

Challenging local realism with human choices

The BIG Bell Test Collaboration*

A Bell test is a randomized trial that compares experimental observations against the philosophical worldview of local realism¹, in which the properties of the physical world are independent of our observation of them and no signal travels faster than light. A Bell test requires spatially distributed entanglement, fast and high-efficiency detection and unpredictable measurement settings^{2,3}. Although technology can satisfy the first two of these requirements^{4–7}, the use of physical devices to choose settings in a Bell test involves making assumptions about the physics that one aims to test. Bell himself noted this weakness in using physical setting choices and argued that human ‘free will’ could be used rigorously to ensure unpredictability in Bell tests⁸. Here we report a set of local-realism tests using human choices, which avoids assumptions about predictability in physics. We recruited about 100,000 human participants to play an online video game that incentivizes fast, sustained input of unpredictable selections and illustrates Bell-test methodology⁹. The participants generated 97,347,490 binary choices, which were directed via a scalable web platform to 12 laboratories on five continents, where 13 experiments tested local realism using photons^{5,6}, single atoms⁷, atomic ensembles¹⁰ and superconducting devices¹¹. Over a 12-hour period on 30 November 2016, participants worldwide provided a sustained data flow of over 1,000 bits per second to the experiments, which used different human-generated data to choose each measurement setting. The observed correlations strongly contradict local realism and other realistic positions in bipartite and tripartite¹² scenarios. Project outcomes include closing the ‘freedom-of-choice loophole’ (the possibility that the setting choices are influenced by ‘hidden variables’ to correlate with the particle properties¹³), the utilization of video-game methods¹⁴ for rapid collection of human-generated randomness, and the use of networking techniques for global participation in experimental science.

Bell tests, like Darwin’s studies of finches and Galileo’s observations of the moons of Jupiter, use empirical methods to address questions previously accessible only by other means, for example, by philosophy or theology¹⁵. Local realism—that is, realism plus relativistic limits on causation—was debated by Einstein and Bohr using metaphysical arguments and recently has been rejected by Bell tests^{4–7} that closed all technical ‘loopholes’. For example, the ‘detection-efficiency loophole’ describes the possibility that the observed statistics are inaccurate owing to selection bias, and is closed by using high-efficiency detection and statistical methods that include all trials in the analysis. Recent work on device-independent quantum information¹⁶ shows how Bell inequality violation (BIV) can also challenge causal determinism¹⁷, another topic that was formerly accessible only by metaphysics¹⁸. Central to both applications is the use of free variables to choose measurements: in the words of Aaronson¹⁹, ‘Assuming no preferred reference frames or closed timelike curves, if Alice and Bob have genuine ‘freedom’ in deciding how to measure entangled particles, then the particles must also have ‘freedom’ in deciding how to respond to the measurements’.

Previous Bell tests used physical devices^{20,21} to ‘decide’ for Alice and Bob, and thus demonstrated only a relation among physical processes: if some processes are ‘free’ in the required sense (see Methods, ‘Freedom’ in Bell tests), then other processes are similarly ‘free’. In the language of

strong Bell tests, this conditional relation leaves open the freedom-of-choice loophole (FOCL), which describes the possibility that ‘hidden variables’ influence the setting choices. Because we cannot guarantee such freedom within local realism, the tests must assume physical indeterminacy in the hidden-variable theory². Laboratory methods can tighten, but never close, this loophole^{2–6}.

Gallicchio, Friedman and Kaiser²² have proposed choosing settings by observing cosmic sources at the edge of the visible Universe. A BIV under such conditions could only be explained within local realism if events across history conspired to produce the measured outcomes^{23,24}. Bell himself argued that human choices could be considered ‘free variables’ in a Bell test⁸ (see Methods, John Stewart Bell on ‘free variables’), and noted the impracticality of using humans with 1970s’ technologies. Here we implement Bell’s idea, using modern crowd-sourcing, networking and gamification¹⁴ techniques. In this BIG Bell Test (BBT), Alice and Bob of Aaronson’s formulation are real people. Assuming no faster-than-light communication, such experiments can prove the conditional premise that if human will is free, there are physical events (the measurement outcomes in the Bell tests) that are intrinsically random, that is, impossible to predict²⁵. We note that although this argument in no way uses the theory of quantum mechanics, it arrives at one of the theory’s most profound claims. Intrinsic randomness supported by a BIV is central to device-independent quantum technologies^{16,26}.

It is perhaps surprising that human choices, which are known to contain statistical regularities²⁷, are suitably unpredictable for a Bell test. Recent works on the statistical analysis of Bell tests^{3,28,29} show that sequence randomness—that is, the absence of patterns and correlations in the sequence of choices—is not, *per se*, a requirement for the rejection of local realism. Rather, statistical independence of choices from the hidden variables that describe possible measurement outcomes is required (see Methods, ‘Freedom’ in Bell tests). This independence can fail in different ways, which are categorized into named loopholes. The ‘locality loophole’ describes the possibility that a choice at one station could influence a measurement result at the other station. The term ‘locality’ reflects one way of blocking this possibility, by space-like separation of the choice and measurement events (see Methods, Use of ‘freedom-of-choice loophole’ and ‘locality loophole’ in this work).

Patterns strongly affect statistical strength in experiments that aim to close the locality loophole by space-like separation—they allow current choices to be predicted from earlier choices, which have had more time to reach the distant measurement. As described below, the BBT tightens the locality loophole by using many independent experiments instead of space-like separation. Furthermore, the human capacity for free choice removes the need for assumptions about physical indeterminism, allowing the FOCL to be closed. Thus, although human choices show imperfect sequence randomness, they nonetheless enable a strong rejection of local realism with the BBT strategy.

A major obstacle to a Bell test with humans has been the difficulty in generating enough choices for a statistically significant test. A person can generate roughly three random bits per second, while a strong test may require millions of setting choices in a time span of minutes to hours, depending on the speed and stability of the experiment. To achieve such rates, we crowd-sourced the basis choices, recruiting about

*A list of authors and their affiliations appears at the end of the paper.

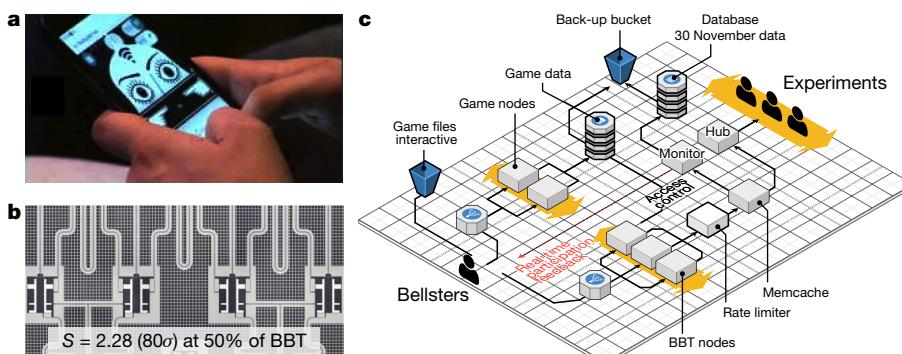


Fig. 1 | Structure of the BBT. **a**, Human participants, or Bellsters, enter 0s and 1s in an online video game that incentivizes sustained generation of unpredictable input. Image: ICFO/Maria Pascual (Kaitos Games). **b**, Experiments use Bellster-generated bits to control measurement-defining elements, such as wave-plates for photons or microwave pulses for matter qubits. Shown is a micrograph of superconducting qubits used in experiment (7), with the measured Clauser–Horne–Shimony–Holt Bell parameter S mid-way through the BBT. **c**, A cloud-based networking

100,000 participants, the ‘Bellsters’, over the course of the project. Each choice made by a participant, encoded as a bit (0 or 1), was entered in an internet-connected device, such as the participant’s mobile phone. Servers relayed these bits to the 13 experiments (see Fig. 1), which used them to determine individual settings without re-use (except experiment (2)). To encourage participants to contribute a larger number of more unpredictable bits, the input was collected in the context of a video game, The BIG Bell Quest (available at <https://museum.thebigbelltest.org/quest/>), implemented in JavaScript to run directly in a device’s web browser.

The BIG Bell Quest was designed to reward sustained, high-rate input of unpredictable bits, while being engaging and informative (see Methods, Gamification). An interactive explanation first describes quantum nonlocality and the role played by participants and experimenters in the BBT. The player is then tasked with entering a given number of unpredictable bits within a limited time. A machine learning algorithm (MLA) attempts to predict each input bit by modelling the user’s input as a Markov process and updating the model parameters using reinforcement learning (see Methods, Prediction engine). Scoring and level completion reflect the degree to which the MLA predicts the player’s input, motivating players to consider their own predictability and take conscious steps to reduce it. However, the MLA does not act as a filter, and all input is passed to the experiments. Bellsters’ input showed unsurprising deviations from ideal randomness²⁷; for example, $P(0) \approx 0.5237$ (bias towards 0), whereas adjacent bits show $P(01) + P(10) \approx 0.6406$ (excess of alternation).

Modern video-game elements were incorporated to boost engagement (animation and sound), encourage persistent play (progressive levels, ‘power-ups’, ‘boss battles’ and leaderboards) and to recruit new players (group formation and postings to social networks). Different level scenarios illustrate the key elements of the BBT (human input, global networking and measurements on quantum systems), and boss battles against the Oracle (see Methods) convey the conceptual challenge of unpredictability. Level completion is rewarded with (i) a report on how many bits from that level were used in each experiment running at that time, (ii) a ‘curious fact’ about statistics, Bell tests or the various experiments and, if the participant is lucky, (iii) one of several videos recorded in the participating laboratories, explaining the experiments. The game and BBT website are available in Chinese, English, Spanish, French, German, Italian and Catalan, making them accessible to roughly three billion speakers of these languages.

To synchronize participant activity with experimental operation, the Bell tests were scheduled to take place on a single day, Wednesday 30 November 2016. The date was chosen so that most schools worldwide would be in session and to avoid competing media events, such as the

system integrates the activities shown in **a** and **b**, serving game elements to Bellsters, distributing input bits to connected laboratories and providing in-game feedback about the experimental use of the player’s input. Through this system, Bellsters are given direct—although brief—control of the experimental apparatus, so that each measurement setting is determined by a single human choice, which is traceable to a given user and time of entry (see Methods, Networking strategy and architecture).

US presidential election. Participants were recruited through a variety of channels, including traditional and social media, as well as school and science museum outreach programmes, with each partner institution handling recruitment in their familiar geographical regions and languages. The media campaign focused on the nature of the experiment and the need for human participants. The press often communicated this with headlines such as “Quantum theory needs your help” (China Daily). A first, small campaign in early October 2016 began spreading the story by word of mouth and a second, large campaign on 29–30 November of that year was made to attract a wide participant base. The media campaign generated at least 230 headlines in the printed and online press, radio and television.

The data networking architecture of the BBT, shown in Fig. 1, includes elements of instant messaging and online gaming and is designed to efficiently serve a fluctuating number of simultaneous users that is not known in advance and could range from 10 to 100,000. A gaming component handles the BBT website, participant account management, delivery of the game code (JavaScript and video), score records and leaderboards. In parallel, a messaging component handles data conditioning, streaming to experiments and reporting of participant choices generated via the game. Horizontal scaling is used in both components; participants do not connect to the servers directly, but connect to dynamic load balancers that spread the input among a pool of servers that are dynamically scaled in response to the load. The timing of input bits (but not their values) was used to identify robot participants and remove their input from the data stream, although game operation was unchanged to avoid alerting the robots’ masters. A single, laboratory-side server received data from the participant-side servers, concatenated the user input and streamed it to the laboratories at laboratory-defined rates (see Methods, Networking strategy and architecture).

According to global time zoning, 30 November defines a 51-h window, from 0:00 UTC (coordinated universal time) plus 14 h (for example, Samoa) to 23:59 UTC minus 12 h (for example, Midway Island). Nevertheless, most participants contributed during a 24-h window centred on 18:00 UTC. The recruitment of participants was geographically uneven, with a notable failure to recruit large numbers of participants from Africa. Despite this, the latitude zones of Asia–Oceania, Europe–Africa and the Americas had comparable participation, which proved important for the experiment. As shown in Fig. 2, input from any single region dropped to low values during the local early morning but was compensated by high input from other regions, resulting in a sustained high global bitrate. Over the 12-h period from 09:00 UTC to 21:00 UTC on 30 November 2016, the input exceeded 10^3 bits per second, allowing the majority of the experiments to run at their full speeds. Several

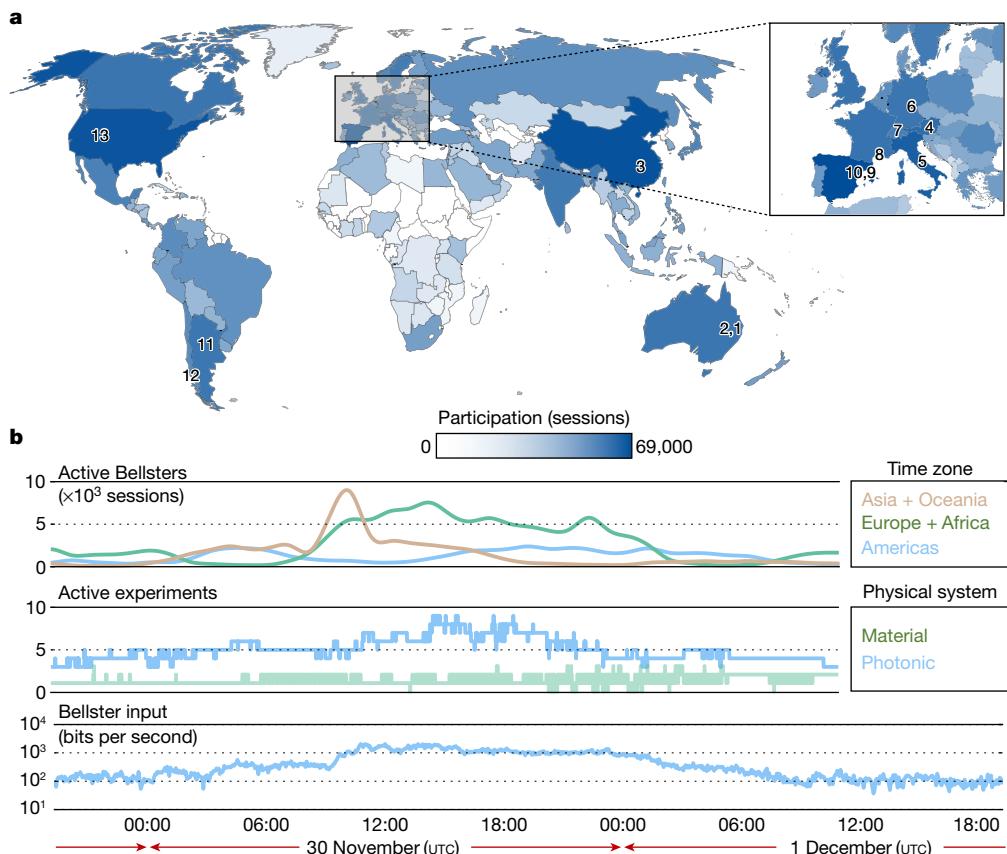


Fig. 2 | Geography and timing of the BBT. **a**, Locations of the 13 BBT experiments, ordered from east to west. The index numbers label the experiments, which are summarized in Table 1. Shading shows total sessions by country. Eight sessions from Antarctica are not shown. Map created by G. Colangelo using data from OpenStreetMaps, rendered in Wolfram Mathematica. **b**, Temporal evolution of the project. The top graph shows the number of live sessions versus time for different-continent groups, which exhibits a large drop in the local early morning in each region. The spike in the participation of the Asian group around 11:00 UTC coincides with a live-streamed event in Barcelona, hosted by

D. Jiménez and the CosmoCaixa science museum, re-broadcast live in Chinese by L.-F. Yuan and the University of Science and Technology of China (USTC). The middle graph shows the number of connected laboratories versus time, divided into experiments using only photons and experiments with at least one material component (such as atoms or superconductors). The bottom graph shows the input bitrate versus time. The data flow remains nearly constant despite regional variations, with Asian Bellsters handing off to Bellsters from the Americas in the critical period 12:00–00:00 UTC. Session data from Google Analytics.

experiments posted their results live on social networks. Owing to their high speeds, experiments (12) and (13) accumulated participants' input to use in short bursts. As determined by separate tests, for example of interference visibility, experiments (9) and (12) were not in condition to observe a BIV on the day of the event; they reported later results using stored input.

Because the Earth has a diameter of only 43 light milliseconds, human choices are too slow to be space-like separated from the measurements. This leaves open the locality loophole regarding the influence of choices on remote detection. Influence of Alice's measurement setting on Bob's detection (and vice versa) is nonetheless excluded by space-like separation in experiments (3) and (13) (see Supplementary Information). To tighten the locality loophole, we employ a strategy that we call the BIG test; namely, the use of many simultaneous Bell tests in widely separated locations and using different physical systems, with each experiment's apparatus constructed and operated by different experimental teams. The only hidden-variable theories that escape this tightening are those in which choices can simultaneously influence hidden variables in many differently constructed experiments to produce a BIV in each one. This strategy is strengthened by using the same bits in many experiments, as described above.

The set of 13 BBT experiments, including true Bell tests and other realism tests requiring free choice of measurement, are summarized in Table 1 and described in Supplementary Information. Experiments (1)–(5), (8) and (11)–(13) used entangled photon pairs, (6) used entanglement between single photons and single atoms, (9) used

entanglement between single photons and atoms and experiment (7) used entangled superconducting qubits. Experiments (7) and (13) used high-efficiency detection to avoid the fair-sampling assumption, thus closing simultaneously the detection-efficiency loophole and the FOCL. Experiment (5) demonstrated a violation of bilocal realism, and (10) violated a Bell inequality for multi-mode entanglement. Experiment (1) demonstrated quantum steering and (2) investigated temporal quantum correlations with a three-station measurement. Experiment (12) closed the post-selection loophole that is typically present in Bell tests based on energy-time entanglement. The analysis of the experimental results of (3) sets bounds on how well a measurement-dependent local model would have to predict Bellster behaviour to produce the observed results³⁰. Experiments (3), (4), (6) and (13) tested whether human-generated measurement choices gave different results from machine-generated ones. Most experiments observed statistically strong violations of their respective inequalities, supporting the rejection of local realism in a multitude of systems and scenarios.

In summary, on 30 November 2016, a set of 13 Bell tests and similar experiments using photons, single atoms, atomic ensembles and superconducting devices, demonstrated strong disagreement with local realism, using measurement settings chosen by tens of thousands of globally distributed human participants. The results also showed empirically that measurement-setting independence—here provided by human agency—is in strong disagreement with causal determinism^{17–19}, a topic formerly accessible only by metaphysics. The experiments reject local realism in a wide variety of physical systems and

Table 1 | Experiments carried out as part of the BBT, ordered by longitude, from east to west

Experiment	Lead Institution	Location	Entangled system	Rate (bps)	Inequality	Result	Stat. sig.
(1)	Griffith University	Brisbane, Australia	Photon polarization	4	$S_{16} \leq 0.511$	$S_{16} = 0.965 \pm 0.008$	57σ
(2)	University of Queensland & EQUS	Brisbane, Australia	Photon polarization	3	$ S \leq 2$	$S_{AB} = 2.75 \pm 0.05$	15σ
(3)	USTC	Shanghai, China	Photon polarization	10^3	PRBLG^{30}	$I_0 = 0.10 \pm 0.05$	N/A
(4)	IQOQI	Vienna, Austria	Photon polarization	1.61×10^3	$ S \leq 2$	$S_{HRN} = 2.639 \pm 0.008$	81σ
						$S_{QRN} = 2.643 \pm 0.006$	116σ
(5)	Sapienza	Rome, Italy	Photon polarization	0.62	$B \leq 1$	$B = 1.225 \pm 0.007$	32σ
(6)	LMU	Munich, Germany	Photon–atom	1.7	$ S \leq 2$	$S_{HRN} = 2.427 \pm 0.0223$	19σ
						$S_{QRN} = 2.413 \pm 0.0223$	18.5σ
(7)	ETHZ	Zurich, Switzerland	Transmon qubit	3×10^3	$ S \leq 2$	$S = 2.3066 \pm 0.0012$	$P < 10^{-99}$
(8)	INPHYNI	Nice, France	Photon time bin	2×10^3	$ S \leq 2$	$S = 2.431 \pm 0.003$	140σ
(9)	ICFO	Barcelona, Spain	Photon–atom ensemble	125	$ S \leq 2$	$S = 2.29 \pm 0.10$	2.9σ
(10)	ICFO	Barcelona, Spain	Photon multi-frequency bin	20	$ S \leq 2$	$S = 2.25 \pm 0.08$	3.1σ
(11)	CITEDEF	Buenos Aires, Argentina	Photon polarization	1.02	$ S \leq 2$	$S = 2.55 \pm 0.07$	7.8σ
(12)	UdeC	Concepción, Chile	Photon time bin	5.2×10^4	$ S \leq 2$	$S = 2.43 \pm 0.02$	20σ
(13)	NIST	Boulder, USA	Photon polarization	10^5	$K \leq 0$	$K = (1.65 \pm 0.20) \times 10^{-4}$	8.7σ

Descriptions of the experiments are given in Supplementary Information. Stat. sig., statistical significance; indicates the number of standard deviations assuming independent and identically distributed trials, unless otherwise indicated. Rate indicates the peak rate (in bits per second, bps) at which bits were used by the experiments. Owing to the limited rate of Bellster input, some experiments had dead times. B , K , S , S_{AB} , S_{BC} , S_{HRN} and S_{QRN} indicate Bell parameters for the respective experiments and S_{16} is the steering parameter (see Supplementary Information). I_0 indicates the minimum Pütz–Rosset–Barnea–Liang–Gisin measure of setting–choice independence, consistent with the observed BIV. USTC, University of Science and Technology of China; EQUS, Centre for Engineered Quantum Systems; IQOQI, Institute for Quantum Optics and Quantum Information; INFYNI, Institut de Physique de Nice; ICFO, Institut de Ciències Fotoniques; LMU, Ludwig-Maximilians-Universität; ETHZ, ETH Zurich; CITEDEF, Institute of Scientific and Technical Research for Defence; UdeC, University of Concepción; NIST, National Institute of Standards and Technology.

scenarios, set the groundwork for Bell-test-based applications in quantum information, introduce gamification to randomness generation and demonstrate global networking techniques by which hundreds of thousands of individuals can directly participate in experimental science.

Online content

Any Methods, including any statements of data availability and Nature Research reporting summaries, along with any additional references and Source Data files, are available in the online version of the paper at <https://doi.org/10.1038/s41586-018-0085-3>.

Received: 31 March 2017; Accepted: 7 March 2018;

Published online: 09 May 2018

1. Bell, J. S. On the Einstein–Podolsky–Rosen paradox. *Physics* **1**, 195–200 (1964).
2. Larsson, J.-Å. Loopholes in Bell inequality tests of local realism. *J. Phys. A* **47**, 424003 (2014).
3. Kofler, J., Giustina, M., Larsson, J.-Å. & Mitchell, M. W. Requirements for a loophole-free photonic Bell test using imperfect setting generators. *Phys. Rev. A* **93**, 032115 (2016).
4. Hensen, B. et al. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature* **526**, 682–686 (2015).
5. Giustina, M. et al. Significant-loophole-free test of Bell’s theorem with entangled photons. *Phys. Rev. Lett.* **115**, 250401 (2015).
6. Shalm, L. K. et al. Strong loophole-free test of local realism. *Phys. Rev. Lett.* **115**, 250402 (2015).
7. Rosenfeld, W. et al. Event-ready Bell test using entangled atoms simultaneously closing detection and locality loopholes. *Phys. Rev. Lett.* **119**, 010402 (2017).
8. Bell, J. in *Speakable and Unspeakable in Quantum Mechanics* Ch. 7 (Cambridge Univ. Press, Cambridge UK, 2004).
9. *The BIG Bell Test* <http://thebigbelltest.org>.
10. Farrera, P., Heinze, G. & da Riedmatten, H. Entanglement between a photonic time-bin qubit and a collective atomic spin excitation. *Phys. Rev. Lett.* **120**, 100501 (2018).
11. Wallraff, A. et al. Strong coupling of a single photon to a superconducting qubit using circuit quantum electrodynamics. *Nature* **431**, 162–167 (2004).
12. Carvacho, G. et al. Experimental violation of local causality in a quantum network. *Nat. Commun.* **8**, 14775 (2017).
13. Scheidl, T. et al. Violation of local realism with freedom of choice. *Proc. Natl Acad. Sci. USA* **107**, 19708–19713 (2010).
14. Sørensen, J. et al. Exploring the quantum speed limit with computer games. *Nature* **532**, 210–213 (2016).
15. Shimony, A. *Bell’s Theorem, The Stanford Encyclopedia of Philosophy* (ed. Zalta, E. N.) <https://plato.stanford.edu/entries/bell-theorem/> (Metaphysics Research Lab, Stanford Univ., Stanford, 2017).
16. Colbeck, R. *Quantum and Relativistic Protocols for Secure Multi-Party Computation*. PhD thesis, Cambridge Univ. (2007).
17. Hoefer, C. *Causal Determinism*, *The Stanford Encyclopedia of Philosophy* (ed. Zalta, E. N.) <https://plato.stanford.edu/entries/determinism-causal/> (Metaphysics Research Lab, Stanford Univ., Stanford, 2016).
18. Acín, A. & Masanes, L. Certified randomness in quantum physics. *Nature* **540**, 213–219 (2016).
19. Aaronson, S. Quantum randomness. *Am. Sci.* **102**, 266–271 (2014).
20. Abellán, C., Amaya, W., Mitrani, D., Pruneri, V. & Mitchell, M. W. Generation of fresh and pure random numbers for loophole-free Bell tests. *Phys. Rev. Lett.* **115**, 250403 (2015).
21. Fürst, M. et al. High speed optical quantum random number generation. *Opt. Express* **18**, 13029 (2010).
22. Gallicchio, J., Friedman, A. S. & Kaiser, D. I. Testing Bell’s inequality with cosmic photons: closing the setting-independence loophole. *Phys. Rev. Lett.* **112**, 110405 (2014).
23. Handsteiner, J. et al. Cosmic Bell test: measurement settings from Milky Way stars. *Phys. Rev. Lett.* **118**, 060401 (2017).
24. Wu, C. et al. Random number generation with cosmic photons. *Phys. Rev. Lett.* **118**, 140402 (2017).
25. Bera, M. N., Acín, A., Mitchell, M. W. & Lewenstein, M. Randomness in quantum mechanics: philosophy, physics and technology. *Rep. Prog. Phys.* **80**, 124001 (2017).
26. Pironio, S. et al. Random numbers certified by Bell’s theorem. *Nature* **464**, 1021–1024 (2010).
27. Bar-Hillel, M. & Wagenaar, W. A. The perception of randomness. *Adv. Appl. Math.* **12**, 428–454 (1991).
28. Bierhorst, P. A robust mathematical model for a loophole-free Clauser–Horne experiment. *J. Phys. A* **48**, 195302 (2015).
29. Elkouss, D. & Wehner, S. (Nearly) optimal P values for all Bell inequalities. *NJP Quantum Inf.* **2**, 16026 (2016).
30. Pütz, G., Rosset, D., Barnea, T. J., Liang, Y.-C. & Gisin, N. Arbitrarily small amount of measurement independence is sufficient to manifest quantum nonlocality. *Phys. Rev. Lett.* **113**, 190402 (2014).

Acknowledgements We are grateful to the many people and organizations who contributed to this project, starting with the Bellsters (details at <http://thebigbelltest.org>). We thank the Departament d’Ensenyament de la Generalitat de Catalunya, Ministerio de Educación, Cultura y Deporte of Spain, CosmoCaixa, Fundación Bancaria “la Caixa”, INTEF, Optical Society of America, European Centers for Outreach in Photonics (ECOP), Muncyt-Museo de Ciencia y Tecnología de Madrid, Investigación y Ciencia, Big Van, Crea Ciència, Tencent news, Micius Salon, Politécnico di Milano, University of Waterloo–Institute for Quantum Computing, Universidad Autónoma de Barcelona, Universidad Complutense de Madrid, Università degli Studi di Milano, Optica, EPS Young Minds, Real Sociedad Española de Física, Ajuntament de Barcelona, Ajuntament de Castelldefels, Università degli studi di Padova, Università degli studi dell’Insubria, CNRIFN Istituto di Fotonica e Nanotecnologia, Istituto d’Istruzione Superiore Carlo Livi, Webbile, and Kaitos Games. We are especially thankful for the recruitment efforts of the outreach departments at our institutions. We acknowledge financial support from CONICET; ANPCyT (Argentina), Australian Research Council Centre for Quantum Computation and Communication Technology (CE110001027, CE170100012); Australian Research Council and the University of Queensland (UQ) Centre for Engineered Quantum Systems (CE110001013, CE170100009). A.G.W. acknowledges the University of Queensland Vice-Chancellor’s Research and Teaching Fellowship (Australia); Austrian Academy of Sciences (OEAW), the Austrian Science Fund (FWF) (SFB F40 (FoQuS); CoQuS No. W1210-N16), the Austrian Federal Ministry of Education, Science and Research (BMBWF) and the University of Vienna (project QUESS) (Austria); MCTIC (Brazil), Barcelona City Hall; Generalitat de Catalunya (SGR 874 and 2014-SGR-1295; CERCA programme) (Catalonia), PIA CONICYT (grants PFB0824 and PAI79160083); FONDECYT (grants 1140635, 115101, 1160400, 3170596 and 11150325; Becas Chile); the Millennium Institute for Research in Optics (MIRO); Becas CONICYT (Chile), the National Fundamental Research Program (grant 2013CB336800); the

National Natural Science Foundation of China (grants 91121022, 61401441 and 61401443); the Chinese Academy of Science (Strategic Priority Research Program (B) XDB04010200); the 1000 Youth Fellowship programme; the National Natural Science Foundation of China (grant 11674193); the Science and Technology Commission of Shanghai Municipality (grant 16JC1400402) (China); ERC (grant agreements AQUMET 280169, 3DQUEST 307783, OSYRIS 339106, ERIDIAN 713682, QITBOX, QUOLAPS, QuLIMA and SuperQuNet 339871); ESA (contract number 4000112591/14/NL/US); FEDER; H2020 (QUIC 641122) and the Marie Skłodowska-Curie programme (grant agreement 748549) (European Commission); German Federal Ministry of Education and Research (projects QuOReP and Q.com-Q) (Germany); CONACyT graduate fellowship programme (Mexico), MINECO (FIS2014-60843-P, FIS2014-62181-EXP, SEV-2015-0522, FIS2015-68039-P, FIS2015-69535-R and FIS2016-79508-P; Ramon y Cajal fellowship programme; TEC2016-75080-R); the ICFOnest+ international postdoctoral fellowship programme (Spain); the Knut and Alice Wallenberg Foundation (project 'Photonic Quantum Information') (Sweden); NIST (USA), AXA Chair in Quantum Information Science; FQXi Fund; Fundació Privada CELLEX; Fundació Privada MIR-PUIG; the CELLEX-ICFO-MPQ programme; Fundació Catalunya-La Pedrera; and the International PhD-fellowship programme 'la Caixa'-Severo Ochoa.

Reviewer information *Nature* thanks S. Maniscalco (with F. Cosco, S. Hamedani Raja, H. Lyyra, J. Nokkala, B. Sokolov, N. Talarico and J. Teittinen), J. Sherson and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Author contributions C.A., instigator; M.W.M. (contact author), project leader. Coordination, gamification and networking (ICFO); S.C., general supervision; M.M.-P., project management; M.G.-M., F.A.B. and C.A., game design; J.T., prediction engine; A.H., M.G.-M. and F.A.B., Bellster recruitment and engagement strategy, design and execution; C.A., web infrastructure and networking; M.W.M., C.A. and J.T., writing of the main manuscript with input from all authors. Experiments: (1) G.J.P., R.B.P. (contact author), F.G.J., M.M.W. and S.S., experiment design and execution. (2) A.G.W. and M.R. (contact author), experiment design and execution. (3) J.-W.P. (contact author) and Q.Z. (contact author), supervision; J.-W.P., Q.Z., X.M., X.Y. and Y.L., experiment conception and design; Z.W., L.Y., H.L. and W.Z., superconducting nanowire single-photon detector (SNSPD) fabrication and characterization; J.Z., SNSPD maintenance; M.-H.L., C.W. and Y.L., photon source design and characterization; J.-Y.G. and Y.L., software design and deployment; X.Y., protocol analysis; X.Y. and Y.L., data analysis; J.-W.P., Q.Z., C.W., X.Y. and Y.L., manuscript preparation with input from all authors. (4) T.S. (contact author), A.Z. and R.U., supervision, conception and coordination; B.L., J.H. and D.R., experiment execution and analysis. (5) G.C., L.S., F.A., M.B. and F.S. (contact author), experiment execution and analysis; R.C., theory support. (6) H.W., W.R. (contact author), K.R., R.G. and D.B., experiment design and execution. (7) J. He, P.K., Y.S., C.K.A., A.B., S.K., P.M., M.O., T.W., S.G., C.E. and A.W. (contact author), experiment design and execution. (8) S.T. (contact author), T.L., F.K., G.S., P.V. and O.A., experiment design and execution. (9) H.d.R. (contact author), P.F. and G.H., experiment design and execution. (10) H.d.R. (contact author), A.S., A.L., M.M., D.R., O.J. and A.M., experiment design and execution; D.C. and A. Ac., theory support. (11) M.A.L. (contact author), coordination and server communication; L.T.K., I.H.L.G. and A.G.M., experiment design and execution; C.T.S. and A.B., input data formatting; L.T.K., I.H.L.G., A.G.M., C.T.S., A.B. and M.A.L., data analysis. (12) G.X. (contact author), coordination; F.T., optical setup; P.G., A. Al., J.F., A. Cu. and G.C., optical setup support; J.C., electronics design and implementation; J.C. and F.T., experiment execution; D.M., software; G.L., P.M. and F.S., experimental support; A. Ca., theory support; J.C., E.S.G. and J.-Å.L., data analysis. (13) L.K.S. (contact author), S.N., M.S., O.S.M.-L., T.G., S.G., P.B., E.K. and R.M., experiment design, execution and analysis.

Competing interests The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41586-018-0085-3>.

Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-018-0085-3>.

Reprints and permissions information is available at <http://www.nature.com/reprints>.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

The BIG Bell Test Collaboration

C. Abellán¹, A. Acín^{1,2}, A. Alarcón^{3,4}, O. Alibart⁵, C. K. Andersen⁶, F. Andreoli⁷, A. Beckert⁸, F. A. Beduini¹, A. Bendersky⁸, M. Bentivegna⁷, P. Bierhorst⁹, D. Burchardt¹⁰, A. Cabello¹¹, J. Cariñé^{3,4,12}, S. Carrasco¹, G. Carvacho⁷, D. Cavalcanti¹, R. Chaves¹³, J. Cortés-Vega^{3,12}, A. Cuevas⁷, A. Delgado^{3,12}, H. de Riedmatten^{1,2}, C. Eichler⁶, P. Farrera¹, J. Fuenzalida^{3,12,14}, M. García-Matos¹, R. Garthoff¹⁰, S. Gasparinetti⁶, T. Gerrits⁹, F. Ghafari Jouneghani^{16,17}, S. Glancy⁹, E. S. Gómez^{3,12}, P. González^{3,12}, J.-Y. Guan^{18,19}, J. Handsteiner^{14,15}, J. Heinsoo⁶, G. Heinze¹, A. Hirschmann¹, O. Jiménez⁴, F. Kaiser⁵, E. Knill⁹, L. T. Knoll^{20,21}, S. Krinner⁶, P. Kurpiers⁶, M. A. Larotonda^{20,21}, J.-Å. Larsson²², A. Lenhard¹, H. Li^{23,24}, M.-H. Li^{18,19}, G. Lima^{3,12}, B. Liu^{14,25}, Y. Liu^{18,19}, I. H. López Grande^{20,21}, T. Lunghi¹⁵, X. Ma²⁶, O. S. Magaña-Loaiza⁹, P. Magnard⁶, A. Magnoni²¹, M. Martí-Prieto¹, D. Martínez^{3,12}, P. Mataloni⁷, A. Mattar¹, M. Mazzera¹, R. P. Mirin⁹, M. W. Mitchell^{1,2*}, S. Nam⁹, M. Oppliger⁶, J.-W. Pan^{18,19}, R. B. Patel^{16,17}, G. J. Pryde^{16,17}, D. Rauch^{14,15}, K. Redeker¹⁰, D. Rieländer¹, M. Ringbauer^{27,28}, T. Roberson^{27,28}, W. Rosenfeld¹⁰, Y. Salathé⁶, L. Santodonato⁷, G. Sauder⁵, T. Scheidl^{14,15}, C. T. Schmiegelow²¹, F. Sciarriño⁷, A. Seri¹, L. K. Shalm⁹, S.-C. Shi²⁹, S. Slussarenko^{16,17}, M. J. Stevens⁹, S. Tanzilli⁵, F. Toledo^{3,12}, J. Tura^{1,30}, R. Ursin^{14,15}, P. Vergyris⁵, V. B. Verma⁹, T. Walter⁶, A. Wallraff⁶, Z. Wang^{23,24}, H. Weinfurter^{10,30}, M. M. Weston^{16,17}, A. G. White^{27,28}, C. Wu^{18,19}, G. B. Xavier^{3,4,22}, L. You^{23,24}, X. Yuan²⁶, A. Zeilinger^{14,15}, Q. Zhang^{18,19}, W. Zhang^{23,24}, J. Zhong²⁹

¹ICFO - Institut de Ciències Fotoniques, The Barcelona Institute of Science and Technology, Barcelona, Spain. ²ICREA - Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain.

³Millennium Institute for Research in Optics, Universidad de Concepción, Concepción, Chile.

⁴Departamento de Ingeniería Eléctrica, Universidad de Concepción, Concepción, Chile. ⁵CNRS UMR 7010, Institut de Physique de Nice (INPHYNI) Université Côte d'Azur, Nice, France.

⁶Departamento de Física, ETH Zurich, Zurich, Switzerland. ⁷Dipartimento di Fisica, Sapienza Università di Roma, Rome, Italy. ⁸Dipartimento de Computación, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires and Instituto de Investigación en Ciencias de la Comunicación (ICC), CONICET, Buenos Aires, Argentina. ⁹National Institute of Standards and Technology, Boulder, CO, USA. ¹⁰Ludwig-Maximilians-Universität, Munich, Germany.

¹¹Departamento de Física Aplicada II, Universidad de Sevilla, Seville, Spain. ¹²Departamento de Física, Universidad de Concepción, Concepción, Chile. ¹³International Institute of Physics, Federal University of Rio Grande do Norte, Natal, Brazil. ¹⁴Institute for Quantum Optics and Quantum Information (IQOQI), Austrian Academy of Sciences, Vienna, Austria. ¹⁵Vienna Center for Quantum Science & Technology (VCQ), Faculty of Physics, University of Vienna, Vienna, Austria. ¹⁶Centre for Quantum Computation and Communication Technology, Griffith University, Brisbane, Queensland, Australia. ¹⁷Centre for Quantum Dynamics, Griffith University, Brisbane, Queensland, Australia. ¹⁸Shanghai Branch, National Laboratory for Physical Sciences at Microscale and Department of Modern Physics, University of Science and Technology of China, Shanghai, China. ¹⁹Shanghai Branch, CAS Center for Excellence and Synergetic Innovation Center in Quantum Information and Quantum Physics, University of Science and Technology of China, Shanghai, China. ²⁰Departamento de Investigaciones en Láseres y Aplicaciones (DEILAP), CITEDEF and CONICET, Buenos Aires, Argentina. ²¹Departamento de Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires and Instituto de Física de Buenos Aires (IFIBA), CONICET, Buenos Aires, Argentina. ²²Department of Electrical Engineering, Linköping University, Linköping, Sweden. ²³State Key Laboratory of Functional Materials for Informatics, Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai, China. ²⁴CAS Center for Excellence in Superconducting Electronics, Shanghai, China. ²⁵School of Computer, NUDT, Changsha, China. ²⁶Center for Quantum Information, Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing, China. ²⁷Centre for Engineered Quantum Systems, University of Queensland, Brisbane, Queensland, Australia. ²⁸School of Mathematics and Physics, University of Queensland, Brisbane, Queensland, Australia. ²⁹Purple Mountain Observatory and Key Laboratory of Radio Astronomy, Chinese Academy of Sciences, Nanjing, China. ³⁰Max-Planck-Institut für Quantenoptik, Garching, Germany. *e-mail: morgan.mitchell@icfo.eu

METHODS

Local realism, Bell parameters, Bell inequalities. In their 1935 article³¹, Einstein, Podolsky and Rosen employed notions of locality (actions or observations in one location do not have immediate effects in other locations) and realism (observables have values even if we do not observe them) to argue that quantum theory was incomplete and could, in principle, be supplemented with information about which outcomes actually occur in any given run of an experiment. Bell formalized these notions by defining local hidden-variable models (LHVMs), a class of non-quantum theories that are simultaneously local and realistic. We consider the simplest case, of two systems measured by two observers, Alice and Bob. We write $x(y)$ to represent Alice's (Bob's) measurement setting, $a(b)$ to represent their measurement outcomes, and λ to represent the hidden variable—something that we cannot measure, but we include in the model to explain why a and b take on particular values. The predictions of any such bipartite LHVM are given by

$$P(a, b|x, y) = \sum_{\lambda} P(a|x, \lambda)P(b|y, \lambda)P(\lambda) \quad (1)$$

where $P(\cdot | \cdot)$ indicates a conditional probability. That is, the probability of getting outcome a and b when Alice and Bob measure x and y , respectively, can be expressed in terms of the local conditional probabilities $P(a|x, \lambda)$ and $P(b|y, \lambda)$. λ is averaged over, because of our ignorance of the value of this hidden variable. If the probabilities $P(a|x, \lambda)$ and $P(b|y, \lambda)$ are restricted to 0 or 1, then we have a deterministic LHVM, in which λ , x and y fully determine the outcomes a and b . Such LHVMs are explicitly realistic in the Einstein–Podolsky–Rosen sense. Locality is also explicit in the model. For example, $P(a|x, \lambda)$ depends on neither b nor y , so that the events at Bob's station have no influence on Alice's measurement outcome a . A mathematical notion of 'freedom' is implicit in the LHVM; x and y are included as free parameters and not, for example, as functions of λ . If $P(a|x, \lambda)$ and $P(b|y, \lambda)$ are allowed to take intermediate values, we speak of a non-deterministic LHVM. Because the unknown $P(\lambda)$ can take on intermediate values, deterministic and non-deterministic LHVMs are equivalent, and from here on we do not use this distinction.

This class of models, which by construction embody the Einstein–Podolsky–Rosen assumptions, was shown by Bell to be incapable of reproducing the predictions of quantum mechanics. For example, if Alice's and Bob's local systems are spin-1/2 particles in a singlet state, then their measurements (assumed to be ideal) will agree—that is, they will show that they are both spin-up or both spin-down—with probability

$$P(\text{agree}|\phi_a, \phi_b) \equiv P(\uparrow, \uparrow|\phi_a, \phi_b) + P(\downarrow, \downarrow|\phi_a, \phi_b) = \sin^2\left(\frac{\phi_b - \phi_a}{2}\right)$$

where $\phi_b - \phi_a$ is the angle between Alice's and Bob's analysis directions. Equation (1) cannot reproduce all the features of this distribution. No choice of $P(a|x, \lambda)$, $P(b|y, \lambda)$ and $P(\lambda)$ can give a probability $P(a, b|x, y)$ that simultaneously depends on the difference $\phi_b - \phi_a$ has high visibility (ranging from 0 to 1) and is sinusoidal.

This difference is efficiently captured by Bell inequalities. A Bell parameter is a linear combination of conditional probabilities $P(a, b|x, y)$, and a Bell inequality indicates the bounds (within the class of LHVMs) of a Bell parameter. Typically, the Bell inequalities of interest are those that are not obeyed by quantum mechanics, that is, those for which quantum correlations can be strong enough to violate the Bell inequality. Bell's theorem shows that there are such inequalities and thus that quantum mechanics cannot be 'completed' by hidden variables.

Bell inequalities also enable experimental tests of local realism. A Bell test is an experiment that makes many spatially separated measurements with varied settings to obtain $P(a, b|x, y)$ estimates that appear in a Bell parameter. If the observed Bell parameter violates the inequality, one can conclude that the measured systems are not governed by any LHVM. We note that this conclusion is always statistical and typically takes the form of a hypothesis test, leading to a conclusion of the form 'assuming nature is governed by local realism, the probability to produce the observed BIV (or a stronger one) is $P(\text{observed or stronger} | \text{local realism}) \leq p$ '. This p value is a key indicator of statistical significance in Bell tests.

'Freedom' in Bell tests. The use of the term 'free' to describe the choices in a Bell test derives more from mathematical usage than from its usage in philosophy, although the two are clearly related. Bell⁸ (see Methods, John Stewart Bell on 'free variables') states that his use of 'free will' reflects the notion of 'free variables', that is, externally given parameters in physical theories, as opposed to dynamical variables that are determined by the mathematical equations of the theory.

Any realistic LHVM is described by equation (1). The mathematical requirements for the relevant 'freedom' are made evident by a more general description, in which the local realistic model includes also x and y and thus specifies the joint probability

$$P(a, b, x, y) = \sum_{\lambda} P(a|x, \lambda)P(b|y, \lambda)P(x, y|\lambda)P(\lambda) \quad (2)$$

Using the Kolmogorov definition of conditional probability, $P(A, B) = P(A|B)P(B)$, we find that equation (2) is reduced to equation (1), provided that $P(x, y|\lambda) = P(x, y)$, that is, provided that the settings are statistically independent of the hidden variables. According to Bayes' theorem, this condition can be written as $P(\lambda|x, y) = P(\lambda)$ and $P(x, y, \lambda) = P(x, y)P(\lambda)$. This condition is known in the literature as the freedom-of-choice assumption, although it implies more than just free choices. A more accurate term might be 'measurement setting–hidden-variable independence'. We note that this condition does not require x to be independent of y , nor does it require $P(x, y)$ to be unbiased. Similar observations emerge from the more complex calculations needed to assign p values to observed data in Bell tests³.

The above describes the sense in which the basis choices should be 'free'. The desideratum is independence from the hidden variables that describe the particle behaviours, keeping in mind that the choices and measurements could, consistent with relativistic causality, be influenced by any event in their backward light-cones. Because the setting choices and the measurements will always have overlapping backward light-cones, it is impossible to rule out the possibility of a common past influence through space-time considerations. If human choices are free, however, such influences are excluded. We also note that complete independence is not required, although the tolerance for interdependence can be low^{30,32,33}. The theory that the entire experiment, including choices and outcomes, is pre-determined by initial conditions is known as superdeterminism. Superdeterminism cannot be tested³⁴.

A very similar concept of 'freedom' applies to the entangled systems measured in a Bell test. A BIV with free choice and under strict locality conditions implies either indeterminacy of the measurement outcomes or faster-than-light communications and thus closed time-like curves^{16,18}. If Bob's measurement outcome is predictable based on information available to him before the measurement, and if it also satisfies the condition for a BIV (namely, a strong correlation with Alice's measurement outcome that depends on his measurement choice), then Bob can influence the statistics of Alice's measurement outcome and thus communicate with her, despite being space-like separated from her. Considering that Bob could in principle have information on any event that occurs in his backward light cone, this implies (assuming no closed time-like curves) that Bob's measurement outcome must be statistically independent of all prior events.

In this way, we see that 'freedom', understood as a behaviour statistically independent of prior conditions, appears twice in a Bell test; first as a requirement on the setting choices and second as a conclusion about the nature of measurement outcomes in entangled systems. These two are linked, in that the second can be demonstrated if the first is present.

Previous tests using physical randomness generators to choose measurement settings have thus demonstrated a relationship between physical processes, showing for example^{4,20} that if spontaneous emission is 'free', then the outcomes of measurements on entangled electrons are also 'free'. By using humans to make the choices, we translate this to the human realm, showing that, in the words of Conway and Kochen³⁵, "if indeed there exist any experimenters with a modicum of free will, then elementary particles must have their own share of this valuable commodity". Here, 'experimenters' are those who choose the settings, that is, the Bellsters. See the main text for a discussion of the locality loophole when using humans.

John Stewart Bell on 'free variables'. A brief but informative source for Bell's positions on setting choices is an exchange of opinions with Clauser, Horne and Shimony in articles titled 'The theory of local beables'⁸ and 'Free variables and local causality'³⁶. In the first of these articles, Bell very briefly considers using humans to choose the measurement settings

It has been assumed [in deriving Bell's theorem] that the settings of instruments are in some sense free variables - say at the whim of experimenters - or in any case not determined in the overlap of the backward light cones.

whereas in the second article, Bell defends this choice of method and compares it against 'mechanical'—that is, physical—methods of choosing the settings

Suppose that the instruments are set at the whim, not of experimental physicists, but of mechanical random number generators. Indeed it seems less impractical to envisage experiments of this kind...

Bell proceeds to consider the strengths and weaknesses of physical random-number generators in Bell tests, offering arguments about why under 'reasonable' assumptions, physical random-number generators might be trusted. Nonetheless, he concludes

Of course it might be that these reasonable ideas about physical randomizers are just wrong - for the purpose at hand. A theory might appear in which such conspiracies

inevitably occur, and these conspiracies may then seem more digestable than the non-localities of other theories.

In summary, Bell distinguishes different levels of persuasiveness, noting that physical setting-generators, while having the required independence in many local realistic theories, cannot be expected to do so in all such theories. In contemporary terminology, what he argues is that physical-setting generators can only tighten, not close, the FOCL.

Bell also defends his use of the concept of ‘free will’ in a physics context, something that had been criticized by Clauser, Horne and Shimony. Bell writes

Here I would entertain the hypothesis that experimenters have free will [...] it seems to me that in this matter I am just pursuing my profession of theoretical physics.

A respectable class of theories, including contemporary quantum theory as it is practiced, have ‘free’ external variables in addition to those internal to and conditioned by the theory. These variables provide a point of leverage for ‘free willed experimenters’, if reference to such hypothetical metaphysical entities is permitted. I am inclined to pay particular attention to theories of this kind, which seem to me most simply related to our everyday way of looking at the world.

Of course there is an infamous ambiguity here, about just what and where the free elements are. The fields of Stern-Gerlach magnets could be treated as external. Or such fields and magnets could be included in the quantum mechanical system, with external agents acting only on the external knobs and switches. Or the external agents could be located in the brain of the experimenter. In the latter case the setting of the experiment is *not* itself a free variable. It is only more or less correlated with one, depending on how accurately the experimenter effects his intention.

It is clear from the last three sentences that Bell considers human intention—that is, human free will—to be a ‘free variable’ in the context of the discussion. That is, he believes human intention fulfils the assumptions of Bell’s theorem, as do experimental settings faithfully derived from human intention.

Use of ‘freedom-of-choice loophole’ and ‘locality loophole’ in this work. As noted above, a statistical condition used to derive Bell’s theorem is $P(x, y, \lambda) = P(x, y)P(\lambda)$, where x and y are choices and λ describes the hidden variables. This statistical condition, known as the freedom of choice assumption, does not distinguish between three possible scenarios of influence: the condition could fail if the choices influence the hidden variables, if the hidden variables influence the choices or if a third factor influences both the choices and the hidden variables^{2,3,13}. According to Bayes’ theorem, equivalent forms are $P(x, y | \lambda) = P(x, y)$, which expresses the fact that knowing λ does not give information about (x, y) , and $P(\lambda | x, y) = P(\lambda)$, which expresses the fact that knowing (x, y) does not give information about λ . The latter relationship makes clear that influence (in either direction) is incompatible with the freedom of choice assumption. The term used for this condition should not be taken literally; the condition can be false even if the choices are fully free, in the sense of being independent of all prior conditions. This occurs, for example, if the choices are made freely but then influence the hidden variable.

By long tradition, ‘locality loophole’ is the term given to the possibility of influence from Alice’s (Bob’s) choices or measurements to Bob’s (Alice’s) measurement outcomes. The term ‘freedom-of-choice loophole’ was introduced in Scheidl et al.¹³ to describe the influence of hidden variables on choices. The exact definition was “the possibility that the settings are not chosen independently from the properties of the particle pair”. We note that this formulation centres on the act of choosing and its independence, which (assuming relativistic causality, an element of local realism) can only be violated by influences from past events, not future events. These loophole definitions employ the concept of influence, which is directional, to explain how the non-directional relation of independence can be broken. Similarly directional definitions have recently been applied to experiments using cosmic sources^{23,24}.

Our use of the term in this paper follows the definition of Scheidl et al.¹³ described above: FOCL refers to the possibility of choices being influenced by any combination of hidden variables or other factors within the backward light cone of the choice, whereas the possibility of choices influencing hidden variables, which necessarily occur in the forward light cone of the choice, is included in the locality loophole. Such a division, in addition to fitting the common-sense notion of free choice, avoids counting a single possible channel of influence in both FOCL and the locality loophole.

Status of the FOCL. After recent experiments simultaneously closed the locality, detection-efficiency, memory, timing and other loopholes^{4–7}, the FOCL remains open. Space-time considerations can eliminate the possibility of such influence from the particles^{5,6,13,37}, or from other space-time regions^{4,7}, to the choices, but not the possibility of a sufficiently early prior influence on both the choices and the particles. To motivate freedom of choice in this scenario, well characterized physical randomizers^{20,21} have been used to choose settings.

In some experiments^{4–6}, the physical assumption is that at least one of (i) spontaneous emission, (ii) thermal fluctuations and (iii) classical chaos²⁰ is uninfluenced

by prior events and thus unpredictable even within local realistic theories. In other experiments^{7,13,37,38}, the physical assumption is that photodetection is similarly uninfluenced. While still requiring a physical assumption, and thus not closing the FOCL, this strategy tightens the loophole in various ways. First, by using space-like separation to rule out influence from certain events (for example, entangled pair creation) and from defined space-time regions. Second, by using well characterized randomness sources, for which the setting choice is known to faithfully derive from a given physical process, it avoids assumptions about the predictability of side-channel processes. Third, in the case of refs.^{4–6,20}, by using a physical variable that can be randomized by each of several processes, the required assumption is reduced from ‘ x is uninfluenced’ to ‘at least one of x, y and z is uninfluenced’.

Prediction engine. Generation of random sequences by humans has been studied in the field of psychology for decades^{27,39}. Early studies showed that humans perform poorly when asked to produce a random sequence, choosing in a biased manner and deviating from a uniform distribution. In ref.⁴⁰ it was shown that humans playing competitive, zero-sum games that reward uniform random choices tend to produce sequences with fewer identifiable biases. One such game is matching pennies: players must simultaneously choose between heads or tails; one player wins if the results are equal and the other player wins if the results are different. This is a standard two-person game used in game theory⁴¹ (see also ref.⁴²) with a mixed-strategy Nash equilibrium: as both players try to outguess each other, by behaving randomly, they do not incentivize the other player to change their strategy.

The BIG Bell Quest reproduces the coin-matching game, with an MLA playing the part of the opponent. The MLA operates on simple principles that human players could employ; it maintains a model of the tendencies of the opponent, noting, for example, that after choosing 0 and 0, the opponent usually chooses 1 as the next bit. The MLA strategy operates with very little memory, mirroring the limited short-term memory of humans.

Formally, we write $x_i \in \{0, 1\}$ for the i th input bit, $S_k \equiv \{x_1, \dots, x_k\}$ for a sequence of k input bits, and $\mathbf{x}_j^{(L)} \equiv \{x_j, x_{j+1}, \dots, x_{j+L-1}\}$ for a length- L sub-sequence of S_k starting from bit j . Given S_k as input, the algorithm predicts the value of $x_{k+1} \in \{0, 1\}$ that maximizes

$$\max_{L \in [1, L_{\max}]} f_L(x_{k+1} | \mathbf{x}_{k-L+1}^{(L)}) \quad (3)$$

where f estimates the probability of x following \mathbf{x} in S_k

$$f_L(x | \mathbf{x}) = \frac{\#\{\mathbf{x}_i^{(L+1)} : \mathbf{x}_i^{(L+1)} = \mathbf{x} || \mathbf{x}, 1 \leq i \leq k-L\}}{\#\{\mathbf{x}_i^{(L)} : \mathbf{x}_i^{(L)} = \mathbf{x}, 1 \leq i \leq k-L\}} \quad (4)$$

where $||$ indicates concatenation and $\#A$ indicates the number of elements in set A . Equations (3) and (4) mean that the prediction algorithm identifies the most frequently input sequence, of length $L_{\max} + 1$ or shorter, that the player can form when adding the bit x_{k+1} , and predicts that the player will produce the bit needed to complete that sequence. L_{\max} is chosen to be 3, reflecting the limited memory of a human opponent in the coin-matching game.

In equation (4), the estimator $f_L(x | \mathbf{x})$ of the probability that x follows \mathbf{x} in S_k is based on modelling the user’s input as a Markov process⁴³. The MLA keeps a running estimate $T_L(z | y)$, updated with each new input bit, of the matrix describing the probabilities of transitions among length- L words, from word y to word z . The estimates are simply the observed frequency of transitions in S_k . The MLA then obtains f_L as a marginal probability distribution: the probability of the first bit of z being x , conditioned on the tail of S_k being y (See Extended Data Fig. 1).

Networking strategy and architecture. The BBT required reliable, robust and scalable operation of two linked networking tasks: providing the BIG Bell Quest video game experience, as well as live aggregation and streaming of user input to the running experiments. From a networking perspective, the latter task resembles an instant messaging service, with the important asymmetry that messages from a large pool of senders (the Bellstars) are directed to a much smaller pool of recipients (the laboratories). The network architecture, shown in Fig. 1c, was implemented using Amazon Web Services IaaS (Infrastructure as a Service) products.

In the messaging component, we employed a two-layered architecture, shown in Fig. 1c. In the first layer, BBT nodes received input bits from the users and performed a real-time health check (described below) to block spamming by robot participants. The data were then sent to the second layer, a single-instance hub node that concatenated all the bits from the first stage and distributed them to the laboratories. The communication between the two layers was implemented using a memcache computation node to maximize speed and to simplify the synchronization between the two layers.

The gaming task was handled by a single layer of game nodes and a database. To protect the critical messaging task from possible attacks on the gaming components, we used separate instances to handle back-end gaming tasks, such as user

information and rankings, and to handle back-end tasks in the messaging chain, such as data logging. Load balancers, networking devices that distribute incoming traffic to a scalable pool of servers, were used in both the gaming and messaging front ends to avoid overloading. This design pattern is known as horizontal scaling, and is a common practice in scalable cloud systems.

This specific architecture was not available as a standard service from web service providers, but was readily constructed from standard component services. The architecture is not specific to the low-bitrate manual input collected for the BBT and could easily be adapted to other data that can be collected by personal devices, for example, audio or acceleration. The architecture was designed to solve a problem specific to time-limited projects with crowd-sourced input: owing to the single-day nature of the BBT, the unknown number and geography of the participants and the possibility of hackers or spammers, it was not practical to test the system under full-load conditions before the event itself. The two-layer architecture helped maintain all the critical servers isolated and independently operating and helped us to scale up the system smoothly when traffic increased. In the event, the traffic surpassed our initial estimates, and we deployed three additional BBT nodes at 09:00 UTC (when Europe was waking up) with no interruption of service. Such scaling up is expected to be critical for projects (for example, ref.⁴⁴) that combine laboratory experimentation, which tends to be time-limited because of stability and resource considerations, with crowd-sourcing, which usually entails unknown and fluctuating demand.

BBT nodes. The first layer of computing resources received data from Bellsters—or, more precisely, from The BIG Bell Quest running in browsers on their computers and devices. A variable number of servers running the same software functionalities were placed behind a pre-warmed load balancer that was prepared to support up to 10,000 simultaneous connections. Users connected to the load balancer via a public URL end-point and sent the data from their browsers using websocket connections. This first layer of servers aggregated the data from each connection (that is, from each user) in independent buffers during a 0.5-s interval.

A simple but important ‘health check’ was performed to identify and block high-speed robotic participants. If a given user contributed more than 10 bits in a single 0.5-s interval, corresponding to a rate of more than 20 keypresses per second, the user account was flagged as being non-human, and all subsequent input from that user was removed from the data stream. No feedback was provided to the users if their account was flagged, to avoid leaking information about the blocking mechanism. This method could potentially ban honest users because of networking delays and other timing anomalies, but was necessary to prevent the greater risk of the data stream being flooded with robotic input.

Hub node. The hub node aggregated the data from all the BBT nodes and also handled the connection to the laboratories. In contrast to the BBT nodes, which had to service connections from an unknown and rapidly changing number of users, the hub node aggregated data from a small and relatively stable number of trusted instances. Overall, the two-layer design simplified the networking task of delivering input from a large and variable number of users to end points (the laboratories) receiving aggregated data streams at variable rates.

Laboratories connected to the hub requested to receive random bits from the Bellsters, which were distributed after aggregating four of the 0.5-s batches from the BBT nodes, that is, in intervals of 2 s. At the end of each interval, bits were sent to each running experiment. If an experiment had requested N bits, it was sent bits $\{x_0, x_1, \dots, x_{N-1}\}$, that is, the earliest bits to arrive in that interval. Thus, the same bits were used simultaneously in many experiments. This helped to tighten the locality loophole, because an influence from the input bits on the measurements would have to operate in the same way in several independent experiments and in several locations. With the exception of experiments (12) and (13), the sent bits were used within the next 2-s interval. To run faster than the Bellster input rate, experiments (12) and (13) operated in burst mode, accumulating bits for a specific time and then rapidly using them. As with the BBT instances, these connections were established using websocket connections. When connecting to the hub node, the laboratories specified their bitrate requirement, which could be dynamically changed. The hub node then sent a stream of Bellster-generated bits at the requested rate. Archived bits from BBT participation prior to 30 November were cached locally at the laboratories, so that experiments could continue to run even if the flow of real-time bits was insufficient or interrupted. During the event, the flux of live bits was sufficient and no experiments used these pre-distributed bits.

Memcache node. The interface between the BBT nodes and the hub instance was implemented using a memcache node. While adding an extra computing resource slightly increased the complexity of the architecture, it added robustness and simplified operations. The memcache node, in contrast to the BBT and hub nodes, had no internet-facing functionality, making its operation less dependent on external conditions. For this reason, both the BBT nodes and the hub node were registered and maintained on the memcache node, allowing the restart of any of these internet-facing instances without loss of records or synchronization.

In addition, as detailed in the next section, there was an additional monitor node in charge of (i) recording all the random bits sent from the Bellsters to the laboratories and (ii) providing real-time feedback to the Bellsters. This functionality was isolated from the operations of the hub node. Again, by splitting the monitor and hub instances, a failure or attack in the public and non-critical real-time feedback functionality had no effect on the main, private and critical random-bit distribution task.

Monitor node. For analysis and auditing purposes, all of the bits passing through the first layer of servers were recorded in a database, together with metadata describing their origin (monitor computing resource in Fig. 1c). In particular, every bit was stored together with the username that created it and the origin timestamp. The random bitstreams sent to the individual laboratories were similarly recorded bit by bit, allowing a full reconstruction of the input to the experiments.

In post-event studies of the input data, we estimated the possible contribution from potentially machine-generated participations that were not blocked by the real-time blocking mechanism. We analysed participants whose contribution were substantial, more than 2 kbit in total, and looked for anomalous timing behaviours—such as an improbably short time spent between missions and improbably large number of bits introduced per mission, both of which are limited by the dynamics of human reactions when playing the game. By flagging participants that contributed such anomalous participations as suspicious and cross-referencing against the bits sent to the experiments, we found that no experiment received more than 0.1% of the bits from the 11 suspicious participants.

In addition to using the monitor computing resource to store all the information that was streamed to the laboratories in a database, we also implemented a real-time feedback mechanism to improve the Bellsters’ participation experience. After accomplishing each mission, users were shown a report on the use of their input at each of the laboratories running at that moment, as illustrated in Extended Data Fig. 2d. The numbers shown were calculated as a binomial random process $B(n, p_i)$ with parameters $n = N$ and $p_i = R_i/R$, where N is the number of bits introduced by a user in his/her last mission, R_i is the number of bits sent to laboratory i , and R is the total number of bits entered in the last 0.5-s interval.

Gamification. The BBT required a large number of human-generated random bits in a short time, thus requiring many participants, rapid input and sustained participation. The gamification strategy was designed to maximize all of these factors. Extended Data Fig. 2 shows screenshots of the game.

While still adhering to common conventions of video games (such as levels, power-ups, boss battles, animations and sound effects), the intended appeal of The BIG Bell Quest is less its entertainment value than the opportunity to contribute to the BBT experiments and to test one’s unpredictability against a computer opponent. The game design incorporated internationalization, connection to social networking, community-building features and a feedback system to inform users about their contribution to the experiment, all considered essential to attract the necessary tens of thousands of participants.

The game has a classic challenge-and-reward incentive structure. The challenge is to produce random bits while avoiding being predicted by the Oracle (see Extended Data Fig. 2a). This reproduces the ‘penny-matching’ game studied in psychology^{41,42} and resembles the well known ‘rock–paper–scissors’ game, thus requiring little explanation. The Oracle is an MLA that predicts player behaviour based on patterns in past input (described in Methods, Prediction engine). Most player time was spent in a rapid ‘speed game’ (see Extended Data Fig. 2b), where the Bellster moves along a road by hitting 0s and 1s. This part of the game requires rapid bit generation (a few bits per second) to complete the level in time. Every 20 bits an indicator shows the player’s ‘unpredictability’—that is, the percentage of unpredicted bits entered thus far—and the final score reflects the number of unpredicted bits, with a power-up multiplier for bits entered during a particular time window.

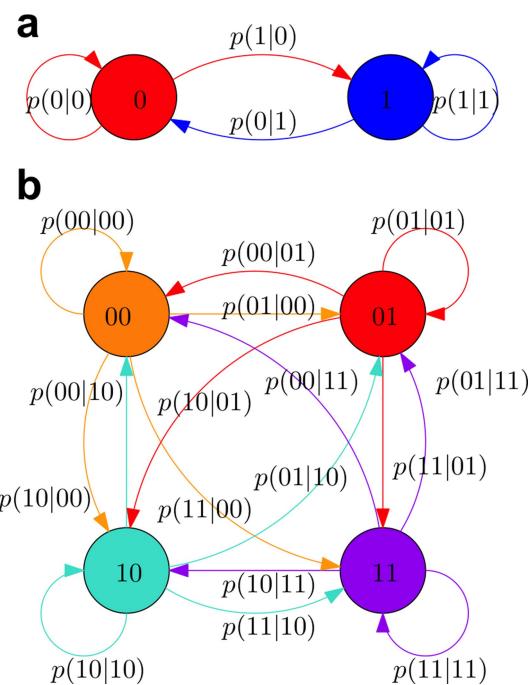
The rewards are multiple: at the individual level, the player is given a score for each level (to encourage a high fraction of unguessed input) and a cumulative score (to encourage repeated play). At the community level, a sharing platform offers rankings and a tool to create groups, so that Bellsters can compare their performance among friends and colleagues. Players can also post their scores to social networks (Facebook, Twitter or Weibo) at the press of a button (see Extended Data Fig. 2e and f). At the scientific level, the game provides a report on which laboratories have used how many of a player’s input and for what purpose (see Extended Data Fig. 2c). Finally, the player is occasionally rewarded with a short video pre-recorded at one of the laboratories, in which experimentalists explain a part of their experiment. User feedback (see Extended Data Fig. 2e) suggests that this approach succeeded in making Bellsters feel meaningfully involved in the project, with a positive effect on retention and propagation.

The speed game–Oracle structure was repeated through three levels of difficulty, or ‘worlds’ (‘Users’, ‘Internet’ and ‘Laboratory’), illustrating the travel of the bits from the fingers of the Bellsters, through the internet, to the laboratories

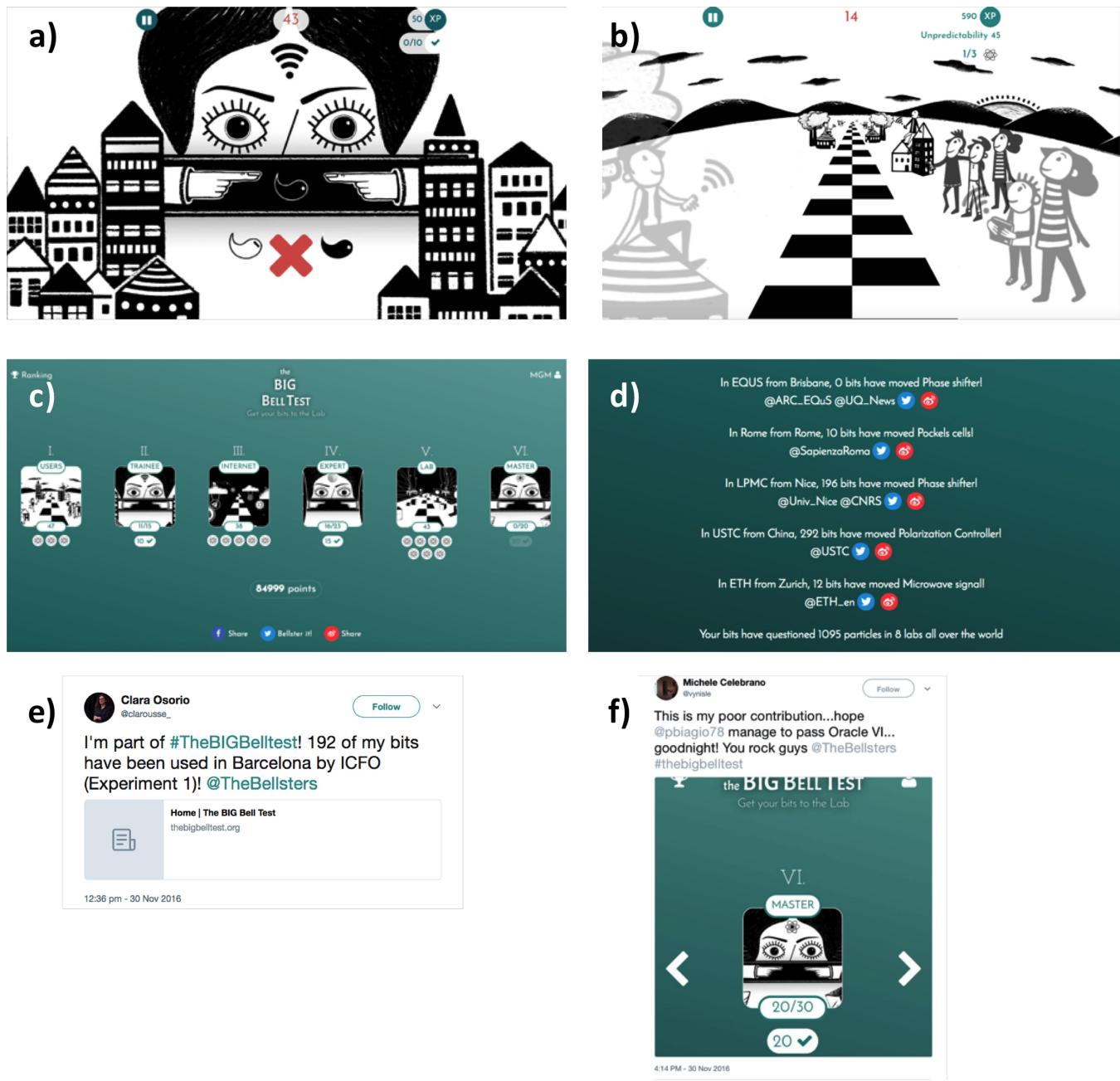
(see Extended Data Fig. 2c). The times, speeds and required unpredicted-input fraction in each of the levels were adjusted with the help of beta testers to avoid offering levels of trivial or impossible difficulty. The final Oracle level was objectively difficult even for an experienced player. To pass, it required $n \geq 20$ unguessed bits in a time period that allowed at most 30 bits to be entered. Even for a sequence of 30 ideal random input bits, the condition $n \geq 20$ occurs less than 5% of the time according to binomial statistics. For 30 bits that are predictable with probability of 0.6, the chance of success drops below 0.003. Nevertheless, several players persisted and completed the game (see Extended Data Fig. 2f).

Data availability. Experimental data are available upon reasonable request from the contact author of each experiment, as indicated in the author contributions. Other project data are available upon reasonable request from the corresponding author.

31. Einstein, A., Podolsky, B. & Rosen, N. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.* **47**, 777–780 (1935).
32. Hall, M. J. W. Relaxed Bell inequalities and Kochen–Specker theorems. *Phys. Rev. A* **84**, 022102 (2011).
33. Barrett, J. & Gisin, N. How much measurement independence is needed to demonstrate nonlocality? *Phys. Rev. Lett.* **106**, 100406 (2011).
34. Bell, J. S., Clauser, J., Horne, M. & Shimony, A. An exchange on local beables. *Dialectica* **39**, 85–110 (1985).
35. Conway, J. & Kochen, S. The free will theorem. *Found. Phys.* **36**, 1441–1473 (2006).
36. Bell, J. in *Speakable and Unspeakable in Quantum Mechanics* Ch. 12 (Cambridge Univ. Press, Cambridge UK, 2004).
37. Erven, C. C. et al. Experimental three-photon quantum nonlocality under strict locality conditions. *Nat. Photon.* **8**, 292–296 (2014).
38. Weihs, G., Jennewein, T., Simon, C., Weinfurter, H. & Zeilinger, A. Violation of Bell’s inequality under strict Einstein locality conditions. *Phys. Rev. Lett.* **81**, 5039–5043 (1998).
39. Wagenaar, W. A. Generation of random sequences by human subjects: A critical survey of the literature. *Psychol. Bull.* **77**, 65–72 (1972).
40. Rapoport, A. & Budescu, D. V. Generation of random series in two-person strictly competitive games. *J. Exp. Psychol. Gen.* **121**, 352–363 (1992).
41. Gibbons, R. *Game theory for applied economists* (Princeton Univ. Press, Princeton, 1992).
42. Mookherjee, D. & Sopher, B. Learning behavior in an experimental matching pennies game. *Games Econ. Behav.* **7**, 62–91 (1994).
43. Serfozo, R. *Basics of applied stochastic processes* (Springer, Berlin, 2009).
44. Heck, R. et al. Remote optimization of an ultra-cold atoms experiment by experts and citizen scientists. Preprint at <https://arxiv.org/abs/1709.02230> (2017).



Extended Data Fig. 1 | Markov chains. **a**, Markov chain for $L=1$. States are represented by circles of different colours, and transitions between states by arrows coloured as the initial state. The last input bit determines the state of the predictor. The probability $p(a \mid b)$ of a transition from state b to state a is estimated from the sequence S_k . **b**, Markov chain for $L=2$, with four states determined by the last two input bits. The transitions model the probability $p(a, b \mid c, d)$ that the user will input bits ab , given that the last two bits were cd . The final prediction is based on the marginal of the next single bit, which is extracted from these estimated probabilities.



Extended Data Fig. 2 | Screenshots from The BIG Bell Quest, illustrating various game elements. a, The Oracle uses an MLA to predict user input. b, The ‘running’ component of the game, in which participants are asked to enter a minimum number of bits with a minimum unpredicted-input fraction in a limited time. c, Sequence of increasing-difficulty levels, interspersed with Oracle challenges. d, In-game feedback

on the use of the user’s input bits in the running experiments. Blue and red buttons allow instant sharing on social networks Twitter and Weibo, respectively. e, A social media post sharing participant results. f, A social media post by a participant who completed the very difficult last Oracle level. The BIG Bell Quest artwork by Maria Pascual (Kaitos Games). See also Methods, Gamification.