16/06/2022

Quizás, quizás, quizás

TXT ARIEL ZYLBERBERG

¿Cuánto es bastante? ¿Qué significa “seguramente”? ¿Es tu quizás mi tal vez?

Hay cosas que sabemos que sabemos, y otras que sabemos que no sabemos. En el medio, cosas que creemos saber, pero no estamos seguros. Muchas lenguas cuentan con múltiples expresiones para comunicar incerteza en nuestras decisiones o conocimiento. Estas expresiones (‘puede ser’, ‘tal vez’, ‘quizas’, ‘seguro’) juegan distintos roles en la toma de decisiones y en la cognición. Por ejemplo, es más probable que salgamos con paraguas si alguien dice ‘estoy seguro de que hoy va a llover’ que si esa misma persona dice ‘quizás hoy llueva’. La confianza afecta qué tan enfáticamente comunicamos nuestras decisiones a otras personas, cómo

ponderamos las opiniones ajenas en relación a nuestras propias decisiones, cómo planeamos acciones futuras y qué tanto peso le damos a las opiniones individuales.

Claro que, cuando conversamos, estamos teniendo en cuenta muchas más cosas que el mero contenido de las palabras que se dicen. Al hablar, nos importa tanto lo dicho como lo no dicho, la relación que tenemos con quien hablamos, el contexto, el tono de las palabras y un sinfín de variables más. Pero, a los fines de este análisis, vamos a quedarnos con el contenido de las palabras. El significado. Desde ese punto de vista, un potencial problema en el uso de expresiones verbales de incerteza es su ambigüedad. Por ejemplo, la expresión 'puede ser' representa distintos grados de certeza para distintas personas. La complejidad de este problema se aprecia en los siguientes ejemplos.

Imaginá que debés decidir si someterte a una operación quirúrgica riesgosa y decidiste que lo harás solo si la probabilidad de éxito es mayor al 80%. Supongamos que consultaste a dos médicas acerca de la probabilidad de éxito de la operación. Una consideró que 'es probable' que la operación sea exitosa, y la otra que 'puede ser' que la operación sea exitosa. ¿Cuál te inspira más confianza? ¿Si sumás ambas opiniones, te está dando una probabilidad mayor a 80%?

O suponé que sos parte de un jurado que debe decidir si condenar o no a un acusado, y el testigo clave dice estar 'bastante seguro' de haber visto al acusado en la escena del crimen. ¿Qué dice este testimonio acerca de la probabilidad de que el acusado haya estado en el lugar? ¿Deberíamos pedirle al testigo que reporte su confianza con expresiones verbales o con números?

En una investigación que estamos llevando adelante, intentamos entender las diferencias y similitudes entre las expresiones verbales de probabilidad y los reportes numéricos de probabilidad a la hora de comunicar grados de incerteza. Para esto, diseñamos un experimento web en el que les pedimos a más de 8.000 participantes que nos dijeran cómo interpretan distintas expresiones verbales de incerteza y que luego las usaran para comunicar la confianza en una decisión.

Antes de avanzar con los resultados preliminares, recomendamos atravesar el experimento en primera persona. No es imprescindible para entender lo que sigue, pero te va a dar una experiencia más rica y una mejor apreciación de todo lo que

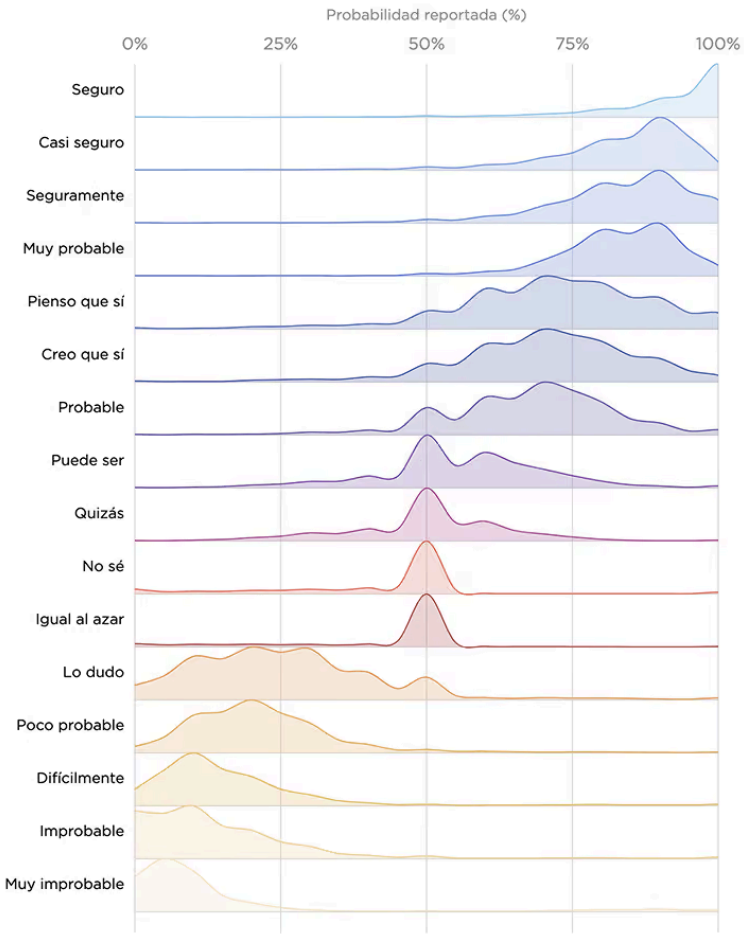
viene a continuación. Si aún no lo hiciste, **podés participar acá**. Lleva sólo 10 minutos.

¿Es tu 'quizás' mi 'tal vez'?

La primera parte del experimento está basada en investigaciones anteriores similares. Como solemos hacer en nuestros estudios, lo primero que buscamos es reproducir algún resultado conocido para entender si nuestra herramienta y metodología de toma de datos funciona como esperamos.

Al inicio, presentamos 16 expresiones verbales de incerteza y les pedimos a las personas que indiquen, para cada una de estas, cuál es el valor de probabilidad que mejor captura su significado. Los participantes asignaron a cada expresión un valor de probabilidad en una escala de 0 a 100%.

A continuación, vemos la distribución de valores numéricos asignados a cada expresión verbal. Las palabras están ordenadas de mayor a menor valor medio (promedio) de probabilidad.



Lo más saliente de este resultado es que existe una gran dispersión respecto a los valores de probabilidad asignados para cada expresión. Es decir, **el 'casi seguro' de algunas personas es el 'puede ser' de otras**.

Esta gran dispersión en los valores numéricos que las personas asignan a cada expresión verbal plantea la pregunta de **cómo es que nos comunicamos eficientemente si la misma palabra significa distintas cosas para personas distintas**.

Una respuesta posible es que, de hecho, no nos comunicamos tan eficientemente como creemos. Otra respuesta, más acorde a lo que diversas teorías (lingüísticas, semióticas, etc.) han desarrollado, tiene que ver con que la comunicación es un fenómeno multidimensional que se produce por varios canales a la vez y que nunca, o casi nunca, puede reducirse a un mero correlato entre la información y el código con el que es brindada.

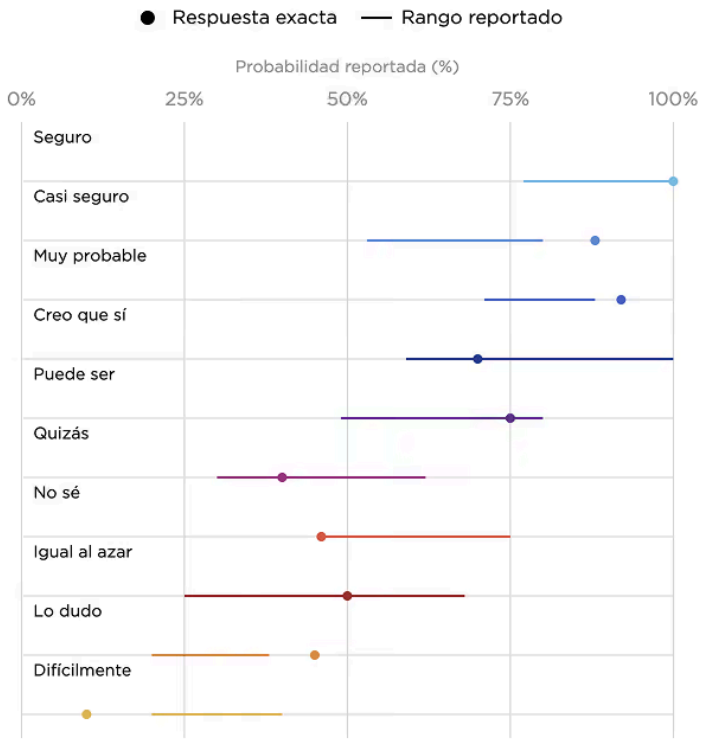
Pero, en casos donde el rigor es necesario, surge la duda: ya que son tan ambiguas, ¿no podrían evitarse las expresiones verbales para comunicar grados de certeza y, en

su lugar, priorizar el uso de números en vez de palabras? Esto, de hecho, se ha intentado en algunos casos específicos. Por ejemplo, **el servicio meteorológico nacional de EE.UU. restringe el uso de expresiones verbales de incerteza a valores específicos de probabilidad**. Sus meteorólogos sólo pueden usar la expresión 'leve chance' (*slight chance*) de precipitaciones cuando la probabilidad está entre el 10% y el 20%, y la expresión 'probable' (*likely*) solo si la probabilidad estimada está entre el 60% y el 70%.

Uno de los objetivos de nuestro experimento era **evaluar la siguiente hipótesis: las expresiones verbales de probabilidad no son interpretadas como valores puntuales de probabilidad, sino como distribuciones de probabilidad.** Según esta hipótesis, la expresión '*puede ser* que X sea cierto' no significa que 'X tiene una probabilidad de *30%* de ser cierto' sino que refiere a **un rango de valores posibles** (o una distribución de probabilidad sobre valores de probabilidad). Si esta hipótesis es correcta, al pedir un reporte puntual de probabilidad para cada expresión verbal, estábamos forzando a los participantes a hacer algo que contradice el uso habitual de estas expresiones.

Para evaluar esta hipótesis, en la siguiente etapa del experimento le pedimos a los participantes que reportaran el rango de valores (mínimo y máximo) que capturase de mejor manera el significado de cada expresión. Incluimos sólo un subconjunto de las expresiones verbales que usamos en la primera parte del experimento.

A continuación se muestra el ejemplo de un participante random, donde el círculo indica el valor puntual de probabilidad que fue reportado en la primera parte del experimento, y la línea horizontal roja indica el rango reportado.



Lo que descubrimos al hacer esto fue que, en general, los rangos reportados son consistentes con el valor puntual de probabilidad asignado a cada palabra por el participante en la primera parte del experimento. Pero, en algunos casos, el valor puntual de probabilidad quedó fuera del rango reportado por el mismo participante. Dicho de otro modo, hubo casos en los que alguien asignó un grado de probabilidad de, por ejemplo 85% y luego, cuando se le pidió un rango para la misma palabra, le asignó un 60% a 80%. En el 81% de los casos, el valor puntual estaba dentro del rango pero ¿por qué a veces no? Asignar un valor de probabilidad explícito a una expresión verbal no es algo que hagamos cotidianamente y puede tener un componente azaroso, no es como recordar un número de teléfono.

No sé, puede ser

La ventaja de pensar que cada palabra representa una distribución de probabilidades es que no solo nos da una idea de qué valor es más probable que podamos tomar de referencia para saber cuál es el "significado" de esa palabra, sino que también nos habla sobre la incerteza alrededor de ese valor. Y la incerteza es

clave cuando nos proponemos combinar distintas expresiones. Por ejemplo, para tomar una decisión cuando una médica dijo que la operación *probablemente* salga bien y otra dijo que es *casi seguro*.

Cuando partimos de dos distribuciones, se tienen en cuenta no sólo los valores medios sino también cuánta dispersión tienen esas distribuciones. El resultado de combinarlas no es un número sino una nueva distribución de probabilidad. La distribución de menor dispersión (la más angosta) es la que más peso tiene sobre la distribución final resultante. El valor central (el valor medio) resultante no es el promedio de los valores centrales originales, sino que está más cerca del valor medio de la distribución más angosta. **Esta forma de combinar distribuciones de probabilidad está dentro de lo que llamamos estadística bayesiana** (por el matemático inglés Thomas Bayes).

Nuestra hipótesis es que lo mismo ocurre cuando les presentamos a las personas combinaciones de expresiones de probabilidad: las combinan de esta forma. Dado que existen múltiples procesos cognitivos que son consistentes con una visión bayesiana, esta en principio sería una hipótesis plausible.

Para evaluar cómo las personas combinan expresiones de probabilidad, les presentamos distintos pares de expresiones verbales y les pedimos que reportaran el valor numérico que mejor capturara la combinación de ambas expresiones. Las expresiones a combinar podían ser, por ejemplo, 'quizás' y 'casi seguro'.

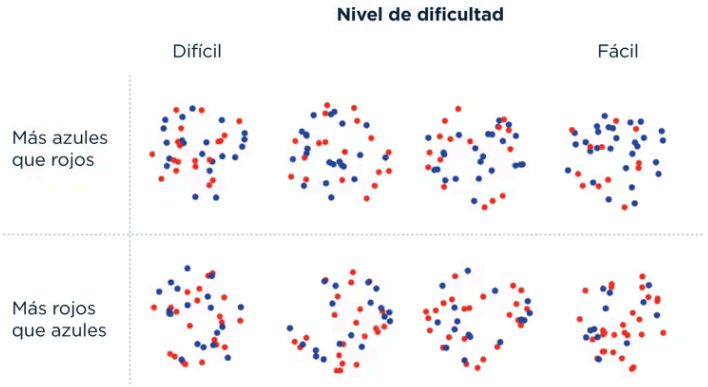
Consistente con la predicción de una combinación bayesiana, encontramos que **el valor de las combinaciones de expresiones verbales reportado tiende a estar más cerca de la expresión que, dentro de cada par, los participantes juzgaron como menos variable (menor rango o dispersión).** En otras palabras, las personas le dieron menos peso en la combinación a la expresión cuyo valor consideraban más incierto.

Parecería ser que **las expresiones verbales de incerteza comunican no sólo un valor medio de probabilidad sino también un grado de variabilidad (o dispersión) alrededor de este valor medio**. Es decir que cada expresión abarca un rango de valores de probabilidad y las personas usan esta información al interpretar el significado de pares de expresiones.

Decisiones, confianza y el peso de las opiniones ajenas

En la tercera y última etapa del experimento, ensayamos el uso de estas expresiones de probabilidad en la práctica, estudiando **cómo la expresiones numéricas y verbales difieren al ser usadas para reportar la confianza que se tiene en una decisión**.

Les presentamos a las personas diferentes estímulos compuestos por puntos azules y rojos. En cada ensayo, el estímulo era presentado por medio segundo. Luego, las personas reportaban si habían más puntos rojos o azules en el estímulo, y la confianza en que la decisión que tomaron era la correcta. En la mitad de los ensayos el color mayoritario era el azul, y en la otra mitad el color mayoritario era el rojo. La dificultad de la decisión se controló cambiando la proporción de puntos del color mayoritario. Exploramos cuatro niveles de dificultad, ilustrados en la siguiente figura.



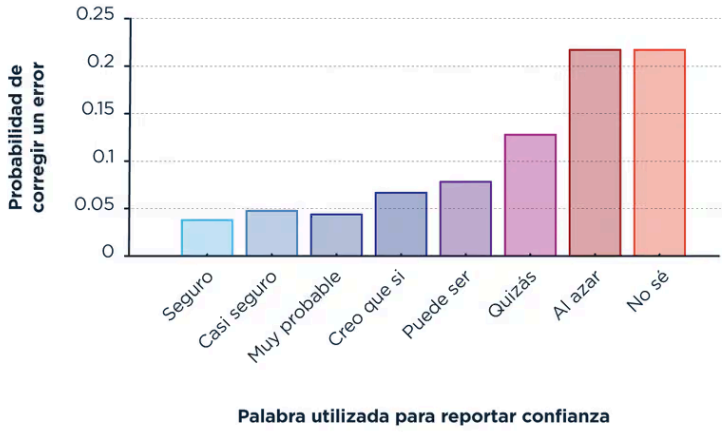
Para reportar la confianza en su decisión, algunas personas debían usar números (probabilidad numérica de que la opción elegida fuese correcta) y otras debían usar expresiones verbales.

Luego de reportar la decisión y la confianza para un ensayo, le presentamos a cada persona la decisión y la confianza de otro participante que vio el mismo estímulo. Finalmente, las personas tenían que volver a reportar su decisión y su confianza, luego de ver la opinión del otro participante, lo que les permitía modificar sus decisiones iniciales si así lo quisieran.

El primer resultado (esperable) es que **hay más decisiones correctas cuando los ensayos son más fáciles**, es decir, cuando la diferencia entre el número de puntos azules y rojos es mayor. Además, la performance no cambia mucho si tomamos la respuesta final dada por los participantes luego de ver la opinión ajena, solo mejora un poco.

Pero ¿es porque la valoración de otros no ayuda tanto o porque **las personas no le dan suficiente peso a las opiniones ajenas al revisar su opinión inicial**? Si en cada caso, ante la repregunta, las personas hubieran elegido la opinión emitida con mayor confianza independientemente de si era la propia o la ajena, la precisión obtenida **sería mucho mayor**. O sea que la opinión ajena ayudaba a tener un mejor resultado pero se le dio menos importancia que lo que debería. En este experimento de puntos y colores, la mejor opción era confiar siempre en quién tuviera más convicción. Pero no siempre tiene que ser así, fuera del laboratorio (virtual en este caso), las personas pueden reportar alta confianza en sus decisiones no solo por estar convencidos de que la opción escogida es correcta, sino para, por ejemplo influenciar a otras personas o ser vistos como personas capaces o asertivas. Por ejemplo, un agente de inversiones que dice no saber si las acciones van a subir o bajar no va a tener muchos clientes. Esto puede hacer que las personas crean menos en la confianza ajena que en la propia, a pesar de que, en nuestro experimento, este comportamiento no sea el óptimo.

Después de encontrar estos resultados, quisimos ver cómo cambiaba la probabilidad de corregir un error según la expresión elegida por el usuario para describir su confianza inicial. Se ve que en todos estos casos, la probabilidad de cambiar de opinión es baja, menor al 25%, consistente con el resultado anterior que sugería que **las personas son menos influenciadas por las decisiones ajenas de lo que deberían** para acertar al color mayoritario.



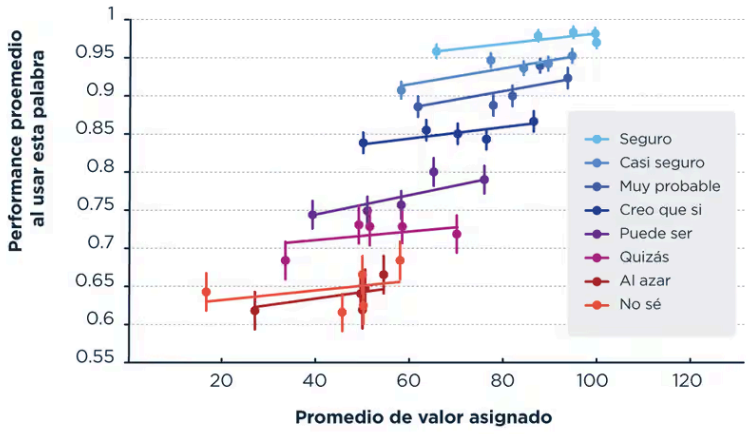
Probabilidad de corregir un error inicial según la expresión elegida para reportar confianza. Solo incluimos decisiones en las que el otro participante reportó la decisión correcta.

Hechos y palabras

En la primera parte del experimento vimos que los valores de probabilidad asignados a cada expresión verbal variaban mucho entre personas. Es decir, a la misma expresión (por ejemplo, 'seguro'), algunas personas le asignan una probabilidad numérica mayor que otras. Ahora vamos a evaluar si esta variabilidad refleja verdaderamente diferencias individuales en el uso 'real' de las expresiones verbales en un contexto concreto. Si una persona considera que la expresión 'seguro' solo representa casos donde la probabilidad de estar en lo correcto en el juego de los puntitos es, por ejemplo, mayor a 95%, deberíamos ver que la precisión de sus decisiones sea muy alta cada vez que elige la palabra 'seguro' para describir su confianza. Por el contrario, si otra persona considera que 'seguro' puede usarse para probabilidades cercanas a, por ejemplo, 70%, entonces en los casos en los que utilice esta expresión va a ser (en promedio) menos precisa que el participante que la use sólo cuando esté más convencido.

Para estudiar esto, dividimos a los participantes en 5 percentiles según la probabilidad asignada a cada expresión en la primera etapa del experimento. Luego, graficamos la performance al usar esta palabra (es decir, cuán precisa fue la decisión) en función del promedio del valor asignado a la palabra en la parte inicial del experimento. Lo que vemos en la figura es justamente lo que esperábamos:

personas que, por ejemplo, asignaron una probabilidad de 100 para la palabra ‘seguro’, luego la usaron para situaciones donde la precisión de su decisión es muy alta. Mientras que personas que le asignaron una probabilidad de 70, la usaron para reportar confianza en decisiones donde la precisión fue menor. Por lo tanto, **la variabilidad individual** (mostrada en la primera parte del experimento) **refleja diferencias individuales en el uso de las expresiones verbales de incerteza.**



Conclusión

Las personas tenemos la capacidad de asignar grados de certeza a nuestras creencias y afirmaciones, lo que nos permite razonar sobre proposiciones que no pueden clasificarse simplemente como verdaderas o falsas. Central a esta capacidad es el concepto de probabilidad como forma de representar grados de creencia. Las personas contamos con dos sistemas para representar y comunicar incerteza. Uno —en principio más preciso y mucho más reciente en la historia de nuestra especie — consiste en usar valores numéricos de probabilidad. Otro —más antiguo y más ambiguo— consiste en utilizar expresiones verbales de incerteza (creo, puede ser, seguro, etc). Con nuestro experimento quisimos explorar las similitudes y diferencias entre estos dos sistemas, en particular cuando deben usarse para expresar la confianza en una decisión.

Replicamos en idioma español un efecto observado originalmente en inglés, en el que se reportó que las expresiones verbales de incerteza se mapean como un valor numérico de probabilidad con alta dispersión. Esto, lógicamente, se interpreta

como evidencia a favor de la ambigüedad de las expresiones verbales. Pero en nuestro caso propusimos, además, una hipótesis según la cual las personas no sólo tienen en cuenta la ambigüedad de las expresiones verbales de incerteza al interpretar su significado, sino que esa consideración adopta la lógica de una combinación bayesiana. Este procedimiento nos debería permitir combinar distintas expresiones verbales de manera óptima, por ejemplo, al formar una opinión propia a partir de varias opiniones ajenas. Si bien los resultados son aún preliminares, encontramos un rasgo distintivo y prometedor: al combinar dos expresiones verbales, los participantes le dieron más peso a aquella expresión cuya variabilidad era menor, es decir, más precisa. Esto es lo esperable si cada palabra transmite no solo un valor puntual de probabilidad sino toda una distribución de valores posibles, y si las personas combinan expresiones verbales siguiendo las reglas de la teoría de probabilidad.

Pero vale señalar que cuando se debe reportar la confianza en que una decisión es correcta, la optimalidad parece irse por la borda. Los participantes le asignaron a sus propias decisiones y juicios de confianza un peso mucho mayor del que deberían. Aún debemos entender mejor este fenómeno, pero en principio tampoco debería sorprendernos, considerando la cantidad de variables que pueden estar influenciando ese proceso: reticencia a cambiar la opinión, sesgos de confirmación, imposibilidad de validar la opinión ajena o hasta cuestiones específicas relacionadas al contexto digital en el que se estaba desarrollando el experimento.

Al final del día, lo único *seguro* es que la comunicación es un fenómeno de lo más complejo. Y que cuando decimos “seguro”, estamos diciendo algo que se parece a eso que estás pensando. Aunque no del todo.

Sumate en
eglc.ar/bancar