nocturna y para adultos.
- La selección se realiza a nivel de unidad educativa. Es decir, si un establecimiento educacional (identificado por un RBD) que, es seleccionado en la muestra tiene más de un código de enseñanza asociado, se consideran todas las especialidades asociadas a los códigos para la aplicación del piloto.

1.4. Metodología de análisis

Una vez rendidas las pruebas, se lleva a cabo el proceso de análisis de los instrumentos aplicados que, de manera general, consta de los siguientes pasos:

1. En primer lugar se realiza la corrección de las bases de datos que contienen las respuestas a las preguntas, mediante la cual, se determina si la respuesta dada a cada ítem es correcta o incorrecta considerando la información del clavijero. A demás, a partir de la corrección se obtiene información a nivel de persona acerca del porcentaje de preguntas omitidas, correctas e incorrectas.
2. Con el objetivo de tener una estimación inicial más limpia de las dificultades de los ítems y habilidades de las personas, no se considera la información de personas que hubieran omitido más del 90 % de las preguntas. Luego se ajusta un modelo Rasch¹ a los datos, estimando de forma libre todos los parámetros de los ítems, con el objetivo de tener una estimación inicial de las habilidades de cada persona.
3. Una vez ajustado el modelo se realiza la verificación de claves, en donde se analiza el comportamiento psicométrico de la clave y los distractores de cada pregunta, por medio de la correlación biserial. Ante una posible clave incorrecta se levanta una alerta, lo cual es posteriormente revisado de manera cualitativa por el comité relativo a la prueba.

Luego de la validación de claves, en caso de ser necesario, se vuelve a corregir la base de datos y se ajusta nuevamente un modelo de Rasch, a partir del cual se obtienen los estadísticos de ajuste Infit y Outfit, tanto para las personas como para las preguntas. Se determina como criterio común para personas y preguntas que los valores de los estadísticos estén en el rango $[0.5, 1.5]$ y en caso de no estarlo (personas y/o preguntas) se decide excluirlos para los análisis posteriores. Por otro lado, si al observar la Curva Característica del Ítem (ICC, por sus siglas en inglés) de alguna pregunta se observa que la clave tiene un comportamiento decreciente, esta tampoco es considerada en los análisis posteriores.

Luego de esto, se realiza la revisión y evaluación de los ítems anclas (denominados así debido a que provienen de procesos oficiales anteriores) en la cual mediante un modelo de regresión lineal se determina si los ítems siguen o no el comportamiento general de la recta ajustada. Dicha recta compara las dificultades obtenidas en el proceso de pilotaje, estimadas de manera libre versus las que se estimaron en el proceso oficial correspondiente. Todos aquellos ítems que se detecten como outliers de la regresión dejan de cumplir su función como ancla entre el banco de ítems calibrado y el proceso de pilotaje. Con esta información se ajusta un nuevo modelo Rasch a los datos, con

¹ver anexo modelo de Rasch 4.1.

la salvedad de que a los ítems con buen comportamiento en el análisis de anclaje (no outliers) no se les estima una dificultad con los datos del pilotaje, sino que dicho parámetro se fija con la estimación del proceso oficial que corresponda. El resto de las dificultades se estima de manera libre.

4. Del modelo de Rasch antes descrito, se obtienen las estimaciones de dificultad de los ítems, la habilidad de las personas y la confiabilidad EAP del instrumento.
5. Finalmente, se evalúan por una última vez los ítems a través de las ICC. Este proceso de revisión y evaluación permite poner los ítems en dos categorías: “Creciente”, en donde la curva de la clave es creciente en todo el rango de habilidad, o “Decreciente”, en donde la curva de la clave no crece en uno o más de los intervalos de habilidad considerados. Posterior a esta categorización, se revisan nuevamente los ítems y se determina su clasificación final para el uso en procesos oficiales posteriores.

1.5. Calibración del banco de preguntas

El fin principal del piloto es nutrir de preguntas de calidad al banco de preguntas. Para esto, es fundamental el uso de preguntas de anclaje entre la aplicación oficial y piloto con el propósito de calibrar el banco en base a una misma escala para cada instrumento de medición.

Para estudiar las diferencias entre los resultados de la aplicación piloto y la oficial junto con la calibración del banco de preguntas, se utilizó para cada disciplina, un bloque de ítems “anclas”. El bloque ancla corresponde a ítems aplicados en las pruebas oficiales de los procesos de Admisión entre 2016 y 2022, con el fin de mantener un enlace entre los pilotos y el proceso oficial.

Para verificar la estabilidad de los parámetros entre la prueba oficial y la aplicación piloto se comparan los resultados obtenidos en los bloques anclas, dado que fueron utilizados en ambas instancias. Realizar el pilotaje de preguntas en una muestra de la población que finalmente participará en la aplicación oficial, permite confeccionar las pruebas oficiales con mayor información y anticipar el comportamiento de las preguntas utilizadas.

Uno de los desafíos al realizar el pilotaje de forma independiente a la aplicación oficial es validar de la manera más precisa posible el comportamiento de las preguntas, ya que, al estar libre del peso inherente de la aplicación oficial debido a las consecuencias individuales de esta para cada examinado, la motivación de los estudiantes para responder correctamente la prueba podría verse afectada. Sin embargo, el tipo de análisis realizado por DEMRE nos asegura la calidad de las preguntas destinadas a procesos oficiales.

2. Datos de rendición

A continuación se muestra la cantidad de estudiantes que rindió cada una de las pruebas piloteadas.

Prueba	Número de Personas
Competencia Lectora	15915
Competencia Matemática 1	15295
Competencia Matemática 2	8919
Historia y Ciencias Sociales	12631
Ciencias	18207

Cuadro 3: Número de personas que respondieron la Simulación PAES

2.1. Rendición por zona

Para los análisis se dividió en tres partes la zona geográfica: en zona norte, centro y sur². La siguiente tabla muestra la distribución por zona geográfica de la población que rindió cada una de las pruebas. Se observa que en la zona central tiene la mayor cantidad de participantes en todas las pruebas.

Zona	Número de personas
Competencia Lectora	
Zona Norte	3930
Zona Centro	7437
Zona Sur	4491
Sin Información	57
Competencia Matemática 1	
Zona Norte	3709
Zona Centro	6559
Zona Sur	4971
Sin Información	56
Competencia Matemática 2	
Zona Norte	1594
Zona Centro	5083
Zona Sur	2095
Sin Información	147
Historia y Ciencias Sociales	
Zona Norte	3124
Zona Centro	6316
Zona Sur	3139
Sin Información	52
Ciencias	
Zona Norte	4085
Zona Centro	9620
Zona Sur	4429
Sin Información	73

Cuadro 4: Distribución por zona geográfica del número de personas que respondieron la Simulación PAES

²La división de regiones en las tres zonas geográficas que se describe en el anexo 4.2.

2.2. Rendición por sexo

La siguiente tabla se puede ver la distribución por sexo de la población que respondió cada una de las pruebas. Se observa que el número de mujeres que realizaron la prueba de Competencia Matemática 2 es menor en comparación con las otras pruebas, mientras que las cantidades de hombres y mujeres son bastante similares en el resto de las pruebas.

Sexo	Número de Personas
Competencia Lectora	
Hombres	7890
Mujeres	8025
Competencia Matemática 1	
Hombres	7311
Mujeres	7984
Competencia Matemática 2	
Hombres	5135
Mujeres	3784
Historia y Ciencias Sociales	
Hombres	6286
Mujeres	6345
Ciencias	
Hombres	8737
Mujeres	9470

Cuadro 5: Distribución por sexo del número de personas que respondieron la Simulación PAES

2.3. Rendición por dependencia

En la siguiente tabla se muestra la distribución por dependencia de la población que rindió cada una de las pruebas. Se observa que en todas la prueba el grupo con mayor número de personas proviene de establecimientos particulares subvencionados, seguido de los establecimientos municipales.

Dependencia	Número de Personas
Competencia Lectora	
Particular Pagado	1273
Particular Subvencionado	8298
Municipal	6287
Sin Información	57
Competencia Matemática 1	
Particular Pagado	1616
Particular Subvencionado	6849
Municipal	6774
Sin Información	56
Competencia Matemática 2	
Particular Pagado	2206
Particular Subvencionado	3252
Municipal	3214
Sin Información	147
Historia y Ciencias Sociales	
Particular Pagado	1492
Particular Subvencionado	6054
Municipal	5033
. Sin Información	52
Ciencias	
Particular Pagado	2232
Particular Subvencionado	8794
Municipal	7108
Sin Información	73

Cuadro 6: Distribución por grupo dependencia del número de personas que respondieron la Simulación PAES

2.4. Rendición por rama educativa

La siguiente tabla muestra la distribución por rama educativa de la población que rindió cada una de las pruebas. Se observa que, en todas las pruebas el grupo que predominó fue el de establecimientos humanista científico.

Rama	Número de Personas
Competencia Lectora	
Humanista Científico	12472
Técnico Profesional	3387
Sin información	56
Competencia Matemática 1	
Humanista Científico	10741
Técnico Profesional	4499
Sin información	51
Competencia Matemática 2	
Humanista Científico	6882
Técnico Profesional	1890
Sin información	147
Historia y Ciencias Sociales	
Humanista Científico	9158
Técnico Profesional	3421
Sin información	52
Ciencias	
Humanista Científico	13618
Técnico Profesional	4517
Sin información	72

Cuadro 7: Distribución por rama educativa del número de personas que respondieron la Simulación PAES

3. Resultados

En esta sección se entregan los resultados, relativos al comportamiento de los ítems, en cuanto al total, como por habilidad y ejes temáticos³, para las pruebas de Competencia Lectora, Competencia Matemática 1, Competencia Matemática 2, Historia y Ciencias Sociales. En el caso de los ítems de Ciencias los resultados se entregan por disciplina y ejes temáticos.

3.1. Competencia Lectora

Las confiabilidades de los instrumentos oscilaron entre 0,86 y 0.88 aproximadamente, entre las formas piloteadas. A continuación se presenta un resumen de la cantidad de ítems utilizados.

Grupo	Número de Ítems
Cantidad formas	18
Cantidad ítems	915
Ítems nuevos	876
Ítems ancla	39
Ítems aceptados	801
Ítems rechazados	75

Cuadro 8: Número de ítems y formas de Simulación PAES en prueba de Competencia Lectora

³Para mayor información ver temario en <https://demre.cl/la-prueba/pruebas-y-temarios/presentacion-pruebas-temarios-paes-invierno-p2024>.

3.1.1. Funcionamiento de ítems anclas

El siguiente gráfico muestra el funcionamiento de los ítems ancla en el piloto de abril donde se puede observar la relación lineal entre las dificultades estimadas en el piloto (eje Y) con las dificultades de la calibración del banco de ítems (eje X). De acuerdo a la detección de outliers, se puede apreciar que se desanclan 2 ítems. El ajuste de esta prueba es de 0,88.

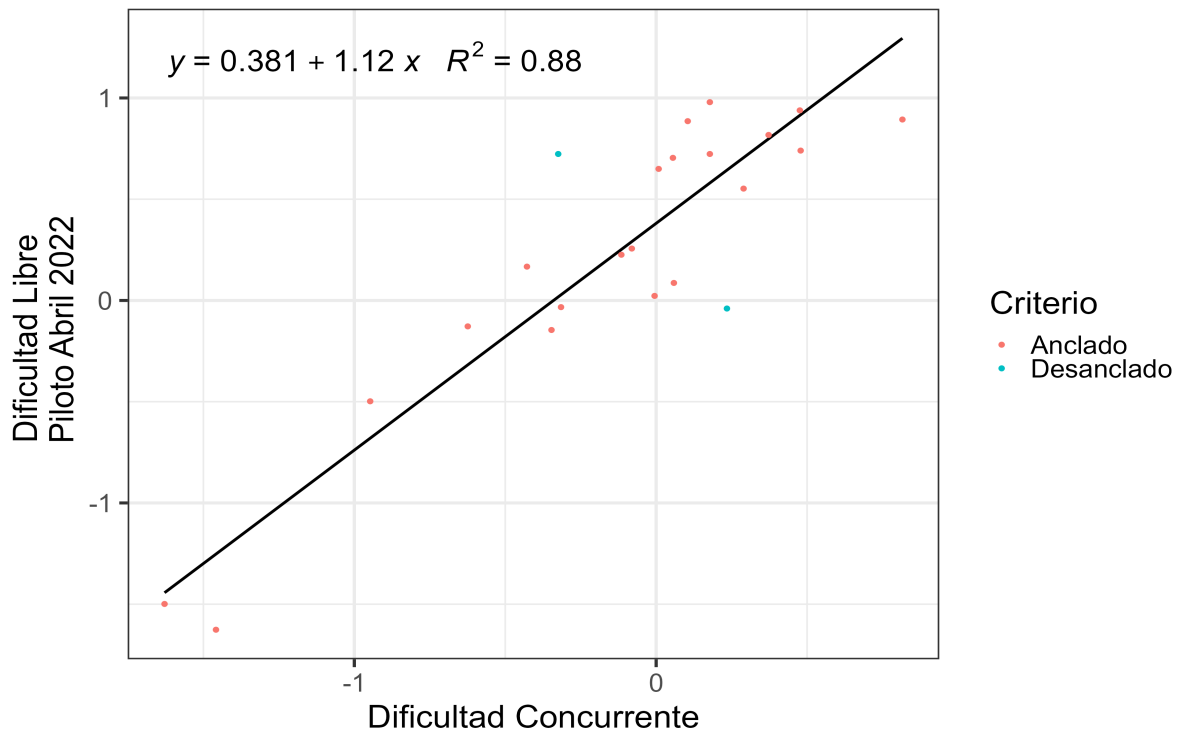


Figura 1: Funcionamiento de Ítems Anclas en el Piloto del Mes de Abril

El siguiente gráfico muestra el funcionamiento de los ítems ancla en el piloto de octubre donde se puede observar la relación lineal entre las dificultades estimadas en el piloto (eje Y) con las dificultades de la calibración del banco de ítems (eje X). De acuerdo a la detección de outliers, se puede apreciar que no se desanclan ítems. El ajuste de esta prueba es de 0,87.

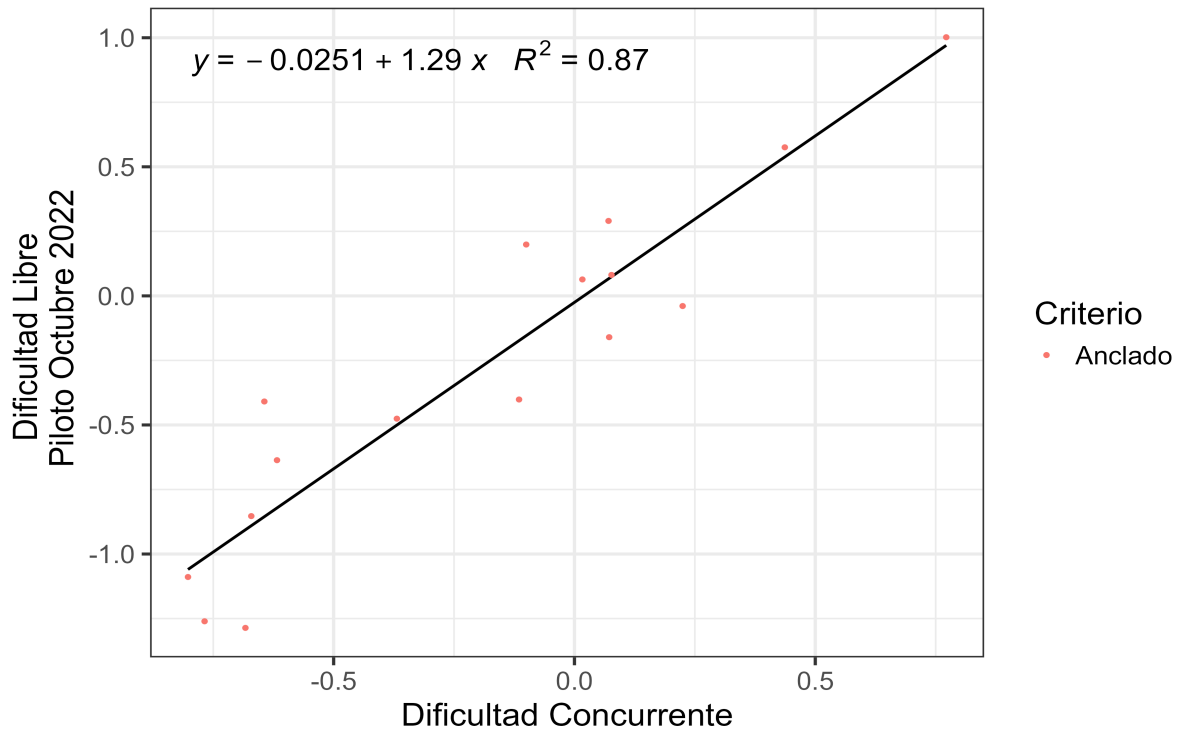


Figura 2: Funcionamiento de ítems anclas en Simulación PAES del mes de octubre

3.1.2. Estado de ítems

En la prueba de Competencia Lectora se pilotearon 876 ítems nuevos, de los cuales se aprobaron 801, lo que corresponde al 91,4% de aprobación.

En cuanto a las habilidades se probaron 412 ítems de Interpretar, de los cuales se aprobó un 90 %, mientras que en la habilidad de Evaluar solo se probaron 158 ítems y se aprobó un 85 %.

Asimismo, la situación de lectura Personal obtuvo una tasa de aprobación superior al 91 %, como se puede apreciar en las figuras 3, 4 y 5.

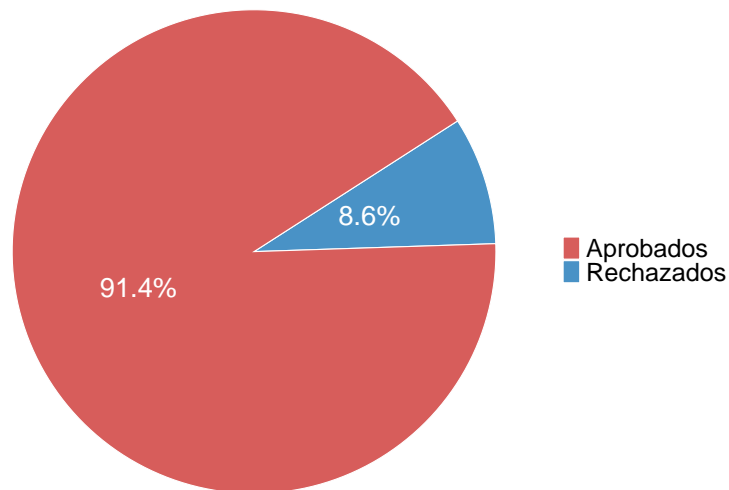


Figura 3: Distribución de ítems aprobados de Competencia Lectora

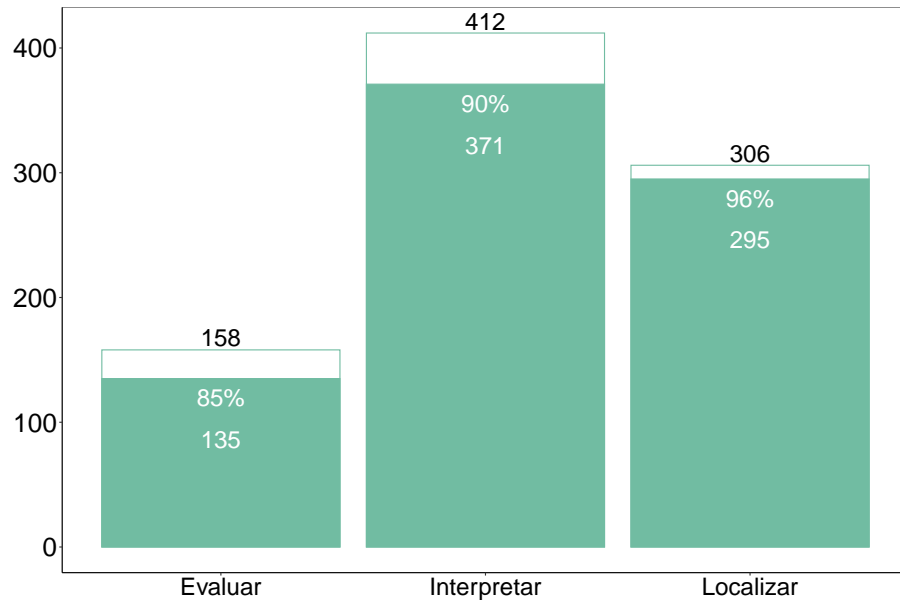


Figura 4: Distribución de ítems aprobados según habilidad

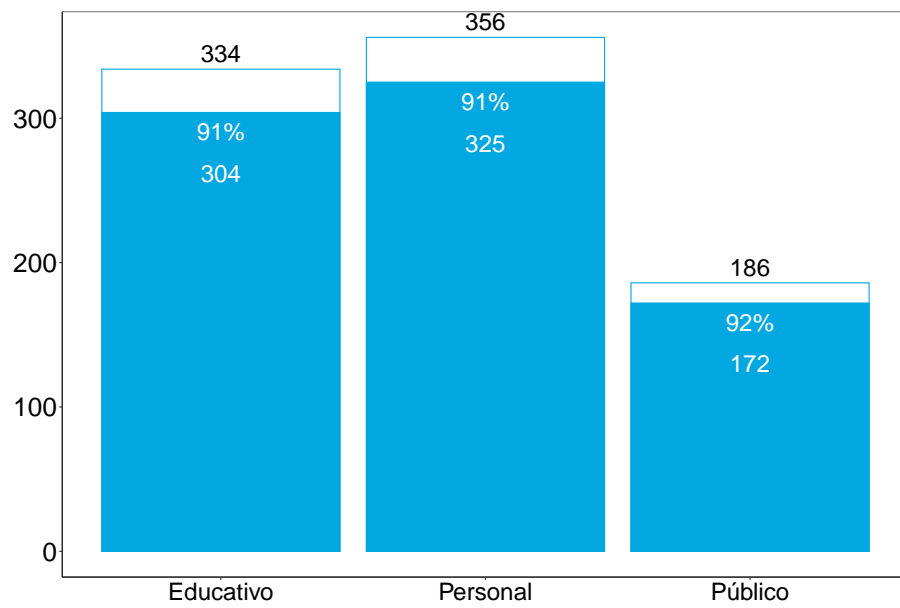


Figura 5: Distribución de ítems aprobados según situación de Lectura

3.2. Competencia Matemática 1 (M1)

Las confiabilidades de los instrumentos oscilaron entre 0.69 y 0.81 aproximadamente, entre las formas piloteadas. A continuación se presenta un resumen de los ítems utilizados.

Grupo	Datos de Ítems
Cantidad formas	18
Cantidad ítems	925
Ítems nuevos	880
Ítems ancla	45
Ítems aceptados	725
Ítems rechazados	180

Cuadro 9: Datos de ítems y formas de Simulación PAES en prueba de Competencia Matemática 1

3.2.1. Funcionamiento de ítems anclas

El siguiente gráfico de Competencia Matemática 1⁴ muestra el funcionamiento de los ítems ancla en el piloto de abril donde se puede observar la relación lineal entre las dificultades estimadas en el piloto (eje Y) con las dificultades de la calibración del banco de ítems (eje X). De acuerdo a la detección de outliers, se puede apreciar que se desanclan 3 ítems. El ajuste de esta prueba es de 0,84.

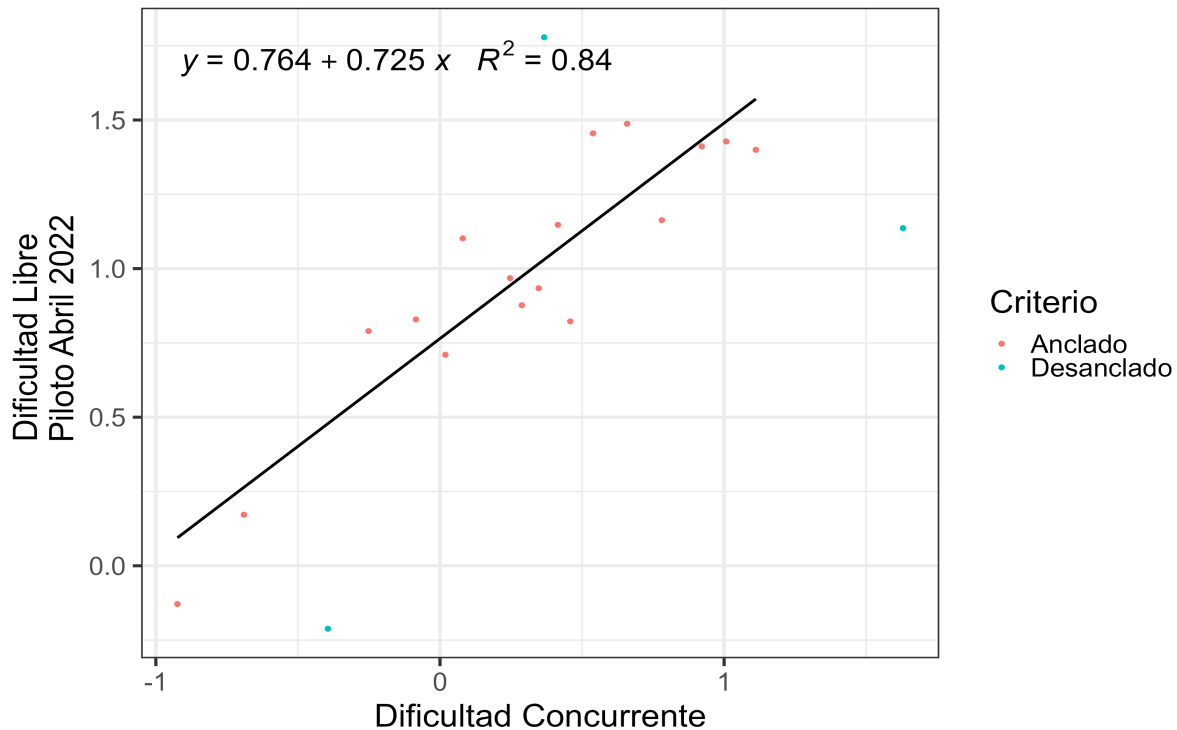


Figura 6: Funcionamiento de ítems anclas en Simulación PAES del mes de abril

⁴Los análisis para las pruebas M1 y M2 se realizaron de manera conjunta pero en este informe se muestran de forma separada.

El siguiente gráfico de Competencia Matemática ⁵ muestra el funcionamiento de los ítems ancla en el piloto de octubre donde se puede observar la relación lineal entre las dificultades estimadas en el piloto (eje Y) con las dificultades de la calibración del banco de ítems (eje X). De acuerdo a la detección de outliers, se puede apreciar que se desanclan 4 ítems. El ajuste de esta prueba es de 0,92.

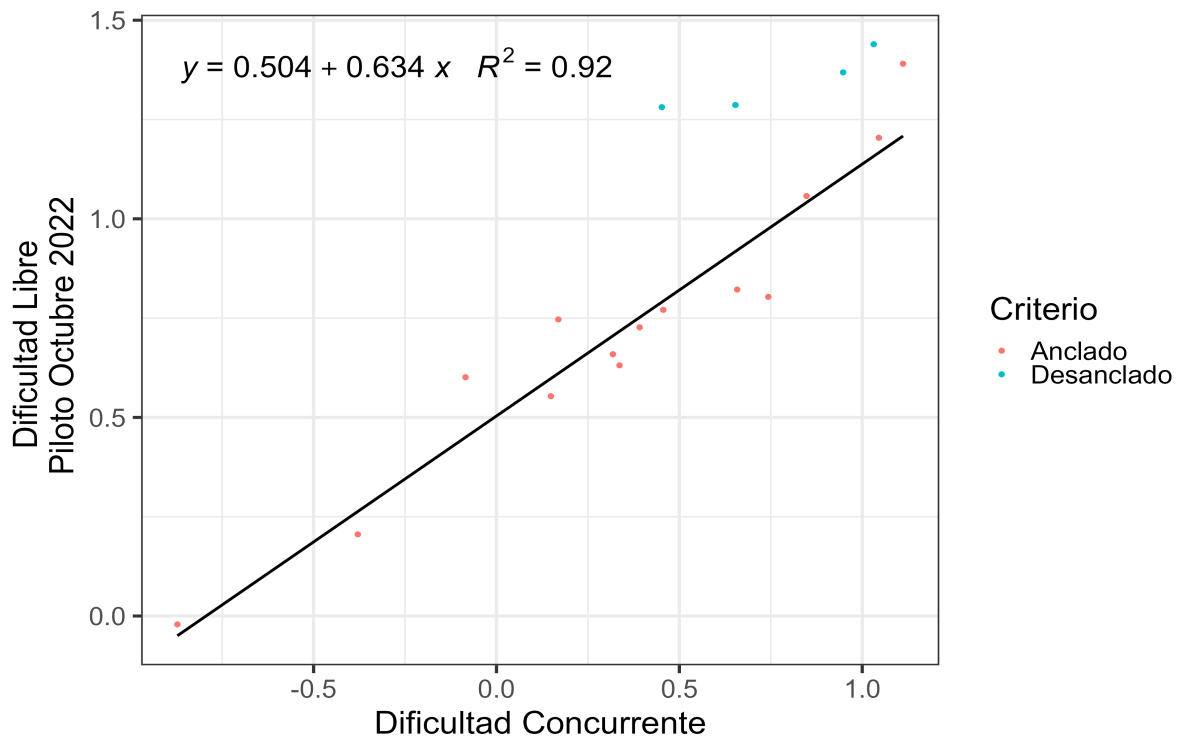


Figura 7: Funcionamiento de ítems anclas en Simulación PAES del Mes de octubre

⁵Los análisis para las pruebas M1 y M2 se realizaron de manera conjunta pero en este informe se muestran de forma separada.

3.2.2. Estado de ítems

En la prueba de Competencia Matemática 1 (M1), se incluyeron 880 nuevos ítems como parte del piloto. De estos, se aprobaron 725, lo que representa un 82,4 % de éxito. El índice de aprobación de los ítems obtuvo un rango del 72 % al 92 %, dependiendo del eje temático abordado.

En cuanto a la evaluación de habilidades, se incluyeron 377 ítems para evaluar la habilidad de resolución de problemas, de los cuales se aprobó un 84 %. Además, se incluyeron 91 ítems para evaluar la habilidad de argumentación, de los cuales se aprobó el 86 %. Se pueden observar más detalles en las figuras 8, 9 y 10.

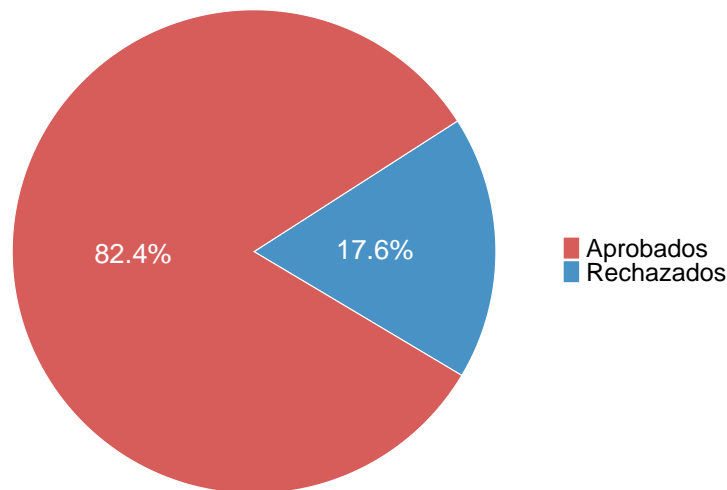


Figura 8: Distribución de ítems aprobados de Competencia Matemática 1 (M1)

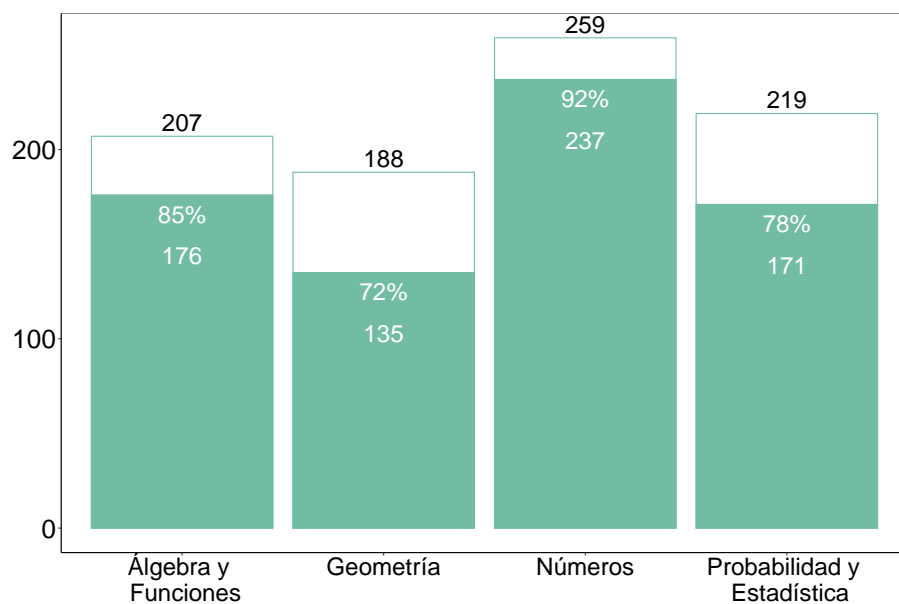


Figura 9: Distribución de ítems aprobados según eje temático

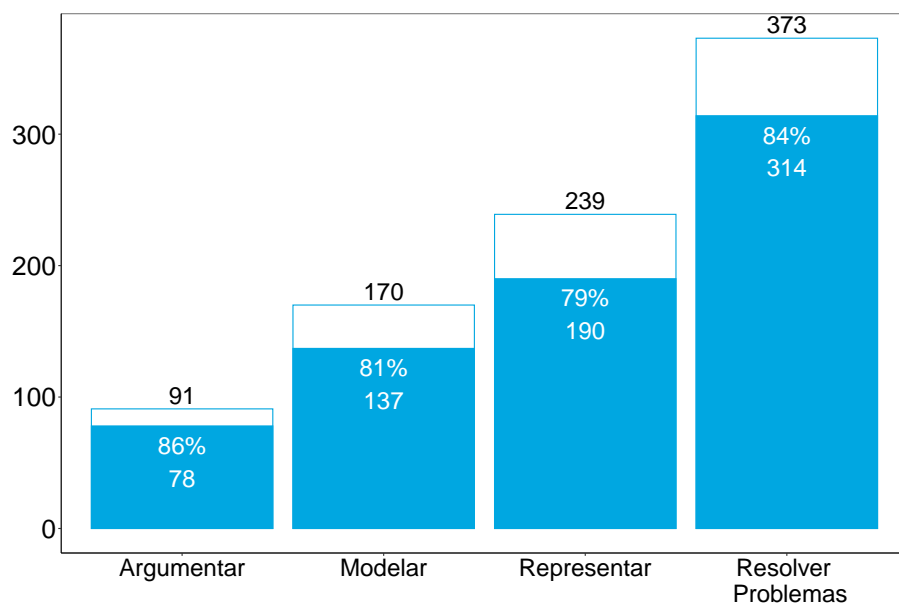


Figura 10: Distribución de ítems aprobados según habilidad

3.3. Competencia Matemática 2 (M2)

Las confiabilidades de los instrumentos oscilaron entre 0.76 y 0.83 aproximadamente, entre las formas piloteadas. A continuación se presenta un resumen de los ítem utilizados.

Grupo	Número de Ítems
Cantidad formas	18
Cantidad ítems	815
Ítems nuevos	783
Ítems ancla	32
Ítems aceptados	666
Ítems rechazados	117

Cuadro 10: Número de ítems y formas de la Simulación PAES prueba de Competencia Matemática 2 (M2)

3.3.1. Funcionamiento de ítems anclas

El siguiente gráfico de Competencia Matemática 2⁶ muestra el funcionamiento de los ítems ancla en el piloto de abril donde se puede observar la relación lineal entre las dificultades estimadas en el piloto (eje Y) con las dificultades de la calibración del banco de ítems (eje X). De acuerdo a la detección de outliers, se puede apreciar que se desanclan 5 ítems. El ajuste de esta prueba es de 0,65.

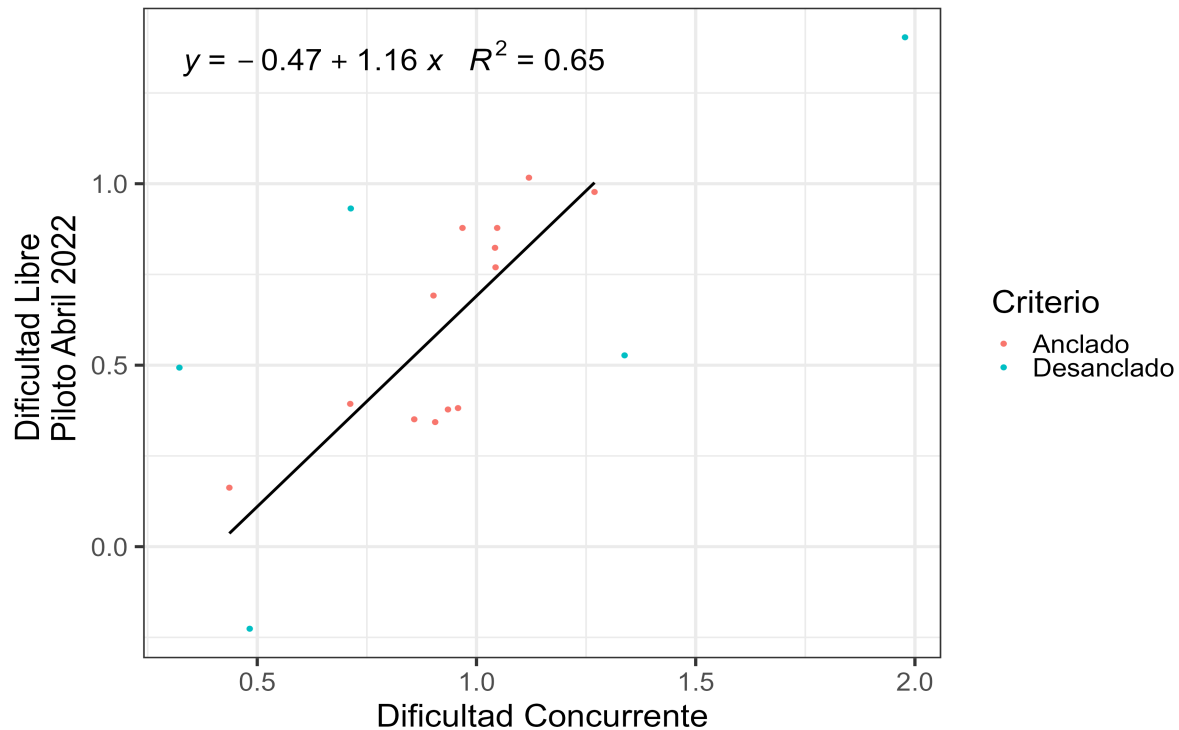


Figura 11: Funcionamiento de ítems anclas en la Simulación PAES del mes de abril

⁶Los análisis para las pruebas M1 y M2 se realizaron de manera conjunta pero en este informe se muestran de forma separada.

El siguiente gráfico de Competencia Matemática 2⁷ muestra el funcionamiento de los ítems ancla en el piloto de octubre donde se puede observar la relación lineal entre las dificultades estimadas en el piloto (eje Y) con las dificultades de la calibración del banco de ítems (eje X). De acuerdo a la detección de outliers, se puede apreciar que se desanclan 6 ítems. El ajuste de esta prueba es de 0,95.

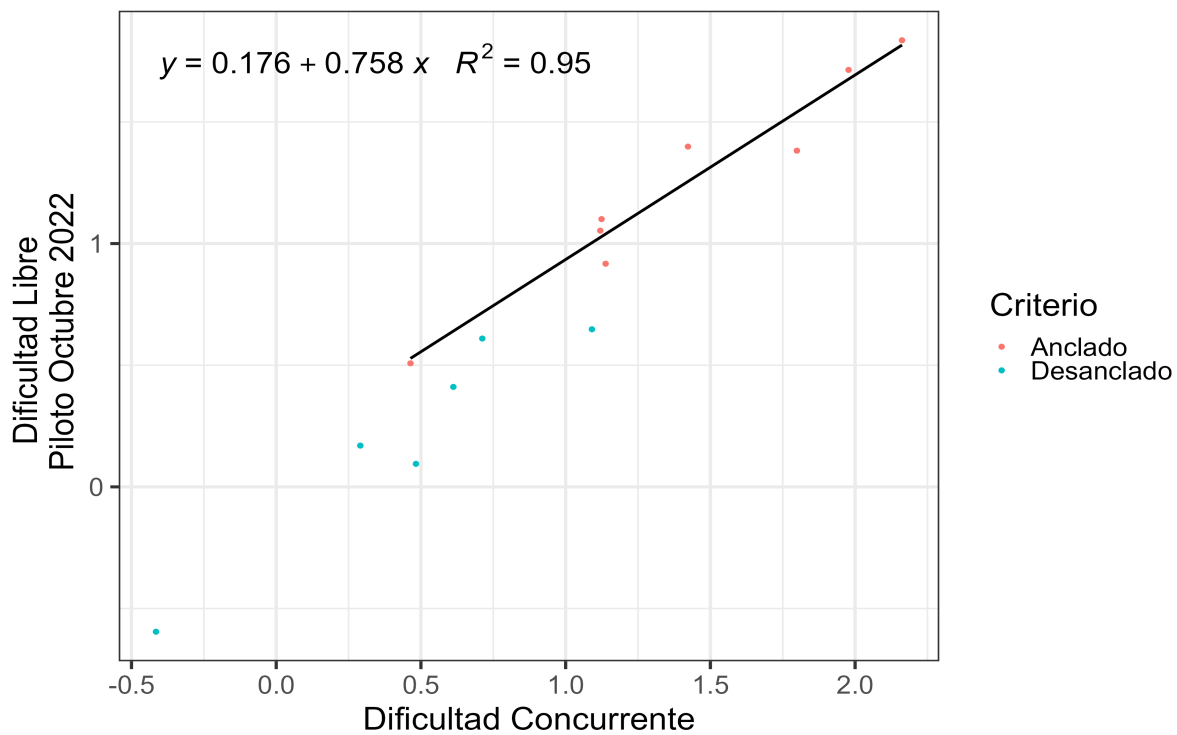


Figura 12: Funcionamiento de ítems anclas en la Simulación PAES del mes de octubre

⁷Los análisis para las pruebas M1 y M2 se realizaron de manera conjunta pero en este informe se muestran de forma separada.

3.3.2. Estado de ítems

En la prueba de Competencia Matemática 2 (M2) se agregaron 783 nuevos ítems en el piloto. De estos, 666 fueron aprobadas, lo que representa una aprobación del 85,1%. La aprobación varía según el eje temático. Álgebra y Funciones, así como Números, tuvieron un rango de aprobación entre el 87% y el 89%. En contraste, Geometría, y Probabilidad y Estadística obtuvieron una aprobación del 79%.

En cuanto a las habilidades para resolver problemas y Representar, se logró un 85% de aprobación. Sin embargo, en cuanto a Argumentar y Modelar, se obtuvieron porcentajes de aprobación de 82% y el 87%, respectivamente. Para obtener más detalles, se pueden observar las figuras 13, 14 y 15.

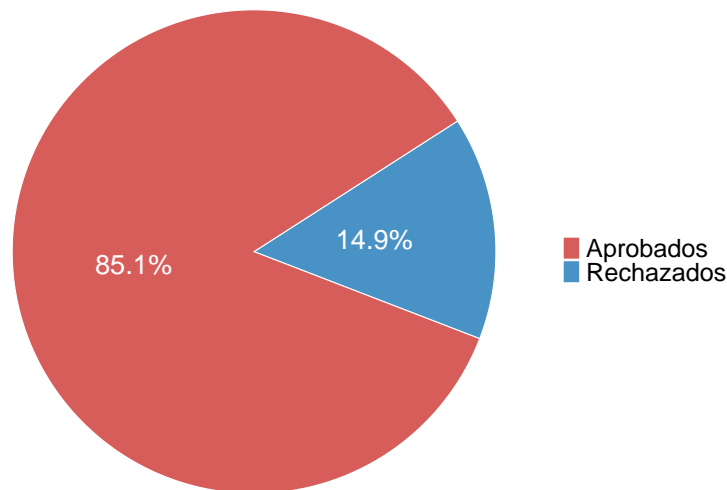


Figura 13: Distribución de ítems aprobados en la prueba de Competencia Matemática 2 (M2)

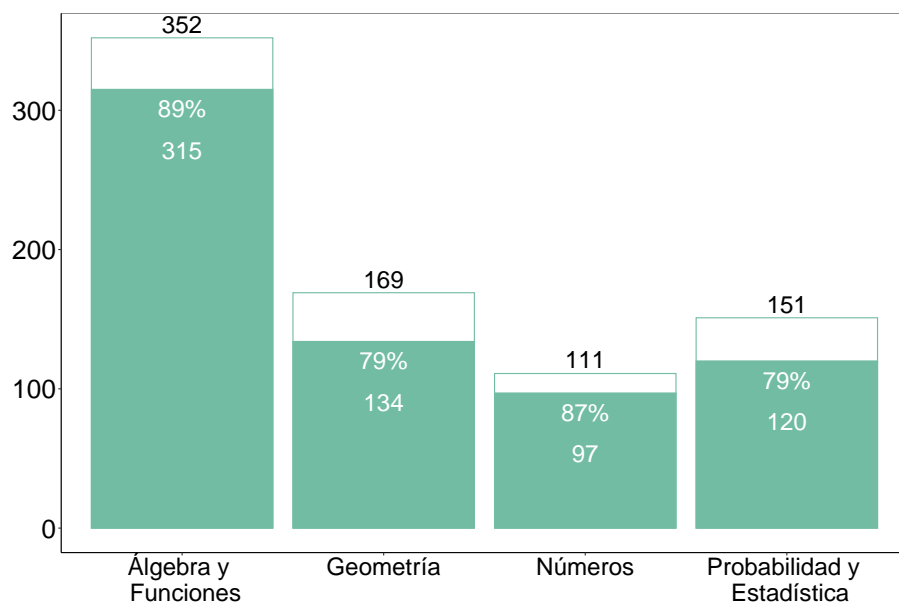


Figura 14: Distribución de ítems aprobados según eje temático

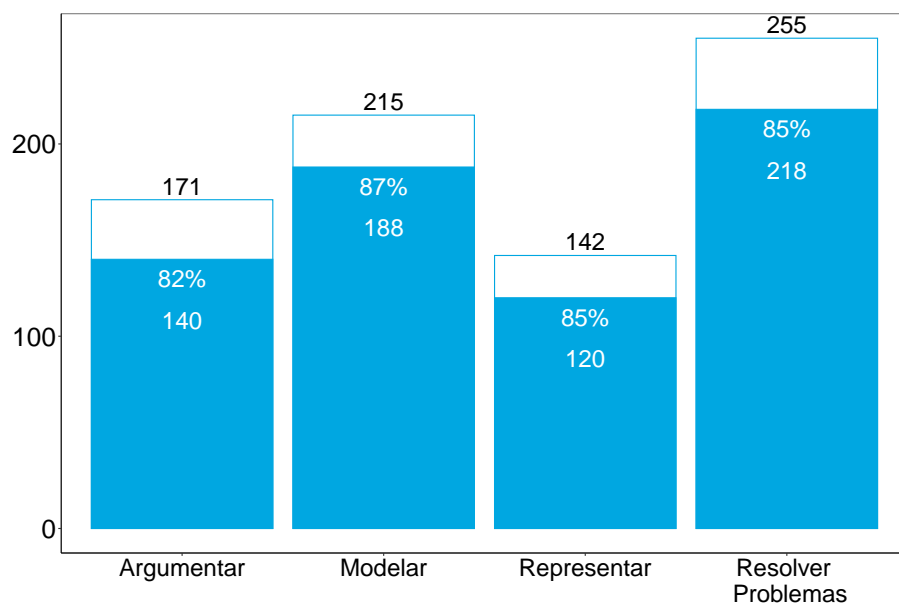


Figura 15: Distribución de ítems aprobados según habilidad

3.4. Historia y Ciencias Sociales

Las confiabilidades de los instrumentos oscilaron entre 0.73 y 0.88 aproximadamente, entre las formas piloteadas. A continuación se presenta un resumen de los ítems utilizados.

Grupo	Número de Ítems
Cantidad formas	19
Cantidad ítems	969
Ítems nuevos	916
Ítems ancla	43
Ítems aceptados	752
Ítems rechazados	164

Cuadro 11: Número de ítems y formas de Simulación PAES en Prueba de Historia

3.4.1. Funcionamiento de ítems anclas

El siguiente gráfico muestra el funcionamiento de los ítems ancla en el piloto de abril donde se puede observar la relación lineal entre las dificultades estimadas en el piloto (eje Y) con las dificultades de la calibración del banco de ítems (eje X). De acuerdo a la detección de outliers, se puede apreciar que se desanclan 10 ítems. El ajuste de esta prueba es de 0,93.

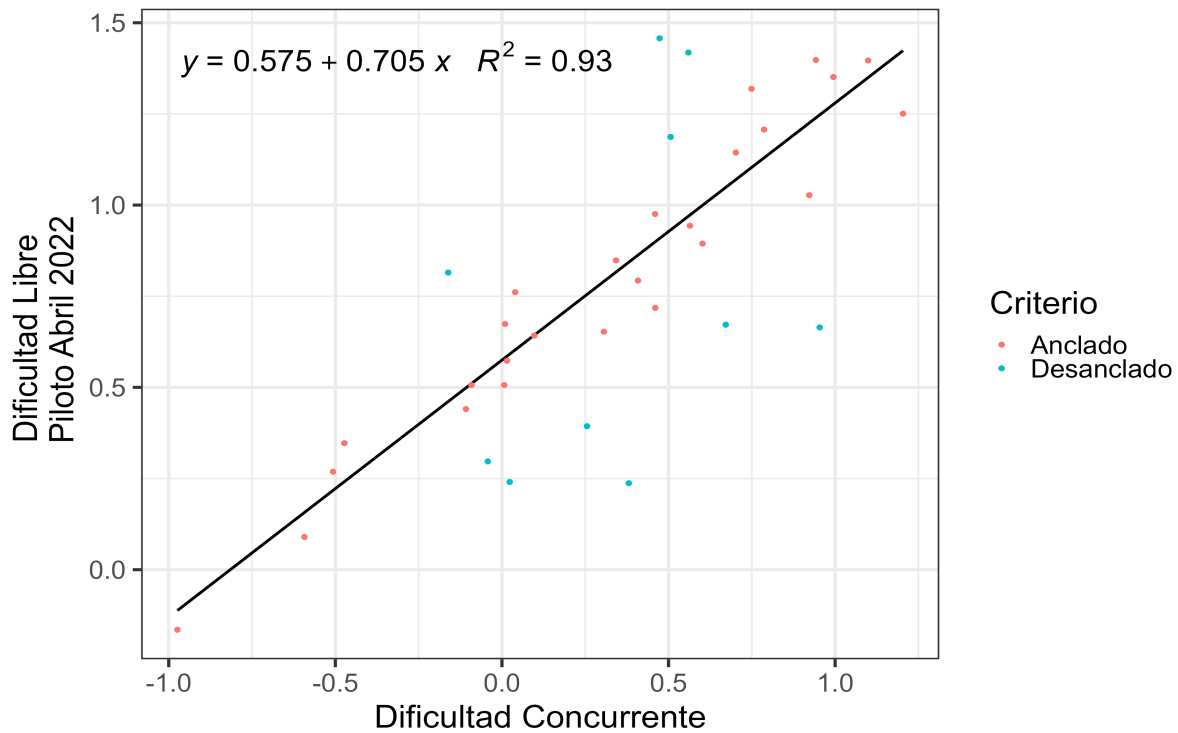


Figura 16: Funcionamiento de ítems anclas en Simulación PAES del mes de abril.

El siguiente gráfico muestra el funcionamiento de los ítems ancla en el piloto de octubre donde se puede observar la relación lineal entre las dificultades estimadas en el piloto (eje Y) con las dificultades de la calibración del banco de ítems (eje X). De acuerdo a la detección de outliers, se puede apreciar que se desancla 1 ítem. El ajuste de esta prueba es de 0,80.

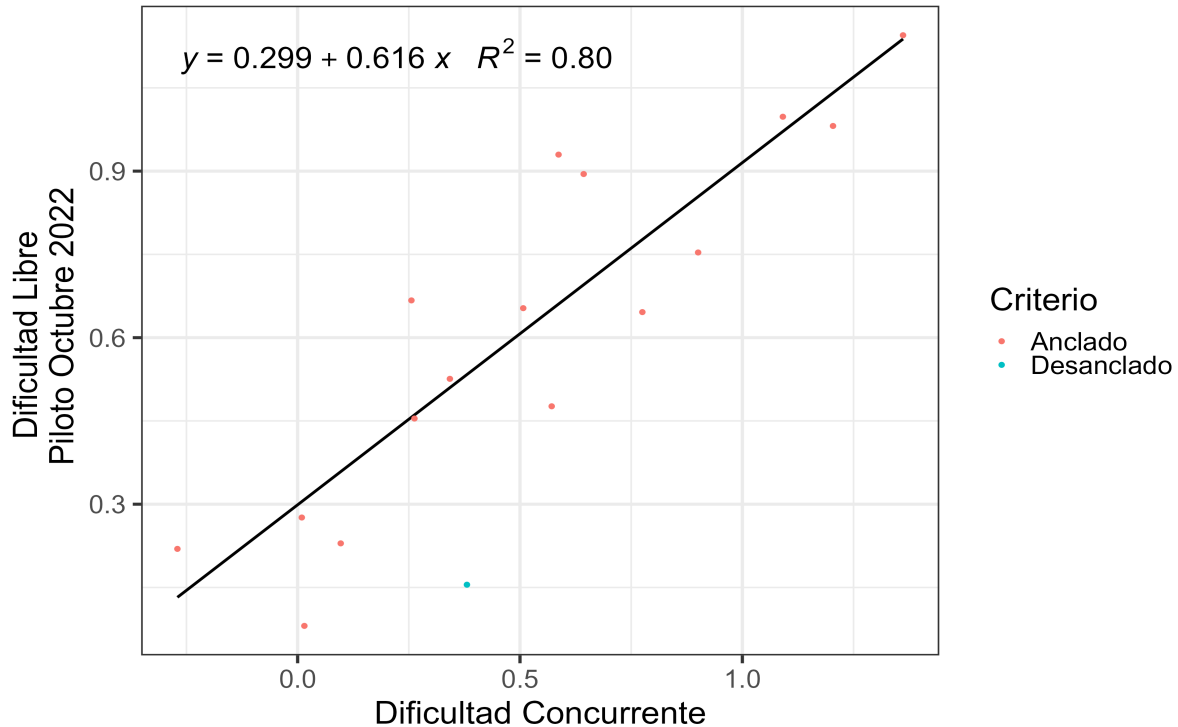


Figura 17: Funcionamiento de ítems anclas en Simulación PAES del mes de octubre

3.4.2. Estado de ítems

En la prueba de Historia y Ciencias Sociales se incluyeron un total de 916 nuevos ítems, de los cuales 752 fueron aprobados, logrando un alto índice de aprobación del 82,1%. En cuanto a los ejes temáticos evaluados, se obtuvo un 76 % de aprobación en Historia: Mundo, America y Chile. En Economía y Sociedad, así como en Formación Ciudadana, se alcanzaron puntajes aún más altos, con un 91 % y un 94 % de aprobación, respectivamente.

Respecto a las habilidades evaluadas, el índice de aprobación varió entre un 75 % y un 88 %, según se puede apreciar en las figuras 18, 19 y 20.

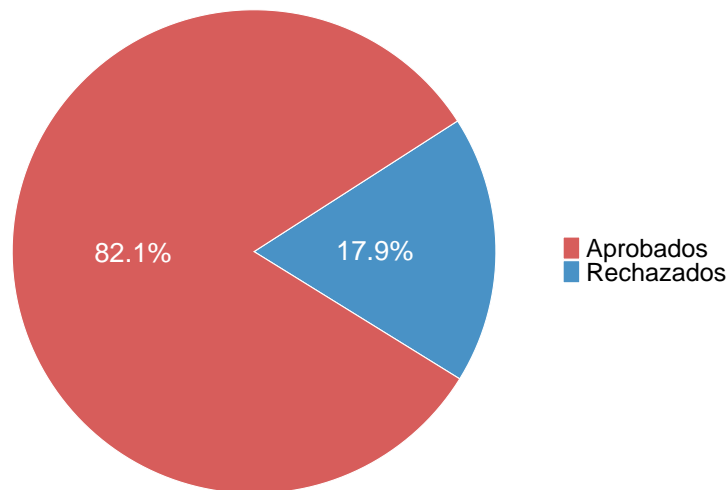


Figura 18: Distribución de ítems aprobados de Historia y Ciencias Sociales

En este punto de medición no se pilotearon ítem del eje temático de Formación Ciudadana, ya que la muestra se limitó únicamente a estudiantes de tercero medio de establecimientos educativos.

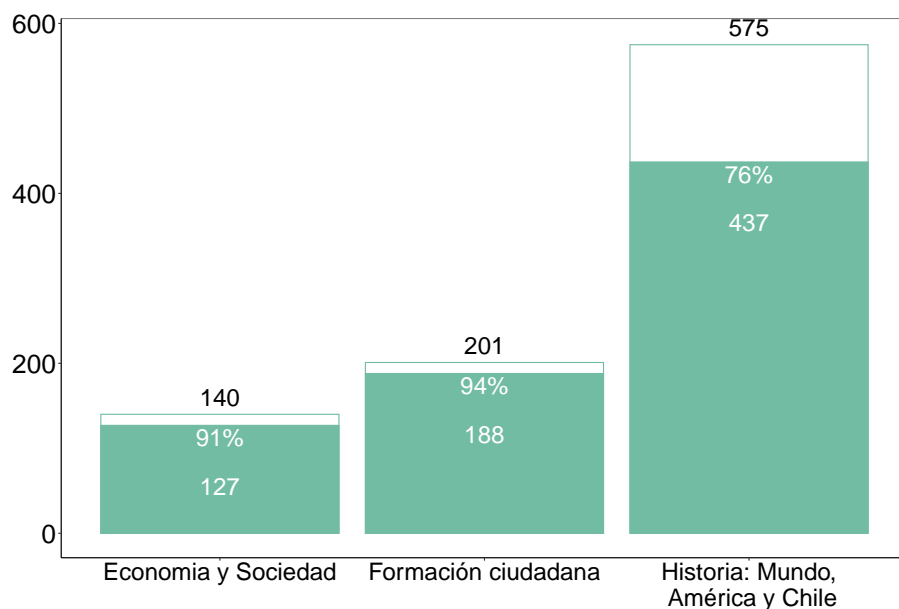


Figura 19: Distribución de ítems aprobados según eje temáticos

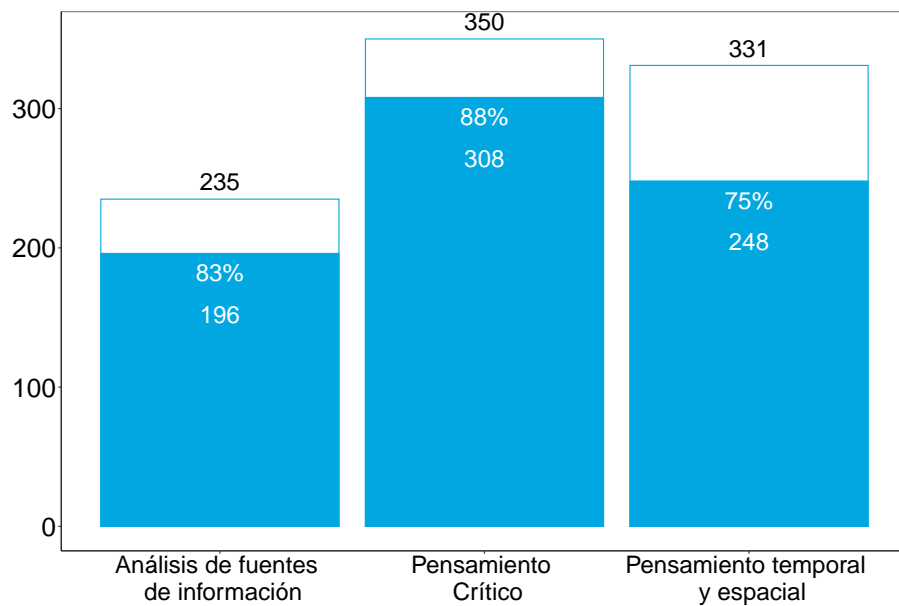


Figura 20: Distribución de ítems aprobados según habilidad

3.5. Ciencias

Las confiabilidades de los instrumentos oscilaron entre 0.78 y 0.88 aproximadamente, entre las formas piloteadas. A continuación se presenta un resumen de los ítems utilizados.

Grupo	Número de Ítems
Cantidad formas	35
Cantidad ítems	1893
Ítems nuevos	1839
Ítems ancla	54
Ítems aceptados	1423
Ítems rechazados	416

Cuadro 12: Número de ítems y formas de Simulación PAES en prueba de Ciencias

3.5.1. Funcionamiento de ítems anclas

El siguiente gráfico muestra el funcionamiento de los ítems ancla en el piloto de abril donde se puede observar la relación lineal entre las dificultades estimadas en el piloto (eje Y) con las dificultades de la calibración del banco de ítems (eje X), De acuerdo a la detección de outliers, se puede apreciar que se desanclan 2 ítems. El ajuste de esta prueba es de 0,90.

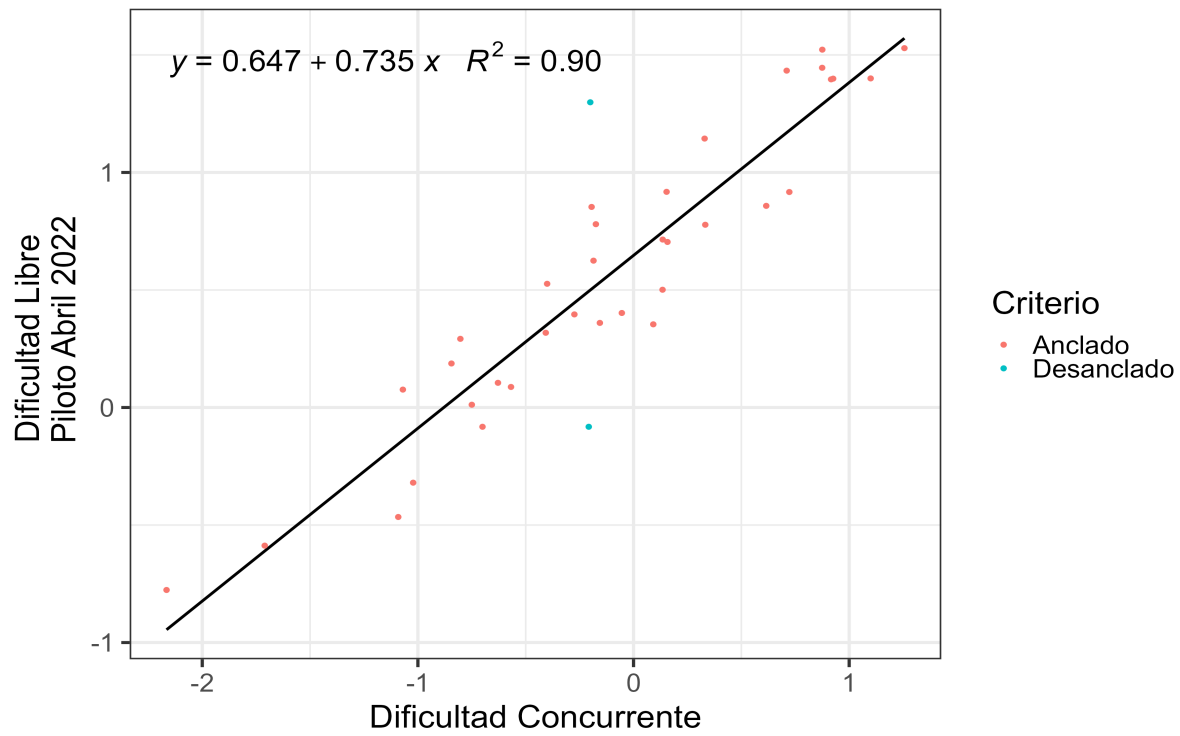


Figura 21: Funcionamiento de ítems anclas en Simulación PAES del mes de abril

El siguiente gráfico muestra el funcionamiento de los ítems ancla en el piloto de octubre donde se puede observar la relación lineal entre las dificultades estimadas en el piloto (eje Y) con las dificultades de la calibración del banco de ítems (eje X). De acuerdo a la detección de outliers, se puede apreciar que se desancla 1 ítem. El ajuste de esta prueba es de 0,96.

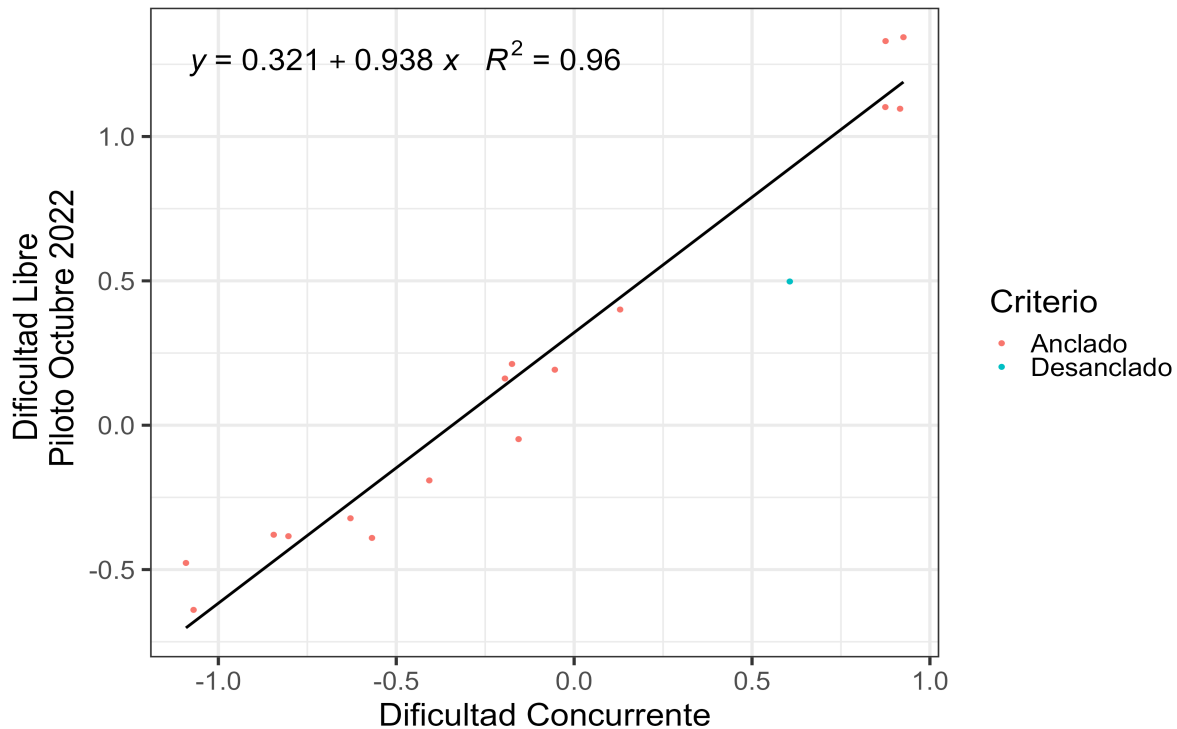


Figura 22: Funcionamiento de ítems anclas en Simulación PAES del mes de octubre

3.5.2. Estado de ítems

En la prueba de Ciencias, se incluyeron un total de 1839 nuevos ítems. De ellas, se aprobaron 1423, alcanzando una tasa de éxito del 77,4%. En cuanto a las disciplinas específicas, se obtuvo un 67% de aprobación en Física, mientras que Biología y Química lograron un 82% y 83% de aprobación, respectivamente.

Dentro del área de Biología, los procesos y funciones biológicas obtuvieron la mayor tasa de aprobación con un 88%, mientras que el resto de las áreas oscilaron entre el 79% y el 81% de aprobación.

En el caso de Física, en el área de mecánica se obtuvo un 61% de aprobación, mientras que electricidad y magnetismo, energía u ondas obtuvieron un 68% de aprobación.

En Química, en el área de química orgánica obtuvo una aprobación de un 79%, mientras que el resto de las áreas obtuvieron un 82% y un 88% de aprobación, tal como se puede apreciar en las figuras 23, 24, 25, 26, y 27.

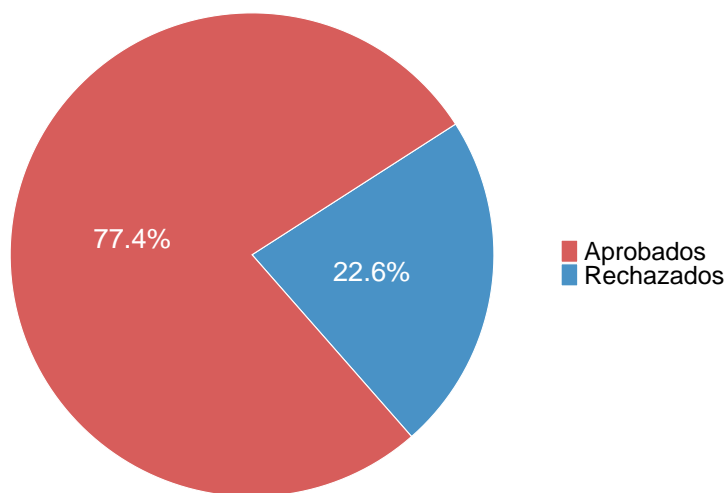


Figura 23: Distribución de ítems aprobados de Ciencias

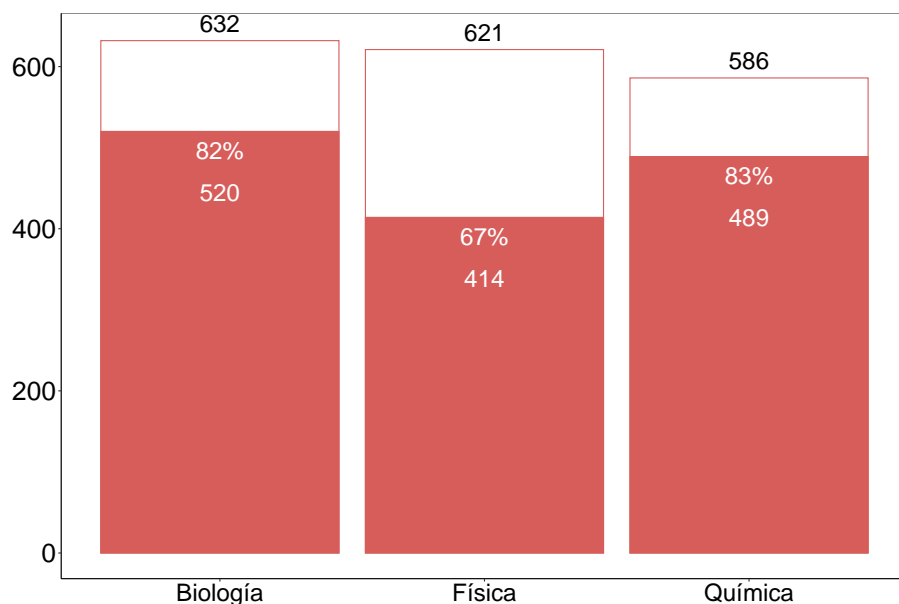


Figura 24: Distribución de ítems aprobados según disciplina

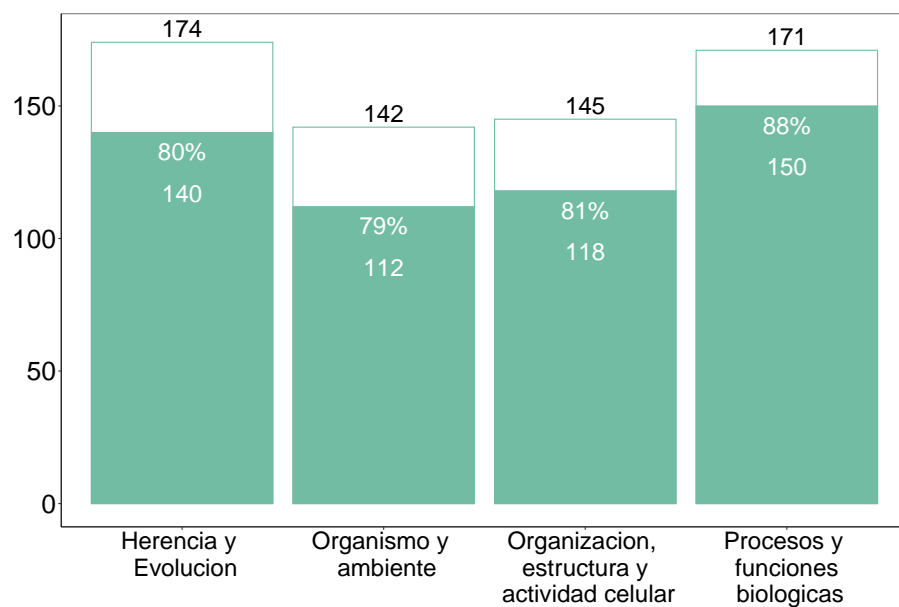


Figura 25: Distribución de ítems aprobados Ciencias Biología

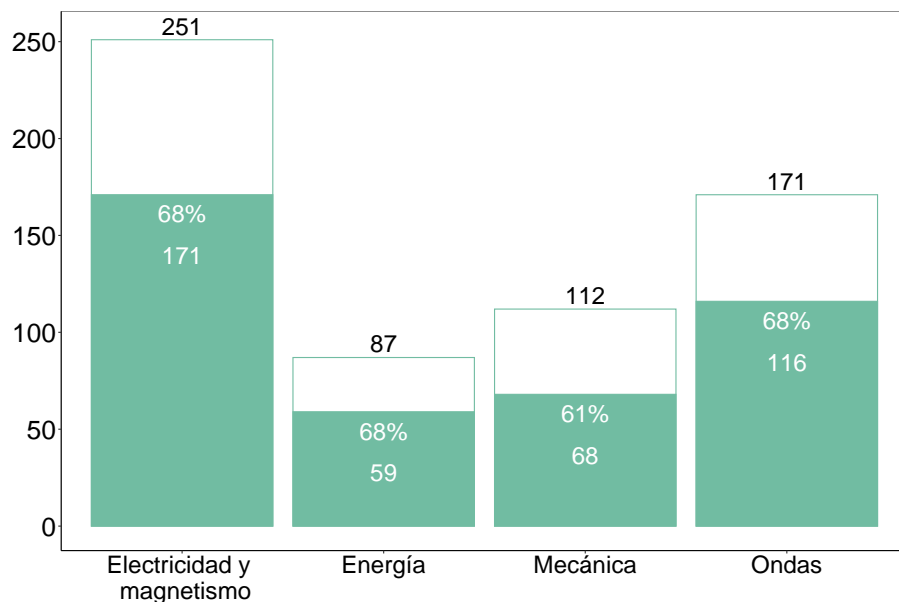


Figura 26: Distribución de ítems aprobados Ciencias Física

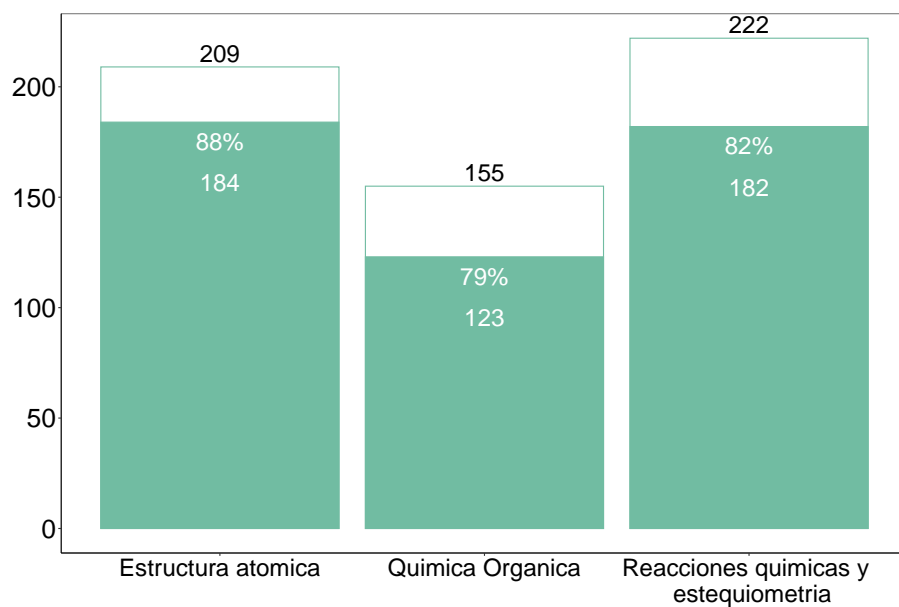


Figura 27: Distribución de ítems aprobados Ciencias Química

4. Anexo

4.1. Anexo A: Modelo de Rasch

El modelo de Rasch al igual que otros modelos, se enmarca en la Teoría de Respuesta al ítem, TRI en español o IRT en inglés. En éste la probabilidad de responder correctamente una pregunta se modela a través de una función de los parámetros de habilidad de las personas y dificultad de los ítems. De esta manera, X_{ij} es una variable aleatoria definida como 1 si el i -ésimo examinado contesta correctamente el j -ésimo ítem y 0 si no, entonces, se define la probabilidad, p_{ij} , de que una persona i responda correctamente el ítem j , como:

$$p_{ij} = p(X_{ij} = 1 \mid \theta_i, \delta_j) = \frac{e^{(\theta_i - \delta_j)}}{1 + e^{(\theta_i - \delta_j)}} \quad (1)$$

donde,

- $i \in \{1, 2, \dots, M\}$, donde M es la cantidad de postulantes.
- $j \in \{1, 2, \dots, L\}$, donde L es la cantidad de ítems.
- θ_i es la habilidad de la i -ésima persona.
- δ_j es la dificultad del j -ésimo ítem.

Comúnmente, en las pruebas educativas los parámetros de los ítems representan la dificultad de estos, mientras que los parámetros de las personas representan la capacidad o el nivel de logro de las personas evaluadas. Cuanto mayor sea la capacidad de una persona en relación con la dificultad de un ítem, mayor será la probabilidad de una respuesta correcta en ese ítem. Dicho de otra manera, si $\theta_i = \delta_j$ en ecuación (1), entonces $p_{ij} = \frac{1}{2}$, mientras mayor sea el valor del parámetro de dificultad del ítem, mayor habilidad es requerida para poder contestarlo de manera correcta. Es decir, para cierta habilidad θ ,

$$si \quad \delta_1 > \delta_2 \Rightarrow p(X_{,1} = 1 \mid \delta_1, \theta) < p(X_{,2} = 1 \mid \delta_2, \theta) \quad (2)$$

4.1.1. Método de estimación de las dificultades

Sea $\tilde{x}_i = (x_{i1}, x_{i2}, \dots, x_{iL})$ el vector de respuesta de un postulante, donde x_{ij} es 1 si la respuesta del i -ésimo postulante al ítem j es correcta y 0 si no. Entonces tenemos que la probabilidad de que se haya observado ese patrón de respuestas correctas dada la habilidad (θ_i) de un estudiante i y el vector de dificultad de los ítems, $\tilde{\Delta} = (\delta_1, \delta_2, \dots, \delta_L)$ es:

$$p(\tilde{x}_i \mid \theta_i, \tilde{\Delta}) = \prod_{j=1}^L p_{ij}^{x_{ij}} (1 - p_{ij})^{1-x_{ij}} \quad (3)$$

De esta forma, podemos obtener la probabilidad marginal de obtener el vector de respuesta que viene dada por:

$$p(\tilde{x}) = \int_{-\infty}^{\infty} p(\tilde{x} | \theta, \tilde{\Delta})g(\theta | \tilde{v})d\theta$$

donde $p(\tilde{x})$ es la función de verosimilitud del vector de respuestas definido en (3) y $g(\theta | \tilde{v})$ representa la función de probabilidad de las habilidades, siendo \tilde{v} el vector de parámetros de las habilidades (localización y escala) que se consideraron libre en la estimación.

Utilizando un método iterativo que combina sucesivas aproximaciones de la habilidad es posible obtener las estimaciones para el vector de dificultades.

4.1.2. Método de estimación de las habilidades

La estimación de la habilidad a través de EAP (Expected a Posteriori) está dada por la esperanza a posteriori:

$$\hat{\theta}_i = E(\theta | \tilde{x}) = \frac{\int_{-\infty}^{\infty} \theta p(\tilde{x} | \theta)g(\theta)d\theta}{\int_{-\infty}^{\infty} p(\tilde{x} | \theta)g(\theta)d\theta} \quad (4)$$

donde, $p(\tilde{x} | \theta)$ es la función de verosimilitud (ver ecuación 1) y $g(\theta)$ es la densidad a priori de las habilidades. En otras palabras, se busca el valor de θ que maximiza la probabilidad de haber observado el vector de respuestas que cada postulante obtuvo.

4.1.3. Supuestos del modelo Rasch

1. **Monotonicidad:** Este supuesto corresponde a que la probabilidad de dar una respuesta correcta a un ítem es una función no decreciente del fenómeno latente que se está midiendo, es decir, de la habilidad en nuestro caso.
2. **Independencia local del ítem:** Este supuesto hace referencia a que la probabilidad de que un examinado responda correctamente a un ítem determinado dada la habilidad es independiente de los demás ítems del test.
3. **Unidimensionalidad:** Este supuesto se refiere a que existe un único factor que explica la respuestas de los examinados, o en otras palabras, una única dimensión o variable latente que se pretende “cuantificar” a partir del conjunto de ítems seleccionados en el test.

4.2. Anexo B: Población que rinde por zona geográfica

Las 16 regiones del país fueron divididas en tres zonas geográficas que se describen a continuación:

- Zona Norte
 - Región de Arica y Parinacota
 - Región de Tarapacá
 - Región de Antofagasta
 - Región de Atacama
 - Región de Coquimbo
 - Región de Valparaíso
- Zona Centro
 - Región de Metropolitana de Santiago
- Zona Sur
 - Región del Libertador General Bernardo O'Higgins
 - Región del Maule
 - Región del Ñuble.
 - Región del Biobío
 - Región de La Araucanía
 - Región de Los Ríos
 - Región de Los Lagos
 - Región de Aysén del General Carlos Ibañez del Campo
 - Región de Magallanes y la Antártica Chilena.

5. Glosario

- **Piloto Tradicional:** En este piloto se seleccionan alumnos de enseñanza media que corresponde a 3° y 4° medio de educación.
- **Piloto Individual:** En este piloto se selecciona estudiantes de la promoción del año y la promoción anterior que inscribieron la prueba de Competencia Matemática 2 y la prueba de Ciencias para en el proceso de admisión 2023 en la región Metropolitana y la Araucanía.
- **DEMRE:** Departamento de Evaluación, Medición y Registro Educativo.
- **PAES:** Prueba de Acceso a la Educación Superior.
- **RBD:** Rol Base de Datos, este código numérico es usado por el MINEDUC para identificar el local educacional.
- **Clavijero:** Corresponde a un archivo que contiene para cada pregunta, su alternativa correcta o clave.



 /demre.uchile  /demre_uchile  /DEMREuchile  /demre.uchile  demre.cl