

COMBATTING DIGITAL DISINFORMATION

AN EVALUATION OF THE WILLIAM AND FLORA HEWLETT FOUNDATION'S DISINFORMATION STRATEGY

Drs. Amy O'Hara and Jodi Nelson

OCTOBER 2020

Contents

Executive Summary	2
Introduction	3
Background	3
Evaluation Design & Limitations	5
A Quick Primer on Disinformation	7
Hewlett's Disinformation Strategy	8
Evaluation Findings	11
1 The team achieved uneven progress according to its own metrics.....	11
2 The Hewlett foundation played a catalytic role in facilitating the growth of a network of scholars and funders, largely because of the energy, skillset and efforts of the lead program officer.	12
3 Hewlett seeded a multi-dimensional grant portfolio that represents the academic diversity needed to understand a complex topic like disinformation.	14
4 There is increased knowledge about the problem of disinformation now compared to when Hewlett started its work, but challenges remain to generating actionable evidence that can inform decisions about policy solutions.	15
5 The assumption that building one field is both possible and desirable is not borne out by our evidence.	16
6 From our interviews and analysis, it seems appropriate to conclude that the “discovery” stage of this portfolio is complete.	17
7 There is an apparent split among scholars who study disinformation as an equity-neutral, technocratic phenomenon and those who study it as integrally related to equity and power.	18
8 The Hewlett team’s commitment to racial equity has not been well communicated.	19
Where do we go from here?.....	21
Implications of our Findings.....	21
Advice from Hewlett stakeholders.....	23
APPENDIX 1. External Literature	25
APPENDIX 2. Key Informants Interviewed	28

Executive Summary

The Hewlett Foundation's decision to add a focus on disinformation to its 2017 US Democracy strategy was motivated by a desire to figure out how best to encourage social media platforms to reduce polarization exacerbated by disinformation.¹ The foundation invested in two major efforts to support this ambition: a multi-funder partnership with Facebook to create data access for independent researchers to study the impact of social media on democracy and elections; and a grant making portfolio conceived to build an academic field to deepen the understanding of disinformation and identify policy solutions that mitigate its spread and impact on society.

This is the second evaluation in a two-part series to investigate the effectiveness of both of these important efforts.² The data we collected indicate that the key Program Officer at the time, Kelly Born, used her voice, networking skills and foresight to create momentum in the philanthropic funding of disinformation research. Among all the tactics at the foundation's disposal, this program officer's own personal skillset was critical to the foundation's ability to execute on its strategic goals. The peer funders and grantees who know the foundation well cite the program officer's role as integral to the foundation's contributions.

In a similar vein, foundation president Larry Kramer's decision to create a multi-funder partnership with Facebook is also perceived as evidence of the philanthropy's ability to take risks, innovate and act decisively to respond to external events. Although the Facebook experiment did not achieve its stated goals, several stakeholders in our interviews noted that it tested essential assumptions about the obstacles preventing social media companies from sharing user data with independent academic researchers in order to contribute to the public good.

Expert opinion holds that there is more knowledge now about disinformation than there was when Hewlett started this portfolio. That said, the quick pace of technological change imperils the search for replicable, generalizable evidence because of what was described by scholars as *temporal validity*. Moreover, although there remain significant gaps in the overall knowledge base, Hewlett grantees and other organizations are already defining and working on solutions to address the impact of disinformation on people's lives and communities. Among our findings is the conclusion that the foundation may now consider complete its intended period of discovery in light of these circumstances.

We found compelling the suggestion that disinformation is better considered a topic than a coherent field, and that multi-disciplinarity brings richness in framing, methods, and applications. Among Hewlett grantees, scholars self-identify as members of different fields that carry with them important epistemological and philosophical differences. We found evidence that there are perceived gaps between researchers, platforms, journalists, policy makers, civil society and others that obstruct the development of policy solutions. The missing competency among many academics to translate research findings for use, and of the need for tighter connections between scholars and decision makers were additional insights we gained about important gaps in the current landscape.

During the two-year grantmaking period, the disinformation portfolio invested in varied policy and academic institutions and supported a diverse group of principal investigators. This is in contrast to the

¹ Changed in Spring 2020, the name of the US Democracy team was the Madison Initiative from its inception in 2014 until 2020. We use the new title of the team in this evaluation.

² Evaluation of the Social Science One – Social Science Research Council – Facebook Partnership (2019).

perceptions we heard about the foundation's lack of a diverse grantee portfolio and approach. We learned from key informants and secondary research that there is an inextricable link between disinformation, racial and social equity and justice. A stronger communication strategy about the foundation team's approach to diversity, equity and inclusion is warranted, as is a deliberate effort to showcase the work of all their grantees and their distinct perspectives. The foundation's significant role in creating momentum for this body of work was widely noted, as was the need for continued leadership and support.

Introduction

Background

In the wake of the 2016 presidential election, the William and Flora Hewlett Foundation decided to help address the challenge of digital disinformation and, in particular, "the role that social media and search platforms play in fostering affective polarization."³

Hewlett joined other philanthropies interested in understanding the potential influence that disinformation had on democracy and elections. While they all recognized the use of propaganda and false information as long standing political tools in many societies, there was something new about the way in which social media platforms (hereinafter referred to as platforms) accelerated the spread of disinformation and risked deepening polarization. Many of the key informants we interviewed from peer foundations describe the steps they all took to "throw several things against the wall" to see what might stick as an effective strategy to prevent these trends from undermining democratic institutions, values and behavior.

The challenge they all faced was a lack of knowledge about the phenomenon itself. Disinformation – or the pervasive online presence of heavily biased, misleading, and/or emotionally salient misinformation and propaganda – was a black box. Key questions about who spreads it, how it spreads, who consumes it and what impact it has on people's beliefs and behaviors needed to be answered before any policy solutions could be identified.

With this in mind, the Hewlett US Democracy team decided to invest in research to understand the problem before determining how the foundation might contribute to fixing it. At the time, scholars studying the topic did so from disparate disciplines, lacked a common set of research questions, and lacked sufficient funding and access to platform data, which was deemed essential to understand the dynamics, mechanisms and spread of disinformation. The foundation's primary assumption was that the connection between basic research and policy solutions would come only when the problem was better understood and a field of study had developed. The team used several tactics to pursue its objective to understand disinformation and build a field of study. Central to their approach was a series of discussions and white papers designed to elicit the core questions of a founding research agenda from a cross-selection of experts and compile what was known to date. With input from a mix of peer funders, platform leaders and scholars, they concluded there was high demand and interest in Hewlett playing a leadership role.

³ Kelly Born and Nell Edgington, "Analysis of Philanthropic Opportunities to Mitigate the Disinformation/Propaganda Problem," Fall 2017.

The Hewlett team began exploratory grantmaking without a dedicated budget in 2017. In March 2018, they requested and received from the foundation's Board of Directors \$5 million in new funding to supplement \$5 million they planned to repurpose from their existing budget for exploratory grantmaking in this area. This \$10 million comprised approximately 22% of the \$45 million Madison Initiative budget at that time. Over the next two years, the team made 23 grants and supported the Facebook-Social Science One partnership that Hewlett Foundation president Larry Kramer launched with 7 peer foundations to test a novel academic-industry model that would make platform user data available to researchers to study the impact of social media on democracy and elections.⁴

The landscape is different just two years later. Scholars across many disciplines study disinformation and are keen to identify solutions to mitigate its spread and impact on people. We use MediaWell – a Hewlett grantee – as a source to represent a proxy measurement of the increase in relevant research published over these same years in Figure 1.⁵ There is new knowledge and new funding for research that did not exist when Hewlett began its work in this area. One key informant shared his view that anyone researching the most pressing threats to US democracy today is working on the causes or consequences of disinformation.

FIGURE 1. MEDIABELL PUBLICATIONS 2017-19



Despite the fact that platforms have not dramatically opened up access to their user data, or increased transparency about their own business processes, platforms engage more now with scholars to study important questions. There are new articles daily that speak to the increased pressure platform leaders face to change their behavior and business models.⁶ As the 2020 presidential election nears, the COVID

⁴ Social Science One was launched as a new research enabling organization and formed a partnership with Facebook to build a new type of academic-industry partnership designed to both increase data access and credible, independent research on the relationship between social media and democracy. The Facebook-Social Science One partnership accounted for nearly \$2 million of the portfolio and required an enormous amount of time and effort by the foundation President and the US Democracy Program Officer in particular. This partnership is described in detail in both Hewlett and partner press releases and in the evaluation of the partnership conducted in 2019.

⁵ MediaWell curates research and news on digital disinformation and misinformation. This project from the Social Science Research Council aggregates content and events from around the web, distilling messages for the scholarly and policy communities.

⁶ Matt Stoller, "Congress forced Silicon Valley to answer for its misdeeds. It was a glorious sight," *The Guardian*, July 31, 2020.

pandemic remains at a peak, and the need for racial justice is front and center, there are indications that platforms are responding to mitigate the impact that disinformation has on these phenomena.⁷

The Hewlett US Democracy team commissioned this evaluation to assess whether and how the foundation's strategy contributed to this progress, take stock of changes in the landscape and inform its decision making about how to revise its approach moving forward.

The report includes four sections.

- **The Introduction** describes the evaluation design and provides a short primer on the issue of disinformation for readers who may not be steeped in the subject.
- **Strategy Overview** describes the foundation's goals, underlying assumptions and the composition of its portfolio of activities.
- **Evaluation Findings** shares the evaluation team's answers to the foundation's primary evaluation questions.
- **Where do we go from here?** presents the evaluators' reflections on the evaluation's implications for the foundation's learning, and advice we heard from Hewlett's key stakeholders about how the team might revise its strategy moving forward.

Evaluation Design & Limitations

This evaluation was designed to assess the work the US Democracy team has done since 2017 to combat digital disinformation. It is the second of two related evaluations. In December 2019, we evaluated a Hewlett-led funding collaborative to support a novel industry-academic partnership between Facebook, Social Science One and the Social Science Research Council. Also part of the foundation's disinformation strategy, the project was conceived to offer academics extensive privacy-preserving data with no pre-publication approval requirements, thereby enabling independent, credible research on social media's impact on democracy.⁸

In this second part of our evaluation series, we assessed two bodies of work: a set of field-building activities and \$10.135 million in grantmaking to date. The primary audience of the evaluation is the US Democracy team, although the foundation will share the final report publicly with the hope that funders, researchers, private and public policy makers and others interested in this topic will learn from the foundation's experience.

Five major evaluation questions guided our data collection and analysis:

- What did the foundation set out to do, and for whom?
- What actually happened – what progress was made?
- Why did things happen as they did – where and how did the Hewlett foundation play a shaping role in the field and what enabled it to play this role?
- What are the lessons to be learned?

⁷ For example, Twitter has factchecked and labeled misleading Trump tweets, Google-owned YouTube placed a World Health Organization banner about COVID at the top of their search results, and Facebook inserted a box in the newsfeed linking to the Centers for Disease Control for facts on coronavirus.

⁸ <https://hewlett.org/wp-content/uploads/2020/02/Facebook-Partnership-Final-Evaluation-Report.pdf>

- What are potential new directions and next steps for this line of grantmaking?

Given these questions and the nature of the evaluation, we used qualitative data collection and analysis to identify patterns in stakeholders' perceptions. We used publicly available and foundation data to strengthen our analysis, and the expert opinion of several scholars to identify changes in the evidence base. We did our own literature review of outside sources, reviewed key foundation documents and used iterative feedback with the Hewlett team to assure our recommendations are both relevant and useful to their decision making.

The sources of our data are presented in table 1 below. We share the external literature used in Appendix 1.

TABLE 1. DATA SOURCES USED IN EVALUATION	
EVALUATION TEAM DATA COLLECTION	FOUNDATION DOCUMENTS (Hyperlinked)
Interviews of 40 stakeholders, reflecting a 89% response rate from the sample identified by the foundation team.	Application summaries, proposals, interim reports, and final reports of disinformation grants
Publicly available data on gender and race of Hewlett grantees	Analysis of Philanthropic Opportunities to mitigate the Disinformation/Propaganda Problem
Foundation data on the allocation of funds across the disinformation portfolio	Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature
External literature review on relevant topics, including press, academic articles, outside conference agendas, blogs of Hewlett grantees and other researchers	What's Next for Digital Disinformation? A Research Roadmap The Madison Initiative Developmental Evaluation by the Center for Evaluation Innovation, August 2016 Update on the Madison Initiative in the Wake of the 2016 Election memo from the Madison Initiative team to Partners and Colleagues, June 2017 Proposal to combat digital disinformation, submitted by Kelly Born, Daniel Stid, Larry Kramer to Hewlett Foundation Board of Directors, March 2018 Madison Initiative Outcomes Framework, January 2020

The foundation's US Democracy team identified a sample of 45 key informants, across the five segments described in the table below. We completed 40 interviews with grantees, funders, and experts from universities, policy organizations, and companies (who are not funders or grantees). The sample of people interviewed is presented in detail in Appendix 2.

TABLE 2. INTERVIEW SAMPLE		
	Identified	Interviewed
Hewlett	6	6
Hewlett Grantees	13	14
Funders	8	8
Academic Experts*	6	9
Industry Experts*	12	3
Total	45	40

*Academic and Industry Experts were not Hewlett grantees

Because the evaluation uses primarily qualitative methods and data, its findings are based on the patterns in perceptions and insights of those interviewed. With this in mind, we note two relevant limitations.

- (1) President Trump's relationship with the social media platforms entered a particularly volatile phase during the evaluation period (May-July 2020). Administration actions threatening to alter the regulatory landscape for companies may have reduced key informants' willingness to engage with this evaluation. We had limited feedback from industry representatives (e.g., Facebook, Google, Microsoft, Twitter, and Reddit) who had participated in earlier Hewlett convenings and are likely to have helpful insights into the evolution, gaps and opportunities for future work.
- (2) We interviewed 23 scholars as part of our data collection – a relatively small sample of the scholars working on disinformation today. We note this as a limitation because it speaks to the evaluation's necessarily limited description of the study of disinformation writ large.

We hope that other partners and funders will follow Hewlett's lead and evaluate their work on disinformation so that our evaluation on the Facebook-Social Science One partnership and this evaluation contribute to a larger knowledge base about this important topic.

A Quick Primer on Disinformation

Disinformation is false or misleading information intentionally spread with ill-intent. This synonym for propaganda usually has a political or economic objective, is intended to influence public attitudes, or hide the truth. Disinformation is used to incite strong feelings. It affects behaviors, including sharing and liking on social media, that expand its spread and potentially its impact. Misinformation is false or misleading information that is spread unintentionally. While it is "intention neutral", it can be incredibly harmful, especially when individuals don't realize they are sharing false information but share it widely.

Although disinformation is not a new phenomenon and many countries have experienced its potential to influence political outcomes, the 2016 election was a wake-up call for the American public.⁹ In earlier political cycles, traditional news media that upheld journalistic standards typically provided the source of newsworthy information and disseminated it to the public. Journalistic principles including truthfulness, accuracy, objectivity, impartiality, fairness and public accountability were largely absent on social media. After the 2016 election, many feared that fake news articles by foreign actors and spread on Facebook swayed the results of the election.

Concerns were exacerbated because of the growing numbers of people using social media and therefore subject to the ostensible influence of the information spread on various platforms. Indeed, social media use grew between the 2012 and 2016 presidential elections. Facebook, for example, had 1.06 billion monthly active users in Q4 2012, 1.86 billion monthly active users in Q4 2016, and stood at 2.6 billion in Q1 2020.¹⁰

⁹ Samantha Bradshaw and Philip Howard, "The Global Disinformation Order: Global Inventory of Organized Social Media Manipulation." Project on Computational Propaganda. 2019.

¹⁰Facebook Investor Relations Press Releases. <https://investor.fb.com/investor-news/press-release-details/2020/Facebook-Reports-First-Quarter-2020-Results/default.aspx>

The prevalence of platform usage combined with the industry's business model means that information can spread to huge numbers of people based on both automated and individual behaviors. The platforms rely on user engagement, using algorithms to determine which content or ads to serve each user. With the wealth of information collected about past likes, shares, locations, and comments, their ability to segment and target users surpasses anything experienced with traditional media.

Platform business models are reliant on ever-increasing screen time and interactive use. Companies such as Facebook, Google, Reddit, and Twitter want to keep users on their sites, connecting with their friends and families and likeminded strangers, searching for information, or generating new content. This business model is largely hands-off on content moderation, with company policies prohibiting only violent or threatening content.¹¹ This enables conspiracies and polarizing content to flourish.

Hewlett's Disinformation Strategy

In the wake of the 2016 presidential elections, the Hewlett Foundation Board of Directors asked the US Democracy team to consider enlarging their strategy. For the next year, program officer Kelly Born investigated where and how Hewlett's funding could impact the problem of digital disinformation, and in particular the role that social media and search platforms play in fostering affective polarization.

"Affective polarization refers to something beyond ideological or policy differences and involves emotional, tribal disdain for—even hatred of—an opposing party. This form of polarization, which has grown rapidly in recent years, is a matter of great concern for the Madison Initiative. This was a primary concern of ours even before the election, but the campaign and its aftermath—including all we have learned about the role of disinformation and internet platforms—crystallized its importance for our efforts."¹²

The foundation team focused its learning during the two-year exploratory period (2018-2020) on the role that social media platforms play in accelerating disinformation and polarization among different parties. Their rationale was that platforms such as Facebook, Google, Twitter and Reddit have "removed traditional media gate keepers" and contributed to decentralizing the media environment. Further, because their business models and algorithms encourage anonymous, widespread engagement, platform companies can be key catalysts of polarization.

After that period, the US Democracy team would evaluate and reassess whether enough was known to identify and therefore invest in or advocate for policy solutions that could influence platform behavior and mitigate the spread of disinformation. In 2018, they felt it was premature to define solutions without a better understanding of the problem itself.¹³

¹¹ Facebook moderators flag violence, spam, harassment, self-harm, terrorism, intellectual property theft, and hate speech, as noted by Josh Constine in "Facebook reveals 25 pages of takedown rules for hate speech and more." TechCrunch, April 24, 2018. <https://techcrunch.com/2018/04/24/facebook-content-rules/>

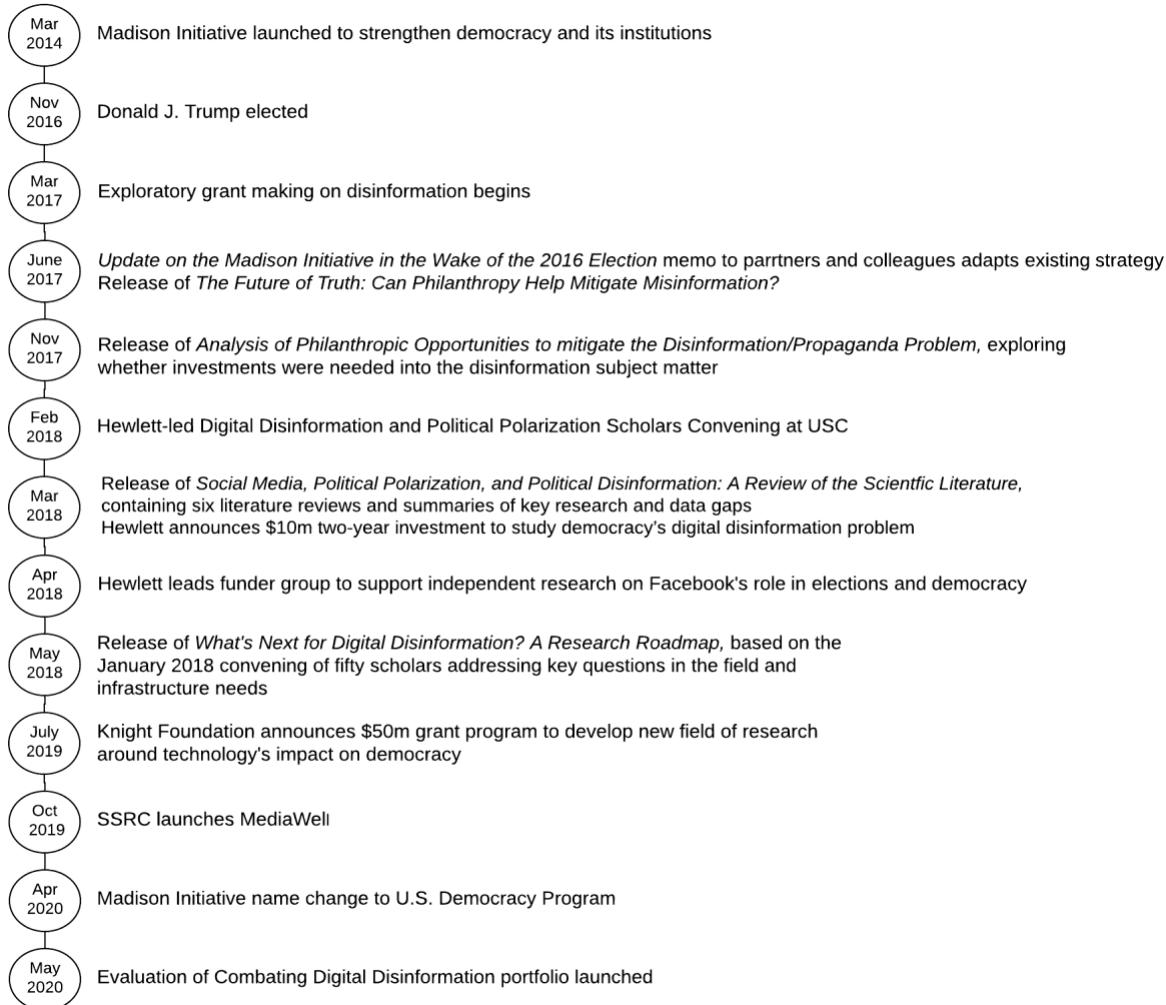
¹² *Funding proposal to combat digital disinformation* submitted by Kelly Born, Daniel Stid, Larry Kramer to Hewlett Foundation Board of Directors, March 6, 2018.

¹³ Ibid.

"We know there is a problem, but we don't really understand its dynamics or contours. In part, this is because the platforms have been reluctant to share data, but it is also because the relevant scholars are in disparate academic disciplines and have yet to come together (or be funded) to think or work as a field. They are like the proverbial group of blind men describing an elephant."¹⁴

The timeline below presents the sequence of events that defined the foundation team's learning through convening, the writing of white papers, a proposal to the foundation Board and launch of relevant grant giving in 2018.

FIGURE 2. TIMELINE



¹⁴ *Funding proposal to combat digital disinformation* submitted by Kelly Born, Daniel Stid, Larry Kramer to Hewlett Foundation Board of Directors, March 6, 2018.

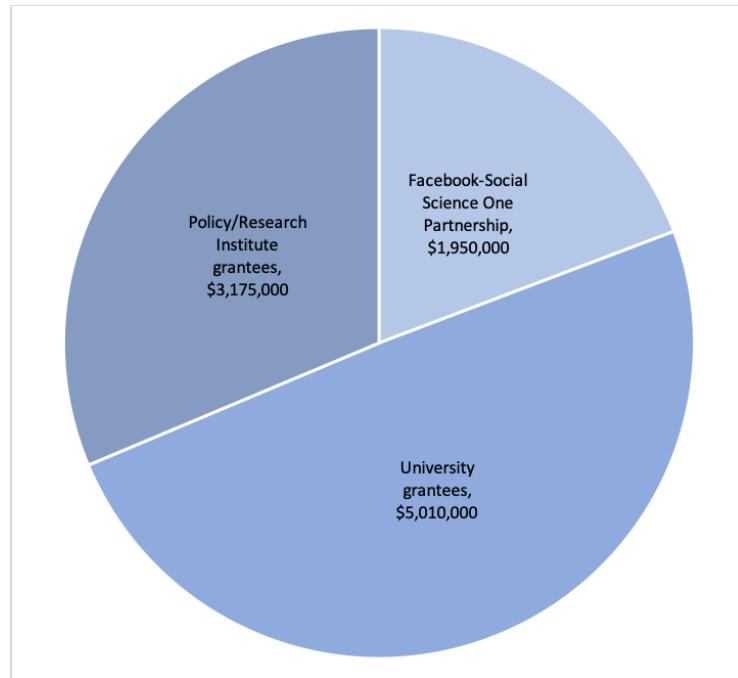
The US Democracy team defined its goal to support research that fit these three criteria:

1. Understanding the problem better, including the supply of and demand for disinformation, and its impact on individuals and society more generally.
2. Understanding potential solutions, e.g., reducing disinformation's negative impact on individuals, alleviating levels of affective polarization in society, elevating quality content, and marginalizing problematic content.
3. Understanding the philosophical, legal, and technical constraints on acting and how these might need to be adjusted to meet demands of the digital age.¹⁵

The team's strategy to achieve this goal was to create a strong digital disinformation field, including data sets and mechanisms to translate academic findings to the real world.¹⁶ They would strengthen this field by supporting research grantees, working with peer funders, identifying key decision makers at the platforms and developing relationships to influence their actions, and building infrastructure deemed necessary for a field to flourish.

To date, the Hewlett team has allocated \$10.135 million in grants to universities, policy and research organizations. The pie chart below summarizes how these funds have been allocated.

FIGURE 3. FUNDING FOR DISINFORMATION GRANTS



¹⁵ *Proposal to combat digital disinformation*, March 2018.

¹⁶ *Madison Initiative Outcomes Framework*, January 2020.

The Hewlett team's decisions were motivated by three main assumptions about how best to address the complex relationship between disinformation and polarization.

1. The first assumption was that what political scientists call “affective polarization” – defined as extreme animosity between political parties – is a pressing threat to our democracy and can be alleviated by stopping the spread of disinformation on social media platforms.¹⁷
2. The second key assumption was that influencing the behavior of social media companies to be “good civic actors” was an essential and feasible recourse for philanthropy to take.¹⁸ The Hewlett team saw the central role the platforms played as channels of disinformation as rationale to support a focus on platform behavior: “Both facts—the platforms’ contribution to political polarization and their potential capacity to alleviate it” were at the heart of the foundation’s thinking.
3. The third assumption was that the translation from basic research to policy solutions would come only when the problem was better understood, and a mature field of study had developed. This assumed further that creating a coherent field of study focused on disinformation was possible, desirable and necessary to identify solutions that could influence platform behavior and mitigate polarization.

We return to these assumptions throughout the report and summarize our conclusions about their validity in section four.

Evaluation Findings

1 The team achieved uneven progress according to its own metrics.

We summarize in table 3 the changes we observe across the team’s defined ‘implementation markers,’ measures the Hewlett Foundation uses to track its strategic progress. The uneven progress may be due in part to the way the markers are defined. It is, for example, unclear whether they pertain to progress made by the foundation or its grantees, making it difficult to identify reliable evidence that the marker was achieved. In addition, some markers were too aggressive for the two year time horizon, like establishing a research agenda or creating access to data that would enable evidence building on the effects of disinformation. Others were unrealistic given the nature of academic study. For example, the Hewlett team could help enable the definition of a research agenda, but could not assign it to academic researchers whose incentives or interests may not align with its contents.

More specific findings related to many of these markers are described throughout the report.

¹⁷ Shanto Iyengar et al, “The Origins and Consequences of Affective Polarization in the United States,” Annual Review of Political Science, Vol. 22:129-146 (2019).

¹⁸ *Analysis of Philanthropic Opportunities to Mitigate the Disinformation/Propaganda Problem*, 2018.

TABLE 3. PROGRESS TOWARDS 2018 MARKERS

Maintaining and growing collaborations	<input checked="" type="checkbox"/> early collaboration with peer funders during the Facebook project <input checked="" type="checkbox"/> support of 2 collaborative projects
A clearer definition of the online disinformation/polarization problem and its effects	<input checked="" type="checkbox"/> on problem definition <input checked="" type="checkbox"/> on defining effects
Joint establishment of a research agenda and identification of key organizations equipped to handle each necessary research component	<input checked="" type="checkbox"/> research questions documented <input checked="" type="checkbox"/> insufficient evidence that scholars used research agenda to guide their work
Establishment of a “learning community” among funders who convene periodically and follow and discuss the research	<input checked="" type="checkbox"/> a call was established <input checked="" type="checkbox"/> insufficient evidence to suggest a learning community of funders convenes to discuss research
Regular engagement with key social media and search platforms to obtain access to data and/or viable workarounds	<input checked="" type="checkbox"/> evidence that Hewlett engaged with Facebook to design the SS1/SSRC partnership <input checked="" type="checkbox"/> grantees engage with Twitter, Facebook, Reddit, Google, Microsoft
Establishment of a communications plan that successfully translates research findings for key government, funder, and platform audiences	<input checked="" type="checkbox"/> no communications plan to share or translate findings

2 The Hewlett foundation played a catalytic role in facilitating the growth of a network of scholars and funders, largely because of the energy, skillset and efforts of the lead program officer.

The Hewlett team launched its disinformation strategy in a collaborative way, canvassing a group of researchers, funders and platform leaders to affirm the need for improved understanding of disinformation and develop a research agenda.

Per the above timeline, the Hewlett team organized an inaugural convening to connect researchers across disciplines and create a forum for funders and platforms to discuss issues and define the contours of the research agenda. Many of the participants describe the convening as the “right thing at the right time” for two main reasons.

1. The mix of participants and intention to create connections among them. The convening professionally connected people who would not otherwise be connected. Hewlett’s ability to bring platforms to the table was critical, enabling policy and research leads from Google, Facebook, Twitter, Microsoft, and Reddit to meet and discuss problems directly with academics. Convenings also included a combination of academics across disciplines - political science, economics, network science, psychology, computer science – as well as practitioners from think tanks and advocacy organizations. As an interviewee explained, “Hewlett, to their credit and benefit, brought people from industry and scholars and funders to try and establish intellectual connections. The goal was to seed the area with connections and make sure the right kinds of people were connected in a semi-structured way.” As one participant describes, the convenings were “a milestone for sure – I learned more in that session than any other convening to date,

but it wouldn't work today – a lot more is known about the space. The actors, threats, funding. It worked because the field was just waking up.”

2. The drive to resolve a perceived knowledge gap facing the community of actors interested in the same topic. Because there was insufficient evidence at the time about disinformation, stakeholders agreed that jumping to policy solutions would be premature. The Hewlett team prioritized the development of a common research agenda that specified what was known and unknown about the cause, spread and impact of disinformation. Convening participants applaud the foundation for being product-oriented in this way, in contrast to subsequent, non-Hewlett convenings on disinformation that engaged the same researchers and often the same studies presented dozens of times without a concrete output.

The team produced the 2017 white paper, “Analysis of Philanthropic Opportunities to Mitigate the Disinformation/Propaganda Problem” and the 2018 report, “What’s next for digital disinformation? A research roadmap” – providing a narrative that was described as a “touchstone used to justify why work needs to be done.” Another key product was the article, “Social Media, Political Polarization and Political Disinformation: A Review of the Scientific Literature.”¹⁹ Compiled and edited by Joshua Tucker, the paper was the first canvassing of knowledge on disinformation; it is described by some key informants as one of the most cited documents in relevant research today.

Both the research roadmap and the literature review contained key questions for an emerging field to address. A high level list of the questions that comprised the research agenda in Tucker et al (2018) is presented in the table below.

TABLE 4. KEY RESEARCH QUESTIONS FROM SOCIAL MEDIA, POLITICAL POLARIZATION, AND POLITICAL DISINFORMATION: A REVIEW OF THE SCIENTIFIC LITERATURE (MARCH 2018)
1. What are the effects of exposure to information and disinformation on individual beliefs and behavior?
2. What are the cumulative effects of having accounts on multiple platforms, and how might such conclusions differ from what we’ve learned from studies of behavior on a single platform?
3. How does the spread of disinformation through images and video differ from the spread of disinformation through text?
4. How do the spread and the effect of disinformation differ across different countries?
5. Do the effects of exposure to disinformation and polarization vary across liberals and conservatives?
6. What are the likely effects of new laws and regulations intended to limit the spread of disinformation?
7. What are the strengths and weaknesses of different methods of bot detection and analysis?
8. What is the role of political elites in spreading disinformation online?

Many key informants characterize the Hewlett team’s lead Program Officer on disinformation as the reason behind these early achievements. Kelly Born’s external engagement was described by Hewlett team members as atypical for foundation program officers. She is described by stakeholders as a visionary, especially inclined to tackle an amorphous and challenging topic like disinformation; “the thoroughness and depth with which Kelly pursued this whole set of ideas was incredible – one of the

¹⁹ Include links to Hewlett papers. The article, authored by Joshua A. Tucker, contains contributions from Andrew Guess, Pablo Barberá, Cristian Vaccari, Alexandra Siegel, Sergey Sanovich, Denis Stukal, and Brendan Nyhan. <https://www.hewlett.org/wp-content/uploads/2018/03/Social-Media-Political-Polarization-and-Political-Disinformation-Literature-Review.pdf>

best kinds of due diligence research jobs I've ever seen a Program Officer do to really understand the state of the field.”²⁰

Born was skilled at weaving the findings of grantees together. For example, in a single opinion piece, she referenced work by grantees at Data and Society, Oxford Internet Institute, Northeastern, NYU, and the Institute For The Future.²¹ She explained that “this research helps us understand where consumption of the most egregious cases of truly “fake news” are concentrated – which has important implications for where and how government, technology and the public decide should focus their attention and resources.”

Born created connections; she matched researchers with funders and convened a regular funders’ call for the community of foundations interested in disinformation to share learnings and priorities with one another. Repeatedly during our data collection, informants expressed the sentiment that one articulated quite clearly: “it’s impossible to imagine the field evolving as it did without Hewlett and Kelly.... There is finally some connectivity in the area that is productive and forming some sort of intellectual community.”

3 Hewlett seeded a multi-dimensional grant portfolio that represents the academic diversity needed to understand a complex topic like disinformation.

The disinformation grant making budget began officially in 2018, after two years of exploration, formulation and convenings. Among the team’s important accomplishments was their investment in different types of organizations and research. Although the foundation’s press releases and Born’s own media focused on the Facebook-Social Science One partnership and the overall problem to solve respectively, we found a rich array of viewpoints, methodological approaches and mandates across the grant portfolio.

We describe here the type of variation that seems especially important when considering the foundation’s objective to understand what was then an ill-defined problem and set of levers to influence political outcomes.

- The portfolio includes grants to both university-based researchers – who are ostensibly more constrained by the parameters of established disciplines, publication norms, and tenure requirements– and researchers who work at independent policy and research institutes and may be less circumspect about traditional academic incentives.
- While many university-based researchers in the portfolio use quantitative methods, focus on social media, derived digital trace data and the tools needed to draw insights from massive data sets, the portfolio also includes many researchers at both universities and policy research

²⁰ Kelly Born joined the Hewlett Foundation in 2012 as a Special Projects Program Officer. She helped launch the Madison Initiative and joined our team in 2014, steering investments in the Combating Digital Disinformation portfolio until late 2019. She joined Stanford University in 2019 as executive director of the Cyber Policy Center, situated in the Freeman Spogli Institute for International Studies.

²¹ “Disinformation Threatens 2020 Election.” April 29, 2019. Morning Consult.

<https://morningconsult.com/opinions/disinformation-threatens-2020-election/>

institutes who use qualitative methods and are more focused on building networks and mutual understanding among researchers, policy makers and civil society groups.

- The research portfolio includes researchers whose work is explicitly linked to platform data and requires platform engagement, and others who use more readily available online data and explicitly reject platform data and engagement as potentially influencing their findings.
- While some research grantees define their success in terms of publication and citation targets, others define their success in terms of whether they translate research to action, working closely with practitioners and policymakers to detect disinformation and mitigate its effects.
- Finally, the variety of outputs produced by grantees included not only research papers, but also careful curation and review of the scientific and policy literature, landscape and network analyses, and products that catalogue the harms that platforms are causing.

4 There is increased knowledge about the problem of disinformation now compared to when Hewlett started its work, but challenges remain to generating actionable evidence that can inform decisions about policy solutions.

Since 2017, researchers have studied many issues surrounding disinformation. Our qualitative assessment comparing the three original foundation goals with the knowledge gained is shown in table 5.

TABLE 5. COMPARING INVESTMENT GOALS TO KNOWLEDGE GAINED	
Investment Goals	Knowledge Gained
Understanding the problem better, including the supply of and demand for disinformation, and its impact on individuals and society more generally	Disinformation and misinformation are better defined and more evidence exists now about the groups most likely to spread disinformation. There is a paucity of evidence on the impact disinformation has on individuals and society more generally.
Understanding potential solutions, e.g., reducing disinformation's negative impact on individuals, alleviating levels of affective polarization in society, elevating quality content, and marginalizing problematic content	Content moderation as an overall solution was rejected, though partnerships with factcheckers are critical to test user behaviors, and takedowns are still crucial for disrupting spread. Scholars and platforms have a better understanding of bots and computational propaganda; “you can no longer use mass fake accounts to get something trending.”
Understanding the philosophical, legal, and technical constraints on acting and how these might need to be adjusted to meet demands of the digital age	The Facebook partnership with Social Science One exposed many of the legal, privacy and technical issues that continue to impede independent research access to user data. Research-policy networks are forming through efforts of several grantees.

Despite these gains, there are barriers to producing the type of knowledge that can actually inform the definition of alternative paths that platforms, policy makers, philanthropy and civil society can take to mitigate the spread and impact of disinformation.

- **Data access remains a central obstacle to certain types of research.** Computational social scientists who rely on large data sets directly from platform companies cite the lack of sufficient, consistent access to platform user data as the central obstacle to increased knowledge. More than a dozen interviewees have projects to conduct analyses or experiments working with the platforms, including Google, Facebook, Twitter, Reddit, and YouTube. Access issues vary across platforms, since some platforms have public-facing data and are simpler in design than others. One interviewee cautioned, “Don’t lump all platforms in with Facebook when discussing the problem or looking for solutions.” We also heard caution from 6 out of 12 experts interviewed that there is a risk presented by over-focusing on platform data.
- **There is an overemphasis on research that is problem rather than solution driven.** Despite the diverse objectives of Hewlett’s grantees, the foundation’s focus on academic research was described to us as problem focused, lacking sufficient attention to identifying and trialing alternative solutions. Key informants identified the possible root cause located in the computational social science reliance on big data from platform companies. One key informant described the approach of lab science taken by these scholars as “fetishizing data” and generating deep and narrow dives into problems with insufficient attention to both translation and solutions.²² “It’s unclear if big data leads us to evidence that is actionable,” explained another key informant: “this is a key problem – you do research and then what do you do with it? How do you show [the platforms or government] that the problem is the right one to address and how?” Several grantees recently co-authored an article specifically noting the need for computational social science to solve real world problems.²³
- **Building a cumulative evidence base may be challenging – or impossible – given how quickly technology and disinformation both change.** Temporal validity refers to the idea that a finding may be internally valid at one point in time, but not hold over time because of unique platform mechanics and disinformation’s speed of change overall. This reality presents a real challenge to the assumption that generalizable evidence can be produced through academic research to understand the problem and solutions for disinformation’s spread and impact.

5 The assumption that building one field is both possible and desirable is not borne out by our evidence.

Field building is a common framing and solution definition underlying Hewlett strategies. The US Democracy team’s founding documents on the disinformation portfolio reference the need and desire to build a field of study to centralize progress towards problem identification and the pursuit of policy

²² Labs address a need for scale when analyzing billions of records, achieving efficiencies in computing infrastructure and the development of methods that properly analyze data. While common in the hard sciences, lab structures were less common in the social sciences until adopted by computational political scientists and others using this type of data and approach.

²³ Grantees Lazer, Freelon, and Nelson co-authored the article, “Computational social science: Obstacles and opportunities,” which also called for strengthening collaborations with companies, investing in new data infrastructures, developing frameworks to address ethical, legal, and social implications, and addressing university incentives and structures. Lazer et al (2020).

solutions. While there is evidence that there is more research and increased knowledge on disinformation now compared to when the Hewlett Foundation launched its disinformation strategy, we did not find evidence that one coherent field of study focused on disinformation has developed.

Our analysis suggests that disinformation is a topic of study across many fields rather than a field in and of itself. Researchers across disciplines self-identify as members of different fields, depending on their objectives, values and priorities. Examples include: Computational Social Science, Public Interest Technology, Internet Studies, Critical Internet Studies. Among scholars, there are concerns of a fracture with political scientists doing computational social science with large amounts of platform data. This critique stems in part from the argument that “trying to know ourselves through platform data tends to yield partial and contorted accounts of human behavior that conceal platform interventions.”

“Such data can be impressive due to their unprecedented granularity and volume, as well as the fact that they are seemingly “unobtrusive” recordings of our activities when no one is watching. These apparent strengths of data for social research are outweighed by a problem in what we call the “measurement conditions”: platform data are platforms’ records of their own behavioral experimentation.”²⁴

In addition to the epistemological split represented by this critique, there are weak incentives for scholars from different academic fields to follow and collaborate on a common research agenda. It’s important to note that while Hewlett staff and some funded scholars lauded the 2018 literature review as a seminal document, other researchers saw it as a contribution but not an articulation of a common research program; some had not heard of it and still others had heard of it but not read it. Key informants reminded us of the reality that academic researchers do not make decisions based on what people tell them to do. In the words of a central key informant, “scholars do whatever they want to do, and pay little attention to what others think they should do.” Several key informants affirmed that researchers who study disinformation continue to engage their peers through their respective journals and conferences and are often incentivized to cite only their disciplinary colleagues.

The lack of one coherent disinformation field may not prove to be a disadvantage for efforts to solve for the most pressing issues today. Indeed, there seems to be an immediate need for foundations to coordinate their approaches. Throughout the remainder of our findings, we highlight the importance that key informants attach to a diversity of perspectives and the opportunity Hewlett has to reframe the problem it is working to solve in the next iteration of this important strategy.

6 From our interviews and analysis, it seems appropriate to conclude that the “discovery” stage of this portfolio is complete.

A key informant who knows the Hewlett team and their strategy well urged the team to “think about what the world would look like in 10 years if you are wildly successful funding a large number of influential publications, but we have not moved closer to addressing the problem. Is that success?” We heard from several key informants that there is “too much at stake” to be purely academic; while

²⁴ “Platform Enclosure of Human Behavior and its Measurement: Using Behavioral Trace Data against Platform Episteme,” Angela Xiao Wu and Harsh Taneja. This argument is summarized by Angela Xiao Wu in the blog, “How to now know ourselves,” <https://points.datasociety.net/how-not-to-know-ourselves-5227c185569>

research has a role to play, describing the problem is no longer needed. Summed up by a grantee, it is time to “amplify the signal of good information, dampen the signal of disinformation, and hold platforms accountable. Take that roadmap and push policy proposals.”

Moreover, research needs structured linkages to policy makers, incentives, and translational competencies if it is to influence platform behavior and government policy. Numerous Hewlett grantees are already working to build relationships, platform competencies and strong links between sectors (research, platform, government, journalism, civil society). Institutions such as the Social Science Research Council, German Marshall Fund, the Center for Humane Technology, Data and Society, and the Carnegie Endowment for International Peace focus on relationship and competency building that extends beyond academic circles. Still, key informants identify critical gaps that need to be closed to advance work that drives toward behavior change and not just understanding. Four of these include:

- Additional efforts to change practice from within platform companies by mobilizing employees and technologists to embrace a spirit of Duty of Care and define standards from the inside.²⁵
- Stronger research competencies in policy communication, relationship building and translation so that success of research findings is both judged and pursued by its translation to concrete, actionable recommendations in a timely manner.
- Collaboration “sprints” between platforms, researchers and policy experts to identify, test and communicate possible interventions.
- A broader coalition that unites researchers, industry, traditional media, human rights and civil society organizations to identify and ensure policy solutions are equitably implemented and not harmful.

Interviewees suggested that Hewlett can and should take greater risks in this dynamic and timely space, and that closing important gaps that currently obstruct developing policy solutions even with incomplete data is a necessary step to take.²⁶

7 There is an apparent split among scholars who study disinformation as an equity-neutral, technocratic phenomenon and those who study it as integrally related to equity and power.

The 2018 research roadmap and the Hewlett commissioned literature review do not include questions on equity, on the potential bias platforms may replicate through their systems or on the relative impact disinformation may have on different groups of people in society. Grantees who know the Hewlett commissioned literature review (Tucker et al, 2018) describe its contents as divorced from a deeper understanding of these issues, pointing out that any solutions, future research or recommendations that derive from this agenda may therefore exclude a significant root cause and consequence of disinformation. In our own review of more recent literature, we found efforts to summarize the current

²⁵ Several key informants referred to the need for a new “duty of care” norm to be embraced throughout social media and government. There is evidence that Hewlett grantee [Center for Humane Technology](#) has made progress working directly “with technologists to create a new definition of success: one that honors human nature, grows responsibly, and helps us live lives aligned with our deepest values”

²⁶ The Facebook-Social Science One partnership was mentioned as a good example of this; it was high risk but worth it because it was testing new ways of working in the future.

field of disinformation that similarly do not include reference to research about the relationship between disinformation, equity and power imbalances in our society.

At the same time, many of Hewlett's own grantees are working at the forefront of research that demonstrates how disinformation intersects directly with race and identity. Examples of congressional findings, recent award winning books and articles that focus on race and technology are illustrative of the type of insights that have been produced in just the last two years.

- In 2018, using Hewlett grantee research, the Senate Select Committee on Intelligence concluded that Russian information operatives were mainly targeting African Americans. The Committee found that no single group of Americans was targeted by Russian information operatives more than African-Americans; race and related issues were the preferred target of the information warfare campaign designed to divide the country in 2016.²⁷
- Also in 2018, Dr. Safiya Umoja Noble, affiliated with Hewlett grantee Data and Society, published her book Algorithms of Oppression: How Search Engines Reinforce Racism. Noble argues that negative biases against women of color are embedded in search engine results and algorithms and that the combination of private interests in promoting certain sites, along with the monopoly status of a relatively small number of Internet search engines, leads to a biased set of search algorithms that privilege whiteness and discriminate against people of color, specifically women of color.
- More recently in 2020, Ruha Benjamin won the Oliver Cromwell Cox Book Award from the American Sociological Association for her book Race after Technology. Similar to Noble, Benjamin seeks to uncover how emerging technologies can reinforce white supremacy and deepen social inequality. Presenting the concept of the “New Jim Code,” she shows how a range of discriminatory designs encode inequity by explicitly amplifying racial hierarchies; by ignoring but thereby replicating social divisions; or by aiming to fix racial bias but ultimately doing quite the opposite.
- Also in 2020, Brandi Collins-Dexter, affiliated with the Shorenstein Center, another Hewlett Grantee, wrote “Canaries in the Coalmine: COVID misinformation in Black Communities.” She describes the report as exposing “how historical oppression, medical mistrust, and healthcare redlining combined with failures by internet platforms and media underreporting have left the Black community with an information vacuum now being filled by dangerous false narratives online. This leaves individuals at great personal risk and imperils democracy by harming Black voters’ ability to be informed on matters of the highest national importance.”²⁸

8 The Hewlett team’s commitment to racial equity has not been well communicated.

Our data collection surfaced a disconnect between the way the Hewlett team and some of its key external stakeholders perceive the foundation’s commitment to racial equity. Although the US Democracy team supports civil rights organizations and scholars who study the relationship between disinformation and race, their written documents do not communicate a coherent vision of the relationship between disinformation, polarization and racial equity.

²⁷ https://www.intelligence.senate.gov/sites/default/files/documents/Report_Volume2.pdf

²⁸ <https://shorensteincenter.org/canaries-in-the-coal-mine/>

Across the US Democracy team's founding documents and updates to the Hewlett Foundation Board of Directors, they define the problem to solve as that of "affective polarization" – "something beyond ideological or policy differences, involving emotional, tribal disdain for – even hatred of – an opposing party." The rationale appears to be that reducing the spread of disinformation can alleviate partisan tensions in the electorate, improve Congress' ability to negotiate and compromise, and therefore strengthen accountability and confidence in liberal democratic institutions. During the 2018-2020 period, among the main questions the team sought to understand was the way disinformation impacts people's political views.

Other funders and many Hewlett grantees come to their interest in disinformation from an explicit assumption that it is a threat to civil rights and social justice.²⁹ We heard from stakeholders who perceive Hewlett to be "out of step with other conversations [because] they don't have as robust a frame of racial justice and harms in terms of democracy." In the words of a Hewlett grantee, this makes partnership with Hewlett a "bit of an uncomfortable fit" for organizations that work on disinformation from the perspective of its intersectionality with race and gender. Another grantee noted that the lack of diversity in the social media companies, and the use of race and immigration as key disinformation tools are among the most important issues that come up in discussions between researchers and policymakers. Several individuals recommended that the foundation strengthen its understanding that race and identity are the main drivers of polarization problems and that social media companies are ill equipped to solve for this in part because disinformation is "a tool of power to act on the powerless."

These perceptions are noteworthy feedback in and of themselves, but also because the US Democracy team invests in civil society organizations and its disinformation portfolio includes grantees that focus on the intersection of equity and technology. We believe the gap stems in part from the contrast between the frequent external communication the Hewlett Foundation and others conducted during the Facebook project (April 2018 -October 2019)³⁰ and the lack of clear, consistent external communication by the foundation team that showcases its own desire and the work of its grantees to advance a deeper understanding of the relationship between disinformation and equity. This is consistent with the fact that most of the key informants we interviewed could not speak to Hewlett's disinformation strategy or its priorities outside of the Facebook-Social Science One-SSRC project.

The lack of consistent external communication is creating a perception that Hewlett's portfolio lacks diversity. Key informants share a perception that Hewlett's portfolio lacks diversity and that Hewlett consolidated a closed, "clubby" network where the funded keep getting funded. We collected relevant secondary data to understand this perception in more detail.

The first perception we heard is that Hewlett supports "a few players" and that these are the same people that get money all the time. A representative quote of this view: "Harvard and Stanford will always get most of the funding. They have great grantmaking apparatus and lots of stars, name recognized scholars. Funders love funding grantees that others have funded."

²⁹ The Ford Foundation and Democracy Fund seek to ensure that "technology advances rather than undermines equity," and to "shift the dynamics of online platforms to support civic participation and safeguard civil rights." They define the public interest and the effectiveness of democratic institutions as necessarily inclusive of racial equity.

³⁰ Scholars from the disinformation and polarization research community were involved with discussions and meetings of Social Science One's twelve subcommittees. The project was known to be Hewlett led, and the high volume of communications about the project (press releases, interviews, tweets, and blogs) likely contributed to perceptions that it was the major thrust of the foundation's disinformation work.

We found that Hewlett's grant portfolio includes public and private universities as well as non-university research and policy institutes. We also refer back to our previous evaluation of the Social Science One-SSRC-Facebook partnership and its frequent press releases and media exposure to highlight the possibility that this perception may be a reflection of the fact that prominent scholars from Harvard and Stanford pursued active external communication as part of their work to create data access and study democracy.

The second perception we heard is that Hewlett's portfolio lacks racial and gender diversity. We collected relevant data to explore this perception, presented in the chart below. As shown in Table 7, Hewlett exceeds parity among genders. Grantees of color have received 29% of funding and 5 out of 23 grants in the portfolio.³¹

TABLE 7. GRANTEE DEMOGRAPHICS

	Number of grants	Percent of grants	Grant dollars	Percent of dollars
Female PI	12	52%	6.510m	64%
Male PI	11	48%	3.625m	36%
Total	23	100%	10.135m	100%
POC PI	5	22%	2.925m	29%
Non-POC PI	18	78%	7.210m	71%
Total	23	100%	10.135m	100%

Where do we go from here?

The Hewlett Foundation's decision to include a focus on disinformation in its 2017 US Democracy strategy was motivated by a desire to figure out how best to encourage social media platforms to help alleviate the polarization exacerbated by disinformation. With the underlying assumptions of the strategy in mind, we use the concluding section here to reflect on the implications of these findings. We received numerous points of advice from Hewlett's stakeholders and use these as a cornerstone of our concluding section *Where do we go from here?*

Implications of our Findings

We identify implications of the lessons learned during these past few years for the foundation to consider as it develops a refreshed disinformation strategy.

1. The foundation's assumption that shifting the behavior of the media platforms is the most pressing outcome to influence through their disinformation strategy is worth revisiting given indications that platforms are beginning to act differently. Companies now acknowledge there is

³¹ In terms of how we collected this data, it's important to note that we used publicly available pictures, interviews and other sources that do not allow the self-definition of gender and / or racial heritage. There may be variation between what we report here and the reality of individuals' preferred gender and or racial identity. These data do not reflect grantees' distribution of sub-grants to scholars or consultants.

a problem with disinformation, and interviewees tell us that was not the case 4 years ago. As one interviewee relayed, “we were saying your recommendation engine is a vast hive of conspiracy theories, they said it was freedom of expression. Not anymore, things have changed.” During the course of this evaluation, platform use of factchecking and content moderation increased with the confluence of the COVID pandemic, the Black Lives Matter movement, and President Trump’s continued use of social media to express his views. In May, Twitter factchecked a Trump tweet for the first time,³² and began labeling tweets that ran afoul of the company’s policy against abusive behavior. When Facebook did not block similar content, Mark Zuckerberg had a near insurrection in his company over what his employees saw as his ambivalence over the platform’s role in spreading racism and violence.³³ In August 2020, Facebook announced a major new partnership to study the 2020 US election with 17 independent academic researchers, including five Hewlett grantees, two of whom (Stroud and Tucker) are the lead external researchers.³⁴

2. The foundation’s assumption that a disinformation field should generate research on both the problem of disinformation and potential policy solutions is worth revisiting. We found that disinformation is better conceived of as a topic of study by different fields and that over-focusing on any of them may unintentionally exclude important perspectives and bodies of work. While perceptions of exclusion and “clubbiness” may be a reflection of weak external communications on the foundation’s part, we believe it is worth considering the unintended consequences of philanthropy seeking to “create” a field of study. The diversity in the foundation’s grantee portfolio might need deeper understanding and investigation on the part of the Hewlett team, as does the perception that there is a fracture among scholars regarding the value of platform data driven research. It seems timely to consider how grantees can best inform the new strategy given that that multi-dimensional solutions to complex problems are likely to come from a proliferation of different ideas rather than a consensus on the best path forward.
3. We were struck by the different paradigms and theories that exist among the philanthropies we interviewed that are committed to supporting US democracy in part by protecting it from the impact of disinformation. Given the importance of this topic, we recommend that Hewlett seize the opportunity to inform their revised strategy with the input of peer funders whose paradigms, priorities and visions of the most pressing threats to democracy might differ. While understandable because philanthropies are all unique, differences in paradigms and priorities may prevent what could be a much more powerful collective approach to spread philanthropic resources across areas of research and action along a shared theory about how to assure US democracy flourishes.
3. The finding that research paradigms on disinformation appear to be either equity neutral or focused explicitly on structural bias and equity suggests the opportunity for the Hewlett team to reflect on whose lives are most impacted by disinformation and why this is of priority interest to

³² <https://twitter.comrealDonaldTrump/status/1265255835124539392> Twitter flagged the tweet with a link to “Get the facts on mail-in ballots.”

³³ One interviewee noted that “Mark Zuckerberg had a near insurrection in company over what is seen as his ambivalence over the platform’s role in spreading racism and violence - that is an accelerant” to change.” Platforms have policies prohibiting hateful or violent content but enforce those policies with discretion. Signs of change are evidence, as already noted with Twitter’s factchecking and flagging of Trump’s tweets and Reddit’s decision to block “The_Donald” on its platform.

³⁴ More detail on the Facebook announcement (Clegg and Nayak 2020) available from Tucker and Stroud (2020).

the Hewlett team. We note the Hewlett Foundation's recent release of its Diversity, Equity and Inclusion commitment as an important reason that is both the "right" and "smart" thing to do moving forward.³⁵ We reference here the two Hewlett commitments that at first blush seem the most pressing:

- Collecting data about the diversity of our grantee pool and pursuing measures to combat the role of implicit and structural biases in our grantee selection
- Ensuring that we have considered the role of diversity, equity, and inclusion in the development of our grantmaking strategies.

Advice from Hewlett stakeholders

Changes in the country and in the disinformation landscape provide a fertile ground for Hewlett to consider whether these implications resonate and how best to move forward. Key stakeholders shared their advice for Hewlett as its US Democracy team considers its future disinformation strategy. We conclude with some of their direct quotes.

INVEST IN SOLUTIONS

"Invest in solutions now, don't wait to research problems fully (they keep changing)"

"Advocate for duty of care – if the platforms are changing constantly and revenue driven, must they be obligated to prevent harms to their users?"

Advocate for consumer rights vis-à-vis technology, "convene interested parties with a constituency to build a movement for the consumer"

"Create new partnerships and advocacy to educate Congress on technical issues, alongside traditional media advocates and organizations and more institutional democracy reform organizations"

"Take a longer view [beyond research] and figure out how policy fits in"

CONTINUE TO BUILD COMMUNITY

"Need to have annual field convening that is bigger than one donor's interests or showcasing particular area;" foundations can convene and produce white paper and books from these

Invite people other than grantees to convenings , "show that doors are open"

"Postdocs are great ways to fund this work, they are cheap and working on ideas that push envelopes more than what senior scholars are, create opportunities for mentorships – put money into someone's career"

Need to link research to a bigger networks of scholars, "looking back over 3-4 years and \$25-50m spent, there are marginally bigger networks"

"Build out networks with civil society actors and think tanks"

"Support study of misinformation overseas, focus on techniques that are hard to translate internationally "

"Engage the newsrooms" - fund data journalism to help explain how disinformation works to lay people, "this might help the public avoid participating in it"

"Give researchers a platform, more context for their work, and introduce them to people on the Hill" set it up like speed dating

Translate research findings for reporters to share, "get understandable info on the local news, these are the trusted voices to many communities"

Seek "working tech-academic intersections" embedding within the firms

³⁵ See the Hewlett Foundation's 2020 website publication of its approach to Diversity, Equity and Inclusion.

PAY ATTENTION TO DRIVERS

"Consider broader questions in law and economics (e.g., incentives, the regulatory environment), less on political science and computational social science"

"Cries about intolerance and calls for openness" are huge opportunities to mobilize now

"Technology governance is the bigger picture," even if disinformation is the problem today

"Focus on the information environment and how to change it"

"Explain the problem in a way that makes it easier to show there are policy roots and policy solutions"

"Data access is the most painful problem, on top of that, the absolute lack of any form of ethical guidelines on how to use it is terrifying"

"Get a better understanding of how information flows, the power of positive and negative memes"

"Shift away from text focus" (despite large investments in getting Facebook and Twitter data by some centers)

"Think about the relationships between the institutions you support and the tech companies, and how to build a more diverse research community that reflect diverse interests"

"Discrimination is central to disinformation. This is an emergency. Break glass."

"We have a president of disinformation – consider how did we get to this point?"

BE MORE DELIBERATE ABOUT DIVERSITY

"What worries me is not just that the field is lacking diversity, but as a result it really warps our understanding of the problem."

"Increase expectations about equity, diversity, and inclusion – bake this into your funder incentives"

Make diverse teams a selection criteria, "where submitting an all-male, all white team works against your chances of funding"

"Fund programs to support more diverse voices within the platforms"

"Understand that race and identity are the main drivers of polarization problems and companies are ill equipped to solve for this"

Improve programs to support researchers at non-elite institutions

Improve programs that support individual scholars working on these questions that may lack resources

Focus on race and COVID and how they have changed the focus for the future; make the "recovery agenda" about addressing social inequalities

"Bring more diversity into the field, the field needs to be seeded internationally... and not just universities"

Understand motivation and impacts on a personal and interpersonal level, "not just through big quant studies"

"Widen your lens beyond disinformation - think more broadly about the types of accountability we need to enforce on technology platforms to deal with hyper targeting and other types of harms beyond disinformation"

"Racial minorities and the global south spread positive messages in ways that other actors use to spread misinformation; consider information spread, not just disinformation spread"

APPENDIX 1. External Literature

Allyn, Bobby. "Reddit Bans The_Donald, Forum Of Nearly 800,000 Trump Fans, Over Abusive Posts." NPR. June 29, 2020. https://www.npr.org/2020/06/29/884819923/reddit-bans-the_donald-forum-of-nearly-800-000-trump-fans-over-abusive-posts

Born, Kelly. "Disinformation Threatens 2020 Election." April 29, 2019. Morning Consult. <https://morningconsult.com/opinions/disinformation-threatens-2020-election/>

Bradshaw, Samantha and Philip Howard, "The Global Disinformation Order: Global Inventory of Organized Social Media Manipulation." Project on Computational Propaganda. 2019. <https://comprop.ox.ac.uk/wp-content/uploads/sites/93/2019/09/CyberTroop-Report19.pdf>

Clegg, Nick and Chaya Nayak. "New Facebook and Instagram Research Initiative to Look at US 2020 Presidential Election." August 31, 2020. Facebook Newsroom. <https://about.fb.com/news/2020/08/research-impact-of-facebook-and-instagram-on-us-election/>

Collins-Dexter, Brandi. "Canaries in the coal mine: COVID-19 Misinformation and Black Communities. Technology and Social Change Project. Harvard Kennedy School Shorenstein Center on Media Politics and Public Policy. <https://shorensteincenter.org/wp-content/uploads/2020/06/Canaries-in-the-Coal-Mine-Shorenstein-Center-June-2020.pdf>

Constine, Josh. "Facebook reveals 25 pages of takedown rules for hate speech and more." TechCrunch, April 24, 2018. <https://techcrunch.com/2018/04/24/facebook-content-rules/>

Facebook Investor Relations Press Releases. <https://investor.fb.com/investor-news/press-release-details/2020/Facebook-Reports-First-Quarter-2020-Results/default.aspx>

Flores-Saviaga, Claudia and Saiph Savage. "Anti-Latinx Computational Propaganda in the United States An analysis of discussions about Latinxs and politics on Reddit prior to the 2018 midterm elections." May 7, 2019. https://www.iftf.org/fileadmin/user_upload/downloads/ourwork/IFTF_Anti_Latinx_comp.prop_W_05.07.19.pdf

Goldzweig, Rafael Schmuziger, Bruno Lupion, and Michael Meyer-Resende. "Social Media Monitoring During Elections: Cases and Best Practice to Inform Electoral Observation Missions." Open Society Foundations. European Policy Institute. 2019. <https://www.opensocietyfoundations.org/uploads/b01958d0-7dce-41d3-94ec-622f992f5efc/social-media-monitoring-during-elections-20190614.pdf>

Institute For The Future. The Human Consequences of Computational Propaganda: Case Studies from the 2018 US Midterm Elections <https://www.iftf.org/disinfoeffects/>

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. The origins and consequences of affective polarization in the United States. Annual Review of Political Science, 22, 129-146. 2019.

Kavanagh, Jennifer, and Michael D. Rich. "Truth Decay: A Threat to Policymaking and Democracy." (2018).

Knight Commission on the Information Needs of Communities in a Democracy, Informing Communities: Sustaining Democracy in the Digital Age, Washington, D.C.: The Aspen Institute, October 2009.

Kornbluh, Karen, and Ellen P. Goodman. "Safeguarding Digital Democracy Digital Innovation and Democracy Initiative Roadmap." Digital Innovation and Democracy Initiative Report. German Marshall Fund. March 2020.

https://www.gmfus.org/sites/default/files/Safeguarding%20Democracy%20against%20Disinformation_v7.pdf

Lazer, David. "Studying human attention on the Internet." Proceedings of the National Academy of Sciences 117.1 (2020): 21-22.

Lazer, David, Alex Pentland, Duncan J. Watts, Sinan Aral, Susan Athey, Noshir Contractor, Deen Freelon, Sandra Gonzalez-Bailon, Gary King, Helen Margetts, Alondra Nelson, Matthew J. Salganik, Markus Strohmaier, Alessandro Vesplignani, and Claudia Wagner
"Computational social science: Obstacles and opportunities." Science 369, no. 6507 (2020): 1060-1062.

Ledger of Harms. "On Systemic Oppression." <https://ledger.humanetech.com/>

McAweeney, Erin, and Palika Makam. "Anti-Immigrant Computational Propaganda in the United States Mapping the flow of image-based alternative narratives in 2018." May 7, 2019.
https://www.iftf.org/fileadmin/user_upload/downloads/ourwork/IFTF_Anti-Immigrant_comp.prop_W_05.07.19.pdf

O'Hara, Amy, and Jodi Nelson. "Evaluation of the Social Science One – Social Science Research Council – Facebook Partnership," December 2019

Social Science Research Council. MediaWell. <https://mediawell.ssrc.org/>

Stoller, Matt. "Congress forced Silicon Valley to answer for its misdeeds. It was a glorious sight." *The Guardian*, July 31, 2020. <https://www.theguardian.com/commentisfree/2020/jul/30/congress-forced-silicon-valley-to-answer-for-its-misdeeds-it-was-a-glorious-sight>

Stroud, Talia, Joshua Tucker, Annie Franco, and Chad Kiewiet de Jonge. "A Proposal for Understanding Social Media's Impact on Elections: Rigorous, Peer-Reviewed Scientific Research." August 31, 2020. Medium. 2020 Election Research Project. https://medium.com/@2020_election_research_project/a-proposal-for-understanding-social-medias-impact-on-elections-4ca5b7aae10.

Swire-Thompson, Briony, Ullrich KH Ecker, Stephan Lewandowsky, and Adam J. Berinsky. "They might be a liar but they're my liar: Source evaluation and the prevalence of misinformation." Political Psychology 41, no. 1 (2020): 21-34.

Trump, Donald J. (@realDonaldTrump). "Mail boxes will be robbed, ballots will be forged & even illegally printed out & fraudulently signed. The Governor of California is sending Ballots to millions of people, anyone....." May 26, 2020 8:17 AM. Tweet.

<https://twitter.comrealDonaldTrump/status/1265255835124539392>

US Senate Select Committee on Intelligence. On Russian Active Measures, Campaigns, and Interference in the 2016 US Election, Volume 2: Russia's Use of Social Media. 116TH Congress, 1st Session.

https://www.intelligence.senate.gov/sites/default/files/documents/Report_Volume2.pdf

Waltzman, Rand, "The Weaponization of Information: The Need for Cognitive Security." Santa Monica, CA: RAND Corporation, 2017. <https://www.rand.org/pubs/testimonies/CT473.html>.

Woolley, Samuel. "Computational Propaganda, Jewish-Americans and the 2018 Midterms: The Amplification of Anti-Semitic Harassment Online." (2018).

<https://www.adl.org/resources/reports/computational-propaganda-jewish-americans-and-the-2018-midterms-the-amplification>

Wu, An Xiao and Harsh Taneja. "Platform Enclosure of Human Behavior and its Measurement: Using Behavioral Trace Data against Platform Episteme," New Media and Society. July 1, 2020.

<https://journals.sagepub.com/doi/10.1177/1461444820933547>

Wu, An Xiao. "How Not to Know Ourselves." Points: Data and Society. July 28, 2020.

<https://points.datasociety.net/how-not-to-know-ourselves-5227c185569>

APPENDIX 2. Key Informants Interviewed

Name	Title	Organization
Yochai Benkler	Professor and Co-director of the Berkman Klein Center for Internet and Society	Harvard University
Adam Berinsky	Professor and Director of the MIT Political Experiments Research Lab	MIT
Kelly Born	Executive Director of the Cyber Policy Center	Stanford University
Renee DiResta	Technical research manager at Stanford Internet Observatory	Stanford University
Sarah Drinkwater	Director, Beneficial Technology	Omidyar
Joan Donovan	Research Director of the Shorenstein Center on Media, Politics and Public Policy	Harvard Kennedy School
Camille Francois	Chief Innovation Officer	Graphika
Deen Freelon	Professor in the School of Media and Journalism	UNC at Chapel Hill
Matthew Gentzkow	Professor of Economics	Stanford University
Daniel Goroff	Vice President	Alfred P. Sloan Foundation
John Sands	Director, Learning & Impact	Knight Foundation
Janet Haven	Executive Director	Data and Society
Philip Howard	Director	Oxford Internet Institute
David Jay	Head of Mobilization	Center for Humane Technology
Jonathan Kartt	Program and Evaluation	Rita Allen Foundation
Jennifer Kavanagh	Director, Strategy, Doctrine, and Resources Program	RAND
Young Mie Kim	Professor of the School of Journalism and Mass Communication	University of Wisconsin, Madison
Karen Kornbluh	Director, Digital Innovation and Democracy Initiative	German Marshall Fund
Larry Kramer	President	Hewlett Foundation
Vidya Krishnamurthy	Director of Communications	Hewlett Foundation
David Lazer	Professor and Co-director of NULab for Texts, Maps, and Networks	Northeastern University
Lillian Mason	Assistant Professor of Government and Politics	University of Maryland
Alice Marwick	Assistant Professor in the Department of Communication	UNC at Chapel Hill
Lori McGlinchey	Director, Technology and Society	Ford Foundation
Dawn Nakagawa	Executive Vice President	Berggruen Institute
Jennifer Nelson	Chief of Staff to the President and Chief Executive Officer	Arnold Ventures
Alondra Nelson	President	Social Science Research Council
Nathaniel Persily	Professor of Law	Stanford University
Nick Pickles	Director, Public Policy Strategy	Twitter
Sarah Ruger	Director of Free Speech Initiatives	Stand Together/Koch Foundation

Laura Rosenberger	Director of the Alliance for Securing Democracy	Alliance for Security Democracy/ German Marshall Fund
Talia Stroud	Professor and Director of the Center for Media Engagement	University of Texas - Austin
Kate Starbird	Professor in Department of Human Centered Design & Engineering	University of Washington
Victoria Smith	Senior Research Analyst at the Partnership for Countering Influence Operations	Carnegie Endowment for International Peace
Daniel Stid	Program Director of U.S. Democracy	Hewlett Foundation
Eli Sugarman	Director of the Cyber Initiative	Hewlett Foundation
Josh Tucker	Professor and Co-director of the NYU Center for Social Media and Politics	NYU
Vanessa Tucker	Program Officer, US Democracy	Hewlett Foundation
Paul Waters	Associate Director	Democracy Fund
An Xiao Wu	Research Affiliate, Berkman Klein Center for Internet and Society	Harvard University
Jonathan Zittrain	Professor	Harvard Law School