



Referentes teóricos del Proyecto de Inclusión de la evaluación estandarizada

Presidente de la República

Iván Duque Márquez

Ministra de Educación Nacional

María Victoria Angulo González

Publicación del Instituto Colombiano
para la Evaluación de la Educación
(Icfes)

© Icfes, 2020.

Todos los derechos de autor
reservados.

Se agradece la colaboración y
participación en la elaboración de
este documento a:

Manuel Alejandro Amado

Rafael Eduardo Benjumea

Martha Jeanet Castillo

María Camila Devia

César Augusto Garzón

Lucy Johana Jiménez

Angie Lorena Valbuena Rojas

Alfonso Cabanzo Vargas

Yuly Paola Martínez

Néstor Andrés Naranjo

Mariam Pinto Heydler

María del Pilar Soler

Santiago Wills Pedraza

Edición

Juan Camilo Gómez Barrera

Diseño de portada y diagramación

Linda Nathaly Sarmiento Olaya

Fotografía portada

Administración Nacional de
Seguridad Social (2011)

Directora General

Mónica Patricia Ospina Londoño

Secretario General

Ciro González Ramírez

Directora de Evaluación

Natalia González Gómez

Director de Producción y Operaciones

John Múnera Estupiñán

Director de Tecnología

Carlos Alberto Sánchez Rave

Subdirector de Diseño de Instrumentos

Luis Javier Toro Baquero

Subdirectora de Estadísticas

Jeimy Paola Aristizabal Rodríguez

ISBN de la versión digital:

978-958-11-0866-4

Bogotá, D. C., junio de 2020

ADVERTENCIA

Todo el contenido es propiedad exclusiva y reservada del Icfes y es el resultado de investigaciones y obras protegidas por la legislación nacional e internacional. No se autoriza su reproducción, utilización ni explotación a ningún tercero. Solo se autoriza su uso para fines exclusivamente académicos. Esta información no podrá ser alterada, modificada o enmendada.

TABLA DE CONTENIDO

Introducción	6
1. El sistema de evaluación inclusivo e intercultural: sus propósitos y alcances	11
1.1 Inclusión	11
1.2 Interculturalidad	12
2. Criterios y principios del diseño universal de la evaluación	17
2.1 Evaluación inclusiva de la población	17
2.2 Constructo definido con precisión	18
2.3 Ítems no sesgados	19
2.4 Susceptibilidad a las acomodaciones	20
2.5 Instrucciones y procedimientos sencillos, claros e intuitivos	20
2.6 Máxima lecturabilidad y comprensibilidad	21
2.7 Máxima legibilidad	21
3. Diseño de pruebas: DCE y patrones de diseño	23

TABLA DE CONTENIDO

4. Estrategias de accesibilidad	28
4.1 Acomodación	28
4.1.1 Pruebas por computador ...	31
4.2 Modificaciones	33
4.3 Estrategias específicas	35
5. Metodologías de validación cualitativa y cuantitativa	51
Glosario	53
Referencias	56

LISTA DE TABLAS

Tabla 1. <i>Formato de patrones de diseño</i>	25
Tabla 2. <i>Acomodaciones relacionadas con el braille y el lector de pantalla</i>	36
Tabla 3. <i>Acomodaciones en la presentación</i>	39

Introducción

El presente documento tiene como objetivo sentar las bases teóricas para el desarrollo e implementación de pruebas accesibles. Este objetivo parte de un proyecto que pretende dar respuesta no solo a las disposiciones consignadas en el decreto 1421 de 2017, sino también a las necesidades de toda la población en términos del acceso, confiabilidad, validez y entrega de resultados de los exámenes que realiza el Instituto Colombiano para la Evaluación de la Educación (en adelante Icfes). Se espera que este trabajo inicial permita consolidar un trabajo sistémico que abarque todas las actividades asociadas a la evaluación, el análisis de resultados, la aplicación de pruebas y la divulgación de la información. Este propósito que pretendemos alcanzar está basado en los principios de participación plena y efectiva, accesibilidad, autonomía y no discriminación.

En el marco de las políticas nacionales para la educación inclusiva e intercultural, el Icfes se acoge a la iniciativa del Ministerio de Educación Nacional (en adelante MEN) en su propósito de lograr “equidad en el acceso a la educación”, de acuerdo con los lineamientos del Plan Estratégico Institucional 2019-2022. Esta directriz subyace a uno de los lineamientos del Plan Decenal de Educación 2016-2026 (PNDE, 2017) que busca “Construir una sociedad en paz sobre una base de equidad, inclusión, respeto a la ética y equidad de género”, a través de “la inclusión de personas con discapacidad, el desarrollo de modalidades propias y pertinentes para grupos étnicos, el reconocimiento de la diversidad y la restitución

de derechos para niños, jóvenes y adultos en condición de vulneración, evitando cualquier forma de marginación y exclusión” (PNDE, 2017, p. 53).

De acuerdo con esa perspectiva, se retoma la noción de educación inclusiva adoptada en Los lineamientos de la Política de Educación Superior Inclusiva (MEN, 2013), entendida como “la capacidad de potenciar y valorar la diversidad (entendiendo y protegiendo las particularidades), [...] lo cual implica [...] garantizar la participación de la comunidad dentro de una estructura intercultural en los procesos educativos” (p. 17-18). Al ubicar la educación inclusiva en el contexto colombiano, las instituciones estatales han reconocido la necesidad de identificar ciertos grupos poblacionales, que, por sus condiciones sociales, económicas, culturales, políticas, físicas o geográficas, están expuestos de forma más directa o son más vulnerables a la exclusión. En los lineamientos de política para atención educativa a poblaciones vulnerables (MEN, 2005, p. 8-26), se indica que los grupos que cumplen esta condición son

las comunidades étnicas (indígenas, afro- colombianos, raizales y el pueblo Rom), los menores con necesidades educativas especiales (con discapacidad o limitaciones o con talentos o capacidades excepcionales), los afectados por la violencia (población en situación de desplazamiento, menores desvinculados de los grupos armados al margen de la ley e hijos en edad escolar de adultos desmovilizados), los menores en riesgo social (menores trabajadores, adolescentes en conflicto con la ley penal y niños, niñas y adolescentes en protección), los habitantes de frontera y la población rural dispersa (p. 8).

En este escenario, y teniendo en cuenta el propósito del sistema de evaluación, el Icfes propone una evaluación inclusiva que busca disminuir las brechas de inequidad y exclusión, asumiendo que, mediante el pleno ejercicio de derechos, las pruebas, a través de medidas técnicas y metodológicas aplicadas en su diseño, brindarán **acceso** a diversos servicios y beneficios a poblaciones vulnerables. En esa medida, la evaluación inclusiva se construye sobre el lineamiento de la accesibilidad, que “promueve la eliminación de todo tipo de exclusión y discriminación, y propone mecanismos de inclusión para las poblaciones tradicionalmente marginadas del derecho a la educación, tales como la accesibilidad material y accesibilidad económica” (PNDE, 2017, p. 12). Cabe indicar que esta consideración es adoptada por el MEN de la propuesta de la Organización de las Naciones Unidas, en su propósito de darle a la educación un enfoque de derecho.

Ahora bien, en el campo educativo, una prueba se considera accesible si la población a la que se aplica demuestra los conocimientos, habilidades o destrezas (en adelante CHD) que constituyen el dominio evaluado, independientemente de su género, etnia, nivel socioeconómico o discapacidad. Esto se logra siempre y cuando se eliminen o reduzcan las barreras que puedan surgir de la interacción entre las características de la prueba y las características de los evaluados (Beddow *et al*, 2011, p. 163).

En consecuencia, al eliminar dichas barreras se proporcionan las condiciones de evaluación óptimas y estandarizadas que permiten la comparabilidad del desempeño de todos los

estudiantes, en relación con los estándares que son exigidos a nivel nacional para determinar la calidad educativa del país. Tornar accesibles las pruebas —para que más personas puedan interactuar con ellas y obtener resultados que tengan un efecto en su educación— implica tener en cuenta elementos de diseño, desarrollo, implementación, calificación y análisis de resultados. Solo así se obtienen datos que pueden ser analizados y utilizados para atender adecuadamente las necesidades que, a partir de la prueba, son explícitas para las diferentes poblaciones.

En este contexto, y con el propósito de presentar los criterios teóricos, técnicos y metodológicos que fundamentan el diseño de la evaluación inclusiva en Colombia (un proyecto prioritario del Icfes a partir de 2018), en el presente documento se desarrollan seis grandes apartados. En el primero, se caracteriza el sistema de evaluación inclusivo e intercultural, sus propósitos y alcances. En el segundo, se define y delimita la población que será evaluada, los grupos que serán considerados de los que no. En el tercero, se presentan los criterios del diseño universal, una herramienta teórico-técnica que provee los principios que favorecen la inclusión y la **validez** de los resultados de la prueba. En el cuarto, se explica el diseño de la prueba, el modelo de precisión del dominio (diseño centrado en evidencias) y los patrones de diseño. En el quinto, se explicitan las estrategias de accesibilidad para prevenir o eliminar todas las posibles barreras cognitivas, sensoriales y físicas que no tengan relación con el constructo por evaluar, es decir, toda la varianza irrelevante que garantiza mayor accesibilidad a la prueba. En el sexto, se explican las

metodologías de validación cuantitativa y cualitativa que utiliza el lcfes como parte del diseño de las pruebas. Adicionalmente, se incluye un glosario y una bibliografía con los términos y las referencias más importantes que se trabajan en este documento; los conceptos referenciados en el glosario se muestran en el cuerpo del documento en negrilla.

1. El sistema de evaluación inclusivo e intercultural: sus propósitos y alcances

El sistema de evaluación inclusivo se fundamenta a partir de dos enfoques: la inclusión y la interculturalidad. El derecho a una educación para todos, genuinamente inclusiva, se ha adelantado en el país mediante la ampliación de la cobertura, el fomento a la permanencia, la asignación del personal de apoyo docente y material respectivo y el énfasis en una educación pertinente y de calidad. Es en este último aspecto que la evaluación estandarizada juega un papel relevante, ya que los resultados que esta provee son un insumo fundamental en la toma de decisiones sobre la política educativa.

1.1 Inclusión

La educación inclusiva es aquella en donde todos los niños, niñas, adolescentes, jóvenes y adultos, según sus necesidades, intereses, posibilidades y expectativas, independientemente de su género, discapacidad, capacidad o talento excepcional, pertenencia étnica, posición política, ideología, visión del mundo, pertenencia a una comunidad o minoría lingüística, orientación sexual, credo religioso, lengua o cultura, asisten y participan de una educación en la que comparten con pares de su misma edad y reciben los apoyos que requieren para que su educación sea exitosa. Es un proceso permanente, cuyo objetivo es promover el desarrollo, el aprendizaje y la participación de todos y todas sin discriminación o exclusión alguna, garantizando los ajustes razonables requeridos en su proceso educativo, atendiendo sus particularidades y

prestando especial énfasis a quienes por diferentes razones están excluidos o en riesgo de ser marginados del sistema educativo (MEN, 2017, p. 11).

Un examen accesible es el complemento de una educación inclusiva y comparte sus principales características: en una educación inclusiva, la población con discapacidad se considera prioritaria, y los esfuerzos para que cada una de las personas sea productiva y autónoma se deben reflejar en la forma en que se diseña una evaluación. En este sentido, el Icfes asume que el desempeño de todos los estudiantes podría ser óptimo si se tienen los apoyos y ajustes pertinentes para sobrepasar las barreras que existen al momento de presentar una prueba.

1.2 Interculturalidad

Colombia se reconoce como un Estado multiétnico y pluricultural en la Constitución política de 1991 (Castillo y Carou, 2002). Por esta razón, tiene sentido considerar el enfoque intercultural en educación, que ha sido desarrollado por medio de diferentes leyes y reglamentaciones: la Ley 121 de 1991 (mediante la que se adopta el Convenio 169 de la OIT), la Ley General de Educación de 1994 en el artículo 55, el Decreto 804 de 1995, el Decreto 1122 de 1998, la Directiva Ministerial N.º 08 de 2003 y la Directiva Ministerial N.º 12 del mismo año.

En la literatura se encuentran dos formas predominantes de concebir la diversidad: el multiculturalismo (Walsh, 2009), o interculturalismo funcional y el interculturalismo crítico (Tubino,

2004). Por un lado, según Walsh (en Quilaqueo & Torres, 2013), el multiculturalismo se caracteriza por su origen occidental; por describir la presencia de diferentes culturas y evitar las relaciones entre dichas culturas, y por fundamentarse en los principios y estructuras del Estado liberal y en una concepción de la tolerancia que oculta las desigualdades sociales. Según Tubino (2004), esta postura se evidencia en dos tipos de políticas: la discriminación positiva y la educación compensatoria. La primera promueve la equidad de oportunidades sin realizar cambios a la estructura distributiva y la segunda mejora la calidad educativa de pocos sectores periféricos de la sociedad, sin solucionar los problemas de fondo.

Por otro lado, el interculturalismo crítico busca transformar las asimetrías culturales y económicas, a partir del reconocimiento de las condiciones del diálogo entre culturas; es decir, se preocupa por poner sobre la mesa las causas de las desigualdades (Tubino, 2004). Bello (2009) complementa lo anterior: “la interculturalidad aborda las condiciones, los mecanismos y las estructuras que están en su base y que permiten la reproducción de las desigualdades derivadas de la cultura en las esferas económicas, políticas y sociales” (p. 64). En términos educativos, la interculturalidad crítica implica relacionar diferentes culturas, así como construir los mecanismos para lograr una comunicación efectiva entre ellas (Bello, 2009).

Con base en lo anterior, el Icfes pretende desarrollar pruebas que tengan en cuenta a todas las poblaciones y que eliminen todas las barreras de **acceso** posibles sin afectar el

dominio evaluado, lo que podría incidir en la **validez** de los resultados. Para optar por criterios técnicos y metodológicos que garanticen la participación de diversas poblaciones en la evaluación, se han acogido los principios del **diseño universal de la evaluación** (en adelante DUE), dado que, como lo afirman Thompson, Johnstone y Thurlow (2002b), la aplicación de estos principios “permite la participación de la mayor variedad posible de estudiantes, y genera inferencias válidas sobre el desempeño de los estudiantes que participan en la prueba”.

Uno de los propósitos de aplicar este diseño a la evaluación es mejorar el **acceso** a las pruebas, reduciendo la necesidad de ajustes razonables, bien sean estas **acomodaciones** o **modificaciones** que algunos estudiantes suelen requerir para poder demostrar sus competencias al presentar una prueba. Es importante resaltar que las pruebas diseñadas universalmente no pretenden cambiar aquello que se espera medir; es decir, no debilitan el constructo ni son más fáciles para ciertos grupos poblacionales. Más bien, como se mencionó en la introducción, minimizan las barreras que impiden que algunos estudiantes demuestren sus CHD al responder la prueba.

Existen diferentes clases de barreras. Algunas pueden ser *físicas*; esto es, aquellas que impiden el acercamiento, la manipulación o la percepción del cuadernillo del examen. Por ejemplo, un cuadernillo con letra muy pequeña, uno con un diseño difícil de plegar o uno con una hoja de respuestas difícil de completar constituiría una barrera física que impide el acceso a estudiantes que desean presentar el examen.

Otras barreras pueden ser *intangibles*. Aunque una prueba pretenda evaluar un conjunto específico de CHD, esto exige tener habilidades adicionales que no se pretenden evaluar, pero que pueden ser requeridas para responder la prueba. Estas habilidades adicionales se conocen como *habilidades de acceso* o *habilidades no centrales al constructo*. Tener visión de color, una amplia capacidad de memoria de trabajo, ser capaz de realizar silogismos complejos o poseer un amplio vocabulario pueden ser habilidades de acceso o habilidades no centrales al constructo. Estas, en algunos casos, son requeridas (sin intención de los diseñadores de prueba) para entender una instrucción o responder acertadamente a una pregunta, a pesar de que no sean habilidades que una prueba pretenda medir. Dichas habilidades no centrales al **dominio** son barreras *intangibles* que impiden el **acceso** a preguntas o a la prueba, pues, aunque no imponen barreras físicas, sí plantean demandas cognitivas que son innecesarias para resolver una tarea o trabajo propuesto (*Ibidem*, 2011, p. 166).

Buena parte de lo que implica desarrollar una prueba accesible reside en la construcción y diseño de ítems. En otras palabras, se deben elaborar ítems cuyos elementos constitutivos (estímulo, enunciado, representación visual, disposición en el cuadernillo y, en el caso de ítems de selección múltiple, opciones de respuesta) tengan la información y estructura esencial que evidencie el **dominio** evaluado sin imponer cargas cognitivas ajenas a lo que se pretende medir (*Ibidem*, 2011, p. 169).

De acuerdo con lo anterior, una prueba accesible también es aquella que reduce al mínimo las barreras físicas y la

interferencia que las habilidades no centrales al **dominio** puedan ejercer sobre la capacidad de los evaluados para exhibir los CHD que evalúa la prueba. Adicionalmente, una prueba accesible permite realizar inferencias mucho más precisas acerca de los CHD de un estudiante, pues, al contrario de una prueba poco accesible, reduce la carga cognitiva innecesaria impuesta por las habilidades no centrales al dominio que pueden afectar su desempeño. Por tanto, los puntajes de una prueba accesible reflejarán con más claridad el grado de posesión de los CHD que constituyen su constructo, lo que implica una mayor **validez y confiabilidad** de la prueba (*Ibídem*, 2011, p. 179). Por lo anterior, el diseño de exámenes accesibles pertenece a la implementación de un sistema inclusivo de evaluación que implica no solo el diseño de ítems o pruebas, sino también la inscripción, aplicación, calificación y la entrega de resultados.

2. Criterios y principios del diseño universal de la evaluación

Según los planteamientos de Thompson, Johnstone y Thurlow (2002b), el diseño universal aplicado a las evaluaciones estandarizadas incluye el análisis de elementos que tradicionalmente contribuyen a la **validez** y **confiabilidad** de las pruebas a través de siete criterios o principios que ayudan a este propósito. El Icfes, al construir sus pruebas a partir del diseño centrado en evidencias (en adelante DCE), incorpora algunas de estos principios —de hecho, aquellos que conciernen a la **validez** y **confiabilidad**, temas que se tratarán en el tercer apartado de este capítulo—. Más allá de esto, a lo largo del desarrollo de las pruebas, el Icfes ha incorporado los principios del DUE para volver las pruebas accesibles. A continuación, se analizarán cada uno de estos principios.

2.1 Evaluación inclusiva de la población

Los exámenes de calidad que hace el Estado deben aplicarse a todo el estudiantado. Por tanto, deben medir el desempeño de los evaluados independientemente de su situación socioeconómica, su identidad cultural o lingüística. De este modo, el Estado se cerciora de que los estudiantes con diversas necesidades de aprendizaje tengan la oportunidad de demostrar sus competencias con respecto al mismo contenido. Para ello, el ideal es que todo estudiante presente un único examen, de acuerdo con su nivel escolar, para que así todos puedan demostrar sus habilidades en relación con una misma evaluación.

No obstante, esto es materialmente imposible. Es factible que algunos estudiantes necesiten una **evaluación alternativa** a la **evaluación general** para poder demostrar sus logros. Para determinar cuáles estudiantes no pueden demostrar sus competencias mediante una **evaluación general** se usan estrategias basadas en evidencias (perfiles psicológicos y médicos de su condición, etc.) y comisiones de expertos que determinen por qué la prueba general no puede evidenciar sus habilidades. Para determinar lo anterior, también se investiga la relación entre la forma de educación y el currículo, y cómo esta influye y determina la forma de la evaluación. Es aconsejable que los equipos conformados por los docentes, docentes de apoyo, familiares y el estudiante (en nuestro contexto, quienes realizan los planes individuales de ajustes razonables, —PIAR—) decidan en cada caso si deben presentar una **evaluación alternativa** (Thurlow *et al.*, 2016).

2.2 Constructo definido con precisión

Para definir con precisión el constructo evaluado, se utiliza el diseño centrado en evidencias (DCE)¹. Con este modelo se busca garantizar la validez de las afirmaciones sobre los CHD de los estudiantes que se hacen a partir de las respuestas de estos a los ítems. Esto permite responder a las siguientes preguntas: ¿cómo una población heterogénea evidencia sus CHD mediante una prueba general?, ¿qué población requiere pruebas alternativas? y ¿cómo se construyen estas pruebas?

1 Para una descripción detallada del DCE, cf. Mislevy, R. *et al* (2003) y Mislevy, R. J., & Riconscente, M. M. (2005) e Icfes (2019).

Las definiciones de los constructos deben tener la precisión adecuada para evitar que sus elementos se conviertan en barreras. El propósito es prevenir o eliminar todas las posibles barreras cognitivas, sensoriales y físicas que no tengan relación con el constructo por evaluar, es decir, toda la varianza irrelevante. El constructo, como modelo global teórico, se hace concreto al tomar decisiones sobre qué evaluar, a quién evaluar, para qué, cómo, a través de qué medios, qué tipos de tareas, formatos de ítems y qué tipo de inferencias hacer a partir de los resultados. La delimitación del constructo permite hablar de **dominio** u objeto de la evaluación.

2.3 Ítems no sesgados

Un ítem se considera *sesgado* si alguna subpoblación tiene una ventaja o desventaja para responderlo correctamente. Estos sesgos se dan por factores *externos* al proceso de educación. Surgen, particularmente, cuando uno o varios ítems indagan por contenidos, habilidades o destrezas que podrían ser familiares, desconocidos u ofensivos, dadas sus características de género, raza, discapacidad, etc., para un sector de la población. Por ejemplo, si un ítem requiere, para ser respondido, el conocimiento de las reglas de un deporte que solo conoce una pequeña población de una región, y si este reglamento no hace parte del contenido curricular enseñado en todo el país, los evaluados de dicha región tendrán una ventaja. Para determinar cuándo hay un **sesgo** se recurre a procedimientos cualitativos y cuantitativos. Los primeros indagan *a priori* el contenido de estos. Los segundos analizan el comportamiento estadístico una vez presentada la

prueba. En las pruebas basadas en los modelos de respuesta al ítem (TRI), se suele usar el análisis diferencial de los ítems. Este modelo busca determinar, entre otras cosas, si diferentes grupos poblacionales con las mismas habilidades tienen una probabilidad diferente de acertar un mismo ítem.

2.4 Susceptibilidad a las acomodaciones

Según los principios del DUE (Thompson *et al.*, 2002b), tanto los instrumentos de evaluación como los ítems permiten ciertos cambios para acceder adecuadamente a la prueba. Estos cambios se conocen como **acomodaciones** (ver capítulo 4.1). De acuerdo con el DUE, los instrumentos y los ítems de la evaluación son susceptibles de **acomodaciones** para garantizar el **acceso** de la población, sin que haya necesidad de efectuar adaptaciones adicionales. Esto quiere decir que, una vez un ítem sea construido, si se han aplicado estos principios en su diseño, podría acomodarse a otros formatos —braille, por ejemplo—. Otro tanto sucede con las gráficas: si el contexto de un ítem construido bajo los principios del DUE presenta imágenes, estas deben ser susceptibles de describirse de manera verbal o textual.

2.5 Instrucciones y procedimientos sencillos, claros e intuitivos

Cada instrumento debe contar con instrucciones y procedimientos que sean entendibles y aplicables a poblaciones diversas. Esto permite el acceso físico a las pruebas. Por tanto, los formatos de la prueba deben responder

a la pregunta: ¿es posible para todos los estudiantes “navegar” de manera independiente la prueba?

2.6 Máxima lecturabilidad y comprensibilidad

La lecturabilidad es una propiedad textual que determina la facilidad para acceder al contenido de los textos. Esta accesibilidad depende de elementos sintácticos, semánticos y pragmáticos. Dentro de los primeros, tenemos, entre otros, las construcciones gramaticales sencillas de la forma sujeto-verbo-predicado; la separación de las ideas con punto seguido y no con comas, y el uso restringido de oraciones subordinadas, palabras cortas y de pocas sílabas. Los elementos semánticos que se deben tener en cuenta son, por ejemplo, el vocabulario no especializado y la elección de palabras de alta frecuencia de uso. Los elementos pragmáticos incluyen evitar el uso del lenguaje figurado, el doble sentido y los juegos de palabras. Solo en la medida en que la prueba busque examinar las habilidades para interpretar estas expresiones, los ítems deben manejar esos elementos pragmáticos. Por esto, es aconsejable evitar tales recursos lingüísticos en pruebas de ciencias y matemáticas.

2.7 Máxima legibilidad

La legibilidad está relacionada con el principio anterior, pero atañe a la forma de presentación de los ítems; refiere a la capacidad para decodificar con facilidad un texto, una gráfica o el formato de respuesta. En el caso de los textos, esto involucra aspectos como el tamaño y tipo de letra, el contraste, etc. Para el caso de las gráficas, tablas e ilustraciones, además

de tener en cuenta estos aspectos, debe verificarse su relación y su pertinencia con respecto al ítem que las contiene. En lo que se refiere al formato de respuesta, no se debe desfavorecer a personas con discapacidades visuales, auditivas o motrices. Para poner un ejemplo, diferentes estudios han revelado que las “burbujas” circulares pequeñas podrían afectar las respuestas de las poblaciones con esas discapacidades y, por tanto, se sugiere ubicar las opciones de respuesta en un formato vertical con óvalos aplanados horizontales (Grise *et al.*, 1982, Willingham *et al.*, 1988, Thompson *et al.*, 2002).

Para sintetizar, es importante aclarar que la aplicación de los principios del DUE en el diseño de la prueba no implica un cambio en los dominios evaluados, definidos en los marcos teóricos de cada área (Tindal *et al.*, 1998; Hollenbeck, 2002). El propósito del DUE es permitir a los estudiantes con discapacidad acceder a la prueba para demostrar su desempeño en los CHD evaluados. Para ello se eliminan, hasta donde sea posible, las interferencias generadas cuando el estudiante con discapacidad interactúa con una prueba estandarizada que no fue diseñada de forma especial para él.

3. Diseño de las pruebas: DCE y patrones de diseño

Una prueba estandarizada, como las que realiza el Icfes, busca elaborar inferencias sobre los conocimientos, habilidades y destrezas (CHD) de los estudiantes a partir de las respuestas a las preguntas que conforman el instrumento.

El diseño de evaluaciones del Icfes se fundamenta en el DCE, el cual es un enfoque para el diseño de pruebas. El objetivo de este diseño es garantizar que exista una línea clara de razonamiento entre las respuestas que dan los estudiantes a los ítems y las afirmaciones que, a partir de esas respuestas, se pretenden realizar sobre los CHD evaluados en la prueba. Dado que las afirmaciones sobre los CHD de los estudiantes deben estar soportadas por evidencias suficientes e independientes, el DCE propone que, para recoger estas evidencias, cada ítem de la prueba demande el cumplimiento de una determinada tarea: una situación específica en la que el estudiante pueda exhibir parte de los CHD evaluados.

Como parte de las estrategias para delimitar los CHD que pertenecen al **dominio** de una prueba y poder garantizar que los ítems que la conforman evalúen lo que se pretende evaluar, el DCE recomienda el desarrollo de **patrones de diseño** (Mislevy *et al.*, 2017, p. 22). Dentro del DCE, los patrones de diseño brindan opciones para obtener evidencias de los CHD a partir del diseño de las tareas desarrolladas para la evaluación. Estos pueden ayudar a lograr los objetivos de las tareas que componen las especificaciones a la luz de

consideraciones como la accesibilidad. Estos patrones permiten presentar los ítems teniendo en cuenta características de los evaluados asociadas a la discapacidad. En otras palabras, un patrón de diseño es una estructura que permite organizar información para llenar los “vacíos” del argumento evidencial de la prueba (Hansen & Mislevy, 2008). Estos vacíos tienen que ver con:

- a) Información sobre los CHD evaluados por un ítem (o un conjunto de ellos).
- b) CHD requeridos para resolver un ítem, pero que son irrelevantes para el constructo evaluado (no centrales al constructo).
- c) Formas de representar el contenido del ítem que se pueden modificar para cambiar la dificultad del ítem, sin alterar el dominio o imponer una carga cognitiva innecesaria.

Actualmente, las pruebas de Matemáticas, Lectura crítica, Sociales y ciudadanas, Ciencias naturales e inglés del examen Saber 11.º están desarrollando patrones de diseño en un formato que especifica la información descrita en los puntos a), b) y c). Las definiciones de lo que se pretende medir y de lo que no es central contribuyen de manera significativa al desarrollo de ítems que sean más accesibles y que, por tanto, también contribuyan a la **validez** de la prueba. A continuación, se muestra un modelo que utilizan las áreas para el desarrollo de los patrones de diseño:

Tabla 1. *Formato de patrones de diseño*

Prueba
Patrón de diseño N.º
Título Nombre corto que referencia el patrón
Resumen Descripción del criterio de agrupación de ítems a los que se aplica el patrón; por ejemplo, ítems agrupados por tema, por tarea, por uso de ciertos tipos de representación (ecuaciones, fórmulas, matrices, tortas, cómics, lecturas, etc.)
Fundamento Descripción de lo común de los CHD por evaluar y cómo se manifiestan. En esta sección se incluye la justificación de la agrupación; p.ej.: la comprensión de la relación entre elementos lingüísticos y no lingüísticos en un texto discontinuo (tabla, infografía, etc.) muestra que un estudiante es capaz de realizar inferencias simples y detectar supuestos implícitos, que son, justamente, habilidades focales del constructo.
Conocimientos, habilidades y destrezas (CHD) evaluadas Descripción de los CHD focales (centrales al constructo) que se evalúan.

Continúa en la siguiente página

Conocimientos, habilidades y destrezas (CHD) adicionales

Descripción de CHD adicionales requeridas para responder el o los ítems, pero que no son focales.

Representación de la información

Especificación de la forma de representación privilegiada para la construcción del ítem; por ejemplo, uso de cifras (55) sobre uso de descripción con palabras (cincuenta y cinco).

Las inferencias que se realicen de las respuestas de los evaluados deben tener dos propiedades la **confiabilidad** y la **validez**. La **confiabilidad** se refiere a la consistencia: una prueba, realizada en diferentes tiempos por diferentes personas y aplicada a poblaciones similares, debe dar resultados consistentes. Esta propiedad se relaciona de manera inversa con el **sesgo**, pues, a medida que aumenta la confiabilidad de una prueba, disminuye su sesgo (Groth-Marnat, 2003). Un **sesgo** se entiende aquí como un direccionamiento, bien sea de la prueba o de uno o varios ítems, que favorece o desfavorece a una parte de la población. Dicha orientación se manifiesta en la presuposición de ciertos CHD necesarios para responder los ítems, típicos de un subgrupo, que no son evaluados explícitamente por la prueba y cuya tenencia o carencia facilita o entorpece las respuestas. Ejemplo de esto son los ítems que requieren conocimiento de elementos culturales. Una prueba confiable, entonces, evita este tipo de sesgos que dependen de las consideraciones del grupo que construye la prueba en determinado momento.

La **validez**, por su parte, se refiere a la propiedad que permite que una prueba evalúe exactamente lo que se desea evaluar; en este caso, los conocimientos, habilidades y destrezas que debe haber desarrollado el estudiante tras el proceso de escolarización. Esto quiere decir que las inferencias sobre los CHD que se extraen a partir de las respuestas de los evaluados deben ser válidas (Groth-Marnat, 2003). En este sentido, se habla de diferentes tipos o etapas del establecimiento de una prueba: la **validez** del contenido, esto es, el establecimiento previo de los CHD que serán medidos; la validez referida al criterio (concurrente y predictiva), o la coincidencia de lo evaluado con los CHD observados *fuera de la prueba*, ya sea con otros *test* similares o con un contexto de aplicación de estos CHD, y la validez del constructo: análisis del rasgo latente, relación con otras variables y medición de si existe o no relación real entre estos aspectos.

Una de las estrategias que tiene el Icfes para el desarrollo de sus pruebas, con miras a mejorar su validez, es justamente el DUE. Este modelo cubre desde la definición de la población que debe ser incluida en la evaluación hasta aspectos específicos del formato en el que se redactan y se presentan los ítems a los evaluados, lo que en últimas contribuye a la validez de la prueba (ver capítulo 2).

4. Estrategias de accesibilidad

4.1 Acomodación

Durante el diseño de cualquier prueba estandarizada se asume un conjunto de CHD generales que permiten el **acceso** a la presentación de la prueba y otro conjunto de CHD que son evaluados directamente por los constructos. La diversidad en el sistema educativo colombiano, sin embargo, requiere que estos CHD asociados al **acceso** sean analizados desde un enfoque inclusivo y diferenciador. Es desde esta perspectiva que surgen las **acomodaciones** y, eventualmente, las **modificaciones**. Dado que una de las características del DUE es permitir las acomodaciones, estas resultan necesarias para garantizar el acceso a la mayoría de la población de estudiantes con discapacidad.

Las **acomodaciones** se definen como las alteraciones sobre las condiciones generales de administración y aplicación de una prueba que no causan un cambio en los CHD que son evaluados y que fueron definidos en el marco teórico de la prueba (Tindal, Helwig y Hollenbeck, 1999; Hollenbeck, 2002). Estas alteraciones surgen de la apropiada identificación y caracterización de las necesidades específicas de los estudiantes durante el proceso de inscripción.

El propósito de las **acomodaciones** es permitir a los estudiantes que lo requieran acceder a la prueba para demostrar su desempeño en los CHD evaluados. Para ello, se eliminan, hasta donde sea posible, las interferencias

generadas cuando un estudiante interactúa con una prueba estandarizada que no fue diseñada de forma especial para él. Cabe indicar que estas **acomodaciones** no deben ser una ventaja injusta respecto al desempeño de la población general (Edgemon *et al.*, 2006). Por tanto, se espera que la acomodación tenga un efecto positivo en las poblaciones que las necesitan, pero que no tenga un efecto significativo para quienes no las requieren.

Las **acomodaciones** incluyen cambios en el formato (por ejemplo, la traducción a braille); en la forma en la que se responde la prueba (por ejemplo, permitirle al estudiante responder en el cuadernillo); o en las condiciones de aplicación (por ejemplo, dar más tiempo durante la aplicación). Las pruebas estandarizadas permiten distintos tipos de acomodaciones, que se dividen en cuatro categorías principales:

▶ *Acomodaciones en la presentación:*

Son aquellas acomodaciones en las que se ajusta la forma de presentar la información de una prueba con el propósito de que los estudiantes tengan acceso a la información a través de los canales sensoriales (Luke y Schwartz, 2007). Este tipo de acomodaciones pueden aplicarse a todo el test, a algunas partes en particular o solo a las instrucciones. Algunos ejemplos son el uso de **macrotipos** y la traducción a lengua de señas colombiana (LSC), entre otros.

▶ *Acomodaciones en la forma de responder:*

Este tipo de acomodaciones les permiten a los estudiantes emplear otras estrategias para responder a una prueba. Un ejemplo es la posibilidad de marcar las respuestas en el cuadernillo y no en la hoja de respuestas (Beech, 2010).

▶ *Acomodaciones en el entorno:*

Estas acomodaciones proveen condiciones específicas en el lugar donde se realizará la aplicación de la prueba, ajustando las condiciones o características del ambiente (Luke y Schwartz, 2007). Ejemplos de estas acomodaciones son las aplicaciones individuales o la ubicación en el primer piso de las zonas de aplicación para los estudiantes con dificultades de movilidad.

▶ *Acomodaciones en la agenda:*

Este tipo de acomodaciones involucra cambios asociados a la programación, asignación y administración del tiempo en la aplicación de la prueba (Beech, 2010). Como ejemplos están los descansos y permitirles a los estudiantes presentar la prueba en un orden distinto al inicialmente definido.

La selección de uno o varios tipos de **acomodaciones** debe realizarse de forma previa a la aplicación de la prueba, por lo que el Icfes identifica en el proceso de inscripción los requerimientos de los estudiantes con discapacidad. En el momento del registro de los estudiantes, es importante contar con información precisa y verídica sobre sus necesidades, que en la mayoría de los casos es provista por las instituciones

educativas, el entorno familiar o cada estudiante. De manera general, las **acomodaciones** y su puesta en marcha se ven favorecidas por el uso de medios electrónicos en la presentación de una prueba. Esto se aborda en detalle en el siguiente apartado.

4.4.1 Pruebas por computador

El uso creciente de computadores y otros aparatos electrónicos, tanto en las escuelas como en los hogares del país, ha impulsado al Icfes a avanzar hacia la implementación de las pruebas en formatos diferentes al papel y lápiz. En esa medida, se han iniciado pilotos en la aplicación de algunas pruebas por computador y **pruebas adaptativas**. Estas últimas presentan ítems de acuerdo con la habilidad que evidencia el evaluado en cada respuesta. Esto permite que las pruebas sean dinámicas y que requieran menos ítems para determinar la habilidad del evaluado.

Respecto a la accesibilidad, existe evidencia que demuestra que las pruebas por computador generan mayor acceso para personas con discapacidad. Thompson, Blount y Thurlow (2002a) revisaron 46 estudios relacionados con **acomodaciones** de pruebas estandarizadas para personas con discapacidad. Los autores identificaron que una de las acomodaciones que mostró un efecto positivo en los puntajes de manera recurrente fue la aplicación de las pruebas por computador.

Muchas de las **acomodaciones** que se utilizan en una prueba de papel y lápiz requieren que se desarrollen múltiples versiones de los cuadernillos o que intervengan personas en diferentes roles de apoyo para los evaluados (lector, intérprete, escriba). Esto no solo genera problemas logísticos, sino que, en algunos casos, puede comprometer la **validez** de los resultados de la prueba. Ello porque la actitud, la idoneidad y la capacitación de las personas de apoyo pueden afectar de manera positiva o negativa el desempeño del estudiante.

Los problemas logísticos y de estandarización se solucionan en gran medida con la aplicación de pruebas por computador. En primer lugar, algunas de las **acomodaciones** se pueden incorporar como parte de la prueba (tamaño de la fuente, contraste, respuesta, audios, lengua de señas, glosarios) manteniendo así un único cuadernillo que permita a los estudiantes “activar” las acomodaciones que requieran para superar las barreras generadas por su condición particular. En ciertos casos, primero, estas acomodaciones pueden reducir la carga en la memoria de trabajo. Segundo, algunos apoyos que requieren la intervención humana se pueden estandarizar al ofrecer el mismo tipo de apoyo a los estudiantes por medio de programas de computador especializados (JAWS, Braille, lectores de texto, etc.) o en formatos de audio y video pregrabados (lectura de textos, traducción a lengua de señas, etc.).

En lo que se refiere a las **pruebas adaptativas** por computador, estas tienen la ventaja adicional de reducir el número de preguntas que presenta un estudiante por prueba, lo que facilita realizar **acomodaciones** de agenda. Este tipo

de formato también reduce la sensación de frustración de los estudiantes —como su nombre lo dice, se adapta al nivel de habilidad de ellos— y, en consecuencia, la desmotivación que pueden sentir al enfrentarse a los ítems, hecho que también afecta la **validez** de la prueba.

Ahora bien, las pruebas por computador y las **pruebas adaptativas** no permiten franquear las barreras de acceso a las que se enfrentan ciertos estudiantes. De hecho, las **acomodaciones** muchas veces no son suficientes, dados los requerimientos educativos del estudiante con discapacidad. Para esos casos, varios autores sugieren el uso de **modificaciones**.

4.2 Modificaciones

Así como las acomodaciones, las **modificaciones** son ajustes a una prueba. A diferencia de las primeras, estas trascienden las estrategias enmarcadas en el DUE. Contrario a las acomodaciones, una modificación es un cambio que altera intencionalmente el constructo que se está midiendo en una prueba (Thurlow, 2015, p. 333-4; Stone, E., Cook L. 2018, p. 64).

Si bien la introducción de **modificaciones** a una prueba puede resultar de la pretensión, a todas luces legítima, de adelantar procesos adecuados de evaluación para la totalidad de la población estudiantil, existe una serie de consideraciones, tanto conceptuales como metodológicas, que es necesario tener en cuenta.

Por un lado, la clasificación de un ajuste como una modificación es relativa al modo en que se fijan las condiciones de identidad del **dominio**: el debate sobre si un ajuste es una acomodación o una modificación es perfectamente razonable allí donde los rasgos esenciales de un dominio no han sido delimitados. Por esto mismo, resulta imperante en el marco del proyecto definir minuciosamente el dominio.

Por lo anterior, un ajuste siempre será una modificación con relación a una prueba específica: un mismo ajuste puede ser una modificación en relación con una prueba y al mismo tiempo una acomodación con relación a otra. De esto se sigue que el manejo de la relación entre **acomodaciones** y **modificaciones** supone haber aclarado, para cada prueba, un conjunto de rasgos estructurales cuya modificación alteraría su identidad y un procedimiento que permita detectar cuándo un ajuste particular genera un cambio en uno de estos rasgos.

Por otro lado, si preservar la identidad de una prueba depende de conservar el **dominio** que esta evalúa, la introducción de una modificación tiene como resultado la creación de una **evaluación alternativa**. Esto implica que, si un proceso de inclusión tiene como objetivo generar las condiciones para que una población lo más amplia posible acceda a un mismo instrumento de evaluación, en dicho proceso debe garantizarse que no se use ninguna modificación. El uso de **modificaciones** tendrá sentido únicamente en procesos de **inclusión** que partan del reconocimiento de la necesidad de contar con más de una prueba para poder brindar información sobre los CHD de la totalidad de la población. Dicho esto, queda abierta la

pregunta sobre cuál es la relación que guarda la prueba que resulta de la introducción de una modificación con la prueba original. Si bien esto es importante al momento de establecer qué es lo que la prueba alternativa mide, solo es posible determinarlo al considerar las modificaciones concretas que hayan sido implementadas.

4.3 Estrategias específicas

Las tablas que se presentan a continuación muestran algunas estrategias y recomendaciones que han sido aplicadas por diferentes consorcios internacionales (Smarter Balanced, NCEO, CCSSO, entre otros), y que han demostrado tener un buen funcionamiento al brindar **acceso** a las pruebas que ellos realizan. Muchas de estas recomendaciones se enfocan en reducir la carga cognitiva de los estudiantes al momento de interactuar con la prueba. Es importante indicar que, para cada una de estas estrategias o recomendaciones, se debe verificar que no se invalide o se altere el **dominio** que se pretende medir en un principio.

Tabla 2. *Acomodaciones relacionadas con el braille y el lector de pantalla*

Diseño Universal	
Braille y lector de pantalla	
Evitar el uso de gráficas irrelevantes al dominio.	
Evitar el uso de texto vertical o diagonal.	
No poner leyendas o información a la izquierda o debajo de los ítems.	
Instrucciones y procedimientos claros intuitivos y simples	
Ítems de muestra previos a la aplicación.	
Ejemplos de manipulación del material.	
Capacitación del personal de apoyo.	
Criterios de calificación conocidos.	
Proporcionar guías de orientación en las que se explique el proceso de aplicación.	
Máxima lecturabilidad y comprensibilidad	
Usar palabras de usos frecuentes y simples. A su vez, eliminar las palabras innecesarias.	
Deben definirse los términos técnicos utilizados.	
Las oraciones complejas deben dividirse en varias oraciones simples en las que las ideas principales se encuentren al principio.	

Continúa en la siguiente página

Introducir ideas, hechos y procesos uno a la vez.

La relación entre nombres y pronombres debe ser clara
(No se deben omitir referentes).

Las ubicaciones temporales y espaciales deben aparecer
siempre al principio de las oraciones.

Las instrucciones deben indicar el orden exacto de
ejecución de un procedimiento.

Los procesos deben ser etiquetados y ubicados cerca
del texto.

Marcar todas las preguntas.

Resaltar con negrilla o subrayado las palabras esenciales
para responder un ítem.

Siempre que sea posible, proveer a los estudiantes las
definiciones de palabras poco comunes.

Redactar el ítem evitando utilizar negaciones.

Redactar el ítem en voz activa.

Todos los elementos que constituyen un ítem deben ser
visibles en una página.

Continúa en la siguiente página

Máxima legibilidad

Etiquetar en las gráficas los ejes y las líneas.

Evitar las escalas de grises.

Usar tamaño de letra de mínimo 14 puntos.

El espacio entre letras, así como el espacio entre palabras, debe ser siempre el mismo.

El espacio entre líneas debe ser entre el 25 % y el 30 % del tamaño de la fuente.

Evitar fuentes de letras que sean cursivas o decorativas.

Evitar textos justificados, dado que son más difíciles de leer.

Las líneas de los textos deben tener entre 8 y 12 palabras.

Demasiados ítems por página son difíciles de leer.

La hoja de respuesta debe utilizar burbujas horizontales grandes.

La hoja de respuesta es difícil para los niños de menor edad. Para esta población se sugiere que las burbujas de respuestas se encuentren en el cuadernillo.

Evitar el texto cuando una imagen sea suficiente. Las ilustraciones deben reemplazar el texto y no duplicar la información.

Continúa en la siguiente página

Usar ilustraciones cuando el razonamiento espacial sea necesario.
Usar etiquetas para facilitar la comprensión de ilustraciones complejas.
Un texto no debe adicionarse a ilustraciones que se expliquen por sí mismas.
Un texto explicativo debe ubicarse lo más próximo posible a su respectiva ilustración.

Tabla 3. Acomodaciones en la presentación

Acomodaciones	
Acomodaciones en la presentación	
Acomodación	Explicación
Visuales	
Impresión grande. Al menos 18 puntos para los que la requieran	El tamaño de letra es una barrera física para población con problemas graves de visión. Referencia Beddow <i>et al</i> , 2011 p. 163
Elementos de magnificación. Manuales o por computador	El tamaño de letra es una barrera física para población con problemas graves de visión. Referencia Beddow <i>et al</i> , 2011 p. 163

Continúa en la siguiente página

<p>Lengua de señas. El material gráfico puede ser descrito, pero debe estar también disponible para su visualización</p>	<p>Las preguntas deben traducirse a lengua de señas usando un video durante la aplicación porque esto facilita la estandarización de la traducción.</p> <p>Referencia Thompson <i>et al</i>, 2005 p. 27</p>
<p>Acomodaciones en la presentación</p>	
<p>Acomodación</p>	<p>Explicación</p>
<p>Táctiles</p>	
<p>Braille</p>	<p>El braille da acceso a la población ciega o con severos problemas de visión a la prueba. No está extendido a toda la población.</p> <p>Referencia Thompson <i>et al</i>, 2005 p. 27</p>
<p>Gráficas, imágenes, ilustraciones en alto relieve</p>	<p>Esta opción permite a la población ciega o con severos problemas de visión acceder a las gráficas sin necesidad de que un intérprete las describa. Esta opción evita que la descripción auditiva revele la respuesta a preguntas que indagan por elementos gráficos que una vez leídos sean obvios, p. ej. gráficas ascendentes o descendentes.</p> <p>Referencia Thompson <i>et al</i>, 2005 p. 27</p>

Continúa en la siguiente página

Acomodaciones en la presentación	
Acomodación	Explicación
Auditivas	
Lector humano	<p>El estudiante que no puede decodificar el texto escrito y no dispone del equipo electrónico necesario requiere de un lector entrenado. Esto es así ya que, al leer, el lector no entrenado puede revelar inconscientemente las respuestas.</p> <p>Referencia Thompson <i>et al</i>, 2005 p. 28</p>
Audio grabaciones	<p>La grabación de los ítems de la prueba garantiza la estandarización de la lectura, pues todos los evaluados reciben la misma lectura con los mismos matices. Si no hay computador puede usarse un CD que facilite retroceder la grabación para volver a escucharla.</p> <p>Referencia Thompson <i>et al</i>, 2005 p. 28</p>

Continúa en la siguiente página

Elementos de magnificación auditiva	Los evaluados que tengan algún tipo de discapacidad auditiva requieren dispositivos de magnificación auditiva para escuchar las instrucciones que imparte el jefe de salón.
	Referencia Thompson <i>et al</i> , 2005 p. 28
Lectores de pantalla	Los dispositivos electrónicos funcionan solo si la prueba se presenta en computador. Son necesarios cuando los evaluados no dominan el braille.
	Referencia Thompson <i>et al</i> , 2005 p. 28

Continúa en la siguiente página

Acomodaciones en la forma de responder	
Acomodación	Explicación
<p>Escriba. Persona que anota las respuestas</p>	<p>Esta acomodación es útil para estudiantes cuya discapacidad se relaciona con dificultades para escribir un texto, (limitación física-movilidad, discapacidad sensorial, discapacidad intelectual). Consiste en asignar una persona que escriba o grabe las indicaciones del estudiante dadas al hablar, al usar la LSC o a través de un instrumento de comunicación aumentativa y alternativa. Este apoyo no debe alterar los comentarios de los estudiantes, pero debe permitirles editarlos, de ser necesario.</p> <hr/> <p>Referencia Beech, 2010 Thompson <i>et al</i>, 2005</p>

Continúa en la siguiente página

<p>Procesador de palabras, convertidor de voz a texto, grabadoras de audio o video</p>	<p>Este apoyo puede ser usado por estudiantes con dificultades para escribir un texto (limitación física-movilidad, discapacidad sensorial, discapacidad intelectual). Son aparatos, programas o aplicaciones que convierten a texto el dictado verbal o señado (LSC) de un estudiante.</p> <hr/> <p>Referencia Beech, 2010 Thompson <i>et al</i>, 2005</p>
<p>Teclados personalizados, señaladores, pantalla táctil, usar teclas especiales, <i>scroll</i> del ratón en computador o tabletas.</p>	<p>Este apoyo puede ser usado por estudiantes con dificultades para escribir un texto (limitación física-movilidad, discapacidad sensorial, discapacidad intelectual). Son aparatos, programas o aplicaciones que simplifican el uso de comandos en un computador para una prueba electrónica.</p> <hr/> <p>Referencia Beech, 2010 Thompson <i>et al</i>, 2005</p>

Continúa en la siguiente página

<p>Responder sobre el cuadernillo</p>	<p>Este apoyo puede ser usado por estudiantes con dificultades para recordar o procesar información en una hoja de respuestas para completar una prueba (discapacidad intelectual). Consiste en permitirles a los estudiantes marcar la respuesta correcta dentro del cuadernillo.</p> <hr/> <p>Referencia Beech, 2010 Thompson <i>et al</i>, 2005</p>
<p>Monitor de respuesta. Persona que vigila que los estudiantes respondan sobre la hoja de respuesta correctamente</p>	<p>Este apoyo puede ser usado por estudiantes con dificultades para recordar o procesar información en una hoja de respuestas para completar una prueba (discapacidad intelectual). Es una persona que supervisa que los estudiantes rellenen y ubiquen correctamente la burbuja de la pregunta que están respondiendo.</p> <hr/> <p>Referencia Beech, 2010 Thompson <i>et al</i>, 2005</p>

Continúa en la siguiente página

<p>Organizadores visuales. Resaltadores</p>	<p>Este apoyo puede ser usado por estudiantes con dificultades para recordar o procesar información (discapacidad intelectual). Los organizadores visuales incluyen papel cuadriculado, marcadores, banderillas, hojas de borrador y plantillas. Los organizadores gráficos incluyen representaciones que elaboran los estudiantes para reorganizar la información de acuerdo con sus necesidades para responder la prueba. Ambos pueden ser empleados en aplicaciones en papel o electrónicas.</p> <hr/> <p>Referencia Beech, 2010 Thompson <i>et al</i>, 2005</p>
<p>Acomodaciones en el entorno</p>	
<p>Acomodación</p>	<p>Explicación</p>
<p>Reducir la distracción de los estudiantes. Sentarlos adelante, lejos de ventanas y puertas, o que presenten la prueba solos o con pocos estudiantes</p>	<p>Los estudiantes deben ser ubicados en el sitio donde se sienta más cómodos, con sillas y mesas adecuadas y con espacio suficiente para trabajar.</p> <hr/> <p>Referencia Elliott, S., <i>et al</i>, 2011</p>

Continúa en la siguiente página

<p>Reducir la distracción hacia otros estudiantes. Algunos estudiantes pueden requerir lectura en voz alta</p>	<p>Si un estudiante recibe una acomodación que pueda distraer a otros estudiantes (escriba, pausas múltiples, etc.), debe ser ubicado en un sitio separado del resto.</p>
<p>Cambiar el sitio de aplicación para mejorar el acceso físico</p>	<p>Esta acomodación permite el acceso físico a todos los estudiantes.</p>
<p>Acomodaciones en la agenda</p>	
<p>Acomodación</p>	<p>Explicación</p>
<p>Tiempo extendido. Usualmente 1,5 del tiempo regular de la prueba</p>	<p>La extensión del tiempo debe administrarse dependiendo el tipo de prueba y la discapacidad del estudiante. Extender el tiempo permite que los estudiantes con dificultades para procesar información tengan suficiente tiempo para responder una tarea. Además, la extensión de tiempo es un complemento necesario para otras acomodaciones como el uso de un lector de prueba o escriba.</p>

Continúa en la siguiente página

<p>Múltiples o frecuentes pausas</p>	<p>Conceder múltiples o frecuentes pausas durante la presentación de la prueba permite que los estudiantes con dificultades para manejar la ansiedad o la frustración puedan responder cómodamente la prueba. Además, permite que los estudiantes con déficit de atención no tengan mayores cargas cognitivas al presentar la prueba.</p> <hr/> <p>Referencias Tindal <i>et al.</i>, 1998, Menlove <i>et al.</i>, 1998 y Thompson <i>et al.</i>, 2005</p>
<p>Cambiar el orden de las pruebas</p>	<p>Dependiendo del nivel de procesamiento que los estudiantes con discapacidad manifiesten, cambiar el orden de las pruebas administradas, según el caso del estudiante, optimiza el rendimiento de ellos sin afectar el dominio de la prueba.</p> <hr/> <p>Referencias Tindal <i>et al.</i>, 1998, Menlove <i>et al.</i>, 1998 y Thompson <i>et al.</i>, 2005</p>

Continúa en la siguiente página

<p>Aplicar en un momento del día determinado</p>	<p>En el caso de estudiantes con problemas para mantener el foco de atención, aplicar la prueba en el momento del día cuando ellos manifiesten un rendimiento óptimo en tareas que requieren atención reduce la carga cognitiva que innecesariamente se impone a los estudiantes.</p> <hr/> <p>Referencias Tindal <i>et al.</i>, 1998, Menlove <i>et al.</i>, 1998 y Thompson <i>et al.</i>, 2005</p>
<p>Aumentar el número de sesiones</p>	<p>Para reducir el efecto de la fatiga que genera responder una prueba con numerosos ítems, especialmente en estudiantes cuya discapacidad involucra dificultades en el procesamiento de información, se recomienda aumentar el número de sesiones de forma que los estudiantes respondan una porción de la prueba cada día.</p> <hr/> <p>Referencias Tindal <i>et al.</i>, 1998, Menlove <i>et al.</i>, 1998 y Thompson <i>et al.</i>, 2005</p>

Las tablas aquí expuestas no son exhaustivas y, dependiendo de los resultados de las validaciones, algunos elementos harán parte de estas o cambiarán de categoría. Las metodologías de validación se presentan en el capítulo 5.

5. Metodologías de validación cualitativa y cuantitativa

El Icfes implementa distintas metodologías de validación como parte del proceso de diseño de una prueba. Dadas las diversas características de las poblaciones a las que se busca dar acceso, se realizan revisiones a la luz de referentes y ejercicios prácticos de tales metodologías. Algunas de las validaciones consideradas desde el aspecto cualitativo son:

▶ *Validación del contenido de los ítems:*

Esta actividad se realiza con docentes expertos en las áreas evaluadas. En este caso, se incluyen docentes de apoyo que realizan el trabajo encaminado a atender las necesidades educativas de los estudiantes con discapacidad y etnoeducadores.

▶ *Validación del proceso de respuesta:*

Esta actividad se realiza con estudiantes. En este caso, se incluyen estudiantes con discapacidad (sensorial, cognitiva-intelectual, TEA, entre otros) y estudiantes pertenecientes a grupos étnicos. Se basa en dos metodologías: las entrevistas cognitivas —que pueden ser grupales o individuales— y ejercicios de pensar en voz alta.

Por su parte, las validaciones cuantitativas se refieren al comportamiento psicométrico de la prueba (ver Tindal (1998), Fuchs y Fuchs (2001) y Shavelson (2018)). Estas validaciones pretenden dar cuenta de la eficiencia de las estrategias

desarrolladas, en especial las **acomodaciones** y los patrones de diseño, para brindar **acceso** a la población con discapacidad. Estas validaciones se basan en el principio conocido como *differential boost*, que establece que una estrategia de **acceso** es eficiente si genera mejoras en los puntajes de una prueba para los estudiantes con discapacidad y no para los estudiantes sin discapacidad. Aunque algunas de estas estrategias podrían no necesitar este tipo de validación, dada la lógica detrás de, por ejemplo, la impresión de un cuadernillo en una fuente de tamaño grande, es importante recopilar estos datos, ya que permiten diferenciar qué tipo de apoyo funciona para algunas poblaciones y pueden dar indicadores de cambios en las habilidades que la prueba pretende medir.

GLOSARIO

- ▶ **Acceso:** oportunidad para que un estudiante demuestre sus habilidades en una prueba.
- ▶ **Acomodaciones:** ajustes que permiten el **acceso** a una prueba sin modificar el dominio evaluado.
- ▶ **Confiabledad:** atributo de una prueba que radica en la consistencia de los resultados obtenidos; por tanto, diferentes aplicaciones de un mismo dominio a poblaciones similares deben dar resultados parecidos.
- ▶ **Dominio:** aquello que se planea evaluar, lo que se va a medir y sobre lo qué se va a emitir un juicio. Por dominio se pueden entender los procesos cognitivos, actitudinales, de ejecución de tareas que se van a tener en cuenta en la prueba.
- ▶ **Diseño universal de la evaluación:** técnica de construcción de instrumentos de evaluación que permite a toda la población estudiantil evidenciar sus competencias con respecto a los CHD que busca desarrollar el proceso educativo que han tenido. El DUE busca que todos los estudiantes tengan la oportunidad de mostrar sus competencias con respecto a un mismo conjunto de CHD definido.

- ▶ *Evaluación alternativa*: prueba que busca evaluar a la población para quienes interacción con los ítems constituye una barrera para demostrar su conocimiento y sus habilidades en evaluaciones generales, aun con acomodaciones. Exige **modificaciones** al dominio.
- ▶ *Evaluación general*: examen de Estado, prueba o módulo que busca evaluar a la población que acaba de culminar un periodo de formación. Evalúa un rasgo latente definido en el dominio.
- ▶ *Inclusión*: conjunto de políticas que buscan permitirles a todos los estudiantes, con y sin discapacidad, participar dentro de cualquier proceso que pertenezca al sistema educativo.
- ▶ *Modificaciones*: cambios al instrumento de evaluación que implican modificar el dominio de evaluación; es decir, dichos cambios incorporan o eliminan CHD al dominio original.
- ▶ *Macrotipo*: tamaño de letra más grande que el usual.
- ▶ *Pruebas adaptativas*: son pruebas aplicadas y calificadas por computador, que cuentan con un banco de ítems calibrado según principios de TRI. Estas pruebas presentan ítems de acuerdo con la habilidad estimada del evaluado a partir de su desempeño en la respuesta de ítems anteriores. Estas pruebas parten del supuesto de ser hechas a la medida del estudiante. Lo anterior permite que la cantidad

de ítems requeridos para estimar el nivel de habilidad del evaluado sea menor que el requerido en las pruebas de lápiz y papel, disminuyendo así la fatiga de los evaluados.

- ▶ **Sesgo:** presuposición de la tenencia, por parte de la población evaluada, de CHD necesarios para responder la prueba, pero que no son centrales al dominio. La tenencia o carencia de estos CHD facilita o entorpece las respuestas de ciertas subpoblaciones.
- ▶ **Validez:** atributo de una prueba que consiste en medir exactamente los CHD que se quieren evaluar; por ejemplo, aquellos que son centrales al dominio.

REFERENCIAS

Administración Nacional de Seguridad Social (2011). “A todos cobijamos” de Mirta Gomez. Flickr. Recuperado de: <https://www.flickr.com/photos/ansesgob/7329817272/in/album-72157630000063096/>

Barton y Finch (2004). Testing assumptions of unidimensionality across groups of students with disabilities, with accommodations and English language learners. In *Poster presented at the National Council for Measurement in Education meetings, San Diego, CA.*

Beddow, Kurz y Frey (2011). Accessibility Theory: Guiding the Science and Practice of Test Item Design with Test-Taker in Mind. En: Elliott, S., Kettler, R., Beddow, P., Kurz, A. (Eds.). (2011). *Handbook of Accessible Achievement Tests for All Students.* Londres: Springer.

Beech (2010). *Accommodations: Assisting students with disabilities.* Florida Department of Education.

Edgemon, Jablonski y Lloyd (2006). Large-scale assessments: A teacher’s guide to making decisions about accommodations. *Teaching Exceptional Children*, 38(3), 6-11.

Bello (2009). Derechos indígenas y ciudadanías diferenciadas en América Latina y el Caribe Implicancias para la educación. *Interculturalidad, educación y ciudadanía perspectivas latinoamericanas.* Editor: López, L.E.. Funproeib Andes & Plural editores: La Paz, Bolivia.

Castillo y Carou (2002). Reinención de la identidad étnica, nuevas territorialidades y redes globales: el Estado multiétnico y pluricultural en Colombia y Ecuador. *Revista Sociedad y Economía*, (3), p. 55-76.

Elliott, Kettler, Beddow y Kurz, (2011). *Handbook of Accessible Achievement Tests for All Students*. Londres: Springer.

Fuchs y Fuchs (2001). Helping teachers formulate sound test accommodation decisions for students with learning disabilities. *Learning Disabilities Research & Practice*, 16(3), pp. 174–181.

Grise, Beattie y Algozzine (1982). Assessment of minimum competency in fifth grade learning disabled students: Test modifications make a difference. *Journal of Educational Research*, 76, 35-40.

Groth-Marnat (2003) *Handbook Of Psychological Assessment*. Minesota: John Wiley.

Hansen y Mislavy (2008). *Design patterns for improving accessibility for test takers with disabilities* (Research Report No. RR-08-49). Princeton, NJ: Educational Testing Service.

Hollenbeck (2002). Determining when test alterations are valid accommodations or modifications for large-scale assessment. *Large-scale assessment programs for all students: Validity, technical adequacy, and implementation*, 395-425.

Icfes (2019). *El diseño centrado en evidencias: teoría y práctica de uso*. Bogotá: Dirección de evaluación, Icfes.

Luke y Schwartz (2007). Assessment and accommodations. *Evidence for Education*, 2(1), pp. 1-11.

Menlove, M., & Hammond, M. (1998). Meeting the demands of ADA, IDEA, and other disability legislation in the design, development, and delivery of instruction. *Journal of Technology and Teacher Education*. 6 (1), 75-85.

Ministerio de Educación Nacional. (2017). Plan Nacional Decenal de Educación, 2016-2026. Recuperado de <http://www.plandecenal.edu.co>.

Ministerio de Educación Nacional. (2013). Lineamientos de Política de Educación Superior Inclusiva. Recuperado de: <https://www.mineducación.gov.co/1759/-article-357277.html?noredirect=1>. pdf

Ministerio de Educación Nacional. (2005). Índice de Inclusión para la Educación Superior (INES). Recuperado de: <https://www.mineducación.gov.co/1759/w3-article-357277.html.pdf>

Ministerio de Educación Nacional. (2017). Orientaciones para la transición educativa de los estudiantes con discapacidad y con capacidades o talentos excepcionales en la educación inicial, básica y media. Recuperado de https://www.mineducacion.gov.co/1759/articles-360294_foto_portada.pdf

Mislevy, R. *et al* (2003). A brief introduction to evidence-centered design. Educational Testing Service, Princeton, NJ. July 2003.

Mislevy y Riconscente (2005). "Evidence-centered design: Layers, structures, and terminology". Menlo Park, CA: SRI International.

Mislevy, Haertel, Riconscente, Rutstein y Ziker (2017). Evidence-centered assessment design. In *Assessing Model Based Reasoning using Evidence Centered Design. A Suite of Research Based Design Patterns* (pp. 19-24). Routledge.

Quilaqueo y Torres (2013). Multiculturalidad e interculturalidad: desafíos epistemológicos de la escolarización desarrollada en contextos indígenas. *Alpha (Osorno)*, (37), 285-300.

Shavelson (2018). *Impact of Accommodations and Access on Students' Test Performance: A Pilot Study*. Manuscrito no publicado.

Sierra-Matamoros, Valdelamar-Jiménez, Hernández-Tamayo y Sarmiento-García, L. M. (2007). Test adaptativos informatizados. *Avances en en Medición*, 5, 157-162.

Sireci, Stephen, Scarpati Stanley y Li, Shuhong (2005). Test Accommodations for Students with Disabilities: An Analysis of the Interaction Hypothesis. *Review of Education Research*.

Stone E., Cook L. (2018). Fair Testing and the Role of Accessibility. *Handbook of accessible instruction and testing practices. Issues, Innovation and Application*. Second edition. Springer

Thompson y Thurlow (2002). Universally designed assessments: Better tests for everyone! (Policy Directions N.º 14). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Thompson, Blount y Thurlow (2002a). A summary of research on the effects of test accommodations: 1999 through 2001 (Technical Report 34). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

_____. (2002b). Universal Design applied to large scale assessments (Synthesis Report 44). Minneapolis, MN: University of Minnesota. NCEO.

Thompson, Morse, Sharpe y Hall (2005). Accommodations manual: How to select, administer, and evaluate the use of accommodations for instruction and assessment of students with disabilities (2nd ed.). Washington, DC: Council of Chief State School Officers.

Tindal (1998). Models for understanding task comparability in accommodated testing. Eugene, OR: Behavioral Research and Teaching.

Tindal, Helwig y Hollenbeck (1999). An update on test accommodations. *Journal of Special Education Leadership*, 12(2), 11-20.

Tubino (2004). La interculturalidad crítica como proyecto ético-político. *Encuentro continental de educadores agustinos*, 24-28.

Thurlow, M. L. (2015). How should we evaluate whether special education works? In B. Bateman, J. W Lloyd & M. Tankersley (Eds.). *Enduring issues in special education: Personal perspectives* (pp. 323 –339). New York, NY: Routledge.

Thurlow, Martha; Lazarus, Sheryl; Christensen, Laurene & Shyyan, Vitaliy (2016) Principios y Características de Sistemas Inclusivos de Evaluación en un Escenario de Evaluación Cambiante. NCEO Report 400). Minneapolis: University of Minnesota, National Center on Educational Outcomes.

Walsh (2009). Interculturalidad crítica y educación intercultural. Ponencia presentada en el Seminario Interculturalidad y Educación Intercultural organizado por el Instituto Internacional de Integración del Convenio Andrés Bello, La Paz, 9-11 de marzo; Walsh et Albán Afolfo. *Interculturalidad y Educación Superior: significados, prácticas y políticas críticas*.



La educación
es de todos

Mineducación

Calle 26 N.º 69-76, Torre 2, Piso 15, Edificio Elemento, Bogotá, D. C., Colombia • www.icfes.gov.co
Líneas de atención al usuario: Bogotá Tel.: (57+1) 484-1460 PBX: (57+1) 484-1410 -
Gratuita nacional: 018000-519535