



LOS LIBERTADORES
FUNDACIÓN UNIVERSITARIA

Análisis de los desempeños de estudiantes de psicología frente a las competencias de razonamiento cuantitativo y competencias ciudadanas de la prueba SABER PRO durante el año 2020 a nivel nacional

Luis Alberto Arias Cohecha, laariasc@libertadores.edu.co

RESUMEN

El objetivo de la presente investigación consiste en analizar los desempeños de estudiantes de psicología frente a la prueba SABER PRO a nivel nacional en relación con variables sociodemográficas, específicamente en los puntajes de las competencias de razonamiento cuantitativo y competencias ciudadanas. Para tal fin se planteó una investigación cuantitativa de tipo no experimental transversal, con implementación de análisis de correspondencias múltiples, componentes principales e implementación de regresión beta para explorar el comportamiento de las variables extraídas de la base de datos de la aplicación de la prueba SABER PRO correspondiente al año 2020. Se observó que el género tiende a mostrar resultados diferenciadores en cuanto a las dos competencias solamente para el género masculino cuando se plantea el modelo de predicción de las puntuaciones de competencias ciudadanas y razonamiento cuantitativo.

Palabras clave: SABER PRO, razonamiento cuantitativo, competencias ciudadanas, psicología.

ABSTRACT

The objective of this research is to analyze the SABER PRO test results of psychology students considering sociodemographic variables, specifically in quantitative reasoning skills and citizenship skills scores. For this purpose, a non-experimental cross-sectional quantitative investigation was proposed, with implementation of multiple correspondence analysis, principal components analysis and beta regression analysis to explore variables interactions extracted from the database of 2020 SABER PRO national test results. It was observed that gender shows differentiating results regarding quantitative reasoning skills and citizenship skills only for male gender when the beta prediction model was proposed.

Key words: SABER PRO, quantitative reasoning, citizenship skills, psychology.

Análisis de los desempeños de estudiantes de psicología frente a las competencias de razonamiento cuantitativo y competencias ciudadanas de la prueba SABER PRO durante el año 2020 a nivel nacional

INTRODUCCIÓN

El presente estudio tiene como finalidad explorar la posible interacción entre variables sociodemográficas y su posible influencia en las puntuaciones de las competencias de razonamiento cuantitativo y competencias ciudadanas de estudiantes de psicología que presentaron la prueba SABER PRO durante el año 2020 a nivel nacional. Para lograr tal fin se ha planteado una investigación de tipo cualitativa descriptiva que permita explorar la información recolectada de una base de datos extraída del banco de datos del ICFES en su aplicación online dataicfes, de modo que se hizo una adecuación de la base de datos para realizar los análisis estadísticos descriptivos, correlacionales y predictivos, tratando de dar una explicación a la información recolectada, con el fin de caracterizar las posibles interacciones entre las variables sociodemográficas y su posible repercusión en los desempeños de estudiantes de psicología.

Dicha suposición se toma como punto de partida teniendo en cuenta lo señalado en el documento de guía de orientación para la prueba SABER PRO (ICFES, 2020, p.13) y lo sugerido por las investigaciones de Rodríguez-Jiménez, Cote-Peña, Malaver-Tolosa, Pineda-Miranda & Sarmiento-Fúquene (2019); Suárez-Rodríguez (2018) y Valbuena & García (2016), en las cuales se señala la importancia de explorar interacciones relevantes entre competencias genéricas de la prueba SABER PRO versus variables sociodemográficas que puedan sugerirse como predictoras del rendimiento de estudiantes de psicología.

Este estudio canaliza el interés de instancias tan importantes como ASCOFAPSI y COLPSIC, entidades encargadas de la toma de decisiones relacionadas con el avance de la profesión de la psicología en Colombia, de modo que el indicador de las pruebas SABER PRO se utiliza como indicador del mejoramiento de la calidad de la educación según lo señala el Ministerio de Educación Nacional en su política de calidad educativa contemplada en el Decreto 1330 de 2019.

La aplicación del examen SABER PRO durante el año 2020 coincidió con la ocurrencia de la primera fase de la pandemia por COVID-19, hecho que afectó la forma tradicional en la cual el instrumento fuera aplicado, ya que implicó que el ICFES dispusiera una plataforma online para la aplicación masiva de exámenes a nivel nacional.

Entre otras modificaciones, se amplió la jornada de presentación de exámenes a 5 días, se redujo la aplicación a una sola sesión, con aplicaciones remotas en casa o en sitio, hecho que implicó que las aplicaciones en sitio fueran vigiladas no solo con los protocolos de evaluación institucionales propios del ICFES sino con las exigencias adicionales propuestas por los Ministerios de Salud y Educación Nacional en cuanto a la incorporación de protocolos de bioseguridad en aplicaciones en sitio. El tiempo de evaluación se redujo a 4 horas y 40 minutos, y en las aplicaciones en sitio se propusieron una mayor cantidad de horarios para evitar las aglomeraciones.

Por otro lado, las aplicaciones remotas en casa dispusieron de una ampliación en despliegue de recursos de soporte tecnológico del ICFES para facilitar el monitoreo, seguimiento y apoyo en tiempo real para evaluados que presentaran inconvenientes con la prueba.

Concretamente se incluyeron medios tales como el chat de soporte, la línea telefónica nacional, el correo electrónico institucional, y el medio de soporte a partir de la plataforma Zendesk y Sumdadi, para brindar apoyo a aproximadamente 211.000 evaluados ubicados en 1090 municipios a nivel nacional. Es necesario indicar que otras de las modificaciones de la aplicación del examen incluyeron firma de un acuerdo de ética por parte de los evaluados y activación de cámara de video para seguimiento en tiempo real, con el fin de aumentar las garantías de transparencia a la hora de la aplicación de los exámenes de estado. Considerando lo anteriormente mencionado, se plantea un contexto en el cual ocurrió la aplicación de la prueba SABER PRO para el año 2020. Dichas contingencias pudieron haber afectado las mediciones que anualmente hace el ICFES respecto de los indicadores de competencia, sin embargo, esta es solo una especulación y no corresponde con la finalidad del estudio aquí propuesto.

DESCRIPCIÓN DEL PROBLEMA DE INVESTIGACIÓN

El problema de investigación consiste en indagar si los factores socioeconómicos asociados a la presentación de la prueba SABER PRO (sexo, estrato, región, procedencia, nivel de competencia, género, modalidad, estado civil), están relacionados e influyen los puntajes de las competencias de razonamiento cuantitativo y competencias ciudadanas evidenciadas en los resultados de la prueba SABER PRO de estudiantes de psicología a nivel nacional, a lo largo del año 2020.

Esta situación configura un problema de investigación que resulta ser relevante para las instituciones de formación en psicología a nivel nacional, dado que se hace necesario verificar si las variables sociodemográficas ejercen algún tipo de influencia sobre los desempeños finales de

los evaluados frente a la prueba, o estos solamente se explican a partir de variables netamente producto de la formación educativa o corresponden a variables contextuales.

La hipótesis que guía este estudio es ad hoc, dado que depende de una observación de la cotidianidad de las instituciones educativas universitarias que tienen programas vigentes de psicología, y que muestran una tendencia similar en los desempeños.

PREGUNTA DE INVESTIGACIÓN

¿Qué variables sociodemográficas generan algún tipo de influencia en los puntajes de las competencias de razonamiento cuantitativo y competencias ciudadanas en estudiantes de psicología que presentaron la prueba SABER PRO, a nivel nacional durante el año 2020?

OBJETIVO GENERAL

Generar un modelo que permita describir la interacción entre variables sociodemográficas y puntajes en las competencias de razonamiento cuantitativo y competencias ciudadanas obtenidas por estudiantes de psicología a nivel nacional durante el año 2020.

OBJETIVOS ESPECÍFICOS

1. Realizar la inspección de datos utilizando estadísticos descriptivos para rastrear tendencias de investigación relevantes con respecto a los datos utilizados.

2. Verificar posibles asociaciones entre las variables de estudio, en especial variables sociodemográficas y variables cuantitativas tales como los puntajes globales de las competencias de razonamiento cuantitativo y competencias ciudadanas

3. Implementar pruebas estadísticas para verificar la hipótesis de que los puntajes de las pruebas de competencias ciudadanas y razonamiento cuantitativo se ven influenciadas por variables sociodemográficas propias de los programas de psicología a nivel nacional.

MARCO TEÓRICO

A continuación, se presentan los lineamientos teóricos que permiten contextualizar al lector sobre las variables utilizadas en el estudio y sus posibles relaciones, de modo que resulte comprensible el análisis de información que se presenta en el apartado resultados. Inicialmente se indican elementos esenciales de la prueba SABER PRO, y posteriormente se plantean características básicas de los procedimientos utilizados en el análisis estadístico.

Descripción de la prueba SABER PRO

La prueba SABER PRO tiene como objetivo realizar una evaluación educativa estandarizada de los desempeños de estudiantes universitarios (independientemente de su formación disciplinar/profesional) a nivel nacional con el fin de estimar su desempeño en cinco competencias básicas tales como razonamiento cuantitativo, competencias ciudadanas, lectura crítica, inglés y comunicación escrita. Dicho objetivo está fundamentado en la Ley 1324 de 2009, dando alcance a la exigencia gubernamental de realizar procesos de seguimiento de la

calidad de la educación superior y producir indicadores de verificación de la idoneidad de la formación disciplinar de los programas educativos ofrecidos en las diferentes IES (Instituciones de Educación Superior) a nivel nacional. El Decreto 3963 de 2009 indica que la prueba SABER PRO tiene como objetivo adicional generar acervos de información educativa estatal que pueda devolverse al Ministerio de Educación Nacional y las IES en virtud del mejoramiento de sus procesos de formación académica. La prueba SABER PRO se aplica a estudiantes que han aprobado mínimo el 75% de los créditos de su respectivo programa de formación. Durante la dinámica del examen, los evaluados responden a dos grupos de pruebas, las primeras denominadas de competencias genéricas, y las restantes, denominadas pruebas específicas. Además de responder a las pruebas mencionadas, los evaluados deben responder a un cuestionario de caracterización sociodemográfica, cuya información se utiliza para realizar los análisis planteados en este trabajo.

Según ICFES (2020) el cuestionario socioeconómico “permite obtener información que podría ayudar a explicar los resultados obtenidos en el examen sobre los procesos de enseñanza y aprendizaje de los estudiantes” (p. 13). La información capturada permite conocer datos de procedencia, contexto socioeconómico, características del núcleo familiar y de domicilio del estudiante, al igual que actividades complementarias a las actividades académicas.

Por otro lado, es necesario señalar que el examen SABER PRO presenta sus resultados de medición de desempeño en una escala de 0 a 300 puntos por cada competencia evaluada al igual que un indicador de desempeño global por competencia, y un indicador de desempeño general frente a la prueba. La primera sesión incluye las subpruebas de las cinco competencias genéricas más el cuestionario de caracterización, por lo cual, el estudiante debe responder a 153 preguntas en un tiempo máximo de 4 horas y 40 minutos, de las cuales 30 corresponden a los módulos de

razonamiento cuantitativo, lectura crítica y competencias ciudadanas, cada uno; 45 preguntas corresponden a inglés; y 1 corresponde a comunicación escrita. Adicionalmente, 17 componen el cuestionario de caracterización socioeconómica. A partir de la plataforma PRISMA las IES reportan si los estudiantes que presentarán el examen tienen necesidades educativas especiales que deban ser tenidas en cuenta a la hora de la aplicación o adecuación del examen. En la segunda sesión los estudiantes responden a las pruebas específicas que implican saberes concretos de su alcance disciplinar, y que son elegibles de acuerdo a la configuración individual o institucional ejecutada al momento del registro de inscripción en las plataformas habilitadas por el ICFES para tal fin. En ese caso, los estudiantes tienen un tiempo estimado de 1 hora y 30 minutos para responder al componente específico.

Adicionalmente, es necesario aclarar que el ICFES compara los desempeños de los evaluados con su grupo de referencia, es decir, una institución es comparada con programas similares en aspectos académicos, estructurales o económicos. En la presentación de los resultados esta información se condensa en los indicadores NBC (núcleo básico de competencia) que presentan una comparación directa con los programas homogéneos por grupo de referencia. De acuerdo a los núcleos comunes, el ICFES realiza combinatorias posibles de formatos de prueba que sean convenientes para la evaluación de las competencias específicas.

En cuanto a la estructura del examen, los reactivos utilizados en la prueba corresponden a la clasificación de ítems tipo pregunta de opción múltiple con única respuesta, excepto en la competencia de comunicación escrita, en el cual la pregunta es abierta y de construcción de un texto argumentativo. La estructura de los reactivos tipo pregunta implica la presentación de un enunciado, que presenta información necesaria y suficiente para responder al cuestionamiento implícito que permitirá evidenciar la competencia; una tarea que implica que el evaluado haga

algo con la información descrita; y finalmente cuatro alternativas de respuesta, de las cuales el evaluado debe escoger la que mejor ejecuta la tarea y da cuenta del enunciado provisto.

Competencias genéricas

Tal como lo define el ICFES (2020) se entiende una competencia como un compendio de habilidades para aplicar saberes en contextos específicos. Esta definición coincide con lo que dice Tobón (2006), quien reconoce que la evaluación de competencias educativas se debe plantear en un contexto evidenciable, de modo que la persona evaluada pueda demostrar que puede hacer algo con el conocimiento adquirido, y no solamente adquirir información que difícilmente podrá aplicar en la vida cotidiana posteriormente. Es necesario recordar que la prueba SABER PRO contempla cinco competencias genéricas, como lo muestra la Figura 1.

Figura 1.

Resumen de competencias genéricas medidas con la prueba SABER PRO



En el caso de Tobón (2008) las competencias evaluadas en el campo educativo tienen una exigencia mayor, dado que deben permitirle al sujeto ubicarse en el medio en el cual interactúa, es decir, teniendo en cuenta los lineamientos del ICFES (2020), el estudiante debe demostrar que

puede escribir un texto coherente, hacer una inferencia lógica, detectar una idea principal e ideas asociadas, tomar postura sobre un problema o interpretar las metacontingencias de un problema por resolver y sus posibles repercusiones.

En este caso el ICFES (2020) denomina *competencias de razonamiento cuantitativo* “al conjunto de elementos de las matemáticas (sean estos conocimientos o competencias) que permiten a un ciudadano tomar parte activa e informada en el contexto social, cultural, político, administrativo, económico, educativo y laboral” (ICFES, 2020, p. 29). Dicha competencia está compuesta por un subconjunto de competencias, a saber: interpretación y representación (34% de la cantidad de preguntas del módulo), que implica que el evaluado comprende y transforma la información para presentarla de diferentes maneras; formulación y ejecución (33% de la cantidad de preguntas del módulo), que implica que el evaluado plantea soluciones a problemas teniendo en cuenta una secuencia lógica de pasos; y argumentación (33% de la cantidad de preguntas del módulo), que implica que el evaluado evalúe procedimientos propuestos para dar cuenta de una adecuada solución de problemas. Se evalúa la competencia en concordancia con nociones básicas de estadística, geometría, álgebra y cálculo aproximando al evaluado a situaciones relacionadas con contextos familiares, personales, laborales, ocupacionales, de divulgación científica, entre otros.

En cuanto a competencias ciudadanas, se entiende como “los conocimientos y habilidades que posibilitan la construcción de marcos de comprensión del entorno, los cuales promueven el ejercicio de la ciudadanía y la coexistencia inclusiva dentro del marco que propone la Constitución política de Colombia” (ICFES, 2020, p. 35). Aquí se hace énfasis en que el ciudadano tenga conocimientos de la Ley Colombiana que le permitan reflexionar sobre las situaciones sociales y problemáticas que afectan la cotidianidad del ejercicio de derechos y

deberes ciudadanos. En este caso se hace fundamental el reconocimiento del contexto social, político, cultural y económico en el cual los evaluados se desenvuelven, y cómo entienden las problemáticas sociales asociadas.

El módulo mide las siguientes subcompetencias, a saber: conocimientos sobre la Constitución Política Colombiana, derechos y deberes y organización del Estado Colombiano (30% de las preguntas del módulo); argumentación sobre enunciados propuestos analizando su relevancia y alcance (20% de las preguntas del módulo); multiperspectivismo, que implica el desarrollo y análisis de diferentes perspectivas frente a solución de problemas sociales (30% de las preguntas del módulo); y finalmente la subcompetencia de pensamiento sistémico, que implica desarrollar una comprensión sobre los problemas y soluciones que implican diferentes niveles de análisis y relaciones entre los elementos del contexto social (20% de las preguntas del módulo).

Investigaciones relacionadas con la prueba SABER PRO en estudiantes de psicología

Según Suárez-Rodríguez (2018), dentro de las competencias de formación en psicología se evidencian tanto las competencias de evaluación en las cuales se trabajan aspectos relacionados con investigación cuantitativa, a la par que se relacionan competencias asociadas al entendimiento del contexto, el compromiso ético, la consideración e importancia de las normas y la articulación entre el ejercicio profesional y el contexto cotidiano de desempeño de la profesión. En síntesis, Suárez-Rodríguez (2018) se permite señalar que las competencias de razonamiento cuantitativo y competencias ciudadanas se incentivan en el proceso transversal de investigación de los formandos en psicología, sin embargo, los resultados en las pruebas SABER PRO no muestran los resultados esperados por los programas de psicología, ni por entidades tan

importantes como COLPSIC (Colegio Colombiano de Psicología) o ASCOFAPSI (Asociación Colombiana de Facultades de Psicología). Esto teniendo en cuenta que dentro de los programas de formación de cualquier programa de psicología a nivel nacional se inscriben estos procesos de formación como obligatorios para obtener el registro calificado.

Por su parte, Castro (2014) señala la necesidad de incentivar la el desarrollo de competencias ciudadanas en la formación de psicólogos y psicólogas dada la necesidad de sensibilizar los aprendizajes de la disciplina a la luz de la realidad sociopolítica colombiana, en la que los futuros profesionales tendrán que enfrentar el reto de participar en los procesos de implementación de acuerdos tan importantes como el Acuerdo de Paz, y la actualización de la política pública actual en materia social.

La información suministrada por Valbuena y García (2016) muestra que los desempeños de los estudiantes de psicología a nivel nacional sobre los puntajes promedio en las competencias de razonamiento cuantitativo ($X=9.73$) y competencias ciudadanas ($X=10.07$) son las más bajas comparados con las restantes competencias medidas por la Prueba SABER PRO (inglés, lectura crítica y comunicación escrita). Entidades importantes como COLPSIC y ASCOFAPSI señalan su preocupación debido a una tendencia al decremento de puntuaciones en las competencias de razonamiento cuantitativo y competencias ciudadanas en estudiantes de psicología a nivel nacional (Suárez-Rodríguez, 2018).

Para Rodríguez-Jiménez, Cote-Peña, Malaver-Tolosa, Pineda-Miranda & Sarmiento-Fúquene (2019) el análisis de los resultados de la prueba SABER PRO debe venir acompañado de una selección de variables tales como género, región, estrato socioeconómico, metodología del programa de formación, valor de matrícula, y carácter académico de la institución.

METODOLOGÍA

Diseño de investigación

Teniendo en cuenta el planteamiento del problema, esta investigación se denomina cuantitativa de alcance descriptivo correlacional, según lo sugieren Hernández, Fernández y Baptista (2014), que se enmarca en la denominación de diseños no experimentales, de corte transversal, dado que no se cumple la condición de manipulación de las variables porque éstas ya han ocurrido o se han medido en un momento determinado en el tiempo. Por otro lado, citando a León y Montero (2003) este estudio cumple con la denominación de diseño *expost-facto* de tipo retrospectivo simple también aludiendo a que las variables ya han ocurrido en el tiempo, y no es posible rastrear el inicio de su influencia recíproca.

Adicionalmente este diseño es ampliamente utilizado en investigación educativa, campo en el cual se deben trabajar variables preexistentes a la toma de datos, por lo cual tampoco se hace indicación de una estrategia de muestreo de datos, por lo cual, se plantea una exigencia mayor al modelo de predicción que se plantee posteriormente, dado que este debe ajustarse a un conjunto de datos preexistente, de modo que se deba modelar todas sus condiciones de interacción entre variables tal como ocurrió originalmente cuando se capturaron los datos.

Descripción de las técnicas de análisis de información

En el caso de este estudio se tuvieron en cuenta técnicas de análisis de información descriptiva, análisis de correspondencias múltiples, análisis de componentes principales y regresión beta. A continuación, se describen brevemente en qué consisten y cuál es su aporte a la investigación.

El análisis de correspondencias múltiples (ACM) tal como lo define Algañaraz (2016) permite analizar comportamientos conjuntos entre variables categóricas, específicamente en interacción de unidades de análisis y propiedades, en este caso modalidades de las variables categóricas-cualitativas que se tienen la matriz de datos obtenida. Dado que la mayoría de las variables de estudio son de orden categórico, tiene sentido utilizar una técnica de rastreo de información que permita graficar las interacciones entre las variables utilizando un plano de 2 ejes en el cual se realizan representaciones de las posibles asociaciones de las variables utilizadas en el cruce de información. Algañaraz (2016) indica que el ACM no solo permite evidenciar el comportamiento de variables cualitativas, sino que también permite rastrear sus posibles jerarquizaciones y diferenciaciones.

Por otro lado, el análisis de componentes principales (ACP) como lo sugiere Macías-Ruano (2015) permite reducir la dimensionalidad de la información contenida en indicadores cuantitativos a partir del análisis correlacional y la suposición de varianza compartida por los indicadores que se ingresan como unidades de análisis, en el caso puntual de esta investigación se toman como equivalentes las denominadas puntuaciones obtenidas por los evaluados en las pruebas de competencias ciudadanas y razonamiento cuantitativo, y puntualmente se utilizaron los indicadores de puntaje NBC (puntaje de núcleo común), para dar inicio con la exploración de la intercorrelación entre las variables tipo puntaje. El objetivo del ACP es organizar las variables iniciales en agrupaciones más reducidas que puedan simplificar la explicación de las condiciones de ocurrencia de las variables de origen.

De manera complementaria, Ferrari & Cribari-Neto (2004) entienden la regresión beta como un procedimiento estadístico que permite modelar datos a partir de una distribución tipo beta, a partir de una parametrización utilizando la media y la dispersión de los datos analizados.

Se aplica cuando la variable dependiente es continua, en este caso los puntajes en las competencias de razonamiento cuantitativo y competencias ciudadanas, y estos valores se distribuyen en el intervalo (0,1) vinculados a variables regresoras. Para el caso de este trabajo se hizo una transformación de la variable puntaje para que cumpliera la condición de la distribución beta, de modo que se adecuó la variable puntaje bruto dividiéndola por el valor de 300 para que pudiera satisfacer el alcance de la regresión beta. Como en otros métodos de regresión se utiliza la estrategia de máxima verosimilitud.

Otra ventaja de la distribución beta es que trabaja con valores positivos, lo cual tiene sentido porque los puntajes de las pruebas también se expresan en términos positivos. Otra de las razones de su elección como modelo de predicción de datos, está relacionado con la flexibilidad de la regresión beta para dar cuenta de la interacción con variables predictoras exógenas y que sus exigencias son menores comparadas con otros modelos de regresión aplicados en la práctica actualmente. Los indicadores claves de la regresión beta están asociados al valor de la media y el valor phi estimado.

Descripción de variables

En la Figura 2, presentada a continuación se inicia con la descripción de algunas variables de interés denominadas sociodemográficas, que se prueban como posibles variables regresoras dentro del modelo de regresión beta, y las variables tipo puntaje que se estiman como variables dependientes, dentro del proceso de análisis.

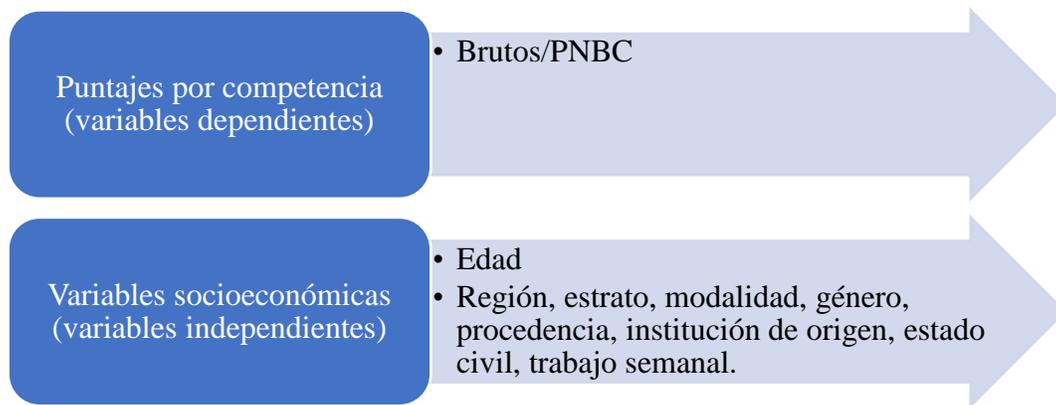
Tabla 1.*Listado de variables utilizadas en el estudio*

Nombre	Tipo	Descripción
Puntaje PNBC	Cuantitativo discreto	Puntaje con variación de 1 a 100
Puntaje Bruto	Cuantitativo discreto	Puntaje con variación de 1 a 300
Nivel de competencia	Cualitativa-ordinal	4 modalidades organizadas jerárquicamente (nivel 1 al 4)
Estrato	Cualitativa-ordinal	7 modalidades (Sin estrato, estratos del 1 al 6)
Género	Cualitativa-nominal	2 modalidades (masculino, femenino)
Región	Cualitativa-nominal	5 modalidades (Amazonía, Orinoquía, Pacífica, Caribe, Andina)
Procedencia	Cualitativa-nominal	2 modalidades (Cabecera municipal y Área rural)
Estado civil	Cualitativa-nominal	4 modalidades (unión libre, casado, soltero, viudo-divorciado)
Trabajo semanal	Cualitativa-nominal	6 modalidades (multicategórica)
Modalidad	Cualitativa-nominal	3 modalidades (presencial, distancia-virtual, distancia)
Institución de origen	Cualitativa-nominal	2 modalidades (oficial y no oficial)

Como se puede observar la mayoría de las variables utilizadas es de orden cualitativo, razón por la cual se debe utilizar una distribución y tipo de regresión que permita el manejo de este tipo de variables, por tal razón se utilizó la regresión beta. Por otro lado, la Figura 2 presenta información sobre la disposición de las variables dependientes e independientes dentro del estudio. Concretamente, las variables cualitativas-categorías se asumen aquí como variables independientes, y las variables puntaje se asumen como variables dependientes, sobre las cuales se montan los procedimientos de regresión tipo beta. A continuación, se presenta la disposición de variables:

Figura 2.

Distribución de variables por tipología dentro de los análisis de regresión.



Procedimiento

Se plantea realizar los análisis requeridos para el ejercicio de comprobación de hipótesis descriptivas - correlacionales y explicativas utilizando el software RStudio para avanzar con la tarea de análisis de las variables, teniendo en cuenta los procedimientos estadísticos antes mencionados.

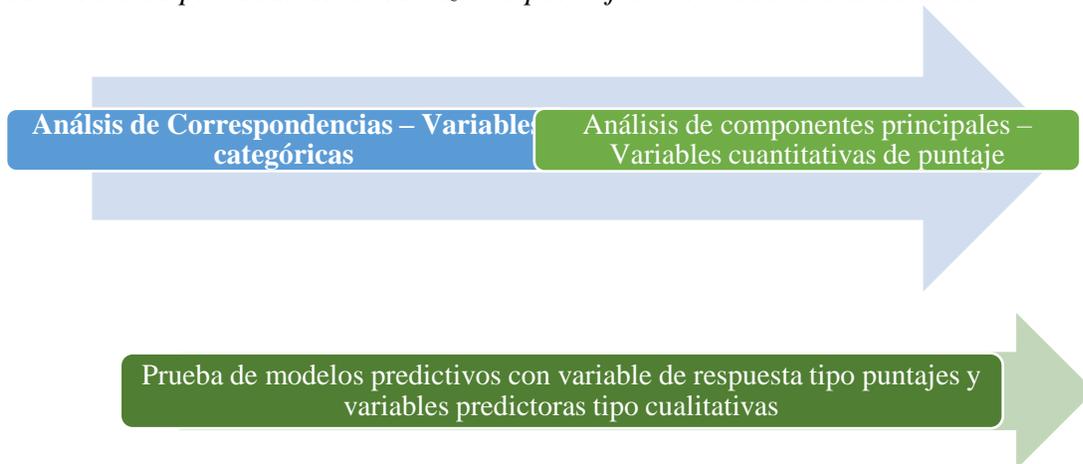
Concretamente se requiere realizar diferentes fases en el análisis de información:

1. Analizar el comportamiento de las variables cualitativas y cuantitativas, de modo que se detecta si ofrecen suficiente variabilidad para poder modelar una interacción que se pueda probar con una distribución teórica de datos.
2. Verificar asociación significativa entre las variables de estudio, de manera que se pueden escoger las variables con la mejor interacción posible para plantear un modelo de asociación-predicción.
3. Plantear una posible estimación de predicción de los puntajes de las competencias de razonamiento cuantitativo y competencias ciudadanas, a partir de las variables sociodemográficas seleccionadas, a partir del uso de un modelo de regresión tipo beta.

El anterior proceso también se puede ejemplificar en el siguiente gráfico de flujo enlistado en la Figura 3 que se presenta a continuación.

Figura 3.

Secuencia de procedimientos realizados para ejecutar el análisis estadístico



Procedimientos para la recolección y captura de datos

Los datos utilizados para la ejecución de esta investigación corresponden a datos derivados de la base de datos del ICFES correspondientes al año 2020. La información se accedió a partir de la página oficial de la institución, consultando el link:

<https://www.icfes.gov.co/investigadores-y-estudiantes-posgrado/acceso-a-basesde-datos>

Una vez se ingresó al portal descrito, se procedió a diligenciar un formulario exigido en el cual se indicaron datos de identificación del investigador y su filiación académica, de modo que posteriormente se recibiera una clave asignada por el sistema dataicfes, con el fin de acceder a la carpeta compartida por la institución. Una vez realizada la gestión de información, se procedió a descargar el archivo por tipología del examen, en este caso, examen SABER PRO, y por año de ejecución de la prueba, en este caso año 2020. Se obtuvo un archivo de texto plano (.txt) que luego fue llevado a Excel para su transformación, y posteriormente cargado al programa R. La base de datos original estaba compuesta de 51 variables, pero se tuvo que implementar un filtro de las 18 variables más relevantes para el estudio como se indicó previamente. Se conformó una nueva base de datos a la cual se le hizo un tratamiento previo de verificación de datos perdidos, lo cual implicó pasar de 16989 casos a 16613 casos con los cuales se realizó el análisis.

RESULTADOS Y DISCUSIÓN

En esta sección se presentan los resultados más sobresalientes del análisis de información recolectada. En principio se presenta el rastreo de estadísticos descriptivos de las variables utilizadas para hacer el análisis de interacciones; posteriormente se presentan los resultados de los análisis de correspondencias múltiples para variables categóricas-cualitativas; luego se presentan los resultados del análisis de componentes principales propuesto con las variables tipo puntaje de las competencias de razonamiento cuantitativo y competencias ciudadanas; y finalmente se presentan los resultados de la estimación de la regresión beta, incluyendo los principales hallazgos.

Análisis descriptivos

En principio es necesario indicar el comportamiento general de las variables de estudio, de modo que se presentan seguidamente los datos obtenidos de las tendencias de las variables cuantitativas, y posteriormente lo observado con las variables cualitativas. En la Figura 4 se presentan los resultados más sobresalientes para las puntuaciones de competencias ciudadanas y razonamiento cuantitativo.

Figura 4

Descriptivos para las variables razonamiento cuantitativo y competencias ciudadanas

RAZCUANTBRUT	CIUDADBRUT
Min. : 67.0	Min. : 58.0
1st Qu.:117.0	1st Qu.:132.0
Median :136.0	Median :152.0
Mean :136.7	Mean :150.9
3rd Qu.:155.0	3rd Qu.:171.0
Max. :300.0	Max. :300.0

A partir de los resultados se puede observar que la media de desempeño en razonamiento cuantitativo (136.7 puntos) es menor comparada con la media de desempeño en competencias ciudadanas (150.9), y las puntuaciones oscilan con un mínimo de 58 y un máximo de 300 en el caso de competencias ciudadanas, y 67 y 300 en el caso de razonamiento cuantitativo. Se puede observar que la dispersión de puntuaciones es mayor en el caso de la variable competencias ciudadanas, dado que sus valores mínimos y máximos superan a los presentados por la variable razonamiento cuantitativo.

Por otro lado, se observa en la Figura 5 el comportamiento de las variables cualitativas distribuidas por sus correspondientes modalidades, en las cuales se presentan las frecuencias de casos, según corresponde, a saber:

Figura 5.

Datos descriptivos de las variables cualitativas sociodemográficas

GENERO	PROCEDENCIA	REGION	ESTADOCIVIL
F:13741	Area Rural : 1930	AMAZONIA : 271	Casado : 1695
M: 2872	Cabecera Municipal:14683	ANDINA :10874	Separado y/o Viudo: 216
		CARIBE : 2708	Soltero :13038
		ORINOQUIA: 561	Unión libre : 1664
		PACIFICA : 2199	
TRABAJOSEMANAL	MODALIDADPROGRAMA	INST_ORIGEN	
0 :2693	DISTANCIA : 4694	NO OFICIAL:13786	
Entre 11 y 20 horas:3325	DISTANCIA VIRTUAL: 1144	OFICIAL : 2827	
Entre 21 y 30 horas:2786	PRESENCIAL :10775		
Más de 30 horas :5300			
Menos de 10 horas :2509			

Se puede observar que la mayor parte de las personas que presentan el examen SABER PRO para el caso de psicología son mujeres con un total de 13741 (82,71%) comparado con los 2872 hombres (17,29%). En cuanto a la variable procedencia se observa que la mayoría de los evaluados pertenecen a la cabecera municipal contando con 14683 casos (88,38%), mientras que un grupo menor de 1930 evaluados proviene del área rural (11,61%).

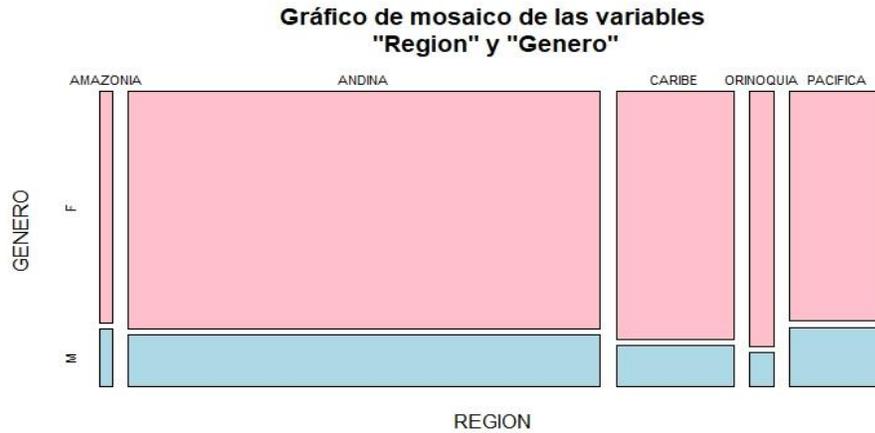
Por otro lado, se observa que la región que aglutina más casos de evaluados es la región andina con una representación del 65.45%, mientras que las regiones caribe y pacífica tienen porcentajes de representación similares correspondientes a 16,30% y 13,23%. Por su parte las regiones Orinoquía y Amazonía contribuyen con una menor cantidad de casos y sus porcentajes de representación son bajos comparados con las demás regiones, 3,37% y 1,63% respectivamente.

En el caso de la variable estado civil se evidencia que la mayoría de evaluados son solteros (78,48%), seguido de las modalidades casado (10,20%), y unión libre (10,01), y con menor cantidad de casos la modalidad separado y/o viudo (1,30%). En cuanto a la variable trabajo semanal se observa que la mayor cantidad de casos corresponde a la modalidad trabajar más de 30 horas semanales (31,90%), comparada con otras modalidades como no trabajar (16,21%), trabajar menos de 10 horas (15,10%), trabajar entre 11 y 20 horas (20,01%), trabajar entre 21 y 30 horas (16,76%). Se observa que la mayoría de los participantes reporta realizar alguna actividad laboral (83,79%).

Se observa que la mayor cantidad de casos de la variable modalidad del programa de estudio corresponde a la modalidad presencial (64,85%), comparado con las modalidades distancia (18,25%) y distancia-virtual (6,88%); y en cuanto a la variable institución de origen, la cual cuenta con dos modalidades, a saber, instituciones oficiales (17,01%) y no oficiales (82,98%), la mayoría de casos se aglutinan en la modalidad no oficial.

Como se observa en el gráfico de mosaico que se presentan en la Figura 6, se puede indicar que las regiones que mayor cantidad de evaluados aportan son las regiones andina, caribe y pacífico, y que el género predominante es femenino, para el caso de estudiantes de psicología.

Figura 6.



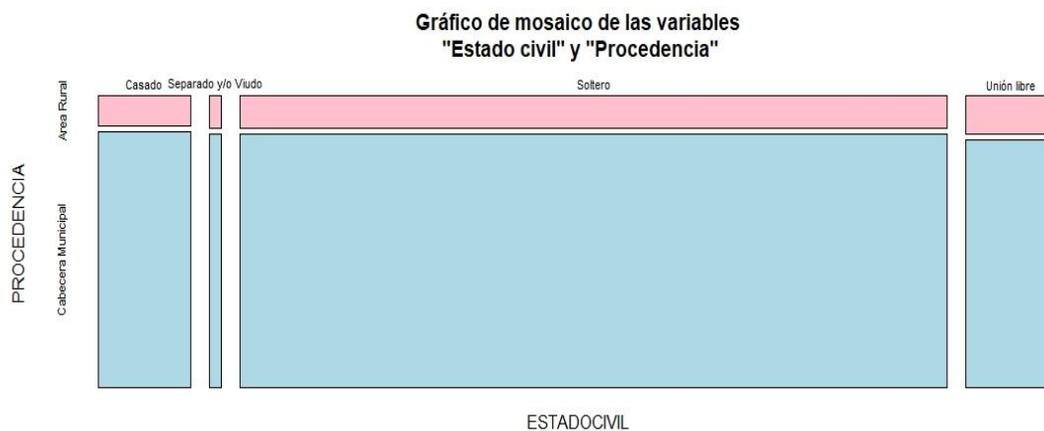
Por su parte, las interacciones entre las variables modalidad del programa e institución de origen indican que hay una predominancia de la educación privada sobre la educación pública en todas las modalidades de formación (presencial, distancia virtual y distancia). En el caso concreto, no hay evaluados que presenten la condición de pertenecer a modalidad distancia virtual y pertenecer a una institución de carácter público.

Figura 7



En cuanto a la Figura 8, es claro que es más común en los evaluados ser soltero y pertenecer al área urbana que presentar los demás estados civiles y encontrarse ubicado en el área rural. La posibilidad de tener estado civil separado y/o viudo, y pertenecer al área rural es poco común dentro de los evaluados que presentaron la prueba SABER PRO.

Figura 8



Análisis de correspondencias múltiples

En esta sección se presentan los resultados de los análisis de correspondencias múltiples establecidos, de acuerdo al análisis de interacción de las variables cualitativas seleccionadas para dar continuidad al ejercicio de investigación. En la Figura 9 se presentan los resultados del análisis de correspondencias para las variables estrato y región. Se observa que los desempeños se asocian con pertenecer a región andina y estratos más altos. Se encuentra asociación entre pertenecer a la región Caribe y vivir en estrato 1. También se puede observar que las regiones Orinoquía y Amazonía no se ven bien asociadas, al igual que la modalidad sin estrato, aunque el pertenecer a la región Caribe está relacionado con vivir en estrato 1. El porcentaje de varianza

explicada para el primer eje corresponde al 96.5% lo cual indica que las variables muestran fuerte asociación, al menos en alguna dimensión relacionada con la varianza explicada por las variables incluidas en el análisis.

Figura 9

Análisis de correspondencias de las variables estrato y región

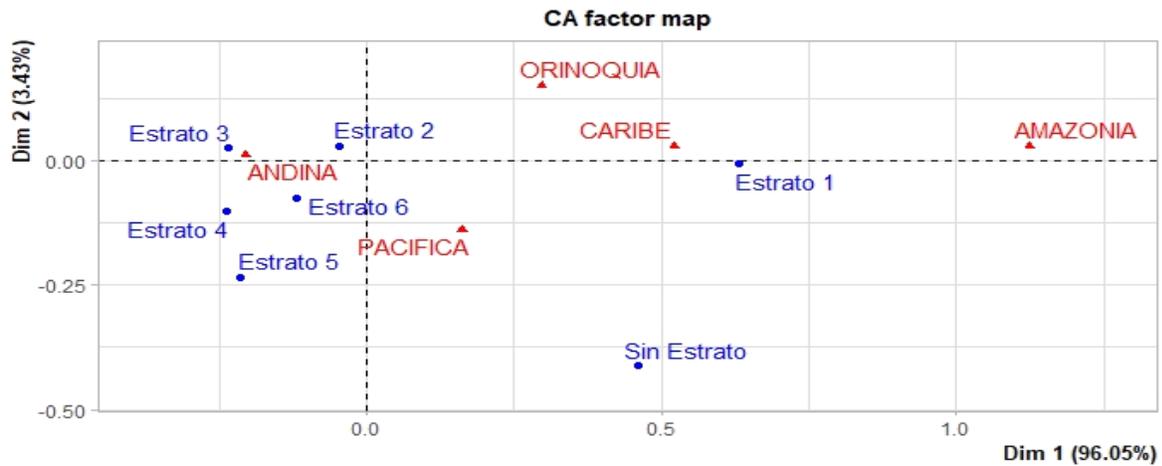
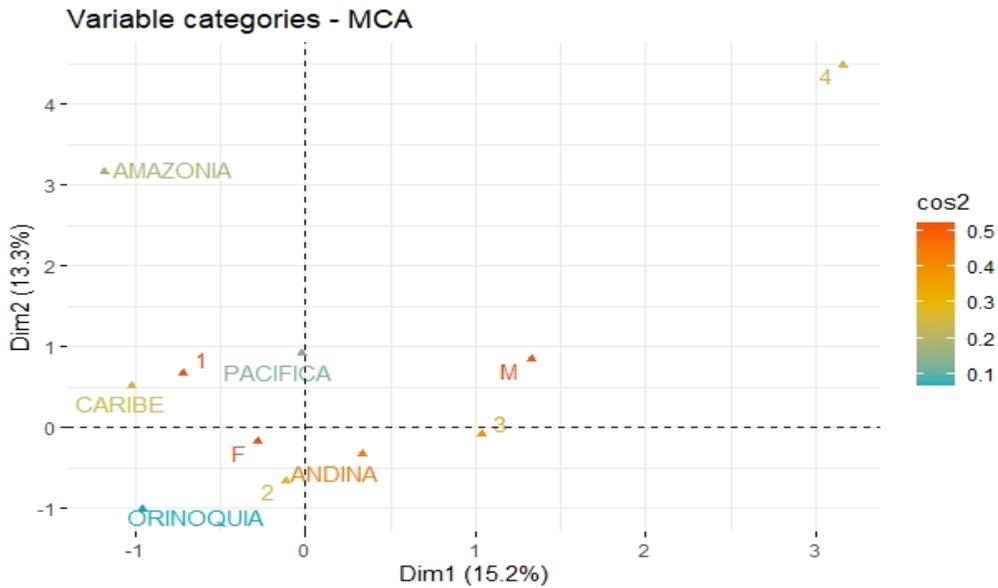


Figura 10

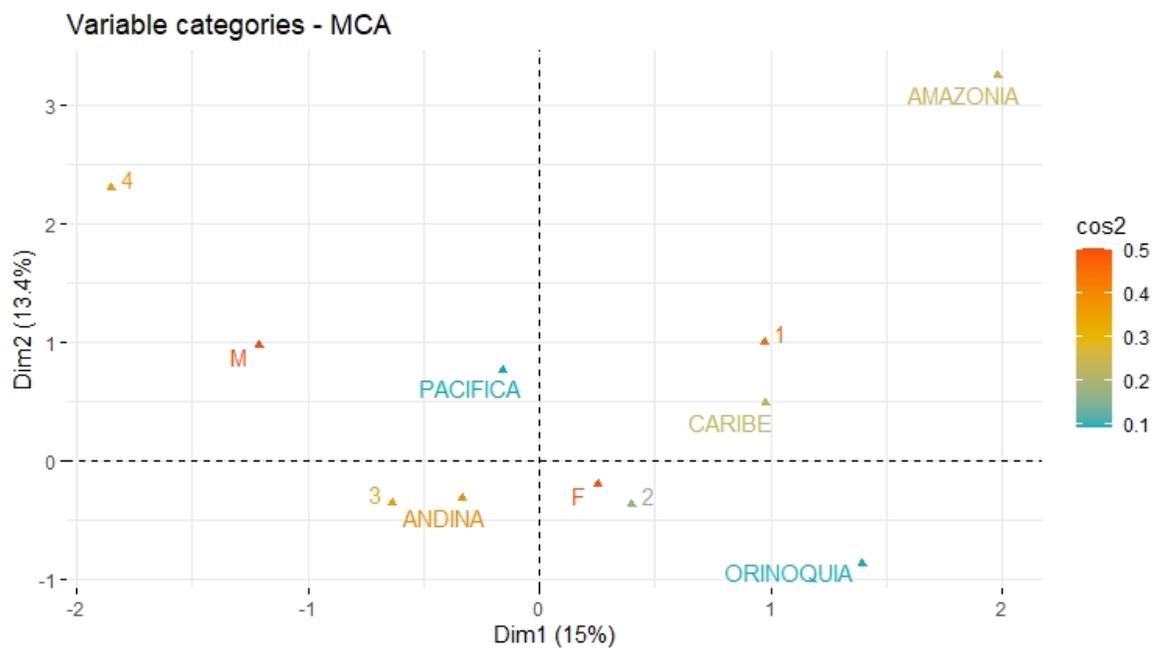
Nivel de competencia de razonamiento cuantitativo vs región y género



En cuanto al nivel de competencia de razonamiento cuantitativo se observa asociación entre ser hombre y el nivel de competencia 3. También pertenecer a la región andina, se mujer y el nivel de competencia 2. Las regiones Orinoquía y Amazonía no aparecen bien representadas en el análisis, al igual que el nivel de competencia 4. Por otro lado, se encuentra asociación entre el nivel de competencia 1 y las regiones Caribe y Pacífica.

Figura 11

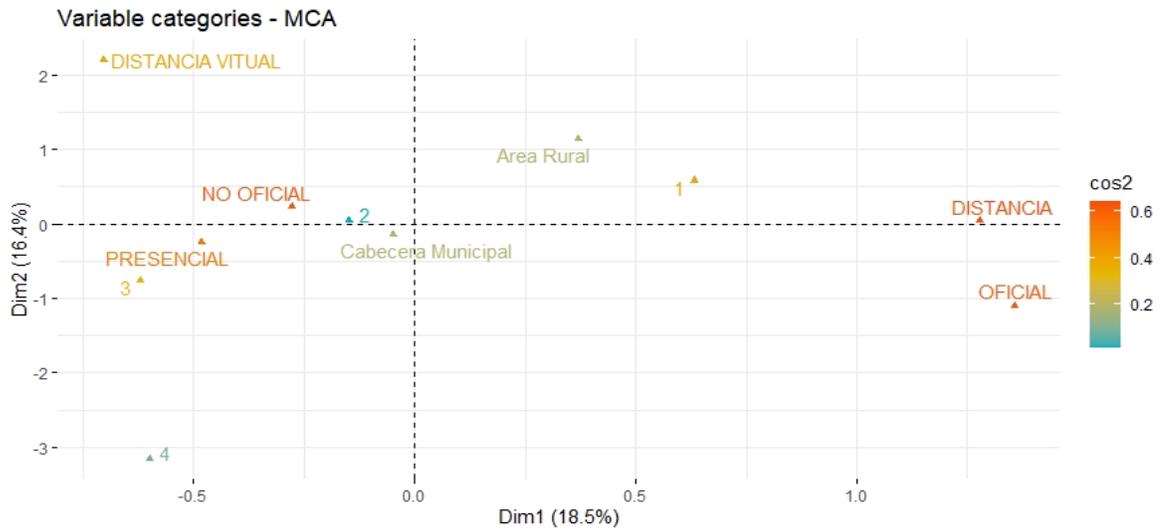
Nivel de competencia de competencias ciudadanas vs género y región



En la Figura 11 se observa asociación entre los niveles de aptitud de competencias ciudadanas y las variables región y género. Puntualmente se observa que pertenecer a la región andina, y el nivel de competencia 3 muestra una fuerte asociación. También se evidencia asociación entre ser mujer y el nivel de competencia 2, al igual que pertenecer a la región Caribe y demostrar el nivel de competencia 1. Las modalidades nivel de competencia 4, ser hombre y las regiones Orinoquía y Amazonía no muestran asociación con otras variables.

Figura 12

Nivel de razonamiento cuantitativo vs modalidad y procedencia



En la Figura 12 se evidencia asociación entre nivel de competencia 2, pertenecer a institución no oficial y estar ubicado en la cabecera municipal. También pertenecer a formación presencial y el nivel de competencia 3. Por otro lado, se muestra asociación entre el nivel de competencia 1 y pertenecer al área rural. Las modalidades tipo de formación distancia virtual, distancia, nivel de competencia 4, y pertenecer a universidad pública no muestran asociación significativa.

Observando la Figura 13 se enlista la asociación de las variables nivel de aptitud en competencias ciudadanas, modalidad y procedencia. Se evidencia que existe asociación entre pertenecer a la modalidad presencial y el nivel de competencia 3. El nivel de competencia 1 tiende a asociarse con la formación a distancia, al igual que el nivel de competencia 2 se asocia con pertenecer al área rural, y pertenecer a la cabecera municipal y pertenecer a la modalidad no oficial.

Figura 13

Nivel de competencias ciudadanas vs modalidad y procedencia

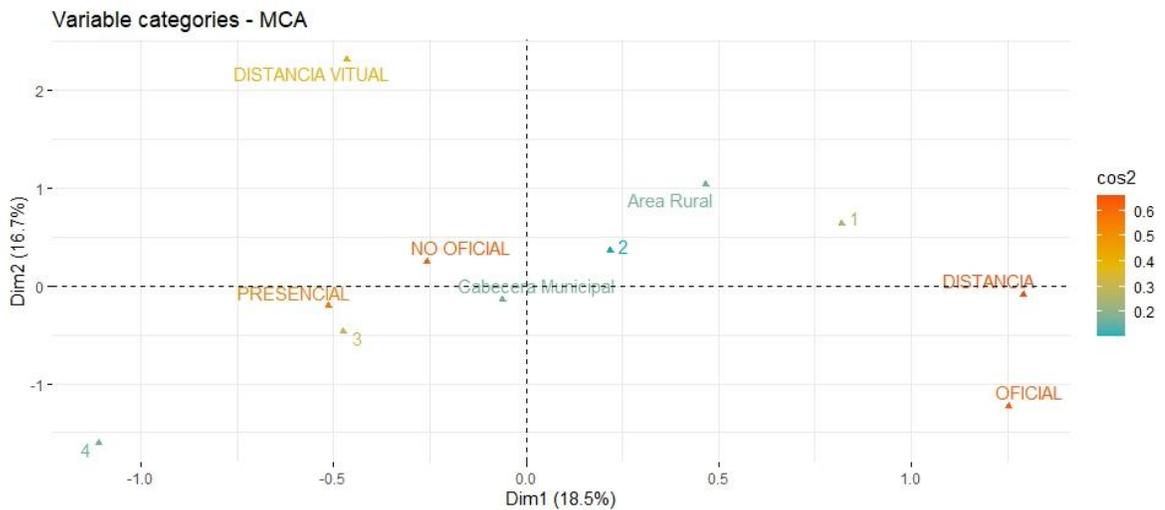
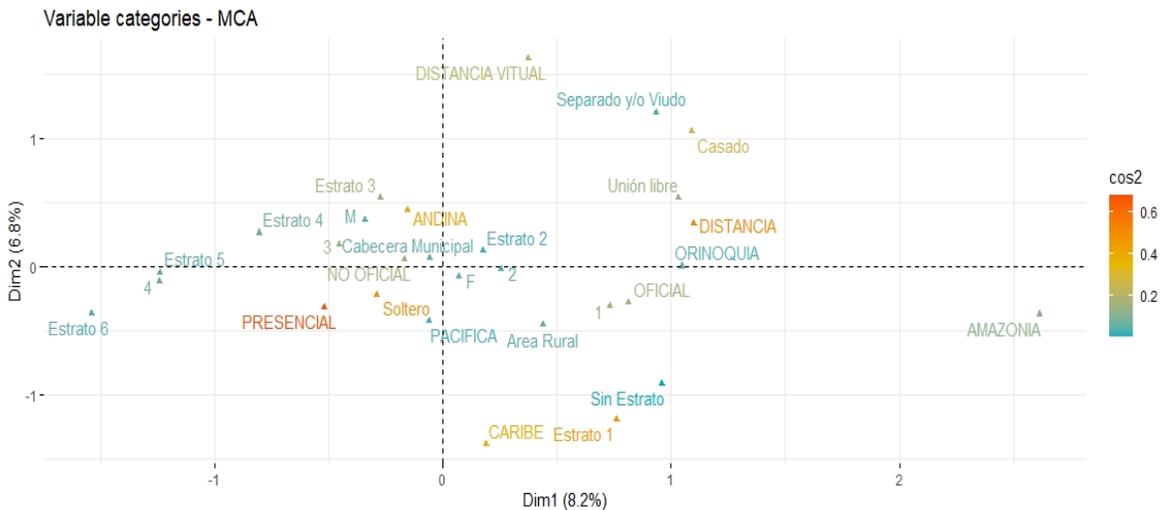


Figura 14

Variables sociodemográficas cualitativas vs nivel de competencia de razonamiento cuantitativo

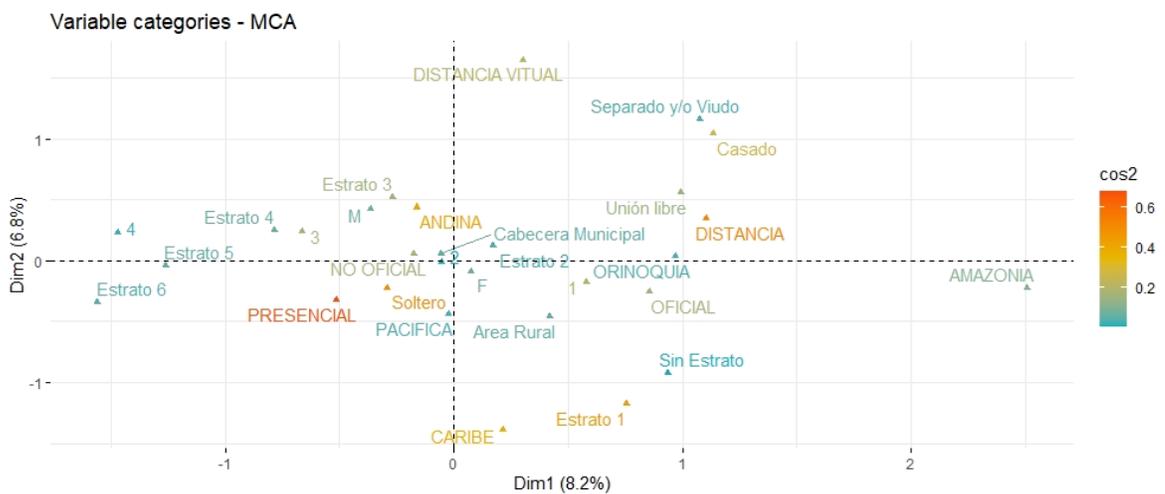


En la Figura 14 se muestra la interacción entre las variables sociodemográficas relevantes para el estudio y el nivel de competencia de razonamiento cuantitativo. Por otro lado, se registra asociación entre el nivel de competencia 3, ser hombre, pertenecer a los estratos 3 y 4, pertenecer

al área urbana y pertenecer a la región Andina. También llama la atención la asociación que existe entre el nivel de competencia 2, ser mujer y pertenecer al estrato 2. Adicionalmente se encuentra asociación entre pertenecer a una universidad pública y el nivel de competencia 1, y en contraste se observa asociación entre pertenecer al estrato 5 y el nivel de competencia 4. Se observa que, entre mayor cantidad de variables incluidas en el análisis, los porcentajes de varianza explicada disminuyen.

Figura 15

Variables sociodemográficas cualitativas vs nivel de aptitud de competencias ciudadanas

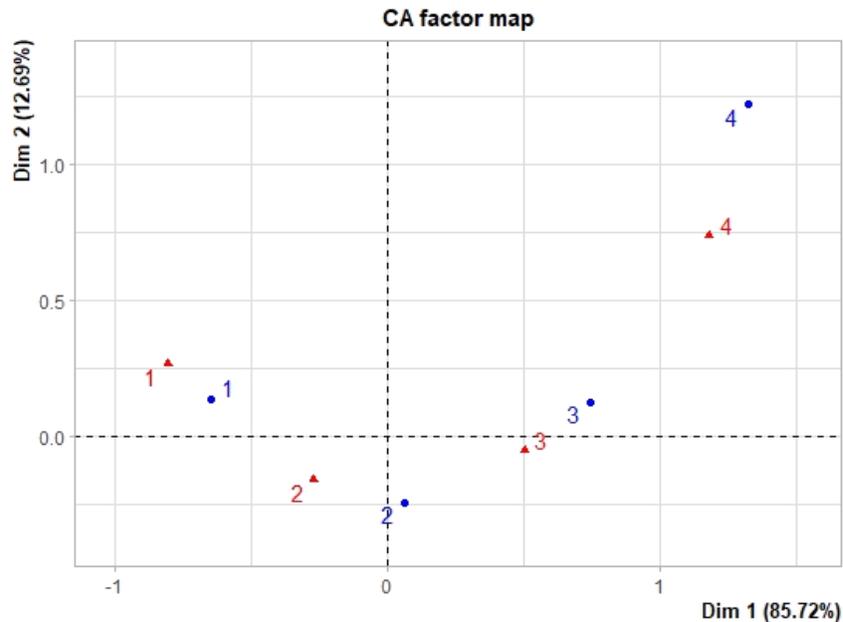


En la Figura 15 se puede rastrear asociaciones entre las variables sociodemográficas seleccionadas para el estudio y el nivel de aptitud en competencias ciudadanas. Se observa fuerte asociación entre obtener el nivel de competencia 3, ser hombre y pertenecer al estrato 4. También se observa que las modalidades nivel de competencia 1, estrato 2, y pertenecer al área rural y a la región Orinoquía tienden a variar conjuntamente y asociarse. El nivel de competencia 4 se ve asociado a estratos altos tales como el estrato 5 y 6. La modalidad presencial aunque bien representada, no tiende a asociarse a un nivel específico de competencia, mientras que el nivel de

competencia 2 se asocia a ser soltero, pertenecer a la región Andina, y estudiar en universidad privada.

Figura 16

Asociación de niveles de competencia de razonamiento cuantitativo y competencias ciudadanas



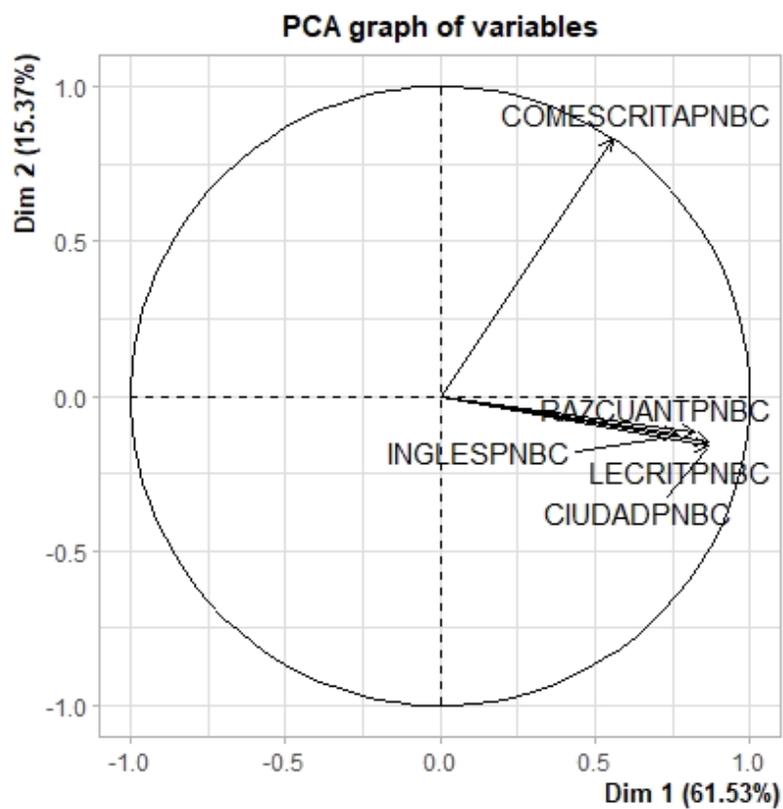
Al comparar los niveles de asociación de las dos competencias centrales de estudio se puede observar homogeneidad en la disposición de los niveles de desempeño de los evaluados con respecto al nivel de exigencia que plantea la prueba. Se observa mayor asociación entre los niveles más bajos de ambas competencias, sin embargo, cuando se analiza la disposición de los desempeños del nivel 4, ya no se muestra asociación fuerte, ni representación significativa en el eje de coordenadas. El nivel de competencia más homogéneo en ambas competencias es el nivel 3, aunque los niveles 1 y 2 también muestran asociación llamativa.

Análisis de componentes principales de variables tipo puntaje

A continuación, se presentan los resultados del análisis de componentes principales (ACP) desarrollado con las variables puntaje de las 5 competencias medidas por la prueba SABER PRO, a saber: competencias ciudadanas, razonamiento cuantitativo, inglés, comunicación escrita y lectura crítica. Cada variable representa un puntaje total obtenido con la prueba. Se presenta el resultado del rastreo de asociaciones entre variables. En la Figura 17 se incluyen los resultados principales del procedimiento ejecutado.

Figura 17

Resultados del ACP para las variables tipo puntaje de las competencias genéricas



Según la Figura 17, se logra explicar aproximadamente un 76,9% con dos dimensiones claramente identificadas que logran agrupar los cinco indicadores de competencias medidas por la prueba SABER PRO para las variables tipo puntaje. En la primera dimensión se tienden a asociar las variables competencias ciudadanas, inglés, lectura crítica y razonamiento cuantitativo, compartiendo un 61,53% de comunalidad, lo cual indica que tienden a asociarse entre sí. Aunque parecen disímiles, las variables aquí incluidas comparten algo en común, y es que obligan al evaluado a poner en juego su competencia lectora, independientemente que sea en español o inglés. La habilidad lectora y la capacidad de interpretación se ponen en juego en este tipo de actividades propuestas en la prueba SABER PRO.

De manera distinta se comporta la competencia de comunicación escrita, en la cual los evaluados deben responder a una sola tarea que implica planear, construir y ejecutar un texto escrito, incluyendo su macroestructura, función, propósito, sintaxis y ortografía para lograr comunicar clara y coherentemente un mensaje con una intencionalidad clara. En el caso del análisis de componentes principales, esta variable tiende a comportarse casi independientemente de las cuatro competencias genéricas restantes alcanzando un 15,37% de explicación de la varianza total, conformando ella sola la dimensión 2.

Llama la atención que la competencia aquí descrita no implica comprensión textual sino argumentación a partir de la elaboración de un texto propio, y teóricamente tiene sentido comparándolo con el resultado de asociación de la dimensión 1. El análisis de componentes principales logra evidenciar que las variables puntaje tienden a asociarse entre sí, y por lo que respecta al ejercicio propio de esta investigación, se podría indicar que las variables razonamiento cuantitativo y competencias ciudadanas tienden a mostrar comportamientos

similares, por lo cual se esperaría que su posibilidad de modelado a partir de la regresión beta mostrara resultados similares.

Resultados de la regresión beta

Como se mencionó en apartados anteriores, se seleccionó la regresión beta dado que ofrece ventajas frente a otros tipos de regresiones, en cuanto a su mayor flexibilidad a la hora de explicar conjuntos de datos que muestran variables exógenas, implicación de una cantidad considerable de variables explicativas, y potencial de explicación de la variable respuesta puntaje en razonamiento cuantitativo y competencias ciudadanas. Los resultados que se presentan a continuación dan cuenta de los modelos planteados para las competencias de razonamiento cuantitativo y competencias ciudadanas. Es necesario recordar que la regresión tipo beta se basa en la distribución beta que implica dentro de sus condiciones generales que los valores medios de la variable respuesta se encuentren en el intervalo $(0, 1)$, y el coeficiente phi debe ser mayor que 0.

Resultados asociados a la predicción del puntaje de razonamiento cuantitativo

Los resultados de la estimación de la regresión beta incluyendo las variables sociodemográficas cualitativas utilizadas en el análisis de interacción con las variables puntaje de competencia de razonamiento cuantitativo se muestran en la Figura 18, en la cual se puede evidenciar que se ejecutó una regresión beta, observando el cumplimiento de las primeras dos condiciones, dado que el valor medio debe estar entre 0 y 1, y en este caso el valor estimado se encuentra alrededor de 0.0023. Por otro lado, el valor del parámetro phi debe ser mayor que 0, y en este caso, el estimado de phi es igual a 32.98. Ferrari & Cribari-Neto (2004) indican que el valor phi es mucho mejor entre tanto sea más grande, dado que tiene un efecto sobre el control

de la varianza de las puntuaciones. En este caso el valor es pequeño, por lo tanto el modelo, aunque resulta significativo según el indicador p-valor sea cercano a 0, indica que la varianza de las puntuaciones no está controlada, factor que podría afectar al modelo estimado.

Figura 18.

Estimación del modelo de regresión beta para puntaje de razonamiento cuantitativo

```

standardized weighted residuals 2:
      Min      1Q  Median      3Q      Max
-3.4750 -0.6670  0.0023  0.6793 16.3227

Coefficients (mean model with logit link):
              Estimate Std. Error z value Pr(>|z|)
(Intercept)   -0.437220   0.025110 -17.412 < 2e-16 ***
GENEROM        0.119526   0.007119  16.791 < 2e-16 ***
PROCEDENCIACabecera Municipal  0.063191   0.008396   7.527 5.20e-14 ***
REGIONANDINA   0.122105   0.021649   5.640 1.70e-08 ***
REGIONCARIBE   -0.032113   0.022575  -1.423 0.15488
REGIONORINOQUIA 0.074018   0.025810   2.868 0.00413 **
REGIONPACIFICA 0.047354   0.022670   2.089 0.03672 *
ESTADOCIVILSeparado y/o viudo -0.057924   0.025094  -2.308 0.02099 *
ESTADOCIVILSoltero  0.020848   0.009537   2.186 0.02881 *
ESTADOCIVILUnión libre -0.034086   0.012017  -2.837 0.00456 **
TRABAJOSEMANALEntre 11 y 20 horas -0.095996   0.009001 -10.666 < 2e-16 ***
TRABAJOSEMANALEntre 21 y 30 horas -0.062942   0.009346  -6.734 1.65e-11 ***
TRABAJOSEMANALMás de 30 horas  -0.014474   0.008381  -1.727 0.08419 .
TRABAJOSEMANALMenos de 10 horas -0.101811   0.009618 -10.585 < 2e-16 ***
MODALIDADPROGRAMADISTANCIA VIRTUAL 0.048619   0.011726   4.146 3.38e-05 ***
MODALIDADPROGRAMAPRESENCIAL  0.177753   0.006807  26.114 < 2e-16 ***
INST_ORIGENOFICIAL  0.070600   0.007425   9.508 < 2e-16 ***

Phi coefficients (precision model with identity link):
      Estimate Std. Error z value Pr(>|z|)
(phi) 32.9880   0.3566   92.5 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Type of estimator: ML (maximum likelihood)
Log-likelihood: 1.735e+04 on 18 Df
Pseudo R-squared: 0.1105
Number of iterations: 23 (BFGS) + 2 (Fisher scoring)

```

Se puede especular que el valor phi pueda verse afectado por la cantidad de variables regresoras independientes que se incluyeron dentro del modelo, y la modalidad cualitativa de las variables, hecho que puede interferir en el hecho de que la varianza de las puntuaciones no esté controlada adecuadamente. Al observar los resultados se encuentra que el modelo no es significativo para las variables región Caribe y trabajar más de 30 horas a la semana. En términos de los indicadores de calidad del modelo se debe tener en cuenta el valor del indicador Pseudo R-

squared que, según Ferrari & Cribari-Neto (2004), debe oscilar entre 0 y 1, entre mayor sea el valor, mayor será la precisión del modelo planteado. En este caso, esta es una debilidad de la estimación, ya que a pesar de que el valor es significativo según el p-valor, el indicador Pseudo R squared corresponde solo al 0.1105. Puede indicarse que el valor de estimación y potencial de alcance del modelo podría haberse visto afectado por la cantidad de variables independientes que se incluyeron en el análisis de regresión, y que algunas de ellas son dicotómicas y otras politómicas, lo cual podría haber afectado su variabilidad. Podría probarse la segmentación de modelos analizando por separado la interacción de variables dicotómicas y politómicas, y verificar los resultados.

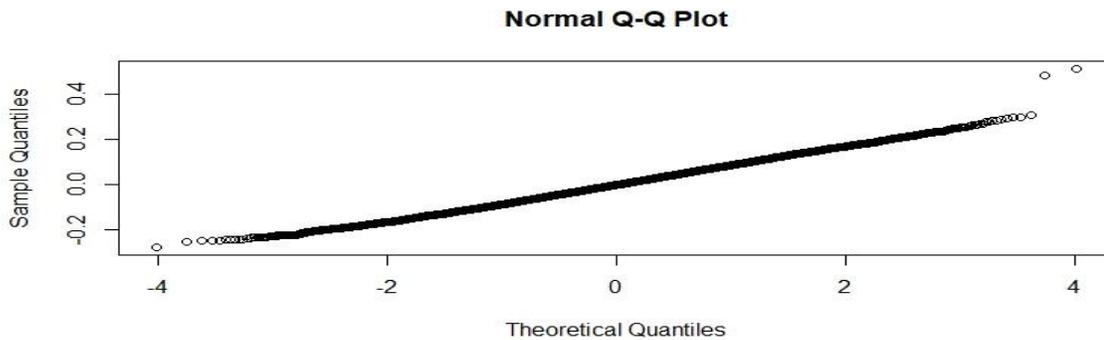
Concretamente se podría expresar la interacción de las variables a partir de la siguiente expresión:

$$g(\mu t) = -\beta_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 - \beta_7 x_7 + \beta_8 x_8 - \beta_9 x_9 - \beta_{10} x_{10} - \beta_{11} x_{11} - \beta_{12} x_{12} + \beta_{13} x_{13} + \beta_{14} x_{14} + \beta_{15} x_{15}$$

Se puede indicar que el modelo propuesto para describir el cambio de las puntuaciones de razonamiento cuantitativo implica que existe una relación negativa entre la variable dependiente y ser separado y/o viudo, vivir en unión libre, e invertir tiempo a la semana en trabajar. El modelo indica una relación positiva entre la variable respuesta y la modalidad de género masculino, pertenecer a la cabecera municipal, pertenecer a la región Andina, Orinoquía o Pacífica, ser soltero, pertenecer a la modalidad presencial y distancia-virtual, y estar estudiando en una entidad pública.

Figura 19

Regresión Beta para modelar interacciones con la variable razonamiento cuantitativo

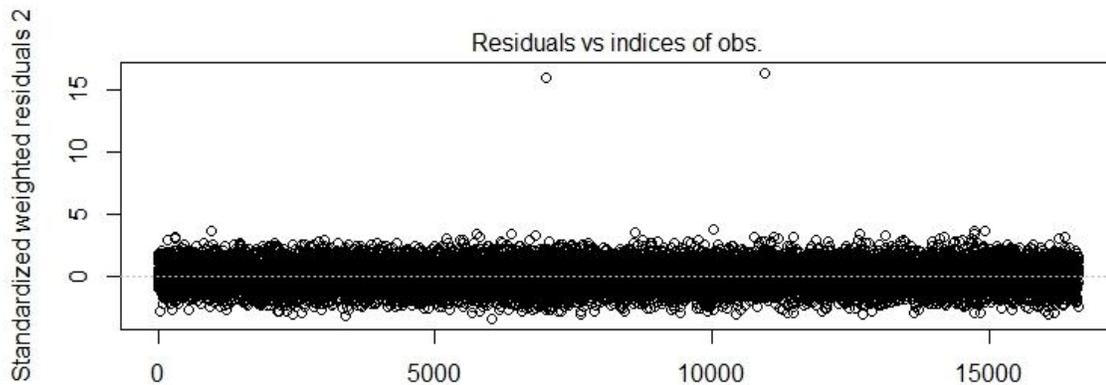


Utilizando el resultado del qqplot se puede observar que la mayoría de datos tienden a ajustarse a la distribución señalada por la recta teórica, hecho que muestra que los datos tienden a agruparse bajo la función establecida, luego de plantear la regresión beta, sin embargo, se presentan aún datos que se escapan de la explicación del modelo. Dichas observaciones podrían constituir un comportamiento no explicado por el modelo, razón por la cual los indicadores de precisión/calidad del modelo se hubiesen visto afectados. Los datos extremos en las colas superior o inferior podrían sugerir asimetría en el comportamiento de los datos de puntaje en la competencia de razonamiento cuantitativo.

El resultado mostrado en la Figura 20, indica que si bien se presentan observaciones que se alejan del modelo propuesto, la mayoría de los datos tiende a comportarse homogéneamente respecto de los residuales, y que el comportamiento de los mismos debe aproximarse a 0. En este comportamiento no se debe encontrar algún tipo de patrón o tendencia de datos, por lo cual se cumpliría con la condición.

Figura 20

Residuales vs índices de observación de puntajes de razonamiento cuantitativo



En las Figuras 21 y 22 se puede observar que existen observaciones influyentes que implicarían tomar decisiones sobre los datos analizados, o bien extender los alcances del modelo o bien estimarlo sin esas observaciones influyentes. Es necesario señalar que son pocas las observaciones que se escapan del modelo, sin embargo, este no alcanza los criterios de calidad que se esperan para un modelo de regresión beta.

Figura 21

Valores de leverage vs valores de predicción en puntajes de razonamiento cuantitativo

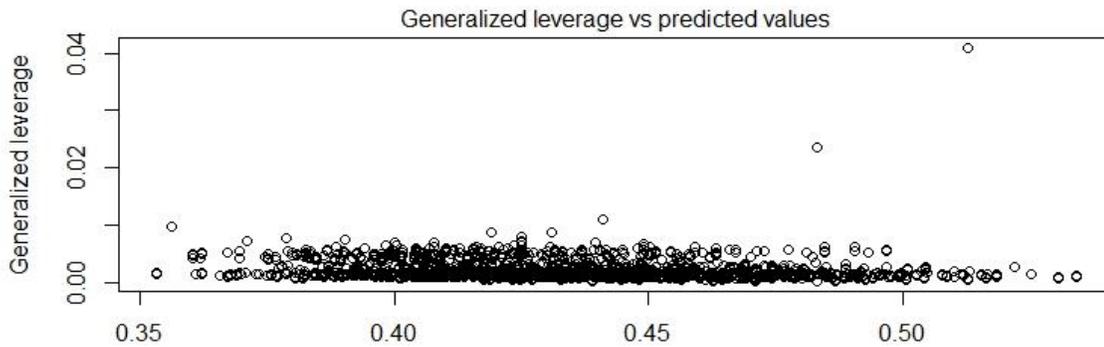
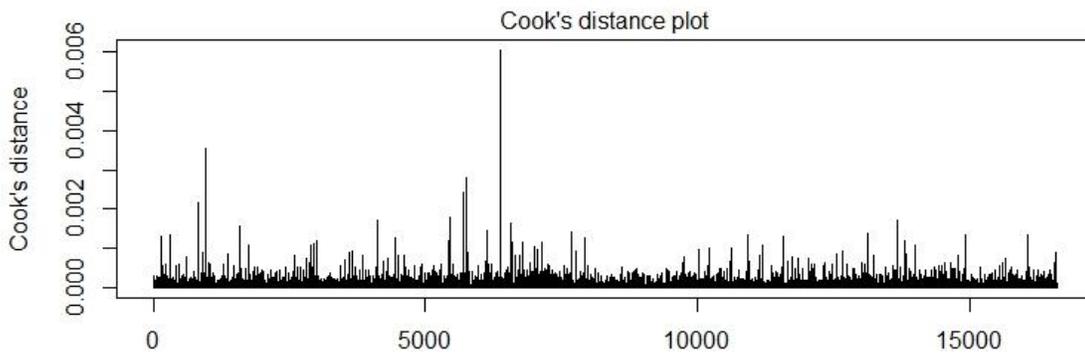


Figura 22

Distancias de Cook asociados al modelo de puntajes de razonamiento cuantitativo



Al observar la información provista por la Figura 22 se puede indicar que las observaciones influyentes se ubican en la primera parte de la base de datos, lo cual confirma la suposición de que el análisis podría ser más efectivo si se segmentara la base de datos y se analizara por segmentos, de modo que se obtuvieran estimaciones con alcance más restringido, pero quizás con indicadores de precisión mejores.

Resultados asociados a la predicción del puntaje de competencias ciudadanas

Los resultados de la estimación de la regresión beta incluyendo las variables sociodemográficas cualitativas utilizadas en el análisis de interacción con las variables puntaje de competencias ciudadanas se muestran en la Figura 23, en la cual se puede evidenciar que se ejecutó una regresión beta, observando el cumplimiento de las primeras dos condiciones, dado que el valor medio debe estar entre 0 y 1, y en este caso el valor estimado se encuentra alrededor de 0.0463. Por otro lado, el valor del parámetro phi debe ser mayor que 0, y en este caso, el estimado de phi es igual a 26.11. Ferrari & Cribari-Neto (2004) indican que el valor phi es mucho mejor entre tanto sea más grande, dado que tiene un efecto sobre el control de la varianza de las puntuaciones. El valor resultante es pequeño, por lo tanto el modelo, aunque resulta significativo según el indicador p-valor sea cercano a 0, indica que la varianza de las puntuaciones no está controlada, factor que podría afectar al modelo estimado.

Se puede especular que el valor phi pueda verse afectado por la cantidad de variables regresoras independientes que se incluyeron dentro del modelo, igual que se planteó anteriormente con el modelo propuesto para la variable razonamiento cuantitativo. Puede influir el hecho de que las variables cualitativas se incluyeron sin tener en cuenta que algunas tienen múltiples alternativas de respuesta, y otras son dicotómicas, factor que pudo haber afectado la varianza. Al observar los resultados se encuentra que el modelo no es significativo para las variables región Caribe y estado civil separado y/o viudo. En términos de los indicadores de calidad del modelo se debe tener en cuenta el valor del indicador Pseudo R-squared que, según Ferrari & Cribari-Neto (2004), debe oscilar entre 0 y 1, se encuentra en 0.1028, un valor menor comparado con la escala, lo cual implica un déficit en los indicadores de precisión del modelo, aún cuando el valor presenta un p-valor significativo.

Concretamente se podría expresar la interacción de las variables a partir de la siguiente expresión:

$$g(\mu t) = -\beta_1 + \beta_2 x t_2 + \beta_3 x t_3 + \beta_4 x t_4 + \beta_5 x t_5 + \beta_6 x t_6 - \beta_7 x t_7 - \beta_8 x t_8 - \beta_9 x t_9 - \beta_{10} x t_{10} - \beta_{11} x t_{11} - \beta_{12} x t_{12} + \beta_{13} x t_{13} + \beta_{14} x t_{14} + \beta_{15} x t_{15}$$

Figura 23

Estimación del modelo de regresión beta para puntaje de razonamiento cuantitativo

```

Standardized weighted residuals 2:
      Min      1Q  Median      3Q      Max
-3.9817 -0.5831  0.0463  0.6305 14.4495

Coefficients (mean model with logit link):
              Estimate Std. Error z value Pr(>|z|)
(Intercept)    -0.296072   0.027919 -10.605 < 2e-16 ***
GENEROM         0.107828   0.007970  13.528 < 2e-16 ***
PROCEDENCIACabecera Municipal  0.087671   0.009340   9.386 < 2e-16 ***
REGIONANDINA   0.199446   0.024058   8.290 < 2e-16 ***
REGIONCARIBE   0.041004   0.025085   1.635 0.102134
REGIONORINOQUIA 0.104096   0.028699   3.627 0.000287 ***
REGIONPACIFICA 0.137428   0.025197   5.454 4.92e-08 ***
ESTADOCIVILseparado y/o viudo -0.006827   0.027818  -0.245 0.806140
ESTADOCIVILsoltero -0.022730   0.010608  -2.143 0.032130 *
ESTADOCIVILUnión libre -0.075372   0.013356  -5.644 1.67e-08 ***
TRABAJOSEMANALEntre 11 y 20 horas -0.130761   0.010050 -13.011 < 2e-16 ***
TRABAJOSEMANALEntre 21 y 30 horas -0.081486   0.010444  -7.802 6.08e-15 ***
TRABAJOSEMANALMás de 30 horas -0.057567   0.009375  -6.141 8.21e-10 ***
TRABAJOSEMANALMenos de 10 horas -0.136182   0.010734 -12.687 < 2e-16 ***
MODALIDADPROGRAMADISTANCIA VIRTUAL 0.035890   0.013043   2.752 0.005928 **
MODALIDADPROGRAMAPRESENCIAL  0.208721   0.007576  27.550 < 2e-16 ***
INST_ORIGENOFICIAL  0.087344   0.008289  10.537 < 2e-16 ***

Phi coefficients (precision model with identity link):
              Estimate Std. Error z value Pr(>|z|)
(phi)  26.1166      0.2812   92.86 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Type of estimator: ML (maximum likelihood)
Log-likelihood: 1.542e+04 on 18 Df
Pseudo R-squared: 0.1028
Number of iterations: 24 (BFGS) + 2 (Fisher scoring)

```

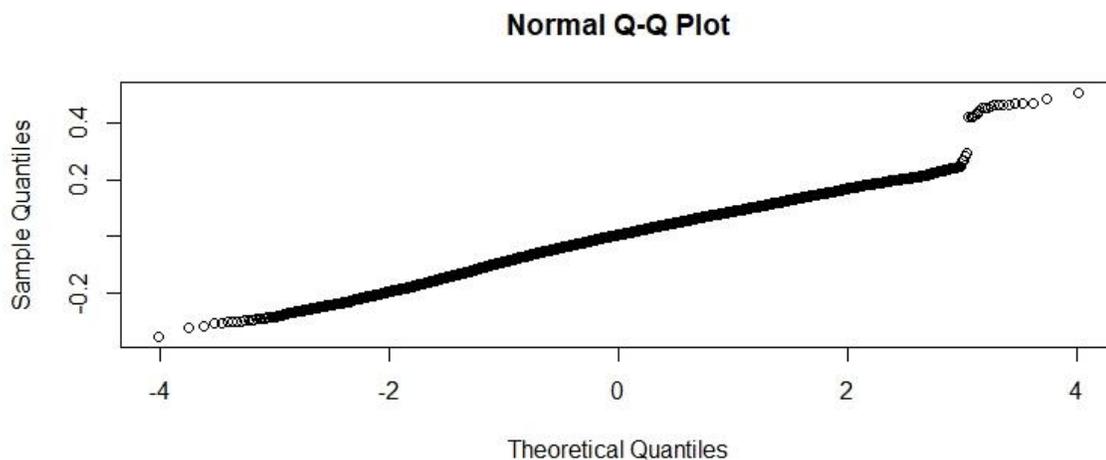
Se puede indicar que el modelo propuesto para describir el cambio de las puntuaciones de competencias ciudadanas implica que existe una relación negativa entre la variable dependiente y ser soltero o vivir en unión libre, e invertir tiempo a la semana en trabajar, Se observa que el modelo indica una relación positiva entre la variable respuesta y la modalidad de género masculino, pertenecer a la cabecera municipal, pertenecer a la región Andina, Orinoquía o

Pacífica, pertenecer a la modalidad presencial y distancia-virtual, y estar estudiando en una entidad pública.

Utilizando el resultado del qqplot presentado en la Figura 24 se puede observar que una buena cantidad de datos tienden a ajustarse a la distribución señalada por la recta teórica, sin embargo, se presentan una mayor cantidad de observaciones que se escapan de la explicación del modelo beta y esto se evidencia en las colas superior e inferior del esquema que implican claramente distribuciones asimétricas de los datos. Dichas observaciones podrían constituir un comportamiento no explicado por el modelo, razón por la cual los indicadores de precisión/calidad del modelo se hubiesen visto afectados.

Figura 24

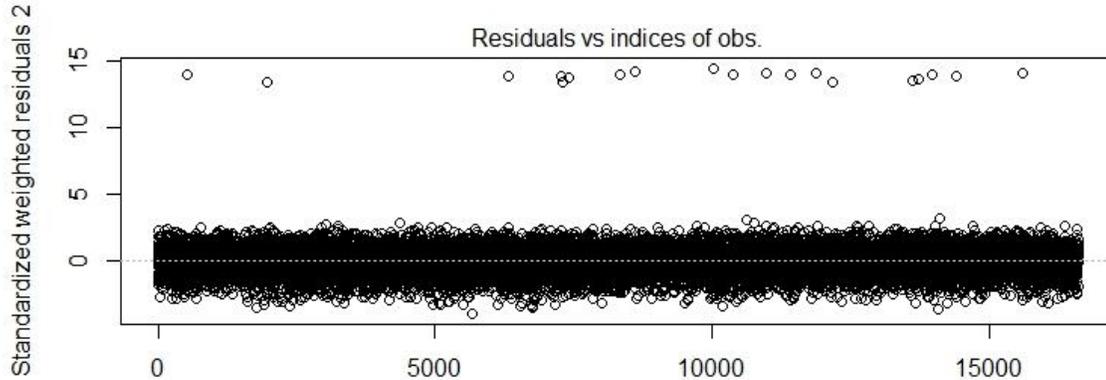
Regresión Beta para modelar interacciones con la variable razonamiento cuantitativo



Como se mencionó previamente la mayor asimetría y discontinuidad en el patrón de regresión se observa en la disposición de los mayores valores, aunque una porción considerable de datos se ajusta a la regresión propuesta.

Figura 25

Comportamiento de los residuales de puntajes de competencias ciudadanas



El resultado mostrado en la Figura 25, indica que el patrón general de datos se ajusta a valores similares a 0, condición que deben respetar los residuales del modelo, sin embargo, se presentan una mayor cantidad de valores no agrupados por el modelo teórico propuesto comparado con el resultado obtenido en la competencia de razonamiento cuantitativo. Aun cuando se evidencian mayor presencia de valores no explicados por el modelo, no se evidencian patrones o tendencias de los residuales.

En las Figuras 26 y 27 se puede detectar que existen una cantidad considerable de observaciones influyentes, comparado con el modelo establecido para razonamiento cuantitativo. Este hecho implica que se deben tomar decisiones sobre los datos analizados, o bien extender los alcances del modelo, realizar la estimación por segmentos, o bien estimarlo sin esas observaciones influyentes. Quizás las observaciones que se escapan del modelo puedan estar afectando los criterios de calidad del modelo de regresión beta.

Figura 26

Resultados de leverage del modelo estimado para competencias ciudadanas

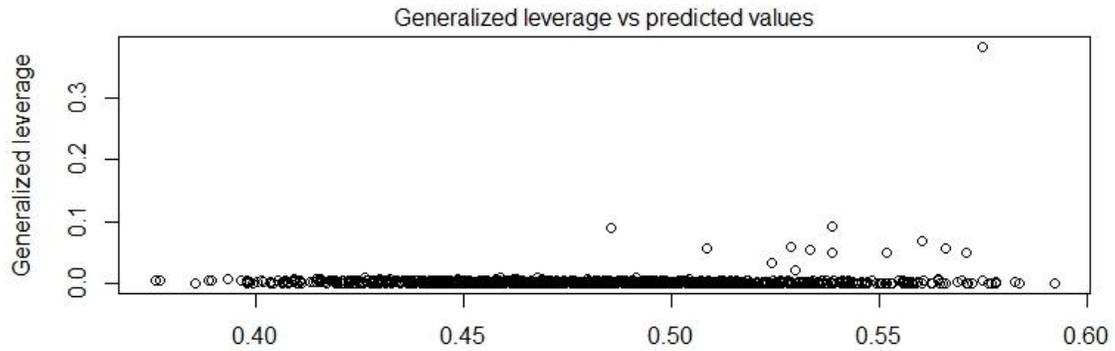
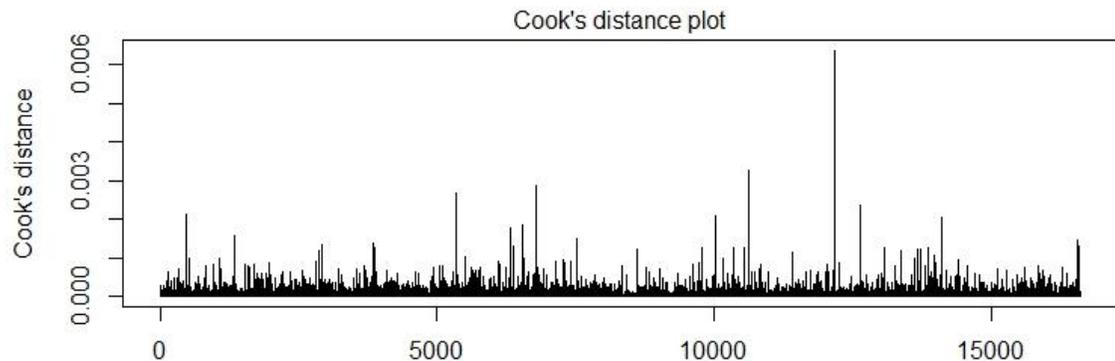


Figura 27

Distancias de Cook del modelo estimado para competencias ciudadanas



Al observar la Figura 27 se puede indicar que las observaciones influyentes se ubican a lo largo de toda la base de datos, hecho que dificulta establecer alguna decisión sobre el ajuste o abordaje de la información. Esta situación podría ameritar realizar análisis segmentando la base de datos, de modo que se obtuvieran estimaciones con alcance más restringido, pero quizás con indicadores de precisión mejores.

CONCLUSIONES

- Se evidencia que los porcentajes de varianza explicada en los análisis de correspondencia tienden a disminuir entre más variables categóricas se sumaron al análisis de interacciones.
- Las regiones Amazonía y Orinoquía no estaban bien representadas en los diagramas de correspondencias múltiples, por tanto, fue difícil describir su comportamiento.
- Se evidencia que las competencias tienden a asociarse de acuerdo a aspectos de comprensión lectora (competencias ciudadanas, razonamiento cuantitativo, lectura crítica, e inglés) comparados con la competencia procedimental de escritura de textos.
- Se observa coincidencia con el estudio de Rodríguez-Jiménez (2019) en cuanto a las variables socioeconómicas utilizadas en este estudio. Puntualmente, estrato, género, procedencia, región y tipo de institución.
- Se observa que el género tiende a mostrar resultados diferenciadores en cuanto a las dos competencias, haciendo significativos los resultados de la predicción de los modelos solamente para el género masculino.
- El nivel de competencia 4 se comporta de manera independiente a las demás covariables incluidas en el estudio. En ambas competencias, este nivel no mostraba suficiente interacción con los cruces de variables estudiadas.

- La regresión beta se ajusta a este tipo de variables cuantitativas, dado que tiene dos ventajas: la primera es que funciona con valores mayores que 0; y la segunda es que permite tener en cuenta variables categóricas.

- Se observan puntos de alto leverage en ambas variables, lo cual indica que hay datos con observaciones influyentes que implicarían tomar decisiones con los datos tales como estudiar el modelo sustrayendo dichas observaciones para verificar si el grado de ajuste del modelo aumenta.

- Los indicadores pseudo r squared son bajos aunque los modelos resultan significativos, lo cual indica que se debe buscar una alternativa más apropiada para modelar los datos. En el ejercicio se hicieron pruebas con otros tipos de modelos (lm,poisson,glm,cuantílica), sin embargo, sus resultados no fueron apropiados para el propósito de investigación. Se podría implementar una segmentación de la variable continua para modelar por intervalos y obtener un mejor resultado.

REFERENCIAS BIBLIOGRÁFICAS

- Algañaraz, V. H. (2016). El análisis de correspondencias múltiples como herramienta metodológica de síntesis teórica y empírica. Su aporte al estudio del locus universitario privado argentino (1955-1983). *Relmecs*, 6, (1), 1-17. ISSN 1853-7863.
- ASCOFAPSI (2018). Compendio de competencias de los psicólogos. Publicaciones ASCOFAPSI.
- Decreto 3963 de 2009. Por el cual se reglamenta el examen de estado de calidad de la educación superior. 14 de octubre de 2009. D.O.
- Ferrari, S. & Cribari-Neto, F (2004) Beta Regression for Modelling Rates and Proportions, *Journal of Applied Statistics*, 31(7), 799-815, DOI: 10.1080/0266476042000214501.
- Hernández, R., Fernández, C., Baptista, M. P. (2014). Metodología de la investigación. Sexta Edición. McGraw-Hill.
- Instituto Colombiano para la Evaluación de la Educación [ICFES], (2020). Guía de orientación Saber Pro 2020: módulos de competencias genéricas.
<https://www.icfes.gov.co/documents/20143/1891934/Guia+de+orientacion+de+Modulos+genericos+Saber+Pro-2020.pdf/a0f24d6f-d82e-cf94-0c1b-2c6fb3acddb0?t=1597776761084>
- León, O. & Montero, I (2003). Métodos de investigación en psicología y educación. McGraw-Hill.

Ley 1324 de 2009. Por la cual se fijan parámetros y criterios para organizar el sistema de evaluación de resultados de la calidad de la educación, se dictan normas para el fomento de una cultura de la evaluación, en procura de facilitar la inspección y vigilancia del Estado y se transforma el ICFES. 13 de julio de 2009. D.O.

Macías-Ruano, M(2015). Aplicación de la técnica de componentes principales para el análisis de la pobreza en Honduras. *Revista Ciencia y Tecnología*, 16 (1), 82-96.

Rodríguez-Jiménez, O., Cote-Peña, L., Malaver-Toloza, J., Pineda-Miranda, B., & Sarmiento-Fúquene, S (2019). Factores asociados al desempeño en la prueba saber-pro 2017: un estudio de los programas de psicología. *Universidad Nacional de Colombia-Sede Bogotá*.

Suárez-Rodríguez, M. A. (2018). Las competencias en la educación superior, perspectiva actual desde los programas de psicología en Colombia. Publicaciones ASCOFAPSI.

Tobón, S (2006). Competencias, calidad y educación superior. Editorial magisterio.

Tobón, S (2008). La formación basada en competencias en la educación superior: el enfoque complejo. Editorial Instituto CIFE.

Valbuena, C., & García G (2016). Factores asociados al desempeño de los estudiantes de psicología en Colombia en las pruebas Saber-Pro. Observatorio para la calidad de la educación superior en psicología - ASCOFAPSI