

# Mainstreaming Race Science

Daniel J. Hicks,<sup>1\*</sup> and Emilio Lobato<sup>1</sup>

<sup>1</sup>University of California, Merced

\*Corresponding author: Daniel J. Hicks, [dhicks4@ucmerced.edu](mailto:dhicks4@ucmerced.edu).

**Background:** While anthropologists and human geneticists have long rejected the biological race concept, both social movements and prominent researchers have purported to justify racial inequality and colonialism by appealing to the epistemic authority of science.

**Methods:** Bibliometric and text mining methods were used to trace the appearance of race science discourse in both mainstream scientific journals and the parascholarly journal *Mankind Quarterly* from 1960 to 2010.

**Results:** Our analysis found two distinguishable lines of race science discourse. One is strongly associated with *Mankind Quarterly*, while the other — emphasizing race and intelligence — is strongly associated with mainstream psychology journals, especially *Intelligence* and *Personality and Individual Differences*.

**Conclusion:** Psychology, and especially intelligence research, has played a significant role in mainstreaming race science.

**One sentence summary:** Bibliometric and text mining analysis indicate that race science discourse appeared in mainstream psychology journals throughout the second half of the twentieth century.

## Introduction

By the end of the Second World War, a combination of technical developments in anthropology and genetics, along with revelations of Nazi atrocities, had led most members of the scientific

community to reject attempts to scientifically justify racial inequalities and colonialism (1). But eugenic thinking and scientific racism persisted throughout the second half of the twentieth century, and recent work has found that white supremacists continue to attempt to use scientific research to justify racial violence (2, 3).

The aim of the study reported here was to trace race science in the scholarly and parascholarly literature, examining the ways that fringe ideas can present themselves as legitimate scholarly inquiry.

We first propose a distinction between scientific racism, race science, and race science discourse. *Scientific racism* refers to the social practice of purporting to justify racial inequality and colonialism by appealing to the epistemic authority of science. *Race science* refers to scientific and/or pseudoscientific research, including research products (journal articles, etc.), that can be utilized for scientific racism. *Race science discourse* refers to a broader category, any treatment of race science as a legitimate area of scientific research. This includes methodological and empirical critiques of race science. For example, the Flynn effect (the finding that IQ scores have increased over time) poses a serious empirical challenge to hereditarian claims of racial differences in intelligence. But technical debates over the extent and causes of the Flynn effect likely create the impression among the general public of ordinary scientific disagreement. In this way technical critiques of race science based on the Flynn effect could potentially legitimize race science. So empirical research supporting the Flynn effect is usually not itself race science, but can still be race science discourse.

These distinction allow us to bracket the intentions and mindset of the researchers involved in any particular piece of research or debate. For the purposes of the current study, it is not important whether, for instance, Arthur Jensen is or is not correctly labelled a scientific racist (4). But it is important that much of Jensen's research was useful and indeed utilized to purportedly justify racial inequality.

## Scientific racism, eugenics, and *Brown*

Histories of scientific racism in the twentieth century have often emphasized the Pioneer Fund and the parascholarly journal *Mankind Quarterly* (MQ) (1, 5–12). Pioneer was formed in 1937, during the waning years of the eugenics movement in North America (1), with the aim of “support[ing] academic research and the ’[sic]dissemination of information, into the ‘problem of heredity and eugenics’ and ‘the problems of race betterment’ ” (5, p.21, quoting Laughlin). MQ was founded in 1960 by biologist R. Ruggles Gates (1882-1962), psychologist Henry Garrett (1894-1973), and non-academic anthropologist G. Robert Gayre (self-styled as “Gayre of Gayre and Nigg”; 1907-1996). Gates’ professional status had risen and fallen with eugenics and the explicit scientific racism of the 1920s, and by the end of the Second World War he was thoroughly marginalized (6). In contrast, Garrett had been president of the American Psychological Association in 1946 and chair of Psychology at Columbia from 1941 to 1955 (6).

In the landmark case *Brown v Board of Education of Topeka* (1954), the US Supreme Court banned *de jure* educational segregation. The Court’s decision relied on expert testimony from psychologists and education researchers; but the segregationists also put forward their own experts, including Henry Garrett (6, 8, 13).

As part of the segregationist reaction to *Brown*, MQ was created as what contemporary historians, philosophers, and sociologists of science call an “echo chamber” (14), a favorable venue for race scientists to publish their views, on the grounds that an “equalitarian dogma” created a censorious “taboo” against their research in mainstream publications (7, 15). Echo chambers are a key part of the “tobacco strategy,” used by numerous regulated industries — most infamously tobacco, but also fossil fuels and chemical manufacturing, among others — to “manufacture doubt” and delay regulation (14, 16–18).

In light of the origins of Pioneer and MQ, the significant attention paid to the journal

by historians of scientific racism, and contemporary research on the “tobacco strategy,” we hypothesized that bibliometric and text mining analyses would confirm that MQ functioned as an echo chamber for Pioneer-funded scientists. This hypothesis was not supported. Instead, mainstream psychology journals served as the echo chambers for race science.



## *Mankind Quarterly* and Pioneer-funded researchers

We identified 16 researchers who had received funding from Pioneer; 14 of these researchers had profiles in the Web of Science (WoS) author search, allowing us identify 13 WoS-indexed journals that had published 6 or more of these authors. See Table 1 and Figure 1.

Table 1: Pioneer-funded researchers. Either identified in (19) or named on an archive copy of Pioneer’s website, along with birth and death dates from Wikipedia, attributed discipline, and WoS author search result counts. Brunetto Chiarelli does not have a Wikipedia page.

Thomas J. Bouchard, Jr. (1937-)	psychology	184
Brunetto Chiarelli (?-?)	anthropology	91
Hans Eysenck (1916-1997)	psychology	661
Robert A. Gordon (1932-)	sociology	0
Linda Gottfredson (1947-)	psychology	71
Garrett Hardin (1915-2003)	ecology	75
Joseph M. Horn (1940-)	psychology	68
Lloyd Humphreys (1913-2003)	psychology	0
Arthur Jensen (1923-2003)	psychology	235
Michael Levin (1943-)	philosophy	96
Richard Lynn (1930-)	psychology	288
R. Travis Osborne (1913-2013)	psychology	59
J. Phillippe Rushton (1943-2012)	psychology	277
Audrey M. Shuey (1900-1977)	psychology	10
Philip A. Vernon (1950-)	psychology	227
Daniel Vining, Jr. (1944-)	demography	33

Figure 1 shows that, while MQ is among the “Pioneer-publishing” journals, a number

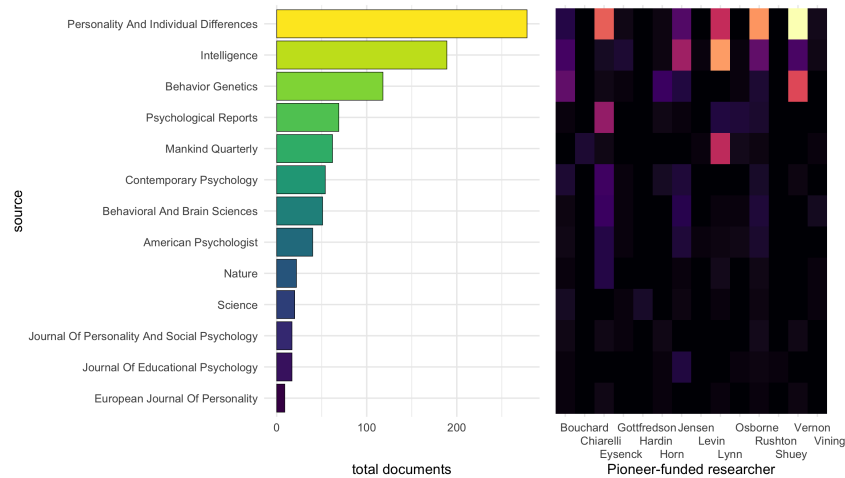


Figure 1: Journals publishing 6 or more Pioneer-funded researchers, Web of Science author search results.

of mainstream journals are more prominent: *Personality and Individual Differences* (PID), *Intelligence* (I), *Behavior Genetics* (BG), and *Psychological Reports* (PR). In addition, only psychologist Richard Lynn appears to have published heavily in MQ. Lynn became an assistant editor of MQ in 1979 (vol. XX, Nos. 1 & 2) and is listed as editor-in-chief on MQ's current website as of 2023-07-21. He has also been president of Pioneer since the death of psychologist J. Phillippe Rushton in 2012 (20). By contrast, a number of Pioneer-funded researchers have published in PID, I, and to a lesser degree BG: Bouchard, Eysenck, Jensen, Rushton, Vernon, and also Lynn.

*Personality and Individual Differences* (PID) was founded in 1980, with Eysenck as editor-in-chief and an editorial board including Jensen and Lynn. In the inaugural editorial, Eysenck identified "studies of the genetic determinants of individual differences in the areas of personality and intelligence" as one of the journal's eight major areas of interest. Eysenck remained editor-in-chief until his death in 1997. In 2005 the editorial board still included Jensen and Lynn. PID was first published by Pergamon Press, a mainstream academic press, and today is published by Elsevier.

*Intelligence* (I) was founded in 1977, with psychologist Douglas Detterman as editor-in-chief from the founding until 2016. Lloyd Humphreys was on the editorial board starting from 1977; by 1990 he had been joined by Jensen and Philip Vernon. Richard Lynn joined the editorial board sometime between 1998 and 2002. (Archive copies of the I editorial board page are not available from the journal’s website from 1999 through 2001.) I has been criticized for including Lynn and Gerhard Meisenberg — who was editor-in-chief of MQ in 2015-18 — on its editorial board until 2018 (9). I is published by Elsevier.

## **Topic model analysis identifies race science discourse**

The fact that Pioneer-funded researchers published heavily in two mainstream psychology journals does not tell us anything about the content of their publications or the claims of their research. We assembled a full-text corpus of articles published in MQ, PID, I, BG, PR, and Behavior and Brain Sciences (BBS) between 1960-2010 and used topic modeling to identify topics discussing race, intelligence, and both.

Figure 2 and this discussion focuses on three topics identified in one of the 24 fitted topic models; see figures S1.3, S1.4, S1.5 and supplemental visualizations S2-S4 for all topics from all models and S5-S7 for selected topics from all models. In the focal model, topic 07 is strongly associated with MQ: MQ published dozens of articles in this topic each year, and no other journal ever published more than a handful. The Silge plot of the top 15 terms (21) contains racial terms (races, whites, negroes and potentially europe, africa, india, and japan) as well as book, likely reflecting the fact that MQ published a number of book reviews, while the psychology journals either did not or these were not available for the corpus.

Topic 22 is strongly associated with I and PID in the same way that 07 is associated with MQ. The Silge plot does contain jensen, as well as a reference to Raymond Cattell, who played a major role in the development of factor analysis and intelligence testing but also advocated for

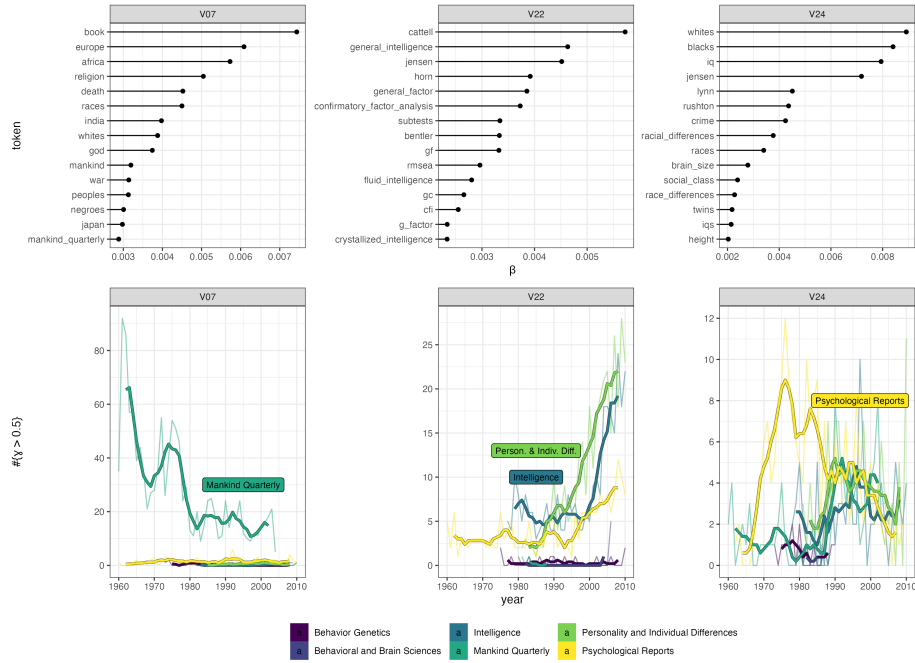


Figure 2: Silge plots (21) and smoothed time series for three focal topics. Top row: top 15 terms by  $\beta$  (term-topic distribution) for each topic. Bottom row: count of articles associated with each topic, by journal and year. Article counts use a threshold approach, with  $\gamma$  (topic-document distribution) greater than 0.5. Thin lines give annual values, thick lines give 5-year running averages. Because *Behavior Genetics* and *Behavioral and Brain Sciences* are not prominent in any panel, neither is given a direct label.

eugenics, fascism, and Nazi race science (22). However, the other authors named in this topic — John Horn and Peter Bentler — do not appear to have contributed to scientific racism. Instead this topic appears to identify “mainstream” (non-race science) intelligence research, especially factor analysis and the debate over whether intelligence is unidimensional or multidimensional. (Three other topics only associated with mainstream intelligence journals and terms were also identified by this topic model.) This topic indicates that the model is not simply lumping race science research together with other intelligence research.

The Silge plot for topic 24, by contrast, suggests a distinct race science discourse topic, with multiple racial terms (`whites`, `blacks`, `racial_differences`, `races`, `race_differences`) and the names of three prominent Pioneer-funded researchers, `jensen`, `lynn`, and `rushton`. Independent qualitative coding of the top 121 papers in this topic (those with  $\gamma > 0.97$ ) confirmed this interpretation of the topic, with 108 (89%) coded as race science discourse by both authors (Cohen’s  $\kappa = 0.86$ ).

In almost all years, most papers in topic 24 were published in mainstream journals rather than MQ. Jensen’s “How Much Can We Boost IQ and Scholastic Achievement?” (23) was published in 1969 in Harvard Educational Review (not included in this study), and the time series indicates that, around this time, there was an increase in articles in topic 24 in both PS and MQ (the only two journals in our corpus that were active at the time). MQ shows another sharp increase in the late 1970s, around the time Lynn was brought on as assistant editor. This finding is consistent with the conclusion of a qualitative analysis that “Psychology became the dominant discipline upon the editorship of Richard Lynn and Gerhard Meisenberg in MQ, where some issues were almost entirely represented by IQ test-based studies” (11, p. 110). PID published multiple papers in topic 24 almost immediately after it was founded, with I showing a more gradual increase between the mid-1970s and mid-1990s.



## **Race science, behavior genetics, and psychology**



It's notable that BG published very few articles in any race-and-intelligence topic identified by any of the 24 topic models. Among both academics and the general public, the field of behavior genetics is strongly associated with race science, and specifically race-and-intelligence research. Panofsky (24) argues that, prior to Jensen's 1969 paper, behavior genetics emphasized the study of non-human animals and intentionally avoided associations with eugenics and public controversy more generally. In response to Jensen, critics such as Lewontin offered broad critiques of behavior genetics as such, and behavior geneticists in turn adopted a radical conception of academic freedom (without any sense of responsibility for the social implications of academic research) and a siege or wartime mentality, as demonstrated by Sandra Scarr's 1986 presidential address to the Behavior Genetics Association (BGA) (25). Scarr's address coincided with the period between 1970-1990 when BG published articles in topic 24 (including Scarr's address itself).

However, the topic model analysis suggests that by the 1990s, behavior genetics may have distanced itself somewhat from race science, albeit without directly repudiating it. This interpretation is further supported by two disciplinary histories of behavior genetics, one published in McGue in 2010 (26) and the other by Plomin in 2023 (27). In addition, the 1995 BGA presidential address by Glayde Whitney — in which he criticized the “Marixist-Lysenkoist denial of genetics” and proposed that differences in murder rates between countries and cities was caused by racial genetic differences in intelligence, empathy, aggression, and impulsivity — was published in MQ rather than BG. As a discipline, behavior genetics may have been prominent in forming an echo chamber for race science in the past, but appears to have been less receptive to such ideas in more recent history.



Instead, our results suggest that psychology has been the dominant discipline in which the

168 race science echo chamber has been maintained. This may be the result of racial essentialization  
169 being maintained more easily in a discipline like psychology than in other disciplines that  
170 contribute to behavior genetics, like genetics and anthropology. The typical practitioners in psy-  
171 chology do not have the degree of formal education about, and therefore knowledge of, relevant  
172 underlying genetic and anthropological research that has led many professional organizations in  
173 genetics and anthropology to formally reject race as a biologically meaningful concept (28–34),  
174 undermining the core assumption of race science. As described by Winston, however, “[s]ince  
175 the 1990s, scientific racism has evolved in ways that sidestep or blunt the charge of essentialist  
176 categorization” (35, p. 427), such as shifting the rhetoric away from race to concepts such as  
177 “national IQ” that are deployed to preserve an essentialist categorization of different human  
178 populations. Professional psychological organizations routinely publish statements rejecting  
179 racism or acknowledging psychology’s complicity in systemic racism (36, 37), but stop short  
180 of rejecting race as a meaningful biological categorization of people, allowing for race science  
181 research to be smuggled into psychology and giving the appearance of continued scientific  
182 legitimacy.

## 183 Discussion

184 Using bibliometric and text-mining methods, this study finds that Pioneer-funded researchers  
185 published heavily in certain mainstream psychology journals, much more than in *Mankind*  
186 *Quarterly*; and that a distinct topic of race science discourse, centered on race and intelligence,  
187 can be identified in these same journals across the period 1960–2010. These findings indicate  
188 that *Mankind Quarterly* (MQ) was less important as an echo chamber for late 20th century race  
189 science than mainstream psychology journals, especially *Intelligence* (I) and *Personality and*  
190 *Individual Differences* (PID). Indeed, MQ published a very different kind of race science from  
191 the race-and-intelligence research published in I and PID.

192 We stress that we are not identifying any particular scientist as a racist, scientific racist,  
193 or race scientist based on the topic model results alone. Scientific research can be utilized to  
194 promote scientific racism — and thus count as race science — independently of the views and  
195 attitudes of the scientist who conducted that research (38–41). But this does not mean that  
196 scientists are not responsible for the downstream social effects of their research, even when  
197 these effects involve deliberate misinterpretation (42–45). Like other citizens, scientists are  
198 responsible for mitigating the reasonably foreseen harms that result from their actions. And we  
199 now have decades of empirical evidence that research on the genetics of intelligence — even  
200 research that does not itself touch on race — can and will be used to promote scientific racism  
201 and indeed racial violence (46). For instance, while our findings suggest that behavior genetics  
202 has attempted to distance itself from scientific racism, it has done little to explicitly repudiate  
203 this research, especially in contrast to statements by biological anthropologists (31) and human  
204 geneticists (47). Behavior genetics research — and research on human genetics more generally  
205 — may be especially susceptible to appropriation by scientific racists, which only strengthens the  
206 obligations these researchers have to prevent their research from causing harm (40, 48).

207 Our distinction between race science and race science discourse highlights an important  
208 “engagement dilemma” for scientists challenging race science and scientific racism. While  
209 the “tobacco strategy” for manufacturing doubt targets mainstream scientists and promotes  
210 fringe scientific views, its primary audiences are lawmakers and the general public. The aim of  
211 the strategy is not to change the opinion of mainstream scientists, but instead to cultivate the  
212 impression that there is legitimate scientific uncertainty about the issue, reasonable disagreement  
213 among qualified experts (14). On the one hand, refusing to engage with fringe scientists — or,  
214 often, pseudoscientists (49) — means that false or misleading statements go unchallenged. On  
215 the other hand, detailed, technical criticism legitimizes the fringe views, creating exactly the  
216 impression of reasonable scientific disagreement that the tobacco strategy is intended to produce.

217        In the case of race science, mainstream psychologists have often taken a non-engagement  
218 approach — whether due to a strong conception of academic freedom or a desire to avoid stoking  
219 controversy — with the result that race scientists have been able to present themselves as the  
220 representatives of “mainstream science on intelligence.” On the other hand, critics — most  
221 prominently Gould, Kamin, and Lewontin — have offered sophisticated technical critiques. But  
222 these tend to spiral off into esoteric debates that are inaccessible to the general public, as with  
223 “Lewontin’s fallacy” (50). Consequently these critiques are likely to be either ignored by the  
224 general public or taken as evidence of legitimate scientific debate.

## References and Notes

## References and Notes

1. E. Barkan, *The Retreat of Scientific Racism: Changing Concepts of Race in Britain and the United States Between the World Wars* (Cambridge University Press, 1992).
2. T. Duster, *Backdoor to Eugenics* (Routledge, 2003).
3. A. Panofsky, K. Dasgupta, N. Iturriaga, How White nationalists mobilize genetics: From genetic ancestry and human biodiversity to counterscience and metapolitics, *American Journal of Physical Anthropology* **175**, 387 (2021).
4. J. P. Jackson Jr., Arthur Jensen, evolutionary biology, and racism, *History of Psychology* pp. No Pagination Specified–No Pagination Specified (2022).
5. B. Mehler, Foundation for fascism: The new eugenics movement in the United States, *Patterns of Prejudice* **23**, 17 (1989).
6. A. S. Winston, Science in the Service of the Far Right: Henry E. Garrett, the IAAEE, and the Liberty Lobby, *Journal of Social Issues* **54**, 179 (1998).
7. W. H. Tucker, *The Funding of Scientific Racism: Wickliffe Draper and the Pioneer Fund* (University of Illinois Press, Urbana, 2002).
8. G. Schaffer, “‘Scientific’ Racism Again?”: Reginald Gates, the Mankind Quarterly and the Question of “Race” in Science after the Second World War, *Journal of American Studies* **41**, 253 (2007).
9. A. Saini, *Superior: The Return of Race Science* (Beacon Press, 2019).

- 245 10. A. S. Winston, *Oxford Research Encyclopedia of Psychology* (Oxford University Press,  
246 2020).
- 247 11. D. Adams, M. Pilloud, The (Mis)appropriation of Biological Anthropology in Race Science  
248 and the Implications for Forensic Anthropology, *FA* (2021).
- 249 12. A. Saini, Draper's Millions: The Philanthropic Wellspring of Modern Race Science,  
250 [https://race.undark.org/articles/drapers-millions-the-philanthropic-wellspring-of-modern-](https://race.undark.org/articles/drapers-millions-the-philanthropic-wellspring-of-modern-race-science?utm_source=Undark:+News+%26+Updates&utm_campaign=68be9b0aaa-RSS_EMAIL_CAMPAIGN&utm_medium=email&utm_term=0_5cee408d66-68be9b0aaa-176017345)  
251 [race-science?utm\\_source=Undark:+News+%26+Updates&utm\\_campaign=68be9b0aaa-](https://race.undark.org/articles/drapers-millions-the-philanthropic-wellspring-of-modern-race-science?utm_source=Undark:+News+%26+Updates&utm_campaign=68be9b0aaa-RSS_EMAIL_CAMPAIGN&utm_medium=email&utm_term=0_5cee408d66-68be9b0aaa-176017345)  
252 [RSS\\_EMAIL\\_CAMPAIGN&utm\\_medium=email&utm\\_term=0\\_5cee408d66-](https://race.undark.org/articles/drapers-millions-the-philanthropic-wellspring-of-modern-race-science?utm_source=Undark:+News+%26+Updates&utm_campaign=68be9b0aaa-RSS_EMAIL_CAMPAIGN&utm_medium=email&utm_term=0_5cee408d66-68be9b0aaa-176017345)  
253 [68be9b0aaa-176017345](https://race.undark.org/articles/drapers-millions-the-philanthropic-wellspring-of-modern-race-science?utm_source=Undark:+News+%26+Updates&utm_campaign=68be9b0aaa-RSS_EMAIL_CAMPAIGN&utm_medium=email&utm_term=0_5cee408d66-68be9b0aaa-176017345) (2022).
- 254 13. J. P. Jackson, Jr., *Science for Segregation: Race, Law, and the Case Against Brown V. Board*  
255 *of Education* (NYU Press, 2005).
- 256 14. M. Fernández Pinto, To Know or Better Not to, *Science & Technology Studies* **30**, 53 (2017).
- 257 15. J. P. Jackson, Jr., A. S. Winston, The Mythical Taboo on Race and Intelligence, *Review of*  
258 *General Psychology* (2020).
- 259 16. N. Oreskes, E. M. Conway, *Merchants of Doubt: How a Handful of Scientists Obscured the*  
260 *Truth on Issues from Tobacco Smoke to Global Warming* (Bloomsbury, New York, NY, USA,  
261 2011).
- 262 17. L. Jasny, J. Waggle, D. R. Fisher, An empirical examination of echo chambers in US climate  
263 policy networks, *Nature Clim Change* **5**, 782 (2015).
- 264 18. J. O. Weatherall, C. O'Connor, J. P. Bruner, How to Beat Science and Influence People:  
265 Policy-Makers and Propaganda in Epistemic Networks, *Br J Philos Sci* (2018).

- 266 19. A. Miller, The Pioneer Fund: Bankrolling the Professors of Hate, *The Journal of Blacks in*  
267 *Higher Education* pp. 58–61 (1994).
- 268 20. H. Beirich, Pioneer Fund Assets Divided; New Leadership Appointed,  
269 [https://www.splcenter.org/hatewatch/2013/10/22/pioneer-fund-assets-divided-new-](https://www.splcenter.org/hatewatch/2013/10/22/pioneer-fund-assets-divided-new-leadership-appointed)  
270 [leadership-appointed](https://www.splcenter.org/hatewatch/2013/10/22/pioneer-fund-assets-divided-new-leadership-appointed) (2013).
- 271 21. *Topic Modeling* ("O'Reilly Media, Inc.", 2017), p. ch. 6.
- 272 22. B. Mehler, Beyondism: Raymond B. Cattell and the new eugenics, *Genetica* **99**, 153 (1997).
- 273 23. A. Jensen, How Much Can We Boost IQ and Scholastic Achievement?, *Harvard Educational*  
274 *Review* **39**, 1 (1969).
- 275 24. A. Panofsky, *Misbehaving Science: Controversy and the Development of Behavior Genetics*  
276 (University of Chicago Press, Chicago, IL, 2014).
- 277 25. S. Scarr, Three cheers for behavior genetics: Winning the war and losing our identity, *Behav*  
278 *Genet* **17**, 219 (1987).
- 279 26. M. McGue, The End of Behavioral Genetics?, *Behav Genet* **40**, 284 (2010).
- 280 27. R. Plomin, Celebrating a Century of Research in Behavioral Genetics, *Behav Genet* **53**, 75  
281 (2023).
- 282 28. O. Anthropology, Aapa statement on biological aspects of race, *American Journal of Physical*  
283 *Anthropology* **101**, 569 (1996).
- 284 29. A. S. of Human Genetics, Ashg denounces attempts to link genetics and racial supremacy,  
285 *American Journal of Human Genetics* **103**, 636 (2018).
- 286 30. A. E. Board, Aaa statement on race, *American Anthropologist* **100**, 712 (1998).

- 287 31. A. Fuentes, R. R. Ackermann, S. Athreya, D. Bolnick, T. Lasisi, S.-H. Lee, S.-A. McLean,  
288 R. Nelson, AAPA Statement on Race and Racism, *American Journal of Physical Anthropol-*  
289 *ogy* **169**, 400 (2019).
- 290 32. N. H. G. R. Institute, Race, <https://www.genome.gov/genetics-glossary/Race> (2023).
- 291 33. C. Rotimi, Ashg statement regarding the warping of genetic knowledge to feed racist  
292 ideology, [https://www.ashg.org/publications-news/ashg-news/ashg-statement-regarding-the-](https://www.ashg.org/publications-news/ashg-news/ashg-statement-regarding-the-warping-of-genetic-knowledge-to-feed-racist-ideology/)  
293 [warping-of-genetic-knowledge-to-feed-racist-ideology/](https://www.ashg.org/publications-news/ashg-news/ashg-statement-regarding-the-warping-of-genetic-knowledge-to-feed-racist-ideology/) (2022).
- 294 34. A. Wynshaw-Boris, American society of human genetics statement regarding con-  
295 cepts of “good genes” and human genetics, [https://www.ashg.org/publications-news/ashg-](https://www.ashg.org/publications-news/ashg-news/statement-regarding-good-genes-human-genetics/)  
296 [news/statement-regarding-good-genes-human-genetics/](https://www.ashg.org/publications-news/ashg-news/statement-regarding-good-genes-human-genetics/) (2020).
- 297 35. A. S. Winston, Why mainstream research will not end scientific racism in psychology,  
298 *Theory & Psychology* **30**, 425 (2020).
- 299 36. A. P. Association, Apology to people of color for apa’s role in promoting, perpetuat-  
300 ing, and failing to challenge racism, racial discrimination, and human hierarchy in us,  
301 <https://www.apa.org/about/policy/racism-apology> (2021).
- 302 37. A. for Psychological Science, Statement on racism, bias, and intoler-  
303 ance, [https://www.psychologicalscience.org/news/statement-on-racism-bias-and-](https://www.psychologicalscience.org/news/statement-on-racism-bias-and-intolerance.html)  
304 [intolerance.html](https://www.psychologicalscience.org/news/statement-on-racism-bias-and-intolerance.html) (2021).
- 305 38. J. Tabery, Why Is Studying the Genetics of Intelligence So Controversial?, *Hastings Center*  
306 *Report* **45**, S9 (2015).
- 307 39. D. Gillborn, Softly, softly: Genetics, intelligence and the hidden racism of the new geneism,  
308 *Journal of Education Policy* **31**, 365 (2016).



- 309 40. J. Carlson, K. Harris, Quantifying and contextualizing the impact of bioRxiv preprints  
310 through automated social media audience segmentation, *PLOS Biology* **18**, e3000860 (2020).
- 311 41. B. M. Henn, E. K. Merchant, A. O'Connor, T. Rulli, Why DNA Is No  
312 Key to Social Equality: On Kathryn Paige Harden's "The Genetic Lottery",  
313 [https://www.lareviewofbooks.org/article/why-dna-is-no-key-to-social-equality-on-](https://www.lareviewofbooks.org/article/why-dna-is-no-key-to-social-equality-on-kathryn-paige-hardens-the-genetic-lottery/)  
314 [kathryn-paige-hardens-the-genetic-lottery/](https://www.lareviewofbooks.org/article/why-dna-is-no-key-to-social-equality-on-kathryn-paige-hardens-the-genetic-lottery/) (2021).
- 315 42. H. E. Douglas, *Science, Policy, and the Value-Free Ideal* (University of Pittsburgh Press,  
316 Pittsburgh, Pa, 2009).
- 317 43. P. Kitcher, An Argument About Free Inquiry, *Noûs* **31**, 279 (1997).
- 318 44. J. A. Kourany, Should Some Knowledge Be Forbidden? The Case of Cognitive Differences  
319 Research, *Philosophy of Science* **83**, 779 (2016).
- 320 45. N. J. Block, G. Dworkin, IQ, Heritability and Inequality, Part 2, *Philosophy & Public Affairs*  
321 **4**, 40 (1974).
- 322 46. M. N. Meyer, P. S. Appelbaum, D. J. Benjamin, S. L. Callier, N. Comfort, D. Conley,  
323 J. Freese, N. A. Garrison, E. M. Hammonds, K. P. Harden, S. S.-J. Lee, A. R. Martin, D. O.  
324 Martschenko, B. M. Neale, R. H. C. Palmer, J. Tabery, E. Turkheimer, P. Turley, E. Parens,  
325 Wrestling with Social and Behavioral Genomics: Risks, Potential Benefits, and Ethical  
326 Responsibility, *Hastings Center Report* **53**, S2 (2023).
- 327 47. C. S. Jackson, D. Turner, M. June, M. V. Miller, Facing Our History—Building an Equitable  
328 Future, *The American Journal of Human Genetics* **110**, 377 (2023).
- 329 48. C. B. Ogbunugafor, DNA, basketball, and birthday luck. A review of The genetic lottery:  
330 Why DNA matters for social equality. By Kathryn Paige Harden, 2021, US: Princeton Uni-

versity Press. 312 pp. ISBN: 9780691190808 (Hardcover), *American Journal of Biological Anthropology* **179**, 501 (2022).

49. S. Bhakthavatsalam, W. Sun, A Virtue Epistemological Approach to the Demarcation Problem, *Sci & Educ* (2021).

50. C. C. Roseman, Lewontin did not commit Lewontin's fallacy, his critics do: Why racial taxonomy is not useful for the scientific study of human variation, *BioEssays* **43**, 2100204 (2021).

51. J. Ooms, Pdftools: Text Extraction, Rendering and Converting of PDF Documents (2023).

52. S. Chamberlain, H. Zhu, N. Jahn, C. Boettiger, K. Ram, Rcrossref: Client for Various 'CrossRef' 'APIs' (2019).

53. J. R. Flynn, Evidence against Rushton: The genetic loading of WISC-R subtests and the causes of between-group IQ differences, *Personality and Individual Differences* **26**, 373 (1998).

54. J. P. Rushton, Secular gains in IQ not related to the g factor and inbreeding depression – unlike Black-White differences: A reply to Flynn, *Personality and Individual Differences* **26**, 381 (1998).

55. J. R. Flynn, Reply to Rushton: A gang of gs overpowers factor analysis, *Personality and Individual Differences* **26**, 391 (1998).

56. H. Wickham, J. Hester, J. Ooms, RStudio, R. F. C. o. R.-p. homepage cached as example), Xml2: Parse XML (2023).

57. spaCy: Industrial-strength Natural Language Processing (NLP) with Python and Cython, Explosion AI (2018).

- 353 58. K. Benoit, A. Matsuo, E. R. Council (ERC-2011-StG 283794-QUANTESS), Spacyr: Wrap-  
354 per to the 'spaCy' 'NLP' Library (2020).
- 355 59. D. Vohra, *Practical Hadoop Ecosystem: A Definitive Guide to Hadoop-Related Frameworks*  
356 *and Tools*, D. Vohra, ed. (Apress, Berkeley, CA, 2016), pp. 325–335.
- 357 60. N. Richardson, I. Cook, N. Crane, D. Dunnington, R. François, J. Keane, D. Moldovan-  
358 Grünfeld, J. Ooms, J. Luraschi, K. D. Werner, J. Wong, A. Arrow, Arrow: Integration to  
359 'Apache' 'Arrow' (2023).
- 360 61. A. Sclafani, Correlation and causation in the study of feeding behavior, *Behavioral and*  
361 *Brain Sciences* **4**, 590 (1981).
- 362 62. D. J. Hicks, Productivity and interdisciplinary impacts of Organized Research Units, *Quantitative Science Studies* **2**, 990 (2021).
- 364 63. K. Rohe, M. Zeng, Vintage Factor Analysis with Varimax Performs Statistical Inference  
365 (2020).
- 366 64. M. E. Roberts, B. M. Stewart, D. Tingley, Stm: An R Package for Structural Topic Models,  
367 *Journal of Statistical Software* **91**, 1 (2019).
- 368 65. D. J. Hicks, Tmfast fits topic models fast (2023).
- 369 66. G. E. Gignac, G. Ekermans, Group differences in EI within a sample of black and white  
370 South Africans, *Personality and Individual Differences* **49**, 639 (2010).
- 371 67. John C. Loehlin (1926–2020), *Behav Genet* **50**, 385 (2020).
- 372 68. I. Waldman, E. Turkheimer, Introduction to a Festschrift for John Loehlin, *Behav Genet* **44**,  
373 547 (2014).

- 374 69. R. Plomin, Genotype-Environment Correlation in the Era of DNA, *Behav Genet* **44**, 629  
375 (2014).
- 376 70. P. L. Roubertoux, Jerry Hirsch (20 September 1922–3 May 2008): A Tribute, *Behav Genet*  
377 **38**, 561 (2008).
- 378 71. J. L. Rodgers, K. Jacobson, E. van den Oord, Obituary: David Christian Rowe, *Behav Genet*  
379 **33**, 627 (2003).
- 380 72. K. C. Jacobson, D. C. Rowe, Genetic and Shared Environmental Influences on Adolescent  
381 BMI: Interactions with Race and Sex, *Behav Genet* **28**, 265 (1998).
- 382 73. S. M. Berman, E. P. Noble, Reduced visuospatial performance in children with the D2  
383 dopamine receptor A1 allele, *Behav Genet* **25**, 45 (1995).
- 384 74. E. Bonilla-Silva, *Racism Without Racists: Color-Blind Racism and the Persistence of Racial*  
385 *Inequality in the United States*, 6th ed. (Rowman & Littlefield Publishers, Lanham, MD,  
386 2022).
- 387 75. J. Saul, *New Work on Speech Acts*, D. Fogal, D. W. Harris, M. Moss, eds. (Oxford University  
388 Press, 2018), p. 0.
- 389 76. C. P. Alderfer, The science and nonscience of psychologists' responses to The Bell Curve,  
390 *Professional Psychology: Research and Practice* **34**, 287 (2003).
- 391 77. M. Wills, Are Clusters Races? A Discussion of the Rhetorical Appropriation of Rosenberg et  
392 al.'s "Genetic Structure of Human Populations", *Philosophy & Theory in Biology* **9** (2017).
- 393 78. L. Trahan, K. K. Stuebing, M. K. Hiscock, J. M. Fletcher, The Flynn Effect: A Meta-analysis,  
394 *Psychol Bull* **140**, 1332 (2014).

## 395 **Supplemental Materials**

396 S1. Materials and methods

397 Figures S2-S7

398 Data S8, S9

## **S1 Materials and methods**

### **S1.1 Corpus assembly**

Corpus assembly is summarized in figure S1.1. The corpus was assembled in two parts, one for *Mankind Quarterly* and the other for mainstream journals.

#### **S1.1.1 Mankind Quarterly**

An electronic archive of *Mankind Quarterly*, from the first issue in 1969 through 2004, is available for free on the open web at the white nationalist website *The Unz Review*. The Python package Beautiful Soup (<https://www.crummy.com/software/BeautifulSoup/>, version not recorded) was used to retrieve every PDF available in this archive. After identifying some gaps (missing issues) in the archive from *The Unz Review*, we manually retrieved PDFs for further articles using ProQuest. Text was extracted using the R package `pdftools` (51, version 3.0.1). This stage of corpus assembly was conducted in Fall 2021 with the assistance of Anthony Sainez.

#### **S1.1.2 Mainstream journals**

To identify suitable mainstream journals for inclusion in the corpus, we first identified academic researchers who had been funded by the Pioneer Fund. Pioneer is an American non-profit organization founded in 1937 to fund eugenics research and propaganda (7). After the *Brown v. Board of Education* ruling in 1954, Pioneer funded various segregationist efforts across the United States, including lectures on eugenics by Stanford physicist William Shockley (13). We reviewed a critical profile of Pioneer (19) as well as an archived page from the organization's own web site (<https://web.archive.org/web/20130103005545/http://www.pioneerfund.org/Grantees.html>), which together listed 16 researchers who had received Pioneer funds.

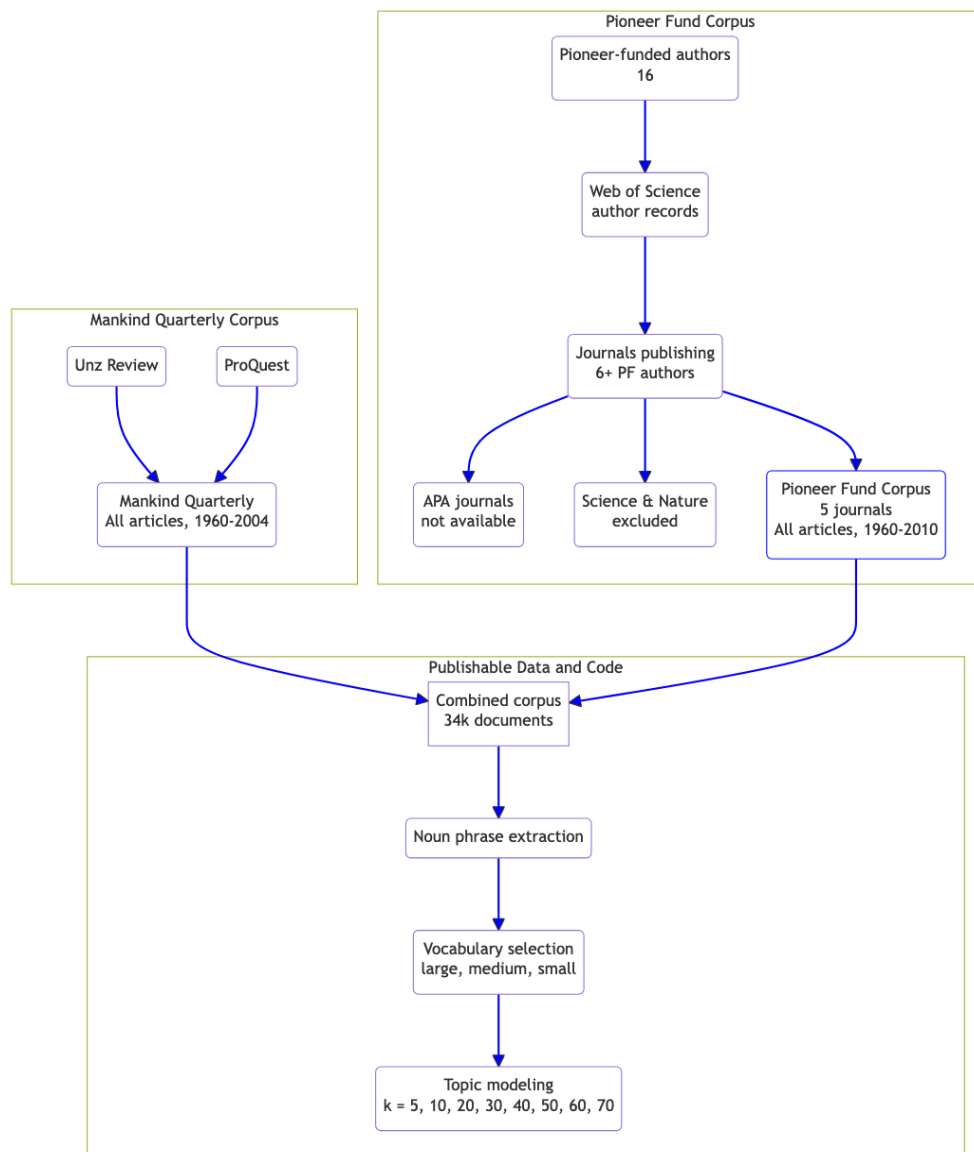


Figure S1.1: Corpus assembly.

We then used the author search tool in Clarivate’s Web of Science platform (<https://www.webofscience.com/wos/author/search>), retrieving publication lists for 14 researchers. These searches were conducted between 2021-09-24 and 2021-10-05 by DJH. After parsing these results, we counted how many of the 14 researchers had published in each journal. Thirteen journals had published 6 or more of the 14 Pioneer-funded researchers. We excluded *Mankind Quarterly* (as already included) as well as *Science* and *Nature* (as too general) from further consideration in this side of the corpus. Three journals published by the American Psychological Association (APA; *American Psychologist*, *Contemporary Psychology*, and *Journal of Educational Psychology*) had to be excluded due to confusion over who could give us permission to use the archives for a text mining project, with both APA and ProQuest asserting that we needed to get permission from the other entity. *European Journal of Personality*, published by SAGE, also had to be excluded because our institutional access only went back to 1999.

The remaining 5 journals are all published by major academic publishers — Elsevier, Springer, or Cambridge University Press — and each item in the entire run of each journal has been assigned a DOI (digital object identifier) for archival purposes. We used the Crossref API and `rcrossref` R interface to this API (52, *version 1.1.0.99*) to retrieve metadata for each item published from 1960-2010 in each of these journals. These metadata included item-level license information — confirming that the text of each item could be used for text mining projects — and a URL to an electronic version of the item. These URLs were used to retrieve a XML or PDF version of each item, except for *Personality and Individual Differences*. This journal has published a relatively large number of non-article documents, such as book reviews and commentaries, that are not available at the URL included in the Crossref metadata. (This is unfortunate, as it was not difficult to find highly relevant documents that we could not automatically retrieve and therefore had to be excluded from our corpus. One set of examples is an exchange between Rushton and Flynn on the Flynn effect (53–55).) Instead we used Elsevier’s



ScienceDirect API (<https://dev.elsevier.com/>) to independently search and retrieve all available items from *Personality and Individual Differences*. This stage of corpus assembly was conducted between 2021-11 and 2022-05.

*Behavioral and Brain Sciences* typically uses a target article + commentary format; in some cases there are dozens of commentaries for a single article. In the Crossref DOI metadata, each target article and individual commentary is given its own DOI, with no distinction between contribution types or metadata links between a commentary and the corresponding target article. But text is only available in aggregate PDFs that bundle together the target article and all of the commentaries. In some cases this results in PDFs that are hundreds of pages long and might be linked from the Crossref metadata 30+ times. In addition, the retrieved PDFs are not perfectly identical, because Cambridge UP's servers add a timestamped watermark to each page when the PDF is requested. We attempted to contact Cambridge for assistance but did not receive a response. We ultimately used a series of ad hoc measures to mitigate text duplication. Approximately 100,000 documents that were 1-2 pages long were excluded before PDF retrieval. After PDF retrieval and text extraction, the timestamped watermarks were removed using a regular expression and the text was hashed using SHA 256. Hashes were used to construct groups of duplicate documents, and a single document (whichever one happened to be first in the dataframe) was chosen from each hash group for inclusion in the corpus. It is plausible that some duplicates made it through this process: where the watermark overlapped with the text, the regular expression likely would have been unable to identify and remove the watermark (a false negative result), and this small difference in the watermarks (not the text) would produce different hashes.

After XML/PDF versions were retrieved, text was extracted using either the `xml2` package (56) or `pdftools`. Figure S1.2 and table S1.1 show the number of fulltext documents included in the corpus. All together, the corpus includes 34,896 documents from 6 journals.

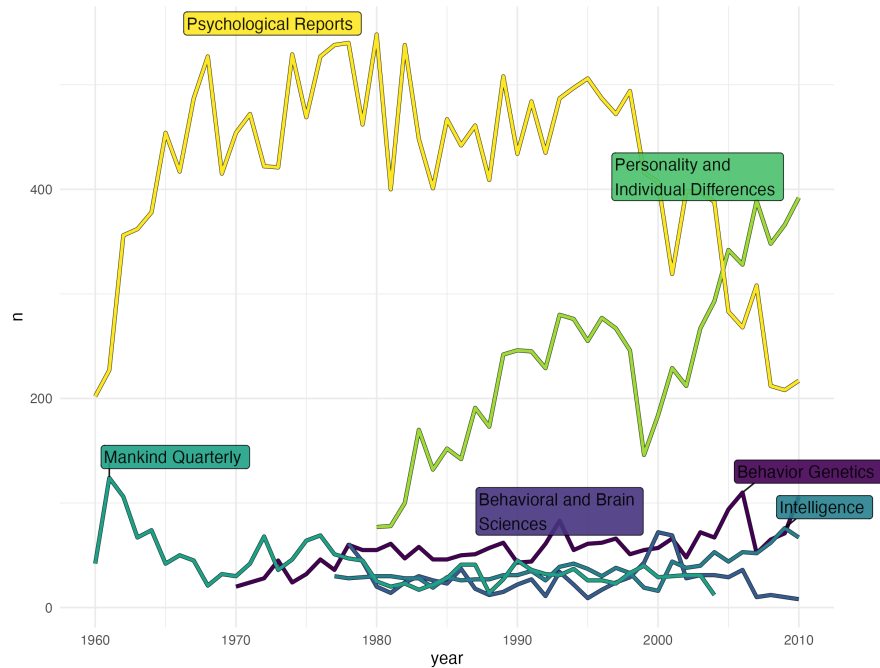


Figure S1.2: Count of documents in the corpus, by journal and year.

Table S1.1: Document counts, by journal, and years included in the corpus.

journal	n	start	end
Behavior Genetics	2268	1970	2010
Intelligence	1237	1977	2010
Mankind Quarterly	1821	1960	2004
Personality and Individual Differences	7274	1980	2010
Psychological Reports	21398	1960	2010
Behavioral and Brain Sciences	898	1978	2010

## S1.2 Text preparation

After document retrieval and text extraction, we pre-processed the text using the `spaCy` NLP (natural language processing) Python library (57, *version 1.9.0*) and the R API `spacyr` (58, *version 1.2.1*). Specifically, we applied regular expressions to remove header/footer copyright

476 notices and hyphenation, used spaCy to annotate and extract noun phrases (eg, “the intelligence  
477 test items”), and then cleaned and standardized these phrases (eg, removing the/an/a, converting all  
478 text to lowercase, and replacing all whitespace with underscores: “intelligence\_test\_items”). We  
479 then counted the occurrence of each noun phrase in the document. The aggregated “document-  
480 term matrix” was stored in Parquet format (59) for performance reasons, and written and read  
481 using the `arrow` package for R (60, version 7.0.0).

482 NLP-extracted noun phrases offer a number of advantages over the more traditional unigram  
483 (“single word”) terms. First, noun phrase extraction removes many standard stopwords (articles,  
484 common verbs, numbers) without relying on a fixed, a priori stopword list. Phrases can be more  
485 informative than single terms, for example, distinguishing “intelligence test” from “hypothesis  
486 test.” While simple n-gram extraction will include numerous phrases that are not especially  
487 meaningful. Consider the sentence “Since our first analyses of feeding patterns in rats, we had  
488 been using a criterion of 40 minutes (Le Magnen & Tallon 1963; 1966)” (61). Bigrams such as  
489 “since our” and “criterion 40” will likely be discarded in vocabulary selection, but significantly  
490 increase the computational cost of vocabulary selection. Noun phrase extraction is therefore  
491 more efficient.

492 After noun phrase extraction, the corpus comprised 43,162,055 total tokens of 6,320,783  
493 distinct phrases.

#### 494 **S1.2.1 Data and code availability**

495 Document metadata retrieved from Crossref does not appear to be covered by any copyright or  
496 other intellectual property restrictions. However, due to copyright restrictions, we are unable to  
497 make document fulltext or Web of Science search results publicly available. Code for the corpus  
498 assembly and text preparation steps described above is available by request.

499 The public analysis repository, <https://github.com/dhicks/race-science> [DOI],

includes document metadata and documentwise counts of NLP-extracted noun phrases (“document-term matrices”), along with code to reproduce the analysis of the following sections.

### S1.3 Vocabulary selection

We took an information-theoretic approach to vocabulary selection (62). Consider a game in which I draw a document from the corpus, then a single token from that document. I tell you the term, and you have to guess which document I picked. Intuitively, highly informative terms (in this project, noun phrase types) are distinctive, allowing you to dramatically narrow down the list of potential documents. This “informativeness” of a term can be quantified as the KL (Kullback-Leibler) divergence from a “baseline” uniform distribution across documents to the distribution conditional on the term, which we abbreviate as  $\Delta H$ . Because the most informative terms tend to be typos and OCR errors — these are unique to a single document — we multiply the KL divergence by the logarithm of the total number of occurrences of the term across the entire corpus. We refer to the resulting measure as  $\log(n)\Delta H$  or ndH.

In the current project, we found that this ndH approach heavily favored recurrent noun phrases in the longest documents. Many of the documents published in *Psychological Reports* are extremely short, 1-2 page brief notices of a single study; while many of the documents published in *Brain and Behavioral Sciences* are book-length collections that include a long review article and sometimes dozens of commentaries. Very generic noun phrases that happen to appear in BBS can occur orders of magnitude more often than highly distinctive phrases in the PS, and so the  $\log(n)$  factor overwhelms the  $\Delta H$  factor.

To address this, we used a different baseline distribution of documents, namely, one in which document probability is proportional to length. This makes phrases from short documents much more “surprising” (much less likely to occur according to the baseline), and hence substantially increases their informativeness. This was more effective at identifying useful phrases from across

the corpus.

A common rule of thumb in topic modeling is that the vocabulary should have about 10 times as many distinct terms as the number of documents in the corpus. However, we had some concerns with computational demands here: the resulting document-term matrix would have roughly  $10 \times 33,000^2$  or 10.9 billion entries; with 10% density this would require on the order of 4 GB of memory just for a single copy of the matrix; and most of the analysis was to be conducted on the authors’ laptops. We therefore chose to work with three smaller vocabularies,  $5\times$ ,  $1\times$  and  $\frac{1}{5}\times$  the number of documents. We refer to these as the “large” (174,480 distinct phrases), “medium” (34,896) and “small” (6,979) vocabulary, respectively. We compare findings across vocabularies as a robustness check.

## S1.4 Topic modeling

To fit topic models, we followed the approach proposed by Rohe and Zeng (63), which uses varimax-rotated partial principal components instead of the variational inference methods used by standard topic model packages such as `stm` (64). This novel approach was implemented in the R package `tmfast` (65). A simulation study of `tmfast` found that it was significantly faster and only slightly less accurate at reconstructing known word-topic and topic-document distributions, compared to `stm` (65).

Topic models were fit for all three vocabularies (large, medium, and small) with  $k = 5, 10, 20, \dots, 70$  (number of topics), resulting in a total of  $24 = 3 \times 8$  models.

Following the approach of (62), we did not attempt to identify a unique best fitted model for further analysis. While the manuscript text focuses on the medium vocabulary,  $k = 40$  model as the “median” among the 24 fitted models, we compare and contrast findings from this model with those from the other models.

#### S1.4.1 Topic model interpretation and quality assessment

After fitting topic models, our first research question was whether we could identify distinctive “race science discourse” topics. Figures S1.3, S1.4, S1.5 show the gamma (topic-document) distributions for each value of  $k$  for the three vocabularies. In these figures, each panel corresponds to a single journal- $k$  combination, each row of cells is a single document in that journal, and each column of cells is a single topic in that particular model. Topics correspond within columns of panels, but not within rows; for example, topic 04 for  $k = 10$  does not necessarily correspond to topic 04 for  $k = 20$ . Color intensity indicates the value of gamma for that particular topic-document combination.

For  $k = 5, 10$  the distributions are quite noisy, difficult to interpret, and generally don’t form very strong visual clusters. As  $k$  increases from 20 through 40, coherent bands start to appear for journal-distinctive topics in MQ, BG, I and, to a lesser degree, PID. Note that the distinctive topics for MQ and I do not obviously coincide. Above about  $k = 50$ , the models appear to identify more fine-grained topics and the coherent bands fade.

We next constructed Silge plots (21), showing the top 15 (highest-probability) terms (noun phrases) from each topic in the model. These are provided as supplemental PDFs S2-S4. Perusing these term lists, we focused on topics in each model that contained racial terms. Every fitted topic model included at least one racial topic. Importantly, in some cases these topics contained terms related to intelligence research, but in other cases they did not.

The gamma distributions and Silge plots suggested that, once  $k$  was sufficiently large, the topic models were distinguishing between two types of race science articles. We therefore examined the prevalence of topics that used racial terms, intelligence-research terms, or both, by journal and across all 24 models. Supplemental PDFs S5-7 provide these visualizations. In these visualizations, each row of facets corresponds to a value of  $k$  (excluding 5 as too noisy), and each column of facets corresponds to one topic. Topics are clustered based on whether the top

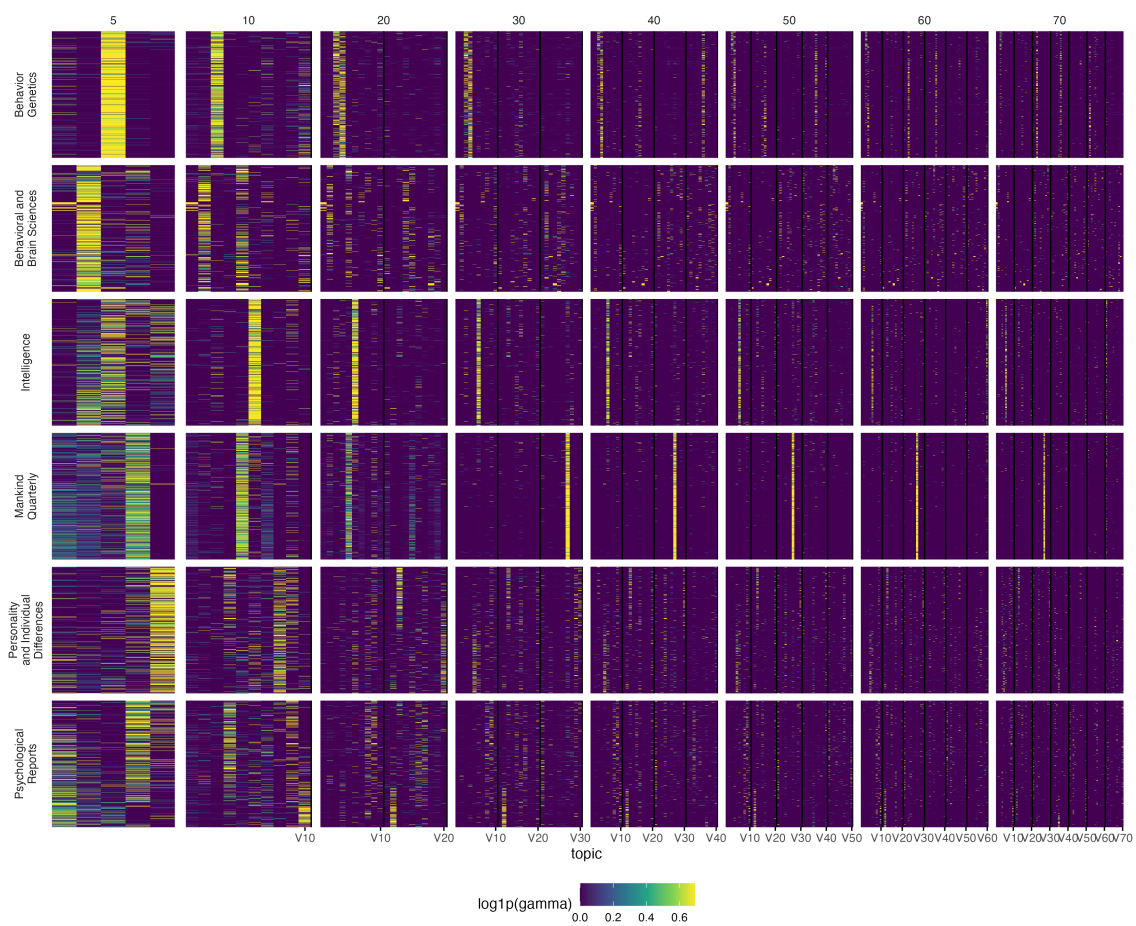


Figure S1.3: Fitted topic models: Gamma (topic-document) distributions, large vocabulary.

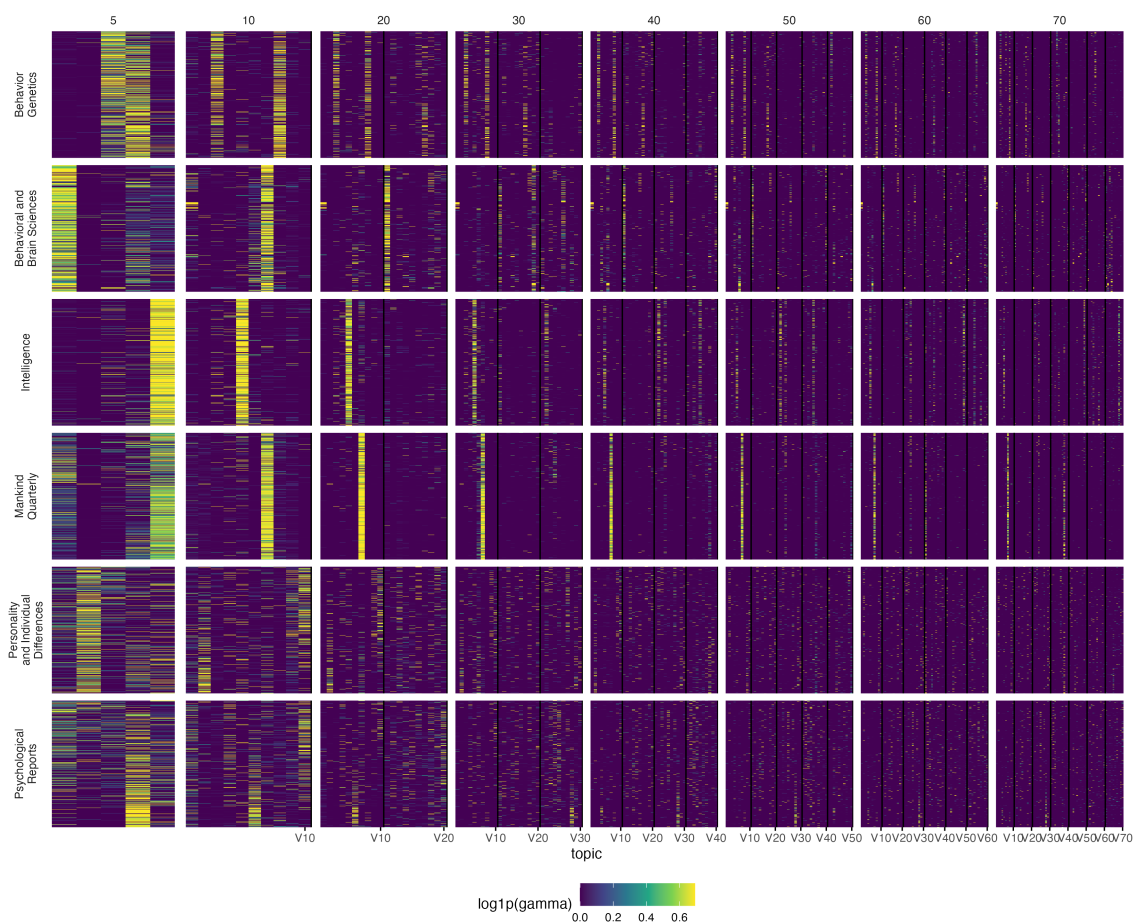


Figure S1.4: Fitted topic models: Gamma (topic-document) distributions, medium vocabulary.



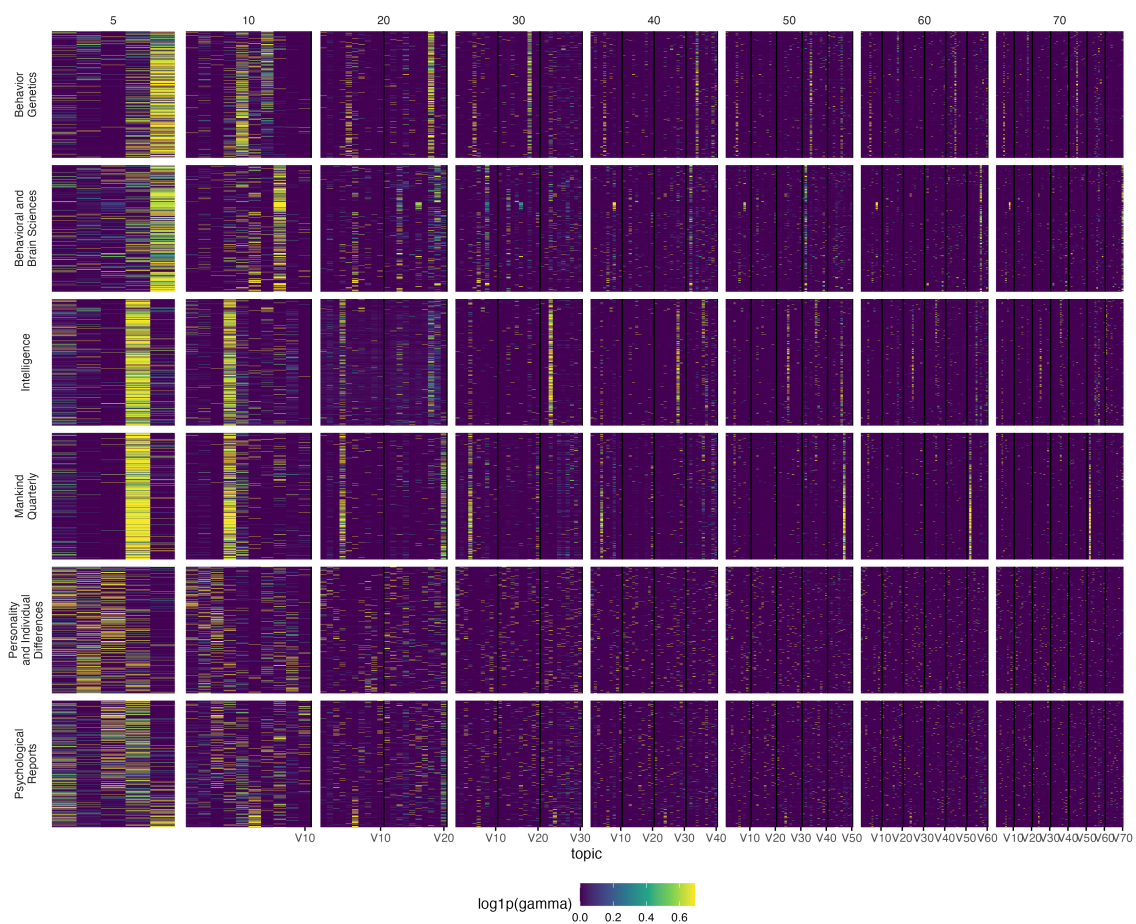


Figure S1.5: Fitted topic models: Gamma (topic-document) distributions, small vocabulary.

15 terms contain racial terms (“race”), terms related to intelligence research (“intelligence”), or both. Line plots show the count of articles with  $\gamma > 0.50$  for the given topic, by journal; thick lines are 5-year running averages and thin lines are raw counts.

Across all three vocabularies, for sufficiently large values of  $k$  the models identify two or more different racial topics, one that appears all but exclusively in MQ and at least one that appears primarily in mainstream journals.

We next conducted a topic quality check of topic 24 from the medium vocabulary,  $k = 40$  model, which appeared to identify race science research on intelligence published in mainstream journals. A spreadsheet of all articles that had a maximum value of  $\gamma$  for this topic was extracted (excluding articles published in *Psychological Reports*), and the top 121 articles ( $\gamma > 0.97$ ) were reviewed manually by both authors. We coded each article as *race science discourse*, or not, using our definition of race science discourse as treating race science as a legitimate area of scientific research; this includes methodological critiques of race science and empirical tests that falsify race science hypotheses.

For the first round of review, both authors worked independently, dichotomously coding each article as race science discourse or not. We calculated the “false positive rate” — documents with maximum  $\gamma$  in this topic that were not race science discourse — and interrater reliability using Cohen’s  $\kappa$ . The consensus true positive rate was 89.3% (108/121), while the consensus false positive rate was 8.3% (10/121), with 98% agreement across the two raters ( $\kappa = 0.86$ ). Coding spreadsheets are included in the supplemental materials S8-S9. For each of the 10 consensus false positive documents, their inclusion in the topic was readily explained, but in different ways for different documents. Some included text from other documents, as when an article started in the middle of a page; others discussed perceptions of race relations or racial animus; one piece in BG discussed differences in the herding behavior of two cattle breeds, referred to as “white” and “black.”

The 3 non-consensus documents reflected essential ambiguity in the operationalization of race science discourse. One is a table of contents from a 1980 issue of I; arguably this document should have been excluded as irrelevant front matter, but arguably it legitimizes the research published in the issue. The second case is a paper on racial differences in emotional intelligence that was classified as race science discourse by one author, while the other author felt it was more focused on psychometrics (66). A finally, a 1990 article in MQ (<https://www.unz.com/print/MankindQuarterly-1990q3-00108/>) proposed to offer an economic explanation for crime. While it has a few pages connecting race/ethnicity to both crime and poverty, it ultimately focuses on unemployment and alcohol sales, rather than race, as the key predictors of crime.

Based on this qualitative review, we judged that the topic model approach was sufficiently sensitive (low “false positive” rate) in identifying race science discourse. However, “false negatives” are still possible, as some documents that human readers would classify as race science discourse might have been associated with other topics.

### **S1.5 Supplementary review of *Behavior Genetics***

To check the finding that very few race-and-intelligence documents had been published in BG, we ran an independent search for documents with the keywords “race” and “intelligence” published in BG from 1972 to 2020 using Springer’s journal search website. This search returned 125 documents. Thirty-three were presentation abstracts from meetings of the Behavior Genetics Association; we did not examine these further. Fourteen of the remaining documents had been published since 2010 (after the scope of the primary study). Among these, one obituary and two papers celebrated the work of John Loehlin, a prominent race-and-intelligence researcher who had been Director of the American Eugenics Society from 1968-1972 (67–69). Another was a retrospective of the work of Lindon Eaves, a geneticist who made at least one notable

contribution to debates on the Scarr-Rowe hypothesis (interactions between heritability, race, and class). None of these 14 documents reported any studies of racial differences.

Twenty-eight documents in this sample from BG were published between 1990 and 2010. One was an obituary of Jerry Hirsch, a critic of hereditarianism in general and Jensen in particular (70), and another was an obituary of David Rowe (71). Only 2 of these 28 documents reported race differences of any kind: one examining interactions among race, sex, and heritability for adolescent BMI (Body Mass Index, used as a measure of overweight/obesity) (72); and the other interactions among race, family history of alcoholism, and visuospatial performance (73). Finally, Philip Vernon — one of the Pioneer-funded scientists we identified earlier — published a critical review of a book that attempted to address hereditarianism, the Flynn effect, race-and-intelligence research, and some related issues.

At the same time, we were unable to find any articles published in BG that acknowledged the involvement of the field with scientific racism with anything like the force of the statements that have been made by biological anthropologists (31) and human geneticists (47). Notably, the ASHG report includes discussions of behavior genetics race science, and in particular is critical of the field and the organization for failing to publicly reject the race-and-intelligence claims made by Shockley and Jensen (47). A search for “eugenics” in BG found articles that presented the issue as part of the distant past, and a search for “racism” turned up studies of racist attitudes. A statement on the Behavior Genetics Association website, dated 2021, acknowledges that “The history of our field is inextricably linked with racism, including the misuse of behavior genetic research to support violent eugenic policies,” but primarily focuses on BGA member demographics ([https://www.bga.org/content.aspx?page\\_id=22&club\\_id=971921&module\\_id=567723](https://www.bga.org/content.aspx?page_id=22&club_id=971921&module_id=567723)).

## S1.6 Limitations

A key limitation of the current study stems from the inability of text mining techniques to draw on a broader cultural context than what is represented in the corpus. Scholars of racism and US race relations have noted a shift in the way racist attitudes are understood and expressed, from overt and direct to a subtle “color-blind” racism (74). Color-blind racism represents a change in the primary mechanisms by which white privilege is maintained, using non-racialized language in furtherance of the racial status quo both institutionally and individually, and increasingly reliant on the invisibility of socio-cultural mechanisms to preserve inequalities. Color-blind racism involves the use of racialized “codes” or “dogwhistles,” language that is superficially non-racial but carries racial implications, such as (in the US context) “violent inner city criminals” or “welfare queen” (75). Alderfer notes that the racist implications of *The Bell Curve* only emerge gradually, with the opening of the book using the color-blind language of “group differences” (76). Panofsky et al. show how the term “human biodiversity,” originally developed for anti-racist purposes by biological anthropologist Jonathan Marks, has been inverted by white supremacists and is now a dogwhistle for biological racial hierarchy (3). These cases suggest further that color-blindness can facilitate white supremacist appropriation of non-racial research — even contrary to the researchers’ own intentions — as scientific racist readers re-interpret racial-neutral language as dogwhistles (39, 77).

Techniques such as topic modeling may be able to account for discursive shifts from explicit to coded racial language, but only if the corpus contains a critical mass of documents using both sets of terms together. This would seem to require either a relatively gradual transition, the inclusion of reflective commentary — such as scholarship identifying and analyzing the transition — or both. It is unlikely that these requirements are met by the corpus used in this study. So it is plausible that the topic models used here have some rate of “false negatives,” documents in the corpus that engage in race science discourse but in a color-blind way that

avoids the use of the explicit racial terms in the race-and-intelligence topic.

The phenomenon of color-blind racism is closely related to ambiguities that cannot be resolved even by human coders familiar with the broader cultural context. Two relevant examples are the “Flynn effect,” a secular increase in intelligence test scores noted by philosopher James Flynn; and studies of “national IQ,” exemplified by the work of Richard Lynn. The Flynn effect implies that environmental factors can create group differences in IQ that are comparable to Black-White differences, and thus directly challenges claims that the racial differences are biological. Using methods widely regarded as ad hoc, cherry-picking, and generally unreliable, Lynn and collaborators have claimed that national IQ averages are correlated with national GDP. Both the Flynn effect and “national IQ” are strongly associated with race-and-intelligence discourse in the US cultural context. But scientific journal articles on these topics may not include any racial language at all. For example, a meta-analysis of the Flynn effect does not discuss racial differences or use terms such as “White” or “Black” outside the reference list (78). Should such articles be coded as race science discourse? We expect that even readers who accept our definitions of race science and race science discourse will reasonably disagree on how to answer this question. Topic models, of course, are not even capable of representing this essential ambiguity.

79. Thanks to Emily Merchant and John Jackson for some initial discussions that helped clarify the scope of the project and identify some key resources related to the Pioneer Fund and *Mankind Quarterly*. Thanks to Anthony Sainez for help retrieving the *Mankind Quarterly* articles. Thanks to Derek Devnich and James Dooley for their work attempting to secure the APA-published articles.

**Funding:** EL’s work on this project was partially supported by UC Merced. DH’s work on this project was not supported by any specific funding.

**Author contributions:** DJH (conceptualization, data curation, formal analysis, investigation,

methodology, project administration, software, validation, visualization, writing-original draft, writing-review & editing); EL (conceptualization, methodology, software, validation, visualization, writing-original draft, writing-review & editing)

**Competing interests:** The authors have no competing interests to declare.

**Data and materials availability:** Web of Science search results and article fulltext are not publicly available due to intellectual property restrictions. Preprocessed analysis data (document-term counts and article metadata) and analysis scripts are available at *[doi]*.