

Mainstreaming Race Science

Daniel J. Hicks,^{1*} and Emilio J. C. Lobato¹

¹University of California, Merced

*Corresponding author: Daniel J. Hicks, 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.

16

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

²³ community to reject attempts to scientifically justify racial inequalities and colonialism (*I*). But
²⁴ eugenic thinking and scientific racism persisted throughout the second half of the twentieth
²⁵ century, and recent work has found that white supremacists continue to attempt to use scientific
²⁶ research to justify racial violence (*2–4*).

²⁷ 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.

³⁰ 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.

⁴³ 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 (*5*). But it is important that much of Jensen's research was useful and indeed utilized to purportedly justify racial inequality.

⁴⁸ **Scientific racism, eugenics, and *Brown***

⁴⁹ Histories of scientific racism in the twentieth century have often emphasized the Pioneer Fund
⁵⁰ and the parascholarly journal *Mankind Quarterly* (MQ) (1, 6–13). Pioneer was formed in 1937,
⁵¹ during the waning years of the eugenics movement in North America (1), with the aim of “sup-
⁵² port[ing] academic research and the ’[sic]dissemination of information, into the ‘problem of
⁵³ heredity and eugenics’ and ‘the problems of race betterment’ ” (6, p.21, quoting Laughlin). MQ
⁵⁴ was founded in 1960 by biologist R. Ruggles Gates (1882–1962), psychologist Henry Garrett
⁵⁵ (1894–1973), and non-academic anthropologist G. Robert Gayre (self-styled as “Gayre of Gayre
⁵⁶ and Nigg”; 1907–1996). Gates’ professional status had risen and fallen with eugenics and the
⁵⁷ explicit scientific racism of the 1920s, and by the end of the Second World War he was thor-
⁵⁸oughly marginalized (7). In contrast, Garrett had been president of the American Psychological
⁵⁹ Association in 1946 and chair of Psychology at Columbia from 1941 to 1955 (7).

⁶⁰ In the landmark case *Brown v Board of Education of Topeka* (1954), the US Supreme Court
⁶¹ banned *de jure* educational segregation. The Court’s decision relied on expert testimony from
⁶² psychologists and education researchers; but the segregationists also put forward their own
⁶³ experts, including Henry Garrett (7, 9, 14). As part of this segregationist reaction, MQ was
⁶⁴ created to provide a favorable venue for race scientists to publish their views, on the grounds
⁶⁵ that an “equalitarian dogma” created a censorious “taboo” against their research in mainstream
⁶⁶ publications (8, 15).

⁶⁷ In light of the scientific racist origins of Pioneer and MQ, and the significant attention MQ
⁶⁸ has received in recent historiography, we hypothesized that bibliometric and text mining anal-
⁶⁹yses would show race science ideas originating in MQ, and from there being disseminated to
⁷⁰ mainstream publications. This hypothesis was not supported. Instead, mainstream psychology
⁷¹ journals provided a venue for a distinct form of race science, based on intelligence research and

⁷² originating outside of MQ.

⁷³ ***Mankind Quarterly* and Pioneer-funded researchers**

⁷⁴ We identified 16 researchers who had received funding from Pioneer; 14 of these researchers
⁷⁵ had profiles in the Web of Science (WoS) author search, allowing us identify 13 WoS-indexed
⁷⁶ journals that had published 6 or more of these authors. See Table 1 and Figure 1.

Table 1: Pioneer-funded researchers. Either identified in (16)
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 (1930-2023)	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

⁷⁷ Figure 1 shows that, while MQ is among the “Pioneer-publishing” journals, a number of
⁷⁸ mainstream journals are more prominent: *Personality and Individual Differences* (PID), *Intell-*
⁷⁹ *ligence* (Int), *Behavior Genetics* (BG), and *Psychological Reports* (PR). In addition, only psy-
⁸⁰ chologist Richard Lynn appears to have published heavily in MQ. Lynn became an assistant

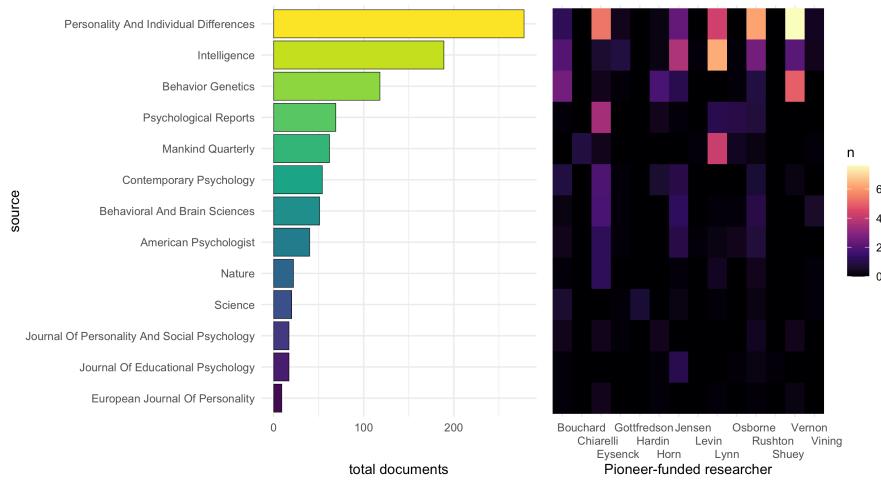


Figure 1: Journals publishing 6 or more Pioneer-funded researchers, Web of Science author search results.

81 editor of MQ in 1979 (vol. XX, Nos. 1 & 2) and is listed as editor-in-chief on MQ's current
 82 website as of 2023-07-21. He has also been president of Pioneer since the death of psychologist
 83 J. Phillippe Rushton in 2012 (17). By contrast, a number of Pioneer-funded researchers have
 84 published in PID, Int, and to a lesser degree BG: Bouchard, Eysenck, Jensen, Rushton, Vernon,
 85 and also Lynn.

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

93 *Intelligence* (Int) was founded in 1977, with psychologist Douglas Detterman as editor-in-
 94 chief from the founding until 2016. Lloyd Humphreys was on the editorial board starting from
 95 1977; by 1990 he had been joined by Jensen and Philip Vernon. Richard Lynn joined the edi-

⁹⁶ torial board sometime between 1998 and 2002. (Archive copies of the Int editorial board page
⁹⁷ are not available from the journal's website from 1999 through 2001.) Int has been criticized
⁹⁸ for including Lynn and Gerhard Meisenberg — who was editor-in-chief of MQ in 2015-18 —
⁹⁹ on its editorial board until 2018 (10). Int is published by Elsevier.

¹⁰⁰ **Topic model analysis identifies race science discourse**

¹⁰¹ The fact that Pioneer-funded researchers published heavily in two mainstream psychology jour-
¹⁰² nals does not tell us anything about the content of their publications or the claims of their
¹⁰³ research. We assembled a full-text corpus of articles published in MQ, PID, Int, BG, PR, and
¹⁰⁴ Behavior and Brain Sciences (BBS) between 1960-2010 and used topic modeling to identify
¹⁰⁵ topics discussing race, intelligence, and both.

¹⁰⁶ Figure 2 and this discussion focuses on three topics identified in one of the 24 fitted topic
¹⁰⁷ models; see figures S1.3, S1.4, S1.5 and supplemental visualizations S2-S4 for all topics from
¹⁰⁸ all models and S5-S7 for selected topics from all models. In the focal model, topic 07 is strongly
¹⁰⁹ associated with MQ: MQ published dozens of articles in this topic each year, and no other
¹¹⁰ journal ever published more than a handful. The “Silge plot,” showing the top 15 terms in
¹¹¹ the topic (18), contains racial terms (races, whites, negroes and potentially europe,
¹¹² africa, india, and japan) as well as book, likely reflecting the fact that MQ published a
¹¹³ number of book reviews, while the psychology journals either did not or these were not available
¹¹⁴ for the corpus.

¹¹⁵ Topic 22 is strongly associated with Int and PID in the same way that 07 is associated with
¹¹⁶ MQ. The Silge plot does contain jensen, as well as a reference to Raymond Cattell, who
¹¹⁷ played a major role in the development of factor analysis and intelligence testing but also ad-
¹¹⁸ vocated for eugenics, fascism, and Nazi race science (19). However, the other authors named
¹¹⁹ in this topic — John Horn and Peter Bentler — do not appear to have contributed to scientific

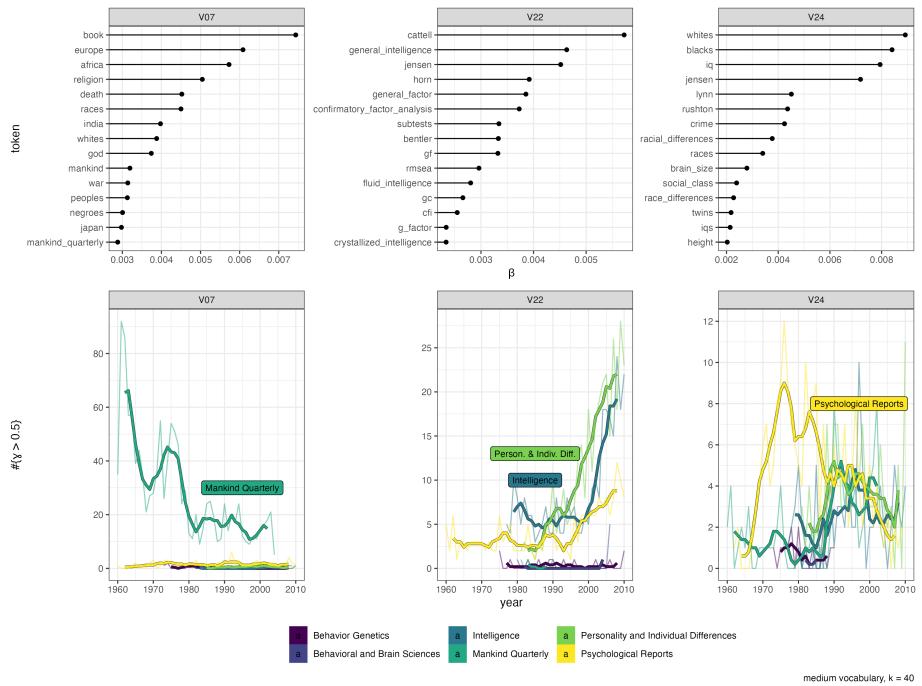


Figure 2: Silge plots (18) and smoothed time series for three focal topics. Top row: top 15 terms by β (term-topic distribution) for each topic. Bottom row: count of articles associated with each topic, by journal and year. Article counts use a threshold approach, with γ (topic-document distribution) 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.

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?” (20) was published in 1969 in *Harvard Educational Review* (not included in this study), and the time series indicates that, in the early 1970s, there was an increase in articles in topic 24 in both PR and MQ (the only two journals in our corpus that were active at the time). MQ shows another sharp increase in the late 1980s; using content analysis, Adams and Pilloud found that psychology was the dominant discipline in MQ in the period 1992-2018 (12). PID published multiple papers in topic 24 almost immediately after it was founded, with Int 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.

¹⁴⁴ Panofsky (21) argues that, prior to Jensen's 1969 paper, behavior genetics emphasized the study
¹⁴⁵ of non-human animals and intentionally avoided associations with eugenics and public contro-
¹⁴⁶ versy more generally. In response to Jensen, critics such as Lewontin offered broad critiques
¹⁴⁷ of behavior genetics as such, and behavior geneticists in turn adopted a radical conception of
¹⁴⁸ academic freedom (without any sense of responsibility for the social implications of academic
¹⁴⁹ research) and a siege or wartime mentality, as illustrated by Sandra Scarr's 1986 presidential
¹⁵⁰ address to the Behavior Genetics Association (BGA) (22). Scarr's address coincided with the
¹⁵¹ period between 1970-1990 when BG published articles in topic 24 (including Scarr's address
¹⁵² itself).

¹⁵³ However, the topic model analysis suggests that, by the 1990s, behavior genetics may have
¹⁵⁴ distanced itself somewhat from race science, albeit without directly repudiating it. This inter-
¹⁵⁵ pretation was supported by a supplemental analysis that focused on BG specifically (SM section
¹⁵⁶ 1.5). In addition, the 1995 BGA presidential address by Glayde Whitney — in which he criti-
¹⁵⁷ cized the “Marxist-Lysenkoist denial of genetics” and proposed that differences in murder rates
¹⁵⁸ between countries and cities were caused by racial genetic differences in intelligence, empathy,
¹⁵⁹ aggression, and impulsivity — was published in MQ rather than BG. As a subdiscipline of psy-
¹⁶⁰ chology, behavior genetics may have been prominent in promoting race science in the past, but
¹⁶¹ appears to have been less receptive to such ideas in more recent history.

¹⁶² Instead, our results suggest that other subdisciplines of psychology have provided the pri-
¹⁶³ mary venue for mainstreaming race science. Over the past few decades, many professional
¹⁶⁴ organizations in genetics and anthropology have made formal statements rejecting race as a bi-
¹⁶⁵ logically meaningful concept (23–30), undermining the core assumption of race science. The
¹⁶⁶ American Psychological Association (APA) — whose members are predominantly clinical psy-
¹⁶⁷ chologists — and the Federation of Associations in Behavioral & Brain Sciences (FABBS) —
¹⁶⁸ an organization whose purpose is to provide policy recommendations — have openly rejected

¹⁶⁹ race as biologically meaningful (31, 32). But research-focused psychological organizations like
¹⁷⁰ the Association for Psychological Science (APS) and Psychonomic Society (PS) stop short of
¹⁷¹ rejecting race as biologically meaningful when they denounce racism (33, 34). Psychology as
¹⁷² a discipline maintains space for several mythological race science narratives, such as allega-
¹⁷³ tions of a taboo against race and intelligence research (35), claims that scientists engaging in
¹⁷⁴ or calling for any kind of anti-racism in the field is an ideological corruption of dispassionate
¹⁷⁵ and value-neutral science (36), or arguments that holding race science to the same evidentiary
¹⁷⁶ standards as other psychological research is a violation of academic freedom (37). And histor-
¹⁷⁷ ically several race scientists have secured important gatekeeping positions within the scientific
¹⁷⁸ community, such as on the editorial boards of Int and PID.

¹⁷⁹ Discussion

¹⁸⁰ Using bibliometric and text-mining methods, this study finds that Pioneer-funded researchers
¹⁸¹ published heavily in certain mainstream psychology journals, much more than in *Mankind*
¹⁸² *Quarterly*; and that a distinct topic of race science discourse, centered on race and intelligence,
¹⁸³ can be identified in these same mainstream journals across the period 1960-2010. These find-
¹⁸⁴ ings indicate that *Mankind Quarterly* (MQ) was less important as venue for late 20th century
¹⁸⁵ race science than mainstream psychology journals, especially *Intelligence* (Int) and *Personality*
¹⁸⁶ and *Individual Differences* (PID). Indeed, until the 1990s, MQ published a very different kind
¹⁸⁷ of race science from the race-and-intelligence research published in Int and PID. During this
¹⁸⁸ same period of time, race scientists served in prominent and influential positions in behavior ge-
¹⁸⁹ netics and psychology, including as society presidents and members of journal editorial boards.
¹⁹⁰ In some cases, some race scientists simultaneously maintained active connections to both the
¹⁹¹ mainstream scientific community and white supremacist organizations (5, 11, 38). Except for
¹⁹² the APA, we are not aware of other major organizations for scientific research in psychology

193 that have issued apologies for the historical contributions of their field to scientific racism.

194 We are not identifying any particular scientist as a racist, scientific racist, or race scientist
195 based on the topic model results alone. Such claims require an analysis of documentary evi-
196 dence that goes beyond the scope of the current study (5). We do assume that our readers, like
197 us, regard white supremacy and scientific racism as morally odious and beyond the scope of
198 reasonable debate (39). However, scientific research can be appropriated to promote scientific
199 racism — and thus count as race science — even when this is contrary to the intentions of
200 the original researchers themselves (40–43). Panofsky et al. show how the term “human bio-
201 diversity,” originally developed for anti-racist purposes by biological anthropologist Jonathan
202 Marks, has been inverted by white supremacists and is now a dogwhistle for biological racial
203 hierarchy (3).

204 Research on the genetics of intelligence has been socially harmful, not just in the somewhat
205 abstract sense of promoting racial stigma, but in the concrete sense of being used to rationalize
206 mass shootings and other acts of racial violence (4, 44). While scientists do not fully control
207 the downstream social effects of their research, like other citizens scientists are responsible for
208 mitigating reasonably foreseeable harms that result from their actions (45–48). These actions
209 include not only individual decisions to research this topic or that, but also collective decisions
210 about who is awarded positions of power or influence. Given the way research on intelligence
211 and behavior genetics has been used historically, these fields may be especially susceptible to
212 appropriation by scientific racists. This only strengthens the obligations of researchers in these
213 areas to prevent their research from causing harm (42, 49).

214 How should psychologists exercise these responsibilities? We believe restorative or repara-
215 tive justice provides a useful model (50, 51). Unlike retributive justice (used by almost all crim-
216 inal justice systems today), the aim of restorative justice is to repair the damaged relationships
217 between victims and perpetrators of injustice. Typically, restorative justice requires perpetrators

²¹⁸ to not only acknowledge their actions and the harms that these actions had on victims, but also
²¹⁹ to work with victims to identify concrete actions that perpetrators can take to redress or mitigate
²²⁰ these harms. From this perspective, explicit apologies for contributions to scientific racism are
²²¹ essential as a first step (27, 31, 52), but not fully sufficient on their own.

²²² **References and Notes**

²²³ **References and Notes**

- ²²⁴ 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).
- ²²⁵ 2. T. Duster, *Backdoor to Eugenics* (Routledge, 2003).
- ²²⁶ 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).
- ²²⁷ 4. M. Pronczuk, K. Ryckewaert, A Racist Researcher, Exposed by a Mass Shooting, *The New York Times* (2022).
- ²²⁸ 5. J. P. Jackson Jr., Arthur Jensen, evolutionary biology, and racism, *History of Psychology* pp. No Pagination Specified–No Pagination Specified (2022).
- ²²⁹ 6. B. Mehler, Foundation for fascism: The new eugenics movement in the United States, *Patterns of Prejudice* **23**, 17 (1989).
- ²³⁰ 7. 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).
- ²³¹ 8. W. H. Tucker, *The Funding of Scientific Racism: Wickliffe Draper and the Pioneer Fund* (University of Illinois Press, Urbana, 2002).
- ²³² 9. 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).

- ²⁴³ 10. A. Saini, *Superior: The Return of Race Science* (Beacon Press, 2019).
- ²⁴⁴ 11. A. S. Winston, *Oxford Research Encyclopedia of Psychology* (Oxford University Press, 2020).
- ²⁴⁶ 12. D. Adams, M. Pilloud, The (Mis)appropriation of Biological Anthropology in Race Science and the Implications for Forensic Anthropology, *FA* (2021).
- ²⁴⁸ 13. A. Saini, Draper's Millions: The Philanthropic Wellspring of Modern Race Science, 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).
- ²⁵³ 14. J. P. Jackson, Jr., *Science for Segregation: Race, Law, and the Case Against Brown V. Board of Education* (NYU Press, 2005).
- ²⁵⁵ 15. J. P. Jackson, Jr., A. S. Winston, The Mythical Taboo on Race and Intelligence, *Review of General Psychology* (2020).
- ²⁵⁷ 16. A. Miller, The Pioneer Fund: Bankrolling the Professors of Hate, *The Journal of Blacks in Higher Education* pp. 58–61 (1994).
- ²⁵⁹ 17. H. Beirich, Pioneer Fund Assets Divided; New Leadership Appointed, <https://www.splcenter.org/hatewatch/2013/10/22/pioneer-fund-assets-divided-new-leadership-appointed> (2013).
- ²⁶² 18. *Topic Modeling* ("O'Reilly Media, Inc.", 2017), p. ch. 6.
- ²⁶³ 19. B. Mehler, Beyondism: Raymond B. Cattell and the new eugenics, *Genetica* **99**, 153 (1997).

- 265 20. A. Jensen, How Much Can We Boost IQ and Scholastic Achievement?, *Harvard Educa-*
266 *tional Review* **39**, 1 (1969).
- 267 21. A. Panofsky, *Misbehaving Science: Controversy and the Development of Behavior Genetics* (University of Chicago Press, Chicago, IL, 2014).
- 268
269 22. S. Scarr, Three cheers for behavior genetics: Winning the war and losing our identity, *Behav*
270 *Genet* **17**, 219 (1987).
- 271 23. O. Anthropology, Aapa statement on biological aspects of race, *American Journal of Phys-*
272 *ical Anthropology* **101**, 569 (1996).
- 273 24. A. S. of Human Genetics, Ashg denounces attempts to link genetics and racial supremacy,
274 *American Journal of Human Genetics* **103**, 636 (2018).
- 275 25. A. E. Board, Aaa statement on race, *American Anthropologist* **100**, 712 (1998).
- 276 26. A. of Social Anthropologists of the UK, Statement on racism by the committee of the ASA,
277 https://www.theasa.org/stmt_racism (2020).
- 278 27. A. Fuentes, R. R. Ackermann, S. Athreya, D. Bolnick, T. Lasisi, S.-H. Lee, S.-A. McLean,
279 R. Nelson, AAPA Statement on Race and Racism, *American Journal of Physical Anthro-*
280 *pology* **169**, 400 (2019).
- 281 28. N. H. G. R. Institute, Race, [https://www.genome.gov/genetics-glossary/
282 Race](https://www.genome.gov/genetics-glossary/Race) (2023).
- 283 29. C. Rotimi, Ashg statement regarding the warping of genetic knowledge to feed
284 racist ideology, [https://www.ashg.org/publications-news/ashg-news/
285 ashg-statement-regarding-the-warping-of-genetic-knowledge-to-feed-racism](https://www.ashg.org/publications-news/ashg-news/ashg-statement-regarding-the-warping-of-genetic-knowledge-to-feed-racism)
286 (2022).

- 287 30. A. Wynshaw-Boris, American society of human genetics statement
288 regarding concepts of good genes and human genetics,
289 <https://www.ashg.org/publications-news/ashg-news/statement-regarding-good-genes-human-genetics/> (2020).
- 291 31. A. P. Association, Apology to people of color for apas role in promoting, perpetuating,
292 and failing to challenge racism, racial discrimination, and human hierarchy in us, <https://www.apa.org/about/policy/racism-apology> (2021).
- 294 32. J. Baron, Re: Omb-2023-0001, "initial proposals for updating omb's race and ethnicity
295 statistical standards", <https://fabbs.org/wp-content/uploads/2022/11/OMB-2023-0001-RFI-FABBS-Response-Submitted-1.pdf> (2023).
- 297 33. A. for Psychological Science, Statement on racism, bias, and intolerance,
298 <https://www.psychologicalscience.org/news/statement-on-racism-bias-and-intolerance.html> (2021).
- 300 34. P. Society, Statement on racism, <https://www.psychonomic.org/page/antiracismstatement> (2020).
- 302 35. J. P. Jackson Jr, A. S. Winston, The mythical taboo on race and intelligence, *Review of
303 General Psychology* **25**, 3 (2021).
- 304 36. S. Roberts, Dealing with diversity in psychology: Science and ideology (2022).
- 305 37. J. M. Herbert, Academic free speech or right-wing grievance?, *Digital Discovery* **2**, 260
306 (2023).
- 307 38. Q. Slobodian, The Unequal Mind: How Charles Murray and Neoliberal Think Tanks Re-
308 vived IQ, *Capitalism: A Journal of History and Economics* **4**, 73 (2023).

- 309 39. S. A. Schroeder, The Limits of Democratizing Science: When Scientists Should Ignore the
310 Public, *Philosophy of Science* **89**, 1034 (2022).
- 311 40. J. Tabery, Why Is Studying the Genetics of Intelligence So Controversial?, *Hastings Center
312 Report* **45**, S9 (2015).
- 313 41. D. Gillborn, Softly, softly: Genetics, intelligence and the hidden racism of the new geneism,
314 *Journal of Education Policy* **31**, 365 (2016).
- 315 42. J. Carlson, K. Harris, Quantifying and contextualizing the impact of bioRxiv preprints
316 through automated social media audience segmentation, *PLOS Biology* **18**, e3000860
317 (2020).
- 318 43. B. M. Henn, E. K. Merchant, A. O'Connor, T. Rulli, Why DNA Is No
319 Key to Social Equality: On Kathryn Paige Harden's "The Genetic Lottery",
320 <https://www.lareviewofbooks.org/article/why-dna-is-no-key-to-social-equality-on-kathryn-paige-harden-the-genetic-lottery/> (2021).
- 322 44. 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 45. N. J. Block, G. Dworkin, IQ, Heritability and Inequality, Part 2, *Philosophy & Public
328 Affairs* **4**, 40 (1974).
- 329 46. H. E. Douglas, *Science, Policy, and the Value-Free Ideal* (University of Pittsburgh Press,
330 Pittsburgh, Pa, 2009).

- ³³¹ 47. P. Kitcher, An Argument About Free Inquiry, *Noûs* **31**, 279 (1997).
- ³³² 48. J. A. Kourany, Should Some Knowledge Be Forbidden? The Case of Cognitive Differences
³³³ Research, *Philosophy of Science* **83**, 779 (2016).
- ³³⁴ 49. C. B. Ogbunugafor, DNA, basketball, and birthday luck. A review of The genetic lottery:
³³⁵ Why DNA matters for social equality. By Kathryn Paige Harden, 2021, US: Princeton Uni-
³³⁶ versity Press. 312 pp. ISBN: 9780691190808 (Hardcover), *American Journal of Biological*
³³⁷ *Anthropology* **179**, 501 (2022).
- ³³⁸ 50. M. Wenzel, T. G. Okimoto, N. T. Feather, M. J. Platow, Retributive and Restorative Justice,
³³⁹ *Law Hum Behav* **32**, 375 (2008).
- ³⁴⁰ 51. K. van Wormer, *Encyclopedia of Social Work* (2013).
- ³⁴¹ 52. C. S. Jackson, D. Turner, M. June, M. V. Miller, Facing Our History—Building an Equitable
³⁴² Future, *The American Journal of Human Genetics* **110**, 377 (2023).
- ³⁴³ 53. J. Ooms, PdfTools: Text Extraction, Rendering and Converting of PDF Documents (2023).
- ³⁴⁴ 54. S. Chamberlain, H. Zhu, N. Jahn, C. Boettiger, K. Ram, Rcrossref: Client for Various
³⁴⁵ 'CrossRef' 'APIs' (2019).
- ³⁴⁶ 55. J. R. Flynn, Evidence against Rushton: The genetic loading of WISC-R subtests and the
³⁴⁷ causes of between-group IQ differences, *Personality and Individual Differences* **26**, 373
³⁴⁸ (1998).
- ³⁴⁹ 56. J. P. Rushton, Secular gains in IQ not related to the g factor and inbreeding depression –
³⁵⁰ unlike Black-White differences: A reply to Flynn, *Personality and Individual Differences*
³⁵¹ **26**, 381 (1998).

- 352 57. J. R. Flynn, Reply to Rushton: A gang of gs overpowers factor analysis, *Personality and*
353 *Individual Differences* **26**, 391 (1998).
- 354 58. H. Wickham, J. Hester, J. Ooms, RStudio, R. F. C. o. R.-p. homepage cached as example),
355 Xml2: Parse XML (2023).
- 356 59. spaCy: Industrial-strength Natural Language Processing (NLP) with Python and Cython,
357 Explosion AI (2018).
- 358 60. K. Benoit, A. Matsuo, E. R. Council (ERC-2011-StG 283794-QUANTESS), Spacyr: Wrap-
359 per to the 'spaCy' 'NLP' Library (2020).
- 360 61. D. Vohra, *Practical Hadoop Ecosystem: A Definitive Guide to Hadoop-Related Frame-*
361 *works and Tools*, D. Vohra, ed. (Apress, Berkeley, CA, 2016), pp. 325–335.
- 362 62. N. Richardson, I. Cook, N. Crane, D. Dunnington, R. François, J. Keane, D. Moldovan-
363 Grünfeld, J. Ooms, J. Luraschi, K. D. Werner, J. Wong, A. Arrow, Arrow: Integration to
364 'Apache' 'Arrow' (2023).
- 365 63. A. Sclafani, Correlation and causation in the study of feeding behavior, *Behavioral and*
366 *Brain Sciences* **4**, 590 (1981).
- 367 64. D. J. Hicks, Productivity and interdisciplinary impacts of Organized Research Units, *Quan-*
368 *titative Science Studies* **2**, 990 (2021).
- 369 65. K. Rohe, M. Zeng, Vintage Factor Analysis with Varimax Performs Statistical Inference
370 (2020).
- 371 66. M. E. Roberts, B. M. Stewart, D. Tingley, Stm: An R Package for Structural Topic Models,
372 *Journal of Statistical Software* **91**, 1 (2019).

- 373 67. D. J. Hicks, Tmfast fits topic models fast (2023).
- 374 68. G. E. Gignac, G. Ekermans, Group differences in EI within a sample of black and white
375 South Africans, *Personality and Individual Differences* **49**, 639 (2010).
- 376 69. John C. Loehlin (1926–2020), *Behav Genet* **50**, 385 (2020).
- 377 70. I. Waldman, E. Turkheimer, Introduction to a Festschrift for John Loehlin, *Behav Genet* **44**,
378 547 (2014).
- 379 71. R. Plomin, Genotype-Environment Correlation in the Era of DNA, *Behav Genet* **44**, 629
380 (2014).
- 381 72. P. L. Roubertoux, Jerry Hirsch (20 September 1922–3 May 2008): A Tribute, *Behav Genet*
382 **38**, 561 (2008).
- 383 73. J. L. Rodgers, K. Jacobson, E. van den Oord, Obituary: David Christian Rowe, *Behav
384 Genet* **33**, 627 (2003).
- 385 74. K. C. Jacobson, D. C. Rowe, Genetic and Shared Environmental Influences on Adolescent
386 BMI: Interactions with Race and Sex, *Behav Genet* **28**, 265 (1998).
- 387 75. S. M. Berman, E. P. Noble, Reduced visuospatial performance in children with the D2
388 dopamine receptor A1 allele, *Behav Genet* **25**, 45 (1995).
- 389 76. E. Bonilla-Silva, *Racism without Racists: Color-Blind Racism and the Persistence of Racial
390 Inequality in the United States* (Rowman & Littlefield Publishers, 2006).
- 391 77. J. Saul, *New Work on Speech Acts*, D. Fogal, D. W. Harris, M. Moss, eds. (Oxford Univer-
392 sity Press, 2018), p. 0.

- 393 78. C. P. Alderfer, The science and nonscience of psychologists' responses to The Bell Curve,
394 *Professional Psychology: Research and Practice* **34**, 287 (2003).
- 395 79. M. Wills, Are Clusters Races? A Discussion of the Rhetorical Appropriation of Rosen-
396 berg et al.'s "Genetic Structure of Human Populations", *Philosophy & Theory in Biology* **9**
397 (2017).
- 398 80. L. Trahan, K. K. Stuebing, M. K. Hiscock, J. M. Fletcher, The Flynn Effect: A Meta-
399 analysis, *Psychol Bull* **140**, 1332 (2014).

400

401 Thanks to Emily Merchant and John Jackson for some initial discussions that helped clar-
402 ify the scope of the project and identify some key resources related to the Pioneer Fund and
403 *Mankind Quarterly*. Thanks to Anthony Sainez for help retrieving the *Mankind Quarterly* ar-
404 ticles. Thanks to Derek Devnich and James Dooley for their work attempting to secure the
405 APA-published articles. Thanks to Kevin Bird, John Jackson, and Manuela Fernández Pinto for
406 comments on a draft of this paper.

407 **Funding:** EJCL's work on this project was partially supported by UC Merced. DJH's work on
408 this project was not supported by any specific funding.

409 **Author contributions:** DJH (conceptualization, data curation, formal analysis, investigation,
410 methodology, project administration, software, validation, visualization, writing-original draft,
411 writing-review & editing); EJCL (conceptualization, methodology, software, validation, visual-
412 ization, writing-original draft, writing-review & editing)

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

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

⁴¹⁷ **Supplemental Materials**

⁴¹⁸ S1. Materials and methods

⁴¹⁹ Figures S2-S7

⁴²⁰ Data S8, S9

⁴²¹ **S1 Materials and methods**

⁴²² **S1.1 Corpus assembly**

⁴²³ Corpus assembly is summarized in figure S1.1. The corpus was assembled in two parts, one for
⁴²⁴ *Mankind Quarterly* and the other for mainstream journals.

⁴²⁵ **S1.1.1 Mankind Quarterly**

⁴²⁶ An electronic archive of *Mankind Quarterly*, from the first issue in 1969 through 2004, is avail-
⁴²⁷ able for free on the open web at the white nationalist website *The Unz Review*. The Python
⁴²⁸ package Beautiful Soup (<https://www.crummy.com/software/BeautifulSoup/>,
⁴²⁹ version not recorded) was used to retrieve every PDF available in this archive. After identifying
⁴³⁰ some gaps (missing issues) in the archive from *The Unz Review*, we manually retrieved PDFs
⁴³¹ for further articles using ProQuest. Text was extracted using the R package *pdftools* (53, ver-
⁴³² sion 3.0.1). This stage of corpus assembly was conducted in Fall 2021 with the assistance of
⁴³³ Anthony Sainez.

⁴³⁴ **S1.1.2 Mainstream journals**

⁴³⁵ To identify suitable mainstream journals for inclusion in the corpus, we first identified academic
⁴³⁶ researchers who had been funded by the Pioneer Fund. Pioneer is an American non-profit or-
⁴³⁷ ganization founded in 1937 to fund eugenics research and propaganda (8). After the *Brown*
⁴³⁸ *v. Board of Education* ruling in 1954, Pioneer funded various segregationist efforts across
⁴³⁹ the United States, including lectures on eugenics by Stanford physicist William Shockley (14).
⁴⁴⁰ We reviewed a critical profile of Pioneer (16) as well as an archived page from the organi-
⁴⁴¹ zation's own web site (<https://web.archive.org/web/20130103005545/http://www.pioneerfund.org/Grantees.html>), which together listed 16 researchers who
⁴⁴² had received Pioneer funds.
⁴⁴³

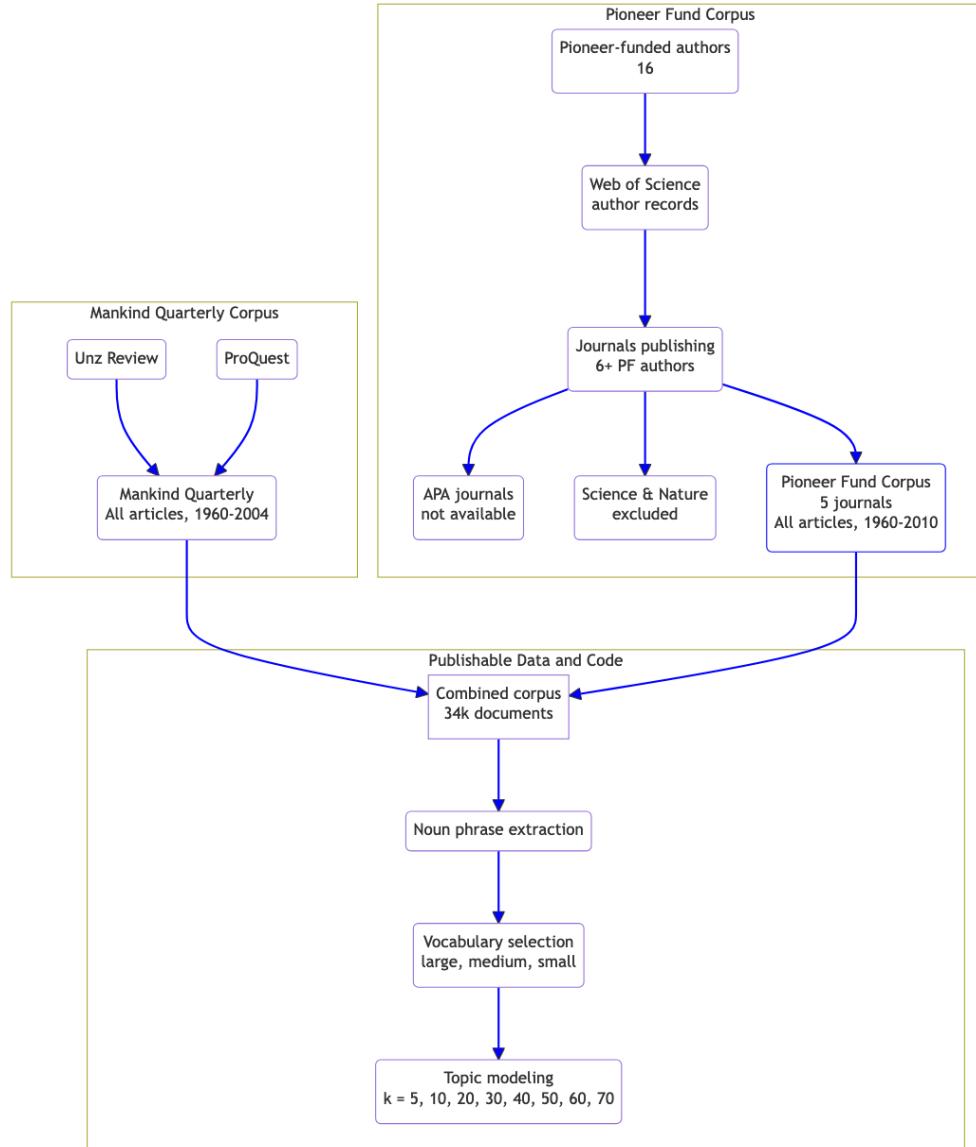


Figure S1.1: Corpus assembly.

444 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 re-
445
446 searchers. These searches were conducted between 2021-09-24 and 2021-10-05 by DJH. After
447 parsing these results, we counted how many of the 14 researchers had published in each journal.
448 Thirteen journals had published 6 or more of the 14 Pioneer-funded researchers. We excluded
449 *Mankind Quarterly* (as already included) as well as *Science* and *Nature* (as too general) from
450 further consideration in this side of the corpus. Three journals published by the American Psy-
451 chological Association (APA; *American Psychologist*, *Contemporary Psychology*, and *Journal*
452 *of Educational Psychology*) had to be excluded due to confusion over who could give us per-
453 mission to use the archives for a text mining project, with both APA and ProQuest asserting that
454 we needed to get permission from the other entity. *European Journal of Personality*, published
455 by SAGE, also had to be excluded because our institutional access only went back to 1999.

456 The remaining 5 journals are all published by major academic publishers — Elsevier, Springer,
457 or Cambridge University Press — and each item in the entire run of each journal has been as-
458 signed a DOI (digital object identifier) for archival purposes. We used the Crossref API and
459 `rcrossref` R interface to this API (54, *version 1.1.0.99*) to retrieve metadata for each item
460 published from 1960-2010 in each of these journals. These metadata included item-level license
461 information — confirming that the text of each item could be used for text mining projects —
462 and a URL to an electronic version of the item. These URLs were used to retrieve a XML
463 or PDF version of each item, except for *Personality and Individual Differences*. This journal
464 has published a relatively large number of non-article documents, such as book reviews and
465 commentaries, that are not available at the URL included in the Crossref metadata. (This is
466 unfortunate, as it was not difficult to find highly relevant documents that we could not automati-
467 cally retrieve and therefore had to be excluded from our corpus. One set of examples is an
468 exchange between Rushton and Flynn on the Flynn effect (55–57).) Instead we used Elsevier's

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

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

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

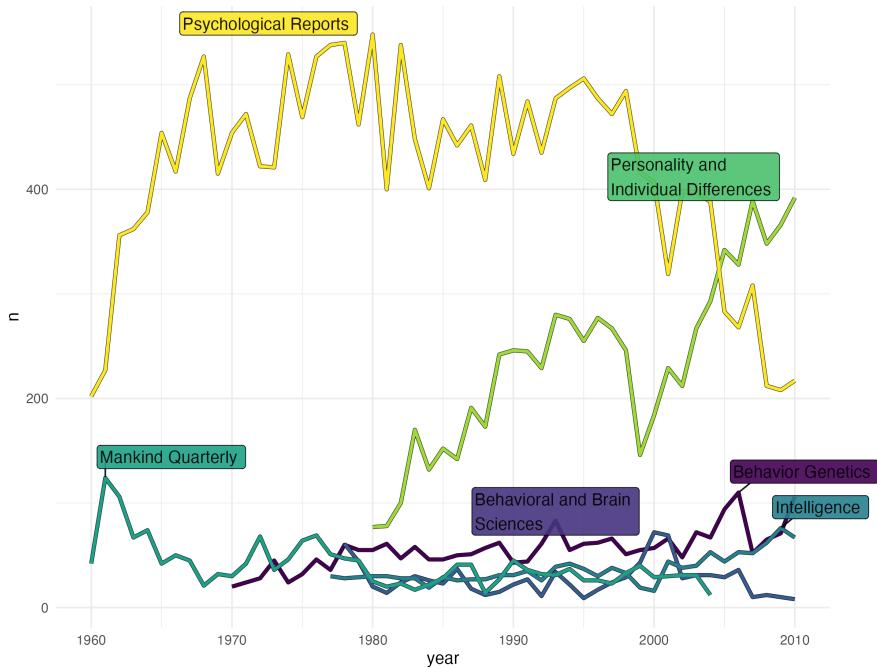


Figure S1.2: Count of documents in the corpus, by journal and year.

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

494 S1.2 Text preparation

495 After document retrieval and text extraction, we pre-processed the text using the spaCy NLP
 496 (natural language processing) Python library (59, *version 1.9.0*) and the R API spacyr (60, *ver-
 497 sion 1.2.1*). Specifically, we applied regular expressions to remove header/footer copyright no-

498 tices and hyphenation, used spaCy to annotate and extract noun phrases (eg, “the intelligence
499 test items”), and then cleaned and standardized these phrases (eg, removing the/an/a, converting
500 all text to lowercase, and replacing all whitespace with underscores: “intelligence_test_items”).
501 We then counted the occurrence of each noun phrase in the document. The aggregated “document-
502 term matrix” was stored in Parquet format (61) for performance reasons, and written and read
503 using the `arrow` package for R (62, *version 7.0.0*).

504 NLP-extracted noun phrases offer a number of advantages over the more traditional unigram
505 (“single word”) terms. First, noun phrase extraction removes many standard stopwords (articles,
506 common verbs, numbers) without relying on a fixed, *a priori* stopword list. Phrases can be more
507 informative than single terms, for example, distinguishing “intelligence test” from “hypothesis
508 test.” While simple n-gram extraction will include numerous phrases that are not especially
509 meaningful. Consider the sentence “Since our first analyses of feeding patterns in rats, we had
510 been using a criterion of 40 minutes (Le Magnen & Tallon 1963; 1966)” (63). Bigrams such as
511 “since our” and “criterion 40” will likely be discarded in vocabulary selection, but significantly
512 increase the computational cost of vocabulary selection. Noun phrase extraction is therefore
513 more efficient.

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

516 **S1.2.1 Data and code availability**

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

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

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

524 **S1.3 Vocabulary selection**

525 We took an information-theoretic approach to vocabulary selection (64). Consider a game in
526 which I draw a document from the corpus, then a single token from that document. I tell you
527 the term, and you have to guess which document I picked. Intuitively, highly informative terms
528 (in this project, noun phrase types) are distinctive, allowing you to dramatically narrow down
529 the list of potential documents. This “informativeness” of a term can be quantified as the KL
530 (Kullback-Leibler) divergence from a “baseline” uniform distribution across documents to the
531 distribution conditional on the term, which we abbreviate as ΔH . Because the most informative
532 terms tend to be typos and OCR errors — these are unique to a single document — we multiply
533 the KL divergence by the logarithm of the total number of occurrences of the term across the
534 entire corpus. We refer to the resulting measure as $\log(n)\Delta H$ or ndH.

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

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

546 across the corpus.

547 A common rule of thumb in topic modeling is that the vocabulary should have about 10
548 times as many distinct terms as the number of documents in the corpus. However, we had some
549 concerns with computational demands here: the resulting document-term matrix would have
550 roughly $10 \times 33,000^2$ or 10.9 billion entries; with 10% density this would require on the order
551 of 4 GB of memory just for a single copy of the matrix; and most of the analysis was to be con-
552 ducted on the authors' laptops. We therefore chose to work with three smaller vocabularies, 5×,
553 $1 \times$ and $\frac{1}{5} \times$ the number of documents. We refer to these as the “large” (174,480 distinct phrases),
554 “medium” (34,896) and “small” (6,979) vocabulary, respectively. We compare findings across
555 vocabularies as a robustness check.

556 **S1.4 Topic modeling**

557 To fit topic models, we followed the approach proposed by Rohe and Zeng (65), which uses
558 varimax-rotated partial principal components instead of the variational inference methods used
559 by standard topic model packages such as `stm` (66). This novel approach was implemented in
560 the R package `tmfast` (67). We conducted a simulation study of `tmfast`, which found that it
561 was significantly faster and only slightly less accurate at reconstructing known word-topic and
562 topic-document distributions, compared to `stm` (67).

563 Topic models were fit for all three vocabularies (large, medium, and small) with $k =$
564 5, 10, 20, ..., 70 (number of topics), resulting in a total of $24 = 3 \times 8$ models.

565 Following the approach of (64), we did not attempt to identify a unique best fitted model for
566 further analysis. While the manuscript text focuses on the medium vocabulary, $k = 40$ model
567 as the “median” among the 24 fitted models, we compare and contrast findings from this model
568 with those from the other models.

569 **S1.4.1 Topic model interpretation and quality assessment**

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

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

583 We next constructed Silge plots (18), showing the top 15 (highest-probability) terms (noun
584 phrases) from each topic in the model. These are provided as supplemental PDFs S2-S4. Perus-
585 ing these term lists, we focused on topics in each model that contained racial terms. Every fitted
586 topic model included at least one racial topic. Importantly, in some cases these topics contained
587 terms related to intelligence research, but in other cases they did not.

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

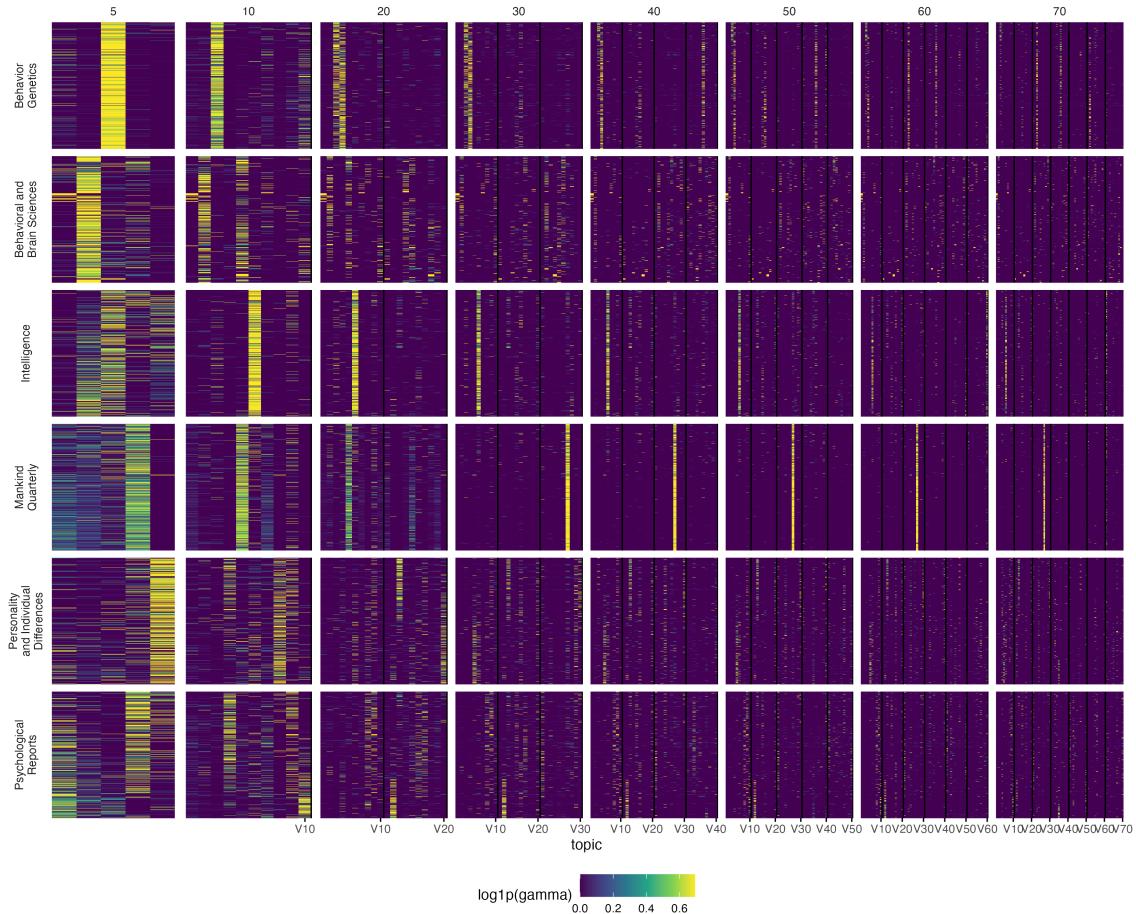


Figure S1.3: Fitted topic models: Gamma (topic-document) distributions, large vocabulary.

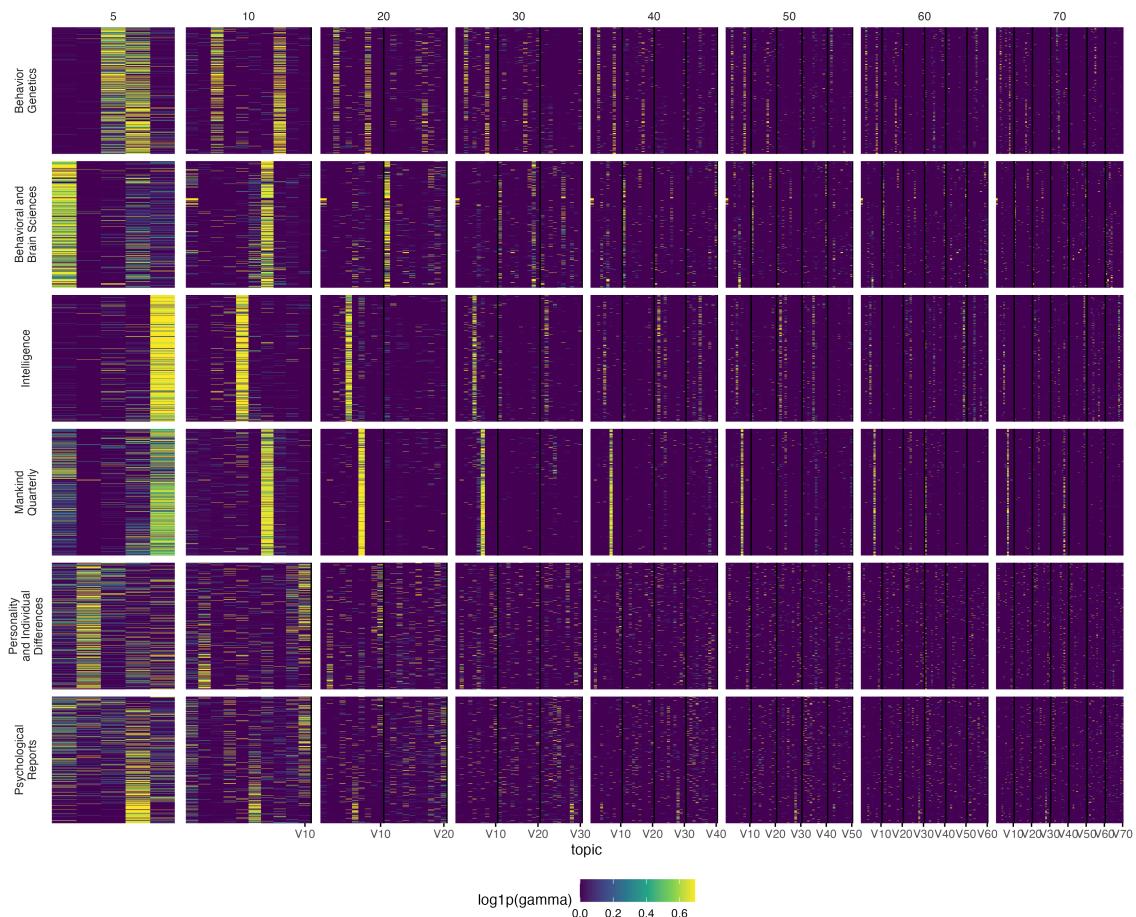


Figure S1.4: Fitted topic models: Gamma (topic-document) distributions, medium vocabulary.

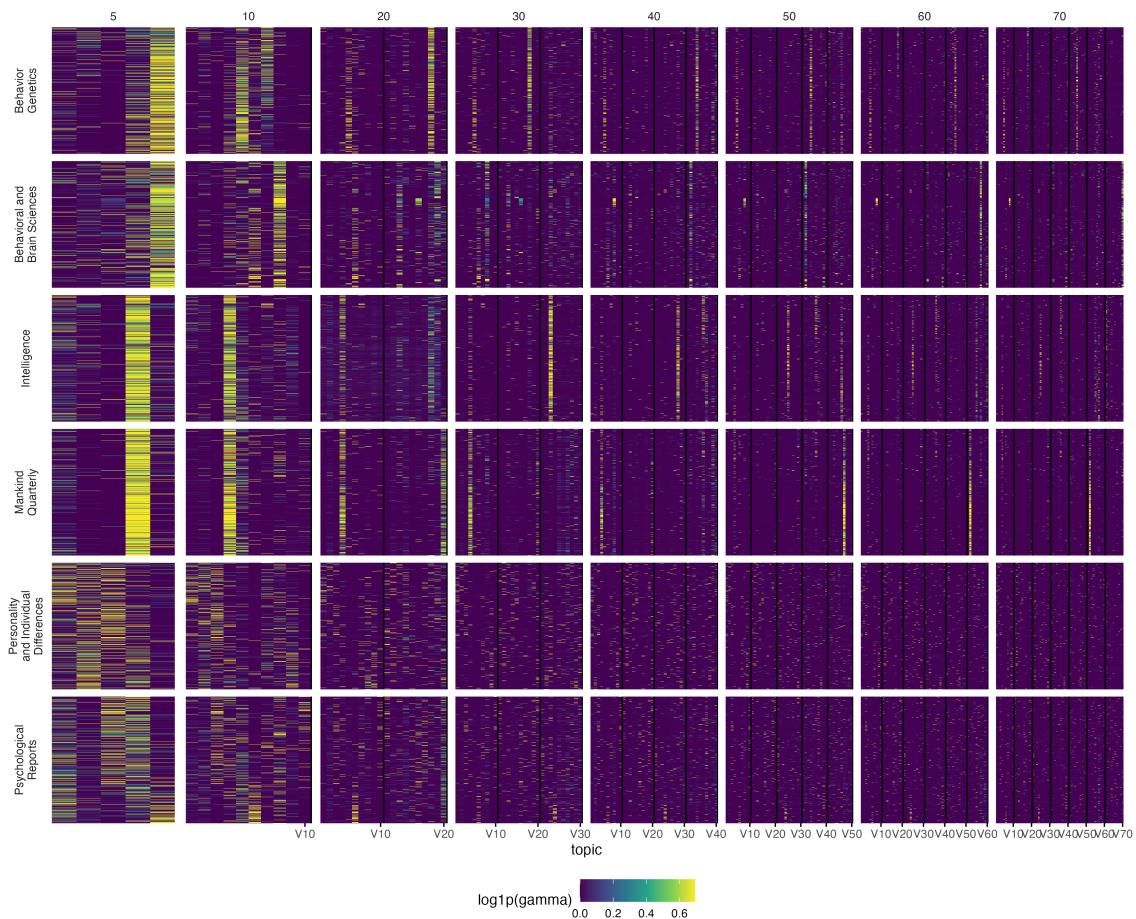


Figure S1.5: Fitted topic models: Gamma (topic-document) distributions, small vocabulary.

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

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

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

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

619 The 3 non-consensus documents reflected essential ambiguity in the operationalization of
620 race science discourse. One is a table of contents from a 1980 issue of Int; arguably this doc-
621 ument should have been excluded as irrelevant front matter, but arguably it legitimizes the
622 research published in the issue. The second case is a paper on racial differences in emotional
623 intelligence that was classified as race science discourse by one author, while the other au-
624 thor felt it was more focused on psychometrics (68). And a 1990 article in MQ (<https://www.unz.com/print/MankindQuarterly-1990q3-00108/>) proposed to offer
625 an economic explanation for crime. While it has a few pages connecting race/ethnicity to both
626 crime and poverty, it ultimately focuses on unemployment and alcohol sales, rather than race,
627 as the key predictors of crime.

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

633 **S1.5 Supplemental analysis of *Behavior Genetics***

634 To check the finding that very few race-and-intelligence documents had been published in BG,
635 we ran an independent search for documents with the keywords “race” and “intelligence” pub-
636 lished in BG from 1972 to 2020 using Springer’s journal search website. This search returned
637 125 documents. Thirty-three were presentation abstracts from meetings of the Behavior Ge-
638 netics Association; we did not examine these further. Fourteen of the remaining documents
639 had been published since 2010 (after the scope of the primary study). Among these, one obit-
640 uary and two papers celebrated the work of John Loehlin, a prominent race-and-intelligence
641 researcher who had been Director of the American Eugenics Society from 1968-1972 (69–71).
642 Another was a retrospective of the work of Lindon Eaves, a geneticist who made at least one

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

645 Twenty-eight documents in this sample from BG were published between 1990 and 2010.
646 One was an obituary of Jerry Hirsch, a critic of hereditarianism in general and Jensen in par-
647 ticular (72), and another was an obituary of David Rowe (73). Only 2 of these 28 documents
648 reported race differences of any kind: one examining interactions among race, sex, and heri-
649 tability for adolescent BMI (Body Mass Index, used as a measure of overweight/obesity) (74);
650 and the other interactions among race, family history of alcoholism, and visuospatial perfor-
651 mance (75). Finally, Philip Vernon — one of the Pioneer-funded scientists we identified earlier
652 — published a critical review of a book that attempted to address hereditarianism, the Flynn
653 effect, race-and-intelligence research, and some related issues.

654 At the same time, we were unable to find any articles published in BG that acknowledged
655 the involvement of the field with scientific racism with anything like the force of the statements
656 that have been made by biological anthropologists (27) and human geneticists (52). Notably,
657 the ASHG report includes discussions of behavior genetics race science, and in particular is
658 critical of the field and the organization for failing to publicly reject the race-and-intelligence
659 claims made by Shockley and Jensen (52). A search for “eugenics” in BG found articles that
660 presented the issue as part of the distant past, and a search for “racism” turned up studies
661 of racist attitudes. A statement on the Behavior Genetics Association website, dated 2021,
662 acknowledges that “The history of our field is inextricably linked with racism, including the
663 misuse of behavior genetic research to support violent eugenic policies,” but primarily focuses
664 on BGA member demographics ([https://www.bga.org/content.aspx?page_id=](https://www.bga.org/content.aspx?page_id=22&club_id=971921&module_id=567723)
665 22&club_id=971921&module_id=567723).

666 S1.6 Limitations

667 A key limitation of the current study stems from the inability of text mining techniques to draw
668 on a broader cultural context than what is represented in the corpus. Scholars of racism and US
669 race relations have noted a shift in the way racist attitudes are understood and expressed, from
670 overt and direct to a subtle “color-blind” racism (76). Color-blind racism represents a change in
671 the primary mechanisms by which white privilege is maintained, using non-racialized language
672 in furtherance of the racial status quo both institutionally and individually, and increasingly
673 reliant on the invisibility of socio-cultural mechanisms to preserve inequalities. Color-blind
674 racism involves the use of racialized “codes” or “dogwhistles,” language that is superficially
675 non-racial but carries racial implications, such as (in the US context) “violent inner city crimi-
676 nals” or “welfare queen” (77). Alderfer notes that the racist implications of *The Bell Curve* only
677 emerge gradually, with the opening of the book using the color-blind language of “group differ-
678 ences” (78). These cases suggest further that color-blindness can facilitate white supremacist
679 appropriation of non-racial research — even contrary to the researchers’ own intentions — as
680 scientific racist readers re-interpret racial-neutral language as dogwhistles (41, 79).

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

689 The phenomenon of color-blind racism is closely related to ambiguities that cannot be re-
690 solved even by human coders familiar with the broader cultural context. Two relevant examples

691 are the “Flynn effect,” a secular increase in intelligence test scores noted by philosopher James
692 Flynn; and studies of “national IQ,” exemplified by the work of Richard Lynn. The Flynn ef-
693 fect implies that environmental factors can create group differences in IQ that are comparable
694 to Black-White differences, and thus directly challenges claims that the racial differences are
695 biological. Using methods widely regarded as ad hoc, cherry-picking, and generally unreliable,
696 Lynn and collaborators have claimed that national IQ averages are correlated with national GDP.
697 Both the Flynn effect and “national IQ” are strongly associated with race-and-intelligence dis-
698 course in the US cultural context. But scientific journal articles on these topics may not include
699 any racial language at all. For example, a meta-analysis of the Flynn effect does not discuss
700 racial differences or use terms such as “White” or “Black” outside the reference list (80). Should
701 such articles be coded as race science discourse? We expect that even readers who accept our
702 definitions of race science and race science discourse will reasonably disagree on how to an-
703 swer this question. Topic models, of course, are not even capable of representing this essential
704 ambiguity.

705 81.