# Investigating the analytical robustness of the social and behavioural sciences

Balazs Aczel[1*], Barnabas Szaszi[2,1*], Harry T. Clelland[1,3], Marton Kovacs[4,1], Felix Holzmeister[5], Don van Ravenzwaaij[6], Hannah Schulz-Kümpel[7], Sabine Hoffmann[7], Gustav Nilsonne[8,9], Livia Kosa[1,4], Zoltan A. Torma[1], Yousuf Abdelfatah[10], Christopher L. Aberson[11], Oguz A. Acar[12], Ensar Acem[13], Matus Adamkovic[14,15,16], Timofey Adamovich[17], Krisna Adiasto[18], Love Ahnström[8], Atakan M. Akil[19], Adil S. Al-Busaidi[20,21], Ali H. Al-Hoorie[22], Casper J. Albers[6], Peter J. Allen[23], Taym Alsalti[24], Micah Altman[25], Shilaan Alzahawi[26], Ettore Ambrosini[27], Saule Anafinova[28], Rahul Anand[29], Martin Angerer[30], Ariadna Angulo-Brunet[31], Alberto Antonietti[32], Jozsef Arato[33], Andreu Arenas[34], Marco M. Aviña[35,36], Flavio Azevedo[37], Marko Bachl[38], Bence Bago[39], Štěpán Bahník[40], Bradley J. Baker[41], Elza Balayan[5], Cassandra L. Baldwin[42], Benjamin Banai[43], Kasia Banas[44], František Bartoš[45], Ernest Baskin[46], Jojanneke A. Bastiaansen[47], Nadège Bault[48], Christopher W. Bauman[49], Quintin H. Beazer[50], Maciej Behnke[51], Theiss Bendixen[52], Sebastian Berger[53], Anna Bernard[54], Ursa Bernardic[55], Paul A. Bloom[56], Annika Boldt[57], Ciril Bosch-Rosa[58], Rotem Botvinik-Nezer[59,60], Adam Bouyamourn[61], Ozge Bozkurt[62], Laurel Brehm[63], Johannes Breuer[64], Ryan Briggs[65,66], Hilmar Brohmer[67], Erin Buchanan[68], Johannes Buckenmaier[69], Jeffrey Buckley[70], Jacek Buczny[71,72], Matthias Burghart[73], Bilal H. Butt[74,75], Nick Byrd[76], Valentina Cafarelli[77], Patrick Callahan[78], Tabaré Capitán[79], Kevin Carriere[80], Andrea M. Cataldo[81,82], Gabriel Cepaluni[83,84], Eugene Chan[85], Jesse J. Chandler[86,87], Chia-chen Chang[88], Xi Chen[89,90], Shirley (Shuo) Chen[91], Fadong Chen[92], Hao Chen[93], Valerii Chirkov[94,95], Daniela Cialfi[96,97], Beth Clarke[98], Sophie G. Coelho[99], Clara Cohen[100], Jason Collins[101], Susan W. Cook[102], Gaia Corlazzoli[103], Jamie Cummins[104,105], Christian Czymara[106,107], Jonathan D'hondt[108,109], Anna Dalla Rosa[110], Abi M. B. Davis[111], Charles P. Davis[112], Martin V. Day[113], Freya De Keyzer[114], Joshua R. de Leeuw[115], Tjeerd Rudmer de Vries[116,117], Ramit Debnath[118,119], Filip Dechterenko[120], Elif E. Demiral[121], Marc Desgroseilliers[122], Dominik Dianovics[4,123], Veronica Diveica[124], Stephan Dochow-Sondershaus[125], Simone Dohle[126], LiChen Dong[127], Jonas Dora[128], Angela R. Dorrough[129,130], Anna Dreber[131,5], Hongfei Du[132,133], John E. Edlund[134], Anita Eerland[18], Emir Efendic[135], Jacob Elder[136], Mahmoud M. Elsherif[137,138], Mareike Ernst[139], Eduardo Estrada[140], Luis Eudave[141], Thomas R. Evans[142,143], Arodi Farrera[144], El Mehdi Ferrouhi[145], Lenka Fiala[146,147,148], Fabrício M. Fialho[149,150], Josh L. Fiechter[151], Miloš Fišar[152], Pablo Ezequiel Flores-Kanter[153,154], Michał Folwarczny[155], Jessica L. Fossum[156], Vithor R. Franco[157], René Freichel[158], Danilo Freire[159,160], Joris Frese[161], Alexander C. Furnas[162,163], Johann D. Gaebler[164], Lisa C. Gajary[165,166], Carl Michael Galang[167,168], Benjamin Ganschow[169], S. Mason Garrison[170], Agata Gasiorowska[171], Bruno Gasparotto Ponne[172], Romain Gauriot[173], Alice Geminiani[174,175], Diogo Geraldes[176,177], Morton Ann Gernsbacher[127], Cinzia Giani[178], Enrico Glerean[179], Vukašin Gligorić[180], Timo Gnambs[181], Amélie Godefroidt[182,183], Bastián González-Bustamante[184,185], Andreas Goreis[186], Lorenz Graf-Vlachy[187,188], Manuel Grieder[189,190], Dmitry Grigoryev[191], Sandra Grinschgl[192,193], David J. Grüning[194], João F. Guassi Moreira[195,196], Clément Guichet[197], Lilas Gurgand[198], Hooman Habibnia[199], Andrew C. Hafenbrack[200], Sebastian Hafenbrädl[201], Carolin Häffner[202], Felix Hagemeister[203], Matthew Haigh[204], Nandor Hajdu[1], Narges Hajimoladarvish[205], Jonathan D. Hall[206], Maik Hamjediers[207], Robert M. Hardwick[208], Mehmet Harma[209,210], Nicholas R. Harp[211], Áron D. Hartvig[212,213], Raphael H. Heiberger[214], Arthur Heim[215,216], Øystein Hernaes[217], Dennis Hernaus[218], Tom Heyman[219], Joshua Hicks[220], Jeremy Hogeveen[221], Julia Höpler[7], Sean Dae Houlihan[222,223], Christoph Huber[224], Conor Hughes[225], Teresa Hummler[226], Karoline Huth[227], Moritz Ingendahl[228], Tatsunori Ishii[229], Ozan Isler[230], Kamil Izydorczak[231], Iain R. Jackson[232], Andrew Jahn[233], Maitri Jain[234], Alexander Jakubow[235], Daisung Jang[236], JunHyeok Jang[237], Marc Jekel[238], Fanli Jia[239], William Jiménez-Leal[240], Rebecca Johnson[241], Alex Jones[242], Sebastian Jungkunz[243,244], Pavol Kačmár[245], Caspar Kaiser[246,247], Yağmur Kalaycı[248], Jaroslaw Kantorowicz[249], Anıl Karabulut[250,251], Julian Karch[252], Hamid Karimi-Rouzbahani[253], Johannes A. Karl[254,255], Austėja Kažemekaitytė[256], Aliaksandr Kazlou[257], Zoltan Kekecs[1], Jin Kim[258], Michael H. Kirchler[259], Bence Kiss-Dobronyi[260], Kai N. Klasmeier[261,262], Jack W. Klein[263,264], Cemal Koba[265], Marta Kołczyńska[266], Pavlos Kolias[267], Matěj Kolouch Grabovský[268], Max Korbmacher[269,270,271], Živa Korda[67], Marta Kowal[272], André Kretzschmar[273], Vladislav Krivoshchekov[274], Angelos-Miltiadis Krypotos[275,276], Marcus Kubsch[277], Yoshihiko Kunisato[278], David Lacko[279], Jan R. Landwehr[280], Martin Lange[281], Hongmi Lee[282], Daniel Lee[283], Sangil Lee[284], Edward P. Lemay, Jr.[285], Daniel Lempert[286], Andrea Leo[287,288], Elise Lesage[289], Joel M. Levin[290,291], Peng Li[292], Jing Lin[293], Luke Lindsay[294], Daria Lisovoj[202], Meng Liu[295], Sihong Liu[296], Tingshu Liu[297], Sergio Lo Iacono[298], Paul Lodder[299], Rubén López-Bueno[300], Ruben Lopez-Nicolas[301], Katharina Loter[299], Nigel Mantou Lou[302], Andrey Lovakov[303], Jackson G. Lu[304], Jonas Ludwig[305], Finn Luebber[306,307], Jiří Lukavský[120], Charles Luo[308], Xuanyu Lyu[309], Esther Maassen[299], Martin Máčel[310], Michael L. Mack[311], Christopher R. Madan[312], Andreas Mädebach[313], Joseph Maffly-Kipp[314,315], Daniel J. Mallinson[316], Igor Marchetti[317], Tyler Marghetis[318], Matteo M. Marini[319], Diego Marino Fages[320], Mayte Martínez[321,322], Mario Martinoli[323], Aidas Masiliunas[324], Sébastien Massoni[325], Kaleb C. Mathieu[326], Stefan Mayer[327], Duncan Mayer[328], Maren Mayer[329], Ethan M. McCormick[330,331], Ian M. McDonough[332,333], Amanda L. McGowan[334,335], Miranda M. McIntyre[336], Paul McKee[112,337], Armando N. Meier[338,339], Pascal F. Meier[340], Helena Melero[341], Christoph Merkle[342,343], Raphael Merz[228], Michalis P. Michaelides[344], Patrik Michaelsen[345,346], Gosia Mikolajczak[347], Wladislaw Mill[348], Philip Millroth[349], Kirill G. Miroshnik[317,350], Michal Misiak[351,352], Youri L. Mora[353], David Moreau[354,355], Chris Moreh[356,357], Coby Morvinski[358], Faisal Mushtaq[359,360], Tamás Nagy[1], Christa Nater[361], Elias Naumann[362], Gorka Navarrete[363], Stephan Nebe[364], Andre Nedderhoff[365], Richard Nennstiel[366], Martin Neugebauer[367], Eliana Nicolaisen-Sobesky[368,369], Yngwie A. Nielsen[370], Guiomar Niso[371], Benjamin Nowak[372], Mehmet Okan[373], Kenneth Ong[374], Adrian I. Onicas[375], Christian Oswald[376], Kasper Otten[377], Shubham Pandey[378,379], Myrto Pantazi[380,381], Paolo Papale[382], Philip Pärnamets[8], Shiva Pauer[383,384], Yuri G. Pavlov[385], Samuel Pawel[386], Jonathan E. Peelle[387], Hannah K. Peetz[18], Anton Peez[388,389], Francesca Pesciarelli[390], Brenton D. Peterson[391], Benjamin Petruželka[392], Jonas Petter[158], Jan Pfänder[393], Gerit

Pfuhl[394], Joseph Phillips[395], Matthew T. Pietryka[396], Angelo Pirrone[397,398], Ilse L. Pit[399,400,401], Anna Plachti[402], Irene Sophia Plank[403], Matteo Ploner[256], Russell A. Poldrack[404], Monique M. H. Pollmann[405], Simon Porcher[406], Patrick Präg[407], Andrew Adrian Y. Pua[408,409], Jessica Pugel[410], Rohan Puri[411,412], Marcell Püski[1], Setayesh Radkani[413], Louis Raes[414], Ismaël Rafaï[415], Klara Raiber[416], Steve Rathje[417], Mikhail Reshetnikov[418], Caleb J. Reynolds[419], James P. Reynolds[420], Kévin Rigaud[421], Charlie Rioux[422], Sebastian Rivera[423,185], Olly Robertson[424], Rafael Román-Caballero[425,426,427,428], Raphael Rehms[78],Ivan Ropovik[429,430], Lukas Röseler[431,244], Robert M. Ross[432], Amanda Rotella[204], Franziska F. Rüffer[433], Felix Rusche[348,434], Massimo Rusconi[435], Irene Russo[436], Alexander H. J. Sahm[437], Janos Salamon[438,421], Margaret Samahita[176], Ali Sanaei[439], Arshiya Sangchooli[440], Alexandra Sarafoglou[158], Michele Scandola[441], Henning Schaak[442], Michael Schaerer[443], Eric Schares[444], Hayden T. Schilling[445], Xenia Schmalz[446], Kathleen Schmidt[447], Tom Schonberg[448], Marcel R. Schreiner[449,450], Joris M. Schröder[451], Anna-Lena Schubert[452], Brendan Schuetze[453], Douglas H. Schultz[454], Lars Schulze[455], Shawn T. Schwartz[404,456], Nicole Schwitter[457,458], Bermond Scoggins[459], Yashvin Seetahul[460], Raffaello Seri[435], David R. Shanks[461], Stacy T. Shaw[462], Joseph Shaw[463], Qiang Shen[464], Christoph Siemroth[465], Martina Sladekova[466], Angela Somo[467], Arjun Sondhi[468], Burak Sonmez[469], Lisa Spantig[470,465], Maarten Speekenbrink[471], Angelos Stamos[472], Lukasz Stasielowicz[473], Leonie C. Steckermeier[474,475], Simon R. Steinkamp[402], Andrea H. Stoevenbelt[476], Chris N. H. Street[477], Jordan W. Suchow[478], Hans Fredrik Sunde[479], James Sundquist[480,481], Vsevolod Suschevskiy[482,483], Scott D. Swain[484], Peter Szecsi[4,1], Raluca D. Szekely-Copîndean[485,486], Ewa Szumowska[487], Alessandro Tacconelli[488], Eli Talbert[489], John P. Tang[490], Jorge N. Tendeiro[491], Martina Testori[492], Enrico Toffalini[493], Aleksandar Tomašević[494], Selin Topel[252,495], Lasse Torkkeli[496,497,498], Leonardo Tozzi[499], Jakub Traczyk[171], Alexander Trinidad[500], Darinka Trübutschek[501,502], Konrad Turek[503], Maximiliane Uhlich[504], Eric L. Uhlmann[505], Karolina Urbanska[506], Jasper Van Assche[507,508], Marcel A. L. M. van Assen[433,509], Noah N. N. van Dongen[158], Kenny van Lieshout[510], Roel van Veldhuizen[511], Marton A. Varga[1], Leigh Ann Vaughn[512], Fruzsina Venczel[1], Michela Vezzoli[513], Paul Vierus[514], Antonino Visalli[515], Emily Voldal[516], Fabio Votta[517], Eric-Jan Wagenmakers[158], Anica Waldendorf[518], Matthew J. Walker[519], Matthew B. Wall[520,521], Henri Wallen[522], Ke Wang[523], Iris Wang[524], Y. Andre Wang[311], Markus Weinmann[525], Martin Weiss[526], Christian Westheide[527,528], Aaron Wichman[529,530], Juliane C. Wilcke[78], Benedict J. Williams[531], David Wisniewski[532,168,533], Thomas K. A. Woiczyk[534], Mateusz Woźniak[535,536], Joshua D. Wright[537], Wu Youyou[538], Jesper N. Wulff[342], Tao Yang[539], Siu Kit Yeung[540], Kenneth S. L. Yuen[541,542], Michał Zawistowski[543], Rizqy A. Zein[544,545], Xian Zhao[546], Zefan Zheng[547,548], Steven Zhou[549], Conrad Ziller[226], David Zimmerman[550], Cristina Zogmaister[551], Ro'i Zultan[552], Nicholas Fox[553], Timothy M. Errington[553], Brian A. Nosek[553,554]

[1]Institute of Psychology, ELTE Eötvös Loránd University, Budapest, Hungary, [2]Institute of Operations and Decision Sciences, Corvinus Institute for Advanced Studies (CIAS), Corvinus University of Budapest, Budapest, Hungary, [3]Department of Psychology, Edge Hill University,

Ormskirk, UK, [4]Doctoral School of Psychology, ELTE Eötvös Loránd University, Budapest, Hungary, [5]Department of Economics, University of Innsbruck, Innsbruck, Austria, [6]Department of Psychometrics & Statistics, University of Groningen, Groningen, The Netherlands, [7]Department of Statistics, Ludwig Maximilian University Munich, Munich, Germany, [8]Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden, [9]Department of Psychology, Stockholm University, Stockholm, Sweden, [10]Department of Politics, Princeton University, Princeton, NJ, USA, [11]The Dissertation Coach, Oak Park, CA, USA, [12]King's Business School, King's College London, London, United Kingdom, [13]Department of Psychology, Kadir Has University, Istanbul, Türkiye, [14]Centre of Social and Psychological Sciences, Slovak Academy of Sciences, Bratislava, Slovakia, [15]Faculty of Humanities and Social Sciences, University of Jyväskylä, Jyväskylä, Finland, [16]Faculty of Education, Charles University, Prague, Czechia, [17]Behavioral Genetics Lab, Psychological Institute of the Russian Academy of Education, Moscow, Russia, [18]Behavioural Science Institute, Radboud University, Nijmegen, The Netherlands, [19]Institute of Psychology, Faculty of Humanities and Social Sciences, University of Pécs, Pécs, Hungary, [20]Innovation & Technology Transfer Center, Sultan Qaboos University, Al-Khoud, Oman, [21]Business Communication Department, Sultan Qaboos University, Oman, [22]Jubail English Language and Preparatory Year Institute, Royal Commission for Jubail and Yanbu, Jubail Industrial City, Saudi Arabia, [23]School of Psychological Science, University of Bristol, Bristol, UK, [24]Personality Psychology and Psychological Diagnostics, University of Leipzig, Leipzig, Germany, [25]Center for Research on Equitable and Open Scholarship, Massachusetts Institute of Technology, Cambridge, MA, USA, [26]Graduate School of Business, Stanford University, Stanford, CA, USA, [27]Department of Neuroscience, University of Padova, Padova, Italy, [28]Department of Technical Education, Faculty of Economic and Social Sciences, Budapest University of Technology and Economics (BME), [29]Department of Management, Aarhus BSS (Aarhus University), Aarhus, Denmark, [30]Liechtenstein Business School, University of Liechtenstein, Vaduz, Liechtenstein, [31]Faculty of Psychology and Educational Sciences, Universitat Oberta de Catalunya, Barcelona, Spain, [32]Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milano, Italy, [33]Vienna Cognitive Science Hub, University of Vienna, Vienna, Austria, [34]Institut d'Economia de Barcelona, University of Barcelona, Barcelona, Spain, [35]Department of Government, Harvard University, Cambridge, MA, USA, [36]James M. and Cathleen D. Stone Ph.D. Scholar in Inequality and Wealth Concentration, [37]Department of Interdisciplinary Social Science, University of Utrecht, Utrecht, The Netherlands, [38]Institute for Media and Communication Studies, Freie Universität Berlin, Berlin, Germany, [39]Department of Social Psychology, Tilburg University, Tilburg, The Netherlands, [40]Department of Management, Prague University of Economics and Business, Faculty of Business Administration, Prague, Czech Republic, [41]Department of Sport, Tourism and Hospitality Management, Temple University, Philadelphia, PA, USA, [42]Department of Psychology, Georgia Southern University, Statesboro, GA, USA, [43]Banai analitika, Osijek, Croatia, [44]Usher Institute, University of Edinburgh, Edinburgh, UK, [45]Department of Psychological Methods, University of Amsterdam, Amsterdam, The Netherlands, [46]Department of Food, Pharma and Healthcare, Saint Joseph's University, Philadelphia, PA, USA, [47]Department of Psychiatry, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands, [48]Brain Research & Imaging Centre, School of Psychology, University of Plymouth, Plymouth, United Kingdom, [49]Paul Merage School of Business, University of California Irvine, Irvine, CA, USA, [50]Department of Political Science, Florida State University, Tallahassee, FL, USA, [51]Cognitive Neuroscience Center, Adam Mickiewicz University, Poznan, Poland, [52]School of Culture and Society, Aarhus University, Aarhus, Denmark, [53]Institute of Sociology, University of Bern, Bern, Switzerland, [54]CUBE Research Center, Catolica Lisbon School of Business and Economics, Lisbon, Portugal, [55]CEPE, ETH Zürich, Zürich, Switzerland, [56]Psychiatry Department, Columbia University, New York, NY, USA, [57]Institute of Cognitive Neuroscience, University College London, London, United Kingdom, [58]Chair of Macroeconomics, Technical University Berlin, Berlin, Germany, [59]Department of Psychology, Hebrew University of Jerusalem, Jerusalem, Israel, [60]Psychological and Brain Sciences, Dartmouth College, Hanover, USA, [61]Charles and Louise Travers Department of Political Science, UC Berkeley, Berkeley, CA, USA, [62]Department of Psychology, Anadolu University, Eskişehir, Türkiye, [63]Department of Linguistics, University of California, Santa Barbara, Santa Barbara, CA, USA, [64]Department of Computational Social Science, GESIS - Leibniz Institute for the Social Sciences, Cologne, Germany, [65]Guelph Institute of Development Studies, University of Guelph, Guelph, ON, Canada, [66]Department of Political Science, [67]Department of Psychology, University of Graz, Graz, Austria, [68]Analytics, Harrisburg University of Science and Technology, Harrisburg, PA, USA, [69]Department of Economics, University of Zurich, Zurich, Switzerland, [70]Department of Technology Education, Technological University of the Shannon: Midlands Midwest, Athlone, Ireland, [71]Experimental and Applied Psychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands, [72]Insitute of Psychology, SWPS University, Sopot, Poland, [73]Department of Criminology, Max Planck Institute for the Study of Crime, Security and Law, Freiburg im Breisgau, Germany, [74]The School of Mathematics and Computer Science (SMCS), Institute of Business Administration, Karachi, Pakistan, [75]Computer Science Department, DHA Suffa University, Karachi, Pakistan, [76]School of Humanities, Arts, and Social Sciences, Stevens Institute of Technology, Hoboken, NJ, USA, [77]Independent Researcher, New York, NY, USA, [78]Institute for Medical Information Processing, Biometry, and Epidemiology, Ludwig-Maximilians-Universität München, Munich, Germany, [79]Department of Economics, Swedish University of Agricultural Sciences, Uppsala, Sweden, [80]Department of Psychology, Stonehill College, Easton, MA, USA, [81]Center for Depression, Anxiety, and Stress Research, McLean Hospital, Belmont, MA, USA, [82]Department of Psychiatry, Harvard Medical School, Boston, USA, [83]Department of International Relations and Public Policy, São Paulo State University, Franca, Brazil, [84]Kellogg Institute for International Studies, University of Notre Dame, Notre Dame, USA, [85]Ted Rogers School of Management, Toronto Metropolitan University, Toronto, ON, Canada, [86]Institute for Social Research, University of Michigan, Ann Arbor, MI, USA, [87]Mathematica, 220 Easy Huron Street, Ann Arbor, MI, 48104, [88]Department of Biological Sciences, National University of Singapore, Singapore, Singapore, [89]Chinese Open Science Network, Shanghai, China, [90]Dept. of Intelligent Perception and Interaction Research, OPPO Shanghai Research Institute, Shanghai, China, [91]Lazaridis School of Business and Economics, Wilfrid Laurier University, Waterloo, ON, Canada, [92]School of Management, Neuromanagement Laboratory & The State Key Laboratory of Brain-machine Intelligence, Zhejiang University, Hangzhou, China, [93]Department of Psychiatry and Psychotherapy, Technische Universität Dresden, Dresden, Germany, [94]Institute for Theoretical Biology, Department of Biology, Humboldt-Universität zu Berlin, Berlin, Germany, [95]Science of Intelligence, Research Cluster of Excellence, Berlin, Germany, [96]Institute for Complex Systems - National Research Council of Italy, Rome, Italy, [97]Enrico Fermi Research Center, [98]Melbourne School of Psychological Sciences, The University of Melbourne, Melbourne, VIC, Australia, [99]Department of Psychology, York University, Toronto, ON, Canada, [100]English Language and Linguistics, School of Critical Studies, University of Glasgow,

Glasgow, UK, [101]UTS Business School, University of Technology Sydney, Ultimo, NSW, Australia, [102]Department of Psychological and Brain Sciences, University of Iowa, Iowa City, IA, USA, [103]Center for Research in Cognition & Neurosciences, Université Libre de Bruxelles, Brussels, Belgium, [104]Department for Psychology of Digitalisation, University of Bern, Bern, Switzerland, [105]Department of Experimental-Clinical and Health Psychology, Ghent University, Ghent, Belgium, [106]Department of Sociology and Anthropology, Tel Aviv University, Tel Aviv, Israel, [107]Department of Sociology, Goethe University Frankfurt, Frankfurt am Main, Germany, [108]Department of Work, Organisation and Society, Ghent University, Gent, Belgium, [109]School of Business and Economics, Vrije Universiteit Amsterdam, Amsterdam, Netherlands, [110]Department of Philosophy, Sociology, Education and Applied Psychology, University of Padova, Padua, Italy, [111]School of Psychology, University of Lincoln, Lincoln, England, [112]Department of Psychology & Neuroscience, Duke University, Durham, NC, USA, [113]Department of Psychology, Memorial University of Newfoundland, St. John's, NL, Canada, [114]Erasmus School of History, Culture and Communication, Erasmus University Rotterdam, Rotterdam, The Netherlands, [115]Cognitive Science, Vassar College, Poughkeepsie, NY, USA, [116]Copenhagen Health Complexity Center, Department of Public Health, University of Copenhagen, Copenhagen, Denmark, [117]Department of Health Sciences, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands, [118]Collective Intelligence and Design Group, University of Cambridge, Cambridge, UK, [119]Climate and Social Intelligence Lab, Caltech, USA, [120]Department of Cognitive Psychology, Institute of Psychology, Czech Academy of Sciences, Prague, Czech Republic, [121]College of Business and Technology, East Tennessee State University, Johnson City, TN, USA, [122]School of Computer Communication Sciences, Ecole Polytechnique Lausanne, Lausanne, Switzerland, [123]Institute of Psychology, ELTE Eötvös Loránd University, Budapest, [124]Department of Neurology and Neurosurgery, Montreal Neurological Institute, Montréal, QC, Canada, [125]Department of Sociology, University of Copenhagen, Copenhagen, Denmark, [126]Institute of General Practice and Family Medicine, University Hospital Bonn, University of Bonn, Bonn, Germany, [127]Department of Psychology, University of Wisconsin-Madison, Madison, WI, USA, [128]Department of Psychology, University of Washington, Seattle, WA, USA, [129]Faculty of Psychology, FernUniversität in Hagen, Hagen, Germany, [130]Department of Psychology, University of Cologne, Germany, [131]Department of Economics, Stockholm School of Economics, Stockholm, Sweden, [132]Department of Psychology, Beijing Normal University at Zhuhai, Zhuhai, China, [133]Beijing Key Laboratory of Applied Experimental Psychology, National Demonstration Center for Experimental Psychology Education (Beijing Normal University), Faculty of Psychology, Beijing Normal University, Beijing, China, [134]Department of Psychology, Rochester Institute of Technology, Rochester, NY, USA, [135]Marketing and supply chain management, Maastricht University, Maastricht, The Netherlands, [136]Independent Researcher, Philadelphia, PA, USA, [137]Psychology and Vision Sciences, University of Leicester, Leicester, England, [138]University of Birmingham, [139]Institute of Psychology, Department of Clinical Psychology, Psychotherapy and Psychoanalysis, University of Klagenfurt, Klagenfurt am Woerthersee, Austria, [140]Social Psychology and Methodology, Universidad Autónoma de Madrid, Madrid, Spain, [141]School of Education and Psychology, University of Navarra, Pamplona, Spain, [142]School of Human Sciences, University of Greenwich, London, UK, [143]Institute for Lifecourse Development, [144]Institute of Anthropological Research, National Autonomous University of Mexico, Mexico city, Mexico, [145]Faculty of Economics and Management, Ibn Tofail University, Kénitra, Morocco, [146]Department of Economics, University of Ottawa, Ottava, ON, Canada, [147]Institute for Replication, University of Ottawa, Ottawa, ON, Canada, [148]Department of Economics, Tilburg University, Netherlands, [149]School of Sociology, HSE University, Moscow, Russia, [150]International Inequalities Institute, London School of Economics and Political Science, London, United Kingdom, [151]Air Force Research Laboratory, Wright Patterson AFB, OH, USA, [152]Department of Public Economics & MUEEL, Masaryk University, Brno, Czech Republic, [153]Universidad Siglo 21, Córdoba, Argentina, [154]Universidad Católica de Córdoba, Argentina, [155]Discipline of Marketing, J.E. Cairnes School of Business & Economics, University of Galway, Galway, Ireland, [156]School of Psychology, Family, and Community, Seattle Pacific University, Seattle, WA, USA, [157]Postgraduate Program of Psychology, São Francisco University, Campinas, Brazil, [158]Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands, [159]Department of Quantitative Theory and Methods, Emory University, Atlanta, GA, USA, [160]School of Social and Political Sciences, University of Lincoln, Lincoln, United Kingdom, [161]Department of Political and Social Sciences, European University Institute, Fiesole, Italy, [162]Center for Science of Science and Innovation, Evanston, IL, USA, [163]Kellogg School of Management, Northwestern University, Evanston, Illinois, USA, [164]Department of Statistics, Harvard University, Cambridge, MA, USA, [165]John Glenn College of Public Affairs, The Ohio State University, Columbus, Ohio, USA, [166]Caspian Strategy & Analytics, Dublin, Ohio, USA, [167]Psychology, Humboldt-Universität zu Berlin, Berlin, Germany, [168]Berlin School of Mind and Brain, [169]Education and Child Studies, Leiden University, Leiden, The Netherlands, [170]Department of Psychology, Wake Forest University, Winston-Salem, NC, USA, [171]Faculty of Psychology in Wroclaw, SWPS University, Wroclaw, Poland, [172]Federal Senate of Brazil, Brasília, Brazil, [173]Department of Economics, Deakin University, Melbourne, VIC, Australia, [174]Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy, [175]Champalimaud Foundation, Lisbon, Portugal, [176]School of Economics, University College Dublin, Dublin, Ireland, [177]Geary Institute for Public Policy, University College Dublin, Dublin, Ireland, [178]Dipartimento di Economia, Università degli Studi dell'Insubria, Varese, Italy, [179]Department of Neuroscience and Biomedical Engineering, Aalto University, Espoo, Finland, [180]Deparrtment of Psychology, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates, [181]Educational Measurement, Leibniz Institute for Educational Trajectories, Bamberg, Germany, [182]Centre for Research on Peace and Development, KU Leuven, Leuven, Belgium, [183]IESEG School of Management, Lille, France, [184]Institute of Public Administration, Leiden University, The Hague, The Netherlands, [185]Universidad Diego Portales, [186]Department of Child and Adolescent Psychiatry, Medical University of Vienna, Vienna, Austria, [187]Department of Business and Economics, TU Dortmund University, Dortmund, Germany, [188]ESCP Business School, Paris, France, [189]Faculty of Economics, UniDistance Suisse, Brig, Switzerland, [190]Zurich University of Applied Sciences (ZHAW), School of Management and Law, [191]Center for Sociocultural Research, HSE University, Moscow, Russia, [192]Institute of Psychology, University of Bern, Bern, Switzerland, [193]Institute of Psychology, University of Graz, Graz, Austria, [194]Center for Adaptive Rationality, Max-Planck Institute, Berlin, Germany, [195]Psychology, University of California, Los Angeles, Los Angeles, CA, USA, [196]University of Wisconsin, Madison, [197]Department of Psychology, University Grenoble Alpes, CNRS UMR 5105 LPNC, Grenoble, France, [198]Département d'études cognitives, Ecole Normale Supérieure Paris, Paris, France, [199]Department of Strategy of Innovation, Institute for Cognition and Behavior, WU Vienna University of Economics and Business, Vienna, Austria, [200]Management & Organization, University of Washington, Seattle, WA, USA, [201]Managing People in Organizations, IESE Business School, Barcelona, Spain,

[202]Department of Psychology: Social Psychology, University of Cologne, Cologne, Germany, [203]Data Science, Sueddeutsche Zeitung Digitale Medien, Munich, Germany, [204]Department of Psychology, Northumbria University, Newcastle-upon-Tyne, United Kingdom, [205]Centre for Social and Behaviour Change, Ashoka University, New Delhi, India, [206]Department of Economics, Finance, and Legal Studies, University of Alabama, Tuscaloosa, AL, USA, [207]Department of Social and Political Science, European University Institute, San Domenico di Fiesole, Italy, [208]Institute of Neuroscience, Catholic University of Louvain, Brussels, Belgium, [209]Faculty of Psychology, University of Akureyri, Akureyri, Iceland, [210]Department of Psychology, Kadir Has University, Istanbul, Turkey, [211]Department of Psychology, University of California, Berkeley, Berkeley, CA, USA, [212]Department of Decision Sciences, Corvinus University of Budapest, Budapest, Hungary, [213]International Institute for Applied Systems Analysis (IIASA), Austria, Laxenburg, [214]Institute for Social Sciences, University of Stuttgart, Stuttgart, Germany, [215]Labour and public policies, Paris school of economics, Paris, France, [216]Statistics, evaluation and research department, Caisse nationale des allocations familiales, Paris, France, [217]Economics, Frisch Centre, Oslo, Norway, [218]Department of Psychiatry & Neuropsychology, Mental Health and NeuroScience (MHeNS) Research Institute, Maastricht University, Maastricht, The Netherlands, [219]Methodology and Statistics Unit, Institute of Psychology, Leiden University, Leiden, The Netherlands, [220]Psychological and Brain Sciences, Texas A&M University, College Station, TX, USA, [221]Department of Psychology, University of New Mexico, Albuquerque, NM, USA, [222]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA, [223]Dartmouth College, [224]Department of Finance, Aalto University, Espoo, Finland, [225]Applied Economics, University of Minnesota, St. Paul, MN, USA, [226]Department of Political Science, University of Duisburg-Essen, Duisburg, Germany, [227]Department of Health Promotion, Care and Public Health Research Institute (CAPHRI), Maastricht University, Maastricht, The Netherlands, [228]Department of Psychology, Ruhr University Bochum, Bochum, Germany, [229]Department of Psychology, Japan Women's University, Bunkyo-ku, Japan, [230]School of Economics, The University of Queensland, St Lucia, Australia, [231]Faculty of Psychology in Wrocław, SWPS University, Wrocław, Poland, [232]Division of Psychology, Communication and Human Neuroscience, University of Manchester, Manchester, UK, [233]Department of Radiology, University of Michigan, Ann Arbor, MI, USA, [234]Department of Psychology, Eastern Michigan University, Ypsilanti, MI, USA, [235]Yale Law School, Yale University, New Haven, CT, USA, [236]Melbourne Business School, Carlton, Australia, [237]Political Science, University of California, Merced, Merced, CA, USA, [238]Department of Psychology, University of Cologne, Köln, Germany, [239]Department of Psychology, Seton Hall University, South Orange, NJ, USA, [240]Department of Psychology, Universidad de los Andes, Bogotá, Colombia, [241]McCourt School of Public Policy, Georgetown University, Washington, DC, USA, [242]School of Psychology, Swansea University, Swansea, UK, [243]Institute of Political Science and Sociology, University of Bonn, Bonn, Germany, [244]University of Bamberg, [245]Department of Psychology, Faculty of Arts, Pavol Jozef Šafárik University in Košice, Košice, Slovakia, [246]Warwick Business School, University of Warwick, Coventry, United Kingdom, [247]University of Oxford, [248]Independent Researcher, Manisa, Turkey, [249]Institute of Security and Global Affairs, Leiden University, The Hague, The Netherlands, [250]MoMiLab, Social and Affective Neuroscience Group, IMT School for Advanced Studies Lucca, Lucca, Italy, [251]Department of Psychology, University of Essex, Colchester, UK, [252]Institute of Psychology, Leiden University, Leiden, The Netherlands, [253]Mater Research Institute, University of Queensland, Brisbane, Australia, [254]School of

Psychology, Dublin City University, Dublin, Ireland, [255]Victoria University of Wellington, New Zealand, [256]Department of Economics and Management, University of Trento, Trento, Italy, [257]Flo Health LTU, UAB, Vilnius, Lithuania, [258]D'Amore-McKim School of Business, Northeastern University, [259]Department of Banking and Finance, University of Innsbruck, Innsbruck, Austria, [260]Institute of Operations and Decision Sciences, Corvinus University of Budapest, Budapest, Hungary, [261]Structural Change and Work Organisation, Federal Institute for Occupational Safety and Health (BAuA), Dortmund, Germany, [262]Chair for Industrial/Organizational Psychology, University of Hamburg, Hamburg, Germany, [263]Department of Psychology, The University of Hong Kong, Hong Kong SAR, China, [264]Hong Kong Institute of Asia-Pacific Studies, The Chinese University of Hong Kong, Hong Kong SAR, China, [265]Computer Vision Group, Sano Centre for Computational Personalized Medicine, Krakow, Poland, [266]Department of Research on Social and Institutional Transformations, Institute of Political Studies of the Polish Academy of Sciences, Warszawa, Poland, [267]Department of Mathematics, Aristotle University of Thessaloniki, Thessaloniki, Greece, [268]Independent Researcher, Brno, Czechia, [269]Department of Health and Functioning, Western Norway University of Applied Sciences, Bergen, Norway, [270]Mohn Medical Imaging and Visualisation Center, Bergen, Norway, [271]Dep. Neurology, Haukeland University Hospital, Faculty of Medicine, University of Bergen, Bergen, Norway, [272]IDN Being Human Lab - Institute of Psychology, University of Wroclaw, Wroclaw, Poland, [273]Department of Psychology, University of Zurich, Zurich, Switzerland, [274]CAPE, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, [275]Department of Clinical Psychology, Utrecht University, Utrecht, The Netherlands, [276]Group of Health Psychology, KU Leuven, Leuven, Belgium, [277]Department of Physics, Freie Universität Berlin, Berlin, Germany, [278]Department of Psychology, Senshu University, Kawasaki, Japan, [279]Department of Personality Psychology and Social Psychology, Institute of Psychology, Czech Academy of Sciences, Brno, Czech Republic, [280]Department of Marketing, Goethe University Frankfurt, Frankfurt am Main, Germany, [281]Labour Markets and Social Insurance, ZEW – Leibniz Centre for European Economic Research, Mannheim, Germany, [282]Department of Psychological Sciences, Purdue University, West Lafayette, IN, USA, [283]Lerner College of Business, University of Delaware, Newark, DE, USA, [284]Department of Psychological and Brain Sciences, Boston University, Boston, MA, USA, [285]Psychology, University of Maryland, College Park, MD, USA, [286]Department of Politics, State University of New York, Potsdam, Potsdam, NY, USA, [287]Department of translational research and advanced technologies in medicine and surgery, University of Pisa, Pisa, Italy, [288]Nuclear Medicine Unit, Department of Diagnostic Imaging, N.O.P. - S. Stefano, U.S.L. Toscana Centro, Prato, Italy, [289]Department of Experimental Psychology, University of Ghent, Ghent, Belgium, [290]Rady School of Management, University of California San Diego, La Jolla, CA, USA, [291]Katz Graduate School of Business, University of Pittsburgh, Pittsburgh, USA, [292]School of Management, University of Bath, Bath, UK, [293]Institute of Developmental Psychology, Beijing Normal University, Beijing, China, [294]Department of Economics, University of Exeter, Exeter, UK, [295]School of English and International Studies, Beijing Foreign Studies University, Beijing, China, [296]Stanford Center on Early Childhood, Stanford University, Stanford, CA, USA, [297]Department of Social, Personality, and Health Psychology, University of Houston, Houston, TX, USA, [298]Department of Sociology, University of Essex, Colchester, UK, [299]Department of Methodology and Statistics, Tilburg University, Tilburg, The Netherlands, [300]Department of Physical Medicine and Nursing, University of Zaragoza, Zaragoza, Spain, [301]Department of Basic Psychology and Methodology, University of Murcia, Murcia, Spain,

[302]Department of Psychology, University of Victoria, Victoria, BC, Canada, [303]German Centre for Higher Education Research and Science Studies (DZHW), Berlin, Germany, [304]MIT Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA, USA, [305]Technische Universität Berlin, Berlin, Germany, [306]Social Neuroscience Lab, Department of Psychiatry and Psychotherapy, University of Lübeck, Lübeck, Germany, [307]Open Science Initiative, University of Lübeck, Lübeck, Germany, [308]Department of Economics, University of Maryland, College Park, College Park, MD, USA, [309]Institute for Behavioral Genetics, University of Colorado at Boulder, Boulder, CO, USA, [310]Department of Psychology, Charles University, Prague, Czech republic, [311]Department of Psychology, University of Toronto, Toronto, ON, Canada, [312]School of Psychology, University of Nottingham, Nottingham, UK, [313]Department of Translation and Language Sciences, Universitat Pompeu Fabra, Barcelona, Spain, [314]Department of Psychological & Brain Sciences, Texas A&M University, College Station, TX, USA, [315]The Ohio State University Medical Center, [316]School of Public Affairs, Penn State Harrisburg, Middletown, PA, USA, [317]Department of Life Sciences, University of Trieste, Trieste, Italy, [318]Cognitive and Information Sciences, University of California, Merced, Merced, CA, USA, [319]Department of Economics , University of Bologna, Bologna, Italy, [320]Department of Economics, Durham University, Durham, UK, [321]Department of Interdisciplinary Life Sciences, Domestication Lab, Konrad Lorenz Institute of Ethology, University of Veterinary Medicine Vienna, Vienna, Austria, [322]Language Research Center, Georgia State University, Atlanta, Georgia, USA, [323]Insitute of Economics & L'EMbeDS, Sant'Anna School of Advanced Studies, Pisa, Italy, [324]School of Economics, University of Sheffield, Sheffield, UK, [325]Université de Lorraine, Université de Strasbourg, CNRS, BETA, Nancy, France, [326]Department of Psychology and Special Education, East Texas A&M University, Commerce, TX, USA, [327]School of Business and Economics, University of Tübingen, Tübingen, Germany, [328]Independent Researcher, Doylestown, PA, USA, [329]Leibniz-Institut für Wissensmedien (Knowledge Media Research Center), Tübingen, Germany, [330]Methodology & Statistics, Universiteit Leiden, Leiden, The Netherlands, [331]Education Statistics and Research Methods, University of Delaware, Newark, USA, [332]Department of Psychology, Binghamton University, Binghamton, NY, USA, [333]Center for Cognitive Applications, Binghamton University, Binghamton, USA, [334]Annenberg School for Communication, University of Pennsylvania, Philadelphia, PA, USA, [335]Annenberg School for Communication, University of Pennsylvania, [336]Department of Psychology, California State University San Bernardino, San Bernardino, CA, USA, [337]Department of Psychology, Southern Connecticut State University, New Haven, USA, [338]Faculty of Business and Economics, University of Basel, Basel, Switzerland, [339]University of Lausanne, [340]Department of Business Administration, University of Zurich, Zurich, Switzerland, [341]Departamento de Psicobiología y Metodología en Ciencias del Comportamiento, Universidad Complutense de Madrid, Madrid, Spain, [342]Department of Economics and Business Economics, Aarhus University, Aarhus, Denmark, [343]Danish Finance Institute, Frederiksberg, Denmark, [344]Department of Psychology, University of Cyprus, Nicosia, Cyprus, [345]Department of Political Science, University of Gothenburg, Gothenburg, Sweden, [346]Department of Management and Engineering, Division of Economics, Jedi Lab, Linköping University, 581 83 Linköping, Sweden, [347]Global Institute for Women's Leadership, Canberra, ACT, Australia, [348]Department of Economics, University of Mannheim, Mannheim, Germany, [349]Department of Psychology, Uppsala University, Uppsala, Sweden, [350]Faculty of Psychology, Saint Petersburg State University, Saint Petersburg, Russia, [351]IDN Being Human, University of Wroclaw, Wroclaw, Poland, [352]University of Oxford,

School of Anthropology and Museum Ethnography, [353]Center for Social and Cultural Psychology, Université libre de Bruxelles [Free University of Brussels], Brussels, Belgium, [354]School of Psychology, University of Auckland, Auckland, New Zealand, [355]University of Auckland, Centre for Brain Research, Auckland, NZ, [356]School of Geography, Politics and Sociology, Newcastle University, Newcastle upon Tyne, United Kingdom, [357]ESRC Centre for Population Change, University of Southampton, UK, [358]Department of Management, Ben-Gurion University of the Negev, Be'er Sheva, Israel, [359]School of Psychology, University of Leeds, Leeds, United Kingdom, [360]Leeds NIHR Biomedical Reseatch Centre, [361]Department of Psychology, University of Bern, Bern, Switzerland, [362]GESIS Leibniz Institute for the Social Sciences, Mannheim, Germany, [363]School of Psychology, Center for Social and Cognitive Neuroscience, Universidad Adolfo Ibáñez, Santiago de Chile, Chile, [364]Zurich Center for Neuroeconomics, Department of Economics, University of Zurich, Zurich, Switzerland, [365]Department of Psychological Methods, Helmut Schmidt University Hamburg, Hamburg, Germany, [366]Department of Sociology of Education, University of Bern, Bern, Switzerland, [367]Karlsruhe University of Education, Karlsruhe, Germany, [368]Institute of Neuroscience and Medicine (INM-7: Brain and Behaviour), Research Centre Jülich, Jülich, Germany, [369]Institute of Systems Neuroscience, Heinrich Heine University Düsseldorf, Düsseldorf, Germany, [370]Department of Linguistics, Cognitive Science and Semiotics, Aarhus University, Aarhus, Denmark, [371]Neuroimaging Group, Instituto Cajal, CSIC, Madrid, Spain, [372]AgroParisTech, INRAE, VetAgro Sup, Université Clermont Auvergne, Lempdes, France, [373]Department of Marketing, Istanbul Bilgi University, Istanbul, Turkey, [374]Department of Economics, University of Maryland College Park, College Park, MD, USA, [375]Department of Neurology, University of Utah, Salt Lake City, UT, USA, [376]Center for Crisis Early Warning, University of the Bundeswehr Munich, Neubiberg, Germany, [377]Asylum and Migration group, WODC - Research and Data Centre, The Hague, The Netherlands, [378]Department of Humanities and Social Sciences, Indian Institute of Technology Bombay, Mumbai, India, [379]University of Osnabrueck, Germany, [380]Center for Social and Cultural Psychology, Université libre de Bruxelles, Brussels, Belgiun, [381]University of Amsterdam, [382]Vision & Cognition, Netherlands Institute for Neuroscience (KNAW), Amsterdam, The Netherlands, [383]Department of Social Psychology, Helmut Schmidt University Hamburg, Hamburg, Germany, [384]Department of Social Psychology, University of Amsterdam, Amsterdam, The Netherlands, [385]University of Tuebingen, Tuebingen, Germany, [386]Department of Biostatistics, University of Zurich, Zurich, Switzerland, [387]Department of Communication Sciences and Disorders, Northeastern University, Boston, MA, USA, [388]Research Center "Normative Orders", Goethe University Frankfurt, Frankfurt am Main, Germany, [389]Peace Research Institute Frankfurt (PRIF), Frankfurt am Main, Germany, [390]Department of Biomedical, Metabolic, and Neurological Sciences, University of Modena and Reggio Emilia, Modena, Italy, [391]The Sababi Institute, Hargeisa, Somaliland, [392]Faculty of Social Studies, Department of Social Work, University of Ostrava, Ostrava, Czech Republic, [393]Département d'études cognitives, ENS, EHESS, PSL University, Institut Jean Nicod, Paris, France, [394]Department of Psychology, Norwegian University of Science and Technology, Trondheim, Norway, [395]School of Law and Politics, Cardiff University, Cardiff, United Kingdom, [396]Political Science, University of Wisconsin–Madison, Madison, WI, USA, [397]Department of Psychology, University of Liverpool, Liverpool, UK, [398]Centre for Philosophy of Natural and Social Science, London School of Economics, London, UK, [399]Institute of Human Sciences, University of Oxford, Oxford, United Kingdom, [400]Calleva Research Centre for Evolution and Human Sciences, Magdalen College, Oxford, United Kingdom, [401]Center for

Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany, [402]Danish Research Centre for Magnetic Resonance, Department of Radiology and Nuclear Medicine, Copenhagen University Hospital - Amager and Hvidovre, Hvidovre, Denmark, [403]Department of Psychiatry and Psychotherapy, Ludwig-Maximilians-Universität München, Munich, Germany, [404]Department of Psychology, Stanford University, Stanford, CA, USA, [405]Department of Communication and Cognition, Tilburg University, Tilburg, The Netherlands, [406]Department of Management, Université Paris Dauphine PSL, Paris, France, [407]Center for Research in Economics and Statistics (CREST), ENSAE, Institut Polytechnique de Paris, Palaiseau, France, [408]Carlos L. Tiu School of Economics, De La Salle University - Manila, Philippines, [409]School of Business, Economics, and Information Systems, University of Passau, Passau, Germany, [410]Evidence-to-Impact Collaborative, Pennsylvania State University, University Park, PA, USA, [411]School of Psychological Sciences, University of Tasmania, Hobart, TAS, Australia, [412]Menzies Institute for Medical Research, University of Tasmania, Hobart, TAS, Australia, [413]Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA, [414]Department of Economics, Tilburg University, Tilburg, The Netherlands, [415]GREDEG, Université Cote d'Azur, Nice, France, [416]Department of Sociology, Radboud University, Nijmegen, The Netherlands, [417]Department of Psychology, New York University, New York, NY, USA, [418]Department of Psychology, National Research University Higher School of Economics, Moscow, Moscow, Russia, [419]Department of Psychology, Willamette University, Salem, OR, USA, [420]School of Psychology, Aston University, Birmingham, United Kingdom, [421]Psychology, Artois University, Arras, France, [422]Department of Interdisciplinary Human Sciences, Texas Tech University, Lubbock, TX, USA, [423]Escuela de Gobierno y Administración Pública, Universidad Mayor, Santiago, Chile, [424]Department of Psychiatry, Department of Experimental Psychology, University of Oxford, Oxford, United Kingdom, [425]Department of Experimental Psychology, University of Granada, Granada, Spain, [426]Mind, Brain and Behavior Research Center (CIMCYC), University of Granada, Granada, Spain, [427]Department of Psychology, Neuroscience and Behaviour, McMaster University, Hamilton, Canada, [428]McMaster Institute for the Music and the Mind, McMaster University, Hamilton, Canada, [429]Faculty of Education, Charles University, Prague, Czech Republic, [430]Institute of Psychology, Czech Academy of Sciences, Brno, Czech Republic, [431]Münster Center for Open Science, University of Münster, Münster, Germany, [432]Department of Philosophy, Macquarie University, Sydney, NSW, Australia, [433]Methodology and Statistics, Tilburg University, Tilburg, The Netherlands, [434]Max Planck Institute for Research on Collective Goods, Bonn, Germany, [435]InsIDE Lab, Dipartimento di Economia, Università degli Studi dell'Insubria, Varese, Italy, [436]ILC, CNR, Pisa, Italy, [437]Department of Psychology, University of Konstanz, Konstanz, Germany, [438]Department of Ergonomics and Psychology, Budapest University of Technology and Economics, Budapest, Hungary, [439]Computational Social Science, University of Chicago, Chicago, IL, USA, [440]Melbourne School of Psychological Sciences, University of Melbourne, Melbourne, VIC, Australia, [441]Department of Human Sciences, University of Verona, Verona, Italy, [442]Department of Economics and Social Sciences, University of Natural Resources and Life Sciences, Vienna, Austria, [443]Lee Kong Chian School of Business, Singapore Management University, Singapore, Singapore, [444]University Library, Iowa State University, Ames, IA, USA, [445]Centre for Marine Science and Innovation, UNSW Australia, UNSW, NSW, Australia, [446]Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital, LMU Munich, Munich, Germany, [447]Department of Psychology, Ashland University, Ashland, OH, USA, [448]School of

Neurobiology, Biochemistry and Biophysics, Faculty of Life Sciences and Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel, [449]Department of Psychology III, Julius-Maximilians-Universität Würzburg, Würzburg, Germany, [450]Department of Psychology, University of Mannheim, Mannheim, Germany, [451]School of Economics, University of Nottingham, Nottingham, United Kingdom, [452]Department of Psychology, University of Mainz, Mainz, Germany, [453]Education Sciences Department, University of Potsdam, Potsdam, Germany, [454]Center for Brain, Biology and Behavior, Department of Psychology, University of Nebraska-Lincoln, Lincoln, NE, USA, [455]Department of Educational Sciences and Psychology, Freie Universität Berlin, Berlin, Germany, [456]Wu Tsai Neurosciences Institute, Stanford University, [457]MZES, University of Mannheim, Mannheim, Germany, [458]Department of Sociology, University of Warwick, Coventry, United Kingdom, [459]School of Politics and International Relations, Australian National University, Canberra, ACT, Australia, [460]Department of Psychology, University of Innsbruck, Innsbruck, Austria, [461]Division of Psychology and Language Sciences, University College London, London, United Kingdom, [462]Department of Social Sciences and Policy Studies, Worcester Polytechnic Institute, Worcester, MA, USA, [463]Faculty of Sport, Technology and Health Sciences, St Mary's University, Twickenham, London, UK, [464]Key Laboratory of Brain-Machine Intelligence for Information Behavior (Ministry of Education and Shanghai), School of Business and Management, Shanghai International Studies University, Shanghai, China, [465]Department of Economics, University of Essex, Colchester, United Kingdom, [466]School of Psychology, University of Sussex, Brighton, United Kingdom, [467]Department of Psychology, San Diego State University, San Diego, CA, USA, [468]Department of Quantitative Intelligence, Feinstein Institutes for Medical Research, New York, NY, USA, [469]Social Research Institute, University College London, London, United Kingdom, [470]School of Business and Economics, RWTH Aachen University, Aachen, Germany, [471]Department of Experimental Psychology, University College London, London, United Kingdom, [472]Department of Management, Society and Communication, Copehagen Bussiness School, Copenhagen, Denmark, [473]Insitute of Psychology, Osnabrück University, Osnabrück, Germany, [474]Department of Social Sciences, Otto von Guericke University Magdeburg, Magdeburg, Germany, [475]Faculty of Social Science, RPTU University of Kaiserslautern-Landau, Kaiserslautern, Germany, [476]Department of Eductional Science, University of Groningen, Groningen, The Netherlands, [477]School of Psychology, Keele University, Keele, United Kingdom, [478]School of Business, Stevens Institute of Technology, Hoboken, NJ, USA, [479]Centre for Fertility and Health, Norwegian Institute of Public Health, Oslo, Norway, [480]Jackson School of Global Affairs, Yale University, New Haven, CT, USA, [481]NSI, Inc., [482]Department of Communication Studies, Northwestern University, Evanston, Il, USA, [483]Department of Computer Science, [484]Department of Marketing, Clemson University, Clemson, SC, USA, [485]Department of Social & Human Research, Romanian Academy, Cluj-Napoca, Romania, [486]Babeș-Bolyai University, [487]Institute of Psychology, Jagiellonian University, Krakow, Poland, [488]Center for Law and Economics, ETH Zürich, Zürich, Switzerland, [489]Education, Research, Statistics, and Evaluation, University of Virginia, Charlottesville, VA, USA, [490]History and Art History, Utrecht University, Utrecht, The Netherlands, [491]Graduate School of Advanced Science and Engineering, Hiroshima University, Higashihiroshima, Japan, [492]Business, Operations and Strategy School, University of Greenwich, London, United Kingdom, [493]Department of General Psychology, University of Padova, Padova, Italy, [494]Department of Sociology, University of Novi Sad, Novi Sad, Serbia, [495]Leiden Institute of Brain and Cognition, Leiden University, Leiden, The Netherlands,

[496]Business, LAB University of Applied Sciences, Lappeenranta, Finland, [497]Visiting Researcher, LUT University, Business School, Finland, [498]Turku School of Economics, University of Turku, [499]Department of Psychiatry and Behavioral Sciences, Stanford University, Palo Alto, CA, USA, [500]Department of Sociology and Social Psychology, University of Cologne, Cologne, Germany, [501]Research Group Neural Circuits, Consciousness and Cognition, Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany, [502]Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands, [503]Human Resource Studies, Tilburg University, Tilburg, The Netherlands, [504]Department of Psychology, University of Basel, Basel, Switzerland, [505]Department of Organizational Behavior, INSEAD, Singapore, Singapore, [506]Independent Researcher, , United Kingdom, [507]Center for Social and Cultural Psychology, Université Libre de Bruxelles, Brussels, Belgium, [508]Optentia Research Unit, North-West University, Vanderbijlpark, South Africa, [509]Department of Sociology, Utrecht University, Utrecht, the Netherlands, [510]Methodology and Statistics unit, Institute of Child and Education Studies, Leiden University, Leiden, The Netherlands, [511]Department of Economics, Lund University, Lund, Sweden, [512]Department of Psychology, Ithaca College, Ithaca, NY, USA, [513]Department of Humanistic Studies, University of Urbino, Urbino, Italy, [514]Department of Political Science, University Duisburg-Essen, Duisburg, Germany, [515]IRCCS San Camillo Hospital, Lido Venice, Italy, [516]Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Center, Seattle, WA, USA, [517]Amsterdam School of Communication Research, University of Amsterdam, Amsterdam, The Netherlands, [518]Nuffield College, University of Oxford, Oxford, United Kingdom, [519]Economics, Newcastle University Business School, Newcastle upon Tyne, United Kingdom, [520]Faculty of Medicine, Imperial College London, London, United Kingdom, [521]Perceptive Discovery, Perceptive, London, UK., [522]Arctic Centre, University of Lapland, Rovaniemi, Finland, [523]Darden School of Business, University of Virginia, Charlottesville, VA, USA, [524]Psychology, University of Michigan, Ann Arbor, MI, USA, [525]Cologne Institute for Information Systems, University of Cologne, Cologne, Germany, [526]Department of Psychology I, Clinical Psychology and Psychotherapy, University of Würzburg, Institute of Psychology, Würzburg, Germany, [527]Accounting and Finance Group, University of Edinburgh Business School, Edinburgh, United Kingdom, [528]Leibniz Institute for Financial Research SAFE, Frankfurt, Germany, [529]Psychological Sciences, Western Kentucky University, Bowling Green, KY, USA, [530]University of Kentucky Medical School, [531]Department of Psychological Sciences, Swinburne University of Technology, Hawthorn, Victoria, Australia, [532]Department of Experimental Psychology, Ghent University, Ghent, Belgium, [533]Department of Psychology, Humboldt University of Berlin, Germany, [534]Department of Business Economics, University of the Balearic Islands, Palma, Spain, [535]Cognitive Science Department, Central European University, Vienna, Austria, [536]Italian Institute of Technology, Genoa, Italy, [537]Data and Analytics, Texas State Technical College, Waco, TX, USA, [538]IOE, UCL's Faculty of Education and Society, Psychology and Human Development, University College London, London, United Kingdom, [539]Cameron School of Business, University of North Carolina Wilmington, Wilmington, NC, USA, [540]Department of Psychology, Chinese University of Hong Kong, Hong Kong SAR, China, [541]Neuroimaging Center (NIC), Focus Program Translational Neuroscience (FTN), Johannes Gutenberg University Medical Center, Mainz, Germany, [542]Leibniz Institute for Resilience Research (LIR), Mainz, Germany, [543]Department of General and Transplantation Surgery, Medical University of Warsaw, Warsaw, Poland, [544]Department of Psychology, Ludwig-Maximilians-Universität München, München, Germany, [545]Universitas Airlangga,

[546]Department of Psychology, Ohio University, Athens, OH, USA, [547]Research Group Neural Circuits, Consciousness, and Cognition, Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany, [548]Institute of Psychology, Goethe University of Frankfurt, Frankfurt am Main, Germany, [549]Department of Psychology, Claremont McKenna College, Claremont, CA, USA, [550]Independent Researcher, Washington, DC, USA, [551]Department of Psychology, University of Milano-Bicocca, Milan, Italy, [552]Department of Economics, Ben-Gurion University of the Negev, Beer Sheva, Israel, [553]Center for Open Science, Charlottesville, VA, USA, [554]University of Virginia

*Balazs Aczel and Barnabas Szaszi are shared first and corresponding authors.

## Abstract

The same dataset can be analysed in different justifiable ways to answer the same research question, potentially challenging the robustness of empirical science[1–3]. In this crowd initiative, we investigated the degree to which research findings in the social and behavioural sciences are contingent on analysts' choices. We examined a stratified random sample of 100 studies published between 2009 and 2018, where for one claim per study, at least five re-analysts independently re-analysed the original data. The statistical appropriateness of the re-analyses was assessed in peer evaluations, and the robustness indicators were inspected along a range of research characteristics and study designs. We found that 34% of the independent re-analyses yielded the same result (within a tolerance region of +/- 0.05 Cohen's *d*) as the original report; with a four times broader tolerance region, this indicator rose to 57%. Regarding the conclusions drawn, 74% of analyses were reported to arrive at the same conclusion as in the original investigation; 24% to no effects/inconclusive result, and 2% to the opposite effect as in the original investigation. This exploratory study suggests that the common single-path analyses in social and behavioural research should not simply be assumed to be robust to alternative analyses[4]. Therefore, we recommend the development and use of practices to explore and communicate this neglected source of uncertainty.

**Keywords:** analytical variability, crowdsourcing science, data analysis, research credibility, robustness, scientific transparency

## Introduction

Over the past decade, social and behavioural scientists have been striving to enhance the robustness, objectivity, and replicability of their findings through systemic reforms in the conduct and communication of empirical research. Practices such as preregistration[5], registered reports[6], multisite replications[7], analytical reproducibility checks[8,9], and automated result validation techniques[10] have been investigated and recommended to produce robust and replicable findings. An important aspect of robustness has yet to be systematically charted across these sciences: the contingency of the results on researchers' analytical choices.

In a typical research pipeline, the collected empirical data are analysed by a single analyst or team, and the published report presents a conclusion based on one analytical path, occasionally accompanied by a few robustness tests. The peer review process aims to ensure that the analysis approach meets the statistical and field-specific standards. However, this procedure does not systematically ascertain whether justifiable alternative analytical choices could have led to different results.

Theories and empirical designs rarely constrain analysts to a single analytical path. Many degrees of freedom exist in how researchers operationalise their variables, process their data, construct their statistical model, select algorithms and software for model estimation, and define their inference criteria; whether they follow frequentist, Bayesian, or likelihoodist analytical approaches; use machine learning or conduct computational modelling to answer the same research question[1,4]. This inherent freedom of the analyst constitutes the so-called *analytical variability* contained within empirical projects, a key component in the robustness of the statistical results. In practical terms, it is the manifested variation among the choices independent scientists consider justified. Fig. 1 lists some sources of analytical variability that can manifest themselves in analysts' statistical results and the conclusions drawn from the results.

One way to explore analytical variability is to employ a *multiverse* methodology[2,11] in which the analyst conducts all combinations of analytic choices they are able to generate across a wide range of reasonable scenarios. Alternatively, in the *multi-analyst* approach, multiple analysts analyse the data following their best judgement. The latter approach requires more organisation, but it takes advantage of alternative expert perspectives without the combinatory expansion of the number of results. A multi-analyst approach also examines naturally occurring variation, empirically answering the counterfactual question of what might have happened if another investigator had considered the same research question using the same data.

Multi-analyst projects[3,12–24] have provided some evidence of the extent to which analysts' individual choices influence the results and conclusions. From economics to neuroscience, these explorations demonstrated that the robustness of empirical findings can be compromised by researcher degrees of freedom[25]. The estimates of previous multi-analyst studies suggest that the variability in effect-size estimates attributable to analytical heterogeneity can exceed the variability one would expect due to sampling error[26].

Do we know how robust published findings are to analytical choices across the social and behavioural sciences? One could argue that multi-analyst projects so far have been purposefully conducted in research areas with little consensus on the best analytical approach or were motivated to demonstrate the potential effect of analytical choices and, therefore, may represent rare cases where alternative analyses produce important differences in results. For example, perhaps the datasets selected afforded greater researcher degrees of freedom than is typical, raising issues of the generalizability of the findings to scientific research more broadly. Differences between academic methodologies and fields also seem plausible - for example, the relatively simple experiments sometimes used in social psychology and behavioural economics may contain fewer analytic decisions than the complex longitudinal observational datasets used in macroeconomics and finance, and thus be more analytically robust in general[22]. To the extent that this is the case, the findings from the existing multi-analyst projects could be biased towards worst-case scenarios, and the traditional analytical practice and review system may not require fundamental adjustments. If, on the other hand, observed results are contingent on the analyst's choices across fields, methodologies, and types of datasets, then the scientific literature could be less robust than is often assumed. If so, the general practices of how we conduct, report, and review empirical analyses should be reformed to address this source of uncertainty.

After conducting 504 re-analyses with the involvement of 457 independent re-analysts on a stratified random sample of 100 social and behavioural studies, we conducted strictly exploratory analyses in order to describe the patterns in the findings. Inspecting the results across different research characteristics and study designs gives rise to a number of hypotheses for future research on how to maximize transparency and address this often-neglected component of scientific uncertainty.

### *Variability of the results*

To explore the robustness of published claims, we selected a key claim from each of our 100 studies, in which the authors provided evidence for a (directional) effect. We presented each empirical claim to at least five analysts along with the original data and asked them to analyse the data to examine the claim, following their best judgement and report only their main result. The analysts were encouraged to analyse those studies where they saw the greatest relevance of their expertise. Therefore, in this study, analytical variability, as a key component of robustness, is defined as the variation among the analytical results when different analysts are provided with the same research questions and the same data.

First, we explored the degree to which the re-analysts produced the same statistics in the re-analysis of each study. We found that in 81% of the studies, the corresponding analysts reported different statistics regarding statistical test families (such as *t*-tests, F-tests, and $\chi 2$ tests) and their values (after rounding them to two decimal places).

A challenge in any multi-analyst project is to find a common metric that allows the results of the different analyses to be compared. A practical solution is to transform the reported point estimates into a standard effect size measure. Although these transformations have limitations and their calculation relies on assumptions that may not hold in all considered analysis settings[25,27–29], for the sake of comparability, we decided to compute Cohen's *d* for each re-analysis, wherever it was feasible. (For an alternative approach, see Supplementary Fig. 1) The methods, materials, analysis plan, peer evaluation, and data management strategy of the project were preregistered on the OSF repository of the project (Deviations from the registered plan are reported and explained in the 'Deviations from preregistration' supplementary document). In our preregistration, we defined that we consider two results qualitatively the same when their effect sizes are within the tolerance region of ±0.05 Cohen's *d.* However, below, we also report analyses with alternative tolerance regions. Our results revealed how far the new estimates were from the original ones (Fig. 2a) and how often the effect sizes of the re-analyses fell within this tolerance region (Fig. 2i).

We found that in 5% (5 out of 95) of the studies for which we could obtain the original effect size all re-analysis effect sizes were inside the tolerance region (+/- 0.05 Cohen's *d*) of the result of the original study (Fig. 2a). Out of the 396 available re-analysis effect sizes, 34% were inside the tolerance region. As a robustness test of our analysis, we explored the degree to which we would observe different results with different tolerance regions. With a four times broader tolerance region (+/- 0.20 Cohen's *d*), in 23% of the studies, all corresponding re-analysis results were inside the tolerance region. Further, out of the available 396 re-analysis effect sizes, 57% (224) were inside of this region (Extended Data Fig. 1a).

Alternatively, we could define the tolerance region as the percentage of the given effect size. As an additional robustness test, we varied the tolerance region between +/-5% and +/-20% but it barely made any difference regarding the percentage of robust studies (Extended Data Fig. 1b).

We next considered whether these robustness results vary by the disciplines of the studies, the study designs, the expertise of the analysts, their prior familiarity with the data, and the sample size in the data. Fig. 2b and Fig. 2c show the results for the major disciplines in our sample (>=10 studies). For Fig. 2c, we created an effect-size estimate range for each study as the numerical difference between the highest and lowest estimate of re-analysis effect sizes. In our reading, the listed disciplines do not yield large differences in the robustness of the results. Still, it is reasonable to think that the level of analytical robustness in different disciplines can be influenced by the type of studies that are commonly conducted there. For example, one could conjecture that empirical claims based on observational data show lower robustness of the conclusions since they likely involve more researcher degrees of freedom in terms of viable analysis paths than experimental research settings. Fig. 2d and Fig. 2e explore this question and indicate that the results of studies with observational study designs have lower analytical robustness in our sample, relative to experimental designs (also see Tables S5 and S6).

Considering the analytical variability found in the statistical results of the re-analyses, one immediate concern is that it could be an artefact of a lack of analytical expertise among some re-analysts. Therefore, we explored whether our robustness results exhibit a different pattern when examined in relation to the self-reported statistical expertise of the re-analysts. Visual inspection of Fig. 2f shows no support for this proposition, as a higher level of expertise corresponds with no increase or decrease in the ratio of the reported results being different from the original ones. It is noteworthy, however, that the level of self-perceived expertise was clustered in the higher end of the scale.

Re-analyzing published studies entails a potential risk of bias if the re-analysts' familiarity with a given study influences their choice of analysis. Re-analysts reported that they were familiar with the original study in only 8% of cases. Moreover, there was no more than 3% difference in robustness between those who were and those who were not familiar with the original study (Fig. 2g). For both groups, around two-thirds of the estimates fell outside our tolerance region. Finally, we were interested to see whether these robustness results would show a different pattern when considering sample size, as one could assume that studies with larger sample sizes could offer more robust results. Fig. 2h does not support this assumption as the density distributions of the sample sizes for results that are within and outside of the tolerance region are virtually the same. Therefore, studies with large sample sizes are not immune to analytical variability.

We next asked whether the re-analyses show a trend or shift in effect sizes compared to the results of the original studies. If the re-analysis effect sizes randomly vary around the original effect size, we would expect that they are larger or smaller than the original ones with an equal chance. Fig. 3a (re-analysis data trimmed at Cohen's *d* =< 5) and 3b (Cohen's *d* =< 1) indicate that re-analysis effect sizes show a tendency to be smaller than the original effect sizes as reflected in their best-fitting (least squares) line. The distribution of original and re-analysis effect sizes also supports this, as the peak of the density distribution of the latter is markedly lower. The mean effect size of the original results is 0.73 (Median = 0.43), whereas for the re-analysis it is 0.49 (Median = 0.35), Cohen's *d*, computed on *d*s =< 5. This result is consistent with the possibilities that original authors were biased to report larger effects than re-analysts, that re-analysts were biased to report smaller effects than original analysts, or both.

***Variability of the conclusions***

Another focal question of our study was whether the re-analysts reached the same qualitative conclusions as the original study analysts. To answer this question, we asked the re-analysts to implement any statistical re-analysis they deemed most appropriate to test the original claim using the original data, with the goal of arriving at a single conclusion. Across all individual re-analyses (*n* = 504), 74% of analyses were reported to arrive at the same conclusion as in the original investigation; 24% to no effects/inconclusive result, and 2% to an effect in the opposite direction as in the original investigation (Fig. 4a).

Out of 100 re-analysed claims, 34% were robust to independent re-analysis, such that all re-analysts reported that they found evidence for the originally reported claim. It is important to note, however, that this result is contingent on the level of agreement we use to define analytically robust findings. With a more liberal definition of analytical robustness, this value was 39% when analytical robustness was defined as >80% re-analysis agreement with the original conclusion, and it was 80% when this definition was >50% (the results with alternative levels of agreement are displayed on Fig. 4j).

We examined whether these results show a different pattern when inspecting them along the earlier-mentioned aspects of the analyses. Fig. 4b and Fig. 4c present the proportions of conclusions that were robust in each of the listed disciplines. Just as in the case of analyses of robustness of the statistical results, the listed disciplines do not manifest large differences in robustness of the conclusions, whereas their robustness may be influenced by the study designs most common in a given field or subfield. Fig. 4d supports this notion as it indicates that nearly half of the conclusions from experimental studies remained robust upon independent re-analysis, whereas less than one-third of observational studies yielded robust conclusions. Moreover, Fig. 4e indicates that, although the majority of re-analyses for both study designs reached the same conclusions as the original study, the figure was 13% higher for experimental studies than for observational studies. Just as for the robustness of the results, we can ask whether the deviation from the originally reported claim in terms of conclusions is explained by the re-analysts' lack of analytical expertise. Fig. 4f shows no support for this conjecture when evaluating the pattern of results as a function of self-reported statistical expertise. The same conjecture can be assessed by considering the quality of the submitted statistical analyses that were evaluated by peer evaluators on a subset of the analyses (see Methods). Fig. 4g shows that the proportion of inferentially robust conclusions is numerically larger for analyses that were rated as medium-quality by peer evaluators than for analyses that were rated as high-quality. Whether this pattern was a result of noise or whether more sophisticated analyses are characterized by greater heterogeneity in approaches and results should be the topic of future metascientific projects.

Just as for the analyses of the robustness of the statistical results, we were interested to see whether these results show a different pattern when inspecting them as a function of analysts' prior familiarity with the dataset. Although those familiar with the original study did report the same conclusion in a higher proportion than those who were not familiar, 17% of their re-analyses still indicated a conclusion different from the original one (Fig. 4h).

Again, we aimed to explore whether these robustness results would show a different pattern when considering sample size. As presented in Fig. 4i, the density distribution corresponding to the analyses with the different conclusion types shows a comparable spread, suggesting that the conclusions of studies with smaller and larger sample sizes appear to be similarly contingent on analytical choices.

For descriptive information about the re-analysts, peer evaluators, and additional robustness analyses, see Extended Data Figs. 2, 3, 4, and Supplementary Information's General descriptives, Demographics of the re-analysts, Peer evaluation, and Robustness analyses sections.

**Limitations**

This study has a number of limitations. First, our collection of 100 articles represents only a tiny fraction of all the empirical work in the social and behavioural disciplines. Despite our efforts to select a representative sample of published articles across disciplines from the investigated time period, we could not include studies when the underlying data were not obtainable, and we excluded studies when our screening attempt to analytically reproduce the original results following the published procedures failed. We cannot exclude the possibility that these prerequisites, in addition to the self-selection of the analysts, led to sampling bias.

Although we conducted more than 500 analyses, our project included only five independent analyses for most datasets, therefore, we do not know to what degree these analyses capture the full variability of analyses and results for the given research question and dataset. Also, since we re-analysed already published studies and the re-analysts were provided with these studies, the original analysis pipeline could have anchored some of the choices of the re-analysts. On the other hand, some analysts could have been motivated to produce alternative results, as it is a basic incentive of scientists to say something new.

While Cohen's $d$ has the advantage of being easy to compute and comparable across different analyses, Kümpel and Hoffmann[39] recently proposed the concept of generalised marginal effects (gMEs), an effect size metric that is both formally applicable and comparable across different statistical models. We had not originally planned to calculate standardised gMEs, and, accordingly, did not collect all required analysis outputs to compute them across the board. Still, we calculated gMEs for a sample of our studies to showcase their potential for future multi-analyst studies (Supplementary Fig. 1).

We presented some exploratory analyses, but there are many other factors to explore that could contribute to analytical variability (e.g., topical expertise). Finally, despite our best efforts to conduct quality checks on the re-analyses to ensure the soundness of the analytic strategies[16], it is possible that some of the discrepancies between the original and the new results are due to weaknesses in the re-analysts' approach rather than equally justifiable alternative analysis decisions. It is likewise possible that there are weaknesses in the original analysts' approaches. It is unknown whether the quality control processes for the re-analysts resulted in better, worse, or similar overall quality of analysis decisions as compared with the quality control processes for original analysts' decisions. The declared statistical expertise of the re-analysts makes us believe that the observed heterogeneity in analyses and outcomes is a good representation of variation in informed analysis decision-making in social-behavioural research.

**Discussion**

Are published results in the social and behavioural sciences robust to independent re-analyses? The present exploration shows considerable variability due to researcher degrees of freedom in statistical choices. Overall, when independent researchers analysed the same research question on the original data, 34% of studies remained robust to independent re-analysis in the sense that all re-analysts arrived at the same conclusion as the original analyst or analyst team. Notably, the new conclusions converged with the original ones in 74% of the individual re-analyses. Our descriptive results suggest a number of hypotheses concerning the circumstances in which we could expect greater analytical variability.

### *Why can there be multiple answers?*

Faced with the variability in the analysts' effect-size estimates and conclusions, one intuitive hypothesis is that the variation must be due to researcher characteristics, such as statistical or field-specific knowledge. Previous multi-analyst studies found little to no effect of researcher-specific characteristics, such as experience in the field or statistical expertise[16,19,28]. Instead, they suggest that analytic results are dependent on the particular choices that the analysts make among similarly acceptable data processing and analysis choices[28]. For example, when 46 independent analyst teams analysed the same speech dataset to answer the same research question, the authors concluded that "depending on the choice of how the speech signal is operationalised, researchers might find evidence for or against a theoretically relevant prediction" (p. 21)[28].

In line with previous findings, our results showed no strikingly different patterns across self-reported statistical expertise and experience in a matching field (see Fig. 2f, Extended Data Fig. 4, Supplementary Table 5, 6, 7, 8, and 9). At the same time, the few analysts who reported that they were familiar with the original article produced alternative results and conclusions at a comparable rate. More importantly, our peer evaluation process did not indicate that the analytical variability of the re-analyses was due to inadequate statistical practices. These results are in line with Menkveld et al.[22] in which the quality assessment of the proposed analysis pipelines did not statistically explain the results.

Another line of thought would suggest that the lack of robustness in the original published results reflects some conceptual ambiguity in the theories or methodology[30]. Research hypotheses are often short verbal expressions that do not force the specifications of the analyses. The underspecification of claims[31] could represent a major source of ambiguity in analytic decisions. We could not test the role of hypothesis ambiguity in a controlled manner, but it is a plausible contributor considering that social science theories often make general claims across many variables, creating theory-laden choice points regarding how constructs are operationalised, and how hypotheses are tested[32].

Regarding methods, we explored our results by separating them by experimental and observational study designs, and observed that the proportions of results and conclusions that were analytically robust were 15-20% higher for the experimental studies. The estimated range of effect sizes was also apparently wider for observational studies compared to experimental ones. This exploratory finding motivates the hypothesis that the increased control over data collection circumstances and the reduced number of variables in experimental versus observational research translate to more limited analytic flexibility. Notably, however, there was still substantial statistical variability among findings from experimental studies.

### *Why do these findings matter?*

Where multiple acceptable analytical paths exist, researchers can use this freedom opportunistically[33,34] and bias the results towards desired findings ("myside bias"[35]). The muchdiscussed credibility challenges in the social and behavioural sciences stem partly from the suspicion that the prevailing incentive systems for publication encourage researchers to report and interpret empirical data to serve non-epistemic goals such as storytelling[36]. Reform initiatives, such as the preregistration of research and analysis plans, aim to decrease researcher degrees of freedom to tweak the analytic method or the research question to the observed data. Would results in these fields become markedly more credible if every study was preregistered? Since preregistration is a protection against overfitting, we hypothesize that it would reduce or eliminate the observed finding that original analyses showed stronger evidence for positive results than re-analyses. However, we also hypothesize that preregistration would have little impact on the observed heterogeneity across alternative analysis strategies since registering and following a single analytic path constrains the analysts only from choosing opportunistically from the alternative analytical paths. Still, it does not confer any unique statistical or epistemic status to the pre-selected analytic path[27]. Unexplored but alternative justifiable analyses applied to the same data could still lead to very different results. The present exploration is clear about the presence of this variability in approaches, results, and inferences in the social and behavioural sciences. Without exploring this variability, authors cannot guarantee consumers of their research that the reported conclusions hold a privileged status over alternative conclusions.

***What can we do?***

The outcomes of this project suggest that the empirical answers to research questions in the social and behavioural sciences depend on the analytic paths taken to pursue them. Therefore, we advocate for the broader adoption of approaches that explore, recognise, and address the uncertainty created by analytical variability.

Two main types of solutions are (1) multi-analyst studies, such as our own, where multiple investigators independently follow their own approach, and (2) the multiverse[2,11,37] approach, where one investigator or team performs numerous analyses across the set of reasonable pipelines. Conducting exploratory studies to identify analytical uncertainties and holding out samples are further advisable practices to tackle analytical variability.

Project leaders aiming to conduct multi-analyst studies can consult various tutorial papers and guidelines. Aczel et al.[38] provide an expert consensus guideline on the entire life-cycle of multi-analyst projects from recruiting suitable analysts, through conducting the project to the reporting of the outcomes. Kümpel & Hoffmann[39] offer a framework for synthesising objective outcome metrics. The Subjective Evidence Evaluation Survey[40] is a tool for systematically exploring and quantifying subjective measures of evidence in multi-analyst studies, allowing analysis teams to subjectively reflect on various aspects of evidence, such as coherence, robustness, and relevance, as well as the quality of the research design and data.

Multiverse analysis is also useful, especially when the dataset cannot be shared with other research groups due to confidentiality reasons or when there are insufficient human resources to recruit several independent analysts. Several guideline papers help researchers conduct and interpret such analyses[2,37,41–43].

Recently, many scholars have called for a stronger focus on replication in science[44]. Similar to preregistration, however, replications are unlikely to help address the robustness of results to multiple analysis strategies as they intentionally repeat the same (or at least a very similar) analysis path. In this sense, replications can help detect bodies of work in which authors may have leveraged their researcher degrees of freedom to generate results that are in line with their own or the journal's expectations. All other things being equal, a severely *p*-hacked literature should contain fewer replicable findings. And yet, replicability does not eliminate analytical variability itself. Nevertheless, having multiple studies creates an opportunity to observe if analytical variability is, itself, replicable. For example, imagine that Study A provides evidence for a claim with Analysis 1 but not with Analysis 2. If several replications also find evidence for the claim with Analysis 1 but not with Analysis 2, then the analytic choices are directly implicated in how evidence for the phenomenon is observed. However, if it is random across replications whether Analysis 1 or Analysis 2 provides evidence for the claim, then the implications of the analytic variability are very different. The combination of replications and robustness investigations will facilitate the advancement of stronger theoretical underpinnings of the topics of study, and could reduce analytical variability in the long run by creating a more direct mapping between theory and measurement[30,11].

All in all, we argue that the scholarly communication system could foster more engagement with systematic and transparent robustness testing. As a starting point, the research data shared openly alongside codebooks and analysis scripts is a prerequisite for any assessment of analytical robustness. Research findings of particular scientific or societal importance could be accompanied by robustness reports[46] that summarise the results of alternative theory-motivated analytic choices by independent analysts. This publication format already provides a platform for analysts to scrutinise the fragility of the findings before they have a major impact on scholarship and policy( see **https://scipost-staging.org/JRobustRep**).

***What did we learn about the robustness?***

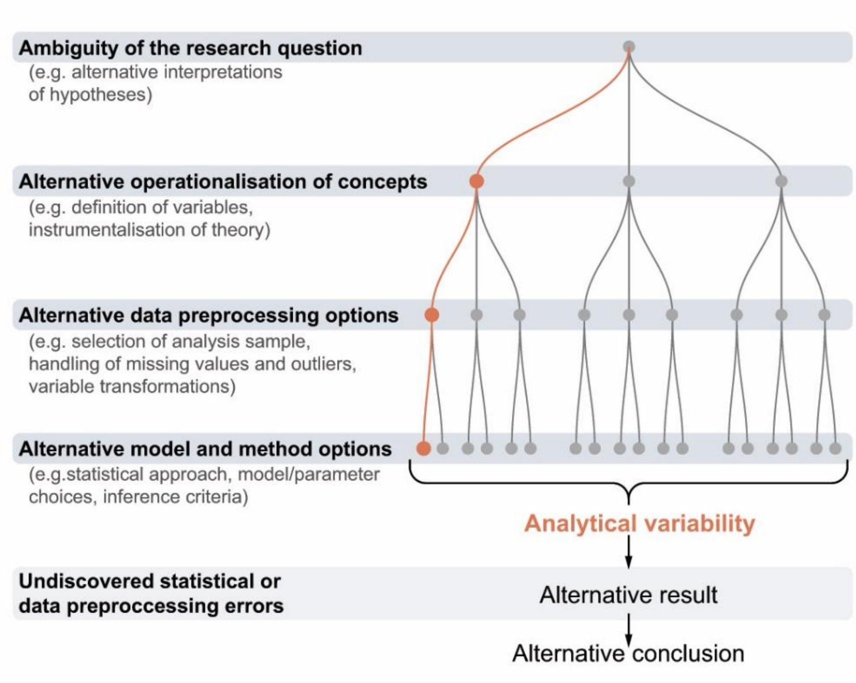
Our results support the view that the results in social and behavioural science studies are contingent on the analyst's choices, and if an analyst reports a single result from a single analytical path, they have not exhausted the possible answers that the dataset can provide. This finding aligns with the conclusions drawn by Wagenmakers, Sarafoglou, and Aczel[4], that the belief that "for any dataset, there exists a single, uniquely appropriate analysis procedure" and "multiple plausible analyses would reliably yield similar conclusions" (p. 424) are no more than statistical myths. Without multi-analyst and multiverse approaches, the fragility of empirical findings remains.

Nonetheless, we emphasise that an optimistic or pessimistic interpretation is a matter of perspective and greatly depends on what evidential support we expect from a given study. Therefore, whether a result is satisfyingly robust will always depend on our epistemic needs and the precision we expect from our results. We caution against using blanket rules in aggregating or interpreting results across different analytical approaches within the same investigation.

Objectivity is a fundamental ideal of science, implying that claims about the world should not be contingent on the predispositions of the claimant. What our results reveal is not that we must distrust or reject the results of the past, including the studies we analysed. Instead, they suggest that we should adopt greater caution about the evidence that single analytical paths can offer to support social and behavioural science claims. We believe that the limitations of "single-shot" analyses cut across numerous scientific disciplines. Methodological innovations, such as multi-lab collaborations, multi-analyst approaches, or multiverse methods, could increase the robustness of the social and behavioural sciences, and perhaps more broadly, in other empirical fields.

## Figure legends

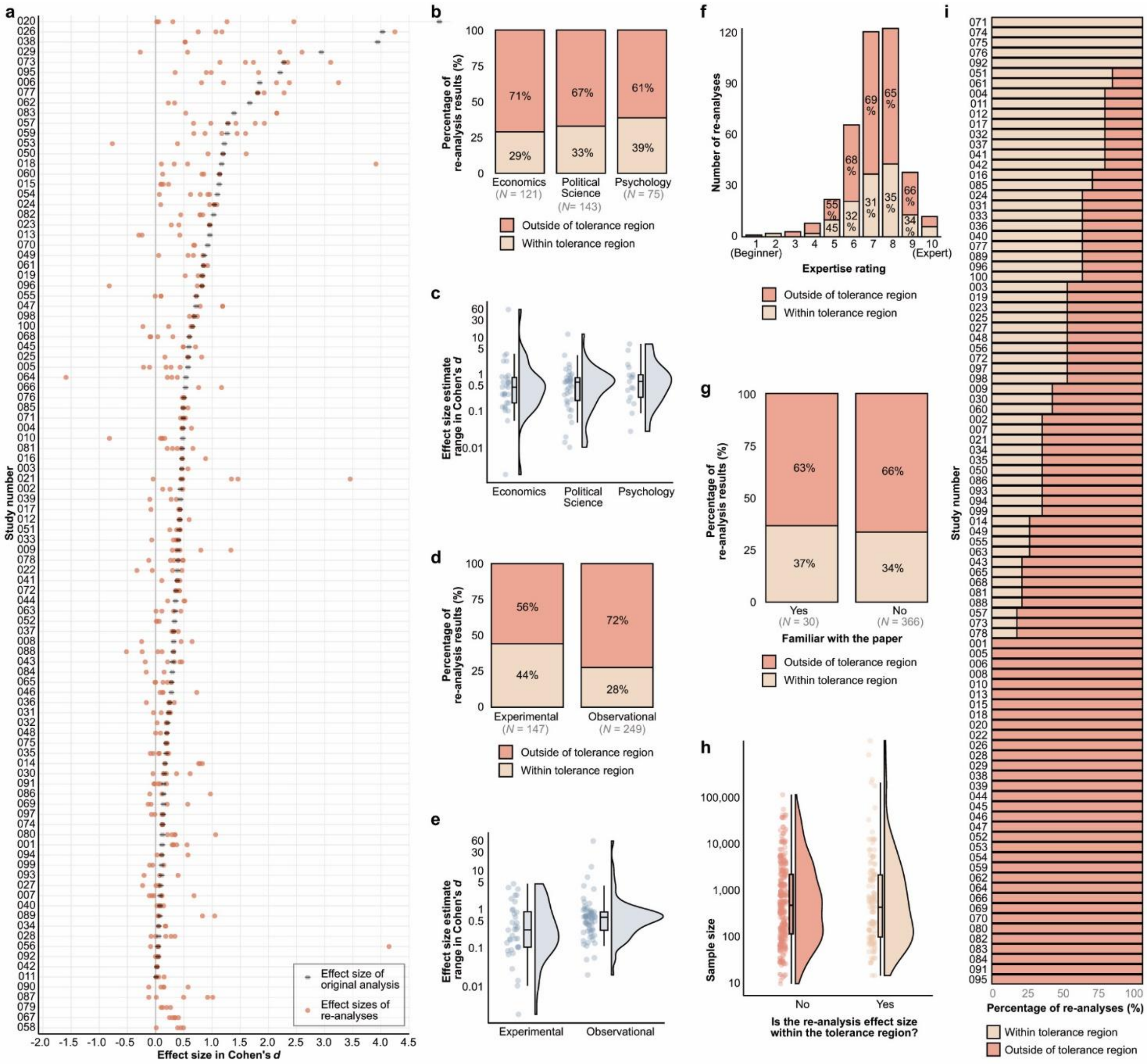
**Fig. 1**
*Major sources of analytical variability.*




*Note.*
Analytical variability can arise from the ambiguity of the research question, the alternative operationalisations of the concepts, variations in data preprocessing options, or model and method choices, as well as from undiscovered statistical or data processing errors.

**Fig. 2**

*Analytical robustness of the statistical results.*

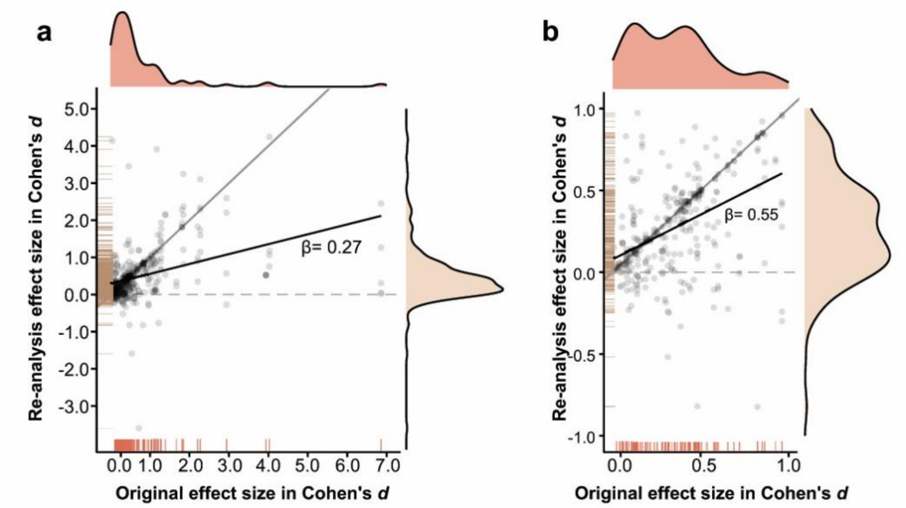



*Note.*

**a,** Effect size of the original analysis (gray square; all represented as positive values) and the effect sizes of the re-analyses (red dot) for each study. The figure displays 415 re-analysis effect-size estimates that were convertible to Cohen's *d* and excludes effect sizes outside the [-2, 4.5] range. For the five studies listed at the bottom of the figure, we could not determine the original effect size due to missing information. Study numbers correspond to studies listed in https://osf.io/mkwhn. The studies are ordered by the size of the original effect size. **b,** Percentage of re-analysis results falling within or outside of the tolerance region of the original results of the studies by major disciplines. The figure displays the count of re-analyses next to each discipline name. **c,** Distributions of effect-size estimate ranges calculated per study for each major discipline. **d,** Proportion of re-analysis results falling within or outside of the tolerance region of the original results of the studies by study type. The figure displays the count of re-analyses next to each discipline name. **e,** Distribution of effect-size estimate ranges calculated per study for observational and experimental studies. **f,** Percentage of re-analysis results falling within or outside of the tolerance region of the original results of the studies by self-rated expertise (1= beginner, 10 = expert). **g,** Percentage of re-analysis results falling within or outside of the tolerance region of the original results of the studies by declared familiarity with the study. **h,** Distribution of sample sizes separately for re-analysis effect sizes falling within or outside of the tolerance region of the original results. **i,** Proportion of effect sizes falling within the preset tolerance range (+/- 0.05 Cohen's *d*) for each study.

**Fig. 3**

*Original study effect size versus re-analysis effect size.*

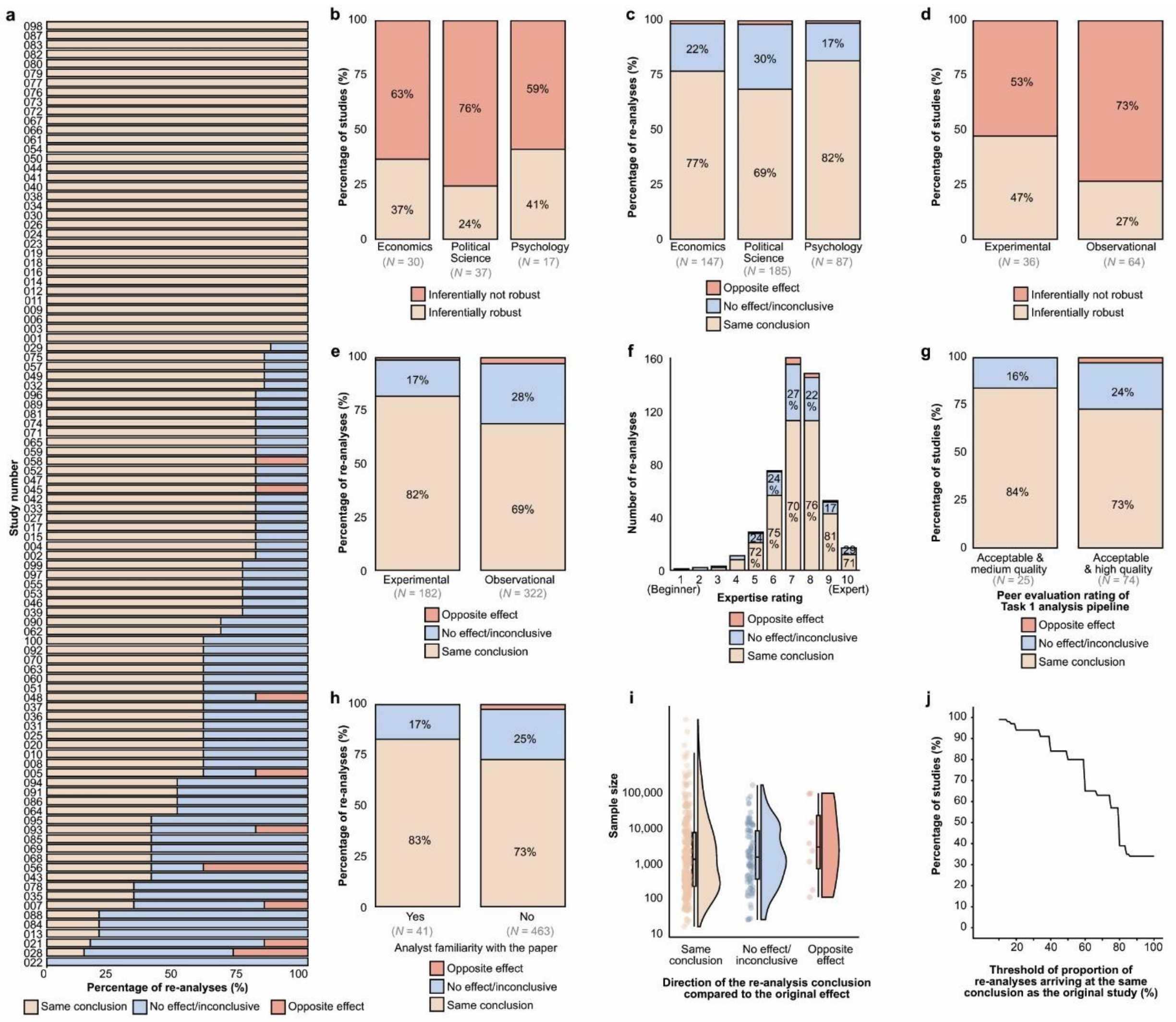



*Note.*

The thin diagonal line represents an ideal case when the re-analysis effect sizes are equal to original effect size, the thick line shows the best-fitting (least squares) line of the displayed dots. Density plots of original ($n$ = 95) and re-analysis ($n$ = 504) effect sizes are parallel to their respective axis. β refers to the regression slope. Figure **a** shows effect sizes Cohen's $d$ =< 5, Figure **b** displays the same for effect sizes Cohen's $d$ =< 1.

**Fig. 4**

*Analytical robustness of the conclusions.*




*Note.*

**a,** Proportion of same conclusion, no effect/inconclusive results, and opposite direction conclusions for each study. Study numbers correspond to studies listed in https://osf.io/mkwhn. **b,** Proportion of inferentially robust results (i.e., all re-analyses arrived at the same conclusion for the given study) by major disciplines (more than 10 studies in our collection: Economics, Political Science, and Psychology). **c,** Proportion of same effect, no effect/inconclusive results, and conclusions in the opposite direction of the original studies by major discipline. The number of re-analyses is displayed below each discipline. **d,** Proportion of inferentially robust results by study design (experimental vs. observational). The number of re-analyses is given below each study design. **e,** Proportion of same conclusion, no effect/inconclusive, and opposite effect of the re-analyses by study type (experimental, observational). **f,** Proportion of same conclusion, no effect/inconclusive, and opposite effect of the re-analyses by self-rated expertise (on a scale of 1 (Beginner) to 10 (Expert)). **g,** Proportion of inferentially robust studies by the acceptability of the analysis pipelines according to the peer evaluators. For this figure, we included only studies with more than one peer evaluation and where the peer evaluators agreed on their rating. The figure shows only the rating options with 5 or more re-analyses in that category. **h,** Proportion of same conclusion, no effect/inconclusive, and opposite effect of the re-analyses by declared familiarity with the study. **i,** Distribution of the sample size of the re-analyses resulting in the same conclusion, no effect/inconclusive, and opposite effects. Sample size values were available for 345 re-analyses. **j,** Percentage of studies with robust conclusions above different levels of re-analysis consensus. Re-analysis consensus refers to the agreement among the conclusions drawn by the original study and the independent re-analyses.

# Methods

All methods and procedures in this study were vetted by a panel of experts with prior experience in multi-analyst studies or who are specialists in the relevant methodology (see Additional Details of Method within Supplementary Information).

## Preregistration

The methods, materials, analysis plan, peer evaluation, and data management strategy of the project were preregistered on the OSF. Deviations from the registered plan are reported and explained in the "Deviations from preregistration" supplementary document.

### *Ethical considerations*

The datasets resulting from this project were not considered human subject research and are covered under an umbrella ethics protocol that was managed by the Center for Open Science (COS) (BRANY SBER IRB protocol #21-056-749), with concurrence from the United States Naval Information Warfare Center Pacific, HRPO. The institutional ethics board of the Faculty of Education and Psychology at Eötvös Loránd University, Budapest, Hungary, determined that the re-analysts are not considered research participants and that the project raises no ethical concerns.

## Materials

### *Selection of studies*

The selection of studies was completed in two stages. In the first stage, the SCORE team created an initial study and claim collection. From this collection, we selected our sample using additional criteria.

In the SCORE project, a stratified random sample of 600 articles was identified from a larger pool of randomly stratified ~30,000 articles from 62 journals, published between 2009 and 2018. The journals covered the main branches of social and behavioural sciences (criminology, economics, education, health-related, marketing/organisational behaviour, management, political science, psychology, public administration, sociology). To obtain the original studies, the following steps were taken: First, the paper was reviewed. If data and/or code were available, they were downloaded and saved into a project on OSF. If data and/or code were not available, the SCORE team attempted to contact the corresponding author to request that they share the data and code used for the original publication. Studies were excluded from the sample if they did not contain at least one inferential test using non-simulated, human data, where human data are defined at any level of human organisation (e.g., the individual person, family, political entity, firm, economic unit). The majority of the studies were tested for analytic *reproducibility* using the original specification, which is to be distinguished from *robustness* to alternative specifications. Analytic reproducibility was tested in cases when both original data and code were available (n = 63), or when the original data were available but the original code had to be adapted by the SCORE team in order to successfully reproduce the result (n = 7). If data were available but the original code was not, SCORE sourced a collaborating lab to generate new analytic code for the reproduction (n = 10). If data and code were not available, the collaborating lab used the secondary source data, which were shared upon request (acquired by SCORE), alongside newly generated analytical code for the reproduction (n = 11). Some reproductions were never attempted (n = 9). If the analytic reproduction failed, the paper was removed from the pool. Therefore, the present project focused solely on robustness to alternative specifications and did not conduct direct reproducibility checks using the original specification, as these had already been carried out by SCORE. Further details of the SCORE methodology (list of journals, selection process, etc.) are available in the original report[45].

In the present work, a further requirement of the selected studies was to contain a single inferential statistical test result that corresponded to the claim with our instructions. Thus, we ensured that given the claim and the instructions, no other statistical result could correspond to the claim in the original article. If all potential claims from the study were too ambiguous and, therefore, could not be linked with a single inferential test statistic with the specification instructions, the study was excluded from our sample. The above-described study selection process was continued until we reached our target number of 100 studies, corresponding claims, and datasets.

The selected studies and all available corresponding data and materials were made available to the re-analysts so that they could fully understand the selected claim and approach. There are trade-offs for how much information to give to the re-analysts to conduct re-analyses. Complete blinding of the original analysis strategy would ensure an entirely independent decision-making process about how to analyse the data. However, in much scientific writing, there is insufficient clarity in the description of the theoretical background, rationale, and specification of the conceptual model to be tested. In some papers, there is a clean break between these and clear hypotheses to test. In other papers, the narrative intermixes theoretical statements and analysis decisions and may not clearly state hypotheses or how they correspond with observed results. As a consequence, attempts to blind papers inevitably lead to variation in what is blinded across papers and many subjective decisions about what should be blinded (because it provides information about analysis strategy) and what can remain unblinded (because it provides information about theory and rationale). A major risk of those blinding decisions is that important information could be removed, which would weaken the re-analysts' ability to conduct a fair re-analysis of the original claim. As such, we opted for complete transparency of the original article so that no potentially important information was missing for the re-analysts, and we instructed re-analysts that they should create an analysis plan based on their own decisions for how best to assess the study's claim. On balance, this increases the risk of dependent decision-making but reduces the risk of misspecification of the hypothesis and rationale of the original research. In this context, we judged the latter to be a more important precondition for conducting an informative study.

### *Claim selection*

Claim selection was built on Phase 1 of the SCORE project effort. The claims identified for Phase 1 of SCORE were executed according to a "single trace" approach, where only a single claim trace was extracted from the article, which corresponded to one statistically significant inferential test result (see https://docs.google.com/document/d/1yKvjMFaIcwCLq1k-02m_Y1G937yMIc__1JK4OfMay40/edit?usp=sharing). Within the current project, first, the lead team ensured that the extractions (i) are understandable, (ii) contain only one claim, (iii) indicate the direction of the effect, (iv) there is a statistical hypothesis test-based result provided in the article that corresponds to the claim; and (v) the claim was phrased on a conceptual and not statistical level. If not, then they extracted the part of the claim that is relevant, or if this could not be achieved, they selected another more suitable sentence from the abstract, or if this could not be achieved, they searched for another suitable sentence from other parts of the article that could satisfy all of our criteria. When none of these steps presented a claim that satisfied the expectations, then the given article was not used in our study (for their list and explanation of dropping, see our data table). Where an expression of a claim has been judged by the lead team as ambiguous or rhetorical, they substituted the expression with an ellipsis mark (e.g., "dramatically increased" to " … increased") while preserving the original wording and the meaning of the claim. Only in cases where, due to the selection, the wording of the claim became complicated, ungrammatical, or contained an ambiguous definition or an unexplained abbreviation, did the core team make necessary (and marked) adjustments in the grammar or wording of the claim while preserving the original meaning of the extraction. For example, the following selection "Three factors increase the salience of the proliferation threat: (1) prior violent militarized conflict…" was changed to "[prior violent militarized conflict] increase[s] the salience of the proliferation threat …". The list of claims can be found at https://osf.io/mkwhn.

### *Analysis instructions*

For the re-analysts' second task, instructions were needed in cases where the original paper contained more than one statistical analysis corresponding to the high-level claim, in order to be able to compare the new result to the one in the original paper. For this, the lead team prepared certain instructions (e.g., data selection, exclusions) that single out only one statistical result in the original paper. The instructions always remained circumstantial (e.g., data selection, exclusions, choice of measurement) and never gave direct instructions to the choice of statistical approach or full specification of the model.

## Procedures

### *Re-analyst recruitment*

Our preregistered aim was to have at least five independent re-analyses carried out for each of the 100 selected studies (Extended Data Fig. 5). Our choice of 5 analyses per study was led by practical considerations, as we judged that recruiting 500 analysts for a project is the limit of our capacity.

Participation in the project was advertised on social media, at conferences, in mailing lists (e.g., SCORE collaborator list), via personal networks, and in research newsletters. As a response to our recruitment call, 1141 researchers signed up to participate in our study. Out of these volunteers, 459 signed up to analyse at least one dataset and submitted their work by the deadline or an extended deadline. From all the eligible volunteers, we selected re-analysts and peer evaluators on a first-come, first-served basis. The expectation of participation in the study was experience with conducting statistical analyses, and this was communicated to the volunteers from the start of the recruitment. Re-analysts were informed that they would qualify as authors on the publication of this study if (1) they completed their analyses and submitted all required materials and the post-analysis survey on time; (2) their analyses passed the peer evaluation, and (3) they reviewed and approved the manuscript in time.

Re-analysts received a flat fee of 100 USD for each of their completed re-analyses (including both Task 1 and Task 2) if they submitted their work before March 2023, the deadline of the grant budget, unless they were from an embargoed country, in which case we were unable to transfer any payment. Peer evaluators received a flat fee of 10 USD per peer evaluation. Any further volunteers were informed that this payment did not apply to them.

Upon joining the project, the volunteers for re-analysis were required to accept the project requirements. They were informed about (a) their tasks and responsibilities; (b) the project confidentiality agreements; (c) the plans for publishing the research report and presenting the data, analyses, and conclusion; (d) the conditions for an analysis to be included or excluded from the study; (e) that their names will be publicly linked to the analyses; (f) the re-analysts' rights to update or revise their analyses; (g) the project time schedule; and (h) the nature and criteria of compensation. Re-analysts were informed that, whereas they could consult other researchers during their analyses, they could not work in teams in this project. Before discussing the details of the analyses with others, the re-analysts were asked to ascertain that the person was not another re-analyst on that dataset. All communication materials of this study are openly available on the public repository of the project at https://osf.io/nvy8a.

***Assignment of analyses and tasks***

The following procedure was first piloted with two analysts to learn about the practical challenges and time demands of the following tasks. As the results of those analyses were not of central interest, we kept no records of them.

First, each re-analyst was asked to assign themselves to one study, but at later rounds of recruitment, we allowed re-analysts to complete analysis on another paper, other than the one they completed earlier. They were asked to choose those studies where they saw the greatest relevance of their expertise. The authors of the original study could not be the re-analysts of that study.

For several practical reasons, the re-analyses were not started at the same time for each study and each analyst. Firstly, it took us several rounds of recruitment to gather the target number of analyses for each study, mainly due to dropouts, delays, unplanned personal difficulties, and a shortage of staff. Secondly, our analysts found it difficult to retrieve, open, or interpret some of the datasets. In some cases, we had to reach out to the original authors, causing further delays in the project.

The task of the re-analysts was to reflect on the corresponding claim (see claim selection) by re-analyzing the corresponding data. The re-analysts were provided with access to the datasets, extracted claims, the original articles, and all the corresponding materials. They were informed that their analyses should be conducted preferably with scripts that could reproduce all their results (including data preprocessing, extraction of test statistics and p-values/Bayes Factors, computing effect-size measures, etc.), but they could use the statistical software of their choice to produce an analysis script. Re-analysts were asked to write and structure their code such that others could understand their analysis scripts (e.g., by annotating the different analysis steps), and they were also informed that the analysis scripts from all analysts would be made publicly available with their names linked to the analyses.

Re-analysts received two main tasks for each study, where Task 2 was given after the completion of Task 1. Once Task 1 was submitted, the analysts could not change the submission of Task 1 unless they were asked by the lead team to provide some missing information from their analysis.

**Task 1** The re-analysts were asked to reflect on the selected claim by re-analyzing the corresponding data. They could conduct and report as many analyses as they wished, but they had to draw a single conclusion from their analysis. They were asked to report their analyses and indicate whether their results provided evidence for the relationship/effect as claimed by the original study.

**Task 2** For this task, the re-analysts had to produce only one statistical result corresponding to the claim they studied in Task 1, which would be compared to a statistical result in the original paper. The lead team provided certain instructions (e.g., data selection, exclusions) for this analysis to be able to compare the new result to one result in the original paper (see Analysis instructions section). Re-analysts were asked to report their results in terms of statistical families of $r$, $z$-, $t$-, $F$-, or $\chi^2$ tests (or their non-parametric versions). In addition, they were asked to report sample sizes (e.g., per group) and the corresponding degrees of freedom. By this means, most results could be translated into standardized coefficients by the coordinators.

The reason for requiring two analyses from the re-analysts was that they served two different aims. The results of Task 1 aimed to answer our first preregistered project question: "Do different analysts arrive at the same conclusions as the analyst of the original study?", whereas the results of Task 2 aimed to answer our second preregistered project question: "Do different analysts arrive at the same effect estimates as the analyst of the original study?" We found that asking only one of the tasks would not have been sufficient to fully address both questions. In Task 1, researchers were not constrained to one analysis, so they could have produced more than one statistical result in order to draw a conclusion from the dataset. Therefore, in Task 1, it was not guaranteed that we would be able to select a single effect size from each analyst in order to answer our second project question. Another challenge to finding an answer to our second question was that in some of the original articles, one claim could have had more than one corresponding statistical result listed. In these cases, we prepared instructions for Task 2 in order to single out only one statistical result in the original paper. For example, if the original study contained two corresponding regression models, one with some exclusions and one with no exclusions, then we chose one of them (e.g., the latter), and instructed the re-analysts not to apply any exclusions to the analysed data. In all other regards, re-analysts were free to conduct their calculations according to their best judgment.

After completing the analysis and writing up the methods, results, and conclusion, re-analysts were expected to upload their analysis code (if available) to the corresponding OSF folder. Their reported methods, results, and conclusions were collected via an online form (see https://osf.io/fjnhz/). When uploading the materials, they were also asked to fill out a post-analysis survey. All major communications between the core project team and re-analysts from the study are openly available on the public repository of the project.

## Peer evaluations

The goal of peer evaluation in this project was to assess whether the applied analytical choices are acceptable and whether the reported conclusion follows from the statistical results. By acceptable, we mean that peer evaluators agree that the analysis pipeline is within the variations that could be considered appropriate by the scientific community in addressing the given analytical task.

The peer evaluation phase did not address potential errors in translating the description of the analytic methodology into analysis scripts. To mitigate potential gross errors in the analysis, peer evaluators were provided with a thorough and standardised description of the results and conclusions obtained using the described analysis, including sample sizes, the effect size, the test statistic, and degrees of freedom. From the description of the dataset, the description of the analysis, and the reported results and conclusions, peer evaluators were able to identify potential flaws in the implementation of the analysis that could stem from errors and/or mismatches.

### *Assignment of the analyses*

When assigning the volunteer peer evaluators to analyses, the initial rule was that they should not evaluate any re-analyses conducted on datasets they had re-analysed as a re-analyst. In practice, for logistical reasons, this rule was applied in all but six cases (i.e., 99% of peer evaluations were carried out on a dataset that was different from the dataset they analysed themselves). They were asked to choose to evaluate those analyses where they see the greatest relevance of their expertise. If, after choosing a study to evaluate, a peer evaluator did not feel sufficiently skilled/experienced to judge whether the proposed analysis was acceptable, he/she was told not to fill out our template and should return the re-analysis to the pool and choose a new one.

### *Peer Evaluation Procedure*

For details, see the corresponding section in the Supplementary Information.

## Analysis methods

This exploratory study contains no inferential statistics. Besides the frequency- and proportion-based summary statistics, we calculated only the effect sizes of the results from the original articles and the re-analyses.

### *Cohen's d effect sizes*

Following our preregistration, we converted all results into Cohen's *d*s wherever possible. For a number of cases, we could not achieve this due to missing information in the original studies or reported statistics that cannot be converted into Cohen's *d* (e.g., logistic regression). All the conversions are listed in the R scripts and the data documentation. All the original effect sizes are listed as positive values, and the re-analysis effect sizes are negative only when they showed an opposite effect compared to the original study.

For further information on methods, see Supplementary Information.

---

## Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

Study data and materials are available on the project OSF (https://osf.io/q5h2c/) and GitHub repositories (https://github.com/marton-balazs-kovacs/multi100/). Archived data include the original datasets or a description of how to gain access to them. Our shared materials include all the survey questions and the general communication texts and instructions that we sent to the re-analysts and peer-evaluators. We excluded from our data files the email addresses of the re-analysts, as well as the records of those analysts who did not comply with the instructions and did not submit all the required analyses by the deadline. For further details about our exclusion criteria and procedure, see our Supplementary Information document.

## Code availability

All analysis codes for this project are available at https://github.com/marton-balazs-kovacs/multi100.

## References


1. Silberzahn, R. & Uhlmann, E. L. Crowdsourced research: Many hands make tight work. *Nature* **526**, 189–191 (2015).
2. Patel, C. J., Burford, B. & Ioannidis, J. P. Assessment of vibration of effects due to model specification can demonstrate the instability of observational associations. *J. Clin. Epidemiol.* **68**, 1046–1058 (2015).
3. Botvinik-Nezer, R. *et al.* Variability in the analysis of a single neuroimaging dataset by many teams. *Nature* **582**, 84–88 (2020).
4. Wagenmakers, E.-J., Sarafoglou, A. & Aczel, B. One statistical analysis must not rule them all. *Nature* **605**, 423–425 (2022).
5. Nosek, B. A., Ebersole, C. R., DeHaven, A. C. & Mellor, D. T. The preregistration revolution. *Proc. Natl. Acad. Sci.* **115**, 2600–2606 (2018).
6. Chambers, C. What's next for registered reports? *Nature* **573**, 187–189 (2019).
7. Coles, N. A., Hamlin, J. K., Sullivan, L. L., Parker, T. H. & Altschul, D. Build up big-team science. *Nature* **601**, 505–507 (2022).
8. Brodeur, A., Dreber, A., Hoces de la Guardia, F. & Miguel, E. Replication games: how to make reproducibility research more systematic. *Nature* **621**, 684–686 (2023).
9. Brodeur, A., Mikola, D. & Cook, N. Mass Reproducibility and Replicability: A New Hope. *SSRN Electron. J.* (2024) doi:10.2139/ssrn.4790780.
10. Nuijten, M. B. & Polanin, J. R. "statcheck": Automatically detect statistical reporting inconsistencies to increase reproducibility of meta-analyses. *Res. Synth. Methods* **11**, 574–579 (2020).
11. Steegen, S., Tuerlinckx, F., Gelman, A. & Vanpaemel, W. Increasing transparency through a multiverse analysis. *Perspect. Psychol. Sci.* **11**, 702–712 (2016).
12. Bastiaansen, J. A. *et al.* Time to get personal? The impact of researchers choices on the

selection of treatment targets using the experience sampling methodology. *J. Psychosom. Res.* **137**, 110211 (2020).

13. Hoogeveen, S. *et al.* A many-analysts approach to the relation between religiosity and well-being. *Relig. Brain Behav.* 1–47 (2022) doi:10.1080/2153599X.2022.2070255.
14. Salganik, M. J. *et al.* Measuring the predictability of life outcomes with a scientific mass collaboration. *Proc. Natl. Acad. Sci.* **117**, 8398–8403 (2020).
15. Schweinsberg, M. *et al.* Same data, different conclusions: Radical dispersion in empirical results when independent analysts operationalize and test the same hypothesis. *Organ. Behav. Hum. Decis. Process.* **165**, 228–249 (2021).
16. Silberzahn, R. *et al.* Many Analysts, One Data Set: Making Transparent How Variations in Analytic Choices Affect Results. *Adv. Methods Pract. Psychol. Sci.* **1**, 337–356 (2018).
17. Starns, J. J. *et al.* Assessing theoretical conclusions with blinded inference to investigate a potential inference crisis. *Adv. Methods Pract. Psychol. Sci.* **2**, 335–349 (2019).
18. Veronese, M. *et al.* Reproducibility of findings in modern PET neuroimaging: insight from the NRM2018 grand challenge. *J. Cereb. Blood Flow Metab.* **41**, 2778–2796 (2021).
19. Breznau, N. *et al.* Observing many researchers using the same data and hypothesis reveals a hidden universe of uncertainty. *Proc. Natl. Acad. Sci.* **119**, e2203150119 (2022).
20. Huntington-Klein, N. *et al.* The influence of hidden researcher decisions in applied microeconomics. *Econ. Inq.* **59**, 944–960 (2021).
21. Trübutschek, D. *et al.* EEGManyPipelines: A Large-scale, Grassroots Multi-analyst Study of Electroencephalography Analysis Practices in the Wild. *J. Cogn. Neurosci.* **36**, 217–224 (2024).
22. Menkveld, A. J. *et al.* Nonstandard Errors. *J. Finance* **79**, 2339–2390 (2024).
23. Sarstedt, M. *et al.* Same model, same data, but different outcomes: Evaluating the impact of method choices in structural equation modeling. *J. Prod. Innov. Manag.* **n/a**,.

24. Schilling, K. G. *et al.* Tractography dissection variability: What happens when 42 groups dissect 14 white matter bundles on the same dataset? *NeuroImage* **243**, 118502 (2021).
25. Gelman, A. & Loken, E. The garden of forking paths: Why multiple comparisons can be a problem, even when there is no “fishing expedition” or “p-hacking” and the research hypothesis was posited ahead of time. *Dep. Stat. Columbia Univ.* **348**, (2013).
26. Holzmeister, F. *et al.* Heterogeneity in effect size estimates. *Proc. Natl. Acad. Sci.* **121**, e2403490121 (2024).
27. Holzmeister, F. *et al.* Heterogeneity in effect size estimates: Empirical evidence and practical implications. Preprint at https://doi.org/10.31222/osf.io/583un (2024).
28. Coretta, S. *et al.* Multidimensional Signals and Analytic Flexibility: Estimating Degrees of Freedom in Human-Speech Analyses. *Adv. Methods Pract. Psychol. Sci.* **6**, 25152459231162567 (2023).
29. van Assen, M. A., Stoevenbelt, A. H. & van Aert, R. C. The end justifies all means: questionable conversion of different effect sizes to a common effect size measure. *Relig. Brain Behav.* **13**, 345–347 (2023).
30. Auspurg, K. & Brüderl, J. Has the credibility of the social sciences been credibly destroyed? Reanalyzing the “Many Analysts, One Data Set” Project. *Socius* **7**, 23780231211024421 (2021).
31. Scheel, A. M. Why most psychological research findings are not even wrong. *Infant Child Dev.* **31**, e2295 (2022).
32. Oberauer, K. & Lewandowsky, S. Addressing the theory crisis in psychology. *Psychon. Bull. Rev.* **26**, 1596–1618 (2019).
33. Brodeur, A., Cook, N. & Heyes, A. Methods matter: P-hacking and publication bias in causal analysis in economics. *Am. Econ. Rev.* **110**, 3634–3660 (2020).
34. Simmons, J. P., Nelson, L. D. & Simonsohn, U. False-Positive Psychology: Undisclosed

Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychol. Sci.* **22**, 1359–1366 (2011).

35. Stanovich, K. E. *The Bias That Divides Us: The Science and Politics of Myside Thinking*. xiv, 241 (The MIT Press, Cambridge, MA, US, 2021). doi:10.7551/mitpress/13367.001.0001.
36. Gernsbacher, M. A. Rewarding research transparency. *Trends Cogn. Sci.* **22**, 953–956 (2018).
37. Simonsohn, U., Simmons, J. P. & Nelson, L. D. Specification curve analysis. *Nat. Hum. Behav.* **4**, 1208–1214 (2020).
38. Aczel, B. *et al.* Consensus-based guidance for conducting and reporting multi-analyst studies. *eLife* **10**, e72185 (2021).
39. Kümpel, H. & Hoffmann, S. A formal framework for generalized reporting methods in parametric settings. *ArXiv Prepr. ArXiv221102621* (2022).
40. Sarafoglou, A. *et al.* Subjective Evidence Evaluation Survey For Multi-Analyst Studies. (2024) doi:10.31234/osf.io/mxje3.
41. Del Giudice, M. & Gangestad, S. W. A Traveler's Guide to the Multiverse: Promises, Pitfalls, and a Framework for the Evaluation of Analytic Decisions. *Adv. Methods Pract. Psychol. Sci.* **4**, 2515245920954925 (2021).
42. Liu, Y., Kale, A., Althoff, T. & Heer, J. Boba: Authoring and visualizing multiverse analyses. *IEEE Trans. Vis. Comput. Graph.* **27**, 1753–1763 (2020).
43. Olsson-Collentine, A., van Aert, R. C. M., Bakker, M. & Wicherts, J. Meta-analyzing the multiverse: A peek under the hood of selective reporting. *Psychol. Methods* (2023) doi:10.1037/met0000559.
44. Zwaan, R. A., Etz, A., Lucas, R. E. & Donnellan, M. B. Making replication mainstream. *Behav. Brain Sci.* **41**, (2018).

## Methods References

45. Abatayo, A. L., Achakulvisut, T., Acuna, D., Aczel, B., Balaji, L., Bandrowski, A. E., & Benjamin, D. M. et al. (2025). Empirical, Human, and Machine Assessments of Research Credibility in the Social and Behavioral Sciences. *Unpublished manuscript.*

46. Bartoš, F. *et al.* Introducing synchronous robustness reports. *Nat. Hum. Behav.* (2025) doi:10.1038/s41562-025-02129-1.

## Acknowledgements

We thank Laura Caron who contributed to the re-analyses but did not wish to be listed as an author. We also thank Dan Sun who contributed to the re-analyses. We are grateful to Peter Czingraber, Julianna Harangozo, and Marton Kolozsvari for their assistance with data review and preparation.We are grateful to Meike Latz for helping with our graphic designs. We say thanks to our reviewers whose recommendations and insights greatly improved this work.

## Author information

D.v.R. was supported by 016.Vidi.188.001; M. Adamkovic was supported by APVV-22-0458, PRIMUS/24/SSH/017; A. Sarafoglou was supported by Amsterdam Brain and Cognition (ABC) project grant (ABC PG 22), 2024 Ammodo Science Award; G. Navarrete was supported by Comisión Nacional de Investigación Científica y Tecnológica (CONICYT/FONDECYT Regular 1211373); D. Lacko was supported by Czech Science Foundation (no. GA24-11974S); B.A., B. Szaszi, L.K., T.M.E., and B.A.N. were supported by Defense Advanced Reseach Projects Agency (DARPA): cooperative agreements No. HR00112020015 and N660011924015; C.B.-R. was supported by Deutsche Forschungsgemeinschaft (DFG) through the CRC TRR 190 "RationalityandCompetition."; H.C. was supported by Deutsche Forschungsgemeinschaft, Grant/Award Number: 402170461 [TRR 265]; S.D.-S. was supported by Deutsche Forschungsgemeinschaft, grant no. 440923825; R.A.Z. was supported by Deutscher Akademischer Austauschdienst (DAAD), ref. no. 91803023; M. Rusconi and R.S. were supported by Dipartimento di Eccellenza 2023-2027, DIECO_DDE_2023-2027; C.S. was supported by ESRC grant ES/T015357/1; D.R.S. was supported by ESRC grant ES/Y002482/1; I.M. was supported by European Union - Next Generation EU, Mission 4 Component 1 CUP J53D23007960006 and J53D23017140001; D.W. was supported by European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 665501, Flemish Science Foundation (FWO, FWO.KAN.2019.0023.01), Special Research Fund of Ghent University; T.N. was supported by Excellence Fund of ELTE Eötvös Loránd University and the János Bolyai research fellowship of the Hungarian Academy of Sciences.; Y.L.M. was supported by F.R.S. - FNRS (Fonds National pour la Recherche Scientifique); J.V.A. was supported by FWO.3E0.2021.0085.01; F.R. was supported by German Research Foundation (CRC TR224, Project A04) and Joachim Herz Stiftung; X.S. was supported by German Research Foundation, grant number SCHM 3460/3-1; G. Niso was supported by Grants RYC2021-033763-I and PID2023-150034OA-I00 from the Spanish Government; D. Grigoryev was supported by HSE University Basic Research Program; E.L.U. was supported by INSEAD

R&D committee; E.C. was supported by Insight Grant $435-2024-0336, Social Sciences and Humanities Research Council (SSHRC) of Canada; S. Alzahawi was supported by Institute for Museum and Library Services #LG-250130-OLS-21; R.M.R. was supported by John Templeton Foundation (grant ID: 62631); R.D. was supported by Keynes Fund [JHVH], CRASSH Grant (climaTRACES lab) and Cambridge Humanities Research Grants; A.D. was supported by Knut and Alice Wallenberg Foundation, Marianne and Marcus Wallenberg Foundation, Jan Wallander and Tom Hedelius Foundation; M.M.E. was supported by Leverhulme Trust Early Career Fellow, RM56G0344; H.D. was supported by MOE (Ministry of Education in China) Project of Humanities and Social Sciences, 24YJA190001; C.M.G. was supported by MSCA-PF; J. Hogeveen was supported by NIAAA (R01AA030283); NSF (BCS 2237795); I.R.J. was supported by NIHR Manchester Biomedical Research Centre (NIHR203308); P. McKee was supported by NSF GRFP DGE 2139754; E.M.M. was supported by NWO (Nederlandse Organisatie voor Wetenschappelijk Onderzoek) Domain Social Sciences and Humanities (SSH) Starter Grant.; P. Papale was supported by NWO Veni VI.Veni.222.217; N.N.N.v.D. was supported by NWO Vici grant no.181.029; A. Gasiorowska was supported by National Science Center, 2021/41/B/HS6/03104; R.L.-B. was supported by Next Generation EU; S. Massoni was supported by Program FUTURE LEADER of Lorraine Université d'Excellence within the program Investissements Avenir (ANR-15-IDEX-04-LUE) operated by the French National Research Agency; F.D. and J. Lukavský were supported by RVO 68081740; J. Pfänder was supported by SCALUP ANR grant ANR-21-CE28-0016-01; M. Schaerer was supported by Singapore Ministry of Education Tier 1 Grant (MSS22B011) and Tier 2 Grant (MM22B03); I. Ropovik was supported by Slovak Research and Development Agency under the Contract no. APVV-23-0421; NPO Systemic Risk Institute (LX22NPO510); S.T. Schwartz was supported by Stanford Center for Mind, Brain, Computation and Technology; M.M.A. was supported by Stone Program in Wealth Distribution, Inequality, and Social Policy at Harvard University; P. Pärnamets was supported by Swedish Research Council 2020-02584; A.N.M. was supported by Swiss National Science Foundation grant PZ00P1_201956; V. Chirkov was supported by The Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2002/1 "Science of Intelligence" – project number 390523135; A. Boldt was supported by The Wellcome Trust 206480/Z/17/Z; B.P. was supported by The article was produced with financial support of the European Union under the REFRESH – Research Excellence For REgion Sustainability and High-tech Industries project number CZ.10.03.01/00/22_003/0000048 via the Operational Programme Just Transition.; F.M. was supported by UKRI BBSRC (BB/X008428/1) and NIHR Leeds Biomedical Research Centre (NIHR203331); G. Corlazzoli was supported by Université Libre de Bruxelles (Mini-ARC); Y.A.W. was supported by a Social Sciences and Humanities Research Council (SSHRC) Insight Development Grant (430-2022-00087); P. Präg was supported by a grant of the French National Research Agency ANR, 'Investissements d'Avenir' (LabEx Ecodec/ANR-11-LABX-0047).; S.R.S. was supported by an Interdisciplinary Exploratory Synergy grant from the Novo Nordisk foundation (NNF: NNF20OC0064869); J.F.G.M. was supported by an NSF SBE Postdoctoral Research Fellowship (award number 2104629); Q.S. was supported by grant 72371165, 71971199 and 71942004 from the National Natural Science Foundation of China.; M.R.S. was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - GRK 2277 - project number 310365261; D.T. was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 101023805; R.R.-C. was supported by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Actions—Global Postdoctoral Fellowship (project Nº 101149355).; M. Kowal was supported by the Foundation for Polish Science (FNP) START scholarship; F.C. was supported by the National Natural Science Foundation of China [Grants 72173113 and 72322009].; A.M.A. was supported by the OTKA FK 146604 research grant; A. Antonietti was supported by the Project “EBRAINS-Italy (European Brain ReseArch INfrastructureS-Italy)” granted by European Union – NextGenerationEU adopted by the Italian Ministry of University and Research, CUP B51E22000150006.; A. Plachti was supported by the Region Hovedstadens Forskningsfond til Sundhedsforskning (grant agreement number A7118); H.F.S. was supported by the Research Council of Norway through its Centres of Excellence funding scheme (#262700); P. Kačmár was supported by the Slovak Research and Development Agency under the Contract no. APVV-22-0458 and APVV-23-0548; A. Arenas was supported by the project Proyectos de Generación de Conocimiento de la Agencia Estatal de Investigación PID2023-150898OB-I00; C. Koba was supported by the project of the Minister of Science and Higher Education "Support for the activity of Centers of Excellence established in Poland under Horizon 2020" on the basis of the contract number MEiN/2023/DIR/379, and the European Union’s Horizon 2020 research and innovation programme under grant agreement No 857533, and by Sano project carried out within the International Research Agendas programme of the Foundation for Polish Science, co-financed by the European Union under the European Regional Development Fund.

**Contributions**

**Conceptualization:** B.A., B. Szaszi, F. Holzmeister, S. Hoffmann, G. Nilsonne, L.K., Z.A.T., E.-J.W., T.M.E., and B.A.N.

**Data curation:** H.T.C., M. Kovacs, M.A.V., and F. Venczel

**Formal analysis:** B.A., B. Szaszi, H.T.C., M. Kovacs, F. Holzmeister, D.v.R., H.S.-K., S. Hoffmann, G. Nilsonne, Y.A., C.L.A., O.A.A., M. Adamkovic, T. Adamovich, K.A., A.M.A., A.S.A.-B., A.H.A.-H., C.J.A., P.J.A., T. Alsalti, M. Altman, S. Alzahawi, E. Ambrosini, S. Anafinova, R.A., M. Angerer, A.A.-B., A. Antonietti, J.A., A. Arenas, M.M.A., F.A., M. Bachl, B. Bago, Š.B., B.J.B., E. Balayan, C.L.B., B. Banai, K.B., F.B., E. Baskin, N. Bault, C.W.B., Q.H.B., M. Behnke, T.B., S.B., A. Bernard, U.B., P.A.B., A. Boldt, C.B.-R., A. Bouyamourn, O.B., L.B., J. Breuer, R.B., H.B., E. Buchanan, J. Buckenmaier, J. Buckley, J. Buczny, M. Burghart, B.H.B., N. Byrd, V. Cafarelli, P.C., T.C., K.C., A.M.C., G. Cepaluni, E.C., J.J.C., C.-c.C., X.C., S.S.C., F.C., H.C., V. Chirkov, D.C., B.C., S.G.C., C. Cohen, J. Collins, S.W.C., G. Corlazzoli, J. Cummins, C. Czymara, J.D.h., A.D.R., A.M.B.D., C.P.D., M.V.D., F.D.K., J.R.d.L., T.R.d.V., R.D., F.D., E.E.D., M.D., V.D., S.D.-S., S.D., L.D., J.D., A.R.D., H.D., J.E.E., A.E., E. Efendic, J.E., M.M.E., M.E., E. Estrada, L.E., T.R.E., A.F., E.M.F., L.F., F.M.F., J.L. Fiechter, M. Fišar, P.E.F.-K., M. Folwarczny, J.L. Fossum, V.R.F., R.F., D.F., J.F., A.C.F., J.D.G., L.C.G., C.M.G., B.G., S.M.G., A. Gasiorowska, B.G.P., R.G., A. Geminiani, D. Geraldes, C. Giani, E.G., V.G., T.G., A. Godefroidt, B.G.-B., A. Goreis, L.G.-V., M.G., D. Grigoryev, S.G., D.J.G., J.F.G.M., C. Guichet, L.G., H.H., A.C.H., S. Hafenbrädl, C. Häffner, F. Hagemeister, M. Haigh, N. Hajdu, N. Hajimoladarvish, J.D.H., M. Hamjediers, R.M.H., M. Harma, N.R.H., Á.D.H., R.H.H., A.H., Ø.H., D.H., T. Heyman, J. Hicks, J. Hogeveen, J. Höpler, S.D.H., C. Huber, C. Hughes, T. Hummler, K.H., M.I., T.I., O.I., K.I., I.R.J., A. Jahn, M. Jain, A. Jakubow, D.J., J.J., M. Jekel, F.J., W.J.-L., R.J., A. Jones, S.J., P. Kačmár, C. Kaiser, Y. Kalaycı, J. Kantorowicz, A. Karabulut, J. Karch, H.K.-R., J.A.K., A. Kažemekaitytė, A. Kazlou, Z.K., J. Kim, B.K.-D., K.N.K., J.W.K., C. Koba, M. Kołczyńska, P. Kolias, M.K.G., M. Korbmacher, Ž.K., M. Kowal, A. Kretzschmar, V.K., A.-M.K., M. Kubsch, Y. Kunisato, D. Lacko, J.R.L., M. Lange, H.L., D. Lee, S. Lee, E.P.L.J., D. Lempert, A. Leo, E.L., J.M.L., P. Li, J. Lin, L.L., D. Lisovoj, M. Liu, S. Liu, T.L., S.L.I., P. Lodder, R.L.-B., R.L.-N., K.L., N.M.L., A. Lovakov, J.G.L., J. Ludwig, F.L., J. Lukavský, C.L., X.L., E.M., M. Máčel, M.L.M., C.R.M., A. Mädebach, J.M.-K., D.J.M., I.M., T.M., M.M. Marini, D.M.F., M. Martínez, M. Martinoli, A. Masiliunas, S. Massoni, K.C.M., S. Mayer, D. Mayer, M. Mayer, E.M.M., I.M.M., A.L.M., M.M. McIntyre, P. McKee, A.N.M., P.F.M., H.M., C. Merkle, R.M., M.P.M., P. Michaelsen, G.M., W.M., P. Millroth, K.G.M., M. Misiak, Y.L.M., D. Moreau, C. Moreh, C. Morvinski, F.M., T.N., C.N., E.N., G. Navarrete, S.N., R.N., M.N., E.N.-S., Y.A.N., G. Niso, B.N., M.O., K. Ong, A.I.O., C.O., K. Otten, S. Pandey, M. Pantazi, P. Papale, P. Pärnamets, S. Pauer, Y.G.P., S. Pawel, J.E.P., H.K.P., A. Peez, F.P., B.D.P., B.P., J. Petter, J. Pfänder, G.P., J. Phillips, M.T.P., A. Pirrone, I.L.P., A. Plachti, I.S.P., M. Ploner, M.M.H.P., S. Porcher, P. Präg, A.A.Y.P., J. Pugel, R.P., M. Püski, S. Radkani, L. Raes, I. Rafaï, K. Raiber, S. Rathje, M. Reshetnikov, C.J.R., J.P.R., K. Rigaud, C.R., S. Rivera, O.R., R.R.-C., I. Ropovik, L. Röseler, R.M.R., A.R., F.F.R., F.R., M. Rusconi, I. Russo, A.H.J.S., J. Salamon, M. Samahita, A. Sanaei, A. Sangchooli, M. Scandola, H.S., M. Schaerer, E. Schares, H.T.S., X.S., K.S., M.R.S., J.M.S., A.-L.S., B. Schuetze, D.H.S., L. Schulze, S.T. Schwartz, N.S., B. Scoggins, Y.S., R.S., S.T. Shaw, J. Shaw, Q.S., C.S., M. Sladekova, A. Somo, A. Sondhi, B. Sonmez, L. Spantig, M. Speekenbrink, A. Stamos, L. Stasielowicz, L.C.S., S.R.S., A.H.S., C.N.H.S., J.W.S., H.F.S., J. Sundquist, V.S., S.D.S., P.S., R.D.S.-C., E. Szumowska, A. Tacconelli, E. Talbert, J.P.T., J.N.T., M.T., E. Toffalini, A. Tomašević, S.T., L. Torkkeli, L. Tozzi, J.T., A. Trinidad, D.T., K.T., M.U., K.U., J.V.A., K.v.L., R.v.V., L.A.V., M.V., P.V., A.V., E.V., F. Votta, A. Waldendorf, M.J.W., M.B.W., H.W., K.W., I.W., Y.A.W., M. Weinmann, M. Weiss, C.W., A. Wichman, B.J.W., D.W., T.K.A.W., M. Woźniak, J.D.W., W.Y., J.N.W., T.Y., S.K.Y., K.S.L.Y., M.Z., R.A.Z., X.Z., Z.Z., S.Z., C. Ziller, D.Z., C. Zogmaister, and R.i.Z.

**Funding acquisition:** B.A., B. Szaszi, T.M.E., and B.A.N.

**Investigation:** L.A.V.

**Methodology:** B.A., B. Szaszi, H.T.C., M. Kovacs, F. Holzmeister, D.v.R., H.S.-K., S. Hoffmann, G. Nilsonne, L.K., Z.A.T., C.J.A., J.A.B., R.B.-N., A.D., M.A.G., M.H.K., R.A.P., A. Sarafoglou, T.S., D.R.S., E.L.U., M.A.L.M.v.A., N.N.N.v.D., E.-J.W., T.M.E., and B.A.N.

**Project administration:** B.A., B. Szaszi, H.T.C., M. Kovacs, L.K., Z.A.T., N.F., and T.M.E.

**Resources:** B.A., B. Szaszi, H.T.C., F. Holzmeister, L.K., L.D., N.F., T.M.E., and B.A.N.

**Software:** M. Kovacs

**Supervision:** B.A., B. Szaszi, H.T.C., F. Holzmeister, D.v.R., S. Hoffmann, G. Nilsonne, C.J.A., J.A.B., R.B.-N., A.D., M.A.G., M.H.K., A.N., R.A.P., A. Sarafoglou, T.S., D.R.S., E.L.U., M.A.L.M.v.A., N.N.N.v.D., E.-J.W., T.M.E., and B.A.N.

**Validation:** X.C., D.D., and S.J.

**Visualization:** B.A., B. Szaszi, H.T.C., and M. Kovacs

**Writing - original draft:** B.A., B. Szaszi, H.T.C., M. Kovacs, and L.K.

**Writing - review & editing:** B.A., B. Szaszi, H.T.C., M. Kovacs, F. Holzmeister, D.v.R., H.S.-K., S. Hoffmann, G. Nilsonne, L.K., Z.A.T., Y.A., C.L.A., O.A.A., E. Acem, M. Adamkovic, T. Adamovich, K.A., L.A., A.M.A., A.S.A.-B., A.H.A.-H., C.J.A., P.J.A., T. Alsalti, M. Altman, S. Alzahawi, E. Ambrosini, S. Anafinova, R.A., M. Angerer, A.A.-B., A. Antonietti, J.A., A. Arenas, M.M.A., F.A., M. Bachl, B. Bago, Š.B., B.J.B., E. Balayan, C.L.B., B. Banai, K.B., F.B., E. Baskin, J.A.B., N. Bault, C.W.B., Q.H.B., M. Behnke, T.B., S.B., A. Bernard, U.B., P.A.B., A. Boldt, C.B.-R., R.B.-N., A. Bouyamourn, O.B., L.B., J. Breuer, R.B., H.B., E. Buchanan, J. Buckenmaier, J. Buckley, J. Buczny, M. Burghart, B.H.B., N. Byrd, V. Cafarelli, P.C., T.C., K.C., A.M.C., G. Cepaluni, E.C., J.J.C., C.-c.C., X.C., S.S.C., F.C., H.C., V. Chirkov, D.C., B.C., S.G.C., C. Cohen, J. Collins, S.W.C., G. Corlazzoli, J. Cummins, C. Czymara, J.D.h., A.D.R., A.M.B.D., C.P.D., M.V.D., F.D.K., J.R.d.L., T.R.d.V., R.D., F.D., E.E.D., M.D., D.D., V.D., S.D.-S., S.D., L.D., J.D., A.R.D., A.D., H.D., J.E.E., A.E., E. Efendic, J.E., M.M.E., M.E., E. Estrada, L.E., T.R.E., A.F., E.M.F., L.F., F.M.F., J.L. Fiechter, M. Fišar, P.E.F.-K., M. Folwarczny, J.L. Fossum, V.R.F., R.F., D.F., J.F., A.C.F., J.D.G., L.C.G., C.M.G., B.G., S.M.G., A. Gasiorowska, B.G.P., R.G., A. Geminiani, D. Geraldes, M.A.G., C. Giani, E.G., V.G., T.G., A. Godefroidt, B.G.-B., A. Goreis, L.G.-V., M.G., D. Grigoryev, S.G., D.J.G., J.F.G.M., C. Guichet, L.G., H.H., A.C.H., S. Hafenbrädl, C. Häffner, F. Hagemeister, M. Haigh, N. Hajdu, N. Hajimoladarvish, J.D.H., M. Hamjediers, R.M.H., M. Harma, N.R.H., Á.D.H., R.H.H., A.H., Ø.H., D.H., T. Heyman, J. Hicks, J. Hogeveen, J. Höpler, S.D.H., C. Huber, C. Hughes, T. Hummler, K.H., M.I., T.I., O.I., K.I., I.R.J., A. Jahn, M. Jain, A. Jakubow, D.J., J.J., M. Jekel, F.J., W.J.-L., R.J., A. Jones, S.J., P. Kačmár, C. Kaiser, Y. Kalaycı, J. Kantorowicz, A. Karabulut, J. Karch, H.K.-R., J.A.K., A. Kažemekaitytė, A. Kazlou, Z.K., J. Kim, M.H.K., B.K.-D., K.N.K., J.W.K., C. Koba, M. Kołczyńska, P. Kolias, M.K.G., M. Korbmacher, Ž.K., M. Kowal, A. Kretzschmar, V.K., A.-M.K., M. Kubsch, Y. Kunisato, D. Lacko, J.R.L., M. Lange, H.L., D. Lee, S. Lee, E.P.L.J., D. Lempert, A. Leo, E.L., J.M.L., P. Li, J. Lin, L.L., D. Lisovoj, M. Liu, S. Liu, T.L., S.L.I., P. Lodder, R.L.-B., R.L.-N., K.L., N.M.L., A. Lovakov, J.G.L., J. Ludwig, F.L., J. Lukavský, C.L., X.L., E.M., M. Máčel, M.L.M., C.R.M., A. Mädebach, J.M.-K., D.J.M., I.M., T.M., M.M. Marini, D.M.F., M. Martínez, M. Martinoli, A. Masiliunas, S. Massoni, K.C.M., S. Mayer, D. Mayer, M. Mayer, E.M.M., I.M.M., A.L.M., M.M. McIntyre, P. McKee, A.N.M., P.F.M., H.M., C. Merkle, R.M., M.P.M., P. Michaelsen, G.M., W.M., P. Millroth, K.G.M., M. Misiak, Y.L.M., D. Moreau, C. Moreh, C. Morvinski, F.M., T.N., C.N., E.N., G. Navarrete, S.N., A.N., R.N., M.N., E.N.-S., Y.A.N., G. Niso, B.N., M.O., K. Ong, A.I.O., C.O., K. Otten, S. Pandey, M. Pantazi, P. Papale, P. Pärnamets, S. Pauer, Y.G.P., S. Pawel, J.E.P., H.K.P., A. Peez, F.P., B.D.P., B.P., J. Petter, J. Pfänder, G.P., J. Phillips, M.T.P., A. Pirrone, I.L.P., A. Plachti, I.S.P., M. Ploner, R.A.P., M.M.H.P., S. Porcher, P. Präg, A.A.Y.P., J. Pugel, R.P., M. Püski, S. Radkani, L. Raes, I. Rafaï, K. Raiber, S. Rathje, M. Reshetnikov, C.J.R., J.P.R., K. Rigaud, C.R., S. Rivera, O.R., R.R.-C., I. Ropovik, L. Röseler, R.M.R., A.R., F.F.R., F.R., M. Rusconi, I. Russo, A.H.J.S., J. Salamon, M. Samahita, A. Sanaei, A. Sangchooli, A. Sarafoglou, M. Scandola, H.S., M. Schaerer, E. Schares, H.T.S., X.S., K.S., T.S., M.R.S., J.M.S., A.-L.S., B. Schuetze, D.H.S., L. Schulze, S.T. Schwartz, N.S., B. Scoggins, Y.S., R.S., D.R.S., S.T. Shaw, J. Shaw, Q.S., C.S., M. Sladekova, A. Somo, A. Sondhi, B. Sonmez, L.

Spantig, M. Speekenbrink, A. Stamos, L. Stasielowicz, L.C.S., S.R.S., A.H.S., C.N.H.S., J.W.S., H.F.S., J. Sundquist, V.S., S.D.S., P.S., R.D.S.-C., E. Szumowska, A. Tacconelli, E. Talbert, J.P.T., J.N.T., M.T., E. Toffalini, A. Tomašević, S.T., L. Torkkeli, L. Tozzi, J.T., A. Trinidad, D.T., K.T., M.U., E.L.U., K.U., J.V.A., M.A.L.M.v.A., N.N.N.v.D., K.v.L., R.v.V., M.A.V., L.A.V., F. Venczel, M.V., P.V., A.V., E.V., F. Votta, E.-J.W., A. Waldendorf, M.J.W., M.B.W., H.W., K.W., I.W., Y.A.W., M. Weinmann, M. Weiss, C.W., A. Wichman, B.J.W., D.W., T.K.A.W., M. Woźniak, J.D.W., W.Y., J.N.W., T.Y., S.K.Y., K.S.L.Y., M.Z., R.A.Z., X.Z., Z.Z., S.Z., C. Ziller, D.Z., C. Zogmaister, R.i.Z., N.F., T.M.E., and B.A.N.

**Corresponding Authors**

Correspondence and requests for materials should be addressed to Balazs Aczel and Barnabas Szaszi; E-mail: balazs.aczel@gmail.com and szaszi.barnabas@gmail.com.

## Ethics declarations

*Competing interests*

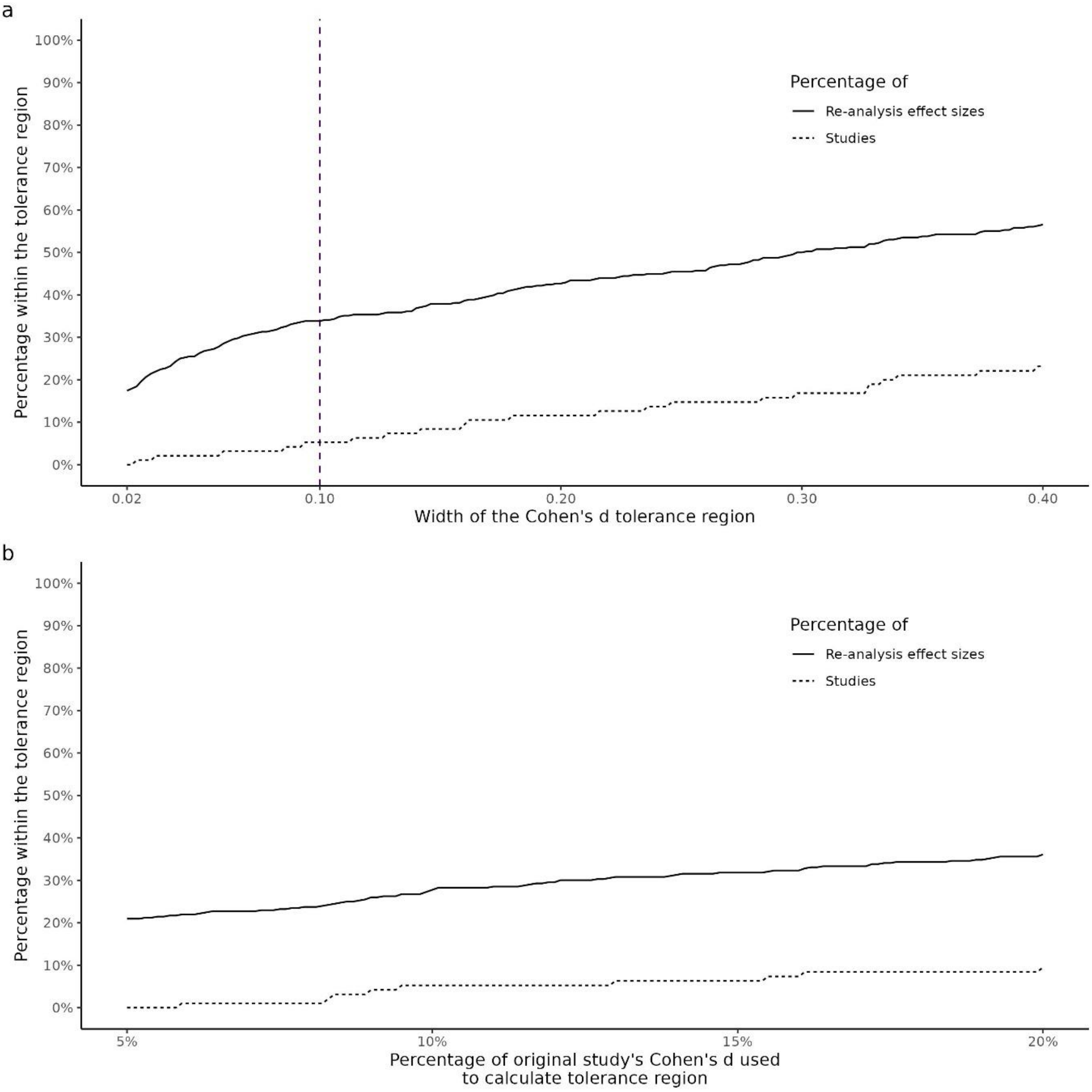
The authors declare no financial or non-financial competing interests.

# SUPPLEMENTARY INFORMATION

| Content | Location |
|---|---|
| Project preregistrations | https://osf.io/4ev9t, https://osf.io/y4evp |
| Deviations from preregistration | https://osf.io/wfdxp |
| Supplementary information | https://osf.io/4ntua |
| Project OSF Repository | https://osf.io/q5h2c/ |
| Project GitHub Repository | https://github.com/marton-balazs-kovacs/multi100 |

## Extended data figures

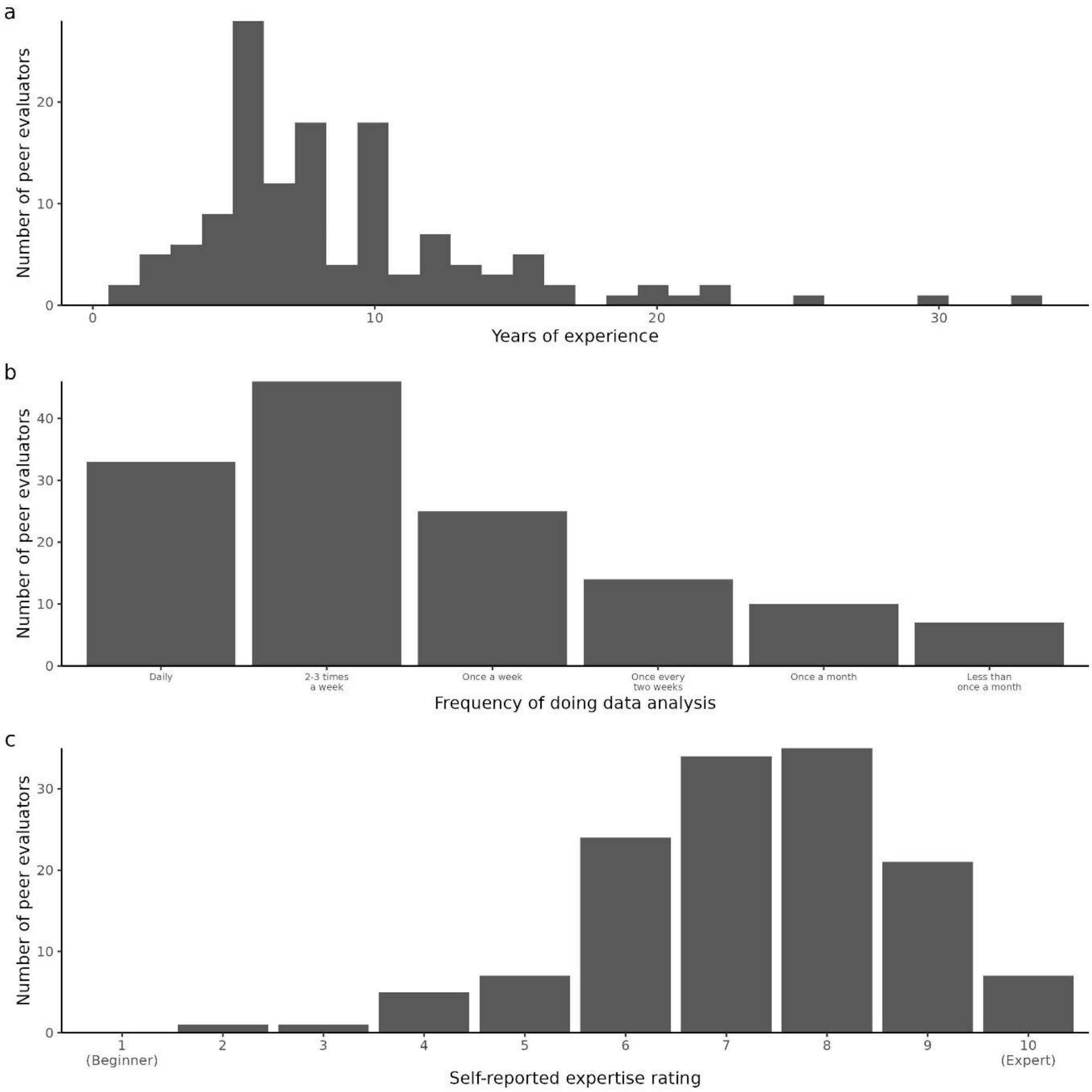
**Extended Data Fig. 1**
*Robustness of the statistical results.*




*Note.*
**a,** Robustness of the statistical results with different widths (+/-[0.01-0.20] Cohen's *d*) of the tolerance region. **b,** Robustness of the statistical results with different percentages (5-20%) of Cohen's *d* as a tolerance region. Calculations on the study and re-analysis levels are shown in different lines.

**Extended Data Fig. 2**

*Descriptive statistics of the peer evaluators.*

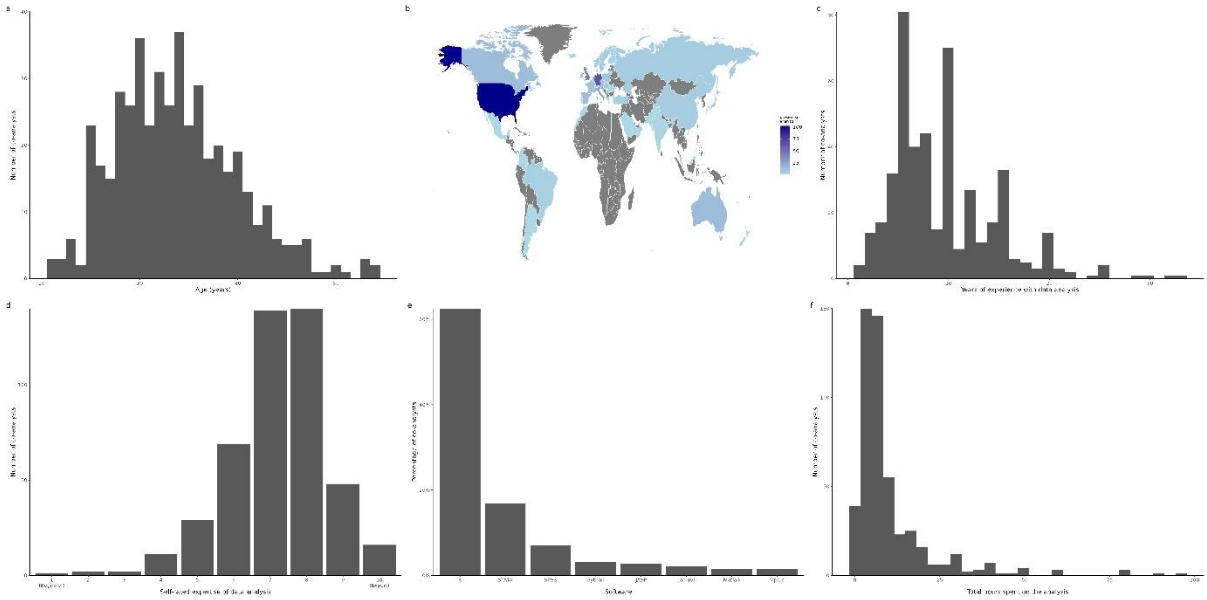



*Note.*

**a,** The peer evaluators' years of experience with data analysis. When a peer evaluator submitted more than one evaluation and a year passed between the responses, we kept only their first response. **b,** The regularity with which peer evaluators perform data analysis. **c,** The peer evaluators' self-rated level of expertise in data analysis. When a peer evaluator submitted more than one re-analysis, we kept only their first response.

**Extended Data Fig. 3**

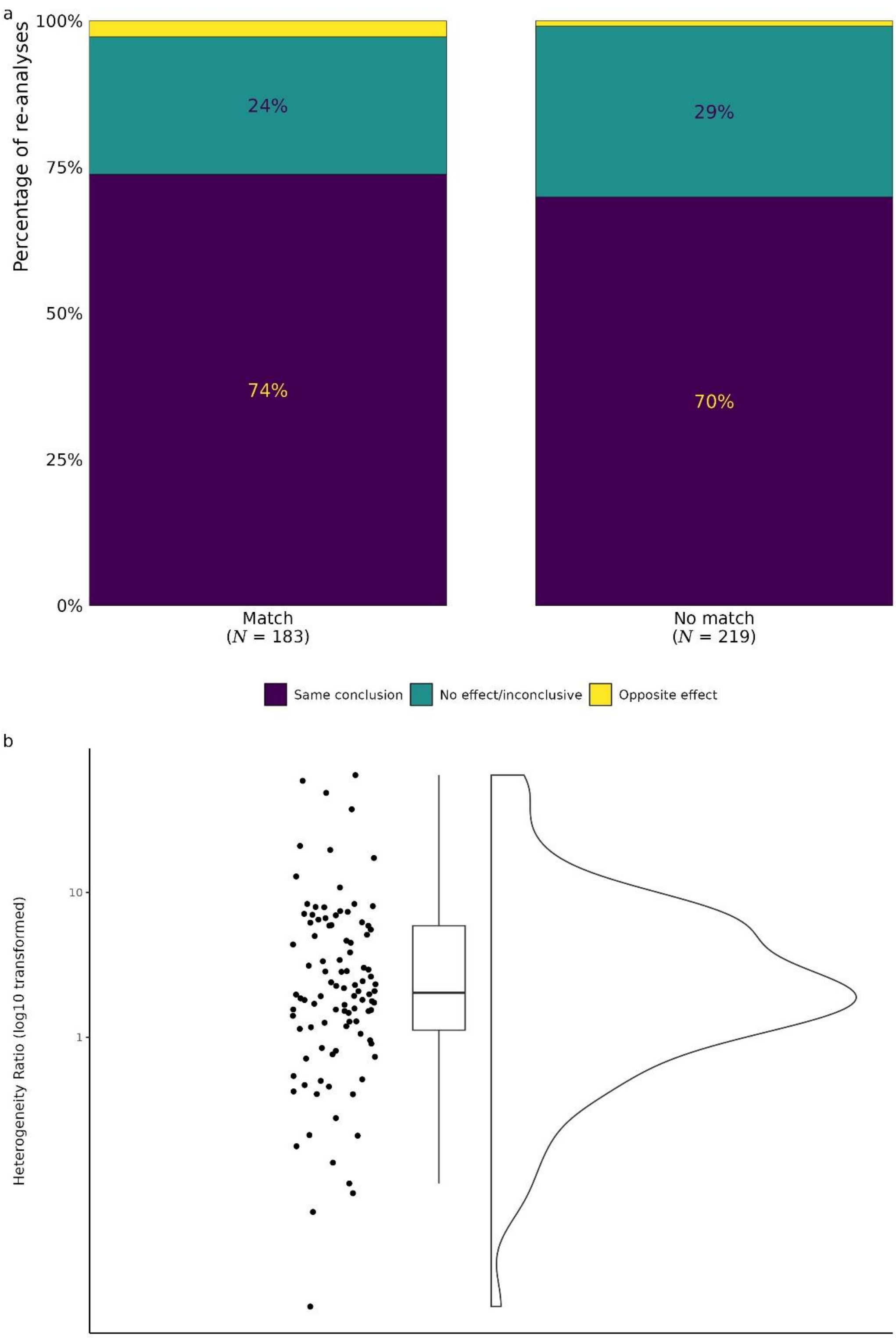
*Descriptive statistics of the analysts and the analyses.*



*Note.*

**a,** The distribution of the analysts' age. When an analyst submitted additional re-analyses with a higher reported age, we kept only their age at the time of their first submission. Moreover, one analyst is not represented in the figure because they did not disclose their age. **b,** The analysts' country of residence. When an analyst submitted more than one re-analysis, and they moved between the submissions, we only kept their first response. **c,** The analysts' years of experience with data analysis. We only kept their first response when an analyst submitted additional re-analyses with a higher reported age. **d,** The analysts' self-rated level of expertise in data analysis. When an analyst submitted more than one re-analysis, we only kept their first response. **e,** The software the analysts used for their re-analysis tasks. In case an analyst completed multiple re-analyses or reported using multiple software applications, we kept all their responses for this figure. The figure displays only software applications used by more than 1% of the analysts. **f,** The reported total hours the analyst spent on Task 1 and Task 2 together. In case an analyst completed multiple re-analyses, we kept all their responses for this figure. One response was excluded due to being an outlier (999 h), which we assumed was an error.

**Extended Data Fig. 4**
*Additional statistical results requested by the reviewers.*

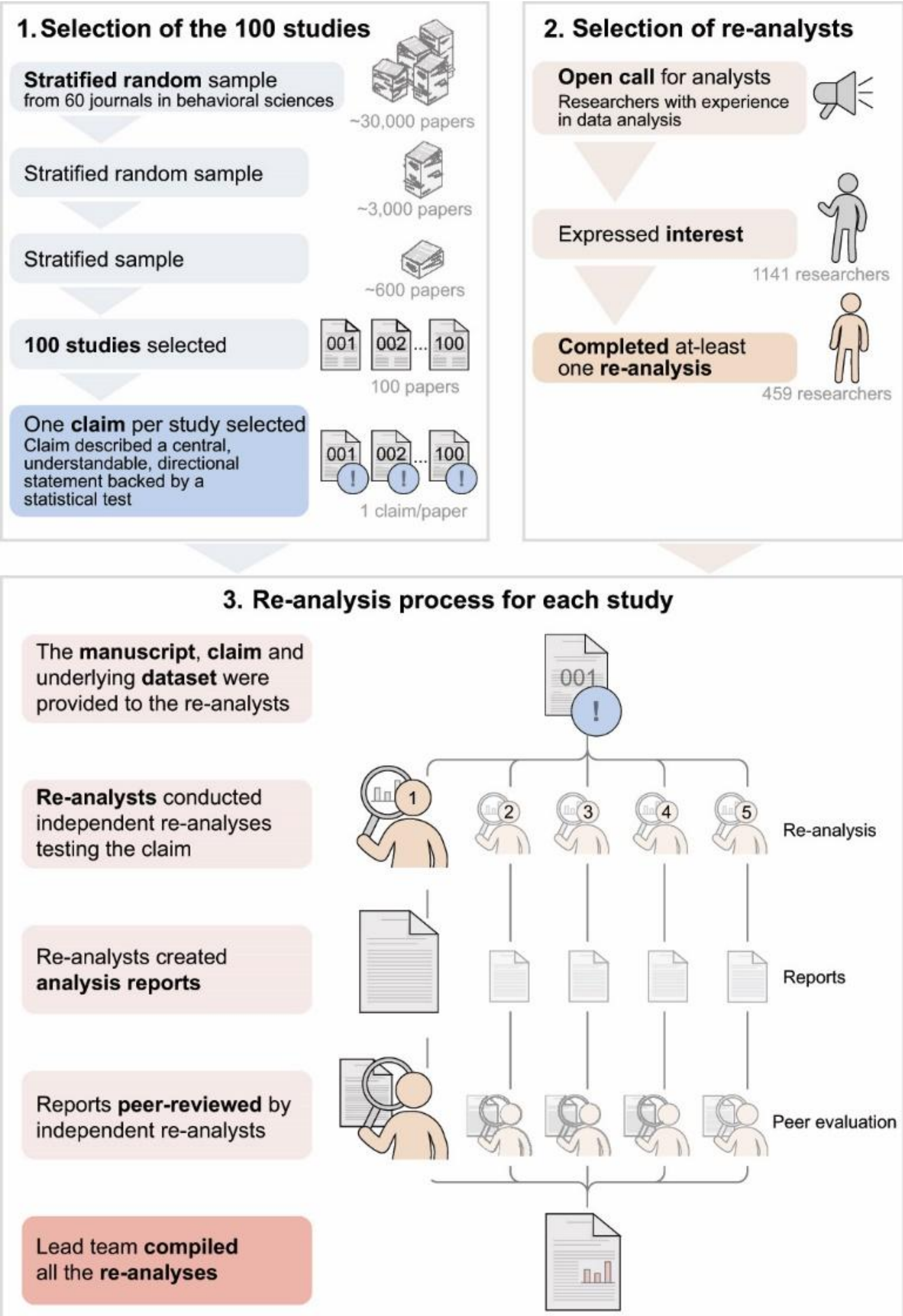



*Note.*
**a,** Proportion of same effect, no effect/inconclusive results, and conclusions in the opposite direction of the original studies, by matches and nonmatches between the discipline of the re-analyst and the

original study. **b,** The distribution of the heterogeneity ratios calculated between the effect size variability over the re-analyses and the sampling variability of the original study effect-size estimates.

**Extended Data Fig. 5**

*Overview of the project procedures.*




*Note.*

The figure depicts the procedural workflow of the selection of the studies (1); the selection of the re-analysts (2); and the re-analysis process for each study (3).