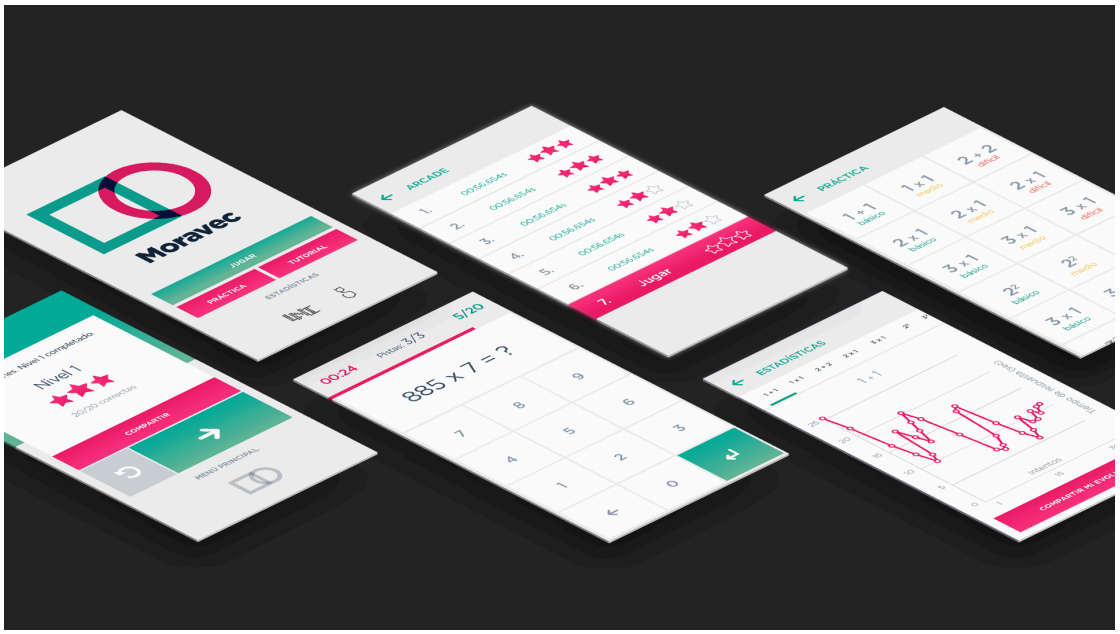


26/12/2018

Recalculando

TXT ANDRÉS RIEZNIK, PABLO A. GONZÁLEZ

¿Cómo se guardan las tablas de multiplicar en el cerebro? ¿Podemos lograr que una cosa tan aburrida como hacer cuentas se convierta en un juego con decenas de miles de usuarios?

*Compartir el siguiente artículo, que está disponible en el Anuario 2017 con leves diferencias (y que pueden conseguir **acá**), es nuestra forma de festejar que este mismo texto por suerte empieza a estar desactualizado por el progreso del proyecto **Moravec**, la app que diseñamos para entrenar la capacidad de cálculo de las personas y como herramienta para estudiar los correlatos neuronales de los procesos mentales durante estas tareas.*

*Hoy estamos celebrando dos lanzamientos: por un lado, **el nuevo y desquiciado Moravec 2.0 para Android** (cuya nueva versión podrán bajar en breve, ya que lo*

vamos a ir liberando progresivamente para cada vez más personas). Por otro lado, algo que muchas personas nos venían pidiendo desde hace un montón: ***la misma versión ya disponible para iOS****.*

Con la primera versión de Moravec aprendimos muchísimo, y hoy podemos empezar una segunda etapa en la que aplicamos todo lo que aprendimos, dada en gran medida por la expansión del equipo y la incorporación de Maxi y Lucho, gracias a quienes pudimos darnos el gusto de desarmar Moravec hasta la base y volverlo a armar. Encuentran al equipo completo y orgulloso al final de la nota.

¿Dónde y cómo guarda nuestro cerebro las tablas de multiplicar? Hace más de 30 años que la neurociencia intenta responder esta pregunta.

Una hipótesis válida sería pensar que están guardadas en forma verbal, como los versos de un poema.

Erre con erre, guitarra.

Erre con erre, barril.

Puedo escribir los versos más tristes esta noche.

Seis por ocho, cuarenta y ocho.

Sin embargo, la velocidad a la que respondemos las diferentes operaciones disminuye con el tamaño de la respuesta. Por ejemplo, respondemos más rápidamente 7×4 que 7×9, y este hecho hace que no cierre la teoría de la memoria puramente verbal.

Algún defensor de esta teoría podría decir que lo que está pasando es que 7×4 nos sale más rápido que 7×9 simplemente porque lo aprendimos antes. Así como nos sabemos mejor los versos que aprendimos en nuestra infancia y repetimos muchas veces, como los del himno nacional, sabemos mejor la tabla del dos porque la aprendimos antes y la repetimos más veces. Pero hay otro hecho bien conocido de las tablas que entra en conflicto con esta explicación: **cuando los dos números a ser multiplicados son iguales (6×6,7×7, etc.), respondemos más rápido, nos 'sabemos mejor' esos resultados.** ¿Por qué? No lo sabemos a ciencia cierta, pero lo que es seguro es que no aprendimos antes esas operaciones con números iguales, ni son más comunes en el día a día ni en los libros en general. O sea que **si guardáramos las tablas sólo en forma verbal, no podríamos explicar** la

diferencia entre la facilidad con la que respondemos cuando los dos números son o no iguales. Teoría descartada.

A medida que quienes hicimos Moravec tratábamos de entender mejor la cognición aritmética, observamos otro hecho importante: **los tiempos de respuesta para operaciones con números iguales se mantienen casi constantes**, aumentando poco con el tamaño del resultado. Esto es completamente distinto de lo que habíamos visto antes, con números no iguales.

Sabemos entonces tres cosas sobre las tablas: primero, que **cuanto mayor es el resultado, más tardamos en responder correctamente**, incluso si entrenamos muchísimo para responder lo más rápido posible.

Segundo, que **si los dos factores son iguales, respondemos más rápido**. Y tercero, que **aunque siempre tardamos más en responder si el resultado es más grande, ese aumento es menor todavía si, de nuevo, los factores son iguales.**

Además, para complicar las cosas, existe un cuarto hecho: **la tabla del 5 se responde mucho más rápido que lo que la tendencia predice**.

Y, en quinto lugar, los errores más comunes son errores dentro de la tabla, no errores numéricos. Por ejemplo, ante la pregunta de cuánto es 7×8 (56), los errores más comunes no son 55 o 57, sino 49 (7×7) o 63 (7×9).

Estos cinco hechos sobre las tablas de multiplicar se lograron establecer después de más de 30 años de investigaciones en las que decenas de participantes eran llevados a laboratorios que podían medir sus tiempos de respuesta con precisión de milisegundos. Durante esos mismos 30 años la gente se mandaba cartas, el gas se pagaba en la caja en vez de por internet y a nadie se le ocurría hacer un *swipe* a derecha o izquierda para expresar voluntad de cita o apareamiento, pero las cosas cambian y cambian rápido.

Esta es la parte donde dejamos de contar las historias de otras investigaciones y empezamos a contar la nuestra. La de nosotros, quienes hacemos **Labs** (la pata de investigación del Gato), con la participación estelar de Fede Zimmerman y Diego Shalom. Una historia que empieza con dos datos simples: los celulares (que básicamente son computadoras) pueden medir tiempos de respuesta con esa

precisión (milisegundos), y los participantes pueden ser contactados vía esta cosa fantástica que es la internet.

Hace dos años nos hicimos una pregunta ahora obvia: **¿podemos reproducir los resultados obtenidos en laboratorio durante 30 años de investigación usando los celulares que todos tienen hoy en día en sus bolsillos?** Creíamos que sí, pero necesitábamos que muchas personas hicieran cuentas en sus celulares y, encima, queríamos que quisieran hacerlas, no sólo por no tener presupuesto para pagarles a los sujetos, sino porque es muy distinto trabajar con alguien que tiene una motivación económica por participar que estudiar cómo aprende alguien que está intrínsecamente motivado. Ese era el desafío principal: **queríamos lograr que tantas personas como fuese posible hicieran algo que parece, en principio, lo más aburrido del mundo: cuentas. Muchas cuentas**. Porque si hay un plan imbatible de sábado a la noche es multiplicar números grandes.

Esta fue la primera vez que nos encontramos de frente con una idea que hoy cimienta todo lo que hacemos en Labs: ¿y si en vez de pedirle a alguien que haga un experimento, lo diseñamos para que sea en sí mismo una experiencia piola? Algo que el otro quiera hacer. Algo entretenido. Y no sólo disfrutable en el momento sino en el gran arco argumental, en la historia. La idea de participar de un experimento sabiendo que lo es, **tomando el compromiso de nuestro lado de que, al terminar, les íbamos a contar exactamente lo que habían ayudado a descubrir**. Justo en la intersección de esas ideas es donde lo mejor de toda la aventura que ha sido ser parte de la construcción de Gato se nos presenta como una epifanía (aunque sin la parte de las estatuas que lloran fluidos corporales): fuimos, entre todos, construyendo una comunidad de personas genuinamente interesadas en compartir ciencia, y si hay una forma última de compartirla es haciéndola.

Teníamos las herramientas para diseñar un experimento/experiencia y un medio para invitar a los usuarios a transitar esa experiencia. Hicimos, entonces, el experimento de hacer un experimento. Pero si lo íbamos a hacer, metámosle todo. Si Gato es <u>investigación, comunicación y diseño</u>, queríamos pensar esta experiencia como una intersección de las tres.

Diseño aplicado a la ciencia, habrase visto

Ningún pibe nace Moravec (bueh, salvo la familia de los Moravec), y este Moravec nació como *Entrenamente*. Creada por Fede Zimmerman (bajo la dirección de Andrés Rieznik y Mariano Sigman), *Entrenamente* arranca como una app para Android. Desarrollarla era parte de la tesis de grado en ingeniería electrónica de Fede, y después de unos meses de trabajo fuerte, *Entrenamente* andaba. Y andaba MUY bien. Tomaba datos de una manera fenomenal y los ofrecía de tal modo que podíamos investigarlos. En ese primer producto encontramos simultáneamente un gran valor para la aplicación, pero también su limitación: la habíamos desarrollado sin pensar en el diseño, en la experiencia estética, ni habíamos armado un plan para dar a conocer nuestra app y que se juegue. Es decir, **habíamos pensado en todo menos en los usuarios**. Las personas.

Sobre lo que sí ya habíamos trabajado mucho en *Entrenamente* era sobre la *jugabilidad*, o lo que nosotros creíamos que esa palabra significaba. En algún lugar de Japón, allá por los '80, existió una conversación que contenía ideas algo así como *'che, ¿primero sale la planta carnívora de adentro del tubo o se viene encima la tortuga pato voladora?'*. Ellos no estaban armando para ellos, sino que estaban pensando en un otro. Un otro (y otra) que tenía que divertirse, aprender, superar desafíos y, por sobre todo, pasar tiempo jugando. Porque era eso. Jugar. **Necesitábamos armar algo jugable**.

Nosotros no nos preguntábamos por plantas carnívoras ni tortugas voladoras (y ningún plomero fue dañado en el proceso), pero nos preguntábamos por sumas, multiplicaciones y potencias. Pero las preguntas se parecían más de lo que esperábamos: ¿En qué orden y cada cuántos niveles teníamos que presentar un enemigo nuevo, un desafío distinto? ¿Cómo hacemos que un usuario sienta un cachito de adrenalina al resolver algo difícil, pero también pueda tener momentos de relajación y disfrute porque viene pisteando como un campeón?

Mejorar *Entrenamente* también implicaba buscar fuentes nuevas, y nosotros (acá nosotros es un Fede trasnochado) encontramos una en Gamasutra, la Biblioteca de Babilonia de los que aman jugar y hacer juegos, y fue ahí donde entendimos que 'Diseño de estructura de niveles' no la íbamos a cursar ni en Ingeniería ni en Exactas. Qué corto te quedaste, Pubmed, en darnos respuestas para estas preguntas. Tan corto que tuvimos que usar la fuerza para solucionar el problema. Bueno, la fuerza no: La Fuerza.

La respuesta era obvia (?): aprendamos como aprenden los grandes. Los héroes. Los Jedis. **Diseñemos los niveles para que progresen igual que las grandes historias**. Nuestro experimento, que ya era experiencia y que ya empezaba a ser juego, tenía que parecerse a Star Wars. Literalmente.

Igual, tensión dramática en serio es besar a tu hermano.

Lo que también necesitábamos ahora era asegurarnos de que los participantes realmente intentaran resolver bien cada una de las operaciones. ¿Cómo le transmitimos al usuario que no nos da lo mismo si pasa el nivel raspando o si mete todas bien una atrás de la otra? En realidad, ¿cómo hacemos que sea obvio para cualquiera que baje la app sin saber que se trata de un experimento? **¿Cómo lo convertimos en un juego?** Con caramelos. Bah, con caramelos, puntos, festejos. Al momento de terminar un nivel, le dábamos al usuario una idea de qué tan bien le había ido en forma de puntos (que luego fueron coquetas estrellitas) en función de las operaciones correctas.

Hasta acá habíamos llegado con *Entrenamente*. El primer paso para empezar a transformarlo en Moravec, ya entendido el mapa general, fue atacar los problemas particulares, y uno eran las subidas en dificultad. Esas subidas incluían hacer cosas nuevas, como multiplicar de izquierda a derecha o descomponer los cuadrados de una manera distinta a la que se enseña en los colegios.

Para pensar en el usuario teníamos también que convertirnos nosotros y nosotras (ya sumado todo el equipo Gato) en usuarios, y teníamos la ventaja de que no todo el equipo era '**matemago**', como Andrés, así que ninguno tenía idea de cómo hacer la mayoría de las operaciones. Eso nos ayudó a desaprender lo que ya sabíamos para ponernos en los zapatos de una persona que se enfrenta a la aplicación por primera vez.

Para presentar las operaciones a los jugadores nuevos, filmamos tutoriales. Tutoriales de YouTube. Porque si podés aprender a perrear, resolver un cubo Rubik o hacer un estuche de teléfono con Oreos, también podés aprender a multiplicar.

"Pipiripipipipipipi pipiripipipipipipi" Andrés Rieznik (Doctor en física, experto en fibra óptica y cognición aritmética)

Listos para hacer cuentas, entramos en el diseño visual. ¿Cómo tenía que verse la app? Acá es donde el diseño, futilidades por las que los científicos no nos preocupamos (!), nos mostró su poder. Elegimos simular una calculadora para que de inmediato el usuario entendiera la interfaz, no necesitara aprender a usar la aplicación y pudiera concentrarse en las operaciones. Punto para los diseñadores.

Pero, ya que habíamos llegado hasta ahí, decidimos ir por todo y lo *gamificamos* de tantas las maneras que se nos ocurrieron. ¿El usuario quiere saber si mejora? Perfecto, acá tenés un registro de tu propio desempeño, con pomposos gráficos de progreso. ¿Querés hablar con otros usuarios? Armemos **un grupo de Facebook para usuarios**, compartamos puntajes, compitamos un poquito, que un WhatsApp de 'estoy en el nivel 135, ¿vos?' nunca mató a nadie (aunque dos miembros del equipo tuvieron una picante escalada que los llevó a multiplicar números de cuatro cifras al cuadrado en menos de un mes de práctica).

Pero todavía nos quedaba pendiente el tema del nombre, y *'Entrenamente'* no nos convencía, de la misma manera que no nos convencen las librerías artísticas que se llaman LiberArte, EnAmorArte, PintArte y DescongestionArte, esa que, además de acuarelas, vende expectorantes. Así fue que nos pusimos a pensar y buscar un nombre que también tuviese una narrativa entretejida.

¿Cómo se llamaba el flaco este? El de inteligencia artificial y robótica. El de la paradoja esa que habla acerca de que las máquinas pueden hacer con facilidad cosas que nosotros humanos no podemos (como multiplicar números enormes), pero les cuesta mucho hacer cosas fáciles para nosotros (como calcular la distancia a un vaso para poder brindar porque encontramos un re buen nombre para la app). ¿Hans? Hans algo.

Hans Moravec. Porque si el nombre encima cuenta una historia, vale doble.

Appaaaaa.

Ciencia en el baño

Gracias a Moravec logramos nuestro objetivo y nos pasamos 4 paradas, porque lo que ocurrió nos sorprendió también a nosotros: **en dos semanas de toma de datos reprodujimos los resultados de 30 años de investigación científica**. En

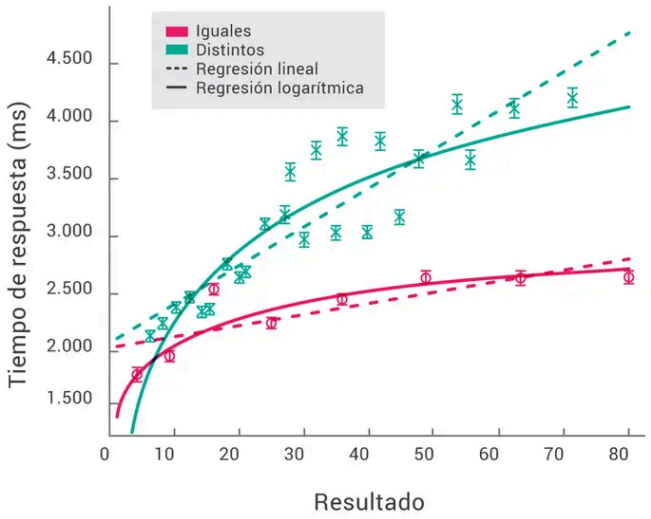
dos semanas, arriba de 500 usuarios manija hicieron más de 90 mil cuentas. Y la cosa siguió: al momento de publicar esta nota (en 2017), los usuarios eran +2500, y las operaciones +500.000, pero nos vamos a permitir decirlo como MEDIO MILLÓN, lo que implica que esta nota pasa de ser 'la de Moravec' a 'la primera de Moravec' y 'el experimento' pasa a ser 'la línea de investigación', porque tenemos datos para hacer escabeche y usuarios tan entusiasmados que estamos pudiendo medir no sólo a muchas personas haciendo pocas operaciones, sino a un grupo cada vez más grande que juega 80 o 100 niveles. O sea que no sólo estamos entendiendo cómo hacemos normalmente las cuentas, sino **cómo progresamos, entrenamos y mejoramos en el tiempo gracias a la práctica**.

Nota del Editor: el update importante es que esa idea de 'medio millón' que era una locura hace un año, hoy está opacada por un número real de casi 12 millones. Una locura v2.0. Y como ya estamos cansados de tener textos desactualizados, lo mejor es que vean un contador que se actualiza, y hasta que vean lo que pasa cuando publicamos esta nota.

Select plot: Cumulative operations

Lo que aprendimos

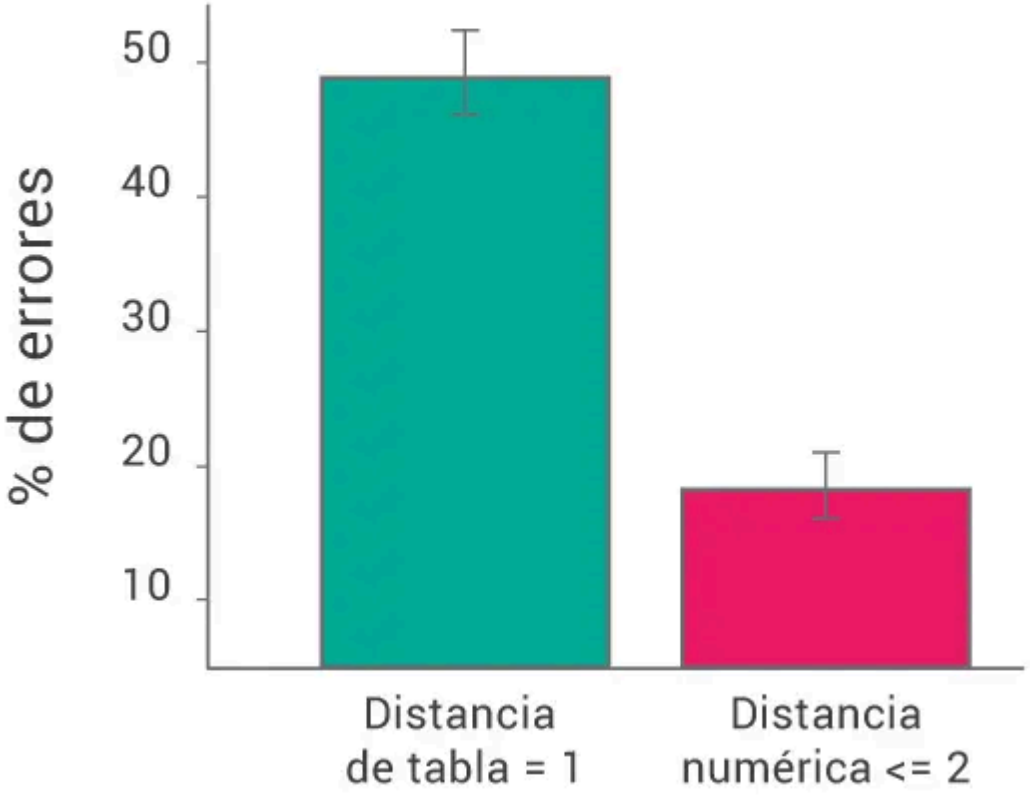
Los tres primeros efectos conocidos de las tablas de multiplicar que explicamos más arriba se reflejan en este gráfico, hecho con los datos de Moravec, donde cada punto es una operación de la tabla de multiplicar. Los puntos rojos corresponden a las ocho operaciones con los dos factores iguales (2×2, 3×3, etc.), y los verdes a las demás operaciones.



Acá podemos ver cómo a medida que el resultado es mayor, aumenta el tiempo de respuesta. Hasta ahí todo bastante intuitivo, pero también podemos ver dos cosas interesantes más: cuando multiplicamos un número por sí mismo (en magenta), el patrón es radicalmente distinto y sube muy de a poco; esto repercute en que, a medida que la operación es mayor, la distancia entre multiplicar dos números distintos y dos iguales crece. También probamos dos modelos que 'ajustan' a los puntos, o sea que tratan de describir el fenómeno subyacente (o de comprimirlo mediante un modelo). Vemos que un modelo logarítmico (en línea llena), 'explica' mejor los puntos que uno lineal (punteado).

En esta primera figura confirmamos que: (1) los tiempos de respuesta aumentan cuando aumenta el resultado de la cuenta; (2) los puntos rojos están más abajo, es decir, son más fáciles esas operaciones y (3) la distancia entre puntos verdes y rojos aumenta para multiplicaciones de mayor resultado.

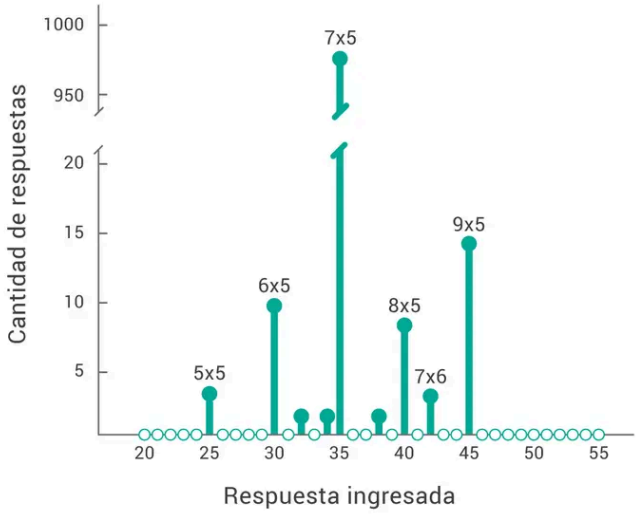
Otra cosa interesante que encontramos es el hecho de que **los errores más comunes son vecinos en la tabla de multiplicar**, y podemos verlo en este gráfico:



En la figura superior podemos ver cómo la mayor cantidad de errores (casi el 50%) cae justo en un número 'vecino de tabla': para 7×5 = 35, los vecinos de tabla son 28, 30, 40 y 42 (cosa que

con otra forma de visualizar los datos se ve más intuitivamente en la figura inferior). Diseño FTW.

Si vemos la figura superior de recién, notamos que casi el 50% de los errores se deben a que los participantes respondieron uno de los cuatro vecinos en la tabla (en la figura inferior vemos que, por ejemplo, los vecinos en la tabla de la respuesta para 7×5 = 35 son las respuestas 28, 30, 40 y 42). En contraste con este valor de 50%, vemos que menos del 20% de los errores corresponden a una de las cuatro respuestas que están a una distancia de uno o dos números del resultado correcto.



Acá podemos ver las respuestas dadas, correctas e incorrectas, a la pregunta ¿cuánto es 7×5? Otra vez, atentos a dónde caen los errores.

Esto tiene sentido porque **así se organiza (en general) nuestra memoria: nos equivocamos por vecinos, ya sea de la tabla, semánticos, sintagmáticos o geográficos**. Análogamente, si nos preguntan cuál es la capital de Finlandia (Helsinki), es más probable que digamos Oslo o Estocolmo, que Quito o Bogotá.

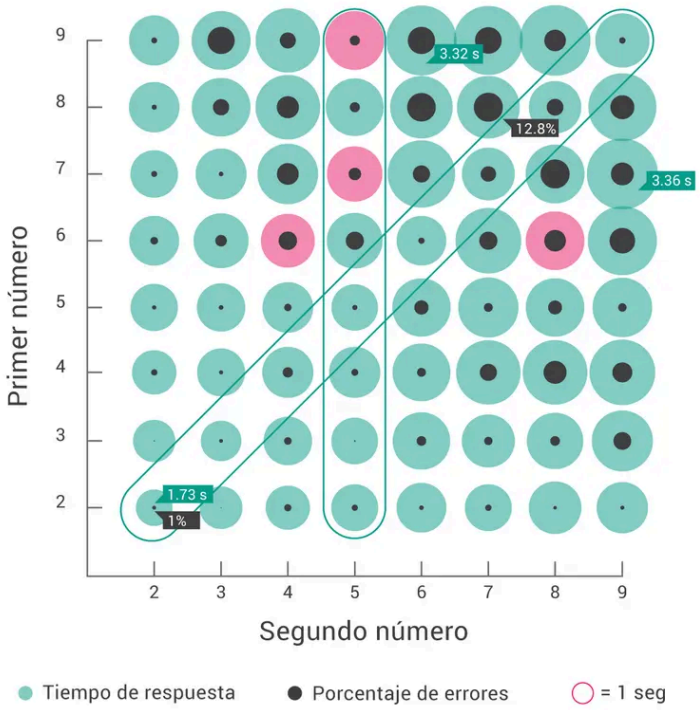
Con todos los datos que teníamos, pudimos ir más allá y encontrar efectos pequeños que nadie había visto antes (porque no tenían tantos datos, y por lo tanto el poder estadístico suficiente). **Pudimos medir diferencias significativas entre operaciones iguales pero conmutadas**. Por ejemplo, 2×3 y 3×2 dan el mismo resultado. Pero tal vez no respondemos igual de rápido en un caso que en el otro. **Descubrimos cuatro operaciones de la tabla en las que se observa una diferencia significativa si se pregunta en un orden o el otro**, y lo divertido es

que, antes de revelar cuáles son esas cuatro operaciones, uno podría intentar adivinarlas.

ESPACIO RESERVADO AL PROCESO DE ADIVINACIÓN

– Donald, ¿vos llegás a ver cuáles son las operaciones con efecto de orden?
– AMEWICAAAAA.
– Donald, eso no fue una respuesta.
– FAKE NEWS.
– ¿Donald? Si sos un rehén y no querés ser más presidente, parpadeá.

La respuesta tiene que ver con lo verbal.



Acá puede verse el tiempo de respuesta promedio para cada operación, así como el porcentaje de errores. De esta figura sale un DATAZO, y es que 8×7 es la operación en la que más pifiamos: 12.8%. UNA BARBARIDAD. También podemos ver 4 operaciones que rompen todos los patrones y comparten una característica en común: en todas, la operación y el resultado riman.

Las operaciones donde la permutación se pone asimétrica son las que tienen rima en un orden, pero no en el otro: 6×4 = 24, 7×5 = 35, 9×5 = 45 y 6×8 = 48. En estos cuatro casos, las operaciones se responden algunas decenas de milisegundos más rápidamente si preguntamos en el orden rimado.

O sea que volvimos al principio: **sí, hay algo de memoria verbal en las tablas**. En la pila de palabras anterior nos encargamos de argumentar exactamente por qué no era así, para ahora decir que un poco así es. Bienvenidos y bienvenidas al maravilloso mundo de la investigación científica. Desconcertante, bello y sexy.

Más allá de estos primeros resultados, que contribuyen desde Argentina y desde un equipo mixtísimo de académicas, diseñadores, desarrolladoras y científicos de garaje a un debate que los mayores laboratorios del mundo están teniendo, Moravec conlleva una importante conclusión: **un grupo entusiasmado de usuarios puede motivarse a ser parte de la construcción de un pedacito de**

conocimiento novedoso si diseñamos el experimento para que sea una experiencia interesante, si contamos exactamente de qué manera con su participación están haciendo una contribución a la ciencia, y si al final del experimento investigadores y usuarios nos sentamos juntos a compartir experiencias y resultados.

Hoy nos toca no sólo festejar lo que ya pasó, sino contar a dónde vamos: por un lado, quienes hacemos Moravec somos cada vez más y hacemos cada vez más cosas, desde desarrollar el sistema que maneja los datos hasta ver cómo se integran con todo Labs, pasando por cómo se siente y se ve la aplicación, cómo extraemos información de un *dataset* cada vez más rico y grande o cómo logramos más usuarios y encontramos formas nuevas de contarles los resultados. Por el otro, las preguntas que aparecen: ¿qué podemos aprender sobre el proceso de aprendizaje? Veamos datos longitudinales. ¿Podemos usar Moravec como herramienta para identificar desórdenes en la comprensión matemática? Probemos usarlo con niños y niñas. ¿Cómo ponemos a prueba una hipótesis que incluye el lenguaje? Hagámoslo en otros idiomas.

Brace yourselves, crisis de replicabilidad y conclusiones con tres sujetos, que somos un equipo de investigación multidisciplinario cada vez más manija y un montón de usuarios queriendo ser parte de la construcción de un pedacito de conocimiento, jugando a hacer ciencia colectiva.

Equipo Moravec

Federico Zimmerman, Diego Shalom, Maximiliano Suppes, Pablo A. Gonzalez, Juan Manuel Garrido, Facundo Alvarez Heduan, Luciano Leveroni, Stanislas Dehaene, Mariano Sigman, Andres Rieznik.

Yapas

Andrés resumiendo Moravec: https://youtu.be/7Ucav9ytpkU

Todos los resultados que contamos acá, entre otros, fueron publicados en una revista científica indexada a principios de 2017 (lo encuentran acá).

elgatoylacaja.com/investigacion/recalculando

Sumate en
eglc.ar/bancar