

## El Big Data desde la perspectiva de una muestra de estudiantes de universidades del norte de Perú

### Big Data from the Perspective of a Sample of Students from Universities in Northern Peru

**Antonio Matas Terrón** 

Universidad de Málaga, Málaga, España  
ORCID: <https://orcid.org/0000-0003-1401-4932>

**Juan José Leiva Olivencia** 

Universidad de Málaga, Málaga, España  
ORCID: <https://orcid.org/0000-0002-2857-8141>

**Nemecio Núñez Rojas\*** 

Universidad Alas Peruanas, Lima, Perú  
ORCID: <https://orcid.org/0000-0002-2376-1051>

**Recibido 09-08-19 Revisado 10-10-19 Aprobado 28-11-19 En línea 03-12-19**

#### \*Correspondencia

Email: [nnunezrojas@hotmail.com](mailto:nnunezrojas@hotmail.com)

#### Citar como:

Matas, A., Leiva, J.J. & Núñez, N. (2020). El Big Data desde la perspectiva de una muestra de estudiantes de universidades del norte de Perú. *Propósitos y Representaciones*, 8(1), e376. doi: <http://dx.doi.org/10.20511/pyr2020.v8n1.376>

## Resumen

El Big Data es una realidad que se está incorporando de forma paulatina y con fuerza en todos los campos de la sociedad, entre ellos en Educación, aunque de forma tardía. El objetivo de este estudio es analizar descriptivamente la perspectiva general que sobre el Big Data aplicado a la Educación tiene una muestra de 104 estudiantes universitarios del norte del Perú. La muestra participante fue no probabilística de carácter incidental. Se llevó a cabo un estudio de encuesta, utilizando la escala VABIDAE para recoger los datos. Se hizo un estudio psicométrico de la escala con los datos recogidos, así como un análisis estadístico descriptivo. Los resultados muestran una percepción positiva de los participantes ante el Big Data así como una predisposición evidente a utilizar los servicios de Big Data en su ejercicio profesional como fuente de información. Igualmente, destaca la ausencia de brecha de género. Finalmente se comentan su posible impacto en la formación del docente.

**Palabras Clave:** Big Data; Enseñanza superior; Docencia; Brecha digital; Género; Psicometría.

## Summary

Big Data is a reality that is gradually being incorporated and strongly in all fields of society, including in Education, although late. The objective of this study is to descriptively analyze the general perspective that a sample of 104 university students of Education of Peru has on the Big Data applied to Education. The participating sample was non-probabilistic incidentally. A survey study was conducted, using the VABIDAE scale to collect the data. A psychometric study of the scale was done with the collected data, as well as a descriptive statistical analysis. The results show a positive perception of the participants before Big Data as well as an obvious predisposition to use Big Data services in their professional practice as a source of information. It also highlights the absence of a gender gap. Finally, its possible impact on teacher training is discussed.

**Keywords:** Big Data; Higher Education; Teaching; Digital Gap; Gender; Psychometry.

## Introducción

En los últimos años ha emergido con fuerza el término de Big Data en distintos contextos y ámbitos profesionales, sociales, culturales y económicos. En pocas palabras, podemos definir el Big Data como un constructo que abarca múltiples posibilidades de generación, reunión, gestión y evaluación de datos e informaciones masivas creadas en distintas esferas de la comunicación y la acción humana (Bogarín, Morales, & Menéndez, 2015). Ni que decir sobre los avances tecnológicos y los rápidos procesos de digitalización de la vida que han supuesto un incremento abrumador de datos que circulan por múltiples canales, fuentes y autopistas de la información (García-Aretio, 2017; Lawal, Zakari, Shuaibu & Bala 2016; Williamson & Piattoeva, 2019). Cabe aquí subrayar el papel tan relevante que han adquirido en los últimos tiempos las redes sociales, así como los canales de comunicación y plataformas telemáticas que han redefinido gran parte del comportamiento humano en prácticamente todas las áreas de desarrollo económico, social, político, cultural y, también, educativo.

Hoy día los centros formativos se encuentran con el reto ineludible de operativizar, gestionar y de emplear de forma inteligente y responsable el Big Data que se genera de forma incesante a partir de las producciones, las comunicaciones, las interacciones y todo el cúmulo de información acerca del rendimiento y los resultados escolares. En este sentido, el fenómeno del Big Data en la educación actual es una constatación de una realidad que ha cambiado con fuerza en las últimas décadas, posibilitando nuevos procesos de interacción y de aprendizaje (Dussel & Reyes, 2018). Ahora bien, los últimos años están demostrando la urgente necesidad de clarificar

las potencialidades y dificultades a la hora de aplicar el Big Data desde parámetros y criterios pedagógicos.

El Big Data en el mundo educativo se refiere, sobre todo, a la ingente cantidad de datos e información que tenemos sobre el propio sistema educativo, el rendimiento académico de los estudiantes (Banerjee, 2002; Williamson, 2017), las necesidades formativas de los docentes, o, el impacto de las medidas educativas que se toman en el ámbito formativo para mejorar la calidad de los procesos educativos en tiempos de globalización, interculturalidad y plurilingüismo (González-Fernández & Maroto, 2017; Sobe, 2018). Así, una de las claves configuradoras del Big Data en el contexto académico se vincula con las analíticas de aprendizaje que se entienden como procesos de valoración continua y procesual que aborda los cambios y la evolución de determinados factores o variables escolares. Así, las potencialidades del Big Data se pueden concretar en los siguientes aspectos (Daniel, 2019; Tejada-Escobar, Murrieta-Marcillo, Villao-Santos & Garzón-Balcázar, 2018; Williamson, 2017):

- Facilidad en el acceso a la información de los distintos agentes educativos, sus producciones, proyecciones y resultados.
- Personalización de las respuestas educativas atendiendo a las singularidades y evolución en sus estilos y/o características psicológicas y educativas.
- Desarrollo de protocolos y pautas de actuación estandarizada en el ámbito pedagógico.
- Aplicación de respuestas personalizadas de aprendizaje en el caso de alumnado con necesidades específicas de apoyo educativo.
- Ajuste de propuestas formativas contextualizadas a las demandas formativas del profesorado y las comunidades educativas.
- Adopción de medidas de prevención y anticipación de dificultades de aprendizaje.
- Gestión operativa de decisiones en el ámbito del liderazgo pedagógico
- Impulso de la creación de recursos adaptados a las necesidades personales de aprendizaje del alumnado.
- Identificación predictiva de problemáticas de aprendizaje y rendimiento académico.
- Aumento del aprendizaje autorregulado y el control en los procesos de diseño, desarrollo y evaluación del currículo.

Las dificultades de la aplicación del Big Data en los escenarios formativos no se relacionan con su propia naturaleza o estructura dinámica, sino en el sentido o carácter pedagógico que pueda impregnar la toma de decisiones. Así, la rapidez de los datos y la diversidad de estos pueden generar disfunciones y, sobre todo, una incapacidad o disfunción significativa si los agentes educativos se sienten abrumados por la cantidad de datos recopilados, o, peor aún, el alto grado de tecnificación o burocratización que puede provocar desánimo o desmotivación entre los docentes y directivos educativos.

Además, uno de los riesgos es la propia situación de inviabilidad en cuanto a analítica de aprendizajes si el profesorado y otros agentes educativos no están formados en la gestión del Big Data en educación. Igualmente, no podemos obviar la sobrecarga de trabajo académico y las dificultades derivadas de una inadecuada planificación o utilización de los tiempos pueden acrecentar la desidia o la no operatividad en la toma de decisiones en la mejora de la calidad docente. Esto debe suponer, necesariamente, la incorporación de nuevos agentes especializados en la gestión masiva de datos y en analítica de aprendizajes con formación digital, pedagógica y organizacional avanzada (Ratner, Andersen, & Madsen, 2019; Williamson, 2016).

Tampoco podemos obviar la existencia de una escasa producción bibliográfica y de corte experimental en torno al Big Data en educación, lo que es un hándicap que será corregido en los próximos años y décadas dada la trascendencia y alcance social, la proyección científica y el impacto educativo real en los sistemas educativos, y, por ende, en los centros escolares y en la

propia configuración de respuestas curriculares y didácticas en múltiples situaciones formativas.

Este trabajo pretende arrojar luz sobre un tema complejo y en permanente construcción, en tanto resulta necesario transitar de forma operativa de ocurrencias a evidencias científicas que permitan una aproximación interdisciplinar, multifacética, globalizada e inclusiva de los profesionales de la educación al Big Data en los escenarios escolares. En este sentido, el objetivo principal de esta investigación es describir la perspectiva general que sobre el Big Data aplicado a la Educación tiene una muestra de estudiantes universitarios y describir qué retos y oportunidades consideran más importantes. En este sentido, se ha propuesto también identificar posibles diferencias entre género en la medida que el género es una variable sociodemográfica que, por razones socioculturales, puede mediar la actitud y opiniones de la ciudadanía con relación a distintos temas socioeducativos (Colás & Jiménez, 2006).

Puesto que se utiliza una escala que está en proceso de validación (ver epígrafe siguiente) otro de los objetivos propuestos fue comprobar las propiedades psicométricas del instrumento con los datos de la muestra.

## **Método**

### **Participantes**

En el estudio participó una muestra de 104 personas, alumnado de cuatro universidades del Norte de Perú, tanto de carácter público (dos de ellas) como privado: Universidad uno (27 participantes); Universidad dos, de carácter público (15 participantes); Universidad tres (50 participantes); y Universidad cuatro, pública (12). La muestra se construyó de forma incidental a partir de la colaboración de varios docentes que solicitaron la participación de su alumnado.

La media de edad fue de 26.30 años (SE.=1.184) siendo la mediana de 21 años, y una desviación típica de 11.90 oscilando entre los 16 y los 67 años. La edad presentó un sesgo positivo de 1.530 puntos. El 54.5% de los participantes era mujer frente al 45.5% de hombres.

### **Diseño, instrumento y procedimiento**

El estudio se desarrolló siguiendo un diseño de encuesta o diseño survey. Para ello, se utilizó la escala de Valoración del Big Data Aplicado a la Educación (VABIDAE) de Borrego, Guardado, Jiménez, Montero, Negro y Matas (2019) que se encuentra en proceso de validación. Esta escala cuenta con 31 ítems organizados en tres bloques: valoración de aspectos positivos, valoración de aspectos negativos, y emociones que induce el Big Data. Cada ítem se acompaña de una escala de apreciación con cinco opciones que varían según el bloque. Así, para bloques de aspectos positivos y negativos la escala es 1 No en absoluto, 2 Creo que no, 3 No lo sé, 4 Creo que sí y 5 Totalmente de acuerdo. Para los ítems de emociones la escala es 1 Nada en absoluto, 2 Casi nada, 3 No lo sé / me resulta indiferente, 4 Algo, 5 Totalmente. En el instrumento se completa con una serie de preguntas sociodemográficas (edad, género, residencia, Universidad, etc.).

Los datos se recogieron en el mes de marzo de 2019. Para ello se solicitó la participación de distintos docentes de universidades del norte del Perú a través de un correo electrónico. Los docentes colaboradores comunicaban los días y horas en que iban a estar con su alumnado en disposición de contestar a la escala. En la fecha y hora indicada se abría la aplicación para que el alumnado pudiese contestar. Concluido el tiempo establecido se cerraba el acceso al instrumento. La aplicación permanecía cerrada el resto del tiempo, impidiendo accesos no controlados.

Con la intención de satisfacer el objetivo del estudio y profundizar en las características

métricas de la escala, tal como se indicó anteriormente, se aplicó un análisis factorial exploratorio, así como un análisis de la fiabilidad a través del coeficiente alpha de Cronbach. Para los análisis se ha utilizado el paquete estadístico SPSS versión 24.

## Resultados

### Fiabilidad y Validez

Se realizó un análisis de la consistencia interna de la escala con todos los ítems a través del procedimiento de Cronbach obteniendo una puntuación de 0.908. Solamente eliminando el ítem “El tema me divierte” el alfa mejoraba a 0.910. En función de la escasa mejora se mantuvieron todos los ítems para el análisis de la estructura latente de los datos.

Para obtener evidencias de la estructura interna o latente de los datos, se llevó a cabo un análisis factorial exploratorio. Inicialmente se comprobaron los supuestos básicos. El determinante de la matriz de correlación era prácticamente cero ( $D= 5,072e11$ ). Por su parte, la medida de adecuación de muestreo de Keiser-Meyer-Olkin fue aceptable (0.85) siendo la prueba de esfericidad de Bartlett significativa (Chi-cuadrado= 2105.774; g.l.= 465;  $p < .0001$ ). Esto sugiere que el análisis factorial es apropiado para estos datos.

Posteriormente se llevó a cabo el AFE por el análisis de componentes principales a través del procedimiento de extracción de máxima verosimilitud. La prueba sugirió seis factores con una varianza explicada total del 62,372% de la varianza después de la extracción (Chi-cuadrado= 338.772; g.l.=294;  $p=.037$ ).

Para facilitar la interpretación, se aplicó una rotación Varimax con normalización Kaiser que convergió en 7 iteraciones. En la tabla se puede ver la carga de los ítems en la matriz de factores rotados. Se ha utilizado un criterio conservador, considerado pesos inferiores a .40 como bajos (Hair, Andrsen, Tatham, & Black, 1998).

**Tabla 1.**  
*Matriz de factores rotados*

	1 Oportunidades	2 Emociones negativas	3 Retos al control	4 Retos a la persona	5 Emociones positivas	6 Empleabilidad
Mejorar la organización de los centros escolares	.836					
Atender mejor las necesidades del alumnado	.834					
Promover la calidad educativa en general	.787					
Mejorar la selección de docentes	.780					
Ayuda a prevenir el fracaso escolar	.764					
Mejorar los resultados académicos	.762					
Personalizar la Educación	.729					
Producir recursos educativos adaptados al alumnado	.702					
Evita el plagio	.689					
Mejorar la empleabilidad	.670					.401
Facilitar la toma de decisiones a nivel político	.491					
Siento vergüenza		.847				
Me hace sentir enojado		.828				
Siento culpa		.827				
Me genera aburrimiento		.688				
Me siento impotente		.665				
Me provoca ansiedad		.655				
Control del sistema educativo por parte de gobiernos			.823			
Manipulación del sistema			.782			
Control del sistema educativo por parte de empresas			.738			
Aumento de poder de políticos			.641			
Aumento del poder de gestores de centros			.541			
Pérdida de privacidad del alumnado				.838		
Pérdida de privacidad del docente				.823		
Ataques informáticos				.583		
Pérdida de socialización propia de la escuela				.564		
Pérdida de funciones del docente			.407	.498		
Me provoca esperanza	.508				.725	
Me hace sentir orgulloso	.500				.593	
El tema me divierte					.540	
Me provoca alivio					.451	
Porcentaje de varianza explicada:	22.387	11.774	10.965	9.204	6.022	2.020

Nota: n=104

**Propósitos y Representaciones**

En base al contenido de los ítems se ha nombrado cada uno de los factores tal como se observa en la primera fila de la tabla 1. En dicha tabla se muestran también los pesos de cada ítem en los factores, así como la varianza explicada por cada factor. Cabe destacar que el sexto factor podría no considerarse en el modelo en función de dos aspectos: 1) solamente tiene un ítem con un peso superior a 0.400 que además satura previamente en el factor 1; 2) la varianza explicada es de sólo el 2.02%.

### Descriptivos

El cuestionario se inicia con el bloque de oportunidades. En general, los encuestados valoran positivamente las posibilidades del Big Data. En este sentido, el 92% de la muestra considera que sí o está totalmente de acuerdo con que se atienden mejor las necesidades del alumnado. Prácticamente el mismo patrón de respuesta se repite ante el ítem sobre el Big Data como recurso para mejorar los resultados académicos y como medio para personalizar la educación (ver tabla 2).

Ante el ítem sobre mejora de empleabilidad, el 21.8% no se posiciona al respecto, mientras que el 34.7% cree que sí la mejora, junto con otro 34.7% que está totalmente seguro de que la mejora.

El posicionamiento a favor de las oportunidades del Big Data se vuelve a repetir cuando se habla de evitar el plagio (38.6% totalmente de acuerdo), mejorar la organización de los centros educativos (48.5% totalmente de acuerdo), producir recursos adaptados al alumnado (55.4% totalmente de acuerdo), promover la calidad educativa (52.8% totalmente de acuerdo), o prevenir el fracaso escolar (52.5% totalmente de acuerdo) tal como puede observarse en la tabla 2.

Dentro de este panorama de unanimidad, destacan solamente tres ítems donde la heterogeneidad aumenta levemente. Así cuando se habla del proceso de selección de docentes, el 15.8% de la muestra elige la opción “No lo sé”. Aparece un posicionamiento contrario cuando el 20.8% afirma que no lo sabe, y el 10.9% se manifiesta en contra de que el Big Data facilite la toma de decisiones a nivel político.

**Tabla 2.**  
*Porcentajes de respuesta para el bloque de oportunidades*

Ítem	1 No en absoluto	2 Creo que no	3 No lo sé	4 Creo que sí	5 Totalmente de acuerdo
Atender mejor las necesidades del alumnado	2	4	2	45.5	45.5
Mejorar los resultados académicos	3	4	3	37.5	52.5
Personalizar la educación	3	4	3	44.5	44.6
Mejorar la empleabilidad	2	6.9	21.7	34.7	34.7
Evitar el plagio	2	12.9	10.9	35.6	35.6
Mejorar la organización de los centros escolares	3	4	8.9	35.6	48.5
Mejorar la selección de los docentes	3	7.9	15.8	28.7	44.6
Producir recursos educativos adaptados al alumnado	0	6.9	5.9	31.7	55.4
Facilitar la toma de decisiones a nivel político	5	10.9	20.8	36.6	26.7
Promover la calidad educativa en general	2	4	3	38.5	52.5
Ayuda a prevenir el fracaso escolar	3	4	6.9	33.6	52.5

Nota: n=104

Con relación a los posibles problemas o retos asociados al Big Data, los participantes se manifiestan indecisos ante el ítem “Pérdida de privacidad del alumnado”. Por su parte, el 44.6% se manifiesta en contra de que el Big Data suponga una amenaza a la privacidad del docente, frente a un 37.5% que estima que si lo puede ser. Un patrón también heterogéneo aparece cuando se les pregunta por la pérdida de la socialización propia de la escuela. En este caso, e 31.7% piensa que no ocurrirá, un 22.8% está indeciso, y un 45.5% cree que sí ocurrirá o está totalmente de acuerdo con la afirmación.

Respecto a los ataques informáticos, la opinión no es contundente. El 20.8% cree que no, el 38.6% cree que sí y un 18.8% se manifiesta claramente indecisa. Este patrón de heterogeneidad alrededor de la opción de indecisión también se presenta en el ítem “Aumento de poder de gestores de centros” (ver tabla 3).

Por su parte, los ítems “Pérdida de funciones del docente”, “Aumento del poder de los políticos”, “Manipulación del sistema”, “Control del sistema educativo por parte de los gobiernos”, y “Control del sistema educativo por parte de empresas” presentan un alto grado de desacuerdo, con un alto porcentaje de los encuestados eligiendo la opción 3 de indecisión o las cercanas, tanto la opción 2 de creo que no, como la 4 de creo que sí tal como se puede ver en la tabla 3.

**Tabla 3.**  
*Porcentajes de respuesta para el bloque de retos*

Ítem	1 No en absoluto	2 Creo que no	3 No lo sé	4 Creo que sí	5 Totalmente de acuerdo
Pérdida de privacidad del alumnado	8.9	24.8	12.9	34.7	18.7
Pérdida de privacidad del docente	10.9	33.7	17.8	25.7	11.9
Pérdida de la socialización propia de la escuela	10.9	20.8	22.8	29.7	15.8
Ataques informáticos	6.9	20.8	18.8	38.6	14.9
Pérdida de funciones del docente	14.9	29.7	16.8	23.8	14.8
Aumento del poder de gestores de centros	5.9	18.8	31.7	31.7	10.9
Aumento del poder de políticos	9.9	14.9	33.7	30.6	10.9
Manipulación del sistema	8.9	15.8	14.9	38.6	21.8
Control del sistema educativo por parte de gobiernos	6.9	14.9	15.8	41.6	20.8
Control del sistema educativo por parte de empresas	5.9	16.8	21.8	35.7	19.8

Nota: n=104

Con relación a las distintas emociones que provoca el tema del Big Data en Educación destaca que un posicionamiento en contra de aspectos negativos, tales como sentirse enojado, ansioso o culpable, así como un posicionamiento claramente a favor de sentirse esperanzado o incluso orgulloso, tal como se puede ver en la tabla 4.

En el caso de “El tema me divierte” presenta un grupo de participantes totalmente identificados o algo identificados con el ítem (76.2%) y un pequeño grupo que no se identifica en absoluto (14.9%). En el resto de ítems la homogeneidad de las respuestas es la tónica. Así, la mayoría afirma que le provoca esperanza y orgullo. En el caso de alivio, el 43.6% eligen la opción “algo” frente a un 15.8% que se identifica con “Nada”.



Con relación a las emociones negativas, la gran mayoría se posicionan en contra como se ha apuntado anteriormente. Así, sólo el 6% afirma estar algo o totalmente enojado sobre el tema, el 19.9% dice sentirse algo o totalmente ansioso ante el Big Data en Educación, frente a un 51.5% que no le despierta nada de ansiedad. Igualmente, sólo el 10.9% afirma sentir culpa, el 21.8% experimentar impotencia o el 8.9% aburrimiento.

**Tabla 4.**  
*Porcentajes de respuesta para el bloque de emociones*

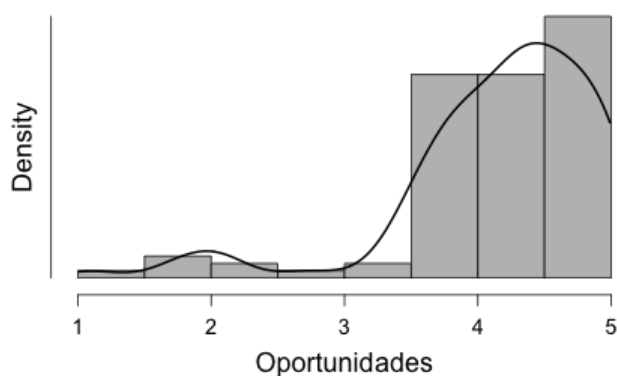
Ítem	1 Nada en absoluto	2 Casi nada	3 No lo sé / indiferente	4 Algo	5 Totalmente
El tema me divierte	14.9	3	5.9	35.6	40.6
Me provoca esperanza	1	6.9	5.9	44.6	41.6
Me hace sentir orgulloso	6.9	9.9	15.8	35.7	31.7
Me hace sentir enojado	69.2	16.9	7.9	3	3
Me provoca ansiedad	51.4	18.8	9.9	14.9	5
Siento vergüenza	69.3	14.8	5	5	5.9
Siento culpa	69.3	12.9	6.9	5.9	5
Me provoca alivio	15.8	8.9	14.9	43.6	16.8
Me siento impotente	50.5	16.8	10.9	14.9	6.9
Me genera aburrimiento	63.4	20.8	6.9	5.9	3

Nota: n=104

Por último, los participantes contestaron a la pregunta “¿Usaría el Big Data en su profesión si fuera posible?” con una escala que numérica desde 1 (no en absoluto) a 5 (Si, por supuesto). La mayoría eligieron los valores 4 o 5 (78.2%) con una media de 4.129 puntos (E.M.= 0.09) y una desviación típica de 0.923 puntos, siendo el sesgo de -0.959 (E.Sesgo=0.240).

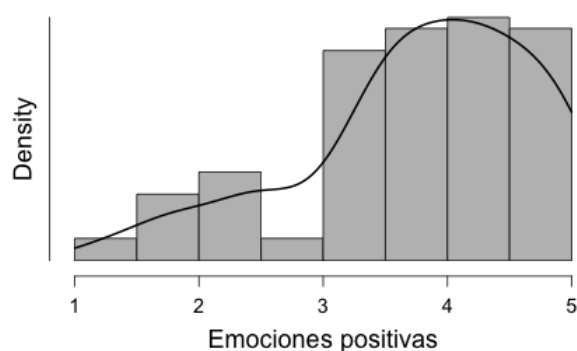
Finalmente, se ha realizado un análisis descriptivo de las puntuaciones generadas en el AFE. Para ello se calcularon los promedios en cada factor a partir de las puntuaciones de los ítems que los componen.

Así, la media del factor de oportunidades es de 4.157 (E.S.= 0.076) con una desviación típica de 0.768. Por su parte, los retos al control presentan un promedio de 3.382 (E.S.= 0.095) y una desviación típica de 0.957, mientras que los retos a la persona tienen un promedio de 3.141 (E.S.=0.098) con una desviación típica de 0.994. En ambos casos, los retos presentan una mayor heterogeneidad que las oportunidades. En la figura 1 se puede observar que además, el factor de oportunidades está sesgado negativamente.

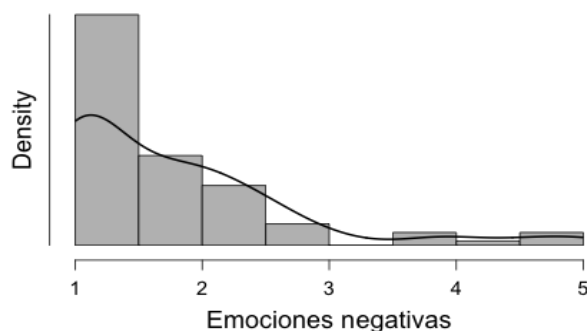


**Figura 1.** Histograma del factor Oportunidades

En relación con las emociones, los positivos tienen una media de 3.787 (E.S.=0.093) presentando una desviación estándar de 0.936, frente a las emociones negativas que tiene una media de 1.765 (E.S.= 0.092) con una desviación típica de 0.925. Ambas factores presentan distribuciones con un relativo alto sesgo, siendo de -0.828 para las emociones positivas y de 1.724 para las emociones negativas, tal como se aprecia en la figura 2 y en la figura 3.



**Figura 2.** Histograma del factor Emociones Positivas



**Figura 3.** Histograma del factor Emociones Negativas

Se ha obviado la presentación del sexto factor puesto que coincide con el ítem “Mejorar la empleabilidad” que se ha comentado anteriormente (ver tabla 2).

### Diferencias en función del sexo

Se analizaron las posibles diferencias entre sexo aplicando pruebas de contraste de hipótesis a los factores generados por el AFE en función del género como factor. Los resultados mostraron diferencias significativas entre hombres y mujeres solamente para las emociones negativas ( $F=9.359$ ;  $g.l.=1$ ;  $p=.003$ ) a un nivel de significación menor a 0.01 y con un tamaño del efecto  $\eta^2$  cuadrado de sólo 0.08.

Los hombres tienen una puntuación promedio superior (2.062) a las mujeres (1.518) siendo el grupo de mujeres más homogéneo en su respuesta (D.T.= 0.769) que el de hombres (D.T.= 1.015). En cualquier caso, estos resultados sugieren que no existen diferencias estadísticamente significativas entre hombres y mujeres de la muestra. Por otro lado, se llevó a cabo un estudio de las diferencias en función de la universidad y de la edad. En ambos casos los resultados no fueron estadísticamente significativos.

### Discusión

En primer lugar, cabe destacar que el instrumento utilizado presenta un buen ajuste, tanto a la naturaleza de los datos, como a la estructura teórica en la que se basa. En este sentido, la escala utilizada podría presentarse como un instrumento estandarizado. Esto resulta de especial interés para impulsar una proyección más potente en términos pedagógicos y también de índole comparativa en diversos contextos formativos.

De la misma forma, subrayamos que la escala llega a explicar un alto porcentaje de la varianza, lo que apoya aún más su validez y utilidad, según los criterios que marcan autores como Hair, Black, Babin y Andersen (2014).

En estudios posteriores se podría tratar de generalizar esta escala como instrumento para evaluar la experiencia de la persona con otras tecnologías e innovaciones tecnológicas, y no sólo con el Big Data. En verdad, esto se concibe como una oportunidad de traslación y de transferencia de aprendizajes que pueden ser sumamente sugerentes para posibilitar una amplia amalgama de posibilidades evaluativas y de innovación pedagógica.

Con relación a los resultados obtenidos, prácticamente todos los participantes afirmaron que no tenían un conocimiento previo de qué es Big Data, según el ítem que se incluye en el VABIDAE al inicio. Este desconocimiento es coherente con lo que apuntan otros autores e investigaciones consultadas (Daniel, 2019). Desconocimiento que implica también una falta de elementos para la reflexividad y concepción crítica, a priori, de sus potencialidades y complejidad desde un punto de vista holístico. Además, esta situación induce a plantearse la necesidad de una reflexión sobre cómo los universitarios son formados con relación a la realidad que les espera tras su formación, teniendo en cuenta la gran cantidad de información disponible en tiempo real, que requieren tener habilidades académicas para gestionarla en cuanto a tres criterios: volumen, variedad y velocidad. La vinculación Universidad – Mercado laboral es un tópico ampliamente debatido, que aquí se pone de manifiesto una vez más.

Por otro lado, la valoración general de los participantes es que el Big Data es una oportunidad para la mejora del proceso de aprendizaje, sobre todo para: atender las necesidades singulares del alumnado, mejorar sus resultados académicos y personalizar la educación. Esto es coherente con la idea de que la enseñanza personalizada mejora los resultados académicos, lo que,

por otra parte, está muy vinculado a cómo los recursos tecnológicos pueden ser un medio para la adaptación del proceso de enseñanza-aprendizaje (Lerís & Sein-Echaluce, 2011).

Igualmente, la escala apuntaba otras oportunidades cuyos resultados muestran una menor unanimidad de criterio. Así, aunque se considera que el Big Data puede ser útil para la empleabilidad, evitar el plagio, seleccionar a los docentes, o mejorar las decisiones sobre la educación, la muestra tiene una opinión más heterogénea. En cierto sentido, esto resulta razonable en la medida en que son cuestiones que influye, sobre todo, a quienes toman las decisiones sobre la organización del sistema formativo. Podría concluirse con que el Big Data se entiende como un recurso que puede ser utilizado por quien toma las decisiones, pero que no fuerza el sentido de dichas decisiones (Provost & Fawcett, 2013). Esto nos suscita interrogantes y cuestionamientos pedagógicos sobre la relevancia de los equipos directivos y de los gestores educativos en los diseños de los propios procesos formativos.

Respecto a los elementos restantes, los participantes también son más heterogéneos en sus respuestas. No obstante, minimizan los posibles aspectos negativos del Big Data. Esto puede deberse a que los participantes no han reflexionado suficientemente sobre las consecuencias que puede tener esta tecnología, y cómo podría suponer un problema para el avance de la educación, si no se sabe utilizar correctamente. En este punto, habrá que realizar estudios similares con participantes que sepan y tengan mayor experiencia con Big Data para que puedan valorar posibles consecuencias negativas de su utilización y proyección pedagógica (Domínguez, Álvarez & Gil-Jaurena, 2016).

El último bloque de cuestiones se centran sobre las emociones. Mayoritariamente son sensaciones positivas las que tienen los participantes, y escasamente negativas. En los resultados sólo cabría señalar la presencia de un pequeño grupo que afirma que no le causa diversión, aunque la mayoría presenta un perfil emocional muy positivo. En estudios posteriores habría que analizar si aparece ese pequeño grupo, y tratar de comprobar si se corresponde con un segmento concreto de la población.

Respecto a la última pregunta, la gran mayoría de los participantes afirma que usaría o usará esta tecnología. Esto lleva a pensar que necesitarán de una formación básica para saber cómo utilizarla. Es necesario señalar que no se han encontrado diferencias entre hombres y mujeres. Esto supone que a priori, parece no existir una brecha en el Big Data (Klasen & Lamanna, 2009). En cualquier caso, esto supone una oportunidad para actuar de forma que siga sin existir dicha brecha, esto es, con una vocación de nítida prevención para la no aparición de discriminaciones o disfunciones que resulten a todas luces inadecuadas desde posicionamientos pedagógicos modernos, inclusivos y democráticos, propios del siglo XXI.

El presente y futuro de los sistemas de formación de todo el mundo van a utilizar Big Data y analíticas de aprendizaje en diversos ámbitos, contextos y factores. Así, y a la luz de lo anteriormente explicitado, cabe subrayar la relevancia e impacto en el incremento de la flexibilidad organizacional en los centros docentes, así como en la optimización y personalización de los procesos de enseñanza-aprendizaje (Carmona & Cabrera, 2016; Lerís & Sein-Echaluce, 2011; Zapata-Ros, 2015). En este sentido, los sistemas educativos más avanzados y efectivos van a reconstruir la educación desde parámetros cada vez más ajustados a las necesidades personales de aprendizaje del alumnado (Asur & Huberman, 2010), así como a la adecuación de recursos didácticos, metodologías y estrategias evaluativas centradas en las analíticas de aprendizaje llevadas a cabo en los centros educativos.

Quizás el reto más cercano que tenemos en los próximos lustros implique un incremento significativo del nivel de sensibilización y de actualización pedagógica máxima en las aulas universitarias (Díaz-Lázaro, 2017; Ruiz, 2019), porque las próximas promociones de docentes

desconocen todavía el sentido y la orientación práctica de lo que supone Big Data para su práctica profesional. Igualmente, desde nuestra perspectiva, será básico la implementación de sistemas de recomendación formativos en la formación permanente del profesorado, aplicando diversos tipos de filtro y cribado atendiendo a múltiples variables contextuales, didácticas, tecnológicas y neurocientíficas susceptibles de su paulatina incorporación previa en distintos recursos e instancias tales como centros o institutos de formación, planes de autoformación, y, en general, en todo tipo de diseño formativo conducente a mejorar la calidad docente (Huda, Maselena, The, Don, Basiron, Jasmi & Ahmad, 2018).

## Conclusiones

El Big Data ha dejado de ser, en muy poco tiempo, una nueva y/o futura tecnología a una realidad más que apreciable y emergente que ya es una realidad en el mundo educativo. Pasar de la mera recopilación aséptica - o neutral- y masiva de datos a una intencionalidad propositiva e integradora de la información educativa no es solo necesario en términos de control de calidad, sino un compromiso moral y ético para mejorar de forma ineludible los sistemas educativos.

Finalmente, es necesario que esta actualización pedagógica no sólo se aplique en la formación permanente, sino especialmente en la formación inicial de los profesionales de la educación. Y es que la inclusión del Big Data en la educación es más que posible y deseable, en la medida en que las próximas promociones docentes tienen unas altas expectativas sobre su transferencia, a la vez que también son progresivamente conscientes de actuar para prevenir cualquier tipo de discontinuidad que potencialmente sea lesiva para la calidad y el éxito inclusivo. La revitalización de la formación inicial del profesorado pasa de forma ineludible por la proyección didáctica de cursos especializados en la gestión inteligente e inclusiva del Big Data. No se trata de estructurar automatizaciones pedagógicas sino de promover de forma participativa el análisis crítico y reflexivo de aquellos aprendizajes que se derivan de las aplicaciones personalizadas de aprendizaje, la generación de respuestas curriculares flexibles y coherentes, a la vez que desterrar toda praxis codificadora o estandarizada o, peor aún, de corte estigmatizadora que resulta contraproducente a todas luces en el marco de las nuevas y modernas tendencias pedagógicas de educación personalizada. Se trata entonces, de diseñar y aplicar estrategias, mecanismos y herramientas para desarrollar el aprendizaje autónomo en los estudiantes teniendo en cuenta el volumen de la información disponible, la variedad de ésta y la velocidad con la que se produce; esta necesidad, requiere de profesores con competencias pedagógicas y didácticas diversificadas para las necesidades de aprendizajes de los estudiantes, que son muy diversas y complejas.

Por último, debe tenerse en cuenta que esta investigación tiene algunas limitaciones, entre otras, contar con una muestra no probabilística de participantes que proceden sólo del norte de Perú. Por otro, aunque el VABIDAE ha tenido unos resultados psicométricos aceptables, debe tomarse con cierta precaución hasta que no sea validada por muestras más amplias y diversas. En próximas investigaciones deberán tenerse en cuenta estas limitaciones. No obstante, los resultados de esta investigación, así como la falta de estudios al respecto, indican que se trata de un tema que debe estudiarse con más profundidad en los años próximos. Además, la analítica de aprendizajes que implica el Big Data puede ser empleado en la optimización de medidas educativas de orientación y de mejora de las estrategias de enseñanza y aprendizaje en el propio contexto universitario. Es más, debe proveer de recursos más efectivos en la propia configuración de itinerarios y menciones formativas más ajustadas a las propias realidades y experiencias, tanto del propio alumnado universitario como del entorno escolar real. Algunas de las líneas que se sugieren son: impacto del Big Data en la organización de centros y desarrollos de políticas educativas, implementación de sistemas de Big Data a nivel ministerial, desarrollo de curadores y sistematizadores del “dato”, desarrollo de modelos de análisis específicos para la realidad educativa, así como impacto en los agentes educativos del Big Data en la realidad.

## Referencias

- Asur, S., & Huberman, B. (2010). Predicting the future with Social Media. *Proceedings of the 2010 IEEE/WIC/ACM International Conference of Web Intelligence and Intelligent Agent Technology* (pp. 492-499). Washington, DC: IEE Computer Society. Recuperado de <https://arxiv.org/pdf/1003.5699.pdf>
- Banerjee, R. (2002). Children's understanding of self-presentational behavior: Links with mental-state reasoning and the attribution of embarrassment. *Merril-Palmer Quarterly*, 48, 378-404. Recuperado de <http://muse.jhu.edu/article/23577>
- Bogarin, A., Morales, C. R., & Menéndez, R. C. (2015). Aplicando minería de datos para descubrir rutas de aprendizaje frecuentes en Moodle. *EDMETIC*, 5(1), 73-92. Recuperado de <https://dialnet.unirioja.es/descarga/articulo/5521471.pdf>
- Borrego, M., Guardado, C., Jiménez, C., Montero, R., Negro, C., & Matas, A. (2019). *Cuestionario Sobre Big Data En Educación: Escala De Valoración Del Big Data Aplicado a la Educación (vabidae)*. Recuperado de <https://vabidae.gitlab.io/vabidae/>
- Carmona, M. G., & Cabrera, A. F. (2016). Uso de Big Data en contextos educativos. In *EDUNOVATIC 2016-I Congreso Virtual internacional de Educación, Innovación y TIC.: Del 14 al 16 de diciembre de 2016. Libro de actas* (pp. 687-689). REDINE. Red de Investigación e Innovación Educativa. Recuperado de <https://dialnet.unirioja.es/servlet/articulo?codigo=5796790>
- Colás, M.P., & Jiménez, R. (2006). Tipos de conciencia de género del profesorado en los contextos escolares. *Revista de Educación*, 340, 415-444. Recuperado de [https://idus.us.es/xmlui/bitstream/handle/11441/24639/file\\_1.pdf](https://idus.us.es/xmlui/bitstream/handle/11441/24639/file_1.pdf)
- Daniel, B. K. (2019). Big Data and data science: A critical review of issues for educational research. *British Journal of Educational Technology*, 50(1), 101-113. doi: <http://dx.doi.org/10.1111/bjet.12595>
- Díaz-Lázaro, J. (2017). Análisis del Aprendizaje Social en la Educación Superior: antecedentes y líneas de actuación futuras. *Hekademos*, 22, 93-104. Recuperado de <https://dialnet.unirioja.es/servlet/articulo?codigo=6280737>
- Domínguez, D., Álvarez, J. F., & Gil-Jaurena, I. (2016). Analítica del aprendizaje y Big Data: heurísticas y marcos interpretativos. *Dilemata*, 22, 87-103. Recuperado de <https://ssrn.com/abstract=2846376>
- Dussel, I., & Reyes, B. F. T. (2018). ¿Nuevas formas de enseñar y aprender? Las posibilidades en conflicto de las tecnologías digitales en la escuela. *Perfiles Educativos*, 40(Especial), 142-178. doi: <http://dx.doi.org/10.22201/iisue.24486167e.2018.Especial.59182>
- García-Aretio, L. (2017). Educación a distancia y virtual: calidad, disrupción, aprendizajes adaptativo y móvil. *RIED. Revista Iberoamericana de Educación a Distancia*, 20(2), 9-25. doi: <http://dx.doi.org/10.5944/ried.20.2.18737>
- González-Fernández, A., & Maroto, A. M. (2017). Big data para la investigación lingüística y la educación bilingüe. *International Journal for 21st Century Education*, 4(1), 33-41. doi: <http://dx.doi.org/10.21071/ij21ce.v4i1.10122>
- Hair, J.F., Jr., Andersen, R.E., Tatham, R.L., & Black, W.C. (1998). *Multivariate data analysis* (5ª Ed.). Upper Saddle River, NJ: Prentice Hall.
- Hair, J.F., Jr., Black, W.C., Babin, B.J., & Andersen, R.E. (2014). *Multivariate data analysis* (7ª Ed.). Edinburgo (Inglaterra): Pearson Education Limited. Recuperado de [https://is.muni.cz/el/1423/podzim2017/PSY028/um/\\_Hair\\_-\\_Multivariate\\_data\\_analysis\\_7th\\_revised.pdf](https://is.muni.cz/el/1423/podzim2017/PSY028/um/_Hair_-_Multivariate_data_analysis_7th_revised.pdf)
- Huda, M., Maselena, A., Teh, K. S. M., Don, A. G., Basiron, B., Jasmi, K. A. & Ahmad, R. (2018). Understanding Modern Learning Environment (MLE) in Big Data Era. *International Journal of Emerging Technologies in Learning (iJET)*, 13(5), 71-85. Recuperado de <https://online-journals.org/index.php/i-jet/article/view/8042/4930>
- Klasen, S., & Lamanna, F. (2009). The impact of gender inequality in education and employment

- on economic growth: new evidence for a panel of countries. *Feminist economics*, 15(3), 91-132. doi: <http://dx.doi.org/10.1080/13545700902893106>
- Lawal, Z., Zakari, R., Shuaibu, M., & Bala, A. (2016). A review: Issues and Challenges in Big Data from Analytic and Storage perspectives. *International Journal of Engineering and Computer Science*, 5(3), 15947-15961. doi: <http://dx.doi.org/10.18535/ijecs/v5i3.12>
- Lerís, D., & Sein-Echaluce, M. (2011). La personalización del aprendizaje: un objetivo del paradigma educativo centrado en el aprendizaje. *Arbor*, 187(Extra 3), 123-134. doi: <http://dx.doi.org/10.3989/arbor.2011.Extra-3n3135>
- Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big data*, 1(1), 51-59. doi: <http://dx.doi.org/10.1089/big.2013.1508>
- Ratner, H., Andersen, B. L., & Madsen, S. R. (2019). Configuring the teacher as data user: public-private sector mediations of national test data. *Learning, Media and Technology*, 44(1), 22-35. doi: <http://dx.doi.org/10.1080/17439884.2018.1556218>
- Ruiz, I. I. B. (2019). Learning Analytics como cultura digital de las universidades: Diagnóstico de su aplicación en el sistema de educación a distancia de la UNAM basado en una escala compleja. *Revista Iberoamericana de Educación*, 80(1), 89-116. doi: <http://dx.doi.org/10.35362/rie8013455>
- Sobe, N. W. (2018). Problematizing Comparison in a Post-Exploration Age: Big Data, Educational Knowledge, and the Art of Criss-Crossing. *Comparative Education Review*, 62(3), 325-343. Recuperado de <https://www.journals.uchicago.edu/doi/pdfplus/10.1086/698348>
- Tejada-Escobar, F., Murrieta-Marcillo, R., Villao-Santos, F., & Garzón-Balcázar, J. (2018). Big Data en la Educación: Beneficios e Impacto de la Analítica de Datos. *Revista Científica y Tecnológica UPSE*, 5(2), 80-88. doi: <http://dx.doi.org/10.26423/rctu.v5i2.424>
- Williamson, B. (2016). Digital education governance: data visualization, predictive analytics, and 'real-time' policy instruments. *Journal of Education Policy*, 31(2), 123-141. doi: <http://dx.doi.org/10.1080/02680939.2015.1035758>
- Williamson, B. (2017). *Big data in education: The digital future of learning, policy and practice*. London: Sage.
- Williamson, B., & Piattoeva, N. (2019). Objectivity as standardization in data-scientific education policy, technology and governance. *Learning, Media and Technology*, 44(1), 64-76. doi: <http://dx.doi.org/10.1080/17439884.2018.1556215>
- Zapata-Ros, M. (2015). Analítica de aprendizaje y personalización. *Campus virtuales*, 2(2), 88-118. Recuperado de <https://dialnet.unirioja.es/servlet/articulo?codigo=5166881>