A Computational Social Science **that makes sense**

Studying the complex dynamics between
social intraction and collective meaning-making

Anna Keuchenius

How do ideas spread and evolve? In what ways do our social circles shape our thinking? What fuels online polarization?

With access to unprecedented levels of digital data and computational tools, Computational Social Science (CSS) has emerged as a powerful interdisciplinary field for studying social dynamics at scale. A complexity-inspired branch of CSS models phenomena such as polarization, misinformation, and belief propagation through system dynamics. It has proven invaluable for understanding emergent and previously difficult-to-explain shifts in collective social outcomes.

Yet the field often prioritizes structural patterns of social interaction over interpretation: ideas are modeled as spreading like viruses, while social interactions are reduced to quantifiable networks. This thesis builds on a growing body of research that studies the intersection of social interaction and interpretation, advancing the field by integrating insights from relational sociology with complexity-inspired CSS to examine how meaning-making shapes idea diffusion and polarization.

Through four case studies, this thesis highlights the role of key individuals in translating concepts across contexts and demonstrates how conflict and disagreement fuel online polarization. This work underscores how collective meaning-making shapes social dynamics, offering insights that are both academically significant and crucial for addressing pressing societal challenges, including misinformation and polarization.

University of Amsterdam
European Commission | Horizon 2020 European Union funding for Research & Innovation
ODYCCEUS

A Computational Social Science **that makes sense**

Anna Keuchenius

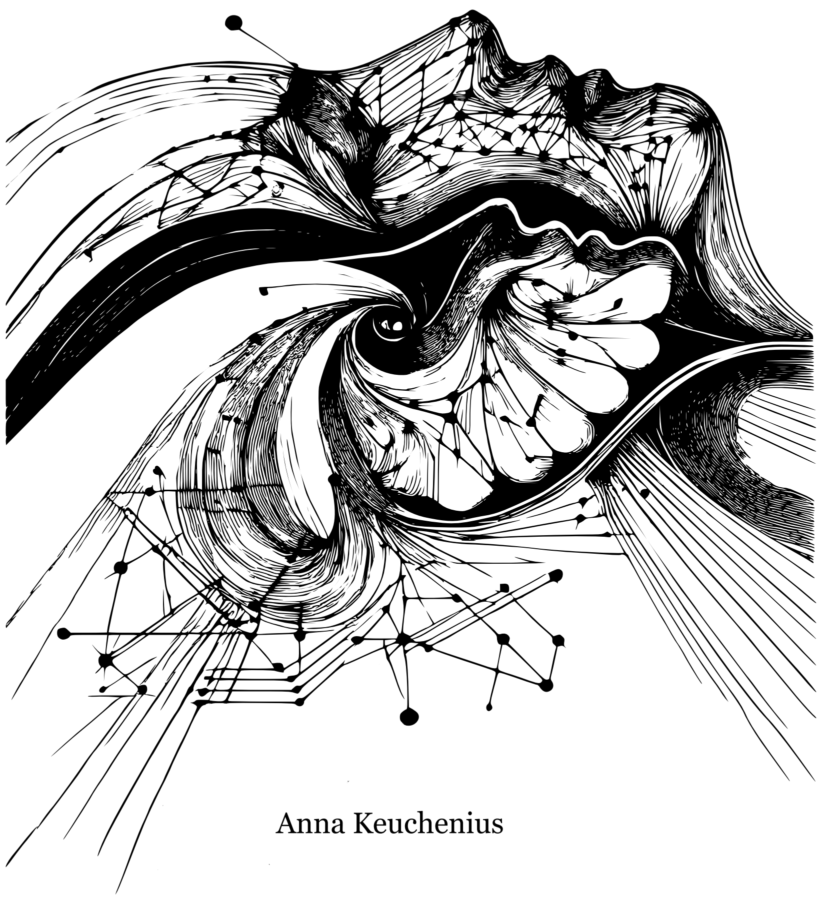
A Computational Social Science **that makes sense**

Studying the complex dynamics between
social interaction and collective meaning-making

Anna Keuchenius

A COMPUTATIONAL SOCIAL SCIENCE THAT MAKES SENSE

Studying the complex dynamics of social interaction and collective meaning-making



Anna Keuchenius

ISBN: 978-94-6473-769-1
Author: A. Keuchenius

This research has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 732942, project ODYCCEUS.

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. Ir. P.P.C.C. Verbeek
ten overstaan van een door het College voor Promoties ingestelde
commissie, in het openbaar te verdedigen in de Agnietenkapel,
op dinsdag 15 April 2025, te 16:00 uur
door Anna Keuchenius, geboren te Purmerend

PROMOTIECOMMISSIE:

Promotor(es):	Prof. dr. J.L Uitermark	University of Amsterdam
Co-promoter(es):	Dr. K.P. Törnberg	University of Amsterdam
Overige leden:	Dr. A. Goldberg	Stanford University
	Dr. A.C. Kroon	University of Amsterdam
	Prof. dr. A. Mears	University of Amsterdam
	Prof. dr. D.C. Trilling	University of Amsterdam
	Dr. D.J. van der Pas	University of Amsterdam
	Prof. dr. E. Leahey	University of Arizona
	Prof. dr. R.L. Breiger	University of Arizona

Faculteit der Maatschappij- en Gedragswetenschappen

TABLE OF CONTENTS

List of Publications

1. Keuchenius, A., Törnberg, P., & Uitermark, J. (2021). Adoption and adaptation: A computational case study of the spread of Granovetter's weak ties hypothesis. *Social Networks*, 66, 10-25.
2. Keuchenius, A., & Mügge, L. (2021). Intersectionality on the go: The diffusion of Black feminist knowledge across disciplinary and geographical borders. *The British Journal of Sociology*, 72(2), 360-378.
3. Keuchenius, A., Törnberg, P., & Uitermark, J. (2021). Why it is important to consider negative ties when studying polarized debates: A signed network analysis of a Dutch cultural controversy on Twitter. *PloS one*, *16*(8), e0256696.
4. Keuchenius, A., Törnberg, P., & Uitermark, J. (2024). Echo Chambers are Defined by Conflict, not Isolation. Conditionally accepted for publication in *Sociological Science*.

In all co-authored articles, Anna Keuchenius was responsible for designing the studies, collecting and analyzing the data, and drafting the majority of the text. The co-authors contributed to refining the theoretical perspectives and the writing of the articles.

Summary

Much of social life takes place in complex interactions that shape collective outcomes, such as the spread of beliefs, the formation of group identities, and societal polarization. Recent advances in computational methods and digitization have enabled unprecedented opportunities to study complex phenomena at a scale and level of detail previously unattainable. These methods have achieved remarkable success in fields like physics and biology, mapping the spread of diseases or neural networks in the brain. Computational Social Science (CSS) has emerged as an interdisciplinary field harnessing digital data and computational tools to study dynamics in the social realm. A prominent branch of CSS, grounded in complexity theory and based on successes in the natural sciences, explores how micro-level interactions give rise to macro-level patterns, examining social phenomena such as polarization, misinformation, and belief propagation through the lens of complex systems. Some scholars in this field aspire for digitization to transform sociology, enabling the precise exploration of social dynamics similar to how the telescope revolutionized physics.

Whilst complexity-inspired CSS yields promising results and opens new research avenues, it has been noticed to prioritize measurable patterns of interaction over the interpretative and sense-making dimensions of human reality: differences in beliefs are modeled to result from differences in people's networks; ideas are conceptualized as spreading like viruses. In an extreme form, social physics compares people explicitly to particles to model system dynamics. This thesis contributes to a growing literature in computational studies that expand the study of social systems to include the significance of meaning-making, focusing on the *intersection of social interaction and meaning-making*. While scholars have begun to explore how associations between beliefs and interpretations emerge and spread, this thesis advances the field by integrating insights from relational sociology with complexity-inspired CSS. Through four case studies, the thesis investigates how sense-making

processes shape idea diffusion and polarization, highlighting the interplay between collective meaning-making and the structural dynamics of social systems.

Objectives and Approach

The overarching goal of this thesis is to integrate the role of meaning-making as a collective process into complexity-inspired computational social science, planting the seed for a *computational social science that makes sense*. This interpretative computational social science would not only excel in understanding the relational dynamics of social processes but also in understanding the role of collective meaning-making in such processes. To illustrate what such research may look like, this thesis offers four original computational case studies that integrate aspects of the role of meaning-making into their research designs. These studies focus on:

1. The diffusion and reinterpretation of ideas within academic communities.
2. The dynamics of conflict and polarization in online social networks.

The thesis demonstrates the value of combining computational tools with an interpretative perspective to reveal the significance of collective sense-making in complex social phenomena.

Key Findings

1. Idea Dissemination in Academia

The first two studies investigate the diffusion of two seminal ideas—Granovetter's *Strength of Weak Ties* and the framework of *Intersectionality*—across academic networks. The research demonstrates that:

- Interpretation Shapes Spread: Ideas are not static entities but are actively interpreted and adapted as they move through communities and these adaptations influence subsequent diffusion pathways.

- Key Translators: Certain scholars act as translators, shaping how ideas are understood and integrated into their communities. Over time, these figures may become symbolic representatives of the ideas, even supplanting their original sources.

These findings illustrate the co-evolution of ideas and the social contexts in which they circulate, highlighting the role of communities in collective sense-making and spreading processes.

2. Polarization in Social Networks

The second set of studies examines the polarized debate surrounding *Zwarte Piet* in the Netherlands, as it played out on Twitter (now named *X*) as a case study. CSS typically analyzes polarization through unsigned network analysis, which aggregates connections without accounting for their interpretative valence. This thesis instead employs signed network analysis distinguishing between positive (agreement) and negative (antagonistic) interactions between users. Key insights include:

- Conflict-Driven Dynamics: Polarization is not merely the result of isolation, as suggested by the echo chamber hypothesis, but is actively fueled by antagonistic interactions between those with differing beliefs.
- Diverse Roles in Conflict: Beyond traditional roles such as hubs or bridges in networks, the signed analysis reveals that some users take structural positions in the network that correspond to other roles in the debate, such as group leaders and scapegoats, shaping the narrative, social coalitions, and divisions in the conflict.

This approach uncovers the central role of conflict in polarized debates, revealing how antagonistic interactions reinforce group cohesion while deepening divides between opposing sides.

Conclusion

This thesis positions itself within the broader movement to study the intersection of meaning and structure in social systems, contributing to the cutting edge of computational social research. It argues that human social life cannot be fully understood without accounting for the interpretive processes through which people collectively construct and contest shared realities.

The findings demonstrate that processes like polarization and the spread of ideas are deeply tied to collective sense-making. Ideas do not merely spread; they are reinterpreted and recontextualized by the communities that adopt them. Additionally, this negotiating of reality may lead to active conflictual polarization, in which expressions of disagreement directed to the outgroup can function to strengthen internal group cohesion.

By integrating relational sociology's understanding of meaning-making and its inseparability from social structure with the analytical tools of CSS, this thesis offers a more comprehensive framework for studying complex social systems. It underscores the centrality of collective sense-making in shaping social dynamics, providing insights that are not only academically significant but also vital for addressing pressing societal challenges, such as misinformation and polarization.

SAMENVATTING

Sociale processen zoals de verspreiding van ideeën, de vorming van groepsidentiteiten en maatschappelijke polarisatie vinden plaats in complexe interacties. Dankzij recente ontwikkelingen in computationele methoden en digitalisering kunnen complexe sociale processen op grotere schaal en met meer precisie worden onderzocht dan voorheen mogelijk was. In vakgebieden zoals de natuurkunde en biologie bewijzen computationele technieken en data analyse hun succes, bijvoorbeeld bij het analyseren van de verspreiding van ziektes of neurale netwerken in de hersenen.

Computational Social Science (CSS) is een interdisciplinair vakgebied dat computationele methoden en de enorme hoeveelheden data van sociale media en andere digitale bronnen gebruikt om sociale dynamieken te bestuderen. Een stroming binnen CSS, geïnspireerd door complexiteitstheorie en voortbouwend op inzichten uit de natuurwetenschappen, onderzoekt hoe micro-interacties leiden tot macro-patronen. Dit perspectief wordt gebruikt om fenomenen zoals polarisatie, desinformatie en de verspreiding van ideeën te analyseren door de bril van complexe systemen. Sommige onderzoekers binnen dit veld zien digitalisering als een potentiële revolutie in de sociale wetenschappen, vergelijkbaar met de impact van de telescoop op de natuurkunde.

Hoewel deze benadering veelbelovende inzichten oplevert, ligt de nadruk vaak op meetbare patronen in sociale interacties, terwijl de rol van interpretatie—de manier waarop mensen ideeën en betekenissen construeren—minder aandacht krijgt. Dit proefschrift maakt deel uit van een bredere wetenschappelijke beweging die interpretatie expliciet meeneemt in de studie van sociale systemen. Dit onderzoek integreert inzichten uit de relationele sociologie met complexiteits-geïnspireerde CSS en verkent hoe interpretatie een rol speelt bij de verspreiding van ideeën en polarisatie. Aan de hand van vier casestudies wordt de wisselwerking tussen collectieve interpretaties en de structurele dynamiek van sociale systemen geanalyseerd.

Doel en aanpak

Het doel van deze dissertatie is om interpretatie als collectief proces te integreren in complexiteits-geïnspireerde computationele sociale wetenschap en zo bij te dragen aan een benadering die niet alleen structurele patronen bestudeert, maar ook begrijpt hoe mensen gezamenlijk betekenissen construeren.

Om dit te onderzoeken, presenteert deze dissertatie vier computationele casestudies waarin interpretatie expliciet wordt meegenomen. De studies richten zich op:

1. De verspreiding en herinterpretatie van academische ideeën.
2. De dynamiek van conflict en polarisatie in online sociale netwerken.

Door computationele methoden te combineren met een interpretatief perspectief, laat dit onderzoek zien hoe collectieve interpretaties en interactie tezamen sociale verandering teweegbrengen.

Belangrijkste bevindingen

1. Verspreiding van ideeën in de academische wereld

De eerste twee casestudies analyseren hoe twee invloedrijke academische ideeën—Granovetter's *Strength of Weak Ties* en het *Intersectionality*-kader—zich binnen academische netwerken verspreiden. Hieruit blijkt:

- Interpretatie beïnvloedt verspreiding: Ideeën worden niet passief overgenomen, maar actief geïnterpreteerd en aangepast wanneer ze in nieuwe contexten terechtkomen. Deze interpretaties bepalen mede hoe ze verder verspreiden.
- Sleutelfiguren als vertalers: Sommige academici spelen een belangrijke rol bij de interpretatie en verspreiding van ideeën. In

bepaalde gevallen worden zij zelfs belangrijker bestempeld voor de doorwerking van een idee dan de oorspronkelijke auteur.

Deze bevindingen laten zien dat ideeën en de sociale context waarin ze circuleren elkaar wederzijds beïnvloeden. Academische gemeenschappen spelen een actieve rol in hoe ideeën worden begrepen en verspreid.

2. Polarisatie in sociale netwerken

De tweede set studies onderzoekt het gepolariseerde debat over Zwarte Piet in Nederland op Twitter (nu X) als casestudy. Traditionele CSS-analyse van polarisatie richt zich vaak op netwerken waarin alleen wordt gemeten of interacties bestaan, zonder te kijken naar de betekenis van de interactie. Dit onderzoek introduceert een *signed network* analysis, waarbij positieve (overeenstemming) en negatieve (antagonistische) interacties worden onderscheiden. Belangrijke inzichten uit dit onderzoek zijn:

- Polarisatie wordt actief versterkt door conflict: In tegenstelling tot de gangbare *echo chamber*-hypothese, die polarisatie verklaart als een gevolg van isolatie, laat dit onderzoek zien dat gepolariseerde netwerken bestaan uit vijandige interacties tussen mensen met tegengestelde overtuigingen.
- Verschillende rollen binnen het conflict: Naast bekende netwerkrollen zoals *hubs* en *bridges*, toont de analyse dat sommige individuen structurele posities innemen die overeenkomen met sociale rollen in het debat, zoals leiders of zondebokken. Deze posities beïnvloeden de manier waarop groepen zich vormen, narratieven worden gevormd en scheidslijnen worden versterkt.

Deze bevindingen laten zien dat polarisatie niet alleen voortkomt uit de afwezigheid van communicatie tussen groepen, maar ook uit de manier waarop groepen zich actief tot elkaar verhouden en uiten. Antagonistische uitingen versterken zowel de interne groepscohesie als de afstand tussen tegengestelde kampen.

Conclusie

Deze dissertatie sluit aan bij een bredere beweging binnen de computationele sociale wetenschappen die zich richt op de relatie tussen interpretatie en sociale structuren. Dit onderzoek laat zien dat processen zoals polarisatie en de verspreiding van ideeën niet volledig begrepen kunnen worden zonder aandacht voor de processen van interpretatie waarmee mensen hun sociale realiteit construeren en betwisten.

De bevindingen tonen aan dat ideeën zich niet simpelweg verspreiden, maar voortdurend opnieuw worden geïnterpreteerd en aangepast, bemiddeld door de gemeenschappen die ze overnemen. Daarnaast suggereren de bevindingen dat polarisatie niet alleen het gevolg is van gescheiden informatienetwerken, maar ook actief wordt gevormd door conflicten, waarbij vijandige interacties met de 'andere groep' de cohesie binnen de eigen groep versterken.

Door de analytische technieken van CSS te combineren met inzichten uit de relationele sociologie, biedt dit proefschrift een bredere benadering voor het bestuderen van complexe sociale systemen. Dit onderzoek laat zien dat interpretatie een centrale rol speelt in sociale dynamieken en levert inzichten die niet alleen wetenschappelijk relevant zijn, maar ook bijdragen aan het begrijpen en aanpakken van maatschappelijke uitdagingen zoals desinformatie en polarisatie.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to ODYCCEUS, the research consortium funded by the EU, for providing the financial support and structures that enabled me to undertake this research, pursue my curiosity, and attend conferences, workshops, and scientific community-building events across the world.

My deepest thanks go to my supervisors. I am especially grateful to Justus Uitermark for offering the best supervision I could have imagined, supporting me during periods of severe disability and illness, and consistently providing ways to continue my research. I also extend my profound gratitude to Petter Törnberg for connecting me to this PhD project and for being one of the quickest thinkers, writers, and most thought-provoking researchers I have encountered and had the privilege to be supervised by.

A special acknowledgment is owed to Javier Garcia, who initiated gatherings for computational scholars within the social science departments of the University of Amsterdam (UvA). Through his innovative ideas—such as hackathons and academic speed dating—he challenged conventional perceptions of how science is conducted, showing that we can do it together and enjoy the process. I would also like to extend my thanks to Damian Trilling, who has been a source of inspiration throughout these years and a model of academic integrity.

I am also deeply appreciative of Eelke Heemskerk, who recognized the creative energy living in Javier, myself, and other computational social scientists and provided instrumental support to establish the Computational Social Science Community, which later became the Computational Social and Behavioral Science Community.

This PhD research project was deeply intertwined with the research practices themselves. I found immense value in collaborating with colleagues from other departments, projects, and disciplines, through which we embodied the interdisciplinary shift that I discuss in this thesis.

Having spent much of my academic career in Science, Technology, Engineering, and Mathematics (STEM) departments, it was an enriching experience to join the sociology department and the wider Amsterdam Institute of Social Science Research (AISSR) which significantly shaped my research. This transition deepened my understanding of science as something inseparable from its authors, as well as the institutional and social contexts in which it is embedded.

I would like to thank Han van der Maas, Joyce Ten Holter, and Peter Sloot for integrating me into the Institute for Advanced Study (IAS). My time at IAS was invaluable, offering experiences that felt akin to a second home. I am also grateful to Patrick Brown for helping me stay connected with the sociology department during my illness and for his continuous trust in my ability to complete my thesis.

A sincere thanks to Jos Bartman, who, early in my PhD journey, helped me realize that the research may never feel perfect, and for guiding me to find joy in embracing this rather than struggling against it.

I am profoundly grateful to the many PhD students and Postdoctoral researchers who have shared this journey with me—Swasti Mishra, David Bokhorst, Nastasja Roels, Phie van Rompu, Fenna Smits, Sander van Haperen, Myra Bosman, Rens Wilderom, Diliara Valeeva, Jonas Haslbeck, Livia van Vliet Teernstra—offering support and celebration throughout. Special thanks to the co-organizers and participants of the Winter School on Complex Systems (2019, 2020, 2021) for the truly interdisciplinary complex systems experiences.

I owe a deep debt of gratitude to my parents for their unconditional support, trust in my decisions, and belief in my abilities throughout this journey.

I would also like to thank Anne-Grietje Franssen, Swasti Mishra, and Roland Bolz, Jean-Babtiste Dupas—dear friends and fellow life travelers—for their invaluable help with the final stages of this PhD. I am honored that Roland edited my thesis and that its cover was shaped with the skill and care of Jean-Babtiste.

To Carolin Goethel, my soul sister, thank you for the weekly peer coaching that empowered me to align with my deepest values and integrate them into my research. I also extend my heartfelt thanks to the Beehive community, MOOS, Imago Space, and the Church of Interbeing in Berlin, where we bring the insights of this thesis to life through gatherings, rituals, and ceremonies, collectively seeking meaning and creating the lives we believe in. Expanding consciousness and care from me, to us, the animate and inanimate.

Finally, I recognize that I am not a fixed entity but a continuous process, evolving through my interactions with those around me. It is in gratitude and honor that I acknowledge the emotional energy, thoughts, and support I have received throughout this journey.

Introduction

Today, social dynamics on a mass scale are reshaping our shared understanding of reality. The digital age facilitates the unrestricted dissemination of information, transcending traditional boundaries and enabling global interconnectedness. Social media platforms are marketed as means to diversify and intensify interpersonal connections. However, rather than fostering cohesion and agreement, prevailing societal narratives underscore a trend to the contrary, highlighting a surge in misinformation, conspiracy theories and polarization. Despite lacking rigorous scientific validation, concepts such as filter bubbles and echo chambers have permeated everyday discourse and are shaping our experience of social life. Echo chambers and filter bubbles point to processes by which small differences between individuals, their beliefs, social network or online behavior, can lead to increasingly divergent and polarized worldviews and distinct camps with conflicting views regarding fact, truth and significance.

At the same time, the proliferation of digital data generated through online interactions has given rise to a burgeoning scientific discipline equipped to dissect mass-scale social processes such as the spread of information, belief propagation, and the increasing polarization observed online. This field, known as computational social science (CSS) (Lazer et al. 2009), harnesses cutting-edge computational advancement and digital traces to explore the patterns and dynamics in complex social phenomena (Bail 2021; Bail et al. 2018; Conte et al. 2012; Ferrara, Cresci, and Luceri 2020; Lazer et al. 2020). Scholars across diverse disciplines, spanning sociology, communication science, computer science, and physics, leverage the enhanced computational capabilities and the abundance of social data for social scientific inquiries at unprecedented scale and detail.

One major branch of CSS specializes in studying emergent effects in collective social behavior: through non-linear processes, the sum of our local actions ultimately leads to substantial shifts and trends within society that are not easily predictable. This system and process view offers a unique contribution that stands in sharp contrast to the dominant variable-based approaches in quantitative social sciences which seek to uncover universal laws governing individual human behavior or societal phenomena through the identification of causal relationships between variables. Typically, the latter approach involves fitting sampled data–often acquired through surveys–into statistical models. However, as the observations of these data are assumed to be independently distributed, this leads to a neglect of the interactions between individuals and the processes taking place between individual micro behavior and collective macro-outcome.

The prominent branch of CSS that is inspired by complexity theory, and which is taken up in this thesis, instead endeavors to shed light on the dynamics of complex social processes, uncovering what happens in that gap between the micro and the macro. Instead of assuming independence between observations, this orientation in computational social science foregrounds the relations between observations, e.g. of individuals, and how these shape the larger system. By employing a diverse toolkit of computational techniques and building on theories from the natural sciences, technology, engineering, and mathematics (STEM), this field offers promising opportunities for understanding how seemingly inconsequential local interactions between individuals can have far-reaching consequences for collective outcomes (Conte et al. 2012; Lazer et al. 2020)–for example: consider how liking or reposting a video online can result in a cascade of viral misinformation.

As an academic discipline, CSS is steadily establishing itself with an increasing number of conferences, educational programs, and journals dedicated to its development (Edelmann et al. 2020; Lazer et al. 2020; Metzler et al. 2016; Törnberg and Uitermark 2021). Beyond the confines of academia, it is widely recognized as a pivotal field with the potential to inform government policies (Conte et al, 2012), attract funding, and shape the future (Lazer et al. 2009; Törnberg and Uitermark 2020).

Its practitioners consider CSS as a revolution in the academic study of complex social processes (Cioffi-Revilla 2017; Hofman et al. 2021; McFarland, Lewis, and Goldberg 2016), offering a promising avenue to influence the course of concerning social phenomena by proposing actionable policy guidance (Cioffi-Revilla 2017; Lazer et al. 2020). CSS is not solely a scientific discipline as it is closely intertwined with industry and government. A seminal publication in *Science* (Lazer et al. 2020) by eminent figures in the field advocates for unprecedented levels of investment from both industry and public sectors, marking a pivotal moment in the history of social science. This call prompts universities to reorganize to foster increased interdisciplinary collaborations within the realm of CSS, ideally culminating in the emergence of colossal research hubs.

This thesis builds on the branch of computational social science that is inspired by complexity theory and that specializes in uncovering system effects in social processes through empirical, simulation and experimental research (Cioffi-Revilla 2010, 2017; Conte et al. 2012; Lazer et al. 2020; Watts 2013). It recognizes this specialization as a groundbreaking advancement in social science, made possible by the recent rise in the accessibility of vast datasets, sophisticated computational tools, and enhanced computing power. Yet, this thesis also identifies a key shortcoming of this field which undermines its potential to fully explain problematic social trends such as the spread of misinformation and polarization, and which stymies its ability to respond to such trends in a way that generates meaningful societal impact. This shortcoming stems from the fact that this branch of research struggles to fully acknowledge and study a crucial aspect of human reality: the role of human interpretation and sense-making, an aspect which profoundly influences our behavior and shapes our understanding of ourselves and the world.

This oversight is due, in part, to the rather high prevalence of technically trained scholars from STEM fields—such as physics, mathematics, and computer science—who in this case have leveraged their methods devised for studying natural systems to widen their purview to include the study of social systems. Advanced methods of complexity research originally developed to study neurons, ants, and flocks of birds are employed to investigate emergent collective behavior of human groups. A popular analogy likens digitalization to the telescope, suggesting that digitalization is to the social sciences what the telescope was to physics (Cioffi-Revilla 2010; Watts 2013). This comparison suggests a similarity of kind between natural systems and human systems. In a more extreme form, "social physics" argues for an explicit comparison of people to particles, as Philip Ball writes in his popular scientific book on critical mass: "to develop a physics of society, we must take a bold step that some might regard as a leap of faith and others as preposterous idealization: particles become people" (2004, p. 110). Not all complexity-inspired computational social scientists adopt this extreme stance, but there appears to be an implicit yet widely held skepticism towards interpretive and qualitative dimensions of data, and a preference to rely on patterns in measurable online traces of behavior.

One of the leading figures and promoters of CSS, Duncan Watts, contributed an extensive piece in the *American Journal of Sociology* (2014) explicitly critiquing the practice of sense-making in social science. Watts highlights the shortcomings associated with relying on individuals' self-reported experiences and motivations as a means to understand behavior. Instead, he advocates prioritizing the extraction and analysis of digital traces, emphasizing their capacity to reveal external and quantifiable determinants that can predict behavior. Because explanation, as Watts argues, is not viable without a level of prediction. He supports this perspective by citing psychological research illustrating how subconscious stimuli can impact behavior without conscious awareness. For instance, one study observed a quantifiable rise in the purchase of German wine when German music was played in the wine store (Watts 2014).

Although it is certainly the case that certain behavioral patterns are better studied through external observation as opposed to self-report, this still leaves the question unanswered as to what role practices of sense-making, interpreting, and representing play in shaping human behavior at the complex systems level. Social scientists and philosophers have long claimed that such interpretive practices are a constitutive dimension of human social life (Berger and Luckmann 1966; Goffman 1956). Philosophical work on complexity and social organization underscores that the study of societal structure cannot be approached in the same way as studying natural complex structures due to ontological, epistemological and relational differences between humans and most other natural entities (Bhaskar 1978; Törnberg 2017). Furthermore, as meaning is ambivalent and humans are narrating and perpetually (re-)negotiating their realities in interaction with each other, processes of meaning-making are contested and may lead to conflictual social dynamics (Collins 1998; Coser 1957).

In contrast to the aforementioned complexity-inspired approaches, computational social scientists that focus on the study of language and culture give insight into the diversity of meanings across times and cultures (Kozlowski et al. 2019). Advancements in text analysis, natural language processing, and more recently, large language models allow for the analysis of the use of words and coevolving meaning at an unprecedented scale and ease, providing insight in social realities such as the identification of biases, inequalities, and polarized discourse (Bail 2016; Bolukbasi et al. 2016; Németh 2023). However, this *content-focused* CSS that delves into the intricacies and effects of language and meaning affords comparatively less attention to the dynamics of social group processes that give rise to these divergent meanings (Bail 2014).

The endeavor of studying precisely *the intersection between meaning and social structure*–i.e. how beliefs, opinions and interpretations emerge from and spread through interactions between individuals, communities and institutions–is at the cutting edge of computational social research.

The mounting research on echo chambers and filter bubbles (Sunstein 1999; Pariser 2011, Bakshy et al. 2015, Barberá et al. 2015a; 2015b, Bail 2018; 2022; Bruns, 2021) is an example of first successful steps towards theorizing and empirically studying this relationship between online social structures and belief formation and propagation. And yet, in their implementation such studies often reduce meaning and beliefs to quantifiable categories or classes, such as sets of opinions or political affiliations such as conservative vs. liberal, considering neither how the individuals in the study might interpret their own beliefs and affiliations, nor the underlying nuanced, fine-grained processes of meaning-making that occur through narratives, storytelling and reconceptualization. Recently, computationally oriented scholars have begun to study individual and collective interpretation from a cognitive scientific perspective and a modeling framework (Goldberg and Singell 2023), offering a complementary, yet distinct approach from this thesis, which advances our understanding of the influence of differing interpretations on the social system.

The overarching goal of this thesis is therefore to integrate the role of meaning-making as a collective process into complexity-inspired computational social science, planting the seed for a *computational social science that makes sense*. This interpretative computational social science would not only excel in understanding the dynamics of social systems and processes but also in understanding the role of collective meaning-making in such processes.

I see a parallel between complexity-inspired CSS and relational sociology, a field that developed in the 1970s. I propose that we can learn from the developments in relational sociology to propel CSS into this more meaningful direction. Like complexity-inspired CSS, relational sociology argued for focusing on processes and relations rather than substantive entities with stable behaviors, introducing a paradigm shift in sociological thinking. Using networks as a primary tool, relational researchers explored how microscopic interactions can lead to macroscopic changes through local interactions (Emirbayer and Goodwin 1994; Granovetter 1973, 1978; Wellman and Berkowitz 1988). However, relational sociology witnessed “a cultural turn” (Breiger 2010; Emirbayer 1997; Fuhse 2009, 2015) during which key relational researchers pointed out that the structuralist approach it had adopted–narrowing down on ties, networks and measurements–risked overlooking the intricate ways that meaning and interpretation play into social dynamics. Consequently, researchers redirected their efforts towards exploring and integrating the meaning-making aspects of human interaction within the relational paradigm (Breiger 2010; Mcfarland and Pals 2005; White, Godart, and Corona 2007). Contemporary research in the tradition of relational sociology emphasizes the mutually constitutive relationship between social structure (network) and meaning, underscoring their inseparability (Fuchs 2001; Fuhse 2009; Pachucki and Breiger 2010).

In the following two sections, I introduce complexity-inspired CSS and relational sociology in more detail. Subsequently, I discuss how CSS may draw from relational sociology as a source of insight which will advance research into the co-evolution of structure and culture that is at play in complex social phenomena. To illustrate what such research may look like, this thesis offers four original computational case studies that integrate aspects of the role of meaning-making into their research designs. The first pair of studies examine the spread of novel ideas and beliefs across networks of individuals and the role of collective sense-making in this process. The second pair of studies is concerned with coalition-forming and communication around polarizing topics on social media. The overarching research motivation behind all four studies is to assess the role of interpretation in a complexity-inspired computational analysis of social processes, which puts it at contrast with studies that overlook or minimize the role of meaning-making in similar contexts. Each case study offers a unique approach and provides its own substantive contribution to the research question addressed in that particular publication.

The four case studies highlight that dominant computational methods for studying the spread of ideas indeed fail to capture important social dynamics that take place during the process of diffusion. The first two studies show that as ideas are spreading, they are adapted and transformed based on the social circles they circulate in. During this process of spreading and transformation, certain actors take on important structural roles, as bridges and brokers in the network, whilst also facilitating the spread through their interpretative efforts translating an idea for a new audience. The second pair of studies incorporate the meaning behind ties in communication networks of polarized debates: are people expressing agreement, disagreement, or even downright hostility? By discerning positive from negative interaction between individuals in networks–a novel endeavor in large scale computational science on polarization–these studies bring to the fore the conflictual nature of controversial topics in society.

Computational Social Science: A Complexity View on Social Phenomena

In the following section I describe the field of CSS in more detail, highlighting its contributions and significance for the study of complex social processes. While I have previously noted a limitation of the dominant complexity-inspired approach–specifically its relative neglect of sense-making and interpretation–this section aims to provide a more comprehensive introduction to the field's merits and also states how the research presented in this thesis builds on literature from this field.

Computational social science (CSS) serves as an overarching term for novel research that capitalizes on enhanced computational capabilities and the vast reservoir of digital data pertaining to social life. This research spans diverse scientific disciplines, ranging from computer science to sociology. CSS is not a mere extension of traditional social science research bolstered by increased data availability and computational power. Instead, its practitioners have described it as a paradigm shift

(Cioffi-Revilla 2017) a watershed moment (McFarland et al. 2016) and a revolution of social sciences (Hofman et al. 2021). The inception of the term "Computational Social Science" can be traced back to a seminal paper published in Science in 2009 by a consortium of 15 researchers, predominantly hailing from STEM fields (Lazer et al. 2009).

Within the broad spectrum of literature encompassed by the label CSS nowadays, this thesis considers the version of CSS that adopts a complexity-oriented approach to examine social phenomena. This branch of CSS investigates how microscopic social actions and interactions aggregate to produce macroscopic outcomes, drawing insights from interdisciplinary research and employing a complex systems lens that specializes in studying phenomena such as emergence and tipping points (Artime and De Domenico 2022). This line of CSS defines itself as follows: "We define CSS as the development and application of computational methods to complex, typically large-scale, human (some-times simulated) behavioral data." (Lazer et al. 2020, p.1060) and "The new field of Computational Social Science can be defined as the interdisciplinary investigation of the social universe on many scales, ranging from individual actors to the largest groupings, through the medium of computation" (Cioffi-Revilla 2017, p.2). In contrast to this complexity-oriented approach to CSS, alternative takes on CSS include studies that focus on advancing any theories of human behavior that leverage computational methodologies on substantial datasets (Edelmann et al. 2020) without this particular micro-to-macro emphasis and thereby give greater emphasis to cultural development through Big Data analysis (Bail 2014), exhibiting a closer connection with the humanities. However, these branches tend to disregard the relational complexities of the social system, diverging from the emergent view adopted by complexity-oriented CSS that is the principle focus of this thesis.

CSS's Departure from Variable-Based Approaches

The revolutionary character ascribed to complexity-inspired CSS stems from its departure from the prevailing quantitative research paradigm in social science, which centers around variables. In this variable-centric world characterized by regression tables, individual behavior is assumed to be governed by intangible variables representing social laws. In this paradigm, theory-informed hypotheses are formulated that revolve around predicting the impact of each variable on the variable under investigation, often referred to as the dependent variable. Data collection, typically via surveys, results in observations (e.g. each observation corresponds to an individual) that are subsequently subjected to statistical analysis. The statistical analysis fits the observed data with a theoretical model which allows the researchers to validate[1] or reject the expectations and hypotheses derived from the theory. For instance, in the context of cooperation, researchers might measure the level of cooperation among individuals in a controlled setting and attempt to explain this through various variables representing influential factors at play, such as the strength of formal institutions and the level of cultural religiosity in the respective society of the subjects (Spadaro et al. 2022).

The foundation of such statistical testing rests on two fundamental assumptions–the well-known i.i.d. requirements: the independence of observations (i.e. no interaction amongst people[2]) and identical distribution (i.e. everyone is subject to the same variable forces). Complexity-inspired CSS strongly deviates from this variable-based paradigm by refraining from assuming independence among observations. Instead, complexity-inspired CSS is inherently interested in the relations between individuals and how these connections may give rise to higher-order patterns. This turns the picture upside down: rather than quantifying the influence of intangible macro-scale forces—referred to as variables in models—that are imagined to operate similarly and independently on individuals, complexity-inspired CSS investigates how relationships and interactions among individuals at a micro-scale may

[1] Strictly speaking, statistical tests never validate any alternative hypothesis, but are merely able to reject a null-hypothesis with a certain level of certainty.
[2] If observations concern entities other than people, such as institutions, it implies a lack of interactions between those entities.

lead to unexpected features at a higher level of organization inspired by earlier complexity research (Axelrod 1997; Conte et al. 2012; Simon 1996; Shelling 1987). To continue with the example of cooperation, complexity-inspired CSS does not strive to quantify the direct influence of variables on the average individual's cooperative tendencies. Its strength lies in studying how cooperation might emerge as a behavioral norm arising from repeated interactions among individuals embedded in social networks (Nowak 2006; Scatà et al. 2016). This branch of CSS thus shifts its focus steadily away from seeking to discern the whole within its constituent parts by attempting to uncover fixed laws governing individual behavior. Instead, it turns attention towards comprehending the collective behavior of the system emerging from numerous and repeated interactions.

Foundational Works for Complexity-inspired CSS in Sociology

This shift in thinking is well illustrated by two foundational yet simple models on collective behavior proposed by Mark Granovetter (1973, 1978). Below, I discuss these models to offer easy-to-follow examples of how to mathematically study complex processes in sociology. Additionally, these models are foundational to the sociological branch of complexity-inspired CSS. They are referenced in subsequent CSS studies and, especially the second model, forms a central part of the literature relevant to the research presented in this thesis.

The first model is a threshold model for collective behavior that was offered to explain the formation of riots, strikes, voting, and migration issues. With the threshold model, Granovetter offers an explanation of paradoxical collective outcomes–that is, group outcomes that are intuitively inconsistent with the intentions of the individuals who generate them. In the threshold model, each individual has a binary decision, e.g. join the riot or not, and a personal threshold that is governing their choice which is dependent on the behavior of others, e.g. how many others are rioting. For example, someone might join a particular riot if at least 10 others are already forming this riot. Granovetter shows that two groups with extremely similar but slightly different threshold distributions may give rise to distinct behavior; in one group everyone ends up joining the riot and in the other group no riot takes place. He illustrates this with an easy-to-understand example of a group with a uniform distribution of thresholds: imagine a group of 100 people in which one person will start the riot irrespective of how many others join (threshold 0), one other person that joins if at least one other is rioting (threshold 1), yet another person that joins if at least two others have joined the riot (threshold 2), etc. until one person that only joins if the riot has size 99 (threshold 99). Now, this group would end up forming a riot of size 100. However, if we would now replace the person with threshold 1 to having threshold 2, there would be no riot; there would be only one person (with threshold 0) rioting alone.

The threshold model shows that averages of two groups might be identical, but slight perturbations influencing the standard deviation can have far reaching and discontinuous effects on the aggregate outcome. This finding, illustrated with a uniform distribution, extends to other distributions, most importantly the normal distribution that is so commonly used in social science. By this, threshold models explain counterintuitive social outcomes as the result of aggregation and the relational dynamics of the situation.

Granovetter introduced a supplementary model that elucidates unforeseen social phenomena by bridging micro-level actions to macro-level outcomes, in a publication titled “The Strength of Weak Ties“ (1973), which became one of the most cited works in the social sciences and has been identified as one of the most cited works that current-day computational social scientists draw upon (Pääkkönen, Nelimarkka, and Reijula 2024). This model centers on the relational configuration among individuals and underscores the significance of tie strength within social networks. Granovetter discerns between weak ties–typified by acquaintances or professional relationships–and strong ties–emblematic of familial or close friendships. Highlighting the broader network overlap engendered by close ties, he empirically and mathematically demonstrates that weak ties often function as bridges between distinct clusters or communities within a network and thereby strongly facilitate the diffusion of information and foster network integration. This was counterintuitive to the then dominant idea that strong ties are most important for network integration and information propagation (Coleman 1964; Jaeckal 1971). This revelation exemplifies yet another paradoxical phenomenon where the aggregation of micro-level social structures and interactions have unexpected emergent outcomes at higher levels.

Whereas Granovetter and his contemporaries used mathematical models and survey data to investigate and elucidate complex emergent patterns in social spreading processes (Coleman, Katz, and Menzel 1957; Crane 1972; Rogers 1983), their ideas have been advanced by computational models that are able to simulate the system under study, thereby laying the groundwork for complexity-inspired CSS as we know it today (Edelmann et al. 2020). Michael Macy and Robert Willer (2002) published a seminal paper titled “From Factors to Actors: computational sociology and agent-based modeling” in which they introduced agent-based models that foreground interactions between agents as an alternative to modeling social processes as interactions amongst variables. They show how these agent-based models (ABMs) can generate familiar but enigmatic global patterns, such as the diffusion of information, behavior or the emergence of norms in sociological context. This was aligned with the general research agenda of the ABM framework that was regarded at that time as the primary field for revitalizing the social and behavioral sciences with foundations rooted in complexity theory (Conte and Paolicci 2014; Bankes et al 2002). Like agent-based models that generate and grow the social phenomena under investigation, a number of network models became highly popular and foundational to complexity-inspired CSS presenting simple processes for growing networks with macroscopic features that are found to be ubiquitous in social networks (Pääkkönen et al. 2024). Examples of such features include the presence of scale-free node distributions, wherein popularity among groups of people is highly uneven (Barabási and Albert, 1999), and the small-world phenomenon, which illustrates that any pair of individuals can be connected through a very limited number of intermediary steps via other individuals (Watts and Strogatz 1998).

When it comes to social spreading processes, these generative models align closely with empirical evidence supporting the concept of social contagion, which elucidates the transmission of beliefs and behaviors among individuals in a manner analogous to viral transmission (Pastor-Satorras and Vespignani 2001). A prominent example is Christakis and Fowler's well-known study on obesity (Christakis and Fowler 2007), which shows the influence of habits and practice that are passed on through social networks contributing to the obesity epidemic. In this social contagion research line, several significant studies were published that served to popularize using agent-based models and network simulations for the study of system effects and the influence of network structures in social spreading phenomena.

Equally in that research line and highly relevant to the progression of the topic of this thesis, Centola and Macy (2007) introduced the concept of complex contagion. They build on Granovetter's suggestion of the Strength of Weak Ties for–what they call–simple contagion but simultaneously recognize that some behavior or information might be costly, risky, or controversial and therefore needs affirmation or reinforcement from multiple sources before being adopted and passed on. Through network simulations, Centola and Macy show that certain network topologies might be very beneficial for the spread of simple information, and yet these same networks turn out not be conducive for the spread of behaviors or opinions that require affirmation from multiple sources. Macy and Centola bring sociological research in conversation with more computationally and mathematically oriented network research by Newman (Girvan and Newman 2006; Newman, Barabási, and Watts, 2000), Watts and Strogatz (1998) and Barábasi (Barabási, Albert, and Jeong 2000; Cohen and Barabási 2002) that explore the influence of network topologies on spreading processes. Through these cross-connections, the mathematical and computer science work within CSS is starting to be integrated with the more sociological literature.

STEM Fields, Data and the Future of CSS

The transition from relying on top-down universal laws, represented by variables dictating micro-level behaviors, to an orientation towards bottom-up mechanisms that elucidate how interacting elements give rise to macro structures, signifies a move towards complexity thinking. Before its development in social science, this complexity perspective has gained solid ground and application supported by suitable methodologies and data in the natural sciences such as biology and physics (Artime and De Domenico 2022). As computational power and digitization rapidly expanded, these fields experienced a surge in research into complex systems and their dynamics. This led to the emergence of thriving areas such as statistical mechanics, computational biology, and computational physics. In addition to the more sociologically embedded CSS described above, a large portion of today's complexity-inspired CSS research is practiced and published in STEM fields without much connection to the social disciplines at all (Cioffi-Revilla 2010, 2017). This research brings in advanced methods of computer simulation, mathematics and data analytics which have demonstrated success in natural fields. STEM researchers are particularly interested in the similarities between social systems to natural systems, leading to propositions such as statistical physics of social dynamics (Castellano 2009) that have shown successful for a number of social systems such as pedestrian flows and other crowd behavior (Cantarella et al. 2014; Helbing, Farkas, and Vicsek 2000). As articulated by social physicist Philip Ball: "Society does not run along the same predictable 'clockwork' lines as the Newtonian universe. It is closer to the kind of complex system that typically preoccupy statistical physicists today: avalanches and granular flows, flocks of birds and fish, networks of interaction in neurology, cell biology and technology" (Ball 2012, ix). One of the defining features of this complexity thinking is that the subject and micro-interactions may be extremely simple yet generate complex patterns. Charged with this insight, humans in this line of research are explicitly compared to ants or worms, and our simplicity is emphasized (Martin 2010; Simon 1996): “Indeed, when we think more clearly, we may realize that the complexity of social life has to be understood as very strong prima facie evidence for our fundamental simplicity. To make this point, let us imagine that we are nematode worms.” (Martin 2010, p.230). As these complexity-inspired computational social scientists turn their attention to social science questions, they hold high expectations for their approach, methodologies, and tools. As previously noted, a limitation of this perspective is its relative neglect for nuanced sense-making and interpretation in human dynamics. The next section on relational sociology will explore the role of meaning and its connection to social structures. Before moving on, however, I will discuss the role of data and the perspectives of this form of CSS on the future of the field.

The growth and evolution of CSS has not solely been catalyzed by the availability of computing power and models for studying system dynamics. It owes much of its impetus to the increasing digitization of social life, a technological development which provides access to a wealth of social data via platforms such as social media sites, blogs, and historical archives—data previously inaccessible to researchers. This Big Data is not only vast in quantity but also possesses a qualitatively distinct nature, often described as revolutionary (Bail 2014; Kitchin 2014; McFarland et al. 2016). Unlike traditional data structured in rows and columns and collected intentionally for research purposes, some computational social scientists consider these data as a treasure trove of "humans in the wild." They are referred to as “digital traces”, “digital breadcrumbs”, or “digital footprints” (Golder and Macy 2014). These data are multifaceted, often relational or transactional in nature, and therefore underscore the interconnectedness between individuals fitting neatly with the complexity lens. Consequently, network analysis stands as a central methodology within CSS (Pääkkönen et al. 2024). The groundbreaking *Science* publication, coining the term "Computational Social Science", first submitted under the title "Life in the Network: The Coming of Age of Computational Social Science," underscores our existence in a networked world (Lazer et al. 2009).

For computational social scientists, the path forward necessarily involves the linking and management of data on an unprecedented scale. Some advocate for the creation of a "social supercollider" (Watts 2013): a facility that integrates diverse data streams, crafting more nuanced portraits of individual behavior and identity while retaining the advantages of massive scale. This envisions the establishment of extensive virtual laboratories (Watts 2013) where experiments can transpire, employing wearable "sociometers" that measure face-to-face interactions between people (Choudhury and Pentland 2003) or mining other social sensor data (Zhang et al. 2020) to track human behavior. In 2020, a consortium of 15 leading researchers, sharing a significant resemblance to the cohort that authored the 2009 article coining the term Computational Social Science (Lazer et al. 2009), published a *Science* article addressing the challenges and opportunities within CSS (Lazer et al. 2020). One of the pivotal hurdles and, consequently, recommendations of the consortium is to radically enhance the data infrastructure paradigm, encompassing both technical and ethical dimensions. This transformation demands a relatively unprecedented investment from public and private sources.

In summary, complexity-inspired CSS provides a relational lens through which to examine social phenomena, accompanied by an expertise in studying relational processes and system effects. CSS harnesses the immense volume of digitized data on human behavior in the wild and today's available computational power. Given its success in other scientific domains, scholars within complexity-inspired CSS hold great expectations for the efficacy of their methods in addressing significant social scientific questions.

Relational Sociology: Investigating the Link between Social Networks and Meaning-Making

The following section maps the field of relational sociology in more detail, highlighting its similarities and differences with complexity-inspired CSS in the study of complex social dynamics. While both fields center on relations and processes and correspondingly utilize network analysis as a key tool, relational sociology places significant emphasis on meaning-making. I discuss how relational sociology investigates the interplay between social structure and meaning, providing insights that will inform how meaning can and is starting to be integrated into complexity-inspired CSS, as discussed in the subsequent sections.

In the late 1980s and early 1990s, a current of sociological research emerged which, similarly to complexity-inspired CSS, focused on relations and processes of the social world, and which defined itself in opposition to the variable-based universe (Abbott 1988; Fuhse 2020)– a tenet that some referred to as the "anti-categorical imperative" (Emirbayer and Goodwin 1994). This relational sociological perspective highlights our tendency to perceive the world in terms of fixed entities, such as when we say "the river flows" rather than acknowledging the ongoing process of water flowing (Emirbayer 1997). However, it challenges these fundamental assumptions about a reality based on stable entities, proposing instead that reality is composed solely of continuous processes.

The central method of relational sociology is social network analysis, which had just started maturing in that time and proving itself a powerful alternative to variable based research and offering a strong new paradigm with far-reaching potential (Wellman, 1988; Mische, 2011): Granovetter's work in social network analysis had shown the importance of weak ties for the spread of information, cultural cohesion (1972) and the general patterns of collective action (1992); Burt's introduction of structural holes as providing competitive advantage lead to an explosion of interest in business and economics (1992, 2004b, 2004a); Wellman's efforts to make an explicit turn to relational structures by taking the social relation as unit of analysis in the study of support networks, communities and social structures in the city, catalyzed a new direction in geography and shaped social network analysis as a field (1983; Wellman and Berkowitz 1988; Wellman and Wortley 1990).

However, although network analysis was described by its practitioners as a form relational and processual thinking, it was quickly critiqued by interpretative scholars and relational sociologists for being overly structuralist, or even substantialist about the relations in the network (Emirbayer 1997). Many social network studies did not delve into the meaning of the ties in the network under examination, but instead considered the ties as external and stable to the relational processes under study (Mische 2011, Pachucki and Breiger 2010). This meant that there remained a sizeable gap between formal network analysis and more interpretively oriented cultural research: "Most cultural theorists saw network analysis as located squarely in the positivist camp, reducing cultural richness to 1s and 0s and lacking attention to processes of interpretation and meaning-construction." (Mische, 2011, p.81). This dichotomy between networks that map social relations and interactions on the one hand, and culture–broadly conceived of as meaning-making and interpretation–on the other hand became the focus area of relational sociologists of that time. Emirbayer (1997) identifies this as a cultural turn in relational sociology, a watershed moment after which relational sociologists turned their attention to the meaning-structure of social networks.

Mische (2011) delineates four primary ways the link between *networks* and *culture* has been conceptualized by relational sociologists and their broader intellectual fields. In the following section, I will provide a brief overview of these conceptualizations, highlighting key research findings relevant to this thesis topic on spreading processes in social networks. Additionally, I will draw insights from these bodies of research that may inform a complexity-inspired CSS approach that is attuned to the nuances of meaning in the relationship between networks and culture.

Subsequently, I will present in more detail the call of Ronald Breiger, a relational sociologist who advocates for transcending the duality of culture and relational structure. To aid this transcending, I present the perspectives of Randall Collins–whose work is also deeply relational and provides a thorough theoretical foundation for the emergence of meaning in social groups through ritual, shared emotional energy, agreement and opposition. These discussions aim to provide insights that may illustrate potential pathways for the advancement of computational social science that is attentive to the complex processes of meaning-making.

Four Approaches to the Link between Networks and Culture

The first approach identified by Mische (2011) for how relational sociologists have conceptualized the link between networks and culture, is by seeing social networks as conduits for social influence. Network ties are here viewed as pipelines along which cultural artefacts such as ideas, attitudes, practices and behaviors can flow. This body of research includes renowned theories such as the two-step flow of information introduced by Katz (1957), according to which information first spreads from mass media to opinion influencers, and then subsequently spreads from these opinion leaders across their networks. This view exemplifies that social network structures facilitate or obstruct the spread of cultural transmission. This perspective demonstrates how social network structures can either facilitate or hinder the spread of cultural transmission. However, these approaches treat cultural elements as entirely external to the networks, a viewpoint similar to that of typical complexity-inspired CSS studies. As a result, this perspective may not be helpful to address the core interest of this research, which is to integrate the influence of collective meaning-making in studying social processes.

A second body of relational sociology considers culture itself as a network of cultural forms, made up of associative relations between nodes which can represent concepts, categories, or practices. For example, Kathleen Carley has been a pioneer in cognitive mapping and extracting mental models from cultural texts (Carley and Kaufer 1993; Carley and Palmquist

1992), demonstrating how “the meaning of a concept for an individual is embedded in its relationship to other concepts in the individual’s mental model” (Carley and Palmquist 1992, p. 602). Zooming out, away from the individual’s mental model towards a meaning system, Andrew Abbott studies patterns in cultural change based on principles of chaos theory and complexity science, and discovers fractal structures, trajectories, bifurcations and turning points in historical processes of cultural evolution that he defines as “a network in time” (2001, 2010). Abbot’s perspective suggests that apart from the individual’s sense-making apparatus, there might be logics inherent in the socio-cultural system of collective sense-making. These two insights, that the meaning of anything is embedded in relation to other things, and that there might be socio-cultural system dynamics that govern individual pathways, are valuable for this thesis’ ambition to include meaning-making into CSS.

As a mirror image of seeing culture as a network, a third branch of relational sociology sees the social network as a cultural construct, constituted by cultural processes of communicative interaction. In this view, ties in networks are merely social constructs, representing a vastly more complex social relation (Fuhse 2009, 2015). White (1992) posits that ties are nothing but the narratives around the relationships, stories which define a social tie by their narratives of ties which is constructed in an attempt to control and form identities. McLean (1998), strongly influenced by Ervin Goffman (1956), shows that both selves and relations are discursively constructed by patronage seekers in Renaissance Florence, where these seekers “key” certain dimensions of their relationships in order for their social network to provide material and social rewards. This body of relational sociology informs us that a relationship does not exist–has no ontological basis–without an interpreter constructing that relationship.

The fourth approach relational sociologists take to study the link between social networks and culture is by seeing networks as shaping culture and vice versa. Research within this approach includes explorations of how network intersections and bridges in networks create coalitions that foster cultural resources and facilitate cultural innovations (Burt, 1992; Mische, 2011). Burt's famous work on innovation shows that innovations are often sparked by some good ideas that are "borrowed" from another network cluster, able to travel through the structural holes (Burt 2004a, 2004b). This approach also includes research on network clusters as incubators of culture exemplified by studies like those of Friedman and McAdam (1992) revealing the presence of strong preexisting social ties in successful social movement mobilization in which actors coalesce around a shared vision. This approach helps us to understand the connections between measurable social relations and the intangible understandings of cultural artefacts without fixating on either one of the two, which may be of value to a complexity-inspired CSS sensitive to meaning. Yet, we also note that in this work, social networks and culture are regarded as separate variables that may influence each other but are fundamentally distinct dimensions of social life (Mische, 2011), a view that has been questioned in later developments of relation sociology.

The Co-construction of Social Networks and Meaning

The angle taken by contemporary relational sociologists is to emphasize the co-evolution and co-construction of social networks and meaning. Breiger advocates to go beyond the duality between structure–signifying social networks–and culture–signifying meanings, local practices, discourse, repertoires, and norms. He argues that "contemporary work on culture (commonly instantiated by, e.g., meanings, local practices, discourse, repertoires, and norms) and social networks (often operationalized by dyadic social ties, homophily, actor nodes, dual networks of persons and groups, and social position) can for important purposes be usefully seen as mutually constitutive and coevolving with common roots in relational thinking." (Breiger and Pachucki, 2010, p. 206). Breiger and Pachucki (2010) introduce the concept of cultural holes, the equivalent of structural holes. Whereas structural holes are structures in the social network that facilitate the flow of information and hence, innovation, Breiger and Pachucki use the term cultural holes to identify contingencies in meaning, practice and discourse that enable, in turn, social structures. They argue that the innovation that was assigned to structural bridging by previous literature, such as in Burt's well-known studies, might have been for a large part caused by cultural holes instead.

A crucial step, I propose, in transitioning beyond the duality between social structure and meaning, is to challenge the implicit assumption that ideas and cultural understandings reside within the individual, imagined as a mental model or stored within patterns in the brain, an assumption that fits neatly in today Western culture's focus on the individual (Lasch 2019). Instead, we may move to an alternative outlook that views meaning as arising only in communication with others, living in the intersubjective (Leydesdorff 2021). Randall Collins brings in this perspective explicitly and presents a novel, yet comprehensive theory of how meaning emerges, transforms, and travels in his work on the history of philosophy, ritual and conflict (Collins 1998, 2004).

Like foundational relational sociologists, Collins holds a radical focus on relations instead of on stable entities. He argues that we should see individuals as "transient fluxes charged up by situations" (Collins, 2004, p. 6), and hence should not ascribe agency to individuals, but to the energy that is appearing in human bodies, that arises in local interactions and that may be charged up from past experiences. With this progressive view, he makes redundant the agency-structure debate that was central in those days, and instead brings attention to the micro-macro dynamics that appear through relationships, emotional energy and ritual. Meaning, according to Collins, is formed through Interaction Rituals (Goffman 1967), which are moments when a group of people come together and mutually focus emotional and cognitive attention producing a momentarily shared reality, generating solidarity and symbols of group membership. Repeated Interaction Rituals–what he calls Iteration Ritual Chains–are needed for meanings and values to stay alive, as otherwise these symbols quickly lose actionable potential and become dead empty vessels.

With this processual lens, Collins views ideas not as static entities, nor as a pattern existing in the human brains. Instead, he perceives ideas as existing in the process of communication between one thinker and another. “Thinkers do not antedate communication and the communicative process creates the thinkers as nodes of the process” (1998, p. 5). As a result, ideas are not shaped within the individual, nor does creativity reside within one human, instead it is the structures of intellectual networks that shape ideas. In *The Sociology of Philosophies* (1998), Collins explores the history of large intellectual and philosophical developments and shows that the history of philosophy is to a considerable extent the history of groups. Creativity flourishes and schools of thought are formed by the evolution of social structures. Conferences and scientific meetings are the interaction rituals of the scientific world. The social structure can be summoned as “an ongoing struggle among chains of persons charged up with emotional energy and cultural capital to fill a small number of centers of attention” (Collins, 1988, p. 14). Within this dynamic, a select few emerge as focal points of focus, their prominence shifting periodically over time. The revered philosopher, often celebrated as an individual genius, is merely a symbolic representation of the social group and its interconnected network across space and time. For instance, the prominent figure of Hegel serves as a symbolic representation for the collaborative Jena-Weimar creative circle, comprising at least thirty individuals (Goodman, Theory, and Mar 2009).

One of the corollaries to Collins theory describing the formation of group solidarity and limited attention spaces is that there emerges an unavoidable opposition and conflict between schools of thought. For this reason, the history of philosophy, he contends, is the history not so much of problems solved as of the discovery of exploitable lines of opposition.

To sum up, relational sociology studies the relations between social structure and collective meaning-making. Different strands of relational sociology provide various perspectives and cues—some of which may conflict with one another—that present both opportunities and challenges for developing a complexity-inspired CSS approach that is sensitive to meaning-making processes:

1. Meaning is embedded: no idea, belief, symbol, word or practice has a meaning that is independent from other concepts, symbols or practices.
2. The development of meaning can be perceived as a socio-cultural system, whose dynamic adheres to particular principles of complexity and chaos theory.
3. Relationships within a social network can be seen as inherently narrative social constructs that lack any ontological status without an interpreter to define them.
4. There is an interaction between the social structure and the meanings that develop, where both mutually influence each other.
5. Beyond an interaction, the social and cultural may be perceived as co-constituting one another and co-evolving.
6. Opposition and conflict are central to the continuous movement of thought and social relationships.

In the studies conducted for this thesis, we primarily follow the first and last two cues, in line with Collins and contemporary relational sociologists, considering social structure and the meanings that develop as mutually constitutive and co-evolving. For the purposes of this research, meanings and social structures may be considered separately to better demonstrate their interrelations. Additionally, the final two papers in this thesis specifically explore the role of conflict in the development of thought and the formation of social coalitions, particularly in the context of polarization.

A Computational Social Science that Makes Sense

Complexity-inspired CSS shares with relational sociology a relational outlook on the world, which focuses on processes over variables and foregrounds the way micro-interactions often have global consequences. Accordingly, one of the dominant methodologies of both fields is network analysis. Both draw in part on the same literature, such as the aforementioned classical contributions of Granovetter and Burt.

What complexity-inspired CSS should take away from relational sociology, I argue, are the different ways that meaning and interpretation play into complex social processes. Although complexity-inspired computational social scientists generally pay attention to the relationship between social structure and the spread and distribution of beliefs and behaviors, such as political preferences, voting, and adoption of innovations, their studies often lack the attention for intricate processes of collective meaning-making that take place through narrative practices (Christakis and Fowler 2007). The conventional approach in complexity-inspired CSS studies takes a network-as-pipelines perspective or social contagion perspective, in which the ties are seen as pipes along which a stable diffusant travel or influences is passed, like a virus that spreads. This perspective, whether explicitly or implicitly, considers humans as similar to ants, cells, neurons or virtual agents and the diffusants as stable entities. But as we have seen, meaning-making is both inherently interpersonal and constitutive of the human social fabric. And meaning in this emphatic sense does not simply reside in the individual's brain but emerges in communication as a feature of the collective field, upheld by groups and networks (Collins, 1987).

Relational sociology offers various perspectives on the link between meaning and networks that could be helpful to aid a computational social science attentive to meaning. Some of these perspectives are already being advanced today by computationally oriented scholars that show the importance of meaning-making using computationally driven models and data analysis. As mentioned, Pachucki and Breiger (2010) advocate to see through the duality of structure and culture, introducing the concept of cultural holes as the counterpart to structural holes. Bail (2013) empirically investigates this cultural boundary spanning in the healthcare sector and finds that bridging a structural hole is most valuable when network clusters are not already rich in heterogeneous knowledge, highlighting the importance of the cultural dimension of the bridging position. In another large empirical study of health advocacy conversations on Facebook, Bail (2016) shows that “organizations which create substantial cultural bridges provoke 2.52 times more comments about their messages from new social media users than those that do not” (p. 11823). Whilst controlling other factors, such as characteristic of the organization and audiences, it is shown that when organizations bring themes within public conversation together that are usually discussed separately–bridging cultural holes–this generates more interest from new social media audiences–possibly bridging structural holes.

Goldberg and Stein present a model that shows the significance of the associations between beliefs or behaviors. They explicitly depart from the social contagion lens– “culture doesn’t spread like a virus” (2018, p. 903)– by introducing a model in which *associations* diffuse through social relations, rather than the beliefs or behaviors themselves. This work shows that cultural heterogeneity–social circles of individuals holding different world views–can come about by associative diffusion and doesn’t merely result from heterogeneities in the social structure (Axelrod, 1997). Recent work of Goldberg and Singell explicitly delves into the question of meaning and proposes a practical perspective in which meaning implies an actor doing meaning, similar to an actor interpreting a stimulus (Goldberg and Singell 2023). Aided by this definition, they offer a model of collective meaning-making as a process in which interpretations are coordinated interpersonally.

Törnberg and Uitermark (2021), my supervisors for this PhD thesis, emphasize the potential of CSS to develop methods that can facilitate interpretation rather than position itself as an external observer that solely measures meaning. They base their work on critical realism (Byrne 1988; Collier 1994) and digital media studies (Couldry and Hepp 2018), both emphasizing that the social world is something accessible to interpretation, and built up, in part, through those interpretations. Therefore, they argue for what they call a heterodox CSS that is critical, pluralist, interpretative and explanative in the sense that it supports critical assessment of social data, allowing for a variety of interpretations to explain and question societal trends.

The Research Presented in this Thesis

This thesis closely aligns with literature presented above, yet advances the insights of relational sociology by presenting four computational social science case studies that each examine the role of meaning-making in a network-driven research design. The overarching research motivation behind all four studies is to assess the significance of incorporating the role of interpretation into the analyses, compared to studies that de-emphasize the role of meaning-making in similar contexts. Each case study offers a unique approach and provides its own substantive contribution to the research question addressed in that particular publication.

The first two studies investigate how a novel idea spread across academia by combining methods from complexity-inspired CSS and interpretative sociology. We picked the academic context to study the spread of novel ideas because references enable the close tracking of the dissemination and recombination with other ideas and narratives. Computational social scientists have recently coined the term “the Science of Science” (Fortunato et al. 2018; Zeng et al. 2017) for computational studies of the spread of scientific ideas. In doing so, they have often made comparisons to the spread of viruses and deployed epidemic models or social contagion views examining how ideas spread through networks of scientists. Conversely, earlier, more interpretative sociological literature on science underscored the significance of conceptual frameworks, paradigms, and shared traditions and culture in fostering the emergence and diffusion of new ideas (Knorr Cetina 1999; Kuhn 1962; Latour 1984). The studies presented in this thesis combine both views by mapping the socio-structural spreading patterns using computational social science methods while also examining how the novel idea is narrated, interpreted and re-interpreted. By integrating insights from the mapped spread with the variety of interpretations, we find that a novel idea is adapted during its diffusion and that research communities function as translation environments, shaping the idea's development in specific directions unique to each community. We observe that certain researchers play pivotal roles in this process, narrated by their communities to be leading the translation work.

The latter two studies concern the study of coalition forming and user-to-user interaction in polarized topics on social media. Complexity-inspired CSS has made strides in studying online polarization and echo chambers in contentious public debates, leveraging large-scale network analysis and natural language processing. However, these studies typically consider only positive ties between individuals, overlooking the crucial role of the valence of interpretations of interactions, thereby also overlooking the role of negative and even antagonistic user-to-user communication. From a relational sociological standpoint, the meaning of (positive and negative) ties is crucial to understanding the social dynamic at play, as conflict lies at the heart of polarized debate (Collins 2012; Coser 1957; Simmel 1904a; 1904b; 1904c). In the two studies presented in the second half of this thesis, we examine a polarized debate using social network analysis, but also integrate sociological views of conflict and interpretation into the research design. Specifically, we explore the debate surrounding Zwarte Piet in the Netherlands from 2017 till 2019, viewed as an innocent children's figure by some, and as a racist colonial legacy by others. In our analysis of this debate, we discriminate between the positive and negative charge of interactions between individuals and reveal a significant amount of antagonism between the sides in this debate. This challenges a leading hypothesis in CSS polarization studies, which poses that a lack of communication between different groups leads to diverging views, which in turn leads to more isolation, giving rise to echo chambers (Pariser 2011; Sunstein 1999). Instead, we reveal a significant level of antagonistic expressions between the two sides of the debate, suggesting that echo chambers are fueled by conflict rather than isolation. Our signed analysis of this debate also sheds light on different roles within this polarized discourse beyond the commonly identified hubs and bridges in networks. These roles include scapegoats–individuals receiving substantial negative expression from the opposing side of the debate without the same level of positive support from their side.

This thesis makes the following theoretical and methodological contributions: theoretically, it makes the case for an interpretative CSS. Methodologically, it presents practical examples of integrating this interpretative approach with complexity-inspired CSS through a combination of methods and perspectives. The empirical evidence of the computational case studies conducted supports the added value of this interpretative approach and generates two primary substantive contributions. First, the first pair of studies finds that the same idea is developed in divergent directions co-evolving with the social circles it diffuses into, thereby underscoring the roles of community and local meaning in the spread of beliefs and ideas in social networks. Second, the latter pair of studies reveal the central role of conflict in online polarization, thereby highlighting the relevance of including negative ties into the research design for studying polarization.

CHAPTER 1.

Adoption and Adaptation: A Computational Case Study of the Spread of Granovetter's Weak Ties Hypothesis

Published as

Keuchenius, A., Törnberg, P., & Uitermark, J. (2021). Adoption and adaptation: A computational case study of the spread of Granovetter's weak ties hypothesis. *Social Networks, 66*, 10-25

Abstract

How do new scientific ideas diffuse? Computational studies reveal how network structures facilitate or obstruct diffusion; qualitative studies demonstrate that diffusion entails the continuous translation and transformation of ideas. This article bridges these computational and qualitative approaches to study diffusion as a complex process of continuous adaptation. As a case study, we analyze the spread of Granovetter's Strength of Weak Ties hypothesis, published in American Journal of Sociology in 1973. Through network analysis, topic modeling and a close reading of a diffusion network created using Web of Science data, we study how different communities in this network interpret and develop Granovetter's hypothesis in distinct ways. We further trace how these communities originate, merge and split, and examine how central scholars emerge as community leaders or brokers in the diffusion process.

Keywords— diffusion, translation, complex networks, meaning, scientific communities

Introduction

In the 1960s and 1970s, the question of how new scientific ideas diffuse was high on the agenda of science studies. Primarily using survey methods, researchers at the time discovered some key dynamics in the spread of ideas. They found that the diffusion of a scientific idea bears similarities to the diffusion of other types of innovation, for example, in that both follow an S-shaped growth curve (Crane, 1972; Holton, 1962; Price, 1963; Mulkay et al., 1975). Research of this time also brought attention to the role of interpretation in science: studies revealed the central role of informal communities—sometimes called "invisible colleges" (Crane, 1972) or "coherent groups" (Griffith and Mullins, 1972)—in the organization of scientific research.

Such communities develop separate vocabularies and narratives through which their members interpret scientific findings (Fisher, 1987). While science studies of the 1960s and 1970s opened a new field of research, scholars faced limitations in their data and methods.

An explosive development in the availability of both data and sophisticated analytical techniques since the 2000s has reinvigorated the field of science studies, allowing researchers to study the development of science at scale (Fortunato et al., 2018; Zeng et al., 2017). But computational analyses come with their own sets of research questions, since they focus on the structural properties of scientific networks while leaving the interpretative work to more qualitative researchers (Pachucki and Breiger, 2010). Combining computational and interpretative analyses in this article, we contend, can help reveal how scientific ideas spread and change in the process of diffusion. This takes us away from what Latour (1984) calls a "diffusion model" of science, in which researchers are passive nodes in a network through which ideas circulate, to what he calls a "translation model," according to which researchers shape the idea to their different projects, resulting in a continuous transformation of the diffusant.

To enable an in-depth and systematic study of how ideas change as they diffuse, we focus on a single idea that has diffused far and wide in academia: Granovetter's (1973) Strength of Weak Ties hypothesis, published in *American Journal of Sociology*. We employ citation network analysis, topic modeling and close reading to study the way this scientific idea was transformed during its spread as a result of the collective behavior and interpretations of scholars. First, we trace the structural spread of Granovetter's hypothesis and analyze its macroscopic patterns using a network representation of citation data. Next, we examine how different communities in this diffusion network developed specific interpretations of Granovetter's hypothesis and focus on the role of individual scholars in this process.

Our work advances the literature in three ways. Theoretically, we develop the notion of a diffusion network and conceptualize how scientific innovations are variably adapted throughout their growth trajectory. Our methodological contribution is to develop an approach that bridges the gap between computational analysis of network properties and the interpretative analysis of meaning (cf. Fuhse, 2009; Pachucki and Breiger, 2010). Finally, our substantive contribution is to show that the spread of scientific ideas entails a complex process of translation in which scholarly communities emerge as meso-level mediators, cultivating divergent interpretations of the diffusing idea in line with the different research projects in which they are engaged. During this process, some scholars—brokers and leaders—perform key roles in translating and introducing the new scientific idea into their circles and across academic boundaries.

The structure of our argument is as follows. The next section outlines the gap between computational and interpretative approaches and suggests how these two types of studies might be combined into an interpretative computational approach. The subsequent section summarizes our methods and explains how we used citation and publication metadata to create a diffusion network. The following three sections analyze: (1) the community structure of this network; (2) the interpretative function of these communities; and (3) the evolution of communities over time, spurred by leading academics with different roles in the diffusion process. The concluding section discusses the implications of our case study for the diffusion of science.

Perspectives on the Diffusion of Science

The groundwork for the study of scientific diffusion was laid in the 1960s and 1970s by scholars, such as Diane Crane (1972), William Goffman (1966), Belver Griffith and Nicholas Mullins (1972), Robert Merton

(1968), Michael Mulkay (1974), Derek J. de Solla Price (1963), and Henry Small (1974). Their research demonstrates that academics are organized in communities[3] that perform pivotal functions in the diffusion and development of ideas. Scholarly communities tend to be organized around one or several academic stars whose status is reinforced through mechanisms of cumulative advantage (Merton, 1968; Newman, 2009; Price, 1976). These star researchers function in their scholarly circles similarly to how opinion leaders function in marketing: recognized as intellectual leaders by the community, they serve as its representatives to the broader scientific world (Collins, 1983; Crane, 1972; Griffith and Mullins, 1972; Price, 1963). This parallel between academic stars and opinion leaders in marketing is in part inspired by Everett M. Rogers's (1983) diffusion of innovations theory and Elihu Katz's concept of the two-step flow of communication, which posits that innovations first spread to opinion leaders, who in turn spread them to consumers (Coleman et al., 1957; Katz, 1957).

While these early science scholars are often credited with formalizing the study of science through the development of mathematical models (Goffman, 1966; Goffman and Newill, 1964; Merton, 1968; Price, 1976), their work contains both quantitative and interpretative insights by addressing the co-evolution of scholarly networks and scholarly cultures. Since then, the study of diffusion has bifurcated. On the one hand, computational scholars have leveraged the explosive growth in the availability of data and sophisticated analytical methods to study the structural properties of academic networks (Fortunato et al., 2018; Zeng et al., 2017). On the other hand, institutional scholars and more qualitatively minded researchers have emphasized the importance of

[3] A multitude of conceptualisations and operationalizations of community is maintained in the literature. The generation of authors discussed here, predominantly looks at relational communities (Emirbayer and Goodwin, 1994), with direct ties between scientists which are typically discovered by means of survey data. However, these authors do not use the concept of community strictly relational, simultaneously trying to get at the cognitive links between members of the same community, see for example (Griffith, 1989; Small and Griffith, 1974).

meaning and interpretation in science and diffusion (Knorr Cetina, 1999; Latour, 1987; Strang and Meyer, 1993). We discuss these research trends separately before exploring how they might be brought into conversation.

Recent computational work on citation and co-author relations has focused on uncovering the relational structures underpinning the development of science and new discoveries (Fortunato et al., 2018). By applying advanced methods to large digital datasets, this research has reaffirmed some of the findings of earlier studies, including that science is organized into communities (Lambiotte and Panzarasa, 2009; Newman, 2001b, 2004a) that revolve around academic stars, who are more likely to receive new references and engage in new collaborations (Barabási et al., 2002; Dahlander and McFarland, 2013; Newman, 2001b, 2004b).

Fortunato et al. (2018) review these and other findings and outline a research field they call “SciSci”—the Science of Science—which uses computational methods, large datasets, and modeling to identify relational structures and mechanisms of discovery in science. A key premise underlying this computational work is that science is a complex system in which interactions on a microscopic level result in non-linear dynamics and the emergence of unintended and unexpected macroscopic patterns (Fortunato et al., 2018; Shi et al., 2015; Zeng et al., 2017). In line with this premise, many scholars working in this field do not discriminate between social and natural systems. They adopt their methods from the natural sciences, drawing parallels between the diffusion of scientific knowledge and evolutionary processes or the spread of diseases (Bettencourt et al., 2008; Goffman and Newill, 1964; Kiss et al., 2010; Morgan et al., 2018; Zeng et al., 2017). Related research uses agent-based simulations in which the behavioral patterns of individuals are translated into simple rules for agents in the simulation, such as “adopt when more than three of my friends adopt,” and interactions between agents determine the speed and reach of diffusion. A common research question in this field is how different network structures obstruct or facilitate diffusion (cf. Centola, 2015; Centola and Macy, 2007; Watts, 2002). This kind of computational work, focused exclusively on the structural aspects of diffusion, generally assumes that the object of diffusion remains constant as it spreads.

At the same time, interpretative studies of the diffusion of science have shown that the spread of scientific ideas entails not just adoption but also adaptation, similar to re-invention (Rogers, 1983) or exaptation (Bonifati, 2010). Knorr-Cetina (1981) describes how the content of knowledge depends on the different subcultures or epistemic communities in which it is practiced. Latour (1984; 1979) similarly sets out how objects and ideas take on different forms and meanings depending on the local context in which they are adopted, and calls for a paradigm shift from the diffusion model to a translation model. Latour describes the spread as a chain, with the diffusing idea as a 'token':

> Each of the people in the chain is not simply resisting a force or transmitting it in the way they would in the diffusion model: rather, they are doing something essential for the existence and maintenance of the token. In other words, the chain is made of actors—not of patients—and since the token is in everyone's hands in turn, everyone shapes it according to their different projects. This is why it is called the model of translation.
>
> (Latour 1984; p.267-268)

In the translation model, not only does the spread come about as a result of collective action, as described in the structural complexity approach, it also involves adaptations of the idea as a consequence of the interpretations and interactions of actors. More recently, Greenhalg's (2005) study of the diffusion of the innovation paradigm shows that different research traditions develop distinct stories and sometimes contradictory interpretations of the same research findings. David Kaiser (2009) examines the development of the Feynman Diagram in postwar physics and illustrates how even the meaning of scientific inscriptions such as diagrams are not "immutable" as Latour (1986) postulates—but depend on the scholarly social circles in which it spreads. Theorists of institutions such as Tammar Zilber (2008), Sarah Soule (1998), David Strang (1993; 1998), and John Meyer (1993) draw attention to the collaborative and interpretative work involved in diffusion. Strang and Meyer (1993) consider diffusion as a sense-making process in which actors must jointly construct an understanding of a practice or idea before they can adopt it. In other words, adoption requires adaptation and largely depends on the social context.

As the study of the diffusion of scientific ideas bifurcated, a divide opened up between structural and interpretative approaches—the former often made use of computational methods and large datasets, the latter tended to be theoretical and privileged case studies. In study of science, some efforts are made more recently to explore the interaction between the—structural—evolution in scientists' networks on the one hand and their—cultural— intellectual advancements (Moody, 2004) on the other hand, theorizing on regularities in the patterns that unify these two dimensions (Abbott, 2001). Scholars in fields such as social network analysis, information theory, opinion dynamics and relational sociology have similarly sought to bridge this broader structural and cultural chasm (Pachucki and Breiger, 2010). One such approach in social network analysis investigates socio-semantic networks designed to capture the joint dynamics of social and socio-semantic structures (Roth and Cointet, 2010). Information theory scholars seek to expand their frameworks to incorporate meaning into the analysis of scientific communication (Leydesdorff et al., 2018, 2017).

For instance, Vilhelna et al. (2014) find that structural holes (cf. Pachucki and Breiger, 2010) and cultural holes overlap but not coincide in science, underlining the importance of studying not only citation networks but also the content of scientific communication. In the fields of opinion dynamics and diffusion modeling, disease as an analogy is under increasing criticism, as scholars seek to incorporate meaning in previously structurally driven models. For example, Goldberg and Stein (2018) advance a model based on associative diffusion, in which the objects of diffusion are associations between beliefs and behaviors, showing how cultural differentiation can arise without relying on structural fragmentation or homophily among agents. Theoretical attempts at bridging the structural-cultural divide in relational sociology have also been made. John Levi Martin (2002) argues for a formal investigation of the relation between beliefs and social structure, while Fuhse (2009; 2015), building on the work of Harrison White, systematically explores the meaning structure of social networks. These contributions all provide clues as to how structural and interpretive methods might be best combined to examine the co-evolution of meaning and social relations.

We build on this literature by developing the notion of a *diffusion network*—the network that maps the spread of a particular innovation, in this case Granovetter's hypothesis on the Strength of Weak Ties, between adopters. Like scholars in Science of Science, we view the diffusion of science as a complex process, and use computational methods and citation-based diffusion networks to study its micro-macro dynamics. However, like interpretative scholars, we consider every citation to involve interpretation and adaptation, as Granovetter's hypothesis is inserted into particular narratives that aid researchers in identifying and answering the questions of interest. As this process of translation is the outcome of collective interpretative work, we hypothesize that researchers self-organize into distinct diffusion communities.

We are interested in the spreading patterns of Granovetter's hypothesis and how this idea is reinterpreted and adapted during the diffusion process. Our main hypothesis is that diffusion networks are comprised of structural communities that advance the same scientific ideas in distinct ways. In addition to testing this general hypothesis, we seek to understand what gives rise to these structural-cultural patterns in the diffusion network. Accordingly, we examine the network's evolution over time and identify the roles of key actors in brokering diffusion and developing specific interpretations of the Strength of Weak Ties.

Data & Methods

Our strategy is to apply network analysis to citation data of Granovetter's hypothesis in order to identify structural patterns in diffusion processes and then to use topic modeling and close reading of publications to understand the interpretative work scholars engage in. While previous research examines aggregate knowledge flows between fields or institutions, confirming the self-organization of science into communities (Rawlings et al., 2015; Noyons and van Raan, 1998; Rosvall et al., 2009), our interest is in the dynamics of the diffusion of a particular scientific idea, shaped by both structural and cultural forces. This entails interest in the specifics of interpretation and therefore requires the type of fine-grained analysis enabled by the in-depth study of a single case of scientific diffusion. We thus conduct what might be thought of as a computational case study. Like computational researchers, we use advanced computational techniques to search for relational structures in the spread of a scientific idea, and like qualitative researchers, we rely on interpretative methods to develop a nuanced understanding of qualitative differences in how Granovetter's hypothesis has been adapted by scholars in different communities.

To construct the diffusion network, we collected data on publications referencing Granovetter (1973) from the Web of Science.[4] For each publication, we retrieved the following metadata: author(s), title, journal, publication date, research areas, keywords, abstract, and references. The dataset contains 8,198 publications from May 1973 until November 2017. We used this data to construct a network that represents the journey of Granovetter's hypothesis through the academic landscape. Previous studies on academic citation networks typically use edges to represent either direct citation (Price, 1965), co-authorship (Newman, 2001a) or co-citation (Small and Griffith, 1974) relationships among scholars. With our edges, we aim to capture the formal scientific communication between

[4] Although the Web of Science's coverage is relatively broad, it primarily includes publications from journals and contains fewer books and book chapters.

authors that involved the idea in question. We therefore combine both co-authorship and direct reference relations between scholars, since both are signals that an exchange of ideas has taken place between these scholars on the Strength of Weak Ties[5]. That is, edges are drawn from scholars new to the Strength of Weak Ties hypothesis to the scholars they cite who have previously used the hypothesis, hence representing influence[6] of prior authors (edge target) to newly adopting authors (edge source). As we are interested more in the spread of the idea than the intensity of its use, we only create outgoing edges for publications in which authors reference Granovetter (1973) for the first time. Similarly, we draw directed edges of authors' first publication that references Granovetter (1973) to their co-authors on that publication, on the assumption that co-authors work together to position their work in relation to others, including Granovetter. For incoming edges, in contrast to outgoing edges, we consider later publications. This procedure generates a diffusion network that includes 8,198 publications, 15,056 scholars (nodes), and 142,227 edges.

To determine whether communities indeed mediate the diffusion of innovation, we first test whether the modularity of the diffusion network is significantly higher than a random network with the same degree distribution and sequence. We then use topic modeling to identify principal themes and frames in the literature (Bail, 2014; DiMaggio et al., 2013), and examine how these relate to the structural diffusion

---

[5] Reference and co-authorship relations might signal a different type of communication about the diffusing idea. References might signal a simple information flow between weak ties in which the edge target informs the edge source about the novel idea, similar to Granovetter's (1973) study on job vacancies. A strong tie coauthorship relation might reveal more about how the novel idea gets embedded in the literature and research methodology by the edge target. However, both types of communication are integral parts of the diffusion, are hard to discern and can take place in both types of relations. We therefore do not discriminate between these two types of relations in our network.

[6] It is difficult to gauge the extent of influence of prior authors upon new authors referencing the Strength of Weak Ties. Some scholars cite articles without reading them; others use cited articles extensively (for an overview of theories of citation, see, for example, Moed (2005)). For our analysis—which focuses upon the meso- or community level rather than upon micro-interactions among scholars—it is sufficient to state that prior authors have 'some influence' over new authors.

communities in the network (as identified through community detection). Finally, we do a close reading of key contributions to investigate how the application and adaptation of Granovetter's hypothesis differs between three large communities. To study how these structural-cultural patterns emerge, we examine the development of communities over time and the role of influential scholars within them. To do so, we ran a temporal community detection algorithm to locate communities in different time slices (1995-2000-2005-2010-2017) (Mucha et al., 2010) and explore the paths of key figures in the diffusion of the Strength of Weak Ties hypothesis. These key scholars play crucial roles in the formation and linking of communities. They do not perform this work on their own, but serve as focal points for scholars who constitute specific communities (Collins, 1998). In other words, their leadership is not an individual property but emerges from the references of numerous scholars in their communities—more precisely, the communities are formed through the references (Collins, 1998). Some communities are quite closed and constructed around key scholars important only to members of that community; other communities have porous boundaries. By examining the role of these key scholars, we form a better idea of the mechanisms by which diffusion communities are constructed as a result of academics' referencing practices.

Communities in the Diffusion Network

A key premise of our argument is that the diffusion network contains clusters corresponding to communities of scholars who collaboratively interpret and cultivate Granovetter's hypothesis in various directions.

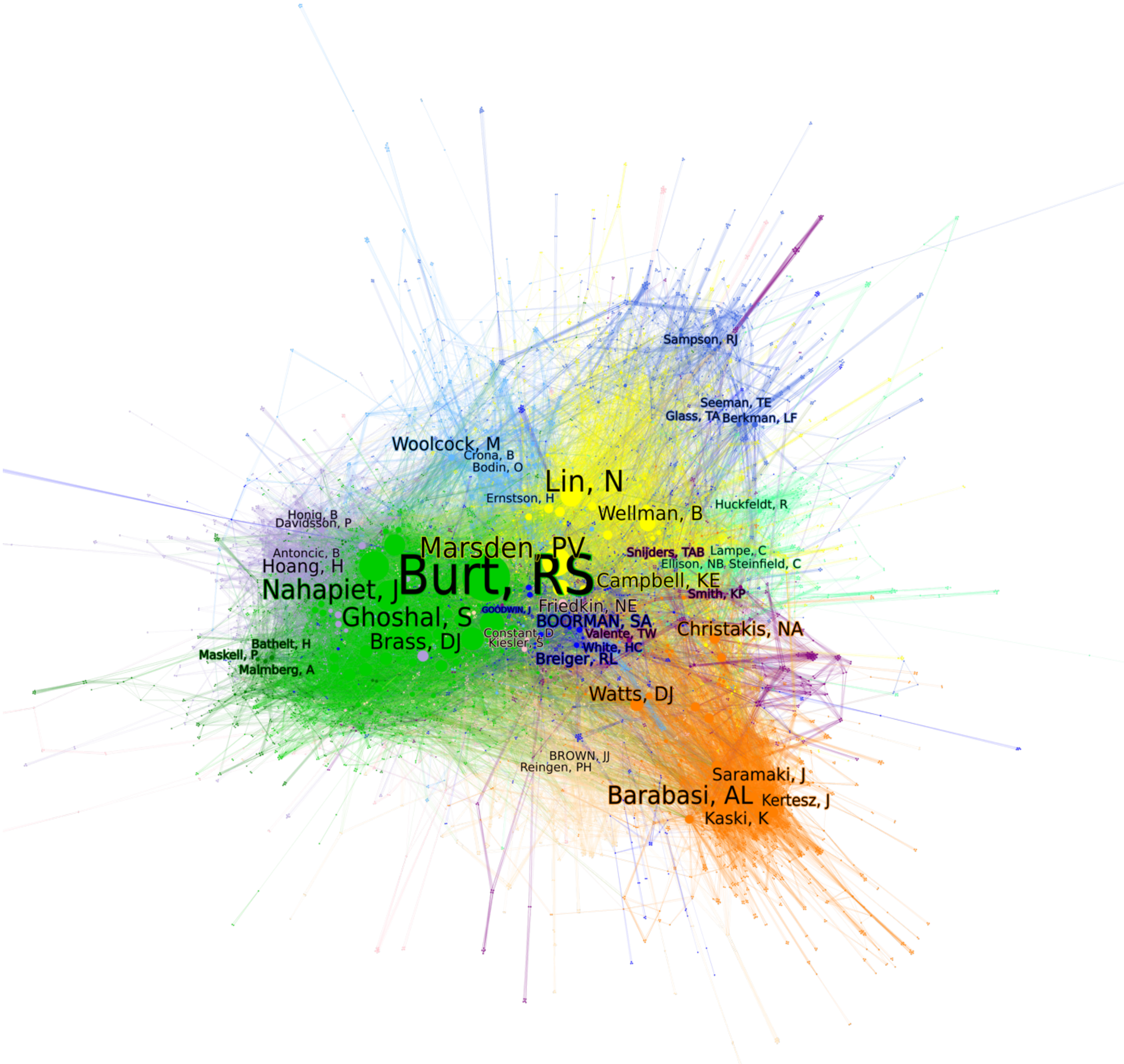


Figure 1: The largest 12 communities of the diffusion network in 2017, containing 10,787 scholars and 121,132 edges. The nodes are colored by their community and the scholars with highest indegree of each diffusion community are labeled. The labels are sized according to their indegree.

Before turning to the question of collaborative interpretation, however, we first need to ascertain that the network indeed exhibits significant clustering. We identify network communities using the Louvain algorithm (Blondel et al., 2008; Traag, 2015), a community detection algorithm which stochastically optimizes modularity. The Louvain algorithm provides slightly different approximations of the optimal partitions in different runs. To improve the robustness of our results, we ran 10,000 instances of the algorithm and compared the resulting community structures by focusing on scholars with a high indegree (>200) (81 scholars representing 0.5% of the sample) and how they are grouped together. We selected an instance where high indegree scholars who are grouped together in the majority of configurations (>60% of 10,000), are grouped together, and high indegree scholars who are never or only rarely grouped together (<10% of 10,000) are not grouped together, as an appropriately robust partition.

When we examine the community structure of the diffusion network (Figure 1), we see that it consists of communities of scholars, defined as groups of scholars with more edges between members of the same community than between members of different communities. We refer to these communities as “diffusion communities.” Figure 2 shows the distribution of the size of the diffusion communities, which is very uneven: the three largest communities comprise 45% of all scholars in the giant component; the largest twelve communities (size >200), 86% of all scholars in the giant component. Our analysis focuses on these 12 communities.

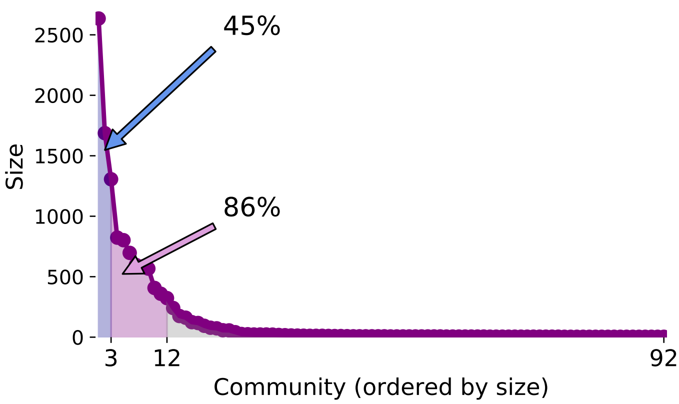


Figure 2: Distribution of community sizes in the diffusion network, with a small number of large communities and a large number of small communities. The largest three and twelve communities consist 45% and 86.4% of all scholars in the giant component of the diffusion network.

To gauge whether this community structure is indeed significant, we need to compare its level of modularity with a plausible benchmark. Since the structure of any network—and particularly networks with an uneven degree distribution—will have some degree of modularity, finding a plausible benchmark is essential. For this, we use an adjusted version of the Havel-Hakimi graph (Hakimi, 1962; Kleitman and Wang, 1973). We compare the modularity of our empirical network to the average modularity of 10,000 Louvain partitions of adjusted Havel-Hakimi networks with an identical degree sequence as the empirical network. We treat reciprocal and singular links separately and match their degree sequences to create our adjusted graphs. This is necessary as our network has notably few reciprocal links, which is not the case in the regular Havel-Hakimi graph. By design and logic of the diffusion network, earlier links are not reciprocated. Only scholars who reference Granovetter (1973) for the first time in a co-authored publication have a reciprocal link in the diffusion network.

The adjusted Havel-Hakimi graph serves as a benchmark for our network, as it represents the hypothesis that the structures of these networks are products of a first-mover advantage (Newman, 2009), positing that the first publications and scholars in a new research area receive citations at a much higher rate than later ones. This hypothesis is modeled as follows: the network grows over time as more scholars discover Granovetter's idea. Each new generation of researchers cites Granovetter as well as previous generations of scholars: the first generation cites only Granovetter; the second cites Granovetter and the first generation; the third cites Granovetter and the first two generations, and so on. This is the process that the Havel-Hakimi algorithm represents: it generates graphs by successively connecting nodes of the highest degree to nodes of the second highest degree, ordering the remaining nodes by degree from high to low, and repeating the process. The Havel-Hakimi graph thus captures how a scientific diffusion network would be structured, were it only organized by the timing of publications and scholars, without scientific communities playing any role in the diffusion process.

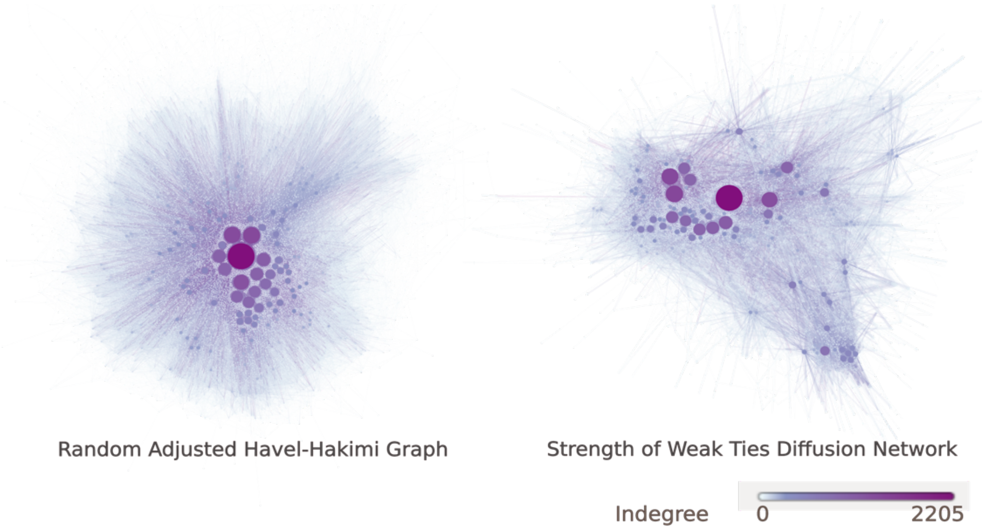


Figure 3: The Strength of Weak Ties diffusion network in 2017 (right) and a random adjusted Havel-Hakimi graph with identical degree distribution (for both reciprocal and singular edges). Both visualizations have identical settings, with nodes sized and colored by their indegree and the same layout algorithm (Gephi's Force Atlas 2). The diffusion network is more clustered (0.623 p-value<0.001) than the adjusted HH graph. The high indegree scholars are highly centered in the HH graph and more spread out over different communities in the diffusion network.

By comparing the Strength of Weak Ties diffusion network with the adjusted Havel-Hakimi graphs, we find that the former has significantly more community structure (0.62, p-value<0.001). Figure 3 shows our diffusion network on the right and a random instance of the Havel-Hakimi graph on the left with identical degree distributions (both for singular and reciprocal links), demonstrating a marked difference in network modularity.

Comparing these networks points to another structural feature that the first-mover advantage model leaves out. Figure 3 shows how scholars with highest indegree are located at the center of the adjusted Havel-Hakimi graph, whereas they are spread out over different communities in the Strength of Weak Ties diffusion network. Scholars with high indegree are authors[7] of publications containing a reference to Granovetter (1973) that

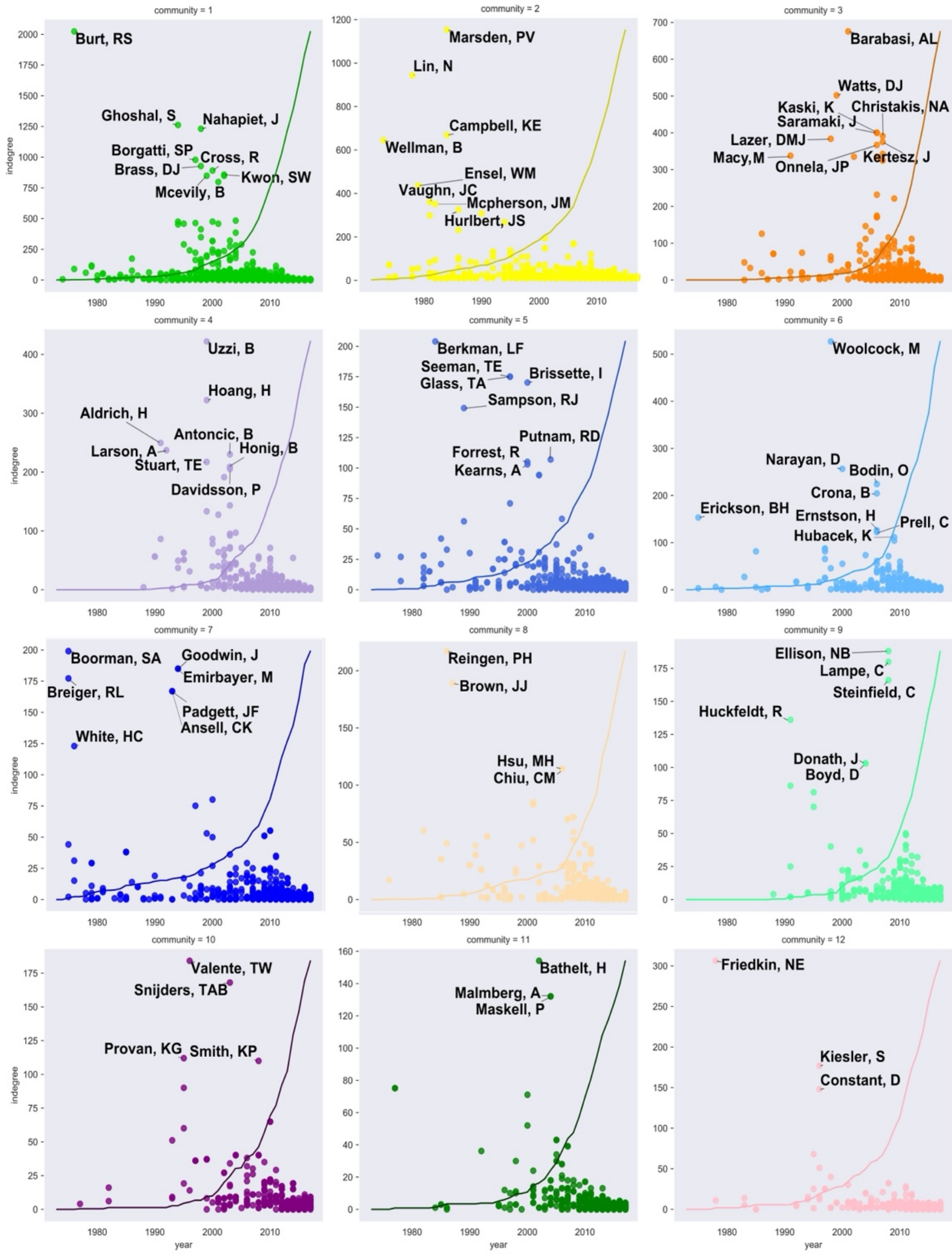
7 References to publications are included in the network as edges to all authors of the referenced work, not only to the first author.

are often referenced by scholars new to the Strength of Weak Ties. Examining the growth of communities and the indegree of scholars over time (Figure 4), we see that the first-mover advantage does not seem to drive the Community sizes diffusion process. Numerous scholars cite Granovetter (1973) much later—for example Brian Uzzi in 1999, Albert-László Barabási in 2000, and Örjan Bodin in 2006, respectively twentysix, twenty-seven, and thirty-three years after Granovetter's publication—but nevertheless receive many citations from the next generation of adopters, making them important figures in the diffusion of Granovetter's hypothesis.

While these academic stars are cited by scholars in the entire network, they are mostly—sometimes even exclusively—cited by scholars from their own communities. These findings show that the spread of Granovetter's idea was not a simple process of contagion, but that scholarly communities containing key figures played an important role in its diffusion to a broader scholarly audience. The distinctive feature of high-indegree scholars may not be simply timing—as the first-mover advantage theory proposes—but their status (Cole, 1970; Morgan et al., 2018; Way et al., 2019) or their ability to apply an existing idea in a novel context, so that it speaks to scholars in other research communities (Lane, 2011). The latter point is part and parcel of the idea that innovation takes place throughout the diffusion process, and not just at its initiation (Lyytinen and Damsgaard, 2011).

Figure 4 (next page): The growth (line) and indegree of researchers (scatter) in each community of the Strength of Weak Ties diffusion network over time. The y-axis for growth—in terms of community size—runs from 0 to 100%, but is not shown for the sake of legibility. The scatter points of the eight scholars with highest indegree per community (and indegree>100) are labeled. Most communities have at least one important high indegree scholar, and the timing of these scholars' first publication referencing Granovetter's hypothesis varies significantly: not all are first movers.

Growth of communities and indegree of researchers over time



Communities' Interpretative Work: The Development of Narratives

We applied topic modeling to the abstracts of publications in the twelve largest communities and explored correlations between topics and communities. The resulting correlation matrix in Figure 5 shows the degree to which scholars in the twelve communities discuss various topics. It reveals that different communities do indeed apply Granovetter's idea to different topics (Chi-squared=2057, df=154, p-value<0.1^5), albeit to different degrees. Communities addressing similar topics tend to be more connected in the citation network (Pearson correlation=0.23, p-value=0.06) (see Figure 6). For example, Community 4's topics are similar to those of Community 11, and these two communities are strongly connected in the diffusion network based on citations (see appendix for details on topics).

We find that communities comprise distinct combinations of scholars from different research fields (Chi-squared=177,432,451, df=1,100, p-value<0.1^5) (Figure 7), with communities closer in their research interests exhibiting stronger connections in the citation network (Pearson correlation=0.39, p-value<0.001) (Figure 8). We can get a sense of a given community simply by looking at topics and disciplinary backgrounds (Figure 9). Scholars in Community 9, for example, appear to be active in the field of communication science, discussing words associated with Topic 12, including "information," "online" and "media."

These findings provide *prima facie* evidence that the diffusion of a scientific idea is mediated by scholarly communities—previously existing or newly formed—with different disciplinary perspectives and research interests. While correlations between topics and research fields do not demonstrate that scholars only cite within their field or that they limit themselves to specialized topics, they do show how the diffusion of a novel idea via citations is closely linked to its contextual understanding and applications. While topic modeling provides us with the contours of interpretative schemas, a close reading of key publications—identified by the number of references they receive in their communities—is necessary to better understand how scholars integrate Granovetter’s hypothesis into their frameworks and apply it in their research. As we shall see, Granovetter’s 1973 article planted a seed for a number of research avenues and understandings of the Strength of Weak Ties, which have each developed and diverged during the diffusion process.

We now turn to a more detailed analysis of the three largest communities in the diffusion network, which each leverage and develop another use case and interpretation of the Strength of Weak Ties. We refer to them as the *Organizational Advantage Community*, the *Ego-Network Community*, and the *Complex Networks Community*.

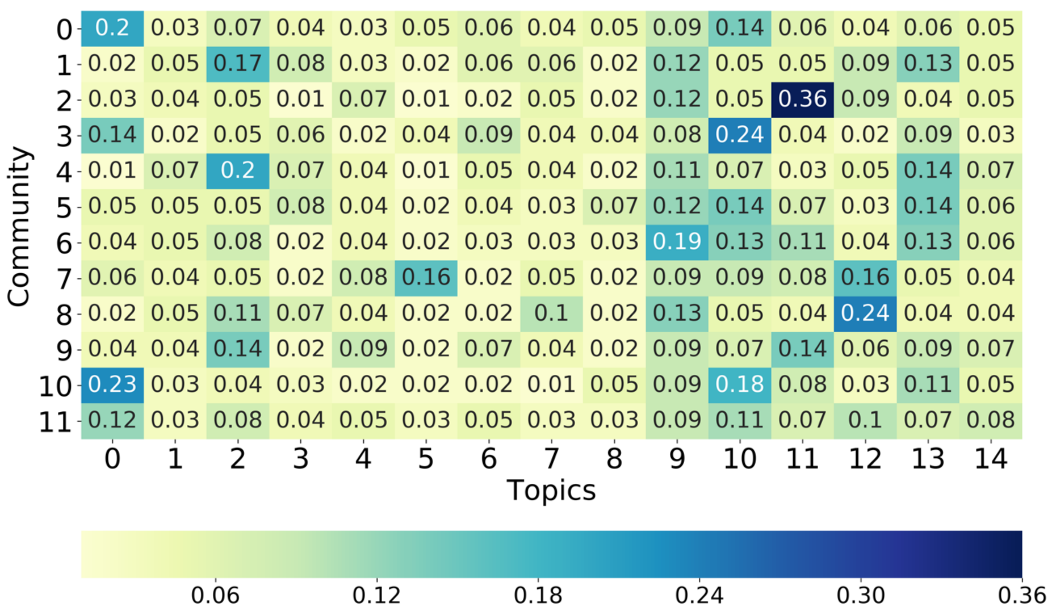


Figure 5: The topics (columns) addressed by communities (rows) in the Strength of Weak Ties diffusion network. Cell numbers indicate coverage by all community publications, e.g. 36% of publications in the Complex Networks Community (community 3) address complex models (topic 11). The parameters for topic modeling are set to find 15 topics and to discard words that occur in less than 30 articles or in more than 80% of articles. See appendix for details of topics.

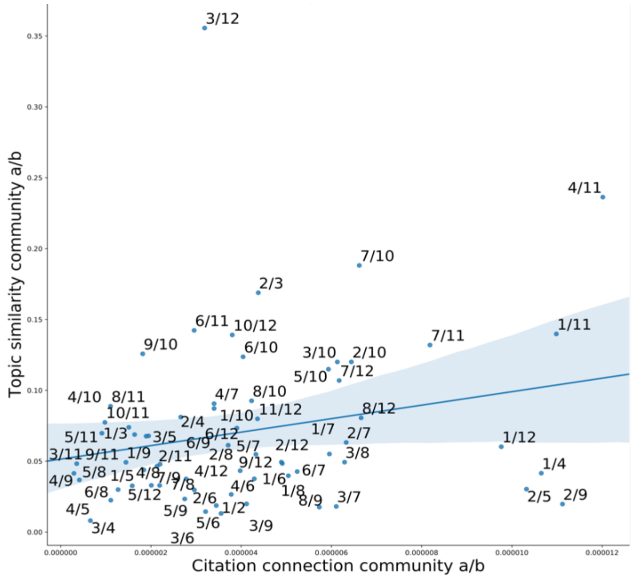


Figure 6: The relation between communities in the Strength of Weak Ties diffusion network expressed by their direct citations (x-axis) vs. their similarity in topic coverage (y-axis), Pearson correlation=0.23, p-value=0.06. The citation relation is calculated as the number of edges between communities a and b, divided by the product of the sizes of communities a and b. The topic similarity is calculated as the correlation between the topics covered by communities a and b.

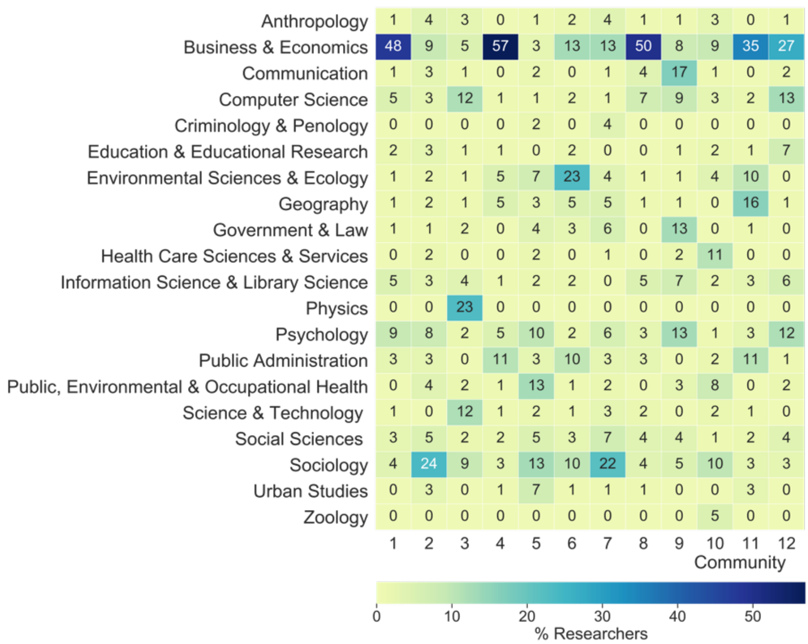
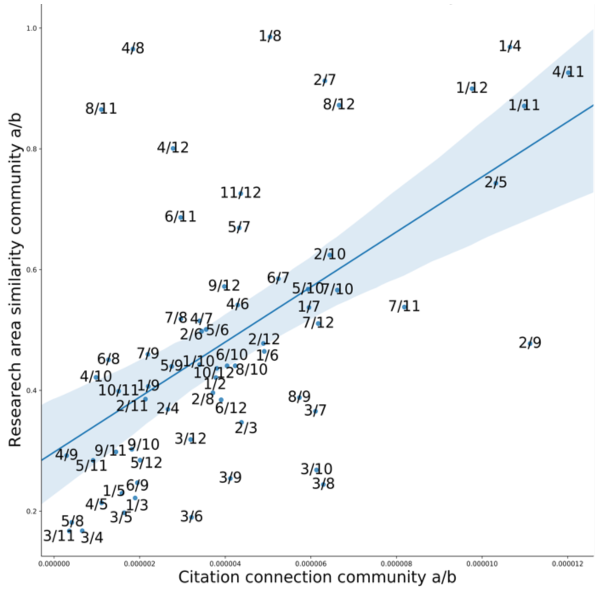


Figure 7: The disciplinary background of communities in the Strength of Weak Ties diffusion network. Each cell value and color represents the percentage of community researchers of a particular field (e.g. 57% of researchers in community 4 publish on business & economics). The figure only contains research fields where at least one community significantly deviates from the overall network (two-sided Z-test) and which involve at least 5% of the community's scholars.

Figure 8: The relation between communities in the Strength of Weak Ties diffusion network expressed by their direct citations (x-axis) vs. their research areas (y-axis), Pearson correlation=0.39, p-value=0.001. The citation relation is calculated as the number of edges between communities a and b, divided by the product of the sizes of communities a and b. The research area similarity is calculated as the correlation between the research areas of communities a and b.

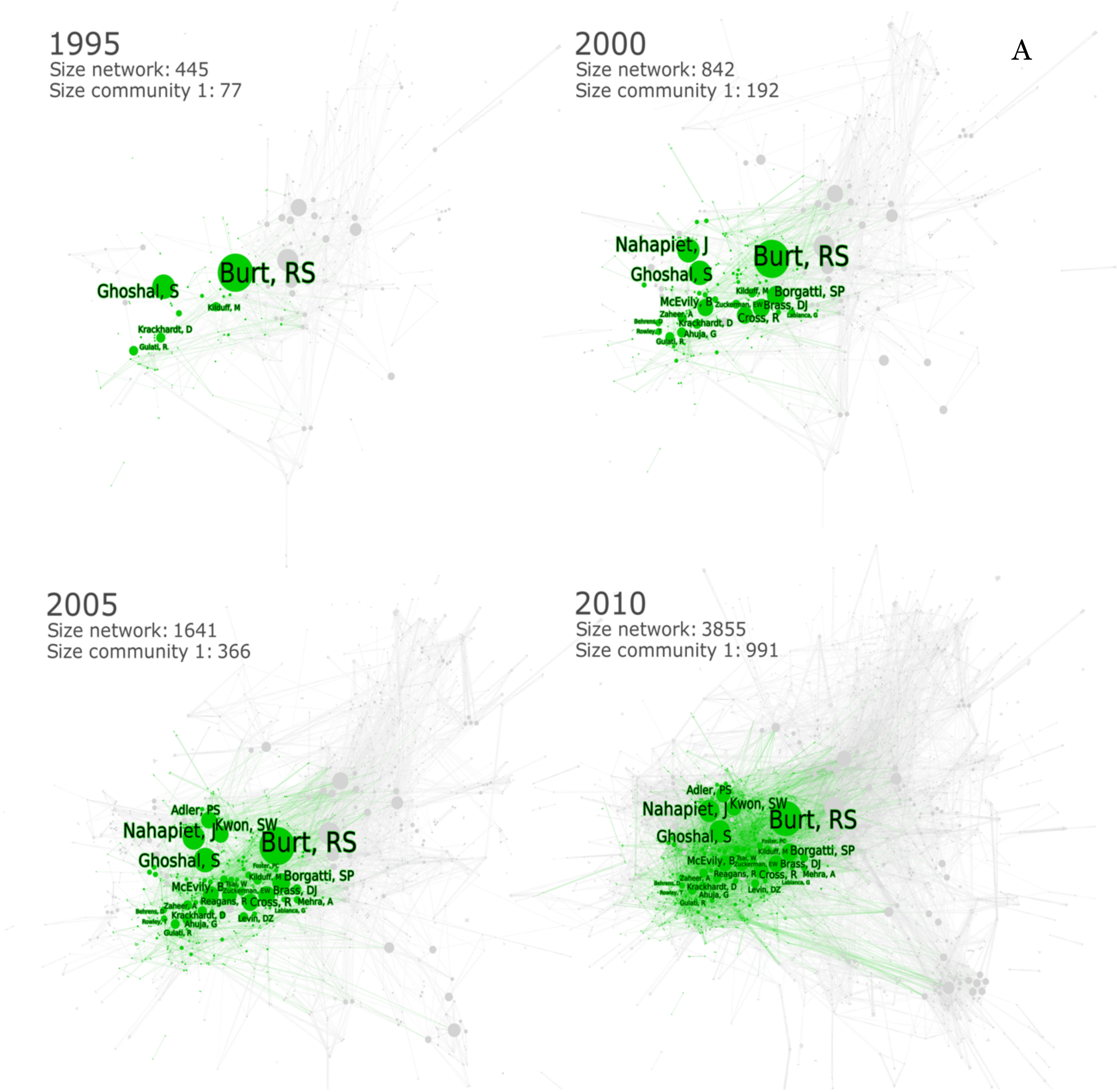


Community	Size	Central Figures	Dominant Research Fields	Dominant Topics
1	2635	Burt, RS | Ghoshal, S | Borgatti, SP	B&E	0 - Organisational Advantage|10 - Enterpreneurship
2	1687	Lin, N | Marsden, PV | Wellman, B	Sociology	2 - Survey Data | 13 - Economic Development
3	1306	Barabási, AL | Watts, DJ | Macy, M	Physics | Science & Tech.	11- Complex Networks | 9- Markets & Politics
4	823	Uzzi, B | Hoang, H | Aldrich, H	B&E	10 - Enterpreneurschip | 0 - Organisational Advantage
5	802	Berkman, LF | Seeman, TE | Glass, TA	Sociology | Public Env. & Occ. Health	2 - Survey Data | 13 - Economic Development
6	697	Woolcock, M | Narayan, D | Bodin, O	Env. sciences & Ecology	10 - Enterpreneurship |13 - Economic Development
7	597	Breiger, RL | Boorman, SA | White, HC	Sociology | B&E	9 - Markets & Politics | 11- Complex Networks
8	584	Reinigen, PH | Brown, JJ	B&E	5 - Methodology | 12 - Communication
9	567	Ellison, NB | Lampe, C | Steinfield, C	Communication	12 - Communication | 9 - Markets & Politics
10	407	Valente, TW | Snijders, TAB	Health Care Sciences	2 - Survey Data | 11 - Complex Networks
11	359	Maskell, P | Bathelt, H | Malmberg, A	B&E | Geography	0 - Organisational Advantage | 10 Enterpreneurship
12	323	Friedkin, NE | Kiesler, D	B&E | Computer Science	0 - Organisational Advantage | 10 Enterpreneurship

Figure 9: Size, central figures, prominent research fields and topics addressed by scholars in each community in the Strength of Weak Ties diffusion network. We have named the topics to capture their essence, see appendix for more details.

Community 1. The Organizational Advantage Community

Granovetter (1973) points out that weak ties are more likely than strong ties to be bridges between socially cohesive clusters, and suggests they are therefore crucial for the flow of information. This observation is taken up by the Organizational Advantage Community in the context of management and organizations. Most scholars in this community publish in the fields of management and organization. The central scholar is Ronald S. Burt, followed by Sumantra Ghoshal, Janine Nahapiet, Daniel



B

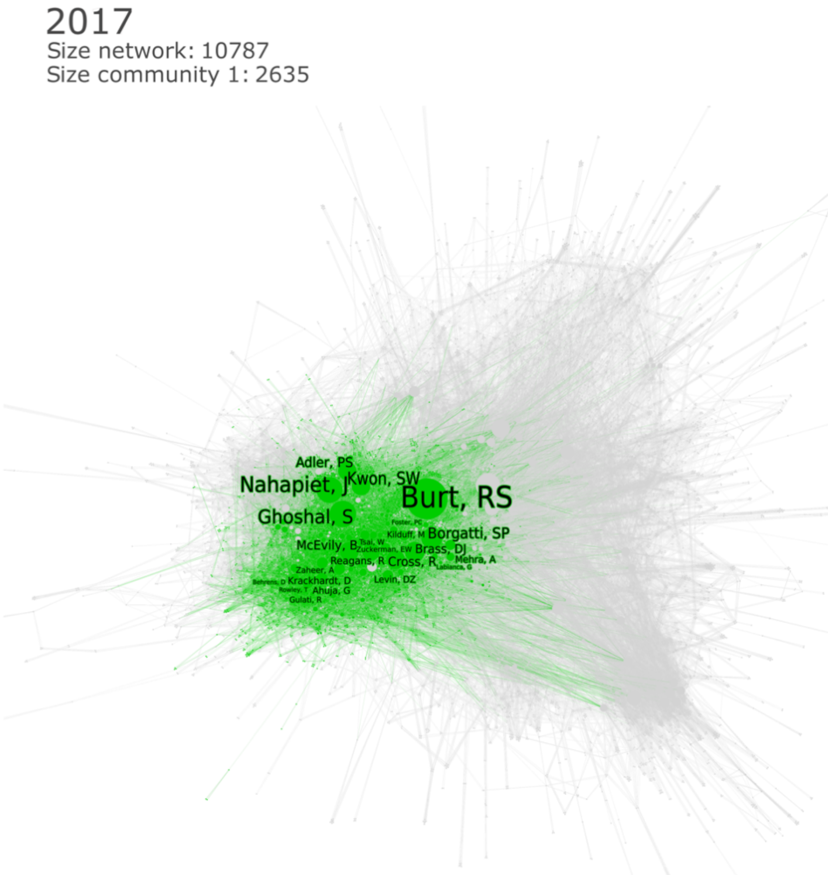


Figure 10.a (previous page) & 10.b: Growth of the organizational advantage community (community 1). All scholars in this community are colored green. Only scholars who received at least 250 citations from future adopters (indegree >= 250) are labeled, sized according to indegree. The community develops around seminal works by central figures such as Burt (1997; 2000; 2004), Nahapiet and Ghoshal (1998), Brass (2004), Cross (2004), Reagans and McEvily (2003) and Borgatti (2003). By 2005, all scholars to be most cited by this community have extensively referred to the Strength of Weak Ties.

J. Brass, Bill McEvily, Rob Cross, Ray Reagans, Stephen P. Borgatti, Seok-Woo Kwon, and Paul S. Adler (see Figure 10 for the structural development and position of scholars in this community).

The vast majority of empirical studies in this community use firm-level data and focus on innovation-based competitive advantage for organizations (e.g. Reagans and McEvily, 2003; Nahapiet and Ghoshal, 1998; Brass et al., 2004; Adler and Kwon, 2002; Burt, 2000). According to scholars in this community, innovation occurs when extant knowledge and experience are combined in new ways and they relate this to the structural patterns within organizations: innovation and good ideas are more likely to appear near structural holes where the knowledge of different social collectives intersects (Burt, 2004). The Strength of Weak Ties is a pillar of knowledge creation in this community, and is the basis for Burt's notion of structural holes: "The structural hole argument draws on several lines of network theorizing that emerged in sociology during the 1970s, most notably, Granovetter (1973) on the Strength of Weak Ties" (Burt, 2000, p. 340). Burt thus interprets, adapts, and extends Granovetter's notion of the Strength of Weak Ties so that it becomes relevant to a community of scholars who seek to understand why some organizations, corporations, and managers have advantages over others.

As this Organizational Advantage Community grows, social capital becomes its most central concept, understood as "the sum of the actual and potential resources embedded within, available through, and derived from the network of relationships possessed by an individual or social unit" (Nahapiet and Ghoshal, 1998, p. 243). "Social Capital, Intellectual Capital, and the Organizational Advantage" by Nahapiet and Ghoshal (1998) is the most frequently cited article by new adopters in this community. While this resonates with the work of scholars such as Robert Putnam and James Coleman, scholars in this community are specifically interested how social capital may confer organizational advantages to corporations or managers. They link concepts such as social capital and weak ties to notions like intellectual capital, knowledge, and innovation, also drawing upon other works of Granovetter such as his writing on embeddedness (1985).

Community 2. The Ego-Network Community

In his 1973 article, Granovetter illustrates his theoretical argument with empirical evidence about job attainment that shows individuals more often find jobs through weak ties than through strong ones. Members of the Ego-Network Community build on this to conceptualize weak ties as a type of individual asset which enhances this individual's status in society. The majority of scholars in this community publish in sociology and are interested in how different types of social relationships can confer advantages to individuals, particularly in terms of status (e.g. Lin et al., 1981; Lin, 1999; Campbell et al., 1986). This focus on individuals corresponds to the main data source for these scholars, namely surveys. The central figures in this community are Nan Lin, Peter V. Marsden, Barry Wellman, and Karen E. Campbell. These scholars laid the groundwork for this community in the 1970s and 1980s and some are directly connected to Granovetter, such as his colleague at Jon Hopkins University, Nan Lin.

In this community, a central research topic is how individuals derive different kinds of benefits from strong and weak ties; some of its most-cited publications are devoted to measuring tie strength using survey questionnaires (Marsden and Campbell, 1984). The central concept of this community is "social resources": different kinds of ties offer different kinds of support to individuals (Wellman and Wortley, 1990). A central theoretical notion is the social resource proposition, explained by Lin and Dumin (1986, p. 366) as: "an individual who uses a contact of higher socioeconomic status should find a better job than someone else whose contact has lower status." Scholars in this community likewise explore the hypothesis that weak ties confer distinct advantages: "for two individuals at the same or similar initial positions, it is hypothesized that the one who uses weak ties rather than strong ties will tend to reach better social resources. This is called the strength of ties proposition" (Lin and Dumin, 1986, p. 367).

Ties are seen as an individual's property, as stated in the following passage from one of the most cited publications in this community: "The friend may use his/her position or network to help ego to find a job. These are 'borrowed' and useful to achieve ego's certain goal, but they remain property of the friend or his/her friends" (Lin, 1999, p. 468). Whereas scholars researching organizational advantage find strength in weak ties by viewing them as a collective property, scholars studying ego-networks consider weak ties as individual property that can strengthen individual status.

Community 3. The Complex Networks Community

The Complex Networks Community shifts the focus from social networks to networks in general. Granovetter (1973) presents the Strength of Weak Ties as part of a broader argument for using structural networks to link micro and macro levels of society. This ties in with the central focus of this community: the study of complex networks, in which individual properties and micro-interactions coalesce into sometimes surprising macro-patterns. This community consists primarily of physicists, science and technology scholars, and computer scientists. The community's main figure is Albert-László Barabási, a physicist interested in detecting and modeling the universal properties of complex networks.

Key words in the community's dominant topic include "model," "structure," "nodes," "properties," "degree," and "complex." The community is driven by data and models as it examines the structural patterns of networks and quantifiable emerging patterns. The first significant scholars who formed this community include Duncan Watts, Michael Macy, and Nicholas A. Christakis, whose work is partly situated in sociology and links the behavior of individuals to collective behavior and network characteristics (e.g. Centola and Macy, 2007; Centola, 2010; Christakis and Fowler, 2007; Kossinets and Watts, 2006). The Strength of Weak Ties diffused from these more social science oriented scholars towards physicists focused on numerical models, such as Albert-László Barabási, Kimmo Kaski, Jari Saramäki, János Kertesz, and Jukka-Pekka Onella (e.g. Karsai et al., 2011; Onnela et al., 2007; Albert and Barabási, 2002). This can be seen in Figure 11, which shows the growth of this community in the network. One of the most referenced works in this community is Barabási and Reka Albert (2002)'s "Statistical Mechanics of Complex Networks," which discusses abstract properties of complex networks. The variety of environments considered in their work—cells, chemicals, and the Internet—speaks to the broad applicability of Granovetter's idea as interpreted by this community.

In contrast to the Organizational Advantage and Ego-Network Communities, which reference both Granovetter's 1973 article and his work on economic life and embeddedness, the Complex Networks Community almost exclusively references the 1973 article. The Strength of Weak Ties idea is disconnected from a social setting and is instead conceptualized as an efficiency principle for diffusion processes in complex networks. Damon Centola writes in his highly-cited article, "The Spread of Behavior in an Online Social Network Experiment": "Evidence in support of the Strength of Weak Ties hypothesis has suggested that networks with high levels of local clustering and tightly knit neighborhoods are inefficient for large-scale diffusion processes" (2010, p. 1197). Similarly, according to Barabási et al. the Strength of Weak Ties hypothesis "states that the strength of a tie between *A* and *B* increases with the overlap of their friendship circles, resulting in the importance of weak ties in connecting communities. The hypothesis leads to high betweenness centrality for weak links, which can be seen as the mirror image of the global efficiency principle" (Onnela et al., 2007, p. 7336). Consistent with an interest in emerging patterns, agent-based simulations are the preferred method of inquiry among scholars in this community.

With a deeper understanding of the research interests of these three communities, we see how different communities of scholars translate and advance a scientific idea in various directions. In the community examining organizational advantage, weak ties are viewed as a collective organizational resource, an antecedent and corollary of Burt's notion of structural holes which enables organizations to innovate. In the Ego-Network Community, weak ties are considered individual property, most notably a resource for individual status attainment. In the Complex Networks Community, the Strength of Weak Ties is first and foremost considered a universal property of complex networks, independent of social context.

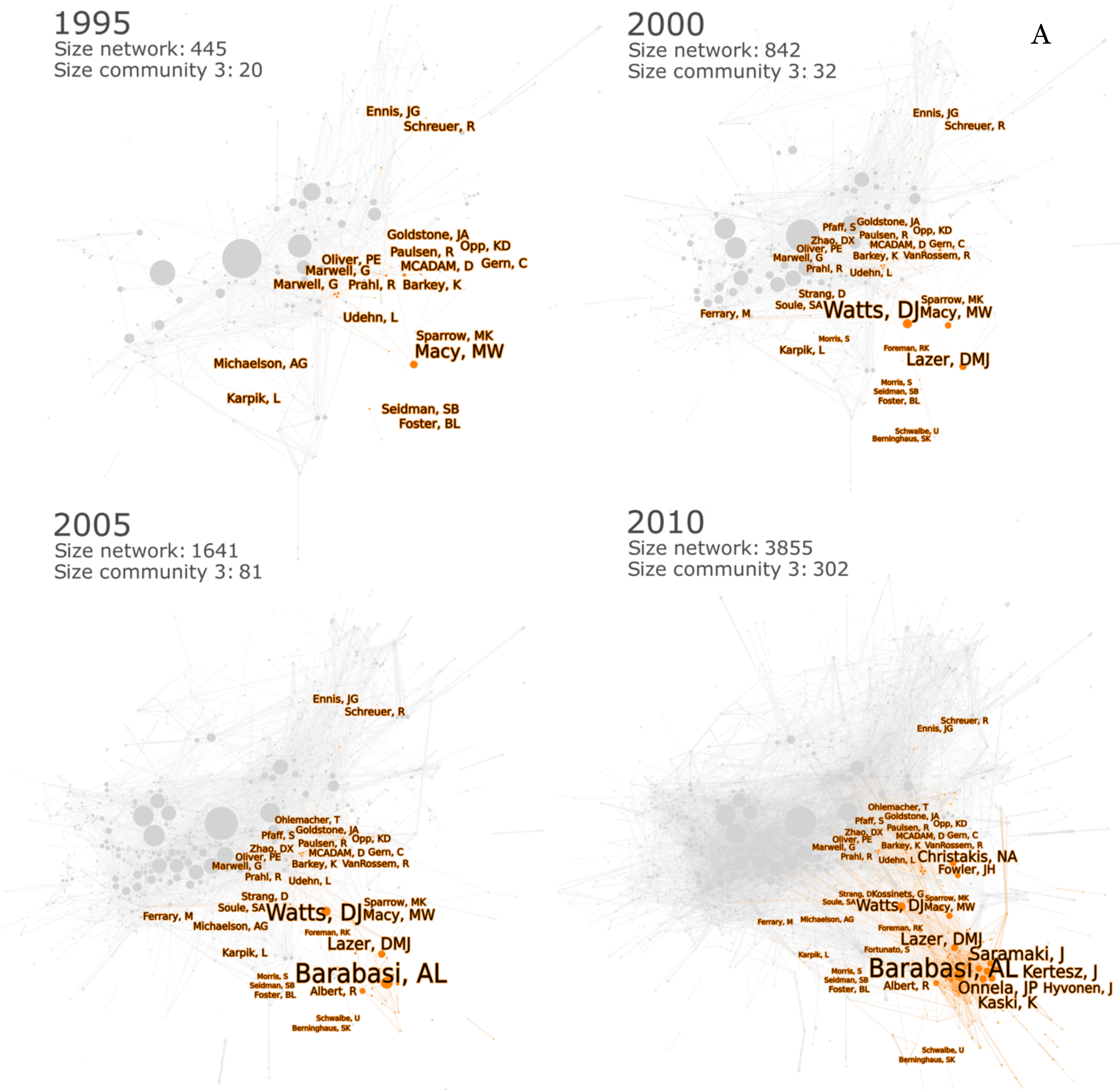
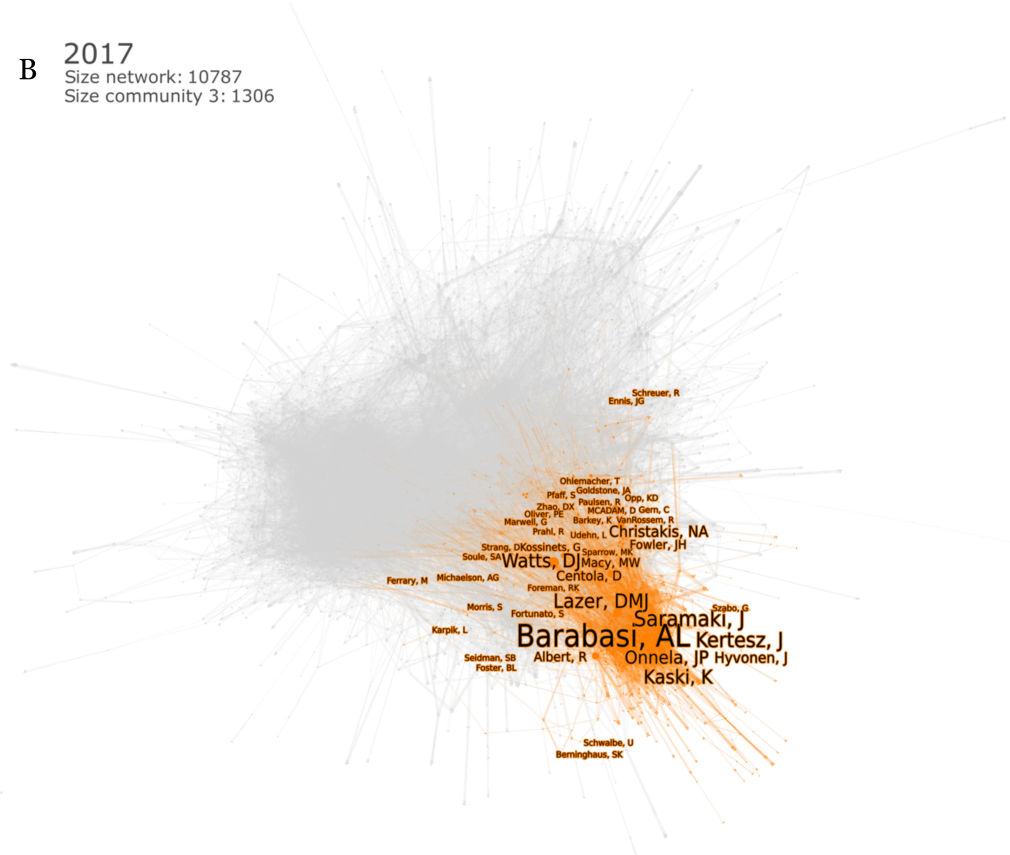


Figure 11.a (11.b, next page): Growth of the Complex Networks Community (community 3). All scholars referencing the Strength of Weak Ties before 2000, who might be considered innovators in this community, are labeled irrespective of indegree. All scholars receiving at least 250 citations by future adopters (indegree >= 250) are also labeled, sized according to indegree. Temporal networks show how this community emerged slowly in 2000, spread due to scholars such as Macy (1991; 1996; 2007) and Watts (1999; 2006), and boomed after Barabási (2002;

2007), Onella, Saramäki, Kaski and Kertesz (2007) began citing Granovetter (1973).

B



Emergence and Growth of Communities

Thus far, we have ascertained that our diffusion network has a community structure; that this structure reflects the development of distinct research cultures which interpret and reuse Granovetter's hypothesis in different ways; and that most communities developed around one or several central researchers active in spreading Granovetter's idea to new audiences. We now turn to the question of what gives rise to these structural and cultural patterns. To do so, we examine the roles individual researchers play as their work collectively shapes the diffusion network over time.

To better understand the forces which shaped the diffusion network over time, we require an historical analysis which considers changes in communities over time.[8] We thus employ a temporal community detection algorithm to find communities in different time slices (19952000-2005-2010-2017) (Mucha et al., 2010), in which nodes in each time slice are weakly linked to the other time slices (interslice weight parameter=0.00001).

Figure 12 shows the evolution of the communities over time (top) and the community paths of key, highly cited scholars (bottom). Some of these hubs—for example Lin, Wellman, and Scott Feld—started out belonging to different communities but later became part of the same community, whereas Ronald Breiger and Burt belonged to the same community and then split into different communities as they are recognized for different contributions to the literature, diffusing the Strength of Weak Ties to different audiences. Burt was acknowledged for his ideas on structural holes within organizational networks (Burt, 1997, 2000, 2004), which became most popular among business and economics scholars interested in innovation (the Organizational Advantage Community). Breiger, alternatively, got known for his contributions on mathematically identifying roles and positions in networks as matrices (Breiger et al., 1975; White et al., 1976). Although his work builds less explicitly on the Strength of Weak Ties, he acknowledges Granovetter, who was also on his thesis committee. Breiger's work is picked up by scholars working in the—at that time—emergent New York School of relational sociology (Emirbayer and Goodwin, 1994; Mische, 2011) who use Breiger's concepts and algorithms for block model analysis. These examples illustrate how the structural communities in the network are related to the interpretative work of the scholars that constitute them but also of scholars citing them at later points in time.

[8] While our analyses have been based on static characterizations of the data, we seek to shed light on a complex and dynamic diffusion process. Thus far, we have defined communities in the diffusion network by the configuration of edges in 2017. Our choice to use a static definition of community was not only technical, but an answer to the ontological question of what communities represent in this case study: by using the full data from 2017, we apply the most recent lens of history as the citation patterns of later researchers are used to identify the community to which earlier contributions belong.

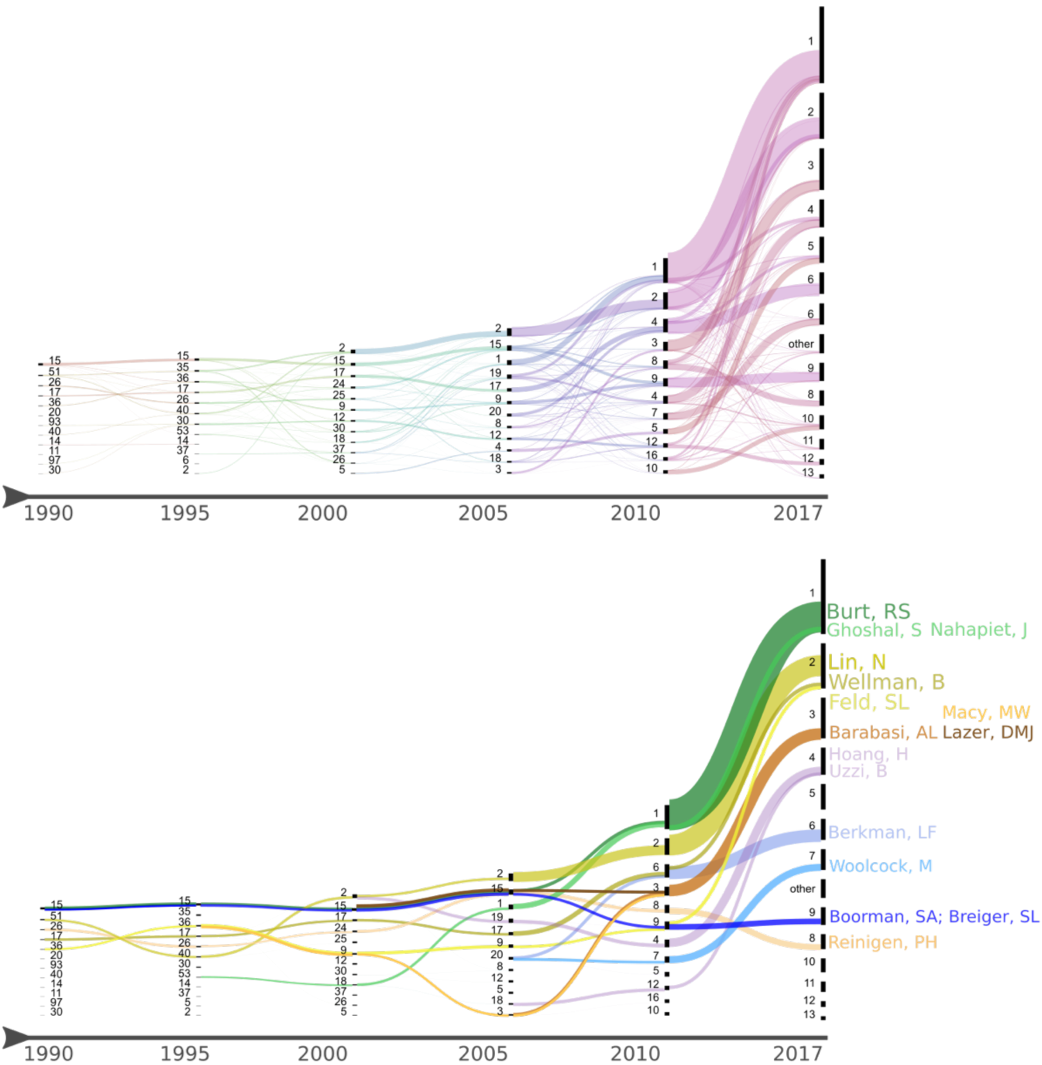


Figure 12: Temporal evolution of communities, detected with the algorithm of Mucha et al (2010), implemented by Vincent Traag in the Louvain Python package, using interslice_weight of value 0.00001 and 1995-2000-2005-2010-2017 time slices. The alluvial diagram shows the largest 13 communities at each time slice. Scholars in smaller communities and scholars who have yet to reference the Strength of Weak Ties in each time slice are omitted. The lower diagram shows the path of important hubs and the splitting and merging of communities over time, arising from both centrifugal and centripetal forces.

We see in these cases centrifugal forces that separate communities and fragment the network, as well as centripetal forces that bring together different communities, integrating the network. The final network structure is a balance of these opposing forces, which emerge from researchers' individual behavior. As researchers navigate the tension between novelty and conventionality, they seek to create new connections, while heeding the common practices of the discipline needed for research to have impact (Foster et al., 2015; Uzzi et al., 2013). When a new idea diffuses, researchers reinterpret it to introduce insights into existing or developing traditions, thus acting as part of the centrifugal force that strengthens the community while fragmenting the larger network. Simultaneously, researchers use new ideas as links or channels to other disciplines and bodies of literature, developing theories that combine different ideas, thereby becoming part of the centripetal force that integrates and draws the diffusion network together.

These competing interests—novelty versus conventionality, tradition versus innovation—become clearly visible if we compare Burt and Barabási's roles in shaping the network. Barabási references Granovetter in a number of highly cited publications (Barabási et al., 2002; Karsai et al., 2011; Onnela et al., 2007), incorporating the Strength of Weak Ties in a complex networks approach, leveraged by the Complex Networks Community. Although Barabási's star rises rapidly, he receives citations almost exclusively (83%) from within his own Complex Networks Community (Figure 13). As Figure 11 shows, this community only took off after 1999 and is primarily organized around Barabási's work, cited by 43% of all new scholars in this community (Figure 13). Like Barabási, Burt is prominent in the diffusion network, but his role is different. Burt theorizes about structural holes and how brokerage enhances creativity and innovation; he is not only the most prominent scholar in the Organizational Advantage Community, but also the most central actor in the diffusion network as a whole (with the highest authority value[9] of

[9] The authority value (Kleinberg, 1999; Langville and Meyer, 2005) measures the centrality of a node by considering the centrality of its neighbors. The focus here is on the incoming edges of nodes, hence the name authority centrality. His high score on this measure thus reflect Burt's centrality in the overall network. Where some individuals are very prominent within their cluster, Burt is influential across the diffusion network as a whole, connecting its different parts.

0.0047). Burt publishes many articles in which he cites Granovetter's Strength of Weak Ties publication (n=26) and offers contributions also beyond the role of structural holes for organizational advantage, such as insights on survey network data (1984) and social capital (1997). In his publications, he draws upon a wide variety of literature. Burt is strongly connected to Ego-Network Community, having been supervised by Lin for his M.A., and his ideas are much influenced by his doctoral advisor James Coleman. Burt receives a large number of citations (of 2.623 unique new scholars in the network) and, in contrast to Barabási, in notable amounts by members of other communities than his own, see Figure 13 for details. Much of Burt's earlier work has become canonical not only in management science and sociology, but also in the interdisciplinary field of network analysis. Like Granovetter, Burt advances ideas that find their way into publications on diverse topics with different theoretical underpinnings and methodologies, in effect serving as a vehicle for network integration. Burt thus diffuses the Strength of Weak Ties across community borders, contributing to connecting networks of scholars. Interestingly, Burt does what he theorizes: he is a broker operating within the structural holes between communities in the academic landscape.

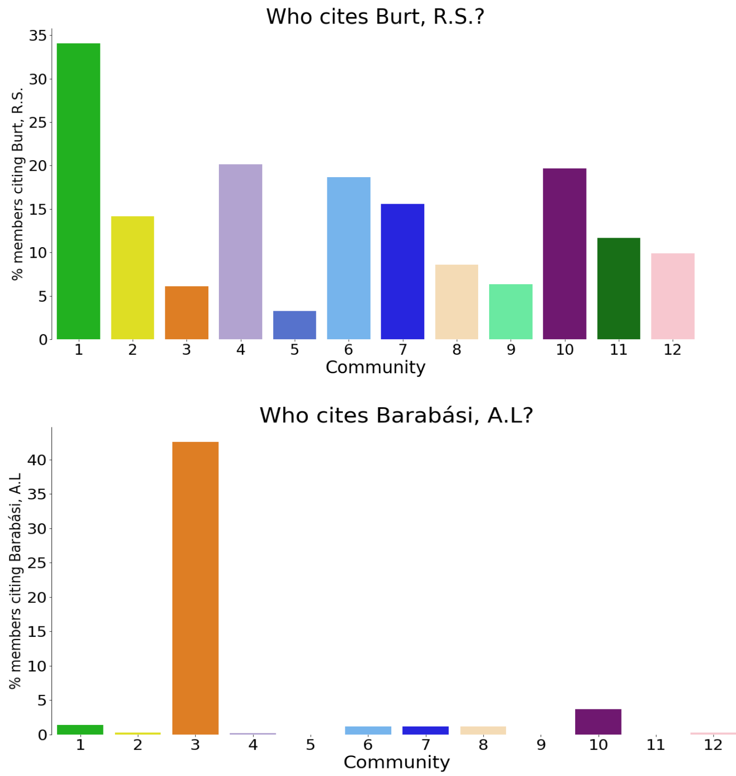


Figure 13: Citations to Ronald S. Burt (top) and Albert-László Barabási (bottom) from scholars in the twelve largest communities in the Strength of Weak Ties diffusion network. The bars represent the percentage of scholars referencing publications by Burt or Barabási on the first occasion they refer to the Strength of Weak Ties. Burt is highly cited in all communities. Barabási is almost exclusively cited by scholars in his own community (by 43% of them).

Our analysis demonstrates how researchers play different roles that together generate countervailing forces which balance fragmentation and integration in the diffusion network. This process is driven by the work of key individuals, backed by collective citing behavior, that either integrates a new idea into existing or developing specializations or fills cultural and structural holes by connecting to other concepts and ideas. Scientific communities are a cultural and structural fabric consisting of strong ties between concepts and individuals, providing a context within which researchers can develop their work and make novel contributions that build on the community's cumulative knowledge. Through the lens cultivated by the Organizational Advantage Community, we see that whereas research communities provide a cultural context for researchers' scientific work to have meaning, the weak ties between research communities are where radical new ideas often emerge as a variety of knowledge is combined in innovative ways (Burt, 2004). The work of researchers is thus simultaneously and inextricably both cultural and structural. Employing the lens of the Ego-Network Community, scholars use both their knowledge and network as resources to advance their work and academic status (Lin, 1999; Lin and Dumin, 1986). Drawing on the Complex Networks Community's focus, we see that they inadvertently fuel the centripetal and centrifugal forces, which shape the cultural and structural network patterns we have analyzed in this article: a diffusion network in which different structural communities interpret and apply Granovetter's hypothesis in diverging ways.

Conclusion

This computational case study has studied the process by which a scientific idea is adopted and adapted as it spreads through scholarship, focusing on the case of Granovetter's (1973) Strength of Weak Ties hypothesis. We found that this hypothesis' diffusion path generates identifiable scientific communities, each of which develops its own interpretation of the hypothesis. Scholars in the various communities focus on different topics, ask different research questions, use distinct vocabularies, and advance the hypothesis in particular ways that fit into their overall research framework. Central figures around whom communities form play pivotal roles in this process; as scholars cite their publications, their work locally becomes a focal point for both the circulation and interpretation of the hypothesis.

Our analysis shows that a spreading idea is unlike viral diffusion or social contagion in that every event of transmission involves interpretation by the adopting scholar, consequently leading to a continuous transformation of the idea. Like a chameleon adopting the colors of its surroundings, the notion of weak ties takes on different guises, advanced by the interests and perspectives of the scholars redeploying and building on it. For some researchers, the Strength of Weak Ties is a universal self-organizing principle of complex networks that is not specific to any social context and can only be understood by considering and modeling the network as a whole. Other scholars find strength in weak ties due to their ability to increase the relative status of individuals in society, conceptualizing weak ties as an asset to an individual ego. Different communities use the same reference to make very different points.

Looking at Granovetter's original article on the Strength of Weak Ties (1973), we can in retrospect see the potential for the different interpretations which later emerged.[10] However, much like the varieties of plants developing from the same seed, the idea progresses in diverging directions as a result of interpretative actions and interactions of numerous scholars. This process of developing distinct interpretations of an idea functions structurally as a centrifugal movement in the diffusion network, fragmenting and separating its communities. This is in line with Burt's intuition: good ideas come about by bridging structural holes in social networks, but spread in ways that divide social groups (2004, p. 394). But we also identify centripetal forces in the diffusion process: several scholars in the network actively work across different communities, tying together ideas and fields, thus integrating the network as a whole.

In line with Latour (1984), this study suggests that translation, according to which both the circulation and the various meanings of an idea result from numerous actions and interactions among individuals, is a better model for the spreading of ideas than diffusion. Our methodology captures both the structural, macroscopic patterns that arise as a result of microscopic actions, namely diffusion communities centered around local

[10] In fact, there are traces of this in other literatures from that time as well, as seems to be the case for most ideas—a phenomenon also referred to as simultaneous invention or multiple independent discoveries (Merton, 1961). In 1972, William Liu and Robert Duff published an article called "The Strength in Weak Ties," proposing an argument similar to Granovetter's, and drawing upon his doctoral thesis.

hubs, and the changes in meaning that follow from numerous individual and collective interpretations and the development of new lines of research. Our results illustrate how these structural and cultural patterns are interrelated. We hope this will motivate researchers to look for other methodologies and approaches that integrate these insights and further our understanding of the mechanisms at play during diffusion-translation processes, in science and beyond.

One open question is to what degree the diffusion communities overlap with already existing scholarly communities or come about as a consequence of the spread and research potential of new ideas. Another avenue for future research would be to look deeper into the roles of influential scholars, to have a better sense of the extent to which they perform unique translation work or receive credit for doing so because of their status. With this article, we hope to suggest that further and more sophisticated development of these ideas will require scholars of a variety of methodological backgrounds.

Chapter 2.

Intersectionality on the Go: the Diffusion of Black Feminist Knowledge across Disciplinary and Geographical Borders

Published as

Keuchenius, A., & Mügge, L. (2021). Intersectionality on the go: The diffusion of Black feminist knowledge across disciplinary and geographical borders. *The British Journal of Sociology, 72*(2), 360-378

Abstract

Kimberlé Crenshaw coined the term "intersectionality" in 1989 as a critique of feminist and critical race scholarship's neglect of—respectively—race and gender. Since then, the concept has been interpreted and re-interpreted to appeal to new disciplinary, geographical and socio-cultural audiences, generating heated debates over its appropriation and continued political significance. Drawing on all 3,807 publications in Scopus that contain the word "intersectionality" in the title, abstract or keywords, we map the spread of intersectionality in academia through its citations. Network analysis reveals the contours of its diffusion among the 6,098 scholars in our data set, while automated text analysis, manual coding, and the close reading of publications reveal how the application and interpretation of intersectional thinking has evolved over time and space. We find that the diffusion network exhibits communities that are not well demarcated by either discipline or geography. Communities form around one or a few highly referenced scholars who introduce intersectionality to new audiences while re-interpreting it in a way that speaks to their research interests. By examining the microscopic interactions of publications and citations, our complex systems approach is able to identify the macroscopic patterns of a controversial concept's diffusion.

Keywords— intersectionality; sociology of knowledge; women's studies; diffusion; feminism; complexity science.

Introduction

Kimberlé Crenshaw coined the term "intersectionality" in 1989 as a critique of feminist and critical race scholarship's neglect of—respectively—race and gender. Focusing exclusively on either, Crenshaw argued, failed to apprehend the experiences of black women inhabiting the intersection of two dimensions of inequality. The idea that black women face different forms of exclusion than white women due to the intersection of sexism and racism was not new (e.g. Collins and Bilge, 2016; Wilson, 1978; Combahee River Collective, 1977; hooks, 1984). Yet, the term was novel. How has intersectionality travelled within academia since its coinage?

"Intersectionality" today is seemingly everywhere. Leslie McCall was already writing in 2005 that the concept "is the most important theoretical contribution that women's studies, in conjunction with other fields, has made so far" (2005, p. 1771). A "buzzword" with dedicated conferences, special issues and journals (Davis, 2008), intersectionality today is part of the standard curriculum of women's studies (Collins, & Chepp, 2013). Intersectionality has also broken out from its original moorings in feminist, legal and critical race scholarship to cross countries and continents, disciplines and subfields. At the time of writing, Google Scholar lists 59,900 publications on intersectionality, while Web of Science counts over 100 distinct research areas under its umbrella. Along its journey, "intersectionality" has been interpreted and re-interpreted to speak to its new disciplinary, geographical, socio-cultural and political surroundings.

Intersectionality is variously understood as a theory, a research paradigm and a strategy to transform power relations (e.g. Hancock, 2016). How the concept has evolved has also been heavily contested. Some argue that intersectionality's newfound popularity comes at the expense of black women, whose voices and knowledge rooted in lived experience has been erased (e.g., Jordan-Zachery, 2007). Others argue that race has been eclipsed by class in the hands of continental European scholars (Carbado et al., 2013) and that "whitewashed" intersectionality has lost its transformative potential (Bilge, 2013). Nash (2018) describes the "intersectionality wars" in which black feminists defend intersectionality from "misuse and abuse". By now nearly everything about intersectionality is contested: "its histories and origins, its methodologies, its efficacy, its politics, its relationship to identity and identity politics, its central metaphor, its juridical orientations, its relationship to 'black woman' and to black feminism" (Nash 2017, pp. 117–118). Scholars have therefore suggested that intersectionality should be defined by what it does, rather than by what it is (Cho et al., 2013).

The aim of this article is not to offer another reflection on what intersectionality is or does. Inspired by Mügge and colleagues (2018) - who trace intersectionality's journey within political science - we broaden the scope and empirically scrutinise how it travelled through networks consisting of thousands of scholars. How is the concept defined and applied across disciplines and geography? What is the role of individual scholars in this process? Drawing on all (n=3,807) publications in Scopus that contain the word "intersectionality" in the title, abstract or keywords, we map the spread of the concept through its citations. We use network analysis to study the citation structure and automated text analysis, manual coding and the close reading of publications to analyse how intersectionality has been interpreted and applied during its spread. Our complex systems approach focuses on the micro-interactions of publications and citations, and how these generate macro patterns of diffusion (Granovetter, 1973; Byrne, 1998) and interpretation (Abbott, 2001). Our contribution is twofold. First, rigorous empirical analysis improves our understanding of the multiple dimensions of intersectionality's spread and incorporation into the mainstream of many disciplines. Second, our study gives detailed insight into the process of the diffusion of scientific concepts and what happens if a new concept takes root in new disciplines. Confirming the worry of critical scholars, we find that interpretations, understandings and applications of intersectionality increasingly diverge from its original meaning and sources as it travels.

This process is similar to the diffusion of academic knowledge more generally.

In what follows, we first review the literature on the diffusion of intersectionality and its relation to the politics of knowledge production and the sociology of knowledge. We then detail our methods. Our findings are organized under four headings: (1) macroscopic patterns in the diffusion network of intersectionality scholars, (2) the role of disciplines and geography, (3) how different diffusion communities use and conceptualize intersectionality, and (4) the role of leading figures in the translation of the concept across disciplines and subfields. We find that how intersectionality is understood changes as the concept travels to new audiences. For example, the largest diffusion community consists of primarily US based scholars who see intersectionality as a tool to empower Black women. While the development of methodological tools to operationalize an intersectional lens to identity is a key concern for a diffusion community of psychologists. Leading figures—whom we call "hubs"—are central in introducing and translating the concept to their peers so that it becomes thematically, theoretically or methodologically interesting. This, at least, is the role they are credited with by scholars who cite them.

Theorizing Intersectionality's Journey

Works addressing the genealogy of intersectionality and the current structure and future prospects of the field contain numerous clues about the diffusion of intersectional thinking and scientific ideas more generally. Many of these works point to the central role of Crenshaw (1989), the role of disciplines and geography, and the politics of academic knowledge production. Our review of the key works generates five expectations about the diffusion of intersectionality.

Genealogies of intersectionality point out that intersectional thinking has a much longer history than the term itself; many refer to the speech "Ain't I a Woman?" by Sojourner Truth at the 1851 Women's Rights Convention in Akron (Brah and Phoenix, 2004). Born into slavery, Truth campaigned for its abolition and for equal rights for women; by focusing on the oppression of black women, she challenged essentialist thinking in single categories. Crenshaw (1989) is often referenced as the foundational article on intersectionality (Nash, 2016), with Crenshaw's location in law and critical race and feminist studies informing how intersectionality subsequently spread in academic publications.

Feminist and critical race scholars have studied the spread of intersectional thinking to other disciplines. Cho, Crenshaw and McCall (2013)—two legal scholars and a sociologist—reflect on two decades of scholarship in their special issue on the emerging "field of intersectionality studies" and describe a loosely connected patchwork of disciplinary islands, which they hope will be bridged to bring greater cohesion to the field. Cho and colleagues distinguish between two ways in which intersectional thinking spreads. The first process is *centrifugal*, when ideas travel and adapt to new disciplines; the second is *centripetal*, when scholars at the margins of their respective disciplines draw on literatures from further afield. The centrifugal process is driven by institutional forces that mould intersectional thinking to the methodological standards, practices and discourses of specific disciplines; centrifugal works include Hancock (2007) in political science, Cole (2009) as well as Purdie-Vaughns and Eibach (2008) in psychology, Choo and Marx Ferree (2010) in sociology, and Walby (2007) in philosophy. Cho and colleagues (2013, p. 807) further point to the relative privilege or marginality of intersectionality scholars, knowing that mainstream disciplinary work is credited more within academic institutions than critical interdisciplinary work.

A 2012 special issue edited by Devon Carbado, Kimberlé Crenshaw (law), Vickie Mays (psychology) and Barbara Tomlinson (literature) on intersectionality's travels highlighted the role played by geography and disciplines in intersectionality's diffusion and conceptualization. In the introduction, the editors emphasize the differences between European and US approaches. European scholars, they argue, often use intersectionality to articulate abstract interactions but are less attentive to race, which is deemed less important than class (cf. Lutz, Herrera Vivar, & Supik, 2011). Bilge (2013) argues that this European treatment has neutralized intersectionality's political potential.

In line with findings from the sociology of science, Carbado and colleagues (2013) find that contextual differences—be it geographies or disciplines—generate alternative engagements with the theory. Kathy Davis (2008) frames the spread of intersectional thinking as a success story, which she attributes to the open-ended ambiguity of the initial theory. Davis draws on the work of sociology of science scholar Murray S. Davis (1971; 1986), who posits that novel scientific theories must be specific enough to be of interest to experts in the field. The theory should also be open and incomplete enough so that scholars in other fields can adjust it to their interests and be encouraged to build on it.

Collins and Chepp (2013) identify six core ideas addressed by intersectional thinking: interrelations between systems of power; the co-construction of knowledge and power; attention to relational processes; the co-construction of knowledge and social relations; the significance of boundaries; and a concern for complexity. Particularly the last three themes are relevant for our study. The co-construction of knowledge and social relations refers to the idea that standpoints and world views—and not just social relations—are relational and construct each other (Collins, 1993; 1990). Following Collins and Chepp, we argue that social relations between academics influence the production and diffusion of knowledge. The role of boundaries refers to the construction of group identities; here the authors argue that intersectionality has been successful in transcending disciplinary boundaries within the academy. The concern for complexity connects intersectionality to complexity science, which can be seen as a diffused field or a "collection of work that addresses fundamental questions on the nature of systems and their changes" (Walby 2007, p. 449). Both intersectionality and complexity science interrogate system complexity, privileging notions such as emergence, the relation between micro-interactions and macro patterns, and non-linearities.

Building on this extant work on the spread of intersectional thinking, we expect the following: first, the trail of intersectionality's spread will appear as clusters of disciplinary communities loosely connected by scholars working at their margins. Second, communities will be tied together geographically. Third, interpretations of intersectionality will correspond to scholars' disciplinary and geographical locations. Fourth, Crenshaw (1989) will be referenced by nearly all scholars and will be the most central scholar in the network. Fifth, each community will have local central scholars like scientific stars (Merton, 1968) or leaders of invisible colleges (Crane, 1972; Carley, 1990)—likely established scholars within their disciplines.

Data and Methods

The diffusion of intersectionality is a complex process of micro-interactions between scholars referencing and building on each other's work. To reveal regularities and exceptions in this process, we construct a network representing the diffusion of intersectionality in terms of citations. We analyse the macroscopic structures of this network and how these relate to geography and disciplines. Consequently, we investigate how intersectionality is used and adapted by individual scholars and communities in the network. This methodology allows us to study the entire trail of intersectionality including its spread among scholars, conceptual journey and how these two relate

Our sample includes data on publications in the Scopus database with "intersectionality" in the keywords, abstract or title. We retrieved: author(s), title, journal, publication date, author research areas, keywords, abstract, and references. Although Scopus has broad coverage, it privileges journal articles over books and book chapters (Mongeon, & Paul-Hus, 2016). We therefore manually included all publications that received more than 30 references from publications in our sample but which were missing from Scopus (see Appendix A). We also retrieved meta-data on journals' subject areas from Scimago Journal & Country Rank. Our dataset contains 3,807 publications authored by 6,098 scholars, published between 1983 and November 2018.

Network analysis enables us to reconstruct intersectionality's journey. Nodes in the network represent authors (n=6,098) who have published on intersectionality. Edges in the network are drawn from new scholars publishing on intersectionality (edge source) to previously published intersectionality scholars whom they cite (edge target). These directed edges represent influence from earlier to later authors. When publications are co-authored by multiple authors publishing on intersectionality for the first time, we draw edges between them. This generates a diffusion network that includes 6,098 scholars (nodes) and 45,264 edges.

For the analysis of the community structure of this network—the degree to which the network can be split into communities of scholars that predominantly reference each other— we use the Leiden algorithm (Traag, Waltman, & van Eck, 2018). We determine statistical significance by comparing the network's community structure to that of a random network with the same degree distribution (see appendix B). Additionally, we analyse the in-degree distribution of the network—the number of incoming edges each scholar obtained—over time and the location of the high in-degree scholars in the network. The in-degree is a proxy for the scholar's importance in diffusing the ideas of intersectionality to their peers. Finally, we investigate the relation between detected communities and their geographical and disciplinary constitution.

To provide an overview of how intersectionality has been adapted by scholars in the network, we use topic modelling followed by the close-reading of key publications (Törnberg, & Törnberg, 2016). Topic modelling is an unsupervised machine learning method that identifies topics in large textual datasets, allowing us to identify principal themes and frames in the data (Bail, 2014; DiMaggio, Nag, & Blei, 2013). We use a topic modelling technique called Latent Dirichlet Allocation (Blei, Ng, & Jordan, 2003; Pritchard, Stephens, & Donnelly 2000) that outputs topics—list of words—present in the abstracts in our dataset. To investigate the relationship between communities and topics, we examine whether scholars of different communities engage with distinct topics. We have set the model parameter for the number of topics, a contested value in the literature, to fifteen, but found similar relations between topics and diffusion communities for higher and lower parameter settings (see appendix C for more details).

Whereas topic modelling provides a bird's eye view of how different scholarly communities narrate intersectionality, the close reading of key publications helps us to see in granular detail how groups of scholars conceptualize intersectionality. We are aware that our own positions as white female researchers employed by a wealthy institution in a western democracy may influence our readings (Labelle, 2020). To circumvent this bias we picked a random sample of publications from each community, between 25 and 100 depending on the community's size, to explore how authors use intersectionality and refer to key publications and scholars in their community. This manual coding consisted of first selecting passages that reference key figures within communities and their publications, and then identifying common themes and narratives within these passages.

Mapping the Structure of Intersectionality's Diffusion

Figure 1 shows that the diffusion network has a clear community structure (modularity value=0.60, p-value<0.01 see Appendix B). Intersectionality did not spread like an oil stain, evenly and outwards from a single centre. Instead, the trail shows multiple centres and local webs within the 6,098 scholars in our dataset, much like the loosely connected arenas theorized in the literature (Carbado et al., 2013). Whereas the network can be categorized into communities, these are not segregated. The three largest communities comprise 42 per cent of all scholars in the giant component, and the largest 12 communities (each size >100), 86 per cent of all scholars (see figure 2). Our analysis focuses on these 12 communities.

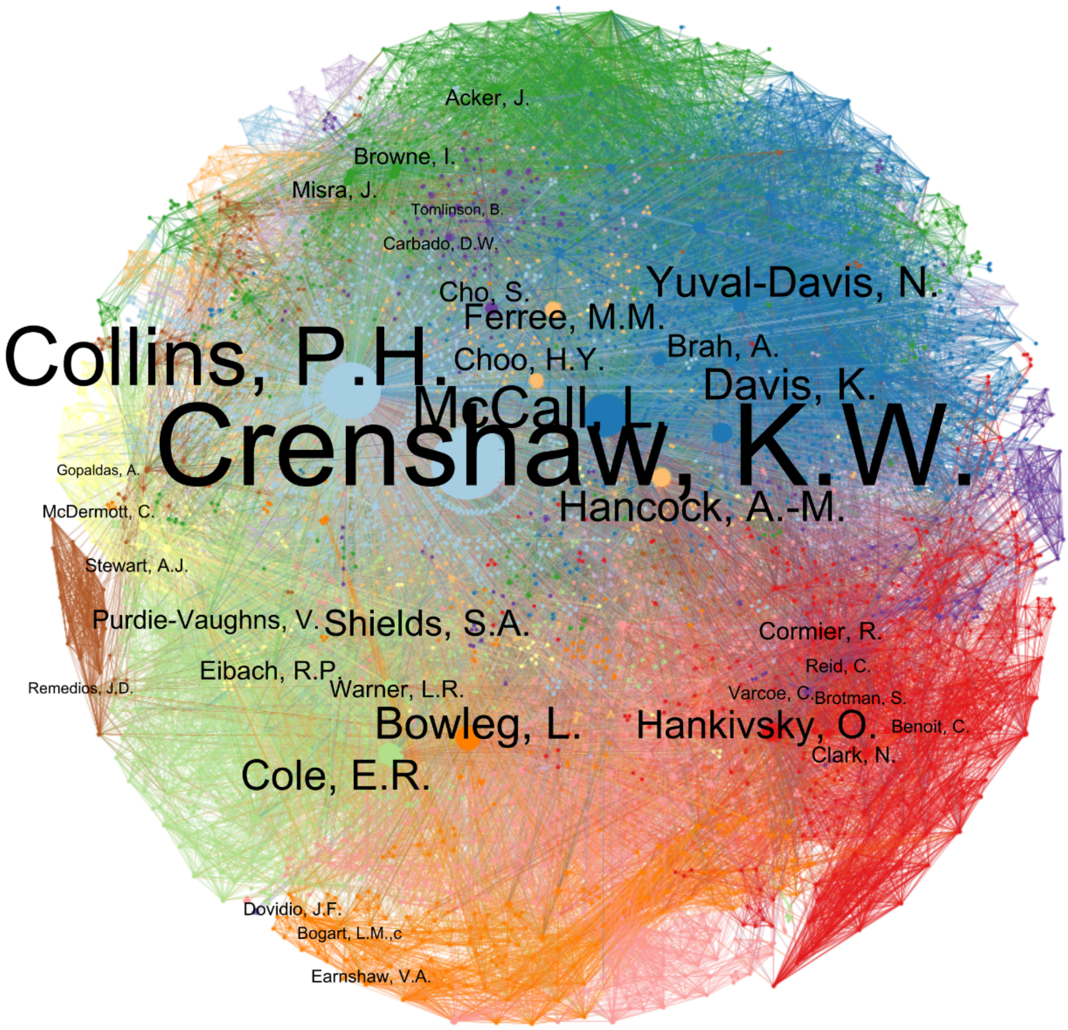


Figure 1: The diffusion network of intersectionality. The nodes are scholars who have published on intersectionality. Directed edges are drawn from scholars publishing on intersectionality for the first time (edge source) to published scholars whom they cite (edge target). The nodes are coloured by community. The most important scholars for the diffusion of intersectionality are labelled, with the labels sized according to their in-degree.

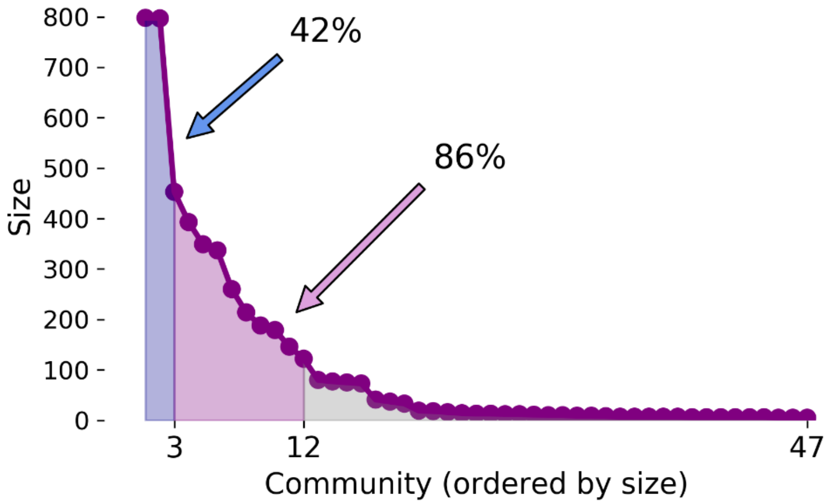


Figure 2: Distribution of community size in the diffusion network, with a small number of large communities and a large number of small communities. The largest 3 and 12 communities respectively contain 42% and 86% of all scholars in the giant component of the diffusion network.

The in-degree—the number of incoming edges—is very unequally distributed in the network. The most influential scholars—e.g., Kimberlé Crenshaw, Patricia Hill Collins, Leslie McCall, Elizabeth Cole, Lisa Bowleg, Kathy Davis, Nira Yuval-Davis, Ange-Marie Hancock and Olena Hankivsky—are each cited by 448 up to 2,320 distinct scholars. Most other scholars (90 per cent) receive less than 13 references. Scholars with high-in degree can be seen as hubs in the diffusion of intersectionality since later scholars reference hubs' works in their first intersectionality publication. We find that the hubs are spread across various communities (see figure 1). In-degree within communities is likewise very unequally distributed. Most communities contain one or several hubs (e.g., Bowleg in community 7, Olena Hankivsky in community 5) cited by between 25 and 58 per cent of community members.

While all communities grow exponentially, the speed and timing of their growth differ (figure 3). Community 0, a US-centred community around Crenshaw and Hill Collins, grows first; community 2, located in psychology, only takes off after 2005. In this growth, the hubs are often forerunners in their respective communities (figure 4).

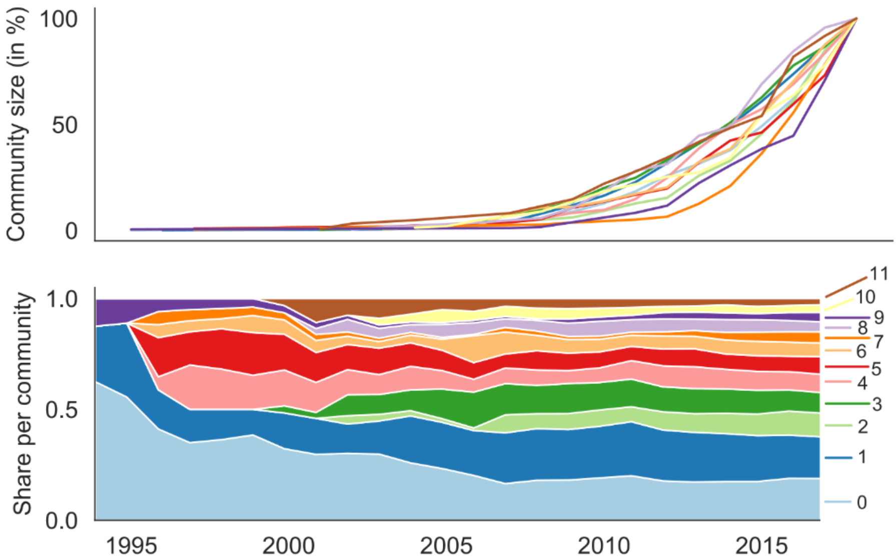


Figure 3: Temporal evolution of the largest 12 communities. The top figure shows the communities' growth curves, most of which are exponential. The bottom figure shows each community's share of total scholars at different points in time. Some communities (0 and 1) emerged early, others (2, 3 and 10) later.

Given their central location and timing, we can view these hubs as scientific opinion leaders in a two-step flow of communication (Katz, 1957). Innovations first spread to a small number of opinion leaders who in turn diffuse it to their followers. Similar leading roles exist in the diffusion of scientific innovations (Crane, 1972; Carley, 1990). But before turning to this, we ask: how do community structures in the diffusion network relate to geography and disciplines?

The Role of Geography and Disciplines

The scholarly communities in the diffusion network of intersectionality are to some extent informed by geography (figure 5). Communities 1 and 3, for example, are dominated by scholars based respectively in continental Europe and the UK. This is in contrast to all other communities, in which the vast majority of scholars—from 57 per cent in community 5 to 88 per cent in community 10—are based in the United States. This overview suggests that while geography has influenced the diffusion of intersectionality, it is far from the only compass. The data do not present neatly demarcated geographical communities.

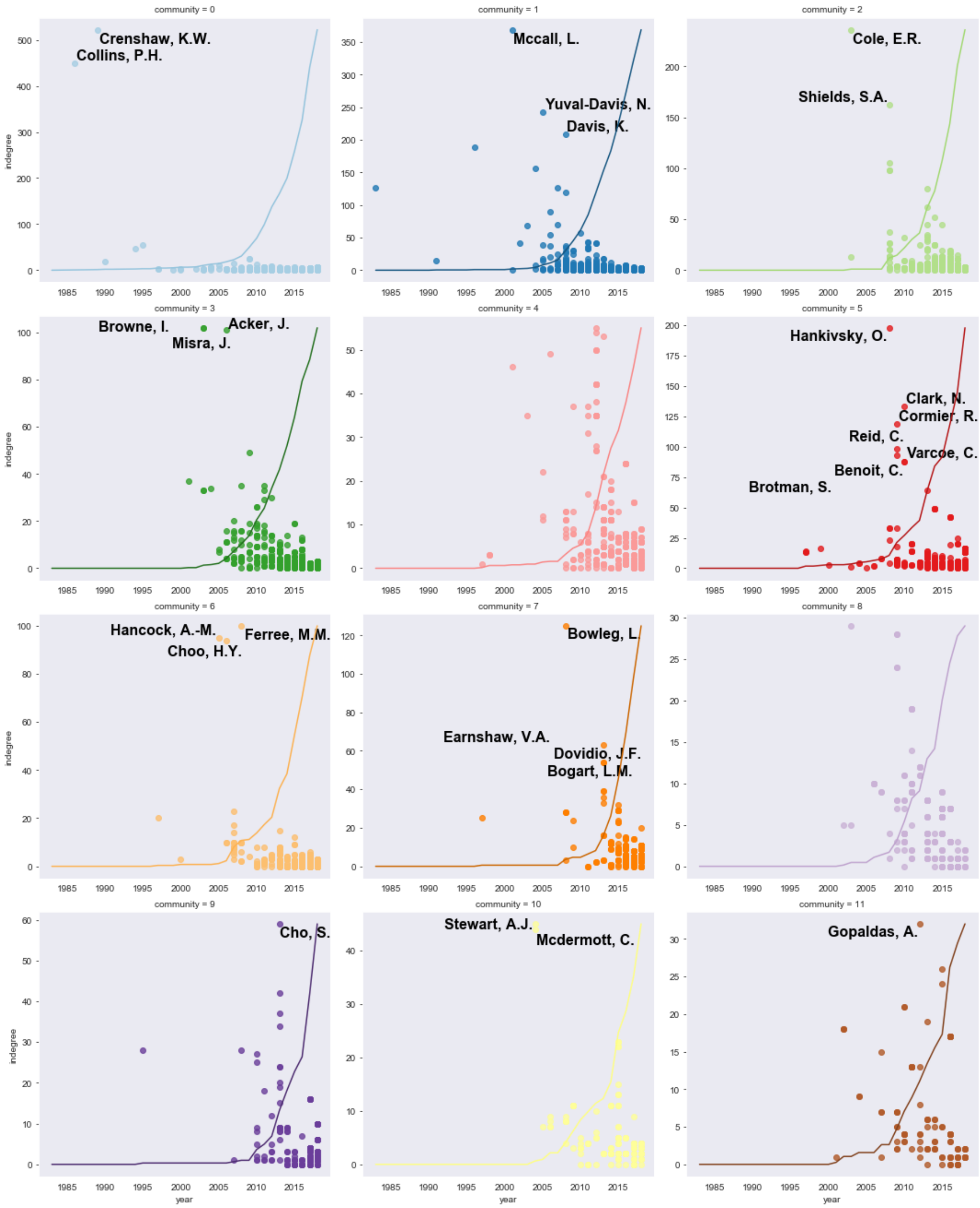


Figure 4 (previous page): The growth (line) and local—within community—in-degree of researchers (scatter) in each community over time. The hidden y-axis for growth runs from 0 to 100%. Local hubs (scholars referenced by more than 25% of their community) are labelled. Most communities have at least one hub, among the first in the community to publish on intersectionality.

georegion	AF	AS	EU	NA	OC	SA	UK	Total
community								
0	2.8%	2.1%	6.9%	75.2%	1.4%	1.6%	10.0%	100 %
1	1.3%	4.6%	35.9%	25.0%	6.4%	2.0%	24.9%	100 %
2	0.9%	0.7%	5.2%	86.5%	4.7%	0.0%	2.0%	100 %
3	2.1%	7.3%	21.3%	36.9%	8.1%	1.0%	23.4%	100 %
4	0.0%	1.5%	7.3%	82.3%	4.1%	0.3%	4.7%	100 %
5	4.9%	10.7%	13.4%	57.3%	3.4%	1.8%	8.5%	100 %
6	0.4%	0.4%	11.3%	79.8%	2.7%	0.8%	4.7%	100 %
7	2.9%	0.5%	1.4%	84.7%	4.3%	2.4%	3.8%	100 %
8	0.0%	4.4%	11.6%	68.0%	9.9%	0.0%	6.1%	100 %
9	1.1%	0.0%	7.5%	76.4%	7.5%	0.0%	7.5%	100 %
10	0.0%	2.1%	2.8%	88.0%	0.7%	0.0%	6.3%	100 %
11	0.0%	6.7%	10.9%	68.9%	1.7%	0.0%	11.8%	100 %

Figure 5: Geography of diffusion communities. Each cell presents the percentage of scholars (row) based in this geographical area (column). Significantly high or low column cells values are coloured green (high) or pink (low), based on a two-sided z-test with alpha=0.05. The geographical areas are Africa (AF), Asia (AS), Europe (EU), North America (NA), Oceania (OC), South America (SA) and the UK separately. Values are based on the location of scholars' current institution, available in Scopus.

How do disciplines tie scholars together? Cho and colleagues (2013) argue that institutional forces pull scholars towards their respective academic disciplines, subjecting intersectionality to established power structures and research practices, while marginalized scholars often remain sceptical of integrating mainstream methods and theories into their intersectional research. The latter act as centripetal forces, rendering the field a more connected and cohesive whole. Academic disciplines have been identified as the main foci around which the work of intersectionality scholars is organized. In our data and diffusion network, this would mean that communities are organized around disciplines, with scholars on the margins forming ties between communities.

Figures 6 and 7 visualise the disciplinary embedding of the communities. On the hand, some communities stand out in terms of their research disciplines. For example, 22 and 16 per cent of scholars in community 7 publish in “Public Health, Environmental and Occupational Health” and “Health Social Science” journals respectively. Scholars of community 11 are unique in publishing on intersectionality within business, econometrics and marketing. On the other hand, some communities are very alike in their disciplinary focus yet remain separate in the diffusion network, such as the two largest communities (community 0 and 1).

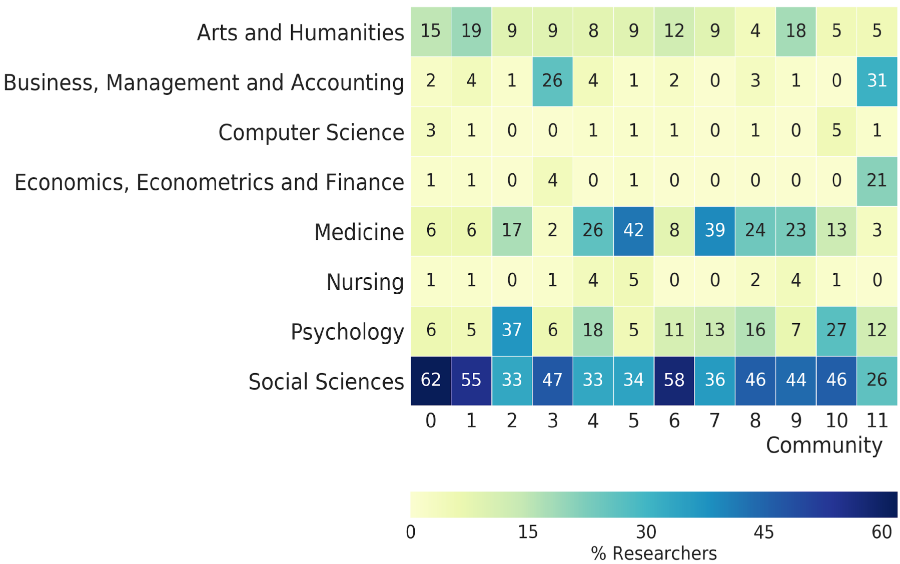


Figure 6: Main research areas of the communities. Each cell value and colour represents the percentage of researchers of that community active in a particular research area (e.g. 26% of researchers in community 3 published in business, management and accounting). The figure only contains research fields for which at least one community significantly deviates from the overall network (two-sided Z-test) and which involve at least 5% of the community’s scholars.

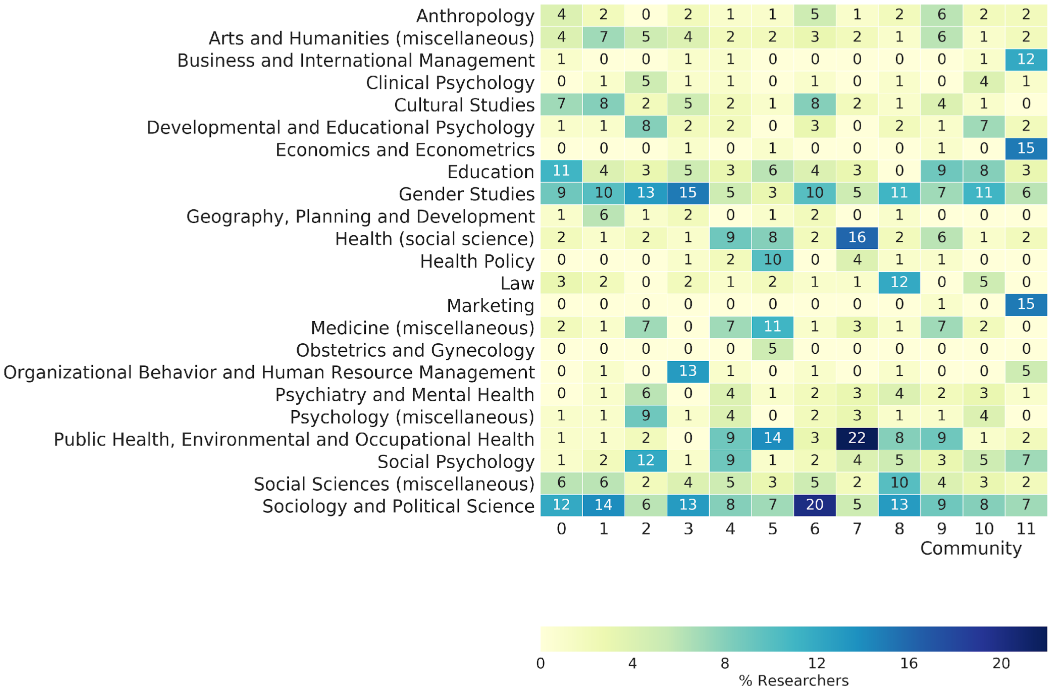


Figure 7: Sub-areas of research for the communities. Each cell value and colour represents the percentage of researchers of that community active in a particular subfield (e.g. 22% of researchers in community 7 published in public health, environmental and occupational health). The figure only contains research fields for which at least one community significantly deviates from the overall network (two-sided Z-test) and which involve at least 5% of the community's scholars.

These results allow for many more detailed observations, but the key take away is that disciplinary forces have indeed shaped the spread of intersectionality but, like geography, cannot fully account for the observed community patterns. To better understand how diffusion communities emerged, we delve deeper into how scholars in these communities narrate intersectionality.

Community-specific Adaptations

Now that we established the structural patterns of intersectionality's spread we explore its conceptual journey. Building on the community structure, we study how scholars in these communities understand and apply intersectionality. To do so we turn to the content of their publications. Aided by topic modelling, we find the intersectionality literature covering topics ranging from migration to domestic violence and stigmatization (see appendix C). Figure 8 shows the relation between the diffusion communities and the topics they write about, illustrating differences in the communities' research narratives and interests. For example, community 7 is interested in "stigmatization" (topic 2), which hardly registers in other communities. Community 2, consisting mostly of social psychologists, focuses on "multiple identities" and "sexual orientation" (topics 11 and 12).

Although the distribution of topics provides general insight into the interests of scholars in the various diffusion communities, it does not yield granular understanding of how intersectionality is interpreted and narrated. We therefore describe in more detail discussions within the network's three largest communities, which we have labelled "The Black Feminist Core", "Categorically Extended Intersectionality" and "The Intersectional Psychologists". Each community has a distinct understanding of intersectionality. For the predominantly US-based scholars of the "Black Feminist Core", improving the lives of black women is central to the intersectional project. Scholars within the "Categorically Extended Intersectionality" community—largely based in continental Europe and the UK—focus on interdisciplinary women's studies and treat intersectionality as an analytical framework and work-in-progress. They bring in more categories than race and gender and tend to focus on ethnicity and migration background rather than race. Finally, "The Intersectional Psychologists" focus on the methodological questions of intersectional research in individual psychology.

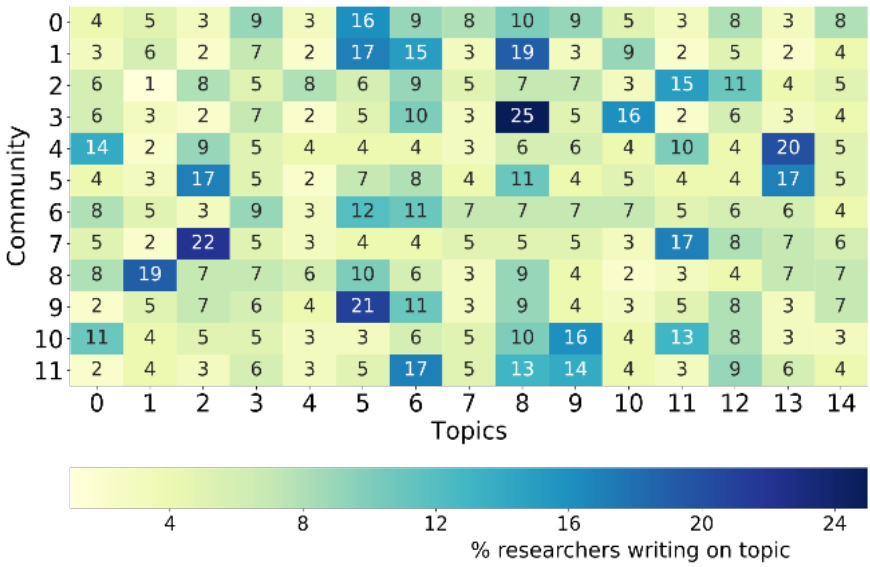


Figure 8: Topics (columns) that the members of communities (rows) cover in their publications. Each cell value and colour represents the percentage of a community's researchers addressing the topic (e.g. 21% of researchers in community 2 address topic 11, sexual identity and orientation).

Community 0: The Black Feminist Core

This community (798 scholars) is centred around the founders of intersectionality: Crenshaw and Hill Collins. Although the three most cited works in this community are canonical and cited by scholars in other communities too, they are particularly frequently referenced by scholars in community 0 (see Appendix D). Hill Collins' book *Black Feminist Thought* (1990) is referenced by 42 per cent and Crenshaw (1989) by 33 per cent of scholars in Community 0. Surprisingly, only a third cite Crenshaw (1989), which many reviews consider to be *the* conceptual birth of intersectionality. Crenshaw (1991) is referenced by 30 per cent of the scholars. Crenshaw's articles are located in law and closely related to critical race theory.

Angela Harris, part of the inner circle of critical race studies scholars who gave birth to intersectionality, argues that the voices of black women are too often ignored in feminist and legal theory and that the gender essentialism in much feminist theory perpetuates the problem. In her critique of second wave feminists espousing a putative "women's experience" (Harris, 1990, p. 588), Harris builds on the work of the American writer, feminist and civil rights activist Audre Lorde (1984). Reflecting on the field in her foreword to *Critical Race Theory* edited by Richard Delgado and Jean Stefancic (2001), Harris recalls a 1989 workshop attended by Derrick Bell, Kimberlé Crenshaw, Richard Delgado, Mari Matsuda and Patricia Williams. Since then, critical race theory has "exploded from a narrow subspecialty of jurisprudence... of interest to academic lawyers into a literature" spanning departments (2001, p. xx).

Characteristic of this community is the view that intersectionality should be used to improve the lives of black women. The majority of publications in this community (75 per cent) are written by North American scholars and focus on the US experience. A strong activist tone suffuses the work of this community, whether it is addressing its research subject of marginalized black women or the current and future direction of intersectionality. In their edited volume, Bonnie Thornton Dill and Ruth Enid Zambrana treat intersectionality as "a systematic approach to understanding human life and behaviour that is rooted in the experiences and struggles of marginalized people" (2009, p. 4). Their mission is to: (1) rethink curricula and promote institutional change in higher education, (2) apply knowledge to create a society in which all voices are heard, and (3) advocate for public policies that are responsive to multiple voices (2009, p. 2). Researchers in this community are generally critical about using intersectionality for pursuits other than empowering black women (Collins and Chepp 2013). In this community, intersectionality is conceptualized as tool to unveil and change systems of oppression, for marginalized black women in particular.

Community 1: Categorically Extended Intersectionality

The second largest community, consisting of 797 scholars, reveals how intersectionality has crossed the Atlantic; it includes, in both absolute and relative terms, the highest number of scholars based in continental

Europe and the UK (see figure 5). The community's main contribution is applying intersectionality to categories beyond race and gender. Its members thereby treat intersectionality as an analytical framework that is not specific to black women per se.

The central figures in this community are sociologists: McCall (based in the US), Kathy Davis (based in the Netherlands) and Nira Yuval-Davis (based in the UK). Their geographical location influences how they frame and apply intersectionality: while scholars based in the US and the UK largely focus on race, those based in continental Europe generally focus on ethnicity, applying the intersectional lens to individuals with migration or multi-ethnic backgrounds (Prins, 2006). Scholars in this community also introduce new disciplinary approaches from political science (e.g., Verloo, 2006), psychology (e.g., Staunæs, 2003) and geography (e.g., Valentine, 2007).

Many European and UK-based scholars apply intersectionality to a wider set of categories. Valentine (2007), for instance, brings in ability, arguing that theories of intersectionality overestimate the ability of individuals to create their own lives. Verloo (2006) analyses how categories are represented in policies and how these are linked to inequality in terms of gender, race/ethnicity, sexuality, and class; she uses intersectionality to show that a one-size-fits-all approach to multiple discrimination, based on the assumption of the sameness of social categories, is inadequate. UK-based scholars (Yuval-Davis, 2006; Brah, & Phoenix, 2004) in this cluster working in the tradition of Crenshaw and Collins see potential for intersectionality in the praxis of feminism and see opportunities for global feminism in the context of global threats.

US-based authors in this cluster are concerned with what intersectionality is and/or does. While Cho and colleagues (2013) focus on collaboration, Nash (2008) is more critical. Drawing on critical race legal scholars such as Harris, Crenshaw and Matsuda central in the "Black Feminist Core", Nash argues that intersectionality aims to disrupt cumulative approaches to identity. "Re-considering intersectionality enables activists to ask under what conditions organizing as ‘women’ or ‘blacks’ or ‘black women’ makes sense, under what conditions temporary coalition-building makes sense” (Nash, 2008, p. 4). Nash (2016) also criticizes scholars in this community for re-reading intersectionality’s inaugural text and rewriting intersectionality as a feminist contribution driven by disciplinary politics.

Interestingly, Nash (2008) and Bilge (2013)—the most ardent critics of the broad appropriation of intersectionality—are part of this community that widens intersectionality's scope. This shows that many scholars are taking notice of their criticisms by citing them. This makes Nash and Bilge, perhaps to their own discomfort, part of this diffusion community. Diffusion communities are far from homogenous academic communities that think alike; their members may indeed be unaware of being part of the clique. Nevertheless, diffusion communities lay bare the trail of how intersectionality has spread. Nash, alongside other high in-degree scholars, has been crucial in diffusing and narrating intersectionality to this community.

Community 2: The Intersectional Psychologists

This community (453 scholars) revolves around intersectionality in psychology. Scholars publish predominantly in the field of psychology including its subfields social psychology and developmental and educational psychology. A key objective is to develop tools to study intersectionality empirically at both the individual and structural levels. “Sexual identity” and “orientation” are among its leading research topics (topic 11). The journal Sex Roles is the community’s preferred outlet, publishing more than half of its top 15 publications and two special issues on intersectionality edited by Stephanie Shields (2008) and Mike Parent and colleagues (2013). The paper “Intersectionality and research in psychology” by Elizabeth R. Cole (2009) is referenced by almost half (46 per cent) of the community’s members, making Cole and Shields its principal hubs.

As psychology largely focuses on individuals and the intersectional lens challenges the discipline's quantitative and empirical orientation, scholars in this community frequently discuss methodological questions. How, for example, can regression analysis be combined with an intersectional approach? Bowleg notes that "the positivist paradigm that undergirds much (but not all) quantitative research appears to be orthogonal to the complexities of intersectionality" (2008, p. 317). Several highly cited publications offer "best practices" for applying intersectionality to psychological research (e.g., Warner, 2008; Purdie-Vaughns, & Eibach, 2008). Members of this community hold fewer meta-discussions about the origins, state and purpose of intersectionality, focusing instead on operationalizing the concept for empirical psychological research.

The detailed descriptions of the largest three communities demonstrate that diffusion communities closely relate to specific interpretations of intersectionality. These interpretations presumably developed in diverging directions and in conjunction with the growth of the communities. To gain a deeper understanding of this interpretation process we take a closer look at the role of communities' hubs.

The Emergence of Figurehead

What is the role of hubs and the two-step flow of communication in the diffusion and interpretation of intersectionality? The latter identifies two phases in the diffusion of new ideas, where the innovation first spreads to opinion leaders and thereafter to their followers. To examine this leading role, we focus on the communities with the most prominent hubs (see table 1) in different disciplines: McCall, Davis and Yuval-Davis in community 1 (sociology); Elizabeth R. Cole and Stephanie Shields in community 2 (psychology); Olena Hankivsky in community 5 (public policy); and Lisa Bowleg in community 7 (psychology). We explore how Collins, Crenshaw, and these hubs are referenced and how their work is narrated based on the coding of a significant number of publications with references to these scholars (table 1).

community	name	nr. edges from community members	% edges from community members	nr. citations analyzed
1	McCall, L.	369	46%	55
	Yuval-Davis, N.	243	30%	39
	Davis, K.	209	26%	40
2	Cole, E.R.	236	52%	96
	Shields, S.A.	162	36%	55
5	Hankivsky, O.	198	59%	47
	Clark, N.	133	39%	
	Cormier, R.	119	35%	
	Reid, C.	98	29%	
	Benoit, C.	93	28%	
	Brotman, S.	88	26%	
	Varcoe, C.	88	26%	
7	Bowleg, L.	125	58%	27
	Earnshaw, V.A.	63	29%	
	Bogart, L.M.,c	54	25%	
	Dovidio, J.F.	54	25%	

Table 1: Hubs in communities 1, 2, 5 and 7, and the number and percentage of first-time intersectionality scholars in their community who reference them.

While Collins and Crenshaw are often referenced when authors write about the origins of intersectionality or are providing a definition of the term, far from every new scholar references Crenshaw or Collins (see appendix D). Different communities also refer to specific contributions by Collins or Crenshaw which speak to their research interests. For example, community 7 references Collins almost exclusively in relation to stigmatization, particularly HIV-related stigma, which is the community's main research topic: "For midlife and older Black women, manifestations of HIV-related stigma intersected with and was compounded by various forms of inequality rendered through ageism, racism, and sexism, what Patricia Hill Collins (1990) has described as a matrix of oppression" (Sangaramoorthy, Jamison, & Dyer 2017, p. 1338). Hubs are often credited for their translation work. For instance, Hankivsky is accredited for introducing intersectionality to health research and public policy, the main interest of community 5: "The paradigm of intersectionality (Crenshaw, 1994[1991]), proposed in the field of women's health, has been highly useful in understanding the interplay between systems of power and oppression on the structural level (Hankivsky et al. 2010)" (Mora-Rios, Ortega-Ortega, & Natera, 2016, p. 698).[11]

Similarly, Cole—community 2's hub—is explicitly praised for her work in translating intersectionality for the field of psychology:

> The construct of intersectionality has been used extensively by feminists, queer theorists, and critical race theorists; however, it has been only recently that scholars within our own fields of counseling and psychology have pointed to intersectionality as a critical analytic tool in understanding the experiences and consequences of holding membership in multiple social identity categories (Cole 2009; Conwill 2010) ... Cole (2009) has provided an excellent guide for how to integrate the rubric of intersectionality into psychological research. Cole also highlighted the bias in the literature on intersectionality toward the investigation of those who experience multiple dimensions of disadvantage.
>
> (Smith, & Shin, 2015, p. 1462)[12]

This last passage also reveals a process of academic positioning, identifying the author and reader as part of "our field of counselling and psychology".

Sometimes Crenshaw and Collins are no longer referenced but eclipsed by the community hub:

> The related concept of intersectionality, which suggests that social categories and identities are not independent but rather multidimensional and linked to structural inequalities (Bowleg et al. 2013), provides a useful reference in understanding how layered stigma works.

[11] This citation is one example of numerous quotes in our sample that reference Hankivsky in this manner in community 5. We analysed a random selection of 25 publications, which contained 47 references to Hankivsky.

[12] Similar to the previous Hankivsky quote, this citation is an illustration. The analysis is based on 90 randomly selected publications by scholars of community 2, which include 96 references to Cole.

> However, while theory and research highlight the importance of understanding layered stigmas and intersectionality in relation to HIV vulnerability among BMSM, these factors have been largely overlooked in most quantitative research.
>
> (Wilson et al., 2016)[13]

Our exploration of the role of community hubs shows that these scholars are not only introducing their peers to the idea of intersectionality but translating the concept in ways that make it relevant to their particular disciplines, fields and subfields. Scholars who reference these leading scholars reinforce their role as hubs by creating narratives that credit them this role. Hubs thus seem to function as scientific opinion leaders or focal points (Collins, 1983) in a chaotic academic landscape that helps stabilize the concept of intersectionality. In extreme cases, we find that the origins of intersectionality have been forgotten as later references to the concept only cite these hubs, who become figureheads for intersectionality in their own communities. This aligns with what Cho and colleagues (2013) describe as the centrifugal process in the diffusion of intersectionality.

Conclusion

Thirty years after its coinage, intersectionality has entered most disciplines that study people in some way. While there is no shortage of critical interventions that question the competing interpretations of what intersectionality is, does, or should be, our study—to the best of our knowledge—is the first systematic empirical attempt to combine quantitative and qualitative methods and a complex systems approach to reconstruct the macro and micro dimensions of intersectionality's spread through the academic literature.

[13] This citation illustrates how the hub Bowleg is referenced alongside a definition of intersectionality, without citing Collins or Crenshaw. Our data showed the same phenomenon for other hubs.

In contrast to extant genealogies of intersectionality that tend to focus on highly cited works and those that are central to specific circles (for an exception in political science see Mügge et al. 2018), our study draws on all (n=3,807) works available on Scopus that include the word intersectionality in the abstract, keywords or title. Based on the extant literature, we formulated five expectations: (1) the trail of intersectionality's spread will resemble clusters of disciplinary communities; (2) communities will be tied together geographically; (3) scholars' interpretations of intersectionality will correspond to their geographical and disciplinary locations; (4) Crenshaw (1989) will be the most referenced work; and (5) each community will have its own scientific star.

Our findings reveal that intersectionality's diffusion trail is made up of highly connected webs within the 6,098 scholars in our data set. Within each of these communities, we find a few central and highly referenced scholars—whom we have referred to as "hubs"—who were crucial in introducing the concept to their peers. While these communities are oriented around disciplines—and to some extent, geography—our analysis suggests that they mostly form around specific narratives of intersectionality. For example, a "Black Feminist Core" of scholars based in the US considers intersectionality primarily as a tool to empower black women, while another large community made up primarily of psychologists seeks to operationalize intersectionality for psychological research on identity. The hubs are influential in creating these narratives of intersectionality for their respective communities, while their roles are recognized and reinforced by other scholars in the community. For example, scholars in the community around Hankivsky credit her for "bringing intersectionality to the field of women's health research". While Crenshaw has the most central position in the overall diffusion network, acknowledged for both coining and defining intersectionality, she is not consistently referenced. At times intersectionality is introduced with a reference to the community's local hub, transforming the hub into a figurehead of intersectionality for this community. Previous studies underline the importance of academic stars or opinion leaders in the diffusion of ideas due to their status and reach (Carley, 1990; Price, 1963; Crane, 1972). Our study reveals that these hubs *also* translate ideas in ways that make sense to their surroundings. Scholars citing these hubs reinforce these new narratives. This way, hubs are credited for their role as translators, and references to the original works in some cases disappear.

The diffusion pattern of intersectionality supports broader findings from the sociology of knowledge, particularly how researchers' social relations inform the knowledge they produce (Collins, & Chepp, 2013). Academics self-organize into social circles (Crane, 1972) or epistemic communities (Knorr Cetina, 1981) that uphold particular stories and knowledge claims. New scientific theories are transformed and redeployed as they traverse academic landscapes (Kaiser, 2009; Keuchenius et al, forthcoming). Generally, scholars in various research communities will agree on the importance of the novel theory, but will–often not knowingly—disagree on the particular content (Kuhn 1970, p.44; Gilbert and Mulkay, 1984). In that light, intersectionality's journey is no exception. Unlike Davis (2008), who ascribes the success of intersectionality to its ambiguous and open-ended nature, we suggest that the multiplicity of perspectives that developed during intersectionality's spread is a precondition and natural consequence of a novel idea that travels far. What is unique to intersectionality, is the—often heated and political—contestation that accompanied the transformations of the concept.

Our analysis focused on the most notable patterns in the diffusion network of intersectionality—its community structure and the existence and role of local hubs—which correspond to the centrifugal spreading process of intersectionality. While the 3,807 publications in Scopus that contain the word "intersectionality" in the title, abstract or keywords represent the visible "elite" within intersectionality studies, we expect that there are many more works produced by scholars of colour and other marginalized groups underrepresented and excluded in academia (Cho et al., 2013).

Although we did not pursue the in-depth analysis of centripetal actors, the diffusion network detected scholars working on the margins of communities and at times bridging them. Future research will be needed to examine their role in the production and diffusion of knowledge. Additionally, we identify a novel research avenue on the emergence of hubs. We analysed their leading role, but the question remains why certain scholars—and not others—acquire a central network position. Finally, our complex systems approach—which focuses on the emergence of macroscopic patterns rooted in microscopic events and interactions—does not explicitly capture power imbalances and racialized hierarchies that influence knowledge production and diffusion. Nevertheless, macro structures including power inequalities and institutional incentives feed back into individual actions and interactions. We hope that our work will inspire scholars to explore methods able to incorporate such feedback loops into systematic empirical research on intersectionality.

Chapter 3

Why it is Important to Consider Negative Ties when Studying Polarized Debates: a Signed Network Analysis of a Dutch Cultural Controversy on Twitter

Published as

Keuchenius, A., Törnberg, P., & Uitermark, J. (2021). Why it is important to consider negative ties when studying polarized debates: A signed network analysis of a Dutch cultural controversy on Twitter. *PloS one*, *16*(8), e0256696

Abstract

Despite the prevalence of disagreement between users on social media platforms, studies of online debates typically only look at positive online interactions, represented as networks with positive ties. In this paper, we hypothesize that the systematic neglect of conflict that these network analyses induce leads to misleading results on polarized debates. We introduce an approach to bring in negative user-to-user interaction, by analyzing online debates using signed networks with positive and negative ties. We apply this approach to the Dutch Twitter debate on 'Black Pete'—an annual Dutch celebration with racist characteristics. Using a dataset of 430,000 tweets, we apply natural language processing and machine learning to identify: (i) users' stance in the debate; and (ii) whether the interaction between users is positive (supportive) or negative (antagonistic). Comparing the resulting signed network with its unsigned counterpart, the retweet network, we find that traditional unsigned approaches distort debates by conflating conflict with indifference, and that the inclusion of negative ties changes and enriches our understanding of coalitions and division within the debate. Our analysis reveals that some groups are attacking each other, while others rather seem to be located in fragmented Twitter spaces. Our approach identifies new network positions of individuals that correspond to roles in the debate, such as leaders and scapegoats. These findings show that representing the polarity of user interactions as signs of ties in networks substantively changes the conclusions drawn from polarized social media activity, which has important implications for various fields studying online debates using network analysis.

Introduction

In recent years, the advent of social media platforms has given researchers access to a wealth of digital data on social relations, behavior, and beliefs (Conte et al. 2012; Lazer et al. 2009). Data from these social media platforms have fueled the growth of new approaches for social research, most notably Computational Social Science, which uses digital data and computational methods to capture and study social dynamics (Conte et al. 2012; Lazer et al. 2009; Lazer et al. 2020). Drawing from the natural and technical sciences, Computational Social Science provides powerful tools and methods for working with large-scale relational data, opening up new avenues into the study of social phenomena such as mass mobilization (Abdul Reda, Sinanoglu, and Abdalla 2021; González-Bailón, Borge-Holthoefer, and Moreno 2013), polarization (Bail et al. 2018b; Tucker et al. 2018), the spread of misinformation (Lazer et al. 2018), political discourse (Garimella et al. 2018), and much more. In this research, social network analysis is among the most powerful and commonly used tools; by representing social interaction as graphs, the network perspective unearths the relational structures emanating from and shaping interactions, allowing researchers to identify communities and central actors (Lazer et al. 2009).

Network studies into polarization have shown that online users sharing ideological affiliation tend to cluster together in terms of interaction (Adamic and Glance 2005; Conover et al. 2011; Conover et al. 2011a), which suggests an understanding of polarization as the simultaneous clustering of allies and repulsion between antagonists (Uitermark, Traag and Bruggeman 2016; Waugh et al. 2009). However, while animosity and conflict are central to this relational angle on polarization, the vast majority of social network studies into polarized online debates have only considered positive relationships, represented as networks with positive ties (Hassan, Abu-Jbara and Radev 20112b; Leskovec, Huttenlocher and Kleinberg 2010b). This has resulted in confounding results, such as finding cohesive network structures in what are known to be polarized debates (Conover et al. 2011; Conover et al 2011a; Guerrero-Solé 2017; Lietz et al. 2014). This paper argues that the systematic neglect of negative, antagonistic, user-to-user interactions in network studies has severe consequences for our understanding of polarized discourse online. We introduce an approach, powered by natural language processing and machine learning, for distinguishing positive from negative interactions between users. This information on the polarity of user interactions allows for the analysis of online debates using a signed network, a network with positive and negative ties. Our analysis shows that the inclusion of negative ties has a profound impact on the findings with respect to the structure of the debate and the positions of actors and communities within it.

We apply this approach to a case study on the divisive Dutch debate over 'Black Pete' ('Zwarte Piet'), a Dutch mythical character with racist connotations. Although black communities and anti-racist activists have long critiqued Black Pete, in recent years, a full-fledged national debate has unfolded about the character and what it might say about racism in Dutch society more broadly (Coenders and Chauvin 2017; Wekker 2016). This debate provides a useful case to show how signed network analysis enhances our understanding of contentious debates.

We use a dataset of the Twitter debate on Black Pete, covering the period from December 2017 to May 2019, comprising roughly 430,000 tweets from 81,700 unique users, with 296,881 unique mentions between users. From this dataset, 10,000 tweets were manually labeled, coding their issue sentiment (*pro* / *neutral* / *anti* / *ambiguous* in relation to the issue of Black Pete), and the sentiment of each user-to-user interaction (*opposition* / *agreement* / *neutral* / *ambiguous* in relation to the targeted user). Using this labeled dataset, we trained a machine learning algorithm to classify the interaction sentiment between users. From that, we constructed a signed network of users in this debate and compared this to the retweet network for the same data, which is typically employed to study polarization on Twitter. The comparison shows that our approach identifies a larger number of actors, finds different communities, and provides greater insight into the diversity of roles that actors play within the debate. In particular, we show important differences between actors that are attacked from all sides (mainly cabinet members and public institutions) and actors that receive support from one side while coming under attack from the other side (mainly parliamentarians and activists).

In the remainder of this introduction, we outline advances in computational Twitter research on polarization and signed network analysis. In the materials and methods section, we first present our case study and dataset on Black Pete and subsequently describe our methods for extracting the sign of user relations. The result section is composed of four subsections. The first two empirical subsections compare the signed network with the retweet network in terms of 1) users included in the network, and 2) the structure and composition of communities. The latter two empirical subsections detail the relations between communities and the roles of individuals in the debate that are laid bare by the (positive and) negative interactions. These results are powered by quantitative as well as qualitative analyses of the data in order to reach meaningful conclusions about their significance. Finally, we conclude what our approach to signed network analysis contributes to the study of polarization.

Literature: Twitter studies and Signed Networks

To capture the structure of interaction, Twitter research has focused on the social networks that are shaped by user interaction through either "retweets" or "mentions," both of which are generally studied through unsigned network analysis. A retweet is a simple act of sharing in which a user shares another user's tweet with their network. Retweets are generally considered endorsements, that is, as positive ties (Metaxas et al. 2015) or as contributing to information flow (Freelon 2014). When studying debates through retweet networks, researchers have found separate user clusters, with limited interaction between political opponents (Barberá 2015a; Barberá 2015b; Conover et al 2011a;

Guerrero-Solé 2017; Himelboim et al. 2016; Soares, Recuero and Zago 2019). The correlation between retweet structure and political ideology is so strong that retweet networks have been used to predict user ideology (Barberá 2015a; Boutet, Kim and Yoneki 2012; Conover et al. 2011a; Guerrero-Solé 2017).

Mentions, on the other hand, are a syntax for targeting a message or part of a message to a specific user, by adding the @-sign to the Twitter username. For example, a user in our dataset mentioned Nadia Bouras, a historian that publicly speaks out against Black Pete with over 29,000 followers on Twitter, in the following tweet: '@NadiaBouras Cry Baby. Black pete stays anyway!!!'. In previous literature, mentions are generally considered neutral, or as expressions of information exchange. Conover et al. (2011, p.6), for instance, suggest that "mentions form a communication bridge across which information flows between ideologically-opposed users". When studying debates through mention networks, researchers generally do not find strict divides between opposing groups, as mentions occur across polarized clusters and party lines (Conover et al 2011; Conover et al. 2011a; Esteve del Valle and Borge Bravo 2018; Feller et al. 2011; Gruzd and Roy 2014; Guerrero-Solé 2017), concluding that information exchange is occurring across political lines.

However, the reliance on unsigned ties in studying Twitter debates means that both retweet and mention network representations have important limitations. Whereas research based on retweet networks ignores interactions across clusters through mentions, research on mentions inadvertently conflates positive and negative interactions. As the example above illustrates, mentions can be used to attack other users, instead of as a tool to share information with them. This has also been found in more qualitative studies on Twitter. Evolvi (2019a), for instance, studied Islamophobic tweets in the aftermath of Brexit, and found that mentions are often used to "belittle others with different ideas rather than invite conversation" (p. 396). Similarly, Moernaut, Mast and Temerman (2020) studied polarized climate discourse on Twitter and found that interactions tend to be antagonistic, aimed at delegitimizing and denaturalizing out-groups. Gruzd and Roy (2014) manually labeled tweets that occur across party lines and found that roughly half of these are hostile. This would suggest that previous studies have conflated very different forms of interaction, mischaracterizing out-group derogation as a form of neutral information flow. Recent computational research that models social group formation via signed networks further suggests that the exclusion of negative ties significantly distorts community structure (Stadtfeld, Takács, and Vörös 2020). The neglect of negative interaction thus appears to have severely limited the capacity of network analysis to accurately represent online debates, with important implications for the many fields relying on this approach.

To date, there is limited research on Twitter debates using signed networks. This is in part due to the difficulty with identifying the polarity of mentions. This is not a simple variable in the data but has to be abstracted from the meaning of the words and position of the mention in the tweet. One approach to identifying the sign of ties between users that has been applied in previous literature is to simply assume that the interaction of users that hold opposite positions will always be negative, while interactions among users with the same position will always be positive (De Stefano and Santelli 2019; Yardi and Boyd 2010). Another approach taken is to focus on online social networks that allow for explicit negative relations between users, such as Epinion or Slashdot. Studies using such data, however, have predominantly been aimed to develop algorithms to predict signs of edges or future link creation rather than answering social scientific questions about polarization or other social processes (for an overview of signed network mining see (Tang et al. 2016).

When signed networks of online data have been studied in relation to social processes, they have typically been used to test theories on social balance (Heider, 1946) and status theory (Guha et al. 2004; Hassan et al. 2012a; Hassan et al. 2012b; Leskovec, Huttenlocher, and Kleinberg 2010;

Sadilek, Klimek, and Thurner 2018; Zheng, Zeng, and Wang 2015. There are examples of studies using signed network representations of offline data to research polarization. Neal (2018) examined the level of polarization in US congress by representing co-sponsorship of bills as a signed network of interactions between congress members. Uitermark et al. (2016) investigated the Dutch debate on minority integration in newspapers through signed network analysis, demonstrating that opposing groups' community structures differ in terms of cohesion and leadership. Traag and Bruggeman (2009) studied international alliances and disputes, establishing the world is divided into six power blocks. While these studies demonstrate the importance and scientific potential of using signed networks, they also illustrate the challenges involved in extracting signed networks from debate data. Scholars either manually classify relations, use niche social platforms, or make strong assumptions on the sign of ties – all of which preclude the use of signed networks in the study of mainstream social media platforms like Twitter. This paper presents a method for moving beyond this impasse by automatically extracting the polarity from online user interaction in large-scale social media debates by using natural language processing and machine learning.

Materials and Methods

Case: is Black Pete Racist?

The celebration of Sinterklaas (Saint Nicholas) is one of the most important traditions of the Netherlands (Rodenberg and Wagenaar 2016a). Saint Nicholas is similar to Santa Claus: he has a long white beard, a red outfit, and he brings presents for children. Saint Nicholas arrives by steamship from Spain every year in early November and is welcomed publicly in almost every Dutch city. A single town is nominated to be the host of the official national welcoming of Sinterklaas, which means having the occasion broadcasted on national television. On the evening of Saint Nicholas, the 5th of December, the Saint visits families across the country, presenting gifts and sweets to children. In the days and weeks leading up to the 5th of December, many shops are decorated with Sinterklaas-themed promotional material.

The part of this tradition that has become an issue of contention are the helpers who accompany Sinterklaas: the “Black Petes” (“Zwarte Pieten”). These are usually represented by white people wearing blackface. The character has periodically become the focus of debate in Dutch society, due to their—for most observers from outside the Netherlands rather striking—racist undertone (Hilhorst and Hermes 2016; Van Der Pijl and Goulordava 2014; Rodenberg and Wagenaar 2016). The current wave of debate started with the arrest of four activists, most notably Quinsy Gario and Jerry Afriyie, for their participation in protests against Black Pete in Dordrecht in 2011 during the official welcoming (Helsloot 2012; Rodenberg and Wagenaar 2016). Since this protest, there have been intense debates in newspapers, on television, in parliament and up to the UN on whether Black Pete embodies a racist stereotype (Helsloot 2012).

This debate intensified from 2013 onwards, with supporters and opponents of Black Pete mobilized online and in the street every year, leading to violent confrontations. In 2017, pro-Black Pete activists blocked a highway in the north of the Netherlands to prevent anti-Black Pete activists from protesting at the official welcoming of Saint Nicholas. The debate about Black Pete has become the focal point in broader debates about Dutch racism, Dutch colonialism, and the Netherlands’ involvement in the transatlantic slave trade (Helsloot 2012; Van Der Pijl and Goulordava 2014). Opponents believe that Black Pete exemplifies Dutch racism whereas opponents see Black Pete as an innocent character and consider criticisms as an attack on their traditions by overdemanding minority groups and arrogant cultural elites (Hilhorst and Hermes 2016; Van Der Pijl and Goulordava 2014; Rodenberg and Wagenaar 2016).

Data

We used a dataset of tweets on the Black Pete debate posted between December 4th, 2017 and May 7th, 2019. The tweets were collected based on keyword matching of various terms related to the debate, such as "Black Pete", "Zwarte Piet" and "KOZP" (abbreviation for "Kick Out Zwarte Piet"), harvested and stored using the Twitter Capture Analysis Toolset (Borra and Rieder 2014). In total, the dataset contains 467,497 unique tweets from 81,700 unique users, with 296,881 unique mentions between users.

Ethics Statement

The data collection process has been carried out exclusively through the Twitter API, which is publicly available, and for the analysis, we used publicly available data (users with privacy restrictions are not included in the dataset). We abided by the terms, conditions, and privacy policies of Twitter. Since this content is publicly published and is frequently discussed in mass media, we regard the debates as a public domain that does not require individual consent for inclusion in research, based on the ethical guidelines for internet research provided by The Association of Internet Researchers (Franzke et al. 2020) and by the British Sociological Association (Anon 2017). We only report on aggregates, and limit reporting on details of individuals to user accounts that belong to public figures or institutions, or that have more than 4,000 followers. The data published along with this research does not include user-ids nor the classification of the sentiment on the Black Pete discussion since this is part of a special category of personal data, formerly known as sensitive data.

Issue and Mention Sentiment Classification

To classify the relationships between users (positive, neutral, negative), we identified, for each tweet (1) the *issue sentiment* – the position expressed on the issue of Black Pete, and (2) the *mention sentiment* – the position toward each mentioned user in the tweets, i.e., whether the tweeting user mentions the other user to express agreement, opposition, or is neutral, such as sharing information. It should be noted that we did not try to classify the overall sentiment of the tweet, for which various existing sentiment analysis algorithms could be deployed, but we specifically targeted the position of the user in relation to Black Pete and the sentiment of the interaction with the mentioned user.

To infer these sentiments, we first manually classified approximately 10,000 tweets randomly selected from the full dataset. By this selection method, we avoid focusing on the most active or popular users which limits the bias towards the vocal minority to the detriment of the (more) silent majority (Mustafaraj et al. 2011). We coded the issue sentiment, whether the tweet expresses a pro, anti, or neutral/ambiguous stance towards Black Pete, as well as the sentiment of each mention in each tweet. We took into consideration that one tweet might contain several mentions, some of which might be intended positively towards the mentioned users and others might be signaling disagreement. These labeling efforts were conducted by four fluent Dutch speakers who were instructed via a coding book designed for this project. The codebook instructions were conservative: if the issue or mention sentiment was not self-evident, the tweet was labeled as ambiguous (see the appendix for more details). The inter-coder agreement was moderate to substantial, measured by a Krippendorf Alpha of 0.72 for the issue sentiment and 0.49 for the mention sentiment, indicating that the classification is a difficult task.

To classify the rest of the data, we applied the following pipeline. First, we classified the issue sentiment of all tweets by the fastText algorithm (Joulin et al. 2017) trained on the manually labeled issue sentiments (see the appendix for more details). Second, we count the number of *pro* and *anti*-tweets of each user and categorized users' stance as pro- or anti-Black Pete by a simple majority rule. That is, if the user posted more pro than anti tweets, we assigned a pro label to the user, and vice versa. Third, we trained the fastText algorithm to classify mention sentiments using the manually labeled mention data. In addition to the tweet text, we provided the fastText model with information about the issue sentiment of the tweet as well as the stance of the tweeting and mentioned users (classified in the previous steps). We additionally constructed two features that might reveal information about the sentiment of the mention: (1) whether the mention takes the form of "via @username"—which are most often neutral, as they are automatically added by the webserver of the media outlet via which the tweet was posted—and (2) whether the mention is located at the start, body, or end of the tweet since that might correlate with the polarity of the mention.

For both the classification of issue sentiment and mention sentiment, we implemented the fastText algorithm for text classification (Joulin et al. 2017), which is informed by advances in word representation learning (Levy, Goldberg, and Dagan 2015; Mikolov et al. 2013). This algorithm uses the training data to construct numerical word vectors for each word in the corpus that represents their relation to other words, thereby capturing (part of) their meaning. To teach the model the basics of the Dutch language, we provided the model word vectors constructed from a Dutch Wikipedia Corpus (Bojanowski et al 2017). The use of such an external corpus enables the machine learning algorithm to discover similarities in words that are missing or infrequent in its training data, thus increasing its vocabulary and subsequent predictive power.

Since the manually labeled data included many more tweets that support than oppose Black Pete (58% expressed a pro position), we balanced the class sizes before classifying the issue sentiment to avoid biasing the algorithm. The fastText classifier categorizes the issue sentiments with sufficiently high accuracy, resulting in 15% (65.314) anti tweets, 48% (225.856) pro tweets and 38% (176.327) tweets with neutral/ambiguous issue sentiment (see the appendix for more details on the issue sentiment classification). Similarly, the labeled mentions were not balanced, containing more negative than positive user mentions. We down-sampled the majority classes to avoid biasing the algorithm, resulting in 1,382 positively annotated mentions, 1,500 negatively annotated mentions and 1,500 neutral/ambiguous mentions. For the mention classification, we filtered the test data on unique tweet text to ensure that the test data included only tweet texts that the classifier had not seen before. We did not filter the training set on unique tweet texts to ensure the classifier learned that one tweet can include several mentions with different mention sentiments.

Since we aim to identify positive and negative mentions, we optimized the algorithm to minimize the risk of incorrectly classifying a negative tweet as positive, and of classifying a positive tweet as negative. We are less concerned with incorrectly classifying a positive or negative tweet as neutral since this will have less impact on distorting the resulting network. To do this, we trained the classifier to maximize the F1 score for all classes, thus attempting to predict all classes well, in both precision and recall. The fastText algorithm gives an indication of how certain the classification is (the softmax probability), valued between 0 and 1 for each prediction. We used this certainty indication to apply a simple rule: classify all mentions with lower certainty (<0.8) as neutral. This procedure reduces the recall for the positive and negative classes, but more importantly, reduces the errors we care most about: classifying positive mentions as negative, and classifying negative mentions as positive.

The classifier—after applying the certainty rule—categorizes the mention sentiments with high accuracy (see Figure 1). There are only 21 cases in which a negative mention is misclassified as positive (0.031 times of all negative mentions and of 0.15 times all positive mention classifications) and 22 cases in which a positive mention is misclassified as negative (0.07 times of all positive mentions and 0.064 times of all negative classifications). To classify user-to-user interaction signs, we considered both the mentions and retweets, where retweets are taken as acts of endorsements, a positive interaction from the retweeting to the retweeted user (Metaxas et al. 2015). Next, we used a majority rule: if most of the user-to-user interactions were positive (negative), we classified the directed sign between these users as positive (negative). This procedure classified approximately 54% of user interactions as neutral (or too ambiguous to categorize), 37% as positive and 8% as negative.

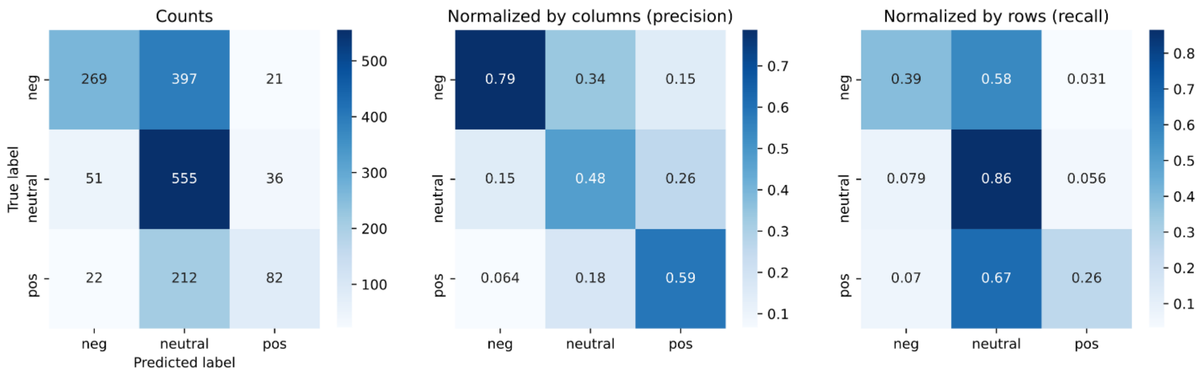


Figure 1: The results of the classifier (parameter values: epoch=25, learning rate=0.7, n-grams=3) after applying the simple certainty rule (neutral if certainty < 0.8): confusion matrix with counts (left), normalized by the true labels (middle) and normalized by the predicted labels (right). The values in the diagonals of the middle matrix are the precision rates, and the values on the diagonals of the right matrix are the recall rates. Recall rates here are reduced due to the certainty rule, but the most important errors (classify positive if the true value is negative and classify negative if the true value is positive) are reduced.

Since we conducted a signed network analysis, we focus on relations that we could with some certainty identify as positive or negative with the procedures described above, while leaving out neutral and ambiguous relations. Most (86%) of the neutral/ambiguous relations were based on only one interaction between the users and were therefore more difficult to classify accurately.

Results

The classification of the sentiment of user interactions (positive, negative) allows us to construct a signed network of this debate and compare that network to the retweet network that is commonly used for studying Twitter debates. In our comparison of the signed network with the retweet network, we focus on (1) differences in the set of users included in both networks; (2) the overall community structure and composition in the networks; (3) the positions of these communities in the network and the debate; and (4) the role of individual actors in the network and debate.

The signed network consists of 94,016 nodes (representing users), 150,555 positive and 33,329 negative relations. The retweet network, which is based on direct retweets ("quote tweets" are treated as mentions), consists 55,758 nodes with 211,669 relations. We consider the edges in the retweet network as positive, in line with previous research with which we aim to compare our signed network results (eg. Barberá et al. 2015a; Conover et al. 2001; Guerrero-Solé 2017; Himelboim et al. 2016; Metaxas et al. 2015).

Missing Users

The first notable difference between the retweet network and the signed network is that they include different actors. In total, there are 38,258 (40%) users in the signed network that are not in the retweet network. These are users that are not being retweeted (because they are not tweeting on the topic in our data) but are receiving mentions on Twitter in the context of the debate on Black Piet. However, many of these users are isolates or not part of the largest connected component of the signed network. We, therefore, focus our subsequent analysis on the largest connected component of each network, as is typically done in network analysis. There are 3,112 users (6%) in the signed network that are not present in the retweet network. In comparison, 559 users (1%) in the retweet network are not part of the signed network. The users that an analysis of the retweet network misses out on tend to be more important in the debate; the users missing in the signed network have very low (all less than 50) indegree, whereas many of the users the retweet network misses out on are prominent in the debate (see Figure 2).

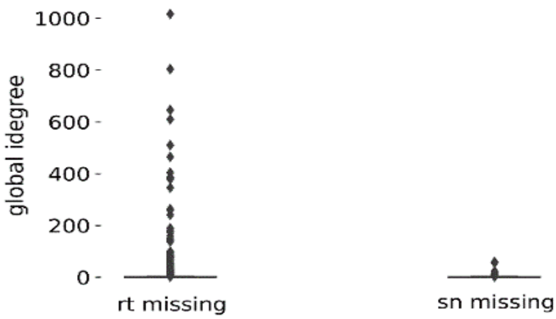


Figure 2: Distribution of the indegree of the users that an analysis of the largest connected component in the retweet network (left) or signed network (right) would miss.

Taking a closer look at users that have a high indegree in the signed network but are missing in the retweet network (see Figure 3), we find that these users are key actors in the debate on Black Pete. For example, the prime minister of the Netherlands Mark Rutte ('minpres' on Twitter) is absent in the retweet network as he did not tweet about the topic. However, his words and actions in this debate are influential and many people mention him on Twitter, giving him a central position in the signed network. Similarly, the public prosecutor (referred to by users as @om) is often mentioned negatively but is absent in the retweet network as this account did not tweet on the topic.

	followers	community	global indegree	pos global indegree	neg global indegree
user_name					
jesseklaver	264,000	4	1017	11	1006
minpres	1,000,000	4	804	4	800
omroepntr	43,700	2	646	8	638
sylvanasimons	12,700	2	610	4	606
om	1,300,000	4	510	0	510
albertheijn	45,600	4	465	0	465
kruidvat	505	2	404	1	403
publiekeomroep	45,200	4	386	0	386
politie	271,300	4	383	34	349
kruidvatservice	7,750	2	379	1	378

Figure 3: The top 10 users of the signed network that are not in the retweet network. These top ten users are not in the retweet network at all—also not in the smaller or isolated communities. The column statistics are based on the signed network. The follower counts are by July 2020.

Community Structure

We next compare the structure and composition of communities in the signed network with those in the retweet network. We detect the community structures in both networks with the Leiden Algorithm (Traag, Waltman, and van Eck 2019a), which maximizes the positive links within communities and minimizes positive links between communities compared to a random network with the same degree distribution. In the case of the signed network, the negative edges are also taken into account but with the reverse logic: minimizing the number of negative edges within communities and maximizing the number of negative edges between communities (Traag and Bruggeman 2009a).

Both networks show similar degrees of modularity: the retweet network has a modularity of 0.45 while the signed network has a modularity of 0.45 in the layer of positive edges and of 0.24 in the layer of negative edges. In both networks, the two largest groups are of a similar size and contain roughly 40% of the nodes, and the largest ten communities together make up roughly 90% of the nodes (see Figure 4).

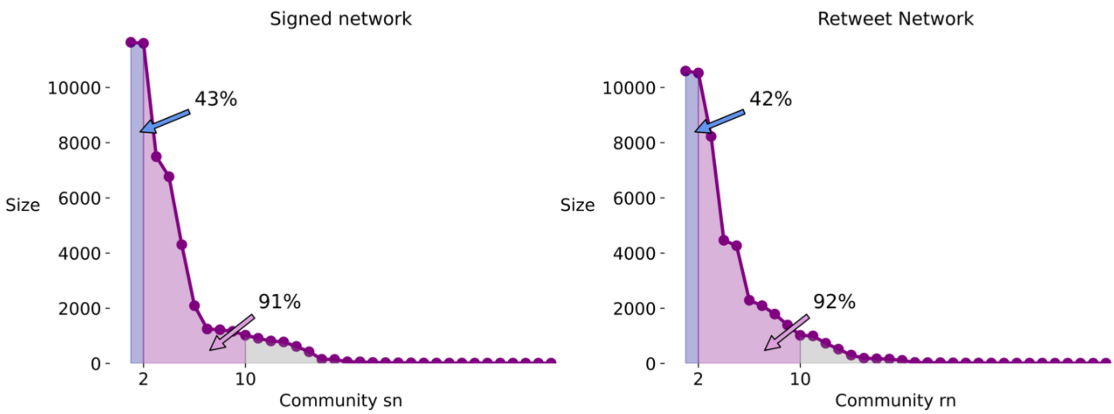


Figure 4: The distribution of sizes of the communities in the signed network (left) and retweet network (right). The tail is cut-off (displaying only communities with more than 10 members) for the sake of legibility. This figure shows that the two networks feature a similar community structure.

Whereas the community structures in both networks are similar in modularity and size, the communities differ significantly in terms of their compositions. Comparing the users and communities of the retweet network directly with those of the signed network reveals that the mentions and their polarity have had a marked influence on the group compositions (see Figure 5). For example, the users classified in community 1 in the signed network, the dominant pro-Pete community, overlap for a substantial part (66%, n=7,784) with the users in the same community in the retweet network. However, they are merged with many other users (n=3,854) from other communities in the retweet network, such as community 5 and 6. At the same time, several users belonging to community 1 in the retweet network are split off into other separate communities in the signed network, community 4 and particularly community 7. This community 7 is a community centered around Geert Wilders, a radical right-wing politician with a strong anti-immigration and anti-Islam agenda.

Zooming in on the twenty most prominent actors (those who receive the highest number of retweets, positive or negative mentions), we find that some of these top users are grouped in different combinations in the retweet network compared to the signed network, illustrated in Figure 6. In sum, taking into account mentions with their signs affects the composition of communities for both rank-and-file as well as prominent actors.

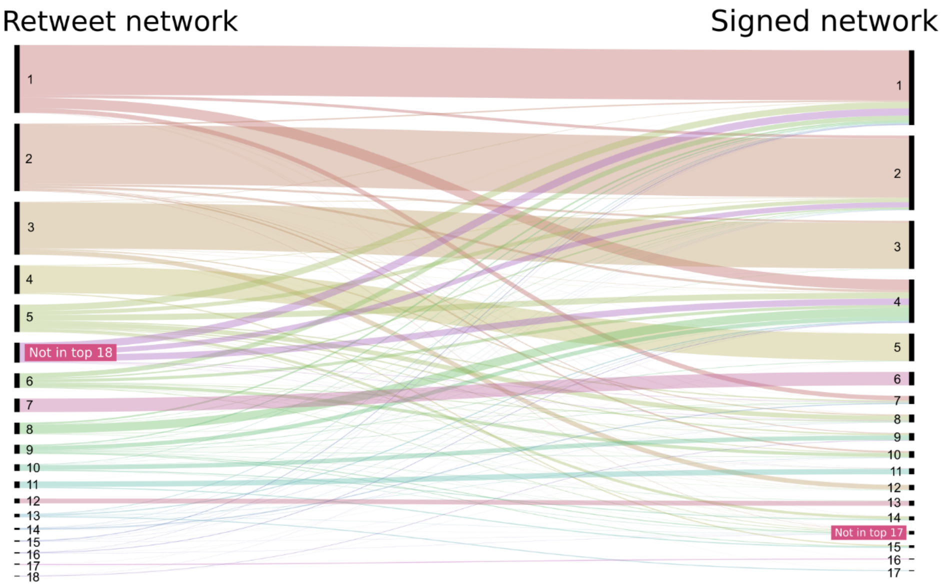


Figure 5: Alluvial graph illustrating the relationships between the group structure in the retweet (left) and signed network (right). The thickness of the lines corresponds to the number of users, and non-horizontal lines indicate differences between the group structures in the two networks. The figure shows considerable differences in the group compositions and illustrates that there are many central users in the signed network (n=3.112) that are not in the top 18 communities of the retweet network.

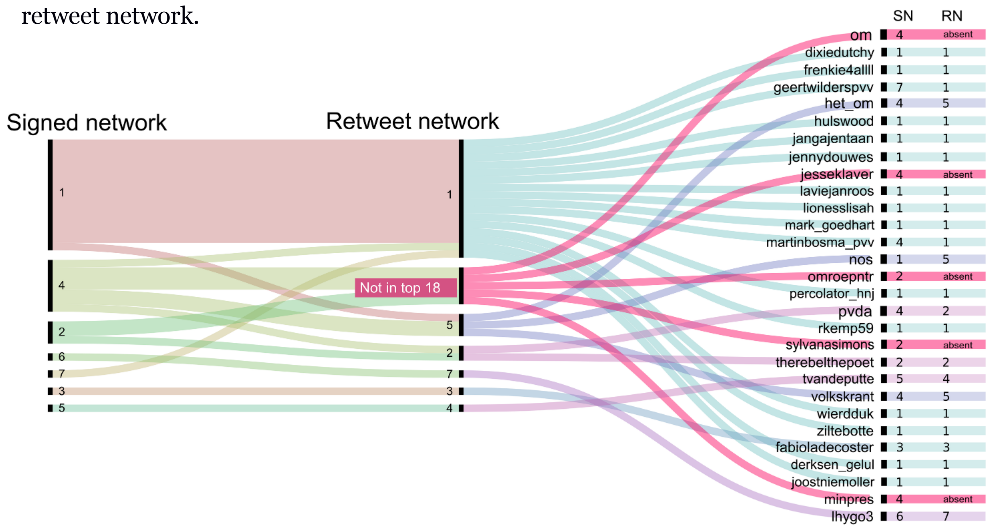


Figure 6: Alluvial graph illustrating the communities of the top 20 actors in the debate in the signed (left) versus retweet network (right). The figure shows considerable differences in the group structures and illustrates that some of these top actors in the debate, such as prime minister Mark Rutte (minpres), are not present in the top 18 communities of the retweet network.

Coalitions and Divisions in the Debate

After exploring the differences in community structure and composition between the two networks, we turn to what the negative interactions contribute to the understanding of these communities' positions in the debate and their relations with one another. Whereas the communities in the retweet network are formed on the basis of separation, the signed network detects groups on the basis of separation and confrontation which leads to richer information on the community relations and their positions in the debate. Figure 7 displays the relationships between the signed network communities in terms of relative positive ties (left) and negative ties (right) and illustrates each communities' dominant stance towards Black Pete (pro/anti/neutral) by a color scale. This shows that some communities send many negative messages to others in the debate, even to communities with a similar aggregate pro/anti-stance on this polarizing topic.

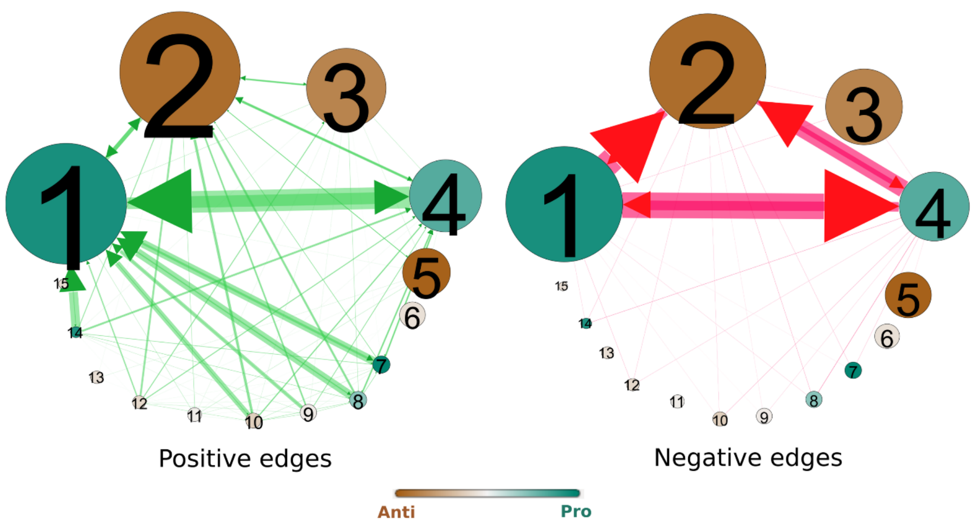


Figure 7: The aggregate network of the communities, with positive relations (left) and negative relations (right). Nodes are sized by the number of users in this community and colored by the average issue sentiment of users in the community. Edges are sized by the absolute count of outgoing edges from source to target community, divided by the source community's size.

The next section provides a more detailed interpretation of these findings. This is based on the community statistics reported in Figure 8, relations between communities in the network visualized in Figure 7, and the users' stance on Black Pete (see Figure 9). We then carry out a qualitative analysis of the main communities in the signed network (those with over 1,000 members), analyzing their twenty most retweeted tweets, all tweets of the community's users with the highest positive indegree, and a random sample (n=100) of other tweets of the community. For this selection of tweets, we examine the themes addressed and the position expressed toward Black Pete.

The signed network is dominated by two large, antagonistic poles: one constituted by pro-Black Pete communities 1 and 4, and another by the anti-Black Pete community 2 (see Figure 9). Community 1 of the pro-Black Pete pole is most vocal (users on average positively referencing almost 9 others in their community) and most confrontational (attacking on average 1.7 users of other groups) (see Figure 4). There are other outspoken pro- and anti-Black Pete communities, such as community 7 and 3 respectively, but there is significantly less antagonistic communication from and to these communities (see Figure 7). Furthermore, some communities (9 and 10) feature pro as well as anti-Black Pete users, together averaging to a neutral position on Black Pete. Popular tweets express exasperation with the debate.

Figure 8 (next page): Community statistics and their top users in the signed network (with size>100). The average issue sentiment is calculated over the sentiment (pro-Black Pete=1, anti-Black Pete=-1) of all users, and reported separately for the top 10 users most often positively related to from within the community. Column 'pos int e fraction' divides the positive edges within the community by the total positive edges outgoing from community members and 'neg out e fr.' divides negative edges within the community by the total negative edges outgoing from community members. The columns 'avg pos int e' and 'density (pos)' divide the total positive edges by the number of users and the possible edges in the community, respectively. The column 'top global negative' lists the users in this community that are most frequently negatively mentioned by other communities.

	size	avg issue sentiment	avg issue sent top users	pos int e	avg pos int e	density (pos)	pos int e fr.	neg out e	neg out e fr.	avg neg ext e	indegree pos	indegree neg	top local positive	top global negative
community														
1	11,638	0.77	1.00	104,011	8.94	0.04	0.87	19,337	0.94	1.66	15,418	5,936	joostniemoller, wierdduk, percolator_hnj, jennydouwes, hulswood	nos, wierdduk, teletekst, eenvandaag, tponl
2	11,603	-0.57	-1.00	35,044	3.02	0.01	0.84	5,738	0.94	0.49	6,158	13,252	therebelthepoet, vicenl, thenuc1, nadiabouras, claricegargard	therebelthepoet, omroepntr, sylvanasimons, erikvmuiswinkel, kruidvat
3	7,496	-0.46	-0.70	11,435	1.53	0.01	0.91	38	1.00	0.01	1,015	70	fabioladecoster, anomalisa__, aranxerini, rani_remy, queentrey_	ninodevries, 2020gwoww, mickjohan, queentrey_, ronnieflex2907
4	6,768	0.58	0.70	8,663	1.28	0.01	0.42	6,103	0.94	0.90	10,877	11,789	martinbosma_pvv, franckentheo, asterberghe, bartdemeulenaer, robinvanhijfte	het_om, jesseklaver, minpres, volkskrant, pvda
5	4,306	-0.63	-0.80	4,660	1.08	0.01	0.95	4	1.00	0.00	622	24	tvandeputte, karenattiah, marcusjdl, tinalasisi, soualiganamazon	karenattiah, tvandeputte, kelechnekoff, redlightvoices, solitudekth
6	2,094	0.00	0.00	2,097	1.00	0.02	1.00	0	0.00	0.00	1	0	lhygo3, lookslikechloe, seongbukrotari, cjsanseo1110, ena_hand_real	lhygo3, wdbb0u39, ena_hand_real, poroli_love, ssc_for_running
7	1,248	0.84	0.50	1,298	1.04	0.04	0.60	10	1.00	0.01	2,184	32	geertwilderspvv, voetbalflitsen, denieuwemaanntr, voetbalinside, defendevropa	geertwilderspvv, denieuwemaanntr, supercemal, voetbalflitsen, saabje96
8	1,222	0.42	0.70	1,394	1.14	0.05	0.57	17	1.00	0.01	1,500	93	danevv, rtlnieuws, adnl, pro_respect, johan_anema	rtlnieuws, adnl, hartvnl, nhnieuws, rtvnoord
9	1,165	0.06	0.20	1,198	1.03	0.04	0.70	4	1.00	0.00	538	28	haralddoornbos, ozcanakyol, evoosterhout	ozcanakyol, haralddoornbos, drbrambakker
10	1,025	-0.08	0.00	1,086	1.06	0.05	0.61	11	1.00	0.01	1,003	16	peter26061980, rloppenheimer, sylviawitteman	sylviawitteman, vice, mariannezw
11	913	0.07	0.10	926	1.01	0.06	0.93	0	0.00	0.00	158	9	ajenglish, redfishstream, rt_com	ajenglish, breakingnlive, rt_com
12	815	-0.03	0.40	1,072	1.32	0.08	0.78	10	1.00	0.01	234	15	lectrr, youssef_kobo, philippedecoene	pensioenspook, vanranstmarc, elhammouchiothm
13	780	-0.00	0.00	780	1.00	0.06	0.99	0	0.00	0.00	3	0	onifinau, madameghana, ajplusfrancais	madameghana, zzaynabtata, adaugo__
14	617	0.75	0.70	647	1.05	0.09	0.55	10	1.00	0.02	656	11	jankoudum, bigbolder, maxdekok1913	rob_rtm, jankoudum, omropfryslan

community	size	avg issue sentiment	avg issue sent top users	pos int e	avg pos int e	density (pos)	pos int e fr.	neg out e	neg out e fr.	avg neg ext e	indegree pos	indegree neg	top local positive	top global negative
1	11,638	0.77	1.00	104,011	8.94	0.04	0.87	19,337	0.94	1.66	15,418	5,936	joostniemoller, wierdduk, percolator_hnj, jennydouwes, hulswood	nos, wierdduk, teletekst, eenvandaag, tponl
2	11,603	-0.57	-1.00	35,044	3.02	0.01	0.84	5,738	0.94	0.49	6,158	13,252	therebelthepoet, vicenl, thenuc1, nadiabouras, claricegargard	therebelthepoet, omroepntr, sylvanasimons, erikvmuiswinkel, kruidvat
3	7,496	-0.46	-0.70	11,435	1.53	0.01	0.91	38	1.00	0.01	1,015	70	fabioladecoster, anomalisa__, aranxerini, rani_remy, queentrey_	ninodevries, 2020gwoww, mickjohan, queentrey_, ronnieflex2907
4	6,768	0.58	0.70	8,663	1.28	0.01	0.42	6,103	0.94	0.90	10,877	11,789	martinbosma_pvv, franckentheo, asterberghe, bartdemeulenaer, robinvanhijfte	het_om, jesseklaver, minpres, volkskrant, pvda
5	4,306	-0.63	-0.80	4,660	1.08	0.01	0.95	4	1.00	0.00	622	24	tvandeputte, karenattiah, marcusjdl, tinalasisi, soualiganamazon	karenattiah, tvandeputte, kelechnekoff, redlightvoices, solitudekth
6	2,094	0.00	0.00	2,097	1.00	0.02	1.00	0	0.00	0.00	1	0	lhygo3, lookslikechloe, seongbukrotari, cjsanseo1110, ena_hand_real	lhygo3, wdbb0u39, ena_hand_real, poroli_love, ssc_for_running
7	1,248	0.84	0.50	1,298	1.04	0.04	0.60	10	1.00	0.01	2,184	32	geertwilderspvv, voetbalflitsen, denieuwemaanntr, voetbalinside, defendevropa	geertwilderspvv, denieuwemaanntr, supercemal, voetbalflitsen, saabje96
8	1,222	0.42	0.70	1,394	1.14	0.05	0.57	17	1.00	0.01	1,500	93	danevv, rtlnieuws, adnl, pro_respect, johan_anema	rtlnieuws, adnl, hartvnl, nhnieuws, rtvnoord
9	1,165	0.06	0.20	1,198	1.03	0.04	0.70	4	1.00	0.00	538	28	haralddoornbos, ozcanakyol, evoosterhout	ozcanakyol, haralddoornbos, drbrambakker
10	1,025	-0.08	0.00	1,086	1.06	0.05	0.61	11	1.00	0.01	1,003	16	peter26061980, rloppenheimer, sylviawitteman	sylviawitteman, vice, mariannezw
11	913	0.07	0.10	926	1.01	0.06	0.93	0	0.00	0.00	158	9	ajenglish, redfishstream, rt_com	ajenglish, breakingnlive, rt_com
12	815	-0.03	0.40	1,072	1.32	0.08	0.78	10	1.00	0.01	234	15	lectrr, youssef_kobo, philippedecoene	pensioenspook, vanranstmarc, elhammouchiothm
13	780	-0.00	0.00	780	1.00	0.06	0.99	0	0.00	0.00	3	0	onifinau, madameghana, ajplusfrancais	madameghana, zzaynabtata, adaugo__
14	617	0.75	0.70	647	1.05	0.09	0.55	10	1.00	0.02	656	11	jankoudum, bigbolder, maxdekok1913	rob_rtm, jankoudum, omropfryslan

Figure 9: The distribution of users' issue sentiment in favor (1) or against (-1) Black Pete per community. The top panel gives the final user issue sentiment by the majority rule (-1,0 or 1 per user). The bottom panel gives the users' issue score, calculated by the sum of users' tweet sentiments. This shows communities 1, 2 and 4 contain a number of highly active users with a strong sentiment on Black Pete.

Community 1, the vocal and confrontational pro-Black Pete community, is one of the largest communities in the network with roughly 11,000 users. Users in this community show strong internal cohesion (users positively reference almost 9 others in the community on average) and heavily attack users from other communities (negatively referencing 1.7 users of other communities on average). Users of this community are vehemently pro-Black Pete and mainly attack users in the anti-Black Pete community 2. This community's stars are Joost Niemoller and Wierd Duk, both journalists and well-known pro-Black Pete supporters. Other prominent figures are the anonymous Twitter account @perculator_hjn (which produces a stream of tweets expressing radical right opinions) and Jenny Douwes, the initiator of a road barricade to block anti-Black Pete protesters in 2017. The main targets of attack are the anti-Black Pete activist Jerry Afriyie (@therebelthepoet, community 2), who often gets scolded for his activism and is told that he should "go back to Ghana," and the public prosecutor (@het_om, community 4) that is accused of being biased against supporters of Black Pete.

Community 2, the activist anti-Black Pete community, also includes roughly 11,000 users and is the main antagonistic pole of community 1. Users in this community tend to hold anti-Black Pete positions and include many of the core anti-Black Pete activists, as well as politicians, newspapers, and national celebrities that have spoken out in favor of changing the appearance of Black Pete. This community is also internally cohesive and externally negative, though less pronounced than the pro-Pete community 1. The majority (62%) of negative references of community 2 are directed towards users of community 1, followed by users of community 4, the second pillar of the pro-Black Pete pole (30%). Community 2's central figures are Jerry Afriyie (@therebelthepoet); ViceNL, a media outlet; the New Urban Collective (@NUC1), an activist social enterprise for inclusivity; and Nadia Bouras (@nadiabouras), a historian working for Leiden University. The main targets of attack are NOS (@nos, community 4), the Dutch Broadcasting Foundation (comparable to the BBC); Wierd Duk (@wierdduk, community 1); and the Dutch Prime Minister (@minpres, community 4).

Community 3 has 7,496 users and is an anti-Black Pete community that stands out for its relatively young and international members. This community is mostly formed around positive internal relations instead of negative external relations. Tweets by these users often feature slang, and particularly English slang ("y'all", "wanna", "trash", etc.) and are often more jovial, for instance discussing the Black Pete issue in relation to dating. Most (75%) of the positive references to other communities are directed towards the activist anti-Black Pete community 2.

Community 4 is overwhelmingly pro-Black Pete and consists of 6,768 users. These users are identified as a community predominantly because of criticisms they direct at others (0.9 per user on average) and that others direct at them (1.75 per user on average). Yet there are also positive connections within the community (1.28 per user on average). This community is mostly in opposition to the activist anti-Black Pete community 2. The relationship with the vocal and confrontational pro-Pete community 1 is more ambiguous: users of community 4 reference community 1 positively as well as negatively, and similarly, users of community 1 reference them positively as well as negatively. This community includes many institutions and institutional actors, such as the public prosecutor (@het_om), the police (@politie), some political parties, and municipalities. These accounts tend to predominantly be subject of negative links from other users. Internal positive edges are centered around one politician, Martin Bosma (@martinbosma_pvv), who is part of the radical right-wing party of Geert Wilders, the PVV. Bosma is very active in the debate, retweeting over 40 distinct users with pro-Black Pete tweets.

Community 5, with 4,306 users, is another anti-Black Pete community that is internationally oriented (most of the tweets are in English) but is older and more academic than community 3. The community is structured around positive internal edges (1.08 per user on average), more so than outgoing or receiving negative critique. The community is organized around Tom van de Putte (@tvandeputte), a Dutch academic who is head of the critical studies department at the Sandberg Institute in Amsterdam the Netherlands. He is retweeted 2,066 times in this community (out of his 5,199 retweets).

Community 6, with 2,094 users, is an exceptionally isolated community whose users neither retweet nor mention users from other communities, and receive in total only one reference from a user in another community. This community predominantly tweets in Korean (98%) and uses no mentions. The community is centered around tweets from one user account (@lhygo3) that has been suspended, but who tweeted both in Korean and English about their surprise about the existence of Black Pete in the Netherlands.

Community 7, the Wilders community, has 1,248 users and is outspoken pro-Black Pete. Users are centered around tweets of Geert Wilders, positively referenced by 63% of them, who uses the Black Pete issue to explicitly call upon users to vote for his political party (PVV). The community is formed around positive internal edges (1.04 per user on average), and less so by negative incoming or outgoing critiques. This community has a remarkably high number of positive incoming edges from other pro-Black Pete users, predominantly from the vocal and confrontational pro-Black Pete community 1 (76%), and to a lesser extent from community 4 (13%). Reciprocally, users in this Wilders community 7 also positively reference users from these two pro-Black Pete communities.

Community 8, the media community, with 1,222 users, receives many positive and negative references from other communities relative to its size (1.30 per user on average), particularly from the activist anti-Black Pete community 2. Many of the top accounts in this community are from news outlets, such as television shows, radio broadcasters and newspapers. The majority of tweets by users in this community express a pro-Black Pete position.

Community 9, one of the neutral communities, consists of 1,165 users with solely positive relations between them. Unlike many other communities, this community is not centered around one highly retweeted or mentioned user. The tweets in this community are not expressively pro-Black Pete, but frame the discussion as irrelevant, making jokes and comparisons with other issues they deem irrelevant. The mayor of the city Emmen, for example, tweets that he would rather deal with creating job opportunities than with issues such as Black Pete, fireworks during New Year's, or some other controversial symbolic political issues in the Netherlands.

Community 10 is another neutral community and has 1,025 users. This community is, similarly to community 9, not expressively anti-Black Pete but its members emphasize they are exasperated by what they see as an overblown discussion. About a third of this community's positive outgoing edges are directed to other communities, most frequently to the vocal and confrontational pro-Black Pete community 1 as well as the activist anti-Black Pete community 2. Users of this neutral community are also positively referenced by both of these communities.

This examination has shown that the inclusion of negative ties not only changes the composition of communities but also reveals a more complex structure of internal fractions and coalitions within and between the supporters and critics of Black Pete. Compared to retweet networks, signed networks enables distinguishing conflict from indifference, which creates a richer understanding of the community structure. For instance, while communities 1, 4 and 7 are all predominantly pro-Black Pete, there is a high level of negative interaction between communities 1 and 4. Between the anti-Black Pete communities, we see few negative interactions. It seems as if users of these communities are subsiding in segregated Twitter spaces, not regularly mentioning, or retweeting each other. This important difference would have been impossible to discover using traditional unsigned network analysis.

Signed Structural Positions and Debate Roles

Unsigned networks have two widely recognized structural positions: *hubs* (users with many ties) (Barabási and Albert 1999; Travers and Milgram 1977) and *bridges* (users that connect otherwise separate communities) (Granovetter 1973). These are widely used in the study of social networks. In the study of political debates, these positions are taken to correspond to roles taken by actors: hubs are central actors or opinion leaders (Lazarsfeld, Berelson, and Gaudet 1968), and bridges play the role of mediators between communities. However, since unsigned networks do not consider the nature of the interaction, these structural positions can be argued to map poorly to roles in polarized debates. Using unsigned networks implies either considering only positive interaction, thereby missing actors that have important roles as subjects of critique, or conflating positive and negative interaction, thereby confusing venerated authorities with hated trolls.

Using signed networks, we can identify a larger number of structural roles, since a given node can be important in terms of negative ties or positive ties, for members of one side of the debate, the other side, or both sides. The argument made here is that these structural network positions more directly map to roles in the debate, by allowing to distinguish popularity from infamy. This provides a central network tool for the examination of polarized debates. We here identify structural positions and their corresponding roles, using the Black Pete debate network for illustrative purposes.

The spectrum of structural positions in a debate with two opposing sides can be represented as a matrix with two axes: on the x-axis, there is the sentiment of one side (operationalized as the number of positive ties minus the number of negative ties) and on the y-axis, there is the sentiment from the other side of the debate. This two-dimensional landscape produces a typology of structural positions based on different regions in the matrix (see Figure 10). We identify five different positions and their corresponding roles:

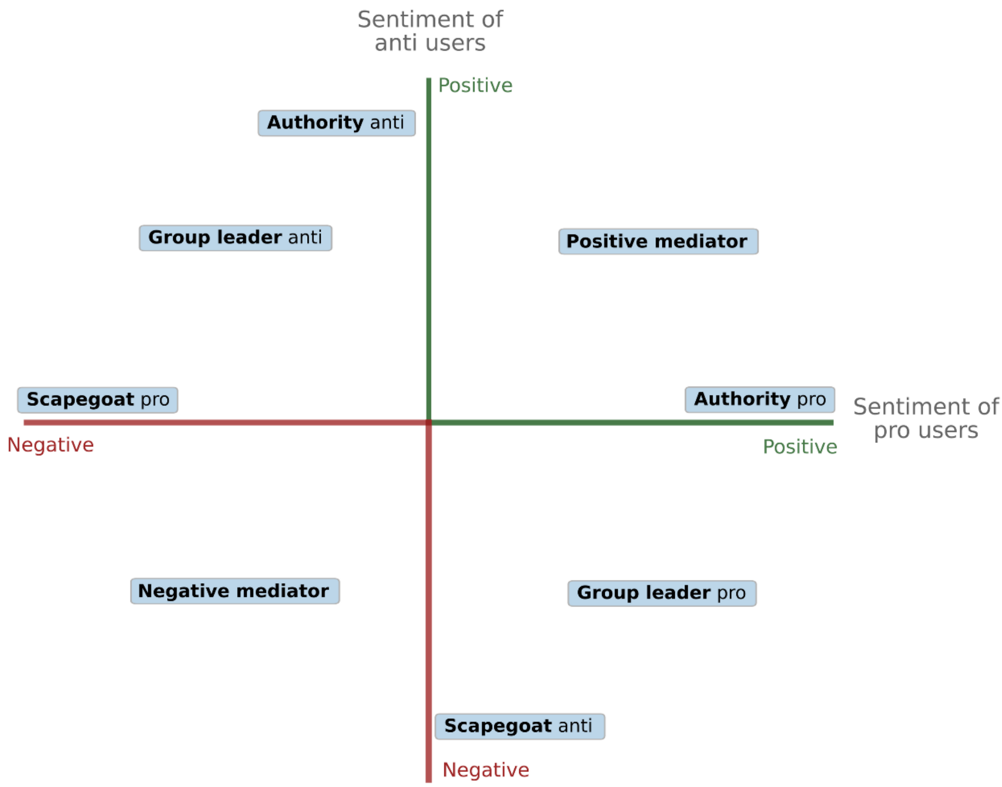


Figure 10: Theoretical landscape of positions in the debate, as defined by the way users of both sides (pro and anti) reference the users. The x-axis and y-axis represent the average sentiment of pro- and anti-users, respectively.

Group leaders receive many positive references from users on their side of the debate, implying that they are recognized as representatives of their cause, and many negative ties from the other side of the debate, implying that the opposing group also views them as important representatives.

Group authorities also receive a lot of positive references by users on their side of the debate but are not attacked by users on the other side, perhaps because they are seen as poor representatives or targets for attacks.

Scapegoats are strongly negatively referenced by the opposing group but ignored or neutrally referenced by users of the side to which they belong. Scapegoats tend to be users that are seen as useful targets of attacks for the opposing group, representing aspects of their outgroup that activates their group solidarity but are not considered as leaders by their ingroup.

Positive mediators are referenced positively by both sides of this debate. Due to this position, positive mediators may function to reduce tensions between the groups.

Negative mediators are referenced negatively by both sides of the debate, though not necessarily for the same reasons. By being the object of dislike from both groups, they potentially bring the groups together by constituting a form of common ground (Heider 1946b).

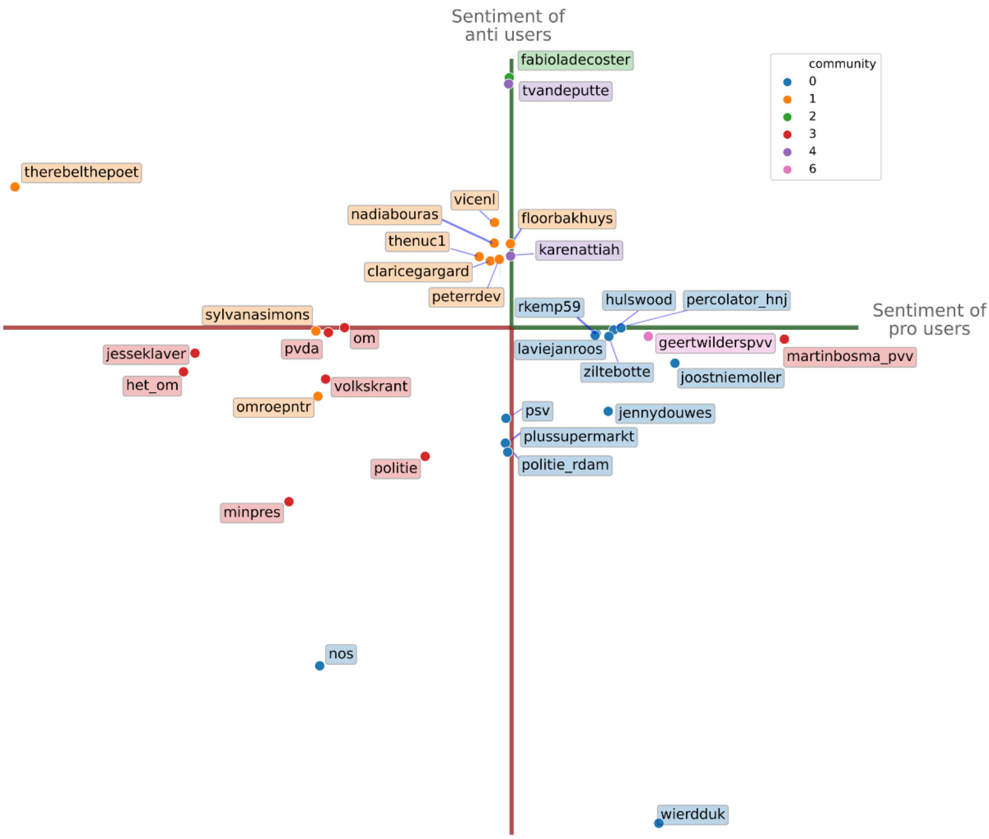


Figure 11: Positions of top users in the Twitter Black Pete debate, as defined by the way users of both sides (pro and anti) reference the users. The x-axis and y-axis represent the average sentiment of pro and anti-users respectively, in which the positive (negative) number of edges to each user is normalized by the total positive (negative) edges of pro (for x-axis) and anti (for y-axis) users in total. Users are colored by their respective communities in the signed network.

Figure 11 shows the structural position landscape of the Black Pete debate, with the users that are most often referenced annotated and colored by their respective communities. The figure identifies the main *group leaders* of both sides: Jerry Afriyie (@therebelthepoet) on the anti-side, and Wierd Duk (@wierdduk) on the pro-side. Both figures are considered by their opponents as radicals who fire up their base to attack the other side.

In the lower-left quadrant in the figure, we see Twitter profiles of *negative mediators* that are heavily attacked by both sides of the debate, most notably NOS (@nos), the Dutch Broadcasting Foundation, which is criticized by both pro and anti-users with claims of biased reporting. Other negative mediators are also predominantly institutions, politicians, and media, such the police (@politie), the prime minister Mark Rutte, the Dutch national television station that broadcasts the children's show on the celebration (@omroepntr), the public prosecutor (@het_om), and one of the main national newspapers (@volkskrant).

On the left side of the horizontal axis, we find two actors with *scapegoat* positions, receiving many negative mentions from pro-users, but few positive mentions from anti-users: Sylvana Simons (@sylvanasimons) and the social-democratic political party PvdA (@pvda). Sylvana Simons is a politician and founder of the anti-racist party Bij1, who has previously been subject to racist threats and hateful attacks. At the lower end of the vertical axis, we find the corresponding scapegoats for anti-users: the football club PSV (@psv) whose supporters allegedly intimidated anti-Black Pete protesters; the supermarket Plus (@plussupermarkt), which—unlike other supermarkets—did not ban the Black Pete characters from products and came under scrutiny for having White employees dressed as Black Pete in its stores; the city of Rotterdam police department (@politie_rdam) which is accused of violating the right of protest of opponents of Black Pete. The scapegoat users thus tend to be institutional actors, often without an official position in the debate, that are targeted as symbolic for the bias of mainstream actors, but lack important discursive roles for the side to which they are taken to belong.

Group authorities are the mirror image of *scapegoats*: they are positively referenced from their side of the debate but receive no or only neutral references from the opposing side. The most important authorities on the pro-side are @percolator_hnj, @hulswood and @rkemp59. These are not public figures such as politicians, journalists, or institutions, but instead are activists on Twitter who have nonetheless built a large following (19,000, 11,000 and 10,000 Twitter followers, respectively). On the anti-side, the most influential accounts are @fabioladecoster (4,000 followers) and @tvandeputte. De Coster does not tweet in a formal capacity; Tom van de Putte is Head of Critical Studies at the Sandberg Institute.

These examples show how the structural positions in signed networks correspond to differentiated social roles in the debate that would not be possible to identify using unsigned network analysis. For example, a highly attacked scapegoat of one side would be missed by a retweet network, or perhaps worse, would be taken as popular figures for the other side by a mention network. The variety of roles in this signed analysis is much broader than can be grasped when negative ties are not taken into account.

Conclusion and Discussion

Twitter has become a central data source for the rapidly growing research on social phenomena using digital data (Bruns and Stieglitz 2014; Tufekci 2014). Data on debates on Twitter have been used to deepen our understanding of a range of phenomena, including mass mobilization (Abdul Reda et al. 2021; González-Bailón et al. 2013), polarization (Bail et al. 2018b; Tucker et al. 2018), the spread of misinformation (Lazer et al. 2018), political discourse (Garimella et al. 2018), and much more. One of the most central methodological pillars underlying this research is the use of social networks to represent interactions between individuals in debates (Lazer et al. 2009).

However, while it is self-evident that in the study of polarized debates it is necessary to distinguish conflict from indifference, leadership from pariahdom, information sharing from insults – doing so has nonetheless been impossible using classical social network methods because these include only one type of interaction: positive interaction. This paper has presented an approach for addressing this limitation by extracting the polarity (positive, negative) of user interaction online and subsequently analyzing the debate using a signed network representation (with positive and negative ties). We applied this approach to Twitter data on the polarized Dutch debate around 'Black Pete,' an annual tradition that has become a lightning rod for the country's culture wars. By processing the tweets on this issue using natural language processing and machine learning, we detect the polarity of user mentions, which we use to extract a signed network of user interaction.

By comparing the resulting signed network with the commonly used unsigned retweet network, the paper showed that signed networks allow for a substantially richer understanding of online debates. First, the signed network captures important and influential users that are missing in the retweet network. Second, the user composition of the identified communities in the signed network differs significantly from the unsigned retweet network. Third, signed networks allow for the identification of not only separate but also conflicting fractions. Our analysis showed that some groups are attacking each other, while others seem to be located in fragmented Twitter spaces – an important distinction that would be impossible to make using unsigned analysis. Fourth, signed networks allow us to distinguish a greater variety of structural positions, which better correspond to roles taken by actors in the debate. Rather than only *hub* and *bridge*, we identified five roles in the debate: *leaders*, *authorities*, *scapegoats*, *positive mediators,* and *negative mediators*.

This shows an important flaw in the existing approaches to studying debates on Twitter and other social media through unsigned networks. These networks with only positive ties systematically neglect or misinterpret negative, antagonistic, sometimes hostile user interactions. We have shown that some of the directed messages to other users (through mentions) do not constitute a “flow of information” (13), but are rather expressions of antagonism, contention and disagreement of the type that sociologists have long argued are central to the process of group formation. These findings have implications for a broad range of research using social media data, suggesting that research needs to begin considering the sign of the interaction when employing network representations of debates.

The primary limitation of the approach introduced in this study is that it requires a labeled set of training data to use supervised machine learning to detect the interaction sentiment in tweets. In contrast to other popular machine learning classification tasks, such as sentiment detection, there are currently no pre-trained classifiers or training data available. However, future research might provide such resources. As this study has focused on a specific debate, embedded in a specific time period, country, and social media platform, future research may study whether the identified patterns hold more broadly, by expanding its approach to study group structures and intergroup communication online in a variety of political debates, countries and platforms. Future research may also focus on what this signed network representation can tell us of the dynamics of political polarization in social media, by shifting our understanding of online polarization from isolation and fragmentation to conflict and confrontation.

Chapter 4.

Echo Chambers are Defined by Conflict, not Isolation

Conditionally accepted

Sociological Science

Abstract

The influential “echo chamber” hypothesis suggests that social media drive polarization through a mutual reinforcement between isolation and radicalization. The existence of such echo chambers has been a central focus of academic debate, with competing studies finding ostensibly contradictory empirical evidence. This paper identifies a fundamental methodological limitation of these empirical studies: they do not differentiate between negative and positive interactions. To overcome this limitation, we develop a method to extract signed network representations of Twitter debates using Machine Learning. Applying our approach to a major Dutch cultural controversy, we show that the inclusion of negative interactions provides a new empirical picture of the dynamics of online polarization. Our findings suggest that conflict, not isolation, is at the heart of polarization.

Main

Politics in many countries has in recent decades entered an era of unprecedented political polarization, with growing divides between political camps and harshening public discourse (Carothers and O’Donohue 2019; Harteveld 2021; Hobolt, Leeper, and Tilley 2021). Scholars have long discussed new media technology as a potential driver of this rise of polarization (Eady et al. 2019; Iyengar and Hahn 2009; Lelkes, Sood, and Iyengar 2017; Pariser 2011; Prior 2007; Sunstein 1999, 2001a; Sunstein and Vermeule 2008), with the so-called “echo chamber” as the most prominent causal link. According to this hypothesis, new media technology facilitates the formation of clusters of likeminded individuals (McPherson, Smith-Lovin, and Cook 2001). When such homogeneous groups are insulated from opposing perspectives, their views and biases are reinforced rather than moderated, resulting in polarization (Bakshy, Messing, and Adamic 2015; Lawrence, Sides, and Farrell 2010; Mutz and Martin 2001; Nikolov et al. 2015; Schmidt et al. 2018; Stroud 2010; Sunstein 1999; Del Vicario et al. 2016). While this hypothesis is intuitive and has been widely adopted by scholars as well as the general public, a growing number of empirical studies find that social media users are in fact engaging in significant interaction across ideological divides (Bakshy et al. 2015; Barberá 2020; Colleoni, Rozza, and Arvidsson 2014; Conover, et al. 2011a; Goel, Mason, and Watts 2010; Gruzd and Roy 2014; Vaccari et al. 2016; Yardi and Boyd 2010; Yoon and Park 2014). Recent literature questions the prevalence of online echo chambers and, by implication, the role of new media technology in driving polarization (Bail 2012; Baberá 2020; Guess et al. 2018).

This paper points to a fundamental methodological limitation of many empirical studies on echo chambers: they do not differentiate between negative and positive interactions. Leveraging the use of networks to study social dynamics online (Lazer et al. 2009), user interactions in this line of research are represented as ties, and patterns in discussions are mapped through graphs (Bakshy et al. 2015; Barberá et al. 2015b; Conover et al. 2011). Typically, researchers do not distinguish between negative and positive interactions: ties can vary in strength, but they are generally assumed to be positive. Such a representation limits our understanding of the role of conflict in social media debates by either ignoring negative interactions or conflating them with positive interactions. In this study, we present a method for bringing conflict into these network representations. Acknowledging the positive or negative sentiment of interactions provides a richer empirical analysis and allows us to go beyond the echo chamber hypothesis towards a more nuanced examination of the driving forces of online polarization.

Taking the Dutch cultural conflict around Zwarte Piet on Twitter as a case, we use Natural Language Processing to classify interactions as either positive or negative. We ask: how do groups emerge out of positive and negative interactions in online debates? The signed network analysis shows that most of the cross-ideological interaction is negative, involving direct attacks or forms of outgroup derogation. By comparing the results of our signed network analysis with traditional unsigned network approaches, we find that acknowledging negative interactions results in qualitatively different understanding of online polarization, revealing different types of actors and different kinds of relationships. Moreover, we show that polarization is asymmetrical, with opposing groups displaying uneven levels of hostility.

Our results provide empirical evidence for an alternative theoretical understanding of echo chambers, by positing *conflict*, not isolation, as the driver of social media polarization. While the traditional echo chamber hypothesis suggests a mutual reinforcement between isolation and radicalization as the core feedback process behind polarization (Sunstein 1999, 2001) the findings of this paper align with a recent model which suggests that social media may in fact drive polarization by *intensifying* conflictual interaction (Tornberg 2022). Such a perspective aligns with classic characterizations of the sociology of conflict, as developed by Simmel (Simmel 1904a; 1904b; 1904c) and elaborated by Coser (Coser, 1957) and Collins (Collins 2012). This provides the starting point for reorienting computational research on polarization to better account for conflict and develop a more complex understanding of the impact of social media on political life.

Echo Chambers on Social Media

The notion of the “echo chamber” suggests that new media technologies enable us to avoid the discomfort of exposure to opposing ideas or opinions by letting us choose to connect with likeminded people and seek out information that confirms our views. Scholars have argued that such “selective exposure” (Garrett 2009) has even become automated, as algorithms personalize our news and information environment, creating “filter bubbles” that shield us from dissenting views (Pariser 2011). Selective exposure can result in a polarization feedback loop, in which partisan media exposure strengthens beliefs which in turn reinforces media selection (Slater 2007; Stroud 2010). This literature thus suggests that polarization fundamentally comes about through lack of exposure to alternative views, resulting in a breakdown of the foundations of democratic pluralism (Quattrociocchi, Scala, and Sunstein 2016; Sunstein 1999, 2001b; Törnberg 2018)

A significant strand of empirical research has set out to test the echo chamber hypothesis by examining the structure of interactions on social media, especially Twitter. Early studies employed network representations of retweets between users – generally understood as a sign of endorsement (Metaxas, 2015) – to examine the structure of political debates. These studies tend to find that users indeed are organized in separate network communities, with limited interconnections across political divides (Barberá et al. 2015b; Conover et al. 2011a; Guerrero-Solé 2017; Himelboim et al. 2016; Recuero, Zago, and Soares 2019). In fact, retweet structures and political ideology are so strongly correlated that user ideology can be predicted from retweet networks (Barberá 2015b; Conover et al 2011b). These findings have been taken as support for the echo chamber hypothesis (McPherson et al. 2001; Mutz and Martin 2001; Stroud 2010; Sunstein 2001; Sunstein and Vermeule 2008)– an idea that has become widely accepted and highly influential in the public debate (D'Costa 2017; El-Bermawy 2016; Hooton 2016).

However, studies that construct their networks based on mentions between users, rather than retweets, typically come to the opposite conclusion: they do not find a significant divide between opposing groups (Barberá 2015b; Conover et al. 2011a; Honeycutt and Herring 2009; Williams et al. 2015; Yardi and Boyd 2010) but suggest that "mentions form a communication bridge across which information flows between ideologically-opposed users" (Conover et al. 2011:6). Interpreting mentions as neutral interaction, or as expressions of information exchange (Borge-Holthoefer et al. 2011), these studies have been used to question the prevalence of echo chambers – and at times even the link between social media and political polarization (Bail 2021; Baerberá et al. 2015b; Bruns 2017, 2019, 2021; Guess et al. 2018; Törnberg et al. 2021).

The assumption that interaction across political divides could be understood as neutral, or as "information flows" (Conover et al. 2011a), can however be questioned based on findings of qualitative research. For instance, Evolvi (2019:396) finds that mentions are often used to "belittle others with different ideas rather than invite conversation." Studying climate change debates on Twitter, Moernaut, Mast and Temerman (Moernaut et al. 2020b) find that climate change sceptics and believers do not engage in constructive debate but rather aim at delegitimizing and dehumanizing one another. Gruzd and Roy (Gruzd and Roy 2014b) manually labelled tweets across party lines in Canadian politics, finding that roughly half are hostile. Williams et al. (Williams et al. 2015) report evidence that mentions express positive ingroup and negative outgroup sentiments. These findings suggest an important limitation at the heart of the current approach to studying echo chambers in social media: this research has either left out negative interactions altogether (studies of retweets) or conflated positive and negative interactions (studies of mentions), thus missing the conflictual dimension of online debates. This points to the need of developing methods that allows bringing conflict into the study of social media polarization.

Signed Networks Analysis

Social network analysis has in recent years grown into one of the foremost means to quantitatively study social structures (Lazer et al. 2009). This body of research almost exclusively represents interactions between individuals as positive or neutral, treating highly connected groups of individuals as belonging to the same community (Harrigan, Labianca, and Agneessens 2020). Recent years have seen the incipient growth of empirical research on signed network, which goes beyond this representation by acknowledging the polarity of ties, with negative ties implying repulsion and positive ties attraction. Such networks have been employed to test social balance theory (Cartwright and Harary 1956; Heider 1946c) and status theory (Guha et al. 2004b; Hassan, Abu-Jbara, and Radev 2012d, 2012b; Leskovec, Huttenlocher, and Kleinberg 2010d; Sadilek, Klimek, and Thurner 2018b; Zheng, Zeng, and Wang 2015b), to measure the impact of negative ties in organizational context (Labianca and Brass 2006) and to predict tie formation (Tang et al. 2016a) (for an overview of research on signed social networks see (Harrigan et al. 2020)). A few studies have employed signed networks to study polarized debates (Neal 2020; Traag and Bruggeman 2009b; Uitermark, Traag, and Bruggeman 2016c). However, the potential of signed network analysis to study polarization on social media has remained underutilized, with a small number of studies that are relatively small in scope.

The limited use of signed network analysis is in large part a result of the relative difficulty involved in extracting signed network data from social media. Scholars have had to either manually classify relations as positive or negative (Gruzd and Roy 2014b; Moernaut et al. 2020b), focus on one of a small number of niche social media platforms that employ negative ties (Guha et al. 2004b; Kunegis, Lommatzsch, and Bauckhage 2009; Leskovec et al. 2010d; Leskovec, Huttenlocher, and Kleinberg 2010a), or make strong assumptions on the sign of ties (De Stefano and Santelli 2019b; Yardi and Boyd 2010b). None of these paths are passable for the application of signed networks to the study of large-scale dynamics of echo chambers on mainstream social media platforms such as Twitter.

This paper presents a method for moving beyond this impasse by automatically extracting signed networks from large-scale social media debates by using Natural Language Processing and Machine Learning. We manually code the sentiment expressed in relation to mentioned users for a large sample of tweets. Each user-to-user mention was classified as expressing agreement (positive), disagreement (negative), or as ambiguous/neutral towards to mentioned user. This data was used to train a Machine Learning algorithm to automatically classify mentions as positive or negative, resulting in a large signed social network of user relation (See the Method section for additional details.)

The Conflictual Debate about 'Zwarte Piet'

'Zwarte Piet' ('Black Pete') is traditionally a key figure in celebration Sinterklaas, a Dutch variant of Christmas. While Sinterklaas looks somewhat like Santa Claus (an old white man dressed in red, often riding a white horse), his helper Zwarte Piet is performed by white people wearing blackface, with exaggerated red lips, a black curly wig, and large golden earrings. Zwarte Piet has in recent years become a lightning-rod in the Dutch culture wars, bringing about a protracted and intense national debate (Chauvin, Coenders, and Koren 2018; Vliegenthart and Zuure 2020; Wekker 2016). Although the debate on Zwarte Piet is complex and sprawling, in essence, opponents of Zwarte Piet see the abolishing of the Zwarte Piet character as one key battle in the struggle against the legacy of colonialism and racism (Helsloot 2009; Rodenberg and Wagenaar 2016b; Vliegenthart and Zuure 2020; Wekker 2016). In contrast, supporters of Zwarte Piet say that the character is not racist and perceive the suggestion to change Zwarte Piet as an attack on a valued Dutch tradition (Helsloot 2012b; Rodenberg and Wagenaar 2016b; Vliegenthart and Zuure 2020).

We use a dataset of tweets on Zwarte Piet to study this contentious debate. The data cover the period from December 2017 to May 2019, comprising roughly 430,000 tweets, from 81,700 unique users, with 296,881 unique mentions between users. To examine how the differentiation of negative, neutral, and positive ties changes the findings of network analysis, we compare the signed network with the unsigned retweet and mention networks that are generally employed to study polarization.

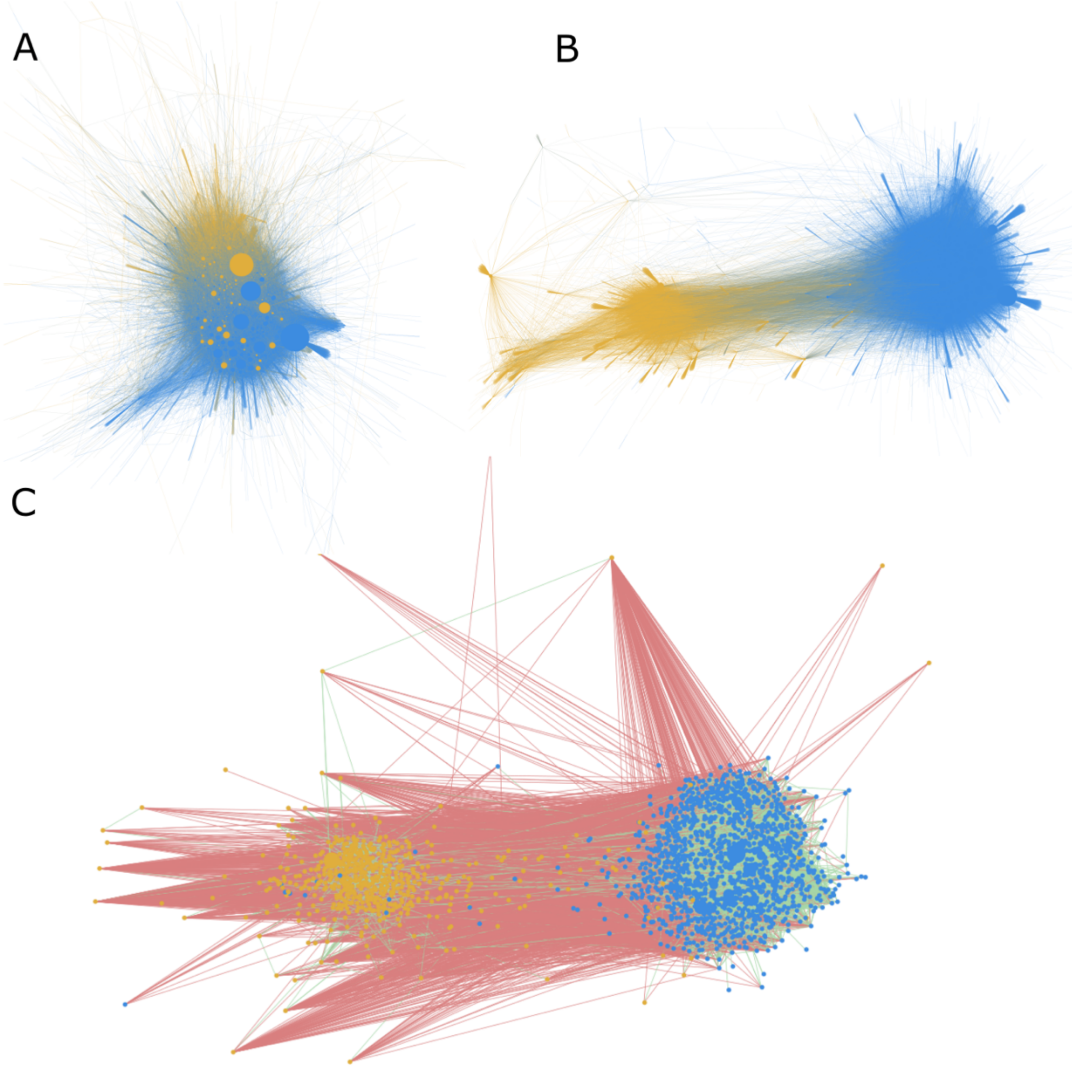


Figure 1: Visualizations of the mention network (A), retweet network (B) and signed network (C) constructed from the Zwarte Piet Twitter data. Users are positioned according to the Force Atlas algorithm and colored by their stance in the debate (proponents blue and opponents yellow). The positive and negative edges in the signed network are colored green and red respectively. In the mention and retweet network the Force Atlas algorithm (implemented in Gephi) draws connected users closer together, whereas in the signed network the adjusted Force Atlas algorithm (implemented in Sigma.js) attracts nodes connected with a positive edge and repulses nodes connected with a negative edge. The Figure shows that the mention network forms a seemingly cohesive whole, the retweet network is split into two opposing groups with little cross interaction and the signed network displays two groups with positive ingroup and negative outgroup interactions.

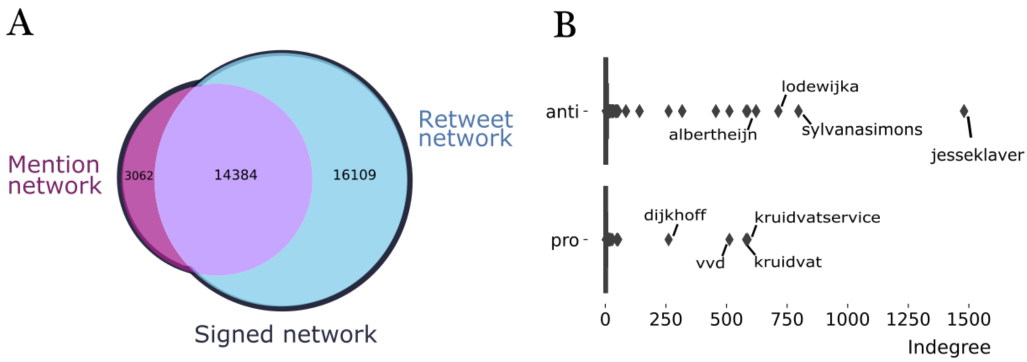


Figure 2: Differences and similarities between the sets of users identified in, respectively, the retweet network, the mention network, and the signed network. The Venn diagram of users sets (Figure 2a) shows that the signed network contains the union of users in the mention and retweet network. The boxplots (Figure 2b) give an impression of the prominence of users that are not included in the retweet network (because they did not send tweets or retweets on the topic). Although the retweet network includes most users (n=30,493 or 90.1%), it misses some prominent users. For example, politician Jesse Klaver, an opponent of Zwarte Piet, received over 1.400 mentions.

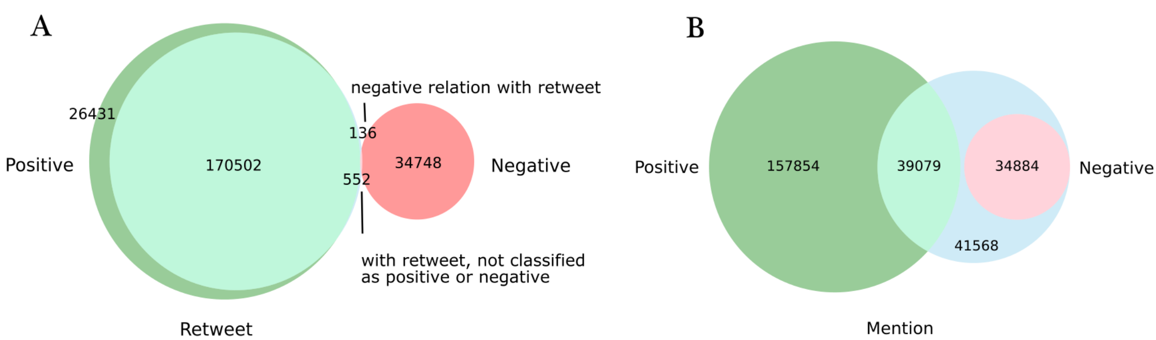


Figure 3: Differences and similarities between the sets of relations identified in, respectively, the retweet network, the mention network, and the signed network. The Venn diagrams demonstrate how the signed relations between users—positive (green) and negative relations (red)—compare to the user-to-user relations based on retweets (Figure 3a) or mentions (Figure 3b). Taking the signed relations as a baseline, the figure shows that retweet relations capture the large majority of positive relations but miss all negative relations between users. The mention relations include 19.8% of the positive relations and include all relations identified as negative but conflate these two types of ties. The mention relations that are not included in the signed relations (n=41.468) are mostly based on only one interaction (85.9%).

Echo Chambers through the Lens of Signed Network Analysis

The included nodes and edges constitute a first notable difference in how the retweet, mention, and signed networks represent the Twitter data: comparing the two unsigned networks in terms of nodes, we find that the retweet network includes more users (30,493; 90.1% of the users in the signed network) than the mention network (17,446; 52.0% of the users in the signed network) (see Figure 2a). While the number of users that the mention network includes but the retweet network misses is relatively small (3,062; 9.1% of the users in the signed network), it includes prominent and important actors in the national debate (see Figure 2b), such as the politicians Lodewijk Ascher (@lodewijka), Klaas Dijkhoff (@dijkhoff), and Jesse Klaver (@jesseklaver). The signed network contains users from both the mention network and the retweet network (33,555). In terms of edges, the most striking feature of the retweet network is that is misses all negative user-to-user interactions. Additionally, it leaves out a substantial fraction of the positive relations (26,413; 13.4%; see Figure 3a.) The mention network misses the majority of positive user-to-user relations (157,854; 80.2% of relations classified as positive; see Figure 2b), and misrepresents a substantial number of negative user-to-user relations (34,884; 30.2% of all mention relations; see Figure 2.)

We now turn to examine what these differing network representations tell us about echo chambers. Focusing on retweets, it seems that the two sides of this debate operate in separate universes, with little cross-ideological interactions (see Figure 4), in line with studies that interpret such a structure as empirical support for the existence of echo chambers (Barberá et al. 2015b; Conover et al. 2011; Conover et al. 2011a; Guerrero-Solé 2017; Himelboim et al. 2016; Soares, Recuero, and Zago 2019). The large majority of retweet relations connect internally either proponents (76.1%) or opponents (20.4%), and only a negligible fraction cuts across the divide (3.5%).

The mention network view on echo chambers, however, suggests a very different picture. This network is highly integrated, with frequent connections between the two opposing sides. In fact, a large majority (64.5%) of mention relations connects users with different stances (see Figure 4.) This is in line with the research that has been seen as providing evidence to question the echo chamber hypothesis, finding that "mentions form a communication bridge across which information flows between ideologically-opposed users" (Conover et al. 2011:6) (Conover et al. 2011; Honeycutt and Herring 2009; Yardi and Boyd 2010).

The signed network provides a third perspective, which posits an explanation for the contradictory findings and interpretations. The signed analysis shows that while there are indeed many connections across the political divide, the majority of such intergroup relations (81.6%) are negative. Intragroup relations, in contrast, are almost exclusively positive (96.6%) (see Figure 4). This perspective provides a qualitatively different view from both the proponents and opponents of the echo chamber hypothesis: there are ideological groupings on social media, but they are not defined by the flow of information, but by active conflict. The sides are engaged in expressing support and solidarity for their ingroup while deriding the outgroup. Members of ideological groups use retweets and mentions to express support and solidarity for the ingroup, and condemnation for the outgroup. Rather than Sunstein's (Sunstein 1999) suggestion of homogeneity being the driver of polarization, this suggests that polarization results from a feedback cycle between external conflict and internal solidarity (Collins 2012; Coser 1957).

So far, we identified groups in the debate based on their issue position (are they in favor or against Zwarte Piet?). Another common method of defining groups is to infer them from interactions by means of community detection. Do our conclusions hold up when we use this method? To answer this question, we apply the Constant Potts Model (Reichardt and Bornholdt 2006) implemented in Leiden algorithm (Traag and Bruggeman 2009; Traag, Van Dooren, and Nesterov 2011; Traag, Waltman, and van Eck 2019) which can use both positive edges (attracting nodes) and negative edges (repulsing nodes), see Methods for details. For the retweet network, we find a clear divide between two groups (CPM quality=162,862), which aligns closely with the two sides of the debate (one group has 94% proponents, the other 95% opponents; see Figure 1.) The mention network is instead structured as a single cohesive community, including users from both sides of the divide (CPM quality=112,367, see Figure 1). The signed network again improves both these representations, showing two sides engaged in conflict, with relations between opposing groups being predominately negative (85%; see Figure 1).

The signed network analysis thus suggests a fundamentally different picture of echo chambers: they are not isolated groups of likeminded people, nor are echo chambers non-existent due to connections across political divides. Instead, they are defined through the conflictual interactions across groups and expressions of support and solidarity within them. This perspective allows to look deeper into the structure of intergroup conflict, bringing to the fore asymmetries between the two sides of debate, the topic to which we now turn.

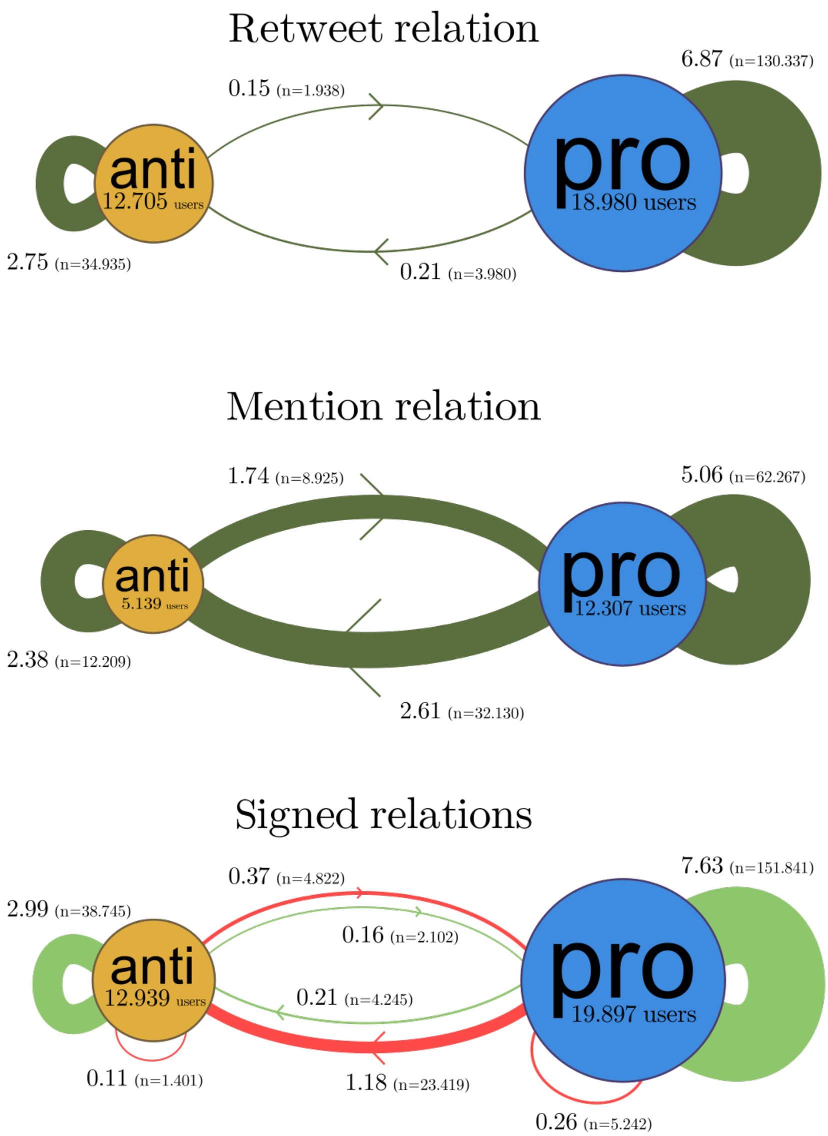


Figure 4: These figures display the retweet, mention, and signed relations between the two sides. The edges are weighted as the count of the type of relation divided by the size of the user-group, e.g., the retweet pro-pro edge has weight 130,337/18,980=6.87, which can be interpreted as an average, e.g. opponents retweet on average 6.87 other opponents. This figure shows that the aggregate interactions between proponents and opponents depends on which types of interaction (retweet, mentions, or signed) are considered.

Structure of Intragroup Conflict

Signed network analysis allows us to cast light on asymmetrical polarization. Such asymmetry is an important theme within the literature on political polarization (Barberá 2020; Grossmann and Hopkins 2016) and signed network analysis enables us to examine asymmetries in the structure and levels of conflictual relations. Since the opposition between supporters and opponents of Zwarte Piet maps onto the opposition between conservatives and progressives, we relate findings from our case study to broader literature on asymmetric polarization.

We should first note that there are remarkable differences in terms of activity, see Figure 5. On average, proponents of Zwarte Piet tweet 2.5 times more than opponents (14.7 versus 5.7 tweets per user) and have roughly 2.5 times as many outgoing edges in the network (9.29 versus 3.64). The proponents are also more confrontational: a third (33.3%) of the proponents has at least one negative outgoing link in the network, compared to only 18% of opponents, and the negative interaction rate (the number of outgoing negative edges divided by total outgoing edges per user) is twice as high for proponents than for opponents (0.07 vs. 0.14). As a result of their higher levels of activity and negativity, proponents have on average three times more negative outgoing links than opposed users (1.44 versus 0.48), most of which are directed towards the opposing side.

Looking at the distribution of negative links per users (Figure 6), we find that the distribution of negative links over users is highly skewed: the proponents of Zwarte Piet include a group of very active and highly confrontational users who account for a large proportion of the negative relations overall. Since most of the negative tweets are directed to people holding a different position on Zwarte Piet, it is unsurprising that the top targets of attack tend to be opponents of Zwarte Piet. A few key figures receive most of the negative ties. Opponents of Zwarte Piet users represent 72.2% of the top-1% (320 users) most negatively referenced users, and 87.1% of the top-0.1% (32 users).

These results suggest, first of all, that there is indeed asymmetrical polarization: supporters of Zwarte Piet are much more active in the debate and are more negative. This means not only that different sides to the debate behave differently but also that they receive different treatment: opponents of Zwarte Piet are much more likely to have negative tweets directed against them. Moreover, we find that a very active and highly negative group of supporters of Zwarte Piet accounts for much of this pattern. These findings suggest that (1) interactions across the political divide often take the form of criticism, derogation or intimidation, and that (2) supporters of Zwarte Piet account for a much larger portion of the negativity than proponents. The results not only confirm asymmetrical polarization, but they are also in line with other recent social media research that shows conservatives interact more across partisan divides (Eady et al. 2019; Wu and Resnick 2021)– but adding the central point that this interaction more often tends to be confrontational.

Issue sentiment	anti	pro
Total (re)tweets	5.72	14.73
Negative links	0.48	1.44
Positive links	3.16	7.84
Total links	3.64	9.29
Positive links to anti	2.99	0.21
Positive links to pro	0.16	7.63
Negative links to anti	0.11	1.18
Negative links to pro	0.37	0.26
Negative interaction rate	0.07	0.14

Figure 5: Average statistics for outgoing links in the signed network for opponents (anti) and proponents (pro) of Black Pete. The negative interaction rate is calculated by the number of negative outgoing links divided by total outgoing links of each user, and thereafter averaged over all users.

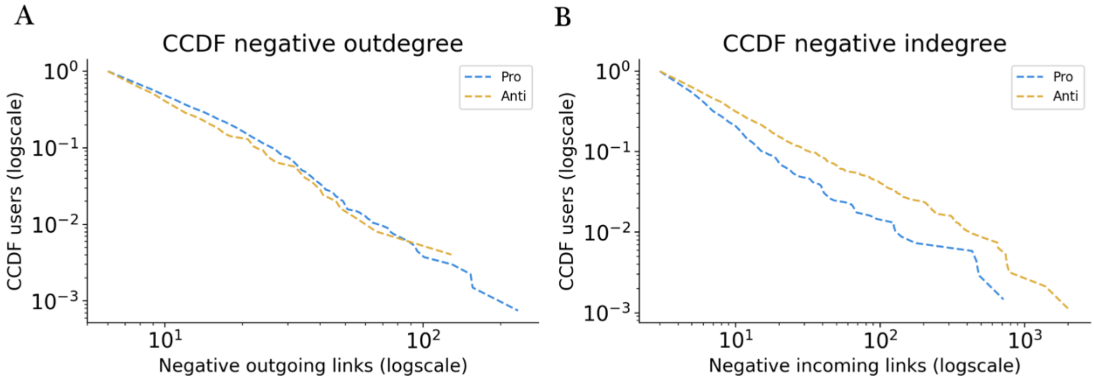


Figure 6: CCDF of negative outgoing links (A) and incoming links (B) per user. The highly skewed distributions of negative outdegree shows that most users have few negative outgoing relations, but there are some users, more among the supporters of Zwarte Piet (blue) than among opponents (orange), with many negative outgoing edges. The different slope between the negative indegree distribution between proponents and opponents shows that latter structurally receive more negativity.

Discussion

Studies of polarization on social media have been limited by their use of unsigned networks, which requires researchers to either leave out negative interaction (by, for example, looking at Twitter retweets only), or to conflate negative and positive relations (by, for example, encoding Twitter mentions as positive network ties). This has led to puzzling empirical findings with respect to the echo chamber hypothesis, as studies either find ideological echo chambers or substantive cross-ideological engagement, depending on whether they choose the former or the latter approach to representing user interaction.

This paper has presented a method for moving beyond this methodological impasse, by introducing a method for extracting signed networks from Twitter data and using this to throw new light on the nature of echo chamber and interaction across the political divide on social media. We compare the picture of a Dutch cultural controversy on Twitter as represented by three forms of network representations: the mention network, the retweet network, and the signed network. In line with previous research, we find that the retweet network reveals two separate echo chambers, with users self-segregated into two isolated groups. The mention network, contrastingly, shows substantial communication across the ideological divide. These seemingly contradictory findings are resolved by using a signed network, which shows that that most of the cross-ideological user interaction is negative, while ingroup interactions are almost exclusively positive. This presents a novel understanding of echo chambers, in which they are defined not by isolation, but by intergroup conflict.

The signed network analysis allows a clearer image of the structure of online polarization, revealing asymmetries between the political fractions. While previous studies have taken the higher level of interaction to suggest that conservatives are less locked into echo chambers than progressives, the signed network representation suggests that conservatives are more active and more prone to attack their outgroup. As some of the most popular target of these attacks include organizations or individuals that have neither stake nor position in the debate – such as supermarkets or political parties – this suggests that the aim of these interactions is not always to convince the opposing group, but also to show allegiance with the ingroup through symbolic attacks against the outgroup. This aligns with research suggesting the interaction with users of different view may not trigger moderation – as the common version of the echo chamber suggests – but may instead further intensify polarization (Bail 2011; Bail et al. 2018). It furthermore contributes important empirical evidence for a recent conflict-driven model of social media polarization, which “turns the echo chamber on its head” (p2) by suggesting that social media may be polarizing by *increasing* interaction across the political divide, rather than isolating political opponents (Tornberg 2022).

These findings reveal how methodological choices of data representation can come to have profound consequences for how we understand social phenomena. Unsigned network representations bring about a false dichotomy between isolation and interconnection by ignoring conflictual interactions. This erases the central role of conflict and solidarity in online polarization, and results in a theoretical foundation in which isolation is taken to be the driver of polarization (Bishop 2009; Pariser 2011; Sunstein 1999; Sunstein and Vermeule 2008). When instead representing social media interaction as signed networks, the resulting understanding of online polarization shifts significantly, revealing echo chambers as defined not by isolation from information flows but through intragroup solidarity and intergroup antagonism. This points to a very different theoretical foundation for the dynamics of polarization, drawing on a long sociological tradition which puts conflict as the chief driver of polarization. Scholars in this tradition, such as Simmel, Coser, and Collins, suggest that polarization results from a feedback loop in which external conflict drives internal solidarity, and vice versa. In this framework, to the extent that social media facilitates polarization, it is not because it isolates opposing communities, but *au contraire*, because it faces them off in contentious confrontation. The approach introduced in this paper captures the conflictual dimension of polarization, allowing more nuanced insight into the underlying social mechanisms that tear social media users apart or pull them together.

Methods

Twitter Data

We gathered tweets on the Black Pete debate by keyword matching of various terms related to the debate[14], such as “Black Pete”, “Zwarte Piet” and “KOZP” (abbreviation for “Kick Out Zwarte Piet”), using the Twitter Capture Analysis Toolset (Borra and Rieder 2014) and removed tweets that were not written in Dutch. The tweets in our dataset were published

[14] The full list of case insensitive keyword matching is ['black pete', 'zwarte piet', BlackPete, BlackPeteIsRacism, BlackPiet, KOZP, ZwartePiet, ZwartePietIsRacisme].

between December 4th, 2017 and May 7th, 2019 and include original tweets and unquoted retweets. In total, the dataset contains 418,421 tweets, from 61,543 unique users, with 174,555 unique mentions between users.

Classifying Users' Stance

In large scale studies on political polarization on Twitter the ideological stance of users is typically inferred from their interactions [cf. 31,91]. In this study we explicitly opt not to do this, in order to investigate how users from the same and different stance interact with one another in the debate. Additionally, we depart from previous approaches by not limiting the analysis to the more active users is the debate which creates a bias towards the vocal minority to the detriment of the more silent majority (Mustafaraj et al. 2011). Instead, we select a method for inferring user's stance that is as inclusive is possible. Our strategy, then, is to classify the position towards Black Pete that is expressed in all the (re)tweets of the user (pro, anti, or neutral/ambiguous) by examining the full tweet texts and use that to deduce a user's stance in the Black Pete.

To classify the position expressed in tweets, we semi-manually classified a sample of tweets and thereafter trained a Machine Learning Algorithm with this data. The sample data consists of two sets of tweets. First, 4,787 (2.7%) tweets were selected at random from the full set of unique tweets (n=179,712). These tweets were manually labeled with the assistance of four fluent Dutch speakers. Each tweet was assigned one label: pro, anti, neutral or ambiguous. The codebook instructions were conservative: if the stance toward Black Pete is not self-evident, the tweet was labeled as ambiguous. From the coding efforts we learned that it was often difficult to distinguish neutral from ambiguous tweets and we found few tweets (n=512) that were coded as expressing a neutral issue sentiment. Therefore, we merged the neutral and ambiguous tweets into one category for subsequent classification purposes. The inter-coder agreement was measured by a Krippendorf Alpha of 0.724.

The second set of the training data consists of tweets by prominent pro and anti-users in the debate. We selected the top 1% accounts of users that were most active either in terms of retweets, mentions or the number of unique users mentioned or retweeted. Similarly, we selected the top 1% of the accounts that received most retweets or mentions, in terms of frequency and in terms of the number of unique source users. These top users (n=329) were manually labeled as having a pro (n=231), anti (n=59), or neutral/ambiguous (n=39) stance in the debate. The tweets of these top pro (anti) users that weren't also retweeted by the opposite side of the debate, were then classified as pro (anti). This resulted in an extra 26,323 labeled tweets.

After splitting the data into a test set (40%) and a training set (60%), we downsampled the tweets with a pro label in the training set to equal the number of anti-tweets in the training set (n=2,607) to avoid biases in the classification. We resampled the test set to have the same distribution of pro, anti and neutral/ambiguous tweets as the original full dataset, respectively 60%, 12% and 28%. Next, we used this data to train the fastText algorithm (Joulin et al. 2017) with pretrained word embeddings on a Dutch Wikipedia Corpus (Bojanowski et al. 2017), maximizing the F1 score for all classes, thus attempting to predict all classes well, in both precision and recall. The fastText algorithm gives an indication of how certain the classification is (the softmax probability), valued between 0 and 1 for each prediction. We use this certainty indication to apply a simple rule: classify all tweets with lower certainty (<0.99) as neutral/ambiguous. This procedure reduces the recall for the pro and anti-classes, but also, more importantly, reduces the errors we care most about: classifying pro tweets as anti and classifying anti tweets as pro.

The classifier—after applying the certainty rule—categorizes the issue stance with sufficiently high accuracy; see Figure 7. There are only 3 cases in which an anti-tweet is misclassified as pro (0.011) times of all anti tweets and 0.0044 times all pro tweet classifications) and 24 cases in which a pro tweet is misclassified as anti (0.019 times of all pro tweets and

0.1 times of all anti classifications). Aggregating tweets per user and applying a simple majority rule results in 14,353 clear anti users (14%), 23,581 (38%) clear pro users and 23,637 (38%) users that we couldn't unambiguously classify as pro or anti. Most of the users with ambiguous stance have just one (re)tweet (75%) and we therefore left them out of the subsequent analysis without jeopardizing the robustness of our results.

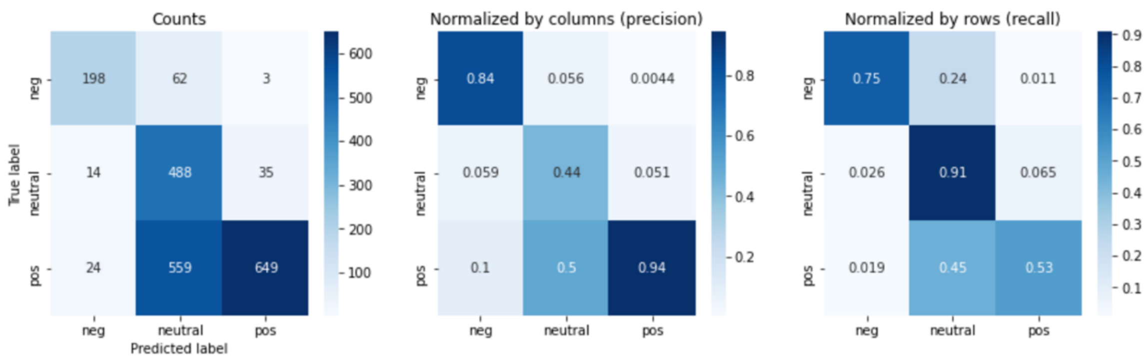


Figure 7: The results of the classifier (parameter values: epoch=10, learning rate=0.7, n-grams=3) after applying the simple certainty rule (neutral if certainty < 0.99): confusion matrix with counts (left), normalized by the true labels (middle) and normalized by the predicted labels (right). The values in the diagonals of the middle matrix are the precision rates, and the values on the diagonals of the right matrix are the recall rates. Recall rates here are reduced due to the certainty rule, but the most important errors (classify positive if true value is negative and classify negative if true value is positive) are reduced.

Classifying signs of interaction

Since the sign of interaction is an integral part of this study, we aim to measure this as accurately as possible. Previous studies have relied on heuristics to infer the sign of interaction, for example using the balance theoretical notion that the enemy of my friend is my enemy (Cartwright and Harary 2956; Heider 1946). Instead, we analyze the text of tweets from one user directed towards another user by the use of mentions (e.g. @username). We classify for each mention whether the source user is expressing endorsement (positive), disagreement (negative) or an ambiguous (neutral) sentiment towards the mentioned user. This type of classification cannot be addressed with existing algorithms for sentiment analysis, since the sentiment towards the mentioned user is not necessarily aligned with the sentiment of the tweet: a tweet expressing a positive sentiment can contain a negative (hostile) mention (e.g. "Hee Sjerrie @TheRebelThePoet that is great news right 😎 long live #BLACKPETE"[15]). Instead, our general strategy is to select a sample of mentions at random from the full dataset, label these manually, and with this sample train a Machine Learning Algorithm. This method is inclusive with respect to different types of users, reducing the known bias in earlier research towards the vocal minority to the detriment of the more silent majority (Mustafaraj et al. 2011).

The random sample data (n=6,056, 3.5%) contains unique user-to-user mentions from users that we identified as pro or anti. These tweets with mentions were then manually labeled with the assistance of four fluent Dutch speakers. Each mention was assigned one label: positive, negative, or neutral/ambiguous. The codebook instructions were conservative: if the interaction sentiment was not self-evident, the mention was labeled as neutral/ambiguous. The inter-coder agreement was measured by a Krippendorf Alpha of 0.42.

After splitting the labeled data into a training set (70%) and test set (40%), we removed all the mentions in the test set that occurred in a tweet that was also included in the training set (since one tweet can contain various mentions). We added features with (1) the predicted stance of the source user, (2) the predicted stance of the mentioned user, (3) whether the mention takes the form of 'via @username'—which are most often neutral, as they are automatically added by the webserver of the media outlet via which the tweet was posted—and (4) whether the mention is located at the start, body or end of the tweet since that might correlate with the polarity of the mention.

[15] The original tweet was in Dutch "Hee Sjerrie @TheRebelThePoet dat is toch geweldig nieuws of niet dan😎 leve #ZWARTEPIET". In this tweet, a proponent of Zwarte Piet poses a rhetorical question to taunt TheRebelThePoet, a prominent opponent of Zwarte Piet. The tweet's sentiment is positive as the user expresses joy about news that is favorable to Zwarte Piet, but the mention to TheRebelThePoet is negative.

Next, we used this input to train the fastText algorithm (Joulin et al. 2017). In order to teach the algorithm the basics of Dutch and Twitter language, we also provided fastText with a word embedding learned from a corpus of approximately 180 million Dutch Tweets posted in 2018 (see Supplementary material). We trained the algorithm to maximize the F1 score for all classes, thus attempting to predict all classes well, in both precision and recall. The classifier categorizes the mention sentiment with sufficiently high accuracy, F-1 score 0.67, see Figure 8.

We aggregated all the user-to-user interaction and labeled them based on a simple majority rule: if most of the user-to-user interactions were positive (negative), we classified the directed sign between these users as positive (negative). Retweets were here considered as positive interaction from the retweeting to the retweeted user (Metaxas, 2015). This procedure classified most user relations as positive (n=216,067, 75%), 12 % as negative and 14% of the relations we could not classify with sufficient certainty. Most of the unclassified relations (86%) are based on one interaction only, and we therefore left them out of the subsequent analysis without jeopardizing the robustness of our results.

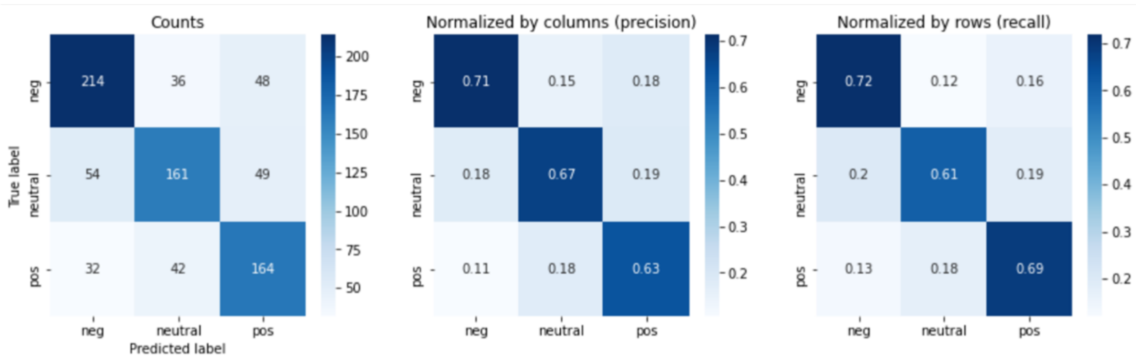


Figure 8: The results of the classifier (parameter values: epoch=20, learning rate=0.65, n-grams=3): confusion matrix with counts (left), normalized by the true labels (middle) and normalized by the predicted labels (right). The values in the diagonals of the middle matrix are the precision rates, and the values on the diagonals of the right matrix are the recall rates. Recall rates here are reduced due to the certainty rule, but the most important errors (classify positive if true value is negative and classify negative if true value is positive) are reduced.

Community detection

Comparing the community structure of the retweet network, mention network and signed network requires a method for community detection that can detect communities based on positive ties as well as negative ties. Additionally, we would like to compare the three networks at the same level of resolution. We therefore used a generalization of the Constant Potts Model (Reichardt and Bornholdt 2006) implemented in Leiden algorithm (Traag and Bruggeman 2009; Traag et al. 2011; 2019). This method for community detection considers the sign of ties by maximizing positive ties within communities and minimizing negative ties within communities and allows to look at different granular scales at the community structure in the network.

We detected the community structure in the retweet network, mention network and signed network with resolution parameter gamma set to equal 0.0001. As reported in the result section in more detail, the users in the mention network are all (99.6%) part of one community on this level, but both the retweet network and signed network feature two communities[16]. These communities in the retweet and signed network are very similar in their composition, see Figure 9. As reported in the result section, some small differences between the community compositions have far reaching implications, since some of the users that the retweet network excludes are frequently attacked.

[16] 99.1% the nodes in the retweet network are part of these two largest communities; 99.0% of the nodes in the signed network are part of the two largest communities.

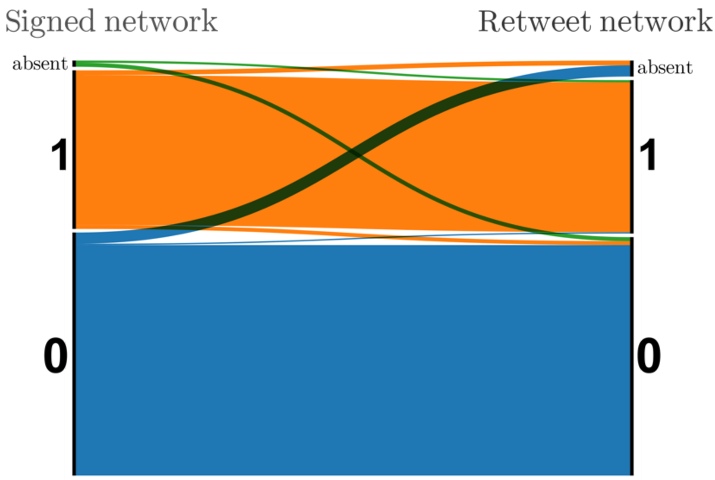


Figure 9: Alluvial graph illustrating the comparison of community compositions between the signed network (left) and retweet network (right). The thickness of lines corresponds to the number of users. The figure shows that the two communities in the networks are similar in composition.

Data Ethics

The data collection process has been carried out exclusively through the Twitter API, which is publicly available, and for the analysis we used publicly available data (users with privacy restrictions are not included in the dataset). We abided by the terms, conditions, and privacy policies of Twitter. Since this content is publicly published and is frequently discussed in mass media, we regard the debates as a public domain that does not require individual consent for inclusion in research, based on the ethical guidelines for internet research provided by The Association of Internet Researchers (Franzke et al. 2020) and by the British Sociological Association (Anon 2017). We only report on aggregates, and limit reporting on details of individuals to user accounts that belong to public figures or institutions, or that have more than 4,000 followers. The data published along with this research does not include user-ids nor the classification of the sentiment on the Black Pete discussion since this is part of a special category of personal data, formerly known as sensitive data.

CONCLUSIONS AND REFLECTIONS

This thesis set out to make the case for a computational social science (CSS) that is attentive to the intricacies of collective sensemaking. This orientation in CSS embraces the strengths of complexity-inspired methods and skills to examine and model measurable relational processes that give rise to complex social patterns, while also integrating the role of meaning-making that plays out within these social processes.

To make its case, the thesis presented four chapters with original case studies that incorporate the role of collective sensemaking into a computational network research design. The main objective of each of the four case studies was to explore the significance of attending to meaning-making by contrasting the study's conclusions with a more classical, primarily structural relational analysis.

Insights from the first two case studies

The first chapter examined the dissemination of Granovetter's concept of the Strength of Weak Ties, a seminal idea that originated in sociology but has since permeated various disciplines within the social sciences, becoming one of the most frequently cited papers to date. Similarly, the second chapter traced the movement of the intersectionality framework from critical race studies to a broad array of social sciences, highlighting its transformation and occasional contentiousness within academic discourse.

These case studies revealed that the spread of novel ideas is a dynamic process characterized by continuous reinterpretation and development within scholarly communities. The features in the diffusion network of both the Strength of Weak Ties and Intersectionality could not be fully explained by structural factors such as disciplinary boundaries, geographical considerations, or the influence of early adopters. What was left out from such structural analyses is the way scholars make sense of the novel idea. Analyzing the contents of publications, we found a significant relation between the community-features of the diffusion networks and the way scholars conceptualize, narrate and bring the novel idea into conversation with other ideas.

To illustrate, for the Strength of Weak Ties, we identified a large community made up by predominantly STEM researchers that consider the Strength of Weak Ties as a universal self-organizing principle of complex networks that is not specific to any social context and can only be understood by considering and modeling the network as a whole. Albert-László Barabási, a physicist interested in detecting and modeling the universal properties of complex networks, is the central figure of this community. One of the most referenced articles in this community is "Statistical Mechanics of Complex Networks" (Barabási and Albert, 2002). In contrast, another large community of scholars find strength in weak ties due to their ability to increase the relative status of individuals in society, conceptualizing weak ties as an asset to an individual person. This community is made up by primarily sociologists and some of its most-cited publications are devoted to measuring tie strength using questionnaires (Marsden and Campbell, 1984). Different communities use the same reference to convey very different points.

Moreover, these studies shed light on the pivotal role played by certain individuals within the translation and diffusion of novel theories. Communities seem to form around one scientific star or a few central figures that become a focal point for both the circulation and interpretation of the novel idea in the respective community. They are narrated to perform important translation work reinforcing their leading role. At times, the references to the original roots of the idea get lost, and these central figures become a local figurehead for the novel theory. To illustrate, a reference of intersectionality that does not credit Crenshaw (1989) or earlier roots:

> The related concept of intersectionality, which suggests that social categories and identities are not independent but rather multidimensional and linked to structural

> inequalities (Bowleg et al. 2013), provides a useful reference in understanding how layered stigma works.
>
> (Wilson et al., 2016)[17]

We conclude that the trajectories of idea propagation across scientific networks are closely aligned with the diverse interpretations and contextual applications adopted by researchers and suggest moving beyond a simplistic "social contagion" perspective, commonly embraced by complexity-inspired computational social scientists, towards a more nuanced understanding of idea dissemination informed by Latour's translation model and Collins' notion of collective meaning-making.

Insights from the second pair of case studies

The second pair of studies presented in this thesis focused on the discourse surrounding a highly polarized topic on social media. Specifically, they examined the communicative dynamics on Twitter (now named X) regarding the contentious Dutch tradition surrounding Black Pete, who some view as an innocent children's figure and others as a symbol of the country’s racist colonial legacy.

These studies deviate from the conventional approach adopted by complexity-inspired computational social scientists, which typically employs an echo-chamber lens to analyze online discourse characterized by positively reinforcing dynamics. Instead, the studies presented here revolve around a signed analysis of the debate, distinguishing between positive communication (indicative of users in agreement) and negative communication (reflective of users in active disagreement), thereby allowing for an exploration of the role of conflict in polarization. An unsigned tie in a network of individuals–such as those based on counts of retweets or mentions–can bring these individuals closer in the clustering of the social network into communities, or lead researchers to perceive

[17] This citation illustrates how the hub Bowleg is referenced alongside a definition of intersectionality, without citing Collins or Crenshaw. Our data showed the same phenomenon for other hubs.

this tie as serving to facilitate information flow. However, such a tie may instead be based on expressions of hostility from the sender to the receiver or function as a signal towards the sender's political ingroup to distance themselves from the receiver. In the two studies presented in this thesis, we compared the results from the conventional unsigned analysis (based on retweets or mentions) with those of the signed analysis to better understand the significance of the meaning of ties and the role of conflict in online polarization.

The empirical results presented in this thesis showed that there is a significant amount of negative interaction manifesting between groups in this debate online—an insight consistent with societal concerns about perceived online conflict. In the proponent-opponent study of the Black Pete debate (Chapter 4), we found that a large majority of intergroup relations are negative (81.6%). Intragroup relations, in contrast, are almost exclusively positive (96.6%). This signed perspective therefore offers a qualitatively different understanding from both supporters and critics of the echo chamber hypothesis: while ideological groups do exist on social media, they are shaped more by active conflict than by the exchange of information. Such groups are involved in showing support and solidarity within their ingroup while criticizing the outgroup.

By examining the coalitions and divisions within the network of this debate that extend beyond the simple categories of proponents and opponents (Chapter 3), we discovered additional nuances in the debate's social dynamics. For example, we found that several communities opposing Black Pete live in different spheres of the online debate, largely unaware of each other, whereas other communities supporting Black Pete explicitly refer to each other, both in positive and negative interactions. Our signed analysis thus allowed to capture more nuances that may prove important for intervention strategies or to inspire further research on the various social drivers and mechanism within polarization processes.

The research from chapters 3 and 4 also provided a detailed analysis of how negative interactions are distributed, revealing the diverse roles of debate participants, surpassing and complementing the classical broker and bridge roles. We recognized that some individuals were turned into scapegoats or opinion leaders, which is explained both by their structural network position as well as by the ways their expressions and influence get narrated by others within the polarizing social dynamics. We presented a model for identifying five roles in the debate: *leaders*, *authorities*, *scapegoats*, *positive mediators,* and *negative mediators*.

Overall, these two studies underscore the importance of considering the meaning and valence of social ties and interactions within online networks, particularly in the context of polarizing discourses. They highlight the nuanced dynamics at play in online debates, emphasizing the need for CSS to account for the diversity of roles that individuals assume in shaping collective narratives and alliances within digital spaces. Substantively, these studies highlight the dynamics of conflictual interactions in online polarized environments, particularly with regards to how individuals show support and solidarity within their ingroup while derogating the outgroup.

The main contribution of this thesis

Although the case studies in this thesis focus on specific contexts—academia, social media, and particular online platforms—and examine particular instances of ideas and practices, they reveal broader patterns of social interaction that shape the collective dynamics behind the formation and spread of beliefs, behaviors, and practices.

Building on interpretative scholarship, theories of the sociology of science, relational sociology and dynamics of conflict (Collins, 2012, 1998; Coser, 1957; Simmel 1904a; 1904b; 1904c), the findings of this thesis highlight a crucial distinction between human social dynamics and the dynamics of most other natural systems: our realities are consciously co-constructed with those around us, through processes of interpretation, alignment and confrontation. This thesis thus underscores the significance of collective meaning-making and conflict in shaping our sense of reality, actions and resulting social landscape. Social phenomena studied by complexity-inspired CSS, such as misinformation and polarization, are deeply rooted in disputes over truth and meaning, making them reliant on such collective meaning-making processes. While complexity-inspired CSS offers valuable insights into system-level dynamics, this thesis argues that integrating the dimension of collective meaning-making is essential to fully understand complex social processes.

The finding that meaning-making is an important explanatory and predictive dimension of social systems is not limited to the subjects of our studies; it also applies to us as computational social scientists. When we approach our research from perspectives that frame humans as particles, ants, or anonymous agents, and use methods consistent with these views, we are more likely to identify patterns and evidence that confirm these perspectives. Thus, the phenomenon of social embeddedness influencing perception extends to our own scientific practices and the frameworks we adopt.

It is my hope that this work will inspire complexity-inspired computational social scientists to recognize the importance of meaning so that we can make a meaningful difference in addressing the fundamental societal questions we face today.

Bibliography

Abbott, Andrew Delano. 2001. *Chaos of Disciplines*. University of Chicago Press.

Abbott, Andrew. 2001. "Time Matters: On Theory and Method." University of Chicago Press.

Abbott, David. 1988. "Transcending General Linear Reality." *Sociological Theory* 6(2):169–86.

Abdul Reda, Amir, Semuhi Sinanoglu, and Mohamed Abdalla. 2021. "Mobilizing the Masses: Measuring Resource Mobilization on Twitter." *Sociological Methods and Research*. doi: 10.1177/0049124120986197.

Adamic, Lada A., and Natalie Glance. 2005. "The Political Blogosphere and the 2004 US Election: Divided They Blog." *Proceedings of the 3rd International Workshop on Link Discovery*. 36–43.

Adler, Paul S. and Seok-Woo Kwon. 2002. "Social Capital: Prospects for a New Concept." *Academy of Management Review* 27:17–40.

Albert, Réka and Albert-László Barabási. 2002. "Statistical Mechanics of Complex Networks." *Reviews of Modern Physics* 74:47–97.

Anon. 2017. *BSA Statement of Ethical Practice*.

Anthias, Floya, and Nira Yuval-Davis. 1983. "Contextualizing Feminism: Gender, Ethnic and Class Divisions." *Feminist Review* 15(1):62–75.

Artime, Oriol, and Manlio De Domenico. 2022. "From the Origin of Life to Pandemics: Emergent Phenomena in Complex Systems." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 380(2227). doi: 10.1098/rsta.2020.0410.

Axelrod, Robert. 1997. "The Dissemination of Culture : A Model with Local Convergence and Global Polarization." 41(2):203–26.

Bail, Christopher A. 2014. "The Cultural Environment: Measuring Culture with Big Data." *Theory and Society* 43(3–4):465–82.

Bail, Christopher A. 2021. *Breaking the Social Media Prism: How to Make Our Platforms Less Polarizing*. Princeton: Princeton University Press.

Bail, Christopher A., Lisa P. Argyle, Taylor W. Brown, John P. Bumpus, Haohan Chen, M. B. Fallin Hunzaker, Jaemin Lee, Marcus Mann, Friedolin Merhout, and Alexander Volfovsky. 2018a. "Exposure to

Opposing Views on Social Media Can Increase Political Polarization." *Proceedings of the National Academy of Sciences of the United States of America* 115(37):9216–21. doi: 10.1073/pnas.1804840115.

Bail, Christopher Andrew. 2016. "Combining Natural Language Processing and Network Analysis to Examine How Advocacy Organizations Stimulate Conversation on Social Media." *Proceedings of the National Academy of Sciences of the United States of America* 113(42):11823–28. doi: 10.1073/pnas.1607151113.

Bakshy, Eytan, Solomon Messing, and Lada A. Adamic. 2015. "Exposure to Ideologically Diverse News and Opinion on Facebook." *Science* 348(6239):1130–32. doi: 10.1126/science.aaa1160.

Ball, Philip. 2004. *Critical Mass: How One Things Leads Into Anoth*er. Farrar, Straus and Giroux.

Ball, Philip. 2012. "Why Society Is a Complex Matter: Meeting Twenty-First Century Challenges with a New Kind of Science." Springer Science & Business Media.

Barabási, Albert László, and Réka Albert. 1999. "Emergence of Scaling in Random Networks." *Science* 286(5439):509–12. doi: 10.1126/science.286.5439.509.

Barabási, Albert László, Réka Albert, and Hawoong Jeong. 2000. "Scale-Free Characteristics of Random Networks: The Topology of the World-Wide Web." *Physica A: Statistical Mechanics and Its Applications* 281(1):69–77. doi: 10.1016/S0378-4371(00)00018-2.

Barabási, Albert-László, and Réka Albert. 1999. "Emergence of Scaling in Random Networks." *Science* 286(5439):509–12.

Barabási, Albert-László, Hawoong Jeong, Zoltan Néda, Erzsebet Ravasz, Andras Schubert, and Tamas Vicsek. 2002. "Evolution of the Social Network of Scientific Collaborations." *Physica A: Statistical Mechanics and its Applications* 311:590–614.

Barberá, Pablo, John T. Jost, Jonathan Nagler, Joshua A. Tucker, and Richard Bonneau. 2015a. "Tweeting From Left to Right: Is Online Political Communication More Than an Echo Chamber?" *Psychological Science* 26(10):1531–42. doi: 10.1177/0956797615594620.

Barberá, Pablo. 2015b. "Birds of the Same Feather Tweet Together: Bayesian Ideal Point Estimation Using Twitter Data." *Political Analysis* 23(1):76–91. doi: 10.1093/pan/mpu011.

Barberá, Pablo. 2020. “Social Media, Echo Chambers , and Political Polarization.” P. 34 in *Social Media and Democracy: The State of the Field, Prospects and Reform.*

Berger, Peter, and Thomas Luckmann. 1966. *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. Harlow, England: Penguin Books.

Bettencourt, Luís M. A., David I. Kaiser, Jasleen Kaur, Carlos Castillo-Chávez, and David E. Wojick. 2008. “Population Modeling of the Emergence and Development of Scientific Fields.” *Scientometrics* 75:495–518.

Bhaskar, Roy. 1978. “On the Possibility of Social Scientific Knowledge and the Limits of Naturalism.” *Journal for the Theory of Social Behaviour* 8(1):1–28. doi: 10.1111/j.1468-5914.1978.tb00389.x.

Bilge, Sirma. 2013. “Intersectionality Undone: Saving Intersectionality from Feminist Intersectionality Studies.” *Du Bois Review* 10(2):402–424. doi: 10.1017/S1742058X13000283.

Bishop, Bill. 2009. *The Big Sort: Why the Clustering of like-Minded America Is Tearing Us Apart*.

Blei, David M., Andrew Y. Ng, and Michael I. Jordan. 2003. “Latent Dirichlet Allocation.” *Journal of Machine Learning Research* 3(1):993–1022.

Blondel, Vincent D., Jean Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. 2008. “Fast Unfolding of Communities in Large Networks.” *Journal of Statistical Mechanics: Theory and Experiment* p. P10008.

Bojanowski, Piotr, Edouard Grave, Armand Joulin, and Tomas Mikolov. 2017. “Enriching Word Vectors with Subword Information.” *Transactions of the Association for Computational Linguistics* 5:135–46. doi: 10.1162/tacl_a_00051.

Bolukbasi, Tolga, Kai-wei Chang, James Zou, Venkatesh Saligrama, and Adam Kalai. 2016. “Man Is to Computer Programmer as Woman Is to Homemaker? Debiasing Word Embeddings.” *Advances in Neural Information Processing Systems 2*.

Bonifati, Giovanni. 2010. “‘More is Different’, Exaptation and Uncertainty: Three Founda- tional Concepts for a Complexity Theory of Innovation.” *Economics of Innovation and New Technology* 19:743–760.

Borgatti, Stephen P and Pacey C Foster. 2003. “The Network Paradigm in Organizational Research: A Review and Typology.” *Journal of management* 29:991–1013.

Borge-Holthoefer, Javier, Alejandro Rivero, Iñigo García, Elisa Cauhé, Alfredo Ferrer, Darío Ferrer, David Francos, David Iñiguez, María Pilar Pérez, Gonzalo Ruiz, Francisco Sanz, Fermín Serrano, Cristina Viñas, Alfonso Tarancón, and Yamir Moreno. 2011. “Structural and Dynamical Patterns on Online Social Networks: The Spanish May 15th Movement as a Case Study.” *PLoS ONE* 6(8):1–8. doi: 10.1371/journal.pone.0023883.

Borra, Erik, and Bernhard Rieder. 2014. “Programmed Method: Developing a Toolset for Capturing and Analyzing Tweets.” *Aslib Journal of Information Management* 66(3):262–78. doi: 10.1108/AJIM-09-2013-0094.

Boutet, Antoine, Hyoungshick Kim, and Eiko Yoneki. 2012. “What’s in Your Tweets? I Know Who You Supported in the UK 2010 General Election.” *ICWSM 2012 - Proceedings of the 6th International AAAI Conference on Weblogs and Social Media* 6(1):411–14.

Bowleg, Lisa, Michelle Teti, David J. Malebranche, and Jeanne M. Tschann. 2013. “‘It’s an Uphill Battle Everyday’: Intersectionality, Low-Income Black Heterosexual Men, and Implications for HIV Prevention Research and Interventions.” *Psychology of Men and Masculinity* 14(1):25–34. doi: 10.1037/a0028392.

Bowleg, Lisa. 2008. “When Black + Lesbian + Woman ≠ Black Lesbian Woman: The Methodological Challenges of Qualitative and Quantitative Intersectionality Research.” *Sex Roles* 59(5):312–325. doi: 10.1007/s11199-008-9400-z.

Brah, Avtar, and Ann Phoenix. 2004. “Ain’t I A Woman? Revisiting Intersectionality.” *Journal of International Women’s Studies* 5(3):75–86.

Brass, Daniel J., Joseph Galaskiewicz, Henrich R. Greve, and Wenpin Tsai. 2004. “Taking Stock of Networks and Organizations: A Multilevel Perspective.” *Academy of Management Journal* 47:795–817.

Breiger, Ronald L. 2010. “Dualities of Cultuer and Structure, s Eeing through Cultural Holes.” *Handbuch Netzwerkforschung* 63–87. doi: 10.1007/978-3-531-92575-2_7.

Breiger, Ronald L., Scott A. Boorman, and Phipps Arabie. 1975. “An Algorithm for Cluster- ing Relational Data with Applications to Social

Network Analysis and Comparison with Multidimensional Scaling." *Journal of Mathematical Psychology* 12:328–383.

Bruns, Axel, and Stefan Stieglitz. 2014. "Twitter Data: What Do They Represent?" *It - Information Technology* 56(5). doi: 10.1515/itit-2014-1049.

Bruns, Axel. 2017. "Echo Chamber? What Echo Chamber? Reviewing the Evidence." *6th Biennial Future of Journalism Conference (FOJ17)*.

Bruns, Axel. 2019. *Are Filter Bubbles Real?* John Wiley & Sons.

Bruns, Axel. 2021. "Echo Chambers? Filter Bubbles? The Misleading Metaphors That Obscure the Real Problem." Pp. 33–48 in *Hate Speech and Polarization in Participatory Society*. Routledge.

Burt, Ronald S. 1984. "Network Items and the General Social Survey." *Social Networks* 6:293–339.

Burt, Ronald S. 1992. "The Social Structure of Competition." *Social Stratification* 659–63. doi: 10.4159/9780674029095-003.

Burt, Ronald S. 1997. "A Note on Social Capital and Network Content." *Social Networks* 19:355–373.

Burt, Ronald S. 2000. "The Network Structure of Social Capital." *Research in Organizational Behavior* 22:345–423.

Burt, Ronald S. 2004. "Structural Holes and Good Ideas." *American Journal of Sociology* 110:349–339.

Burt, Ronald S. 2004a. "Brokerage and Closure An Introduction to Social Capital."

Burt, Ronald S. 2004b. "Structural Holes and Good Ideas." *American Journal of Sociology* 110(2):349–339.

Byrne, David. 1988. *Complexity Theory and the Social Sciences: An Introduction*. New York: Psychology Press.

Campbell, Karen E., Peter V. Marsden, and Jeanne S. Hurlbert. 1986. "Social Resources and Socioeconomic Status." *Social Networks* 8:97–117.

Cantarella, G. E., S. De Luca, M. Di Gangi, R. Di Pace, and S. Memoli. 2014. "Macroscopic vs. Mesoscopic Traffic Flow Models in Signal Setting Design." *2014 17th IEEE International Conference on Intelligent Transportation Systems, ITSC 2014* 73(October):2221–26. doi: 10.1109/ITSC.2014.6958032.

Carbado, Devon W., Kimberlé Williams Crenshaw, Vickie M. Mays, and Barbara Tomlinson. 2013. “Intersectionality: Mapping the Movements of a Theory.” *Du Bois Review* 10(2):303–312. doi: 10.1017/S1742058X13000349.

Carley, Kathleen M., and David S. Kaufer. 1993. “Semantic Connectivity: An Approach for Analyzing Symbols in Semantic Networks.” *Communication Theory* 3(3):183–213.

Carley, Kathleen, and Michael Palmquist. 1992. “Extracting, Representing, and Analyzing Mental Models.” *Social Forces*.

Carley, Kathleen. 1990. “Structural Constraints on Communication: The Diffusion of the Homomorphic Signal Analysis Technique through Scientific Fields.” *The Journal of Mathematical Sociology* 15(3–4):207–46. doi: 10.1080/0022250X.1990.9990070.

Carothers, Thomas, and Andrew O’Donohue. 2019. *Democracies Divided: The Global Challenge of Political Polarization*. Brookings Institution Press.

Cartwright, Dorwin, and Frank Harary. 1956. “Structural Balance: A Generalization of Heider’s Theory.” *Psychological Review* 63(5):277–93. doi: 10.1037/H0046049.

Centola, Damon and Michael Macy. 2007. “Complex Contagions and the Weakness of Long Ties.” *American Journal of Sociology* 113:702–734.

Centola, Damon. 2010. “The Spread of Behavior In an Online Social Network Experiment.” *Science* 329:1194–1197.

Centola, Damon. 2015. “The Social Origins of Networks and Diffusion.” *American Journal of Sociology* 120:1295–1338.

Chauvin, Sébastien, Yannick Coenders, and Timo Koren. 2018. “Never Having Been Racist: Explaining the Blackness of Blackface in the Netherlands.” *Public Culture* 30(3):509–26.

Cho, Sumi, Kimberlé Williams Crenshaw, and Leslie McCall. 2013. “Toward a Field of Intersectionality Studies: Theory, Applications, and Praxis.” *Signs: Journal of Women in Culture and Society* 38(4):785–810. doi: 10.1086/669608.

Choo, Hae Yeon, and Myra Marx Ferree. 2010. “Practicing Intersectionality in Sociological Research: A Critical Analysis of Inclusions, Interactions, and Institutions in the Study of Inequalities.” *Sociological Theory* 28(2):129–149.

Choudhury, Tanzeem, and Alex Pentland. 2003. “Sensing and Modeling Human Networks Using the Sociometer.” *Proceedings - International Symposium on Wearable Computers, ISWC* 216–22. doi: 10.1109/iswc.2003.1241414.

Christakis, Nicholas A. and James H. Fowler. 2007. “The Spread of Obesity in a Large Social Network over 32 Years.” *New England Journal of Medicine* 357:370–379.

Cioffi-Revilla, Claudio. 2010. “Computational Social Science.” *Wiley Interdisciplinary Reviews: Computational Statistics* 2(3):259–71. doi: 10.1002/wics.95.

Cioffi-Revilla, Claudio. 2017. *Introduction to Computational Social Science: Principles and Applications*.

Coenders, Yannick, and Sébastien Chauvin. 2017. “Race and the Pitfalls of Emotional Democracy: Primary Schools and the Critique of Black Pete in the Netherlands.” *Antipode* 49(5):1244–62. doi: 10.1111/anti.12328.

Cohen, Eliot A., and Albert-László Barabási. 2002. “Linked: The New Science of Networks.” *Foreign Affairs* 81(5):204. doi: 10.2307/20033300.

Cole, Elizabeth R. 2009. “Intersectionality and Research in Psychology.” *American Psychologist* 64(3):170–180. doi: 10.1037/a0014564.

Cole, Stephen. 1970. “Professional Standing and the Reception of Scientific Discoveries.” *American Journal of Sociology* 76:286–306.

Coleman, James Samuel. 1964. *Introduction to Mathematical Sociology*. UK: Collier-Macmillan.

Coleman, James, Elihu Katz, and Herbert Menzel. 1957. “The Diffusion of an Innovation Among Physicians.” *Sociometry* 20(4):253. doi: 10.2307/2785979.

Colleoni, Elanor, Alessandro Rozza, and Adam Arvidsson. 2014. “Echo Chamber or Public Sphere? Predicting Political Orientation and Measuring Political Homophily in Twitter Using Big Data.” *Journal of Communication* 64(2):317–32. doi: 10.1111/jcom.12084.

Collier, Andrew. 1994. *Critical Realism: An Introduction to Roy Bhaskar’s Philosophy*. London: Verso.

Collins, Harry M. 1983. “The Sociology of Scientific Knowledge: Studies of Contemporary Science.” *Annual Review of Sociology* 9:265–285.

Collins, Patricia Hill, and Sirma Bilge. 2016. *Intersectionality*. Cambridge: Polity Press.

Collins, Patricia Hill, and Valerie Chepp. 2013. "Intersectionality." Pp. 1–32 in *Intersectionality*, edited by G. Waylen, K. Celis, J. Kantola, and S. L. Weldon. New York, NY: Oxford University Press.

Collins, Patricia Hill. 1990. *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*. New York, London: Routledge.

Collins, Patricia Hill. 1993. "Toward a New Vision: Race, Class, and Gender as Categories of Analysis and Connection." *Race, Sex & Class* 1(1):25–45.

Collins, Randall. 1998. *The Sociology of Philosophies : A Global Theory of Intellectual Change*. Cambridge, MA: Harvard University Press.

Collins, Randall. 2004. *Interaction Ritual Chains*. Princeton University Press.

Collins, Randall. 2012. "C-Escalation and d-Escalation: A Theory of the Time-Dynamics of Conflict." *American Sociological Review* 77(1):1–20. doi: 10.1177/0003122411428221.

Combahee River Collective. 1977. *The Combahee River Collective Statement*.

Conover, Michael D., Bruno Gonçalves, Jacob Ratkiewicz, Alessandro Flammini, and Filippo Menczer. 2011a. "Predicting the Political Alignment of Twitter Users." Pp. 192–99 in *Proceedings - 2011 IEEE International Conference on Privacy, Security, Risk and Trust and IEEE International Conference on Social Computing, PASSAT/SocialCom 2011*. Boston, MA, USA.

Conover, Michael D., Jacob Ratkiewicz, M. Fransisco, B. Concalves, Alessandro Flammini, and Filippo Menczer. 2011a. "Political Polarization on Twitter." in *Proceedings of the 5th International AAAI Conference on Weblogs and Social Media*.

Conte, R., N. Gilbert, G. Bonelli, C. Cioffi-Revilla, G. Deffuant, J. Kertesz, V. Loreto, S. Moat, J. P. Nadal, A. Sanchez, A. Nowak, A. Flache, M. San Miguel, and D. Helbing. 2012a. "Manifesto of Computational Social Science." *European Physical Journal: Special Topics* 214(1):325–46. doi: 10.1140/epjst/e2012-01697-8.

Conwill, William Louis. 2010. “Domestic Violence among the Black Poor: Intersectionality and Social Justice.” *Journal for the Advancement of Counselling* 32(1):31–45.

Coser, Lewis A. 1957. “Social Conflict and the Theory of Social Change.” *The British Journal of Sociology* 8(3):197–207.

Couldry, Nick, and Andreas Hepp. 2018. *The Mediated Construction of Reality*. John Wiley & Sons.

Crane, Diana. 1972. *Invisible Colleges: Diffusion of Knowledge in Scientific Communities*. Chicago: University of Chicago Press.

Crenshaw, Kimberle. 1989. “Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics.” *University of Chicago Legal Forum* 1989(1):139–167.

Crenshaw, Kimberle. 1991. “Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color.” *Stanford Law Review* 43(6):1241–1299.

D’Costa, Krystal. 2017. “A Nation Divided by Social Media.” *Scientific American Blog Network*. Retrieved June 29, 2022.

Dahlander, Linus and Daniel A. McFarland. 2013. “Ties that Last: Tie Formation and Persis- tence in Research Collaborations over Time.” *Administrative Science Quarterly* 58:69–110.

Davis, Kathy. 2008. “Intersectionality as Buzzword: A Sociology of Science Perspective on What Makes a Feminist Theory Successful.” *Feminist Theory* 9(1):67–85. doi: 10.1177/1464700108086364.

Davis, Murray S. 1971. “That’s Interesting! Towards a Phenomenology of Sociology and a Sociology of Phenomenology.” *Philosophy of the Social Sciences* 1(2):309–44.

De Stefano, Domenico, and Francesco Santelli. 2019. “Combining Sentiment Analysis and Social Network Analysis to Explore Twitter Opinion Spreading.” Pp. 1–6 in *Proceedings - 2019 28th International Conference on Computer Communications and Networks, ICCCN*. Vols. 2019-July. Valencia, Spain: IEEE.

Del Vicario, Michela, Gianna Vivaldo, Alessandro Bessi, Fabiana Zollo, Antonio Scala, Guido Caldarelli, and Walter Quattrociocchi. 2016. “Echo Chambers: Emotional Contagion and Group Polarization on Facebook.” *Scientific Reports* 6(37825):1–12. doi: 10.1038/srep37825.

Dill, Bonnie Thornton, and Ruth E. Zambrana. 2009. *Emerging Intersections: Race, Class, and Gender in Theory, Policy, and Practice*. New Brunswick, NJ: Rutgers University Press.

DiMaggio, Paul, Manish Nag, and David Blei. 2013. "Exploiting Affinities Between Topic Modeling and the Sociological Perspective on Culture: Application to Newspaper Coverage of U.S. Government Arts Funding." *Poetics* 41:570–606.

Dirk Helbing, Illés Farkas, and Tamás Vicsek. 2000. "Simulating Dynamical Features of Escape Panic." *Nature* 407(July):487–90.

Eady, Gregory, Jonathan Nagler, Andy Guess, Jan Zilinsky, and Joshua A. Tucker. 2019. "How Many People Live in Political Bubbles on Social Media? Evidence from Linked Survey and Twitter Data." *SAGE Open* 9(1):1–21. doi: 10.1177/2158244019832705.

Edelman, Joe. 2024. "Meaning Alignment Institute." Retrieved May 5, 2024 (https://www.meaningalignment.org/)

Edelmann, Achim, Tom Wolff, Danielle Montagne, and Christopher A. Bail. 2020. "Computational Social Science and Sociology." *Annual Review of Sociology* 46:61–81. doi: 10.1146/annurev-soc-121919-054621.

El-Bermawy, Mostafa M. 2016. "Your Filter Bubble Is Destroying Democracy." *WIRED*, November 18.

Emirbayer, Mustafa and Jeff Goodwin. 1994. "Network Analysis, Culture , and the Problem of Agency." *American Journal of Sociology* 99:1411–1454.

Emirbayer, Mustafa. 1997. "Manifesto for a Relational Sociology." *American Journal of Sociology* 103(2):281–317. doi: 10.1086/231209.

Esteve del Valle, Marc, and Rosa Borge Bravo. 2018. "Echo Chambers in Parliamentary Twitter Networks: The Catalan Case." *International Journal of Communication* 12(2018):1715–35.

Evolvi, Giulia. 2019. "#Islamexit: Inter-Group Antagonism on Twitter." *Information, Communication and Society* 22(3):386–401. doi: 10.1080/1369118X.2017.1388427.

Feller, Albert, Matthias Kuhnert, To Sprenger, and Im Welpe. 2011. "Divided They Tweet: The Network Structure of Political Microbloggers and Discussion Topics." Pp. 474–77 in *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*. Barcelona, Catalonia, Spain.

Ferrara, Emilio, Stefano Cresci, and Luca Luceri. 2020. "Misinformation, Manipulation, and Abuse on Social Media in the Era of COVID-19." *Journal of Computational Social Science* 3(2):271–77. doi: 10.1007/s42001-020-00094-5.

Fisher, Walter R. 1987. *Human Communication as Narration: Toward a Philosophy of Reason, Value, and Action*. Columbia: University of South Carolina Press.

Flache, Andreas and Michael W Macy. 1996. "The Weakness of Strong Ties: Collective Action Failure in a Highly Cohesive Group." *Journal of Mathematical Sociology* 21:3–28.

Fortunato, Santo, Carl T. Bergstrom, Katy Börner, James A. Evans, Dirk Helbing, Staša Milojević, Alexander M. Petersen, Filippo Radicchi, Roberta Sinatra, Brian Uzzi, Alessan- dro Vespignani, Ludo Waltman, Dashun Wang, and Albert-László Barabási. 2018. "Science of Science." *Science* 359:eaao0185.

Foster, Jacob G., Andrey Rzhetsky, and James A. Evans. 2015. "Tradition and Innovation in Scientists' Research Strategies." *American Sociological Review* 80:875–908.

Franzke, aline shakti, Anja Bechmann, Michael Zimmer, Charles Ess, and The Association of Internet. 2020. *Internet Research: Ethical Guidelines 3.0*.

Freelon, Deen. 2014. "On the Interpretation of Digital Trace Data in Communication and Social Computing Research." *Journal of Broadcasting and Electronic Media* 58(1):59–75. doi: 10.1080/08838151.2013.875018.

Friedman, Debra, and Doug McAdam. 1992. "Collective Identity and Activism: Networks, Choices, and the Life of a Social Movement." *Frontiers in Social Movement Theory*.

Fuchs, Stephan. 2001. *Against Essentialism: A Theory of Culture and Society*. Harvard University Press.

Fuhse, Jan A. 2009. "The Meaning Structure of Social Networks." *Sociological Theory* 27(1):51–73. doi: 10.1111/j.1467-9558.2009.00338.x.

Fuhse, Jan A. 2015. "Culture and Social Networks." *Emerging Trends in the Social and Behavioral Sciences* 1–15. doi: 10.1002/9781118900772.ETRDS0066.

Fuhse, Jan A. 2015. “Theorizing Social Networks: The Relational Sociology of and around Harrison White.” *International Review of Sociology* 25:15–44.

Fuhse, Jan A. 2020. “The Field of Relational Sociology.” *Digithum* 2020(26):1–10. doi: 10.7238/d.v0i26.374145.

Gaisbauer, Felix, Armin Pournaki, Sven Banisch, and Eckehard Olbrich. 2021. “Ideological Differences in Engagement in Public Debate on Twitter.” *PLoS ONE* 16(3 March):1–18. doi: 10.1371/journal.pone.0249241.

Garimella, Kiran, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. 2018. “Political Discourse on Social Media: Echo Chambers, Gatekeepers, and the Price of Bipartisanship.” Pp. 913–22 in *The Web Conference 2018 - Proceedings of the World Wide Web Conference, WWW 2018*. Association for Computing Machinery, Inc.

Garrett, Kelly R. 2009. “Politically Motivated Reinforcement Seeking: Reframing the Selective Exposure Debate.” *Journal of Communication* 59(4):676–99. doi: 10.1111/J.1460-2466.2009.01452.X.

Gilbert, G. Nigel, and Michael Mulkay. 1985. “Opening Pandora’s Box: A Sociological Analysis of Scientists’ Discourse.” *Contemporary Sociology* 14(1):98. doi: 10.2307/2070471.

Girvan, M., and M. E. J. Newman. 2006. “Modularity and Community Structure in Networks.” *Proceedings of the National Academy of Sciences of the United States of America* 99(12):7821–26. doi: 10.1073/pnas.122653799.

Goel, Sharad, Winter Mason, and Duncan J. Watts. 2010. “Real and Perceived Attitude Agreement in Social Networks.” *Journal of Personality and Social Psychology* 99(4):611–621. doi: 10.1037/a0020697.

Goffman, Erving. 1956. *The Presentation of Everyday Life*. Garden City, NY: Doubleday Anchor Books.

Goffman, Erwin. 1967. “Interaction Ritual: Essays in Face-to-Face Behavio.” Anchor Books.

Goffman, William and Vaun A. Newill. 1964. “Generalization of Epidemic Theory: An Application to the Transmission of Ideas.” *Nature* 204:225–228.

Goffman, William. 1966. “Mathematical Approach to the Spread of Scientific Ideas: The History of Mast Cell Research.” *Nature* 212:449–452.

Goldberg, Amir and Sarah K. Stein. 2018. “Beyond Social Contagion: Associative Diffusion and the Emergence of Cultural Variation.” *American Sociological Review* 83:897–932.

Goldberg, Amir, and Madison H. Singell. 2023. “The Sociology of Interpretation.” *Annual Review of Sociology* 1–24.

Goldberg, Amir, and Sarah K. Stein. 2018. “Beyond Social Contagion: Associative Diffusion and the Emergence of Cultural Variation.” *American Sociological Review* 83(5):897–932. doi: 10.1177/0003122418797576.

Golder, Scott A., and Michael W. Macy. 2014. “Digital Footprints: Opportunities and Challenges for Online Social Research.” *Annual Review of Sociology* 40:129–52. doi: 10.1146/annurev-soc-071913-043145.

González-Bailón, Sandra, Javier Borge-Holthoefer, and Yamir Moreno. 2013. “Broadcasters and Hidden Influentials in Online Protest Diffusion.” *American Behavioral Scientist* 57(7):943–65. doi: 10.1177/0002764213479371.

Goodman, Douglas, Source Sociological Theory, and No Mar. 2009. “What Collins’s The Sociology of Philosophies Says about Sociological Theory. *Sociological Theory* 19(1):92–101.

Granovetter, Mark S. 1973. “The Strength of Weak Ties.” *American Journal of Sociology* 78(6):1360–80. doi: 10.1086/225469.

Granovetter, Mark. 1985. “Economic Action and Social Structure: The Problem of Embed- dedness.” *American Journal of Sociology* 91:481–510.

Granovetter, Mark. 1992. “Economic Institutions as Social Constructions: A Framework for Analysis.” *Acta Sociologica* 35(1):3–11. doi: 10.1177/000169939203500101.

Granovetter. 1978. “Threshold Models of Collective Behavior.” *American Journal of Sociology* 83(6):1420–43.

Greenhalgh, Trisha, Glenn Robert, Fraser MacFarlane, Paul Bate, Olympia Kyriakidou, and Richard Peacock. 2005. “Storylines of Research in Diffusion of Innovation: A Meta- Narrative Approach to Systematic Review.” *Social Science and Medicine* 61:417–430.

Griffith, Belver C. 1989. “Understanding Science.” *Communication Research* 16:600–614.

Griffith, Belver C. and Nicholas C. Mullins. 1972. “Coherent Social Groups in Scientific Change.” *Science* 177:959–964.

Grossmann, Matt, and David A. Hopkins. 2016. *Asymmetric Politics: Ideological Republicans and Group Interest Democrats*. Oxford, UK: Oxford University Press.

Gruzd, Anatoliy, and Jeffrey Roy. 2014. “Investigating Political Polarization on Twitter: A Canadian Perspective.” *Policy and Internet* 6(1):28–45. doi: 10.1002/1944-2866.POI354.

Guerrero-Solé, Frederic. 2017. “Community Detection in Political Discussions on Twitter: An Application of the Retweet Overlap Network Method to the Catalan Process Toward Independence.” *Social Science Computer Review* 35(2):244–61. doi: 10.1177/0894439315617254.

Guess, Andrew, Benjamin Lyons, Brendan Nyhan, and Jason Reifler. 2018. *Avoiding the Echo Chamber about Echo Chambers: Why Selective Exposure to like-Minded Political News Is Less Prevalent than You Think*.

Guha, R., Prabhakar Raghavan, Ravi Kumar, and Andrew Tomkins. 2004a. “Propagation of Trust and Distrust.” Pp. 403–12 in *WWW ’04: Proceedings of the 13th international conference on World Wide Web*. New York, New York, USA: Association for Computing Machinery.

Hakimi, S. Louis. 1962. “On Realizability of a Set of Integers as Degrees of the Vertices of a Linear Graph. I.” *Journal of the Society for Industrial and Applied Mathematics* 10:496–506.

Hancock, Ange Marie. 2007. “When Multiplication Doesn’t Equal Quick Addition: Examining Intersectionality as a Research Paradigm.” *Perspectives on Politics* 5(1):63–79. doi: 10.1017/S1537592707070065.

Hancock, Ange-Marie. 2016. *Intersectionality: An Intellectual History*. New York: Oxford University Press.

Hankivsky, Olena, Colleen Reid, Renee Cormier, Colleen Varcoe, Natalie Clark, Cecilia Benoit, and Shari Brotman. 2010. “Exploring the Promises of Intersectionality for Advancing Women’s Health Research.” *International Journal for Equity in Health* 9(1):5–19. doi: 10.1186/1475-9276-9-5.

Harrigan, Nicholas M., Giuseppe (Joe) Labianca, and Filip Agneessens. 2020. "Negative Ties and Signed Graphs Research: Stimulating Research on Dissociative Forces in Social Networks." *Social Networks* 60:1–10. doi: 10.1016/j.socnet.2019.09.004.

Harris, Angela P. 1990. "Race and Essentialism in Feminist Legal Theory." *Stanford Law Review* 42(3):581–616.

Harris, Angela. 2001. "Critical Race Theory: The Cutting Edge (Foreword)." edited by R. Delgado and J. Stefancic. Philadelphia, Pennsylvania: Temple University Press.

Harteveld, Eelco. 2021. "Fragmented Foes: Affective Polarization in the Multiparty Context of the Netherlands." *Electoral Studies* 71:102332. doi: 10.1016/J.ELECTSTUD.2021.102332.

Hassan, Ahmed, Amjad Abu-Jbara, and Dragomir Radev. 2012a. "Detecting Subgroups in Online Discussions by Modeling Positive and Negative Relations Among Participants." Pp. 59–70 in *EMNLP-CoNLL '12: Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*. Jeju, Republic of Korea: Association for Computational Linguistics.

Hassan, Ahmed, Amjad Abu-Jbara, and Dragomir Radev. 2012b. "Extracting Signed Social Networks From Text." *Proceedings of the TextGraphs-7 Workshop at ACL* (July):6–14.

Heider, Fritz. 1946. "Attitudes and Cognitive Organization." *The Journal of Psychology* 21(1):107–12. doi: 10.1080/00223980.1946.9917275.

Heider, Fritz. 1946b. "Attitudes and Cognitive Organization." *The Journal of Psychology: Interdisciplinary and Applied* 21(1):107–12. doi: 10.1080/00223980.1946.9917275.

Helsloot, John. 2009. "Is Zwarte Piet Uit Te Leggen ?" *Sinterklaas Verklaard* (november 2005):76–85.

Helsloot, John. 2012. "Zwarte Piet and Cultural Aphasia in the Netherlands." *Quotidian. Journal for the Study of Everyday Life* 3.

Hilhorst, Sacha, and Joke Hermes. 2016. "'We Have given up so Much': Passion and Denial in the Dutch Zwarte Piet (Black Pete) Controversy." *European Journal of Cultural Studies* 19(3):218–33. doi: 10.1177/1367549415603381.

Himelboim, Itai, Kaye D. Sweetser, Spencer F. Tinkham, Kristen Cameron, Matthew Danelo, and Kate West. 2016. “Valence-Based Homophily on Twitter: Network Analysis of Emotions and Political Talk in the 2012 Presidential Election.” *New Media and Society* 18(7):1382–1400. doi: 10.1177/1461444814555096.

Hobolt, Sara B., Thomas J. Leeper, and James Tilley. 2021. “Divided by the Vote: Affective Polarization in the Wake of the Brexit Referendum.” *British Journal of Political Science* 51(4):1476–93. doi: 10.1017/S0007123420000125.

Hofman, Jake M., Duncan J. Watts, Susan Athey, Filiz Garip, Thomas L. Griffiths, Jon Kleinberg, Helen Margetts, Sendhil Mullainathan, Matthew J. Salganik, Simine Vazire, Alessandro Vespignani, and Tal Yarkoni. 2021. “Integrating Explanation and Prediction in Computational Social Science.” *Nature* 595(7866):181–88. doi: 10.1038/s41586-021-03659-0.

Holton, Gerard. 1962. “Scientific Research and Scholarship Notes toward the Design of Proper Scales.” *Science and Technology in Contemporary Society* 91:362–399.

Honeycutt, Courtenay, and Susan C. Herring. 2009. “Beyond Microblogging: Conversation and Collaboration via Twitter.” Pp. 1–10 in *Proceedings of the 42nd Annual Hawaii International Conference on System Sciences, HICSS*.

hooks, bell. 1984. *Feminist Theory: From Margin to Center*. New York and London: Routledge.

Hooton, Christopher. 2016. “Social Media Echo Chambers Gifted Donald Trump the Presidency | The Independent | The Independent.” *Independent*, November 10.

Iyengar, Shanto, and Kyu S. Hahn. 2009. “Red Media, Blue Media: Evidence of Ideological Selectivity in Media Use.” *Journal of Communication* 59:19–39. doi: 10.1111/J.1460-2466.2008.01402.X.

Jaeckal, Martin. 1971. “Coleman’s Process Approach.” *Sociological Methodology* 3:236–75.

Jordan-Zachery, Julia S. 2007. “Am I a Black Woman or a Woman Who Is Black? A Few Thoughts on the Meaning of Intersectionality.” *Politics and Gender* 3(2):254–63. doi: 10.1017/S1743923X07000074.

Joulin, Armand, Edouard Grave, Piotr Bojanowski, and Tomas Mikolov. 2017. “Bag of Tricks for Efficient Text Classification.” *Proceedings of the*

15th Conference of the European Chapter of the Association for Computational Linguistics 2:427–31. doi: 10.18653/v1/e17-2068.

Kaiser, David I. 2009. *Drawing Theories Apart: The Dispersion of Feynman Diagrams in Postwar Physics*. University of Chicago Press.

Karsai, Márton, Mikko Kivelä, Raj Kumar Pan, Kimmo Kaski, János Kertész, Albert-László Barabási, and Jari Saramäki. 2011. "Small but Slow World: How Network Topology and Burstiness Slow Down Spreading." *Physical Review E* 83:25102.

Katz, Elihu. 1957. "The Two-Step Flow of Communication: An Up-To-Date Report on an Hypothesis." *Public Opinion Quarterly* 21(1):61–78. doi: 10.1086/266687.

Kiss, Istvan Z., Mark Broom, Paul G. Craze, and Ismael Rafols. 2010. "Can Epidemic Models Describe the Diffusion of Topics across Disciplines?" *Journal of Informetrics* 4:74–82.

Kitchin, Rob. 2014. "Big Data, New Epistemologies and Paradigm Shifts." *Big Data and Society* 1(1):1–12. doi: 10.1177/2053951714528481.

Kleinberg, Jon M. 1999. "Authoritative Sources in a Hyperlinked Environment." *Journal of the ACM* 46:604–632.

Kleitman, Daniel J. and Da-Lun Wang. 1973. "Algorithms for Constructing Graphs and Digraphs with Given Valences and Factors." *Discrete Mathematics* 6:79–88.

Knorr Cetina, Karin. 1981. *The Manufacture of Knowledge: an Essay on the Constructivist and Contextual Nature of Science*. Oxford: Pergamon.

Knorr Cetina, Karin. 1999. *Epistemic Cultures: How the Sciences Make Knowledge*. Cambridge, MA: Harvard University Press.

Kossinets, Gueorgi and Duncan J. Watts. 2006. "Empirical Analysis of an Evolving Social Network." *Science (New York, N.Y.)* 311:88–90.

Kuhn, Thomas S. 1962. *The Structure of Scientific Revolutions*. University of Chicago Press.

Kunegis, Jérôme, Andreas Lommatzsch, and Christian Bauckhage. 2009. "The Slashdot Zoo: Mining a Social Network with Negative Edges." Pp. 741–50 in *8th International World Wide Web Conference (WWW '09)*. Madrid: ACM.

Labelle, Alexie. 2020. "Bringing Epistemology into Intersectional Methodology." *European Journal of Politics and Gender* Online.

Labianca, Giuseppe, and Daniel J. Brass. 2006. "Exploring the Social Ledger: Negative Relationships and Negative Asymmetry in Social Networks in Organizations." *Academy of Management* 31(3):596–614. doi: 10.5465/AMR.2006.21318920.

Lambiotte, Renaud and Pietro Panzarasa. 2009. "Communities, Knowledge Creation, and Information Diffusion." *Journal of Informetrics* 3:180–190.

Lane, David A. 2011. "Complexity and Innovation Dynamics." In *Handbook on the Economic Complexity of Technological Change*, pp. 63–80. Cheltenham: Edward Elgar Publishing.

Langville, Amy N. and Carl D. Meyer. 2005. "A Survey of Eigenvector Methods for Web Information Retrieval." *SIAM Review* 47:135–161.

Lasch, Christopher. 2019. "The Culture of Narcissism." *American Social Character* 241–67.

Latour, Bruno and Steve Woolgar. 1979. *Laboratory Life: The Construction of Scientific Facts*. Beverly Hills: SAGE.

Latour, Bruno. 1984. "The Powers of Association." *Sociological Review* 32(S1):264–80. doi: 10.1111/j.1467-954X.1984.tb00115.x.

Latour, Bruno. 1986. "Visualisation and Cognition: Drawing Things Together." *Knowledge and Society: Studies in the Sociology of Culture Past and Present* 17:813.

Latour, Bruno. 1987. *Science in Action: How to Follow Scientists and Engineers Through Society*. Cambridge, MA: Harvard University Press.

Lawrence, Eric, John Sides, and Henry Farrell. 2010. "Self-Segregation or Deliberation? Blog Readership, Participation, and Polarization in American Politics." *Perspectives on Politics* 8(1):141–57. doi: 10.1017/S1537592709992714.

Lazarsfeld, Paul F., Bernard Berelson, and Hazel Gaudet. 1968. *The People'S Choice*. Columbia University Press.

Lazer, David M. J., 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 Vespignani, and Claudia Wagner. 2020. "Computational Social Science: Obstacles and Opportunities." *Science (New York, N.Y.)* 369(6507):1060–62. doi: 10.1126/science.aaz8170.

Lazer, David M. J., Matthew A. Baum, Yochai Benkler, Adam J. Berinsky, Kelly M. Greenhill, Filippo Menczer, Miriam J. Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, Michael Schudson, Steven A. Sloman, Cass R. Sunstein, Emily A. Thorson, Duncan J. Watts, and Jonathan L. Zittrain. 2018. "The Science of Fake News: Addressing Fake News Requires a Multidisciplinary Effort." *Science* 359(6380):1094–96. doi: 10.1126/science.aao2998.

Lazer, David, Alex Pentland, Lada Adamic, Sinan Aral, Albert-László Barabási, Devon Brewer, Nicholas Christakis, Noshir Contractor, James Fowler, Myron Gutmann, Tony Jebara, Gary King, Michael Macy, Deb Roy, and Marshall Van Alstyne. 2009. "Computational Social Science." *Science* 323(5915):721–23. doi: 10.1126/science.1167742.

Lelkes, Yphtach, Gaurav Sood, and Shanto Iyengar. 2017. "The Hostile Audience: The Effect of Access to Broadband Internet on Partisan Affect." *American Journal of Political Science* 61(1):5–20. doi: 10.1111/ajps.12237.

Leskovec, Jure, Daniel Huttenlocher, and Jon Kleinberg. 2010a. "Predicting Positive and Negative Links." Pp. 641–50 in *International World Wide Web Conference*. Raleigh, USA: ACM.

Leskovec, Jure, Daniel Huttenlocher, and Jon Kleinberg. 2010a. "Signed Networks in Social Media." *Conference on Human Factors in Computing Systems - Proceedings* 2:1361–70. doi: 10.1145/1753326.1753532.

Levin, Daniel Z. and Rob Cross. 2004. "The Strength of Weak Ties You Can Trust: The Mediating Role of Trust in Effective Knowledge Transfer." *Management Science* 50:1477– 1490.

Levy, Omer, Yoav Goldberg, and Ido Dagan. 2015. "Improving Distributional Similarity with Lessons Learned from Word Embeddings." *Transactions of the Association for Computational Linguistics* 3:211–25. doi: 10.1162/tacl_a_00134.

Leydesdorff, Loet, Alexander M. Petersen, and Inga Ivanova. 2017. "Self-Organization of Meaning and the Reflexive Communication of Information." *Social Science Information* 56:4–27.

Leydesdorff, Loet, Mark Johnson, and Inga Ivanova. 2018. "Toward a Calculus of Redun- dancy: Signification, Codification, and Anticipation in Cultural Evolution." *Journal of the Association for Information Science and Technology* 69:1181–1192.

Leydesdorff, Loet. 2021. *The Evolutionary Dynamics of Discursive Knowledge*.

Lietz, Haiko, Claudia Wagner, Arnim Bleier, and Markus Strohmaier. 2014. "When Politicians Talk: Assessing Online Conversational Practices of Political Parties on Twitter." in *Proceedings of the International AAAI Conference on Web and Social Media*. Vol. 8.

Lin, Nan and Mary Dumin. 1986. "Access to Occupations through Social Ties." *Social Networks* 8:365–385.

Lin, Nan, Walter M Ensel, and John C Vaughn. 1981. "Social Resources and Strength of Ties: Structural Factors in Occupational Status Attainment." *American Sociological Review* 46:393.

Lin, Nan. 1999. "Social Networks and Status Attainment." *Annual Review of Sociology* 25:467–487.

Liu, William T. and Robert W. Duff. 1972. "The Strength in Weak Ties." *American Associ- ation for Public Opinion Research* 36:361–366.

Lutz, H., M. T. Herrera Vivar, and L. Supik. 2011. "Framing Intersectionality: An Introduction." Pp. 15–36 in *Framing Intersectionality: Debates on a Multi-Faceted Concept in Gender Studies*, edited by H. Lutz, M. T. Herrera Vivar, and L. Supik. Farnham: Ashgate.

Lyytinen, Kalle and Jan Damsgaard. 2011. "Inter-Organizational Information Systems Adop- tion - A Configuration Analysis Approach." *European Journal of Information Systems* 20:496–509.

Macy, Michael W., and Robert Willer. 2002. "From Factors to Actors: Computational Sociology and Agent-Based Modeling." *Annual Review of Sociology* 28(Resnick 1997):143–66. doi: 10.1146/annurev.soc.28.110601.141117.

Macy, Michael. 1991. "Chains of Cooperation: Threshold Effects in Collective Action." *Amer- ican Sociological Review* pp. 730–747.

Marsden, Peter V. and Karen E. Campbell. 1984. "Measuring Tie Strength." *Social Forces* 63:482–501.

Martin, John Levi. 2002. "Power, Authority, and the Constraint of Belief Systems." *American Journal of Sociology* 107:861.

Martin, John Levi. 2010. “Life’s a Beach but You’re an Ant, and Other Unwelcome News for the Sociology of Culture.” *Poetics* 38(2):229–44. doi: 10.1016/j.poetic.2009.11.004.

Mastropietro, Christopher, and John Vervaeke. 2020. “What Is the Meaning Crisis?” Retrieved February 26, 2024 (https://thesideview.co/journal/diagnosing-the-current-age/).

McCall, Leslie. 2005. “The Complexity of Intersectionality.” *Signs* 30(3):1771–1800. doi: 10.4324/9780203890882.

McFarland, Daniel A., Kevin Lewis, and Amir Goldberg. 2016. “Sociology in the Era of Big Data: The Ascent of Forensic Social Science.” *The American Sociologist* 47(1):12–35. doi: 10.1007/s12108-015-9291-8.

Mcfarland, Daniel, and Heili Pals. 2005. “Motives and Contexts of Identity Change: A Case for Network Effects.” *Social Psychology Quarterly* 68(4):289–315. doi: 10.1177/019027250506800401.

McLean, Paul D. 1998. “A Frame Analysis of Favor Seeking in the Renaissance: Agency, Networks, and Political Culture.” *American Journal of Sociology* 1(104):51–91.

McPherson, Miller, Lynn Smith-Lovin, and James M. Cook. 2001. “Birds of a Feather: Homophily in Social Networks.” *Annual Review of Sociology* 27:415–44.

Merton, Robert K. 1961. “Singletons and Multiples in Scientific Discovery: A Chapter in the Sociology of Science.” *Proceedings of the American Philosophical Society* 105:470–486.

Merton, Robert K. 1968. “The Matthew Effect in Science.” *Science* 159(3810):56–63.

Metaxas, P. T., E. Mustafaraj, K. Wong, L. Zeng, M. O’Keefe, and S. Finn. 2015. “What Do Retweets Indicate? Results from User Survey and Meta-Review of Research.” Pp. 658–61 in *Proceedings of the 9th International Conference on Web and Social Media, ICWSM 2015*.

Metzler, Katie, David A. Kim, Nick Allum, and Angella DenMan. 2016. “Who Is Doing Computational Social Science? Trends in Big Data Research.” *Sage* 1–21. doi: 10.4135/wp160926.

Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. “Efficient Estimation of Word Representations in Vector Space.” *1st International Conference on Learning Representations, ICLR 2013 - Workshop Track Proceedings* 1–12.

Mische, Ann. 2011. “Relational Sociology, Culture, and Agency.” In *The SAGE Handbook of Social Network Analysis*, edited by John Scott and Peter Carrington, volume 1, pp. 80–98. Sage Publications Sage UK: London, England.

Moed, Henk F. 2005. “What do References and Citations Measure?” In *Citation analysis in research evaluation*, pp. 193–208.

Moernaut, Renée, Jelle Mast, Martina Temerman, and Marcel Broersma. 2020a. “Hot Weather, Hot Topic. Polarization and Sceptical Framing in the Climate Debate on Twitter.” *Information, Communication and Society* 1–20. doi: 10.1080/1369118X.2020.1834600.

Mongeon, Philippe, and Adèle Paul-Hus. 2016. “The Journal Coverage of Web of Science and Scopus: A Comparative Analysis.” *Scientometrics* 106(1):213–28. doi: 10.1007/s11192-015-1765-5.

Moody, James. 2004. “The Structure of a Social Science Collaboration Network: Disciplinary Cohesion from 1963 to 1999.” *American Sociological Review* 69:213–238.

Mora-Rios, Jazmín, Miriam Ortega-Ortega, and Guillermina Natera. 2016. “Subjective Experience and Resources for Coping with Stigma in People with a Diagnosis of Schizophrenia: An Intersectional Approach.” *Qualitative Health Research* 26(5):697–711. doi: 10.1177/1049732315570118.

Morgan, Allison C., Dimitrios J. Economou, Samuel F. Way, and Aaron Clauset. 2018. “Prestige Drives Epistemic Inequality in the Diffusion of Scientific Ideas.” *EPJ Data Science* 7:40.

Mucha, Peter J., Thomas Richardson, Kevin Macon, Mason A. Porter, and Jukka Pekka Onnela. 2010. “Community Structure in Time-Dependent, Multiscale, and Multiplex Net- works.” *Science* 328:876–878.

Mügge, Liza, Celeste Montoya, Akwugo Emejulu, and S. Laurel Weldon. 2018. “Intersectionality and the Politics of Knowledge Production.” *European Journal of Politics and Gender* 1(1):17–36. doi: 10.1332/251510818X15272520831166.

Mulkay, Michael J., G. N. Gilbert, and S. Woolgar. 1975. “Problem Areas and Research Networks in Science.” *Sociology* 9:187–203.

Mulkay, Michael. 1974. “Conceptual Displacement and Migration in Science: A Prefatory Paper.” *Science Studies* 4:205–234.

Mustafaraj, Eni, Samantha Finn, Carolyn Whitlock, and Panagiotis T. Metaxas. 2011. "Vocal Minority versus Silent Majority: Discovering the Opionions of the Long Tail." Pp. 103–10 in *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing*. Boston, MA, USA.

Mutz, Diana C., and Paul S. Martin. 2001. "Facilitating Communication across Lines of Political Difference: The Role of Mass Media." *The American Political Science Review* 95(1):97–114.

Nahapiet, Janine and Sumantra Ghoshal. 1998. "Social Capital, Intellectual Capital, and the Organizational Advantage." *Academy of Management Review* 23:242–266.

Nash, Jennifer C. 2008. "Re-Thinking Intersectionality." *Feminist Review* 89(1):1–15.

Nash, Jennifer C. 2016. "Feminist Originalism: Intersectionality and the Politics of Reading." *Feminist Theory* 17(1):3–20. doi: 10.1177/1464700115620864.

Nash, Jennifer C. 2017. "Intersectionality and Its Discontents." *American Quarterly* 69(1):117–129. doi: 10.1353/aq.2017.0006.

Neal, Zachary P. 2018. "A Sign of the Times? Weak and Strong Polarization in the U.S. Congress, 1973–2016." *Social Networks* 60(2020):103–12. doi: 10.1016/j.socnet.2018.07.007.

Németh, Renáta. 2023. "A Scoping Review on the Use of Natural Language Processing in Research on Political Polarization: Trends and Research Prospects." *Journal of Computational Social Science* 6(1):289–313. doi: 10.1007/s42001-022-00196-2.

Newman, Mark E. J. 2001a. "Scientific Collaboration Networks. I. Network Construction and Fundamental Results." *Physical Review E* 64:016131.

Newman, Mark E. J. 2001b. "Scientific Collaboration Networks. II. Shortest Paths, Weighted Networks, and Centrality." *Physical Review E* 64:016132.

Newman, Mark E. J. 2004a. "Coauthorship Networks and Patterns of Scientific Collabora- tion." *Proceedings of the National Academy of Sciences* 101:5200–5205.

Newman, Mark E. J. 2004b. *Who Is the Best Connected Scientist?A Study of Scientific Coauthorship Networks*, pp. 337–370. Berlin: Springer.

Newman, Mark E. J. 2009. “The First-Mover Advantage in Scientific Publication.” *EPL (Europhysics Letters)* 86:68001.

Newman, Mark, Albert-László Barabási, and Duncan J. Watts. 2000. *The Structure and Dynamics of Network*. Princeton University Press.

Nikolov, Dimitar, Diego F. M. Oliveira, Alessandro Flammini, and Filippo Menczer. 2015. “Measuring Online Social Bubbles.” *PeerJ Computer Science* 2015(12):1–14. doi: 10.7717/peerj-cs.38.

Nowak, Martin A. 2006. “Five Rules for the Evolution of Cooperation.” *Science* 314(5805):1560–63. doi: 10.1126/science.1133755.

Noyons, E.C.M. and Anthony F.J. van Raan. 1998. “Monitoring Scientific Developments From a Dynamic Perspective: Self-Organized Structuring to Map Neural Network Re- search.” *Journal of the American Society for Information Science* 49:68–81.

Onnela, Jukka-Pekka, Jari Saramäki, Jorkki Hyvönen, György Szabó, David Lazer, Kimmo Kaski, János Kertész, and Albert-László Barabási. 2007. “Structure and Tie Strengths in Mobile Communication Networks.” *Proceedings of the National Academy of Sciences of the United States of America* 104:7332–6.

Pääkkönen, Juho, Matti Nelimarkka, and Samuli Reijula. 2024. “The Emergence of Computational Social Science: Intellectual Integration or Persistent Fragmentation?” 1–28.

Pachucki, Mark A., and Ronald L. Breiger. 2010. “Cultural Holes: Beyond Relationality in Social Networks and Culture.” *Annual Review of Sociology* 36:205.

Palla, Gergely, Albert-László Barabási, and Tamás Vicsek. 2007. “Quantifying Social Group Evolution.” *Nature* 446:664–667.

Parent, Mike C., Cirleen DeBlaere, and Bonnie Moradi. 2013. “Approaches to Research on Intersectionality: Perspectives on Gender, LGBT, and Racial/Ethnic Identities.” *Sex Roles* 68(11):639–645. doi: 10.1007/s11199-013-0283-2.

Pariser, Eli. 2011. *The Filter Bubble: What the Internet Is Hiding from You*. Penguin UK.

Price, Derek J. De Solla. 1963. *Little Science, Big Science*. New York: Columbia University Press.

Price, Derek J. De Solla. 1965. “Networks of Scientific Papers.” *Science* 169:510–515.

Price, Derek J. De Solla. 1976. “A General Theory of Bibliometric and Other Cumulative Advantage Processes.” *Journal of the American Society for Information Science* 27:292– 306.

Prins, Baukje. 2006. “Narrative Accounts of Origins.” *European Journal of Women’s Studies* 13(3):277–290. doi: 10.1177/1350506806065757.

Prior, Markus. 2007. *Post-Broadcast Democracy: How Media Choice Increases Inequality in Political Involvement and Polarizes Elections*. New York, NY: Cambridge University Press.

Pritchard, Jonathan K., Matthew Stephens, and Peter Donnelly. 2000. “Inference of Population Structure Using Multilocus Genotype Data.” *Genetics* 155(2).

Purdie-Vaughns, Valerie, and Richard P. Eibach. 2008. “Intersectional Invisibility: The Distinctive Advantages and Disadvantages of Multiple Subordinate-Group Identities.” *Sex Roles* 59(5):377–391. doi: 10.1007/s11199-008-9424-4.

Quattrociocchi, Walter, Antonio Scala, and Cass R. Sunstein. 2016. “Echo Chambers on Facebook.” *ArXiv:1411.2893* 51(2015):1–12. doi: 10.1145/2740908.2745939.

Rawlings, Craig M., Daniel A. McFarland, Linus Dahlander, and Dan Wang. 2015. “Streams of Thought: Knowledge Flows and Intellectual Cohesion in a Multidisciplinary Era.” *Social Forces* 93:1687–1722.

Reagans, Ray and Bill McEvily. 2003. “Network Structure and Knowledge Transfer: The Effects of Cohesion and Range.” *Administrative Science Quarterly* 48:240.

Recuero, Raquel, Gabriela Zago, and Felipe Soares. 2019. “Using Social Network Analysis and Social Capital to Identify User Roles on Polarized Political Conversations on Twitter.” *Social Media + Society* 5(2):205630511984874. doi: 10.1177/2056305119848745.

Reichardt, Jörg, and Stefan Bornholdt. 2006. “Statistical Mechanics of Community Detection.” *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 74:1–14. doi: 10.1103/PHYSREVE.74.016110/FIGURES/13/MEDIUM.

Rodenberg, Jeroen, and Pieter Wagenaar. 2016. “Essentializing ‘Black Pete’: Competing Narratives Surrounding the Sinterklaas Tradition in the

Netherlands." *International Journal of Heritage Studies* 22(9):716–28. doi: 10.1080/13527258.2016.1193039.

Rogers, Everett M. 1983. *Diffusion of Innovations 3rd Edition*. New York: The Free Press.

Rosvall, Martin, Daniel Axelsson, and Carl T. Bergstrom. 2009. "The Map Equation." *The European Physical Journal Special Topics* 178:13–23.

Roth, Camille and Jean-Philippe Cointet. 2010. "Social and Semantic Coevolution in Knowledge Networks." *Social Networks* 32:16–29.

Sadilek, Maximilian, Peter Klimek, and Stefan Thurner. 2018. "Asocial Balance — How Your Friends Determine Your Enemies: Understanding the Co-Evolution of Friendship and Enmity Interactions in a Virtual World." *Journal of Computational Social Science* 1(1):227–39. doi: 10.1007/s42001-017-0010-9.

Sangaramoorthy, Thurka, Amelia Jamison, and Typhanye Dyer. 2017. "Intersectional Stigma among Midlife and Older Black Women Living with HIV." *Culture, Health and Sexuality* 19(12):1329–43. doi: 10.1080/13691058.2017.1312530.

Scatà, Marialisa, Alessandro Di Stefano, Aurelio La Corte, Pietro Liò, Emanuele Catania, Ermanno Guardo, and Salvatore Pagano. 2016. "Combining Evolutionary Game Theory and Network Theory to Analyze Human Cooperation Patterns." *Chaos, Solitons and Fractals* 91:17–24. doi: 10.1016/j.chaos.2016.04.018.

Schmidt, Ana Lucía, Fabiana Zollo, Antonio Scala, Cornelia Betsch, and Walter Quattrociocchi. 2018. "Polarization of the Vaccination Debate on Facebook." *Vaccine* 36(25):3606–12. doi: 10.1016/J.VACCINE.2018.05.040.

Schrei, Joshua. 2023. "So You Want to Be a Sorcerer in the Age of Mythic Powers…". The Emerald. Retrieved May 5, 2024 (https://podcasts.apple.com/us/podcast/so-you-want-to-be-a-sorcerer-in-the-age-of/id1465445746?i=1000620936715)

Shi, Feng, Jacob G. Foster, and James A. Evans. 2015. "Weaving the Fabric of Science: Dynamic Network Models of Science's Unfolding Structure." *Social Networks* 43:73–85. Small, Henry and Belver C. Griffith. 1974. "The Structure of Scientific Literatures I: Identifying and Graphing Specialties." *Science Studies* 4:pp. 17–40. 37

Shields, Stephanie A. 2008. "Gender: An Intersectionality Perspective." *Sex Roles* 59(5):301–311. doi: 10.1007/s11199-008-9501-8.

Simmel, Georg. 1904a. "The Sociology of Conflict. I." *American Journal of Sociology* 9(4):490–525. doi: 10.1086/211234.

Simmel, Georg. 1904b. "The Sociology of Conflict. II." *American Journal of Sociology* 9(5):672–89. doi: 10.1086/211248.

Simmel, Georg. 1904c. "The Sociology of Conflict. II." *American Journal of Sociology* 9(5):672–89. doi: 10.1086/211248.

Simon, Herbert A. 1996. *The Sciences of the Artificial*. MIT Press.

Slater, Michael D. 2007. "Reinforcing Spirals: The Mutual Influence of Media Selectivity and Media Effects and Their Impact on Individual Behavior and Social Identity." *Communication Theory* 17(3):281–303. doi: 10.1111/J.1468-2885.2007.00296.X.

Smith, Lance C., and Richard Q. Shin. 2015. "Negotiating the Intersection of Racial Oppression and Heteronormativity." *Journal of Homosexuality* 62(11):1459–84. doi: 10.1080/00918369.2015.1073029.

Soares, Felipe Bonow, Raquel Recuero, and Gabriela Zago. 2019. "Asymmetric Polarization on Twitter and the 2018 Brazilian Presidential Elections." Pp. 67–76 in *Proceedings of the International Conference on Social Media and Society (SMSociety)*.

Spadaro, Giuliana, Caroline Graf, Shuxian Jin, Sakura Arai, Yukako Inoue, Eleanor Lieberman, Maria Isabela Rinderu, Mingliang Yuan, Caspar J. van Lissa, and Daniel Balliet. 2022. "Cross-Cultural Variation in Cooperation: A Meta-Analysis." *Journal of Personality and Social Psychology* 123(5):1024–88. doi: 10.1037/pspi0000389.

Stadtfeld, Christoph, Károly Takács, and András Vörös. 2020. "The Emergence and Stability of Groups in Social Networks." *Social Networks* 60:129–45. doi: 10.1016/j.socnet.2019.10.008.

Staunæs, Dorthe. 2003. "Where Have All the Subjects Gone? Bringing Together the Concepts of Intersectionality and Subjectification." *Nordic Journal of Feminist and Gender Research* 11(2):101–110. doi: 10.1080/08038740310002950.

Strang, David and John W Meyer. 1993. "Institutional Conditions for Diffusion." *Theory and society* 22:487–511.

Strang, David and Sarah A Soule. 1998. “Diffusion in Organizations and Social Movements: From Hybrid Corn to Poison Pills.” *Annual review of sociology* 24:265–290.

Stroud, Natalie Jomini. 2010. “Polarization and Partisan Selective Exposure.” *Journal of Communication* 60(3):556–76. doi: 10.1111/j.1460-2466.2010.01497.x.

Sunstein, Cass R. 1999. “The Law of Group Polarization.” *John M. Olin Program in Law and Economics Working Paper* 91(91).

Sunstein, Cass R. 2001a. *Designing Democracy: What Constitutions Do.* Oxford University Press.

Sunstein, Cass R. 2001b. *Echo Chambers: Bush v. Gore, Impeachment, and Beyond.* Princeton University Press Princeton, NJ.

Sunstein, Cass R., and Adrian Vermeule. 2008. *Conspiracy Theories.* doi: 10.1088/2058-7058/33/2/26.

Tang, Jiliang, Yi Chang, Charu Aggarwal, and Huan Liu. 2016. “A Survey of Signed Network Mining in Social Media.” *ACM Computing Surveys* 49(3):1–37. doi: 10.1145/2956185.

Törnberg, Anton, and Petter Törnberg. 2016. “Muslims in Social Media Discourse: Combining Topic Modeling and Critical Discourse Analysis.” *Discourse, Context & Media* 13:132–42. doi: 10.1016/J.DCM.2016.04.003.

Törnberg, Petter, and Justus Uitermark. 2021. “For a Heterodox Computational Social Science.” *Big Data and Society* 8(2). doi: 10.1177/20539517211047725.

Törnberg, Petter, Claes Andersson, Kristian Lindgren, and Sven Banisch. 2021. “Modeling the Emergence of Affective Polarization in the Social Media Society.” *PLOS ONE* 16(10):e0258259. doi: 10.1371/JOURNAL.PONE.0258259.

Törnberg, Petter. 2017. “Worse than Complex.”

Törnberg, Petter. 2018. “Echo Chambers and Viral Misinformation: Modeling Fake News as Complex Contagion.” *PloS One* 13(9):e0203958.

Tornberg, Petter. 2022. “How Digital Media Drive Affective Polarization through Partisan Sorting.” *Proceedings of the National Academy of Sciences* 119(42):e2207159119. doi: 10.1073/pnas.2207159119.

Traag, V. A., and Jeroen Bruggeman. 2009. "Community Detection in Networks with Positive and Negative Links." *Physical Review E* 80(3):036115.1-036115.6. doi: 10.1103/PhysRevE.80.036115.

Traag, V. A., L. Waltman, and N. J. van Eck. 2019. "From Louvain to Leiden: Guaranteeing Well-Connected Communities." *Scientific Reports* 9(1):1–12. doi: 10.1038/s41598-019-41695-z.

Traag, V. A., P. Van Dooren, and Y. Nesterov. 2011. "Narrow Scope for Resolution-Limit-Free Community Detection." *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 84(1):1–9. doi: 10.1103/PhysRevE.84.016114.

Traag, Vincent A. 2015. "Faster Unfolding of Communities: Speeding up the Louvain Algo- rithm." *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 92:032801.

Travers, Jeffrey, and Stanley Milgram. 1977. "An Experimental Study of the Small World Problem." Pp. 179–97 in *Social Networks*. Elsevier.

Tucker, Joshua A., Andrew Guess, Pablo Barberá, Cristian Vaccari, Alexandra Siegel, Sergey Sanovich, Denis Stukal, and Brendan Nyhan. 2018. *Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature*.

Tufekci, Zeynep. 2014. *Big Questions for Social Media Big Data: Representativeness, Validity and Other Methodological Pitfalls*.

Uitermark, Justus, Vincent A. Traag, and Jeroen Bruggeman. 2016. "Dissecting Discursive Contention: A Relational Analysis of the Dutch Debate on Minority Integration, 1990–2006." *Social Networks* 47(2016):107–15. doi: 10.1016/J.SOCNET.2016.05.006.

Uzzi, Brian, Satyam Mukherjee, Michael Stringer, and Ben Jones. 2013. "Atypical Combi- nations and Scientific Impact." *Science* 342:468–472.

Vaccari, Cristian, Augusto Valeriani, Pablo Barberá, John T. Jost, Jonathan Nagler, and Joshua A. Tucker. 2016. "Of Echo Chambers and Contrarian Clubs: Exposure to Political Disagreement among German and Italian Users of Twitter." *Social Media and Society* 2(3):1–24. doi: 10.1177/2056305116664221.

Valentine, Gill. 2007. "Theorizing and Researching Intersectionality: A Challenge for Feminist Geography." *The Professional Geographer* 59(1):10–21. doi: 10.1111/j.1467-9272.2007.00587.x.

Van Der Pijl, Yvon, and Karina Goulordava. 2014. “Black Pete, ‘Smug Ignorance,’ and the Value of the Black Body in Postcolonial Netherlands.” *NWIG New West Indian Guide* 88(3–4):262–91. doi: 10.1163/22134360-08803062.

Verloo, Mieke. 2006. “Multiple Inequalities, Intersectionality and the European Union.” *European Journal of Women’s Studies* 13(3):211–228. doi: 10.1177/1350506806065753.

Vilhena, Daril, Jacob Foster, Martin Rosvall, Jevin West, James Evans, and Carl T. Bergstrom. 2014. “Finding Cultural Holes: How Structure and Culture Diverge in Net- works of Scholarly Communication.” *Sociological Science* 1:221–238.

Vliegenthart, Rens, and Jasper Zuure. 2020. *Doen, Durven of de Waarheid? Democratie in Digitale Tijden*. edited by J. de Ridder, R. Vliegenthart, and J. Zuure. Amsterdam University Press.

Walby, Silvia. 2007. “Complexity Theory, Systems Theory, and Multiple Intersecting Social Inequalities.” *Philosophy of the Social Sciences* 37(4):449–70. doi: 10.4135/9781412983907.n1952.

Warner, Leah R. 2008. “A Best Practices Guide to Intersectional Approaches in Psychological Research.” *Sex Roles* 59(5):454–463. doi: 10.1007/s11199-008-9504-5.

Watts, Duncan J. 1999. “Networks, Dynamics, and the Small-World Phenomenon.” *American Journal of Sociology* 105:493–527.

Watts, Duncan J. 2002. “A simple Model of Global Cascades on Random Networks.” *Pro- ceedings of the National Academy of Sciences* 99:5766–5771.

Watts, Duncan J. 2013. “Computational Social Science: Exciting Progress and Future Directions.” *The Bridge on Frontiers of Engineering* 43(4):5–10.

Watts, Duncan J. 2014. “Common Sense and Sociological Explanations.” *American Journal of Sociology* 120(2):313–51. doi: 10.1086/678271.

Watts, Duncan J., and Steven H. Strogatz. 1998. “Collective Dynamics of ‘Small-World’ Networks.” *Nature* 393(6684):440–42. doi: 10.1038/30918.

Waugh, Andrew Scott, Liuyi Pei, James H. Fowler, Peter J. Mucha, and Mason A. Porter. 2009. “Party Polarization in Congress: A Network Science Approach.”

Way, Samuel F., Allison C. Morgan, Daniel B. Larremore, and Aaron Clauset. 2019. “Productivity, Prominence, and the Effects of Academic Environment.” *Proceedings of the National Academy of Sciences of the United States of America* 166:10729–10733.

Wekker, Gloria. 2016. *White Innocence: Paradoxes of Colonialism and Race*. Duke University Pres.

Wellman, Barry, and Scot Wortley. 1990. “Different Strokes from Different Folks: Community Ties and Social Support.” *American Journal of Sociology* 96(3):558–88. doi: 10.1086/229572.

Wellman, Barry, and Stephen D. Berkowitz. 1988. *Social Structures: A Network Approach*. Cambridge University Press.

Wellman, Barry. 1983. “Network Analysis: Some Basic Principles.” *Sociological Theory* 1:155–200.

White, Harrison C. 1992. *Identity and Control: A Structural Theory of Social Action*. Princeton University Press.

White, Harrison C., Frédéric C. Godart, and Victor P. Corona. 2007. “Mobilizing Identities: Uncertainty and Control in Strategy.” *Theory, Culture & Society* 24(8):181–202. doi: 10.1177/0263276407084703.

White, Harrison C., Scott A. Boorman, and Ronald L. Breiger. 1976. “Social Structure from Multiple Networks . I . Blockmodels of Roles and Positions.” *American Journal of Sociology* 81:730–780.

Williams, Hywel T. P., James R. McMurray, Tim Kurz, and F. Hugo Lambert. 2015. “Network Analysis Reveals Open Forums and Echo Chambers in Social Media Discussions of Climate Change.” *Global Environmental Change* 32:126–38. doi: 10.1016/j.gloenvcha.2015.03.006.

Wilson, A. 1978. *Finding a Voice: Asian Women in Britain*. London: Virago Press.

Wilson, P. A., Valera, P., Martos, A. J., Wittlin, N. M., Muñoz-Laboy, M. A., & Parker, R. G. (2016). “Contributions of qualitative research in informing HIV/AIDS interventions targeting black MSM in the United States. “ *Journal of Sex Research*, 53(6), 642–654.

Wu, Siqi, and Paul Resnick. 2021. “Cross-Partisan Discussions on YouTube: Conservatives Talk to Liberals but Liberals Don’t Talk to Conservatives.” *Icwsm*.

Yardi, Sarita, and Danah Boyd. 2010. “Dynamic Debates: An Analysis of Group Polarization Over Time on Twitter.” *Bulletin of Science, Technology & Society* 30(5):316–27. doi: 10.1177/0270467610380011.

Yoon, Ho Young, and Han Woo Park. 2014. “Strategies Affecting Twitter-Based Networking Pattern of South Korean Politicians: Social Network Analysis and Exponential Random Graph Model.” *Quality and Quantity* 48(1):409–23. doi: 10.1007/s11135-012-9777-1.

Yuval-Davis, Nira. 2006. “Intersectionality and Feminist Politics.” *European Journal of Women’s Studies* 13(3):193–209. doi: 10.1177/1350506806065752.

Zeng, An, Zhesi Shen, Jianlin Zhou, Jinshan Wu, Ying Fan, Yougui Wang, and H. Eugene Stanley. 2017. “The Science of Science: From the Perspective of Complex Systems.” *Physics Reports* 714–715:1–73. doi: 10.1016/J.PHYSREP.2017.10.001.

Zhang, Jun, Wei Wang, Feng Xia, Yu Ru Lin, and Hanghang Tong. 2020. “Data-Driven Computational Social Science: A Survey.” *Big Data Research* 21:100145. doi: 10.1016/j.bdr.2020.100145.

Zheng, Xiaolong, Daniel Zeng, and Fei Yue Wang. 2015. “Social Balance in Signed Networks.” *Information Systems Frontiers* 17(5):1077–95. doi: 10.1007/s10796-014-9483-8.

Zilber, Tammar B. 2008. “The Work of Meaning in Institutional Process and Thinking.” In *The SAGE Handbook of Organizational Institutionalism*, pp. 151–169. London: SAGE.

Appendix for "Adoption *and* Adaptation: A Computational Case Study of the Spread of Granovetter's Weak Ties Hypothesis"

Section A: Topic Modeling

We used an LDA topic model[18] (Blei, Ng, and Jordan 2003; Pritchard, Stephens, and Donnelly 2003) on a corpus consisting of the abstracts of 5,696 scientific publications and comprising 30,134 unique words.[19] The LDA is an unsupervised model that identifies topics—consisting of frequently co-occurring words—in documents. Documents can contain numerous topics while the same word can belong to multiple topics. In addition to a standard list of English stop words,[20] we composed a stop word list specific to our corpus, including words such as: paper, result, finding, argue, relationship, literature, different, examine, investigate, network. We instructed the model to find exactly 15 topics.[21] In addition, we instructed the model to omit words that occur in fewer than 30 documents, or in more than half of them. Table 1 displays the most important 30 words of the 15 topics. We then labeled these topics to capture their essence.

Table 1 (next pages): Top 30 words for the 15 topics detected by LDA

[18] We used LDA as implemented in python scikit-sklearn (Pedregosa et al. 2011).
[19] More precisely, 30,134 tokens. We used python nltk tokenize to count all the unique tokens in the abstracts.
[20] We used a combination of the python scikit-learn English stop words list and the python NLTK stop words list.
[21] We explored the results with more and fewer topics, and found similar patterns for correlations between topics and diffusion communities as well as for topic similarity between communities.

Topic: Label	Words
0: Organizational Advantage	knowledge; innovation; firms; firm; organizational; information; learning; technology; industry; external; performance; transfer; management; structural; relationships; partners; technological; inter; process; strategic; capabilities; development; organizations; knowledge transfer; innovations; managers; competitive; internal; activities; acquisition
1: Labor Market	community; communities; job; organizations; participation; members; agents; civic; local; movement; organizational; conflict; engagement; information; organization; identity; groups; population; farmers; structure; sense; indirect; reciprocity; job search; leader; risk; related; associated; crime; criminal
2: Survey Data	data; effects; effect; survey; individual; level; associated; individuals; neighborhood; positive; status; likely; size; characteristics; friendship; relationships; negative; levels; low; lower; friends; variables; contacts; regression; contact; ethnic; age; outcomes; income; gender
3: Civic Bridging and Bonding	capital; social capital; bridging; bonding; immigrants; students; associations; creative; bridging social; voluntary; united states; human capital; united; participation; forms; capital social; states; bonding social; dimensions; bonding bridging; positive; positively; development; school; activities; teachers; concept; self; resources; associated
4: Health Behavior	health; influence; diffusion; behavior; adoption; human; decision; individual; individuals; behaviors; systems; model; peer; making; decision making; process; social influence; evidence; mechanisms; norms; behavioral; methods; wiley; spread; effects; interventions; personality; opinion; processes; probability

5: Marketing & Design	purpose; value; service; design; methodology; marketing; relationships; supply; originality; authors; originality value; design methodology; product; customer; customers; relational; model; practical; influence; exchange; managers; consumers; limitations; supplier; chain; commitment; services; consumer; supply chain; purpose
6: Family Support	resources; women; family; team; project; support; relationships; access; members; men; teams; care; entrepreneurs; work; term; venture; financial; projects; entrepreneurial; informal; services; program; investment; public; long; long term; benefits; advice; success; gender
7: Ties	ties; weak; strong; tie; strength; weak ties; political; social ties; strong ties; tie strength; effect; collective action; discussion; influence; information; strong weak; non; civil; kin; weak tie; party; collective; action; relationships; larger; effects; participation; formation; individuals; positive
8: Governance	performance; cross; institutional; environmental; local; embeddedness; organizational; regional; sector; coordination; innovative; organizations; regions; level; government; data; actors; response; organization; national; firm performance; public; region; governance; units; environment; transactions; orientation; conditions; managers
9: Politics & Markets	relations; interaction; power; processes; groups; interactions; change; search; political; actors; people; market; action; collective; labor; recent; time; outcomes; behavior; changes; society; structures; individual; information; strategies; conditions; control; dynamics; various; digital

10: Entrepreneurship	business; management; policy; development; theoretical; entrepreneurship; networking; empirical; value; review; firms; growth; governance; organizational; small; regional; entrepreneurial; economic; processes; conceptual; concept; model; context; alliance; strategic; entrepreneurs; global; case; concepts; resource
11: Complex Networks	model; structure; models; world; time; nodes; properties; degree; complex; structural; links; dynamics; data; small; evolution; real; learning; methods; number; centrality; systems; link; empirical; scale; method; process; structures; clustering; distribution; global
12: Online Communication	information; online; communication; media; support; users; social media; people; internet; Facebook; self; social support; personal; friends; data; user; mobile; face; participants; online social; content; individuals; sites; interaction; relationships; related; twitter; offline; students; technologies
13: Economic Development	economic; local; cultural; development; interviews; market; urban; integration; informal; practices; migration; international; career; mobility; rural; migrants; markets; leadership; case; countries; qualitative; professional; culture; labor; production; leaders; work; place; areas; city
14: Scientific Community	trust; group; groups; diversity; data; collaboration; level; cohesion; scientific; core; researchers; scientists; measure; collaborative; academic; members; structure; measures; diverse; brokerage; structural; actors; general; authors; science; people; characteristics; positions; membership; elite

References

Blei, David M., Andrew Y. Ng, and Michael I. Jordan. 2003. “Latent Dirichlet Allocation.” *Journal of Machine Learning Research* 3(Jan):993–1022.

Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, and Mathieu Blondel. 2011. "Scikit-Learn: Machine Learning in Python." *The Journal of Machine Learning Research* 12:2825–30.

Pritchard, Jonathan K., Matthew Stephens, and Peter Donnelly. 2003. "Inference of Population Structure Using Multilocus Genotype Data." *Genetics* 164(4):1567–87.

Section B: Community Descriptions

To show how scholars of the diffusion communities conceptualize the Strength of Weak Ties and embed the hypothesis in their communities' research interests and theoretical frameworks, the main paper contains in-depth descriptions of the largest three communities. As an extension, this appendix offers descriptions of the subsequent largest seven communities in the network, that each consist of more than 300 scholars.

These results are based on a close reading of the key contributions in each community, combined with the previously reported results on the communities' main research areas (figure 7), topics (figure 5) and high-in degree scholars (figure 9). We also take the relations between the communities, visualized in Figure 1 of this appendix below, into consideration. Based on our reading of the literature, we cite one statement or two statements per community that we find typical for how members of the cluster interpret the Strength of Weak Ties hypothesis. Although such a selection is of course arbitrary to a certain extent, it gives a sense of the kind of arguments that are made within the community. We select these statements from articles that are frequently cited by scholars within the respective communities. To complete the overview, we provide for each community a table with the local in-degree (refs from within the community), global in-degree (refs from entire network) and research areas[22] for the scholars with highest local in-degree.

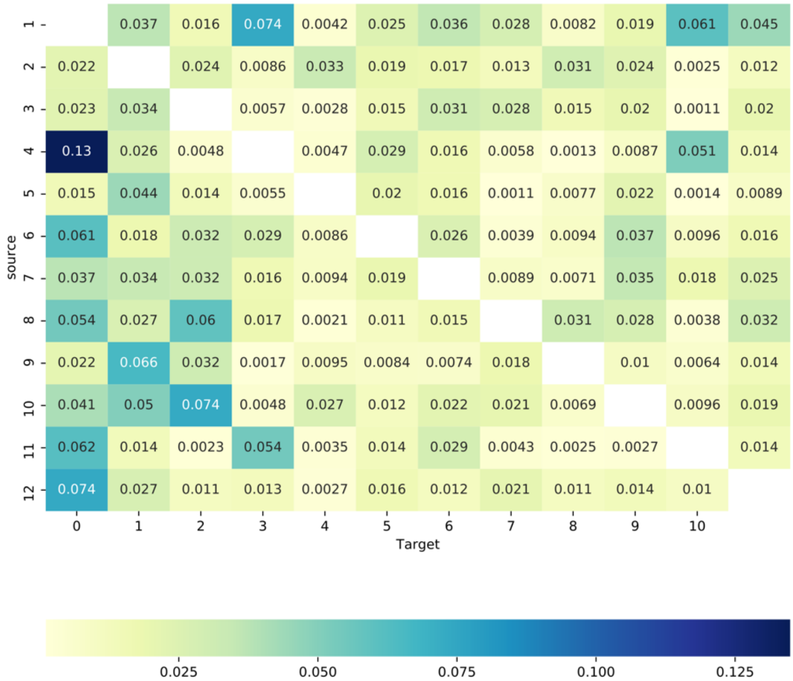


Figure 1: Directed relations between the communities. The directed relations (cell numbers) are calculated as the number of edges from the source community (rows) to the target community (columns) divided by the total amount of outgoing edges from the source community times the total amount of incoming edges to the target community. By construction, this matrix is not symmetric.

Community 4: Entrepreneurial Networking Community

This community embeds the Strength of Weak Ties in entrepreneurial context. The community in strongly linked to the Organizational Advantage Community (community 1) in the diffusion network (see Figure 1) which is in line with their common interest in businesses. However, this community operates in the niche of research on

[22] Each of the publications in our Web of Science data is categorized in one or several research fields. The research fields per author are composed by aggregating the research fields of authors' publications, ordered by their frequency. We like to emphasize that authors' research areas are thus based only on their publications that reference Granovetter (1973) and not on their other works.

entrepreneurs, start-ups and new ventures or nascent firms. Popular works includes studies into the effects of the activity of networking for entrepreneurs' effectiveness (Dubini and Aldrich 1991) and loan conditions (Uzzi 1999). Whereas the focus of the Organizational Advantage Community is predominantly on structures of networks between or within corporates, this community takes a more ego-centric network approach, for instance by studying how properties of the ego network of an entrepreneur or young firm affect success (Davidsson and Honig 2003; Stuart, Hoang, and Hybels 1999). Granovetter's concepts of strong and weak ties are typically introduced in the context of the entrepreneurial network, see for example this reference by Hoang and Antoncic "Granovetter's (1973) notion of weak ties, in particular, describes the extent to which actors can gain access to new information and ideas through ties that lie outside of their immediate cluster of contacts. For example, Hansen and Witkowski (1995) found that entrepreneurs who had network ties that extended outside of the US at the time of start-up were more likely to continue to conduct business abroad." (2003:171).

	Name	In-degree local	In-degree global	Research areas
1	Ha Hoang	222	322	Business & Economics
2	Bostjan Antoncic	181	230	Business & Economics
3	Benson Honig	180	209	Business & Economics
4	Per Davidsson	176	205	Business & Economics
5	Howard E. Aldrich	175	249	Sociology, Business & Economics
6	Toby E. Stuart	115	217	Business & Economics, Sociology, Mathematics
7	Tom Elfring	115	143	Business & Economics
8	Brian Uzzi	82	422	Sociology. Science & Technology, Business & Economics
9	Ralph C.Hybels	79	133	Business & Economics
10	Paola Dubini	73	86	Business & Economics

Table 2: An overview of the scholars of community 4 that are most often referenced within their community (in-degree local). The table also lists how often they are referenced in the entire network (in-degree global), and their main research areas, ordered by their relevance in this dataset. These stats are based on the data from Web of Science.

References

Davidsson, Per and Benson Honig. 2003. “The Role of Social and Human Capital Among Nascent Entrepreneurs.” *Journal of Business Venturing* 18(3):301–31.

Dubini, Paola and Howard Aldrich. 1991. “Personal and Extended Networks Are Central to the Entrepreneurial Process.” *Journal of Business Venturing* 6(5):305–13.

Hoang, Ha and Bostjan Antoncic. 2003. “Network-Based Research in Entrepreneurship: A Critical Review.” *Journal of Business Venturing* 18(2):165–87.

Stuart, Toby E., Ha Hoang, and Ralph C. Hybels. 1999. "Interorganizational Endorsements and the Performance of Entrepreneurial Ventures." *Administrative Science Quarterly* 44(2):315–49.

Uzzi, Brian. 1999. "Embeddedness in the Making of Financial Capital: How Social Relations and Networks Benefit Firms Seeking Financing." *American Sociological Review* 64(4):481–505.

Community 5: The Neighbourhood Cohesion Community

Scholars of the Neigbourhood Cohesion Community study networks in the context of social cohesion and the their effects on health (e.g. Berkman et al. 2000) and crime (e.g. Sampson and Groves 1989). The dominant level of analysis is the neighborhood (e.g. Forrest and Kearns 2001; Sampson, Morenoff, and Gannon-Rowley 2002), and most empirical results stem from surveys. Scholars in this community are active in various research areas, but their dominant research fields are sociology and public, environmental and occupational health, see Figure 7 in the main paper. Similar to the Ego Networks Community (community 2), to which this community is strongly linked in the diffusion network (see Figure 1 of this appendix), this community explores social capital. Yet, social capital in this literature is conceptualized not as a resource of individuals to attain jobs or status, like it is in the Ego Networks Community, but rather as an asset of neighborhoods or countries that encapsulates trust, cohesion or solidarity (Putnam 2007). In line with their interest in cohesion and trust, scholars in this community explain weak ties in related terms, describing weak ties for example as 'superficial' (Forrest and Kearns 2001:2138) or 'less intimate contacts' (Berkman et al. 2000:850; Sampson et al. 2002:459), see for instance: "Moreover, weak ties—less intimate connections between people based on more infrequent social interaction—maybe essential for establishing social resources such as job referrals because they integrate the community by bringing together otherwise disconnected subgroups (Granovetter 1973, Bellair 1997)" (Sampson et al. 2002:459).

	Name	In-degree local	In-degree global	Research areas
1	Lisa F. Berkman	144	204	Public, Environmental & Occupational Health, Biomedical Social Sciences
2	Robert J. Sampson	125	149	Sociology
3	Thomas A. Glass	124	175	Public, Environmental & Occupational Health, Biomedical Social Sciences
4	Teresa E. Seeman	124	175	Public, Environmental & Occupational Health, Biomedical Social Sciences
5	Ian Brissette	120	170	Public, Environmental & Occupational Health, Biomedical Social Sciences
6	Thomas Gannon-Rowley	79	94	Sociology
7	Jeffrey D. Morenoff	79	94	Sociology
8	Ray Forrest	78	105	Environmental Sciences & Ecology, Urban Studies, Sociology
9	Ade Kearns	76	103	Environmental Sciences & Ecology, Urban Studies
10	Robert D. Putnam	51	107	Public, Environmental & Occupational Health, Government & Law

Table 3: An overview of the scholars of community 5 that are most often referenced within their community (in-degree local). The table also lists how often they are referenced in the entire network (in-degree global), and their main research areas, ordered by their relevance in this dataset. These stats are based on the data from Web of Science.

References

Berkman, Lisa F., Thomas Glass, Ian Brissette, and Teresa E. Seeman. 2000. "From Social Integration to Health: Durkheim in the New Millennium." *Social Science and Medicine* 51(6):843–57.

Forrest, Ray and Ade Kearns. 2001. "Social Cohesion, Social Capital and the Neighbourhood." *Urban Studies* 38(12):2125–43.

Sampson, Robert J. and W. Byron Groves. 1989. "Community Structure and Crime: Testing Social-Disorganization Theory." *American Journal of Sociology* 94(4):774–802.

Sampson, Robert J., Jeffrey D. Morenoff, and Thomas Gannon-Rowley. 2002. "Assessing 'Neighborhood Effects': Social Processes and New Directions in Research." *Annual Review of Sociology* 28(1):443–78.

Putnam, Robert D. 2007. "E Pluribus Unum: Diversity and Community in the Twenty-First Century the 2006 Johan Skytte Prize Lecture." *Scandinavian Political Studies* 30(2):137–74.

Community 6: Resource Governance Community

Scholars in this Resource Governance Community are concerned with the role of social capital and networks in transitional countries and underdeveloped communities from a policy and governance perspective. They are interested in questions regarding natural resource management and collective action dilemmas, developing the method of stakeholder analysis to study how stakeholders' network topology influences the management of natural resources (Bodin and Crona 2009) and how to use this data in order to target the right actors (Prell, Hubacek, and Reed 2009). With this focus on management and development, this community interprets social capital from a collaborative action perspective. For example, Woolcock and Narayan define social capital as the "the norms and networks that enable people to act collectively" (2000:225). Similarly, the weak ties are deemd important due to their ability to potentially solve collective action dilemmas. For example, Crona and Bodin write "Bridging ties, on the other hand, provide access to external resources of various kinds, and are often needed to help actors initiate or support collective action (e.g. Granovetter, 1973; Newman and Dale, 2007; Lin, 2002), both of which are vital for resource governance" (2006:369).

	Name	In-degree local	In-degree global	Research areas
1	Örjan Bodin	217	224	Environmental Sciences & Ecology, Geography. Science & Technology
2	Beatrice I. Crona	197	204	Environmental Sciences & Ecology, Geography, Business & Economics
3	Michael Woolcock	181	527	Business & Economics, Public Administration, Sociology
4	Henrik Ernstson	126	126	Environmental Sciences & Ecology, Engineering, Water Resources
5	Christina Prell	115	122	Environmental Sciences & Ecology, Sociology, Public Administration
6	Deepa Narayan	113	256	Business & Economics, Public Administration
7	Klaus Hubacek	104	111	Environmental Sciences & Ecology, Public Administration, Sociology
8	Mark S. Reed	97	104	Environmental Sciences & Ecology, Public Administration, Sociology
9	Ryan R.J. McAllister	61	61	Environmental Sciences & Ecology, Biodiversity & Conservation, Government & Law
10	Jens Newig	58	59	Environmental Sciences & Ecology

Table 4: An overview of the scholars of community 6 that are most often referenced within their community (in-degree local). The table also lists how often they are referenced in the entire network (in-degree global), and their main research areas, ordered by their relevance in this dataset. These stats are based on the data from Web of Science.

References

Bodin, Örjan and Beatrice I. Crona. 2009. "The Role of Social Networks in Natural Resource Governance: What Relational Patterns Make a Difference?" *Global Environmental Change* 19(3):366–74.

Crona, Beatrice and Örjan Bodin. 2006. "What You Know Is Who You Know? Communication Patterns Among Resource Users as a Prerequisite for Co-Management." *Ecology and Society* 11(2).

Prell, Christina, Klaus Hubacek, and Mark Reed. 2009. "Stakeholder Analysis and Social Network Analysis in Natural Resource Management." *Society and Natural Resources* 22(6):501–18.

Woolcock, Michael and Deepa Narayan. 2000. "Social Capital: Implications for Development Theory, Research, and Policy." *World Bank Research Observer* 15(2):225–49.

Community 7: Relational Sociology Community

The basis of this community formed in the early days of network analysis in sociology by relational sociologists such as Harrison White, Ronald Breiger, Scott A. Boorman, Mustafa Emirbayer, and Jeff Goodwin. Part of this work is theoretical, questioning the ontological basis of networks and social structure (Emirbayer and Goodwin 1994; White, Boorman, and Breiger 1976). At the same time, it offers applicable and detailed techniques for the analysis of relational data, most prominently block model analysis (White et al. 1976) and hierarchical clustering (Breiger, Boorman, and Arabie 1975). Granovetter takes a more prominent position being mentioned as key figure in network analysis (Emirbayer and Goodwin 1994:1412) and in some of the 'acknowledgments' (Padgett and Ansell 1993) in this community. Later authors of this community link to this relational sociological literature, using block model analysis and other network techniques to study a variety of historical and sociological processes, publishing predominantly in sociological journals (see Figure 7 in the main paper). By the theoretical works, the Strength of Weak Ties are mentioned as being part of relational sociology, see for example

Emirbayer and Goodwin “Relational analysis, however, also demonstrate that ‘weak’ ties indirectly connecting individuals or bridging the ‘structural holes’ between isolated social groups may be crucial for many important social processes, such as locating employment opportunities (e.g. Granovetter, 1973; Burt 1992)” (1994:1419). In the more technical applications of block models, Granovetter’s weak ties are mentioned as limitation to these models which are fit for strong ties only, see White, Boorman and Breiger: “There is an important limitation in the viewpoint urged thus far...The contrast between weak and strong ties should be a major factor in connectivity analysis for large populations (Rapoiport and Horvath 1961; Granovetter 1973; Boorman 1975). Blockmodels as developed thus far deal chiefly with strong ties” (White et al. 1976:773–74).

	Name	**In-degree local**	**In-degree global**	**Research areas**
1	Ronald L. Breiger	114	177	Sociology, Mathematics, Mathematical Methods In Social Sciences
2	Scott A. Boorman	102	199	Sociology, Mathematics, Psychology
3	Harrison C. White	76	123	Sociology, Mathematics, Mathematical Methods In Social Sciences
4	Mustafa Emirbayer	76	185	Sociology
5	Jeff Goodwin	76	185	Sociology
6	John F. Padgett	75	167	Sociology
7	Christopher K. Ansell	75	167	Sociology
8	Naomi B. Rosenthal	36	38	Sociology

9	Michele Ethier	36	38	Sociology
10	Meryl Fingrutd	36	38	Sociology

Table 5: An overview of the scholars of community 7 that are most often referenced within their community (in-degree local). The table also lists how often they are referenced in the entire network (in-degree global), and their main research areas, ordered by their relevance in this dataset. These stats are based on the data from Web of Science.

References

Emirbayer, Mustafa and Jeff Goodwin. 1994. "Network Analysis, Culture and the Problem of Agency." *American Journal of Sociology* 99(6):1411–54.

White, Harrison C., Scott A. Boorman, and Ronald L. Breiger. 1976. "Social Structure from Multiple Networks . I . Blockmodels of Roles and Positions." *American Journal of Sociology* 81(4):730–80.

Breiger, Ronald L., Scott A. Boorman, and Phipps Arabie. 1975. "An Algorithm for Clustering Relational Data with Applications to Social Network Analysis and Comparison with Multidimensional Scaling." *Journal of Mathematical Psychology* 12(3):328–83.

Padgett, John F. and Christopher K. Ansell. 1993. "Robust Action and the Rise of the Medici, 1400-1434." *American Journal of Sociology* 98(6):1259–1319.

Community 8: Word-of-Mouth Marketing Community

This community of scholars studies the process of word-of-mouth communication with the objective to better understand consumer behavior and improve marketing techniques. Scholars employ a variety of research methods, ranging from qualitative interviews (e.g. Brown, Broderick, and Lee 2007; Muniz and O'Guinn 2001) and network analysis of electronic referral networks (e.g. Brown and Reingen 1987) to complex systems modeling to investigate how low-level interactions aggregate to macro-patterns (e.g. De Bruyn and Lilien 2008; Goldenberg, Libai, and Muller 1987). In line with this diversity in their methodological toolset, the community is connected strongly to both the Organizational Advantage Community (community 1) as well as the Complex Systems Community (community 3), see Figure 1. The Strength of Weak Ties takes a prominent position in this literature, because it provides hypotheses that can be straightforwardly tested by networks of consumer referrals. For example, Brown and Reinigen derive 6 separate hypotheses based on Granovetter's 1973 paper, which they subsequently test in 'who-told-whom networks' (1987:354) of piano teachers. Goldenberg et al. pose that the Strength of Weak Ties "offers one of the most important conceptual explanations of the process by which micro-level interactions affect macro-level phenomena." (1987:213).

	Name	**In-degree local**	**In-degree global**	**Research areas**
1	Peter H. Reingen	174	217	Business & Economics
2	Jacqueline J. Brown	151	189	Business & Economics
3	Gary L. Lilien	61	72	Business & Economics
4	Arnaud De Bruyn	61	72	Business & Economics
5	Amanda J. Broderick	60	70	Business & Economics
6	Nick Lee	60	70	Business & Economics
7	Jo Brown	60	70	Business & Economics
8	Jacob Goldenberg	57	85	Business & Economics
9	David Godes	55	59	Business & Economics
10	Eitan Muller	55	83	Business & Economics

Table 6: An overview of the scholars of community 8 that are most often referenced within their community (in-degree local). The table also lists how often they are referenced in the entire network (in-degree global), and their main research areas, ordered by their relevance in this dataset. These stats are based on the data from Web of Science.

References

Brown, Jo, Amanda J. Broderick, and Nick Lee. 2007. “Word of Mouth Communication within Online Communities : Conceptualizing the Online Social Network.” *Journal of Interactive Marketing* 21(3):2–20.

Brown, Jacqueline J. and Peter H. Reingen. 1987. “Social Ties and Word-of-Mouth Referral Behavior.” *Journal of Consumer Research* 14(3):350–62.

De Bruyn, Arnaud and Gary L. Lilien. 2008. “A Multi-Stage Model of Word-of-Mouth Influence through Viral Marketing.” *International Journal of Research in Marketing* 25(3):151–63.

Goldenberg, Jacob, Barak Libai, and Eitan Muller. 1987. “Talk of the Network: A Complex Systems Look at the Underlying Process of Word-of-Mouth.” *Marketing Letters* 12(3):211–23.

Muniz, Albert M. and Thomas C. O’Guinn. 2001. “Brand Community.” *Journal of Consumer Research* 27(4):412–32.

Community 9: The Social Media and Political Engagement Community

This Social Media and Political Engagement Community studies the role of networks in political engagement and opinion formation. Scholars use survey data (Huckfeldt et al. 1995; Huckfeldt and Mendez 2004; McClurg 2003; Mutz and Martin 2001; Steinfield, Ellison, and Lampe 2008), and social media data (e.g. Sosik and Bazarova 2014) to study how users connect with one another and how this influences their social capital, political engagement and views. The—at that time upcoming—social networking sites (SNSs) were fertile ground for researchers to study the functions of connecting for individuals (Donath and Boyd 2004; Kietzmann et al. 2011; Steinfield et al. 2008), in which the focus of attention is placed at people’s information seeking activities and the role of strong and weak ties therein (Ellison, Steinfield, and Lampe 2011). Huckfeldt et al (1995) explain in their publication highly cited in this community: “Thus the focus in on communication rather than influence, and our particular concern is with how social networks of political communication serve as micro environmental filters on the macro environmental flow of information (Granovetter 1973; Burt 1987). To what extent do cohesive social groups and weak social ties serve to advance or retard the communication and dissemination of public opinion in the larger environment?". This community hosts a relatively high number—particularly in contrast to other communities—of communication scientists (17%), computer scientists (9%) and government and law scholars (13%), see Figure 7 in the main paper.

	Name	**In-degree local**	**In-degree global**	**Research areas**
1	Robert Huckfeldt	116	136	Government & Law, Psychology, Geography
2	Nicole B. Ellison	111	188	Psychology, Communication, Sociology
3	Cliff Lampe	111	180	Communication, Psychology, Information Science & Library Science
4	Charles Steinfield	102	166	Computer Science. Psychology, Information Science & Library Science
5	Paul A. Beck	85	86	Government & Law. Communication, Social Sciences
6	Russell J. Dalton	80	81	Government & Law
7	Jeffrey Levine	69	70	Government & Law
8	Danah Boyd	60	103	Engineering, Telecommunications
9	Judith Donath	60	103	Engineering, Telecommunications
10	Ian P. McCarthy	33	50	Business & Economics

Table 7: An overview of the scholars of community 9 that are most often referenced within their community (in-degree local). The table also lists how often they are referenced in the entire network (in-degree global), and their main research areas,

ordered by their relevance in this dataset. These stats are based on the data from Web of Science.

References

Donath, Judith and Danah Boyd. 2004. "Public Displays of Connection." *BT Technology Journal* 22(1):71–82.

Ellison, Nicole B., Charles Steinfield, and Cliff Lampe. 2011. "Connection Strategies: Social Capital Implications of Facebook-Enabled Communication Practices." *New Media and Society* 13(6):873–92.

Huckfeldt, Robert, Paul A. Beck, Russell J. Dalton, and Jeffrey Levine. 1995. "Political Environments, Cohesive Social Groups and the Communication of Public Opinion." *American Journal of Political Science* 39(4):1025–54.

Huckfeldt, Robert and Jeanette Morehouse Mendez. 2004. "Disagreement, Ambivalence and Engagement: The Political Consequences of Hetergenous Networks." *Political Psychology* 25(1):65–95.

Kietzmann, Jan H., Kristopher Hermkens, Ian P. McCarthy, and Bruno S. Silvestre. 2011. "Social Media? Get Serious! Understanding the Functional Building Blocks of Social Media." *Business Horizons* 54(3):241–51.

McClurg, Scott D. 2003. "Social Networks and Political Participation: The Role of Social Interaction in Explaining Political Participation." *Political Research Quarterly* 56(4):449–64.

Mutz, Diana C. and Paul S. Martin. 2001. "Facilitating Communication across Lines of Political Difference: The Role of Mass Media." *American Political Science Review* 95(1):97–114.

Sosik, Victoria Schwanda and Natalya N. Bazarova. 2014. "Relational Maintenance on Social Network Sites: How Facebook Communication Predicts Relational Escalation." *Computers in Human Behavior* 35(1):124–31.

Steinfield, Charles, Nicole B. Ellison, and Cliff Lampe. 2008. “Social Capital, Self-Esteem, and Use of Online Social Network Sites: A Longitudinal Analysis.” *Journal of Applied Developmental Psychology* 29(6):434–45.

Community 10: Health and Behavioral Organization Community

This community concerns itself with behavior influencing health and the social organization of groups. It brings together scholars studying animal populations (e.g. Fewell 2003; Krause, Lusseau, and James 2009; Wey et al. 2008) and human populations (e.g. Smith and Christakis 2008; Valente et al. 1997; West et al. 1999). This community acknowledges the importance of the ego-centric network approaches and sociological network literature, referencing works such as Marsden (1987) and Fischer (1982; 1977) (Smith and Christakis 2008:406), the theoretical arguments by the Relational Sociology Community (Steglich, Snijders, and Pearson 2010:330) and the structural holes theories of Burt (1991) (West et al. 1999:634). However, the distinctive feature of the methods and theories developed by this community is their emphasis on global networks dynamics and properties instead of local (ego) network approaches. They collaborate with and reference scholars of the Complex Networks Community and adopt the maxim that “...networks have emergent properties not explained by the constituent parts and not present in the parts (Watts 2004). Understanding such properties requires seeing whole groups of individuals and their interconnections at once.” (Smith and Christakis 2008:407–8; see also Fewell 2003 and Steglich et al. 2010). Granovetter’s Strength of Weak Ties are seen as an example of how seemingly weak interactions can, embedded in the aggregate global network typology, have significant effect on people’s health (e.g. West et al. 1999:642; Wey et al. 2008:335).

	Name	In-degree local	In-degree global	Research areas
1	Thomas W. Valente	68	184	Anthropology, Public, Environmental & Occupational Health, Sociology
2	Tom A.B. Snijders	51	168	Sociology, Anthropology
3	Kirsten P. Smith	45	110	Sociology
4	Ferenc Jordan	38	40	Zoology, Behavioral Sciences
5	Weiwei Shen	38	40	Zoology, Behavioral Sciences
6	Tina Wey	38	40	Zoology, Behavioral Sciences
7	Daniel T. Blumstein	38	40	Behavioral Sciences, Zoology, Environmental Sciences & Ecology
8	Susan C. Watkins	36	51	Sociology, Demography, Public, Environmental & Occupational Health
9	Gabriel Ramos-Fernandez	35	38	Zoology, Behavioral Sciences, Environmental Sciences & Ecology
10	John N. Newton	35	37	Public, Environmental & Occupational Health, Biomedical Social Sciences

Table 8: An overview of the scholars of community 10 that are most often referenced within their community (in-degree local). The table also lists how often they are referenced in the entire network (in-degree global), and their main research areas, ordered by their relevance in this dataset. These stats are based on the data from Web of Science.

References

Burt, Ronald S. 1991. *Structural Holes*. Cambridge, MA: Harvard University Press.

Fewell, Jennifer E. 2003. “Social Insect Networks.” *Science* 301(5641):1867–71.

Fischer, Claude S. 1982. *To Dwell Among Friends: Personal Networks in Town and City*. Chicago: University of Chicago Press.

Fischer, Claude S., Robert Max Jackson, C. Ann Stueve, Kathleen Gerson, Lynne McCallister Jones, and Mark Baldassare. 1977. *Networks and Places: Social Relations in the Urban Setting*. New York: Free Press.

Krause, Jens, David Lusseau, and Richard James. 2009. “Animal Social Networks: An Introduction.” *Behavioral Ecology and Sociobiology* 63(7):967–73.

Marsden, Peter V. 1987. “Core Discussion Networks of Americans.” *American Sociological Review* 52(1):122.

Smith, Kirsten P. and Nicholas A. Christakis. 2008. “Social Networks and Health.” *Annual Review of Sociology* 34(1):405–29.

Steglich, Christian, Tom A. B. Snijders, and Michael Pearson. 2010. “Dynamic Networks And Behavior: Separating Selection From Influence.” *Sociological Methodology* 40(1):329–93.

Valente, Thomas W., Susan C. Watkins, Miriam N. Jato, Ariane Van Der Straten, and Louis-Philippe M. Tsitsol. 1997. “Social Network Associations with Contraceptive Use among Cameroonian Women in Voluntary Associations.” *Social Science & Medicine* 45(5):677–87.

West, Elizabeth, David N. Barron, Juliet Dowsett, and John N. Newton. 1999. “Hierarchies and Cliques in the Social Networks of Health Care Professionals: Implications for the Design of Dissemination Strategies.” *Social Science & Medicine* 48(5):633–46.

Wey, Tina, Daniel T. Blumstein, Weiwei Shen, and Ferenc Jordán. 2008. “Social Network Analysis of Animal Behaviour: A Promising Tool for the Study of Sociality.” *Animal Behaviour* 75(2):333–44.

Community 11: Regional Economic Growth Community

This community of scholars operates in the field of economic geography, studying the role of strong and weak ties to explain the behavior of economic actors and their capabilities for innovation. This literature has a strong spatial focus, building on the notion that regional clusters and agglomerations are important sites for economic growth and innovation (e.g. Bathelt, Malmberg, and Maskell 2004; Gordon and McCann 2000; Hauser, Tappeiner, and Walde 2010). One of the most referenced publications, by Torre and Rallet (2005), is fully dedicated to detailing the differences between the notions of proximity and localization. In line with the Organizational Advantage Community (community 1), to which this community is strongly linked in the diffusion network (see Figure 1), scholars agree that knowledge creation is crucial for innovation, particularly with the rise of the knowledge based economy (e.g. Hauser et al. 2010). Weak ties are therefore seen as conduits for information exchange and knowledge creation that can be used strategically by firms selecting partners for establishing or strengthening knowledge pipelines, see for example Bathelt, Malberg and Maskell "Alternatively, firms can also scan their environment through a mobilization of 'weak ties' (Granovetter, 1973) or use regular conventions and trade fairs to establish contact with potential partner which they have known through former such events." (2004:44). The absence of weak ties is identified as one of the explanations for organizations' limited adaptation capabilities (Hannan and Freeman 1977). In addition to his contibutions on weak ties, Granovetter's work on embeddedness (1983) takes a central position in this community's literature positing that "geographical proximity is not so much an economic cause of agglomeration as a social effect of the embeddedness of economic relations" (Torre and Rallet 2005:52). In line with this community's focus on strategic networking and developing businesses, it has strong ties with the Enterpreneurial Networking Community (community 3), see Figure 1.

	Name	In-degree local	In-degree global	Research areas
1	Harald Bathelt	80	154	Geography, Business & Economics, Environmental Sciences & Ecology
2	Peter Maskell	77	132	Geography, Environmental Sciences & Ecology, Urban Studies
3	Anders Malmberg	77	132	Geography, Environmental Sciences & Ecology, Urban Studies
4	Philip McCann	44	71	Business & Economics, Geography, Environmental Sciences & Ecology
5	Ian R. Gordon	36	52	Business & Economics, Geography Environmental Sciences & Ecology
6	Michael T. Hannan	29	75	Sociology
7	John Freeman	29	75	Sociology
8	André Torre	26	43	Environmental Sciences & Ecology, Business & Economics, Geography
9	Alain Rallett	26	43	Business & Economics, Environmental Sciences & Ecology, Geography
10	Tomi Tura	22	34	Environmental Sciences & Ecology, Geography, Urban Studies

Table 9: An overview of the scholars of community 10 that are most often referenced within their community (in-degree local). The table also lists how often they are referenced in the entire network (in-degree global), and their main research areas, ordered by their relevance in this dataset. These stats are based on the data from Web of Science.

References

Bathelt, Harald, Anders Malmberg, and Peter Maskell. 2004. "Clusters and Knowledge: Local Buzz, Global Pipelines and the Process of Knowledge Creation." *Progress in Human Geography* 28(1):31–56.

Gordon, Ian R. and Philip McCann. 2000. "Industrial Clusters: Complexes, Agglomeration and/or Social Networks?" *Urban Studies* 37(3):513–32.

Granovetter, Mark. 1983. "The Strength of Weak Ties: A Network Theory Revisited." *Sociological Theory* 1:201.

Hannan, Michael T. and John Freeman. 1977. "The Population Ecology of Organizations." *American Journal of Sociology* 82(5):929–64.

Hauser, Christoph, Gottfried Tappeiner, and Janette Walde. 2010. "The Learning Region: The Impact of Social Capital and Weak Ties on Innovation." *Regional Studies* 41(1):75–88.

Torre, André and Alain Rallet. 2005. "Proximity and Localization." *Regional Studies* 39(1):47–59

Community 12: Education and Technological Innovation Community

This community of scholars studies collaborations between and within organizations from a policy perspective. Part of the literature operates in the niche of social networks in education, and includes studies on teacher collaborations (Coburn and Russell 2008), scientific collaborations (Cummings and Kiesler 2005) and school leadership structures and performance (Friedkin and Slater 1994). Correspondingly, this community contains the highest amount of scholars in the field of education and education research, see Figure 7. Yet, another line of research, typically by computer scientists and business & economists, see figure 7, focuses on the more technical aspects of these collaborations, studying electronic referral chains or other electronic communication networks (Cho et al. 2005; Constant, Sproull, and Kiesler 1996). One of the more highly cited works is a non-academic article that introduces the

ReferralWeb, a websystem for reconstructing and searching social networks on the Web (Kautz, Selman, and Mehul 1997). The distinguishing feature of the literature in this community is that it less theoretical and more geared towards implications for policy or practical applications and technology for organizations. Granovetter's 1973 paper is referenced in the context of collaboration and information dissemination on the work floor, see for example "People tend to be motivated to share information, and provide each other with early, frequent access to resources available *within* their initial social circle (Granovetter, 1973; Krackhardt, 1992). In other words, people seek information that is the most easily accessed (such as asking co-workers), rather than searching for the best information (O'Reilly 1982)." (Cho et al. 2005:438).

	Name	**In-degree local**	**In-degree global**	**Research areas**
1	Sara Kiesler	59	177	Business & Economics, History & Philosophy of Science
2	David Constant	43	148	Business & Economics
3	Noah E. Friedkin	42	306	Sociology, Anthropology, Mathematics
4	Sue Newell	36	68	Business & Economics, Information Science & Library Science, Computer Science
5	Jacky Swan	26	51	Business & Economics, Social Sciences
6	Mehul Shah	24	25	Computer Science
7	Henry Kautz	24	25	Computer Science
8	Bart Selman	24	25	Computer Science
9	Jennifer L. Russell	23	32	Education & Educational Research

10	Cynthia Coburn E.	23	32	Education & Educational Research

Table 10: An overview of the scholars of community 12 that are most often referenced within their community (in-degree local). The table also lists how often they are referenced in the entire network (in-degree global), and their main research areas, ordered by their relevance in this dataset. These stats are based on the data from Web of Science.

References

Cho, Hichang, J. S. Lee, Michael Stefanone, and Geri Gay. 2005. “Development of Computer-Supported Collaborative Social Networks in a Distributed Learning Community.” *Behaviour and Information Technology* 24(6):435–47.

Coburn, Cynthia E. and Jennifer Lin Russell. 2008. *District Policy and Teachers’ Social Networks*. Vol. 30.

Constant, David, Lee Sproull, and Sara Kiesler. 1996. “The Kindness of Strangers: The Usefulness of Electronic Weak Ties for Technical Advice.” *Organization Science* 7(2):119–35.

Cummings, Jonathan N. and Sara Kiesler. 2005. “Collaborative Research across Disciplinary and Organizational Boundaries.” *Social Studies of Science* 35(5):703–22.

Friedkin, Noah E. and Michael R. Slater. 1994. “School Leadership and Performance: A Social Network Approach.” *Sociology of Education* 67(2):139–57.

Granovetter, Mark S. 1973. “The Strength of Weak Ties.” *American Journal of Sociology* 78(6):1360–80.

Kautz, Henry, Bart Selman, and Shah Mehul. 1997. “Referral Web: Combining Social Networks and Collaborative Filtering.” *Communications of the ACM* 40(3):63–65.

Section C: Temporal Community Detection

We detected temporal communities by using the algorithm of Mucha et al (2010) implemented by Vincent Traag in the Louvain Python package. To compare results with the 2017 community grouping used in the paper, we selected a temporal grouping in which the configuration of hubs in the 2017 time slice matches the original, static 2017 grouping. While there are slight differences, overall the communities are very similar. Figure 1 shows how the communities from the static 2017 grouping map onto the temporal 2017 communities. It shows groupings are 76% similar for the largest 12 communities and 81% similar for the largest 3 communities. The biggest differences are found in the smaller communities. In particular, the temporal community detection algorithm identifies a community (community 12) of organizational scholars with Michael T. Hannan, John H. Freeman, David Strang and Sarah A. Soule as important figures. In the static 2017 network, these scholars are grouped into other communities (Hannah and Freeman in community 11, Strang and Soule in community 3). We can see in the evolution of communities (figure 2) that this community (12) emerged early and persisted, with both Freeman and Hannah's paths being: 97>30>12>12>12.

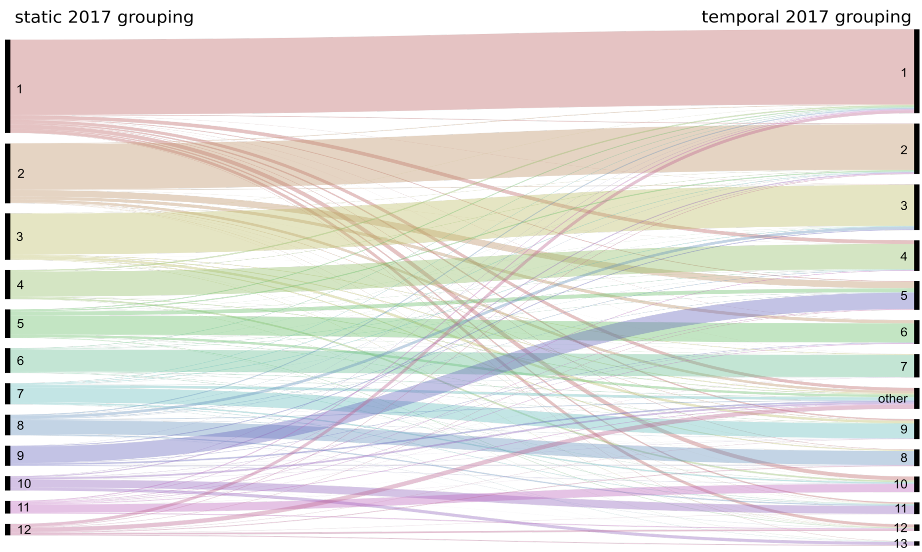


Figure 1 (previous page): Differences between the community structure found in the static 2017 Louvain detection (left) and the temporal community detection slice of 2017 (right). The groupings are 76% similar, while the largest 3 communities are 80% similar.

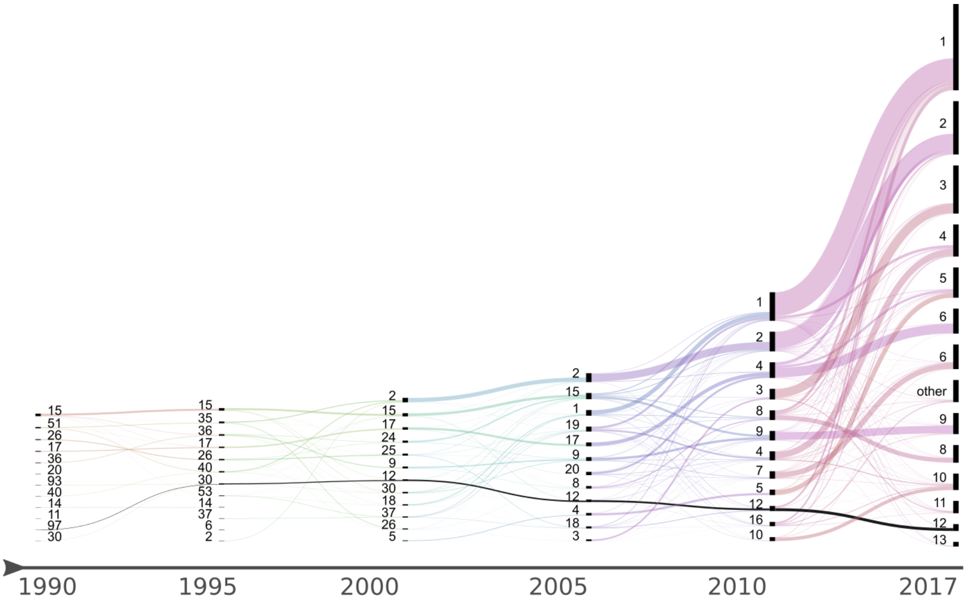


Figure 2: Evolution of communities. The emergence of community 12 is highlighted.

References

Mucha, Peter J., Thomas Richardson, Kevin Macon, Mason A. Porter, and Jukka Pekka Onnela. 2010. "Community Structure in Time-Dependent, Multiscale, and Multiplex Networks." *Science* 328(5980):876–78.

Appendix for "Intersectionality on the Go: The Diffusion of Black Feminist Knowledge across Disciplinary Boundaries and Geographical Borders"

Section A: Improving the quality of the Scopus data

Although Scopus has broad coverage, important publications on intersectionality are not part of its database (cf. Mongeon, & Paul-Hus, 2016). To ensure their inclusion, we did the following. We selected all publications (n=170) referenced at least 30 times by publications in our Scopus dataset. These can be considered part of the intersectionality canon; we therefore included those not yet in our dataset, which resulted in 96 extra publications. We also noticed that Scopus sometimes enters the same publication twice. This can happen when a publication is referenced in different styles or when the reference contains small mistakes. We selected the most cited 100 publications and corrected their mistakes manually.

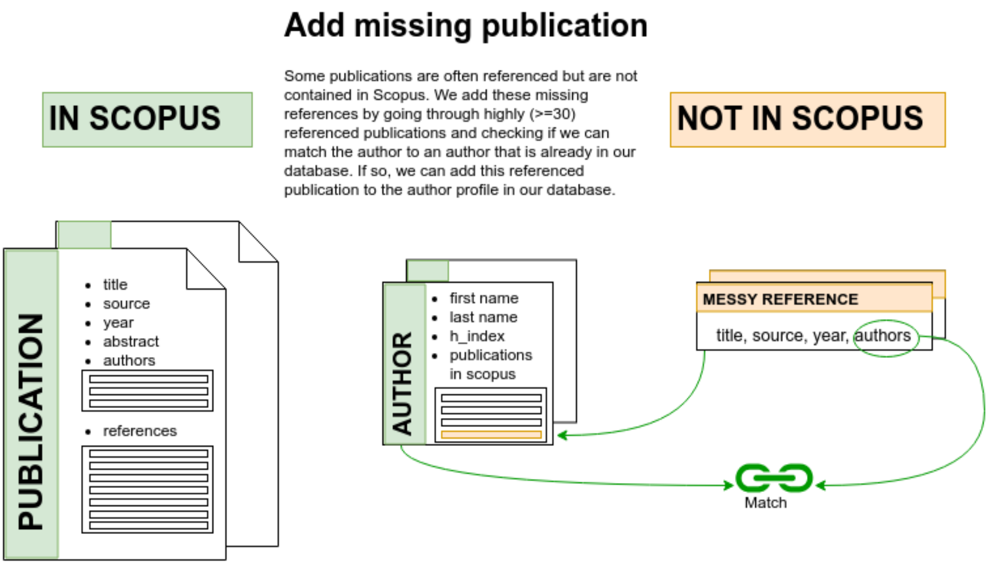


Figure 1: Schematic visualization of our procedure for adding publications that were not included in the Scopus collection.

Section B: Adjusted Havel Hakimi graph as benchmark

To test whether the community structure of our diffusion network is significant, we need to compare its modularity with a plausible benchmark. Since the structure of any network—particularly those with an uneven degree distribution—will have some modularity, finding a plausible benchmark is essential. For this, we use an adjusted Havel Hakimi graph (Hakimi, 1962; Kleitman, & Wang, 1973; Keuchenius et al., forthcoming). We compare the modularity of our diffusion network to the average modularity of 10,000 adjusted Havel Hakimi networks with an identical degree sequence. We treat reciprocal and singular links separately and match their degree sequences when creating our adjusted Havel Hakimi graphs. This is necessary as our network, in contrast to the regular Havel Hakimi graph, has few reciprocal links. By design and logic of the diffusion network, earlier links will not be reciprocated later (only scholars who publish on intersectionality for the first time in a co-authored publication will have reciprocal links).

The adjusted Havel Hakimi graph serves as a benchmark for our network as it represents the hypothesis that the network's structure is the product of first-mover advantage (Newman, 2009). The Havel Hakimi network thus captures how a scientific diffusion network would be structured were it organized only around its central actors, without scientific communities playing any role in the diffusion process.

Section C: Topic modelling

We used an LDA topic model[23] (Blei, Ng, & Jordan 2003; Pritchard, Stephens, & Donnelly 2000) to reduce the complexity of our data, which consists of the abstracts of 2,827 scientific publications. LDA is an unsupervised model that identifies topics—consisting of frequently co-occurring words—in documents. Documents can contain numerous topics while the same word can belong to multiple topics. In addition to a

[23] We used LDA as implemented in python scikit-sklearn.

standard list of English stopwords,[24] we composed a stopword list specific to our corpus including words such as: paper, result, finding, argue, relationship, literature, different, examine, investigate and intersectionality. We instructed the model to find exactly 15 topics.[25] We also instructed the model to omit words that occur in less than 30 documents, or in more than half of them. Table 1 displays the most important 30 words for the 15 topics. We then labelled these topics to capture their essence.

Table 1: Top 30 words for the 15 topics detected by LDA

Topic: Label	**Words**
0: Measuring Effects of Ethnicity	ethnic; racial; minority; ethnicity; children; race; racial ethnic; race ethnicity; differences; white; data; status; age; minorities; inequality; ethnic minority; high; stereotypes; child; school; discrimination; black; effects; parents; representation; non; compared; models; outcomes; majority
1: Violence & Law	violence; law; discrimination; legal; domestic; citizenship; state; rights; intimate; criminal; victims; domestic violence; policy; court; cases; equality; crime; indigenous; case; european; framework; partner; race; structural; control; policies; gendered; india; claims; immigrant
2: Stigmatization	health; care; stigma; hiv; mental; health care; mental health; structural; services; discrimination; experiences; healthcare; treatment; data; people; methods; factors; barriers; cultural; multiple; outcomes; interventions; needs; conducted; support; life; populations; living; characteristics; access

[24] We used a combination of the python scikit-learn English stopwords list and the python NLTK stopwords list.
[25] Setting the number of topics is a contested issue in the literature on topic modeling. We do not argue that 15 is the optimal number for our dataset. We explored the results with more and fewer topics, and found similar patterns for correlations between topics and diffusion communities as well as for topic similarity between communities.

Topic: Label	Words
3: Class	class; race; race class; black; work; white; identity; sexuality; middle; ethnicity; racial; experiences; middle class; intersections; working; family; identities; whiteness; intersection; categories; race ethnicity; cultural; class race; working class; shape; ethnicity class; status; people; color; location
4: Disability	disability; states; united; united states; racial; race; psychological; black; disabled; racism; discrimination; disabilities; gendered; people; identity; differences; oppression; psychology; centrality; role; science; work; racial identity; stigma; history; reported; intersection; physical; education; bias
5: Movements of Feminism	feminist; political; feminism; rights; politics; critical; concept; movement; power; human; scholarship; movements; essay; activism; contemporary; racism; work; sexuality; human rights; history; scholars; difference; western; feminists; field; world; oppression; important; colonial; critique
6: Policy	political; equality; diversity; theoretical; inequalities; feminist; critical; politics; policy; issues; approaches; inequality; sociology; methodological; articles; development; race; special; psychology; global; scholarship; anti; attention; current; debates; perspectives; context; economic; particular; work
7: African-American Experience	american; african; african american; mothers; experiences; narrative; experience; interviews; black; asian; americans; work; qualitative; community; motherhood; states; family; united; united states; personal; support; african americans; cultural; lived; narratives; lives; participants; structured; color; immigrants
8: Categories	power; work; categories; life; processes; concept; migration; migrant; identity; cultural; practices; perspective; gendered;

Topic: Label	Words
	belonging; focus; process; context; migrants; structures; theoretical; purpose; professional; everyday; order; inequality; empirical; masculinity; experiences; complex; value
9: Education	black; students; white; education; race; female; higher; leadership; engineering; male; educational; experiences; college; academic; higher education; color; faculty; school; racial; support; student; university; perceived; diversity; career; data; black white; asian; effects; significant
10: Labour Market	young; girls; work; workers; age; market; people; employment; labour; experiences; gendered; labor; young people; muslim; data; policies; migration; interviews; labour market; discourses; body; family; female; national; explores; role; context; experience; public; physical
11: Sexual Orientation	sexual; gay; orientation; identity; lesbian; bisexual; transgender; hiv; sexual orientation; discrimination; lgbt; experiences; queer; minority; people; identities; gay bisexual; participants; sex; lesbian gay; color; sexual minority; individuals; trans; risk; self; heterosexual; lgbtq; stress; bisexual transgender
12: Multiple Identities	identity; identities; experiences; students; multiple; cultural; religious; teaching; participants; multiple identities; issues; self; interviews; individuals; development; framework; collective; learning; student; action; challenges; education; individual; pedagogy; provide; intersecting; lgbtq; role; classroom; privilege
13: Public Health Inequalities	health; policy; discrimination; multiple; risk; inequalities; public; immigrant; framework; factors; disparities; data; status; public health; inequality; income; canada; family; model; models; determinants; policies; population; low; inequities;

Topic: Label	Words
	equity; socioeconomic; health disparities; disadvantaged; immigration
14: Justice	justice; community; communities; south; work; oppression; practice; environmental; privilege; making; rural; power; reproductive; cultural; latina; case; urban; africa; inclusive; south africa; critical; culture; indian; city; experiences; public; issues; framework; professionals; practices

Section D: Canon analysis

Four publications are referenced by more than 1,400 unique scholars. The fifth most frequently referenced publication is far behind, referenced by 698 unique scholars. We therefore consider the top four as the canon for intersectionality scholars: “Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory and antiracist theory” by Crenshaw (1989), “Mapping the margins: Intersectionality, identity politics and violence against women of color” also by Crenshaw (1991), *Black feminist thought: Knowledge, consciousness and the politics of empowerment* by Collins (1990) and “The complexity of intersectionality” by McCall (2005). Although all 12 communities reference this canonical work, they do so to varying degrees. Table 2 gives an overview of the centrality of these works in the 12 communities.

		percentage			
	publication	Collins (1990)	Crenshaw (1989)	Crenshaw (1991)	McCall (2005)
community	community size				
0	**798**	42 %	33 %	30 %	9 %
1	**797**	14 %	22 %	19 %	37 %
2	**453**	24 %	17 %	16 %	9 %
3	**393**	25 %	10 %	20 %	28 %
4	**349**	21 %	14 %	22 %	21 %
5	**337**	14 %	13 %	17 %	17 %
6	**260**	29 %	20 %	30 %	27 %
7	**214**	18 %	32 %	15 %	13 %
8	**188**	28 %	28 %	25 %	17 %
9	**179**	22 %	30 %	15 %	4 %
10	**146**	15 %	15 %	27 %	23 %
11	**122**	24 %	27 %	18 %	15 %

Table 2: Percentage of scholars in each community referencing the canonical works of Collins, Crenshaw and McCall in their first publication on intersectionality.

APPENDIX FOR "WHY IT IS IMPORTANT TO CONSIDER NEGATIVE TIES WHEN STUDYING POLARIZED DEBATES: A SIGNED NETWORK ANALYSIS OF A DUTCH CULTURAL CONTROVERSY ON TWITTER"

Section A: Issue Sentiment Classification

To analyze the debate on Black Pete, we want to know whether tweets in our dataset express a pro, anti, or neutral stance towards Black Pete. This *issue sentiment* is different from the general sentiment of the tweet, since a tweet with a negative tone of voice can be expressing a (positive) pro Black Pete statement (e.g. 'Makes me so angry hea?!!!! Hands off of our tradition!!!! Let black pete be black!!!! Don't make children cry!!!!! '[26]). Similarly, a tweet with positive tone of voice can be expressing a (negative) anti Black Pete sentiment (e.g. 'Infinite respect for everyone that fought today for an inclusive sinterklaas celebration for everyone ❤️ #KickOutZwartePiet #kozp'[27]. This implies that existing sentiment analysis algorithms cannot be used to identify the issue sentiment expressed in tweets. We therefore tailored a solution for the task at hand. This solution, a form of supervised machine learning, consists of two steps. First, we first manually classified the issue sentiment for roughly 5.300 unique tweets (2.7% of all unique tweet text) and manually identified the stance of the top prominent users. Second, we used these data as input for training a machine learning algorithm to classify the issue sentiment of the unlabeled tweets.

The issue sentiment of 5.300 unique tweets was labeled manually with the assistance of four fluent Dutch speakers. Each tweet was assigned one label: pro, anti, neutral or ambiguous. The codebook instructions were conservative: if the issue sentiment is not self-evident, the tweet was

---

[26] This tweet was originally in Dutch and the link was here omitted for brevity: "https://t.co/vWND8ngJnT Word daar zo boos van he ?!!!!! Blijf van onze traditie!!!!! Laat zwarte piet zwart blijven!!!!! Laat kindjes alle sinds niet huilen!!!!!"
[27] This tweet was originally in Dutch: "Oneindig respect voor iedereen die vandaag streed voor een inclusief sinterklaasfeest voor iedereen ❤️ #KickOutZwartePiet #kozp"

labeled as ambiguous. Table 1 below lists some examples of pro, anti, neutral, and ambiguous tweets in the data. The inter-coder agreement, measured by a Krippendorf Alpha of 0.72, was substantial. From the coding efforts we learned that it was often difficult to distinguish neutral from ambiguous tweets and we found few tweets (n=512) that were coded as expressing a neutral issue sentiment. Therefore, for further classification purposes, we merged the neutral and ambiguous tweets into one category. This manual classification identified 58% tweets with pro Black Pete sentiment, 13% tweets with anti-Black Pete sentiment and 30% tweets with neutral/ambiguous issue sentiment.

In addition to labeling this subset of tweets, we also manually labeled users' issue stance for the users that belong to the 1% top-most frequently retweeted, mentioned and tweeting users by analyzing their tweets and their role in the media and the offline debate. Again, our codebook was conservative: if the stance of the user was not self-evident, the user was assigned an ambiguous stance towards Black Pete. Subsequently, the tweets of clear pro and anti-users (n=732) were classified as pro and anti respectively, increasing our classified tweet data to a total of 51.014 unique tweets.

Next, we applied a preprocessing pipeline to the labeled and unlabeled tweets to clean the data and prepare features for the machine learning. This pipeline consisted of: shortening urls to their main domain; transforming emoticons and emojis to strings; removing single characters, multiple spaces and linebreaks; substituting the @-sign followed by a username with 'at_' and the #-sign followed by a username with 'hashtag_'; and substituting '…' for 'dotdotdot' (since this has a particular meaning). We do not tokenize the text since this step is already integrated into the fastText algorithm.

To train the classifier, we split all labeled data into a training set (70%) and a test set (30%), ensuring that the test set contained no duplicates of the training set. Thereafter, we downsampled the training set to 3.000 pro, 3.000 anti, and 2.000 neutral tweets (the total of neutral tweets available) to ensure the classifier would not be biased towards a pro or anti sentiment.

Using the fastText algorithm and the labeled training data, a classifier was trained to classify the issue sentiment of tweets by maximizing the F1 score for all classes, thus attempting to predict all classes well, in both precision and recall. The fastText algorithm gives an indication of how certain the classification is (the softmax probability), valued between 0 and 1 for each prediction. We use this certainty indication to apply a simple rule: classify all tweets with lower certainty (<0.9) as neutral/ambiguous. This procedure reduces the recall for the pro- and anti-classes but also, more importantly, reduces the errors we care most about: classifying pro tweets as anti and classifying anti tweets as pro.

The classifier—after applying the certainty rule—categorizes the issue sentiments with sufficiently high accuracy; see Figure 1. There are only 28 cases in which an anti-tweet is misclassified as pro (0.012 times of all anti tweets and 0.16 times all pro tweet classifications) and 80 cases in which a pro tweet is misclassified as anti (0.007 times of all pro tweets and 0.13 times of all anti classifications). Classifying the full dataset, we find 15% anti tweets (n=65.314), 48% pro tweets (n=225.856) and 38% tweets with neutral/ambiguous issue sentiment (n=176.327).

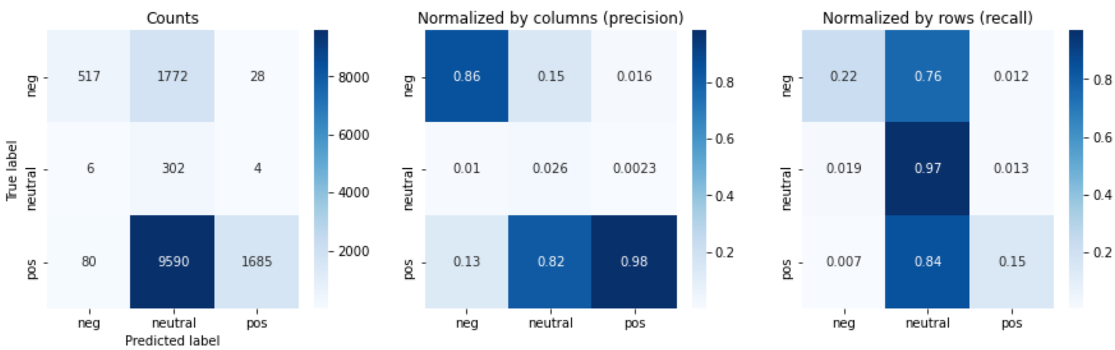


Figure 1: The results of the classifier (parameter values: epoch=10, learning rate=0.7, n-grams=3) after applying the simple certainty rule (neutral if certainty < 0.9): confusion matrix with counts (left), normalized by the true labels (middle) and normalized by the predicted labels (right). The values in the diagonals of the middle matrix are the precision rates, and the values on the diagonals of the right matrix are the recall rates. Recall rates here are reduced due to the certainty rule, but the most

important errors (classify positive if true value is negative and classify negative if true value is positive) are reduced.

Tweet text Dutch	Tweet text English (author's translation)	Labeled/ predicted
Pro sentiment		
Een grote meerderheid van Nederlanders kun je ook kwetsen @albertheijn Wij hebben ook gevoel! Stop met de afbraak van Nederlandse tradities zoals #zwartepiet en #Kerst. U verdient een #BoycotAH Twitteraars boos: geen 'kerst' maar 'winter' https://t.co/zJ065VbUdR via @telegraaf	You can hurt a large majority of the Dutch too @albertheijn We have feelings too! Stop with the destruction of Dutch traditions such as #blackpete and #Christmas. U deserve a #BoycotAH Twitterers angry: no 'christmas' but 'winter' https://t.co/zJ065VbUdR via @telegraaf	labeled
@LodewijkA Het enige doel van #KOZP is kinderen terroriseren. @LodewijkA heeft blijkbaar ook een hekel aan kleine kinderen.	@LodewijkA The only aim of #KOZP is to terrorize children. @LodewijkA apparently hates small children too.	predicted
Yak Laat zwarte piet gewoon zwarte piet blijven.!! Handen af van een kinderfeest.!!	Yikes Let black pete be black pete.!! Hands off of a childrens' party.!!	predicted
@SylvanaSimons en die andere lamlul @TheRebelThePoet Hier een boodschap voor jullie #ZwartePiet https://t.co/e8n9f1PQXx	@SylvanaSimons and the other dump TheRebelThePoet Here a message for you #BlackPete https://t.co/e8n9f1PQXx	predicted
@NadiaBouras Huiliehuilie. Zwarte piet blijft toch!!!	@NadiaBouras Cry baby. Black pete stays anyway!!!	predicted
Anti sentiment		
Zwarte Piet is racisme.	Black Pete is racism	predicted

Wat een raar frame? Anti-pietbetogers raakten in Rdam & Eindhoven niet 'slaags' met omstanders maar werden belaagd en bedreigd door honderden hooligans. In andere steden werden die opgepakt, niet de vreedzame activisten @NOS @Teletekst #falsebalance #kozp https://t.co/vx9YPtoIVr	What a weird frame? Anti-black pete supporters didn't get into a fight in Rdam & Eindhoven with bystanders but were attacked and threatened by hundreds of hooligans. In other cities these were arrested, but not the peaceful activists @NOS @Teletekst #falsebalance #kozp https://t.co/vx9YPtoIVr	labeled
Wierd Duk fabriceert leugens over #KOZP in de krant, en een paar dagen later wordt #KOZP belaagd door hooligans en neonazi's. Trots op jezelf, @wierdduk? https://t.co/kvDhA7O8nv	Wierd Duk creates lies about #KOZP in the paper, and a few days later #KOZP is attacked by hooligans and neonazi's. Are you proud, @wierdduk? https://t.co/kvDhA7O8nv	labeled
De tirannie van de meerderheid wint in meer en meer steden. Gesteund door @Politie, @MinPres, politici en burgemeesters. Democratie? Vrijheid van meningsuiting? Gelijkheid in grondrechten? Waar? Nederland, je lelijkste gezicht is nu goed zichtbaar. #KOZP #ZwartePietIsRacisme https://t.co/ik7bFQYSnL	The tyranny of the majority wins in more and more cities. Supported by @Police, @Minpres, politicians and majors. Democracy? Freedom of speech? Equality in rights? Where? Netherlands, your most ugly face is now clearly visible. #KOZP #BlackPeteIsRacsim https://t.co/ik7bFQYSnL	labeled
#NeemAfstandVanWierdDuk #KOZP https://t.co/xG4Y4qGMMM	#DistanceYourselfFromWierdDuk #KOZP https://t.co/xG4Y4qGMMM	predicted
Hoi @jennydouwes @wierdduk @geertwilderspvv - Minder mensen kiezen voor traditionele Zwarte Piet | NOS https://t.co/N5hbaVjTmw	Hoi @jennydouwes @wierdduk @geertwilderspvv – Less and less people are choosing for the traditional Black Pete | NOS https://t.co/N5hbaVjTmw	predicted
Ambiguous or neutral sentiment		

Beste pro- én anti zwarte piet demonstranten... HOU EENS OP MET DAT GEZEIK EN GA GEWOON EEN GEZELLIGE SINTERKLAASPERIODE MAKEN!	Dear pro- and anti black piet protesters... PLEASE STOP YOUR WHINING AND JUST GO HAVE A NICE SINTERKLAAS SEASON!	predicted
Sinterklaas is weer in het land. #weekend #Sinterklaas #sinterklaasintocht #Zwarte Piet https://t.co/vjyJ4oT88q https://t.co/wDnxyrJMoY	Sinterklaas is back in the country. #weekend #Sinterklaas #sinterklaasintoch #Black Pete https://t.co/vjyJ4oT88q https://t.co/wDnxyrJMoY	predicted
DENK wil Zwarte Piet de nek omdraaien https://t.co/qciqfh8nOw	DENK wants to kill Black Pete too https://t.co/qciqfh8nOw	predicted
@Boargemaster @DilanYesilgoz Onzin, die uitspraak kwam helemaal niet vanuit de KOZP hoek	@Boargemaster @DilanYesilgoz Nonsense, that statement didn't come from the KOZP at all	predicted
in NL op straat tegen idioot hoge brandstofprijzen ??? nee man op straat voor of tegen zwarte piet.... massaal JA #prioriteiten lekker op n rijtje NL..... #koekoek	in NL on the streets against idiotic high fuel prices ???? No man in the street for or against black pete...massive YES #priorities straight in NL.... #cuckoo	predicted
De discussie over Zwarte Piet kabbelt voort, maar het draagvlak brokkelt langzaam af https://t.co/WS33B2aReJ via @volkskrant	The discussion about Black Pete ripples on, but support is slowly crumbling https://t.co/WS33B2aReJ via @volkskrant	labeled

Table 1: Examples of tweets expressing a pro, anti or neutral/ambiguous sentiment on Black Pete. The column predicted/labeled indicates whether this tweet was classified by manual coding (labeled) in our dataset or predicted by the classifier (predicted).

Appendix for "Beyond echo chambers: the role of conflict in social media polarization"

Section A: Twitter word embedding

To infer the sentiment of a mention (endorsement, opposition or neutral/ambiguous) we used the fastText algorithm with pretrained word vectors constructed from a large Twitter dataset of Dutch tweets in 2018. We use a word embedding based on a Twitter corpus instead of a more general corpus, such as a Wikipedia corpus, because this embedding can encode particularities of the ways users discursively express themselves on Twitter, in particular towards others. Users express sentiments towards other users on Twitter with specific words and expressions that are expected to linguistically differ from the word use and sentences found in more general corpora of Dutch language, such as Wikipedia, news articles or books.

The dataset of Dutch tweets was compiled by the Netherlands eScience Center in the scope of the TwiNL project which has as its ambition to collect all tweets on Twitter from the Netherlands (Sang & van den Bosch, 2013). Selecting the tweets from this dataset published in 2018 (n=179,789,348) and tokenizing with the tokenize package of the python nltk python library gave roughly 6 million unique tokens. Vectors with dimension 300 were created with the use of the skipgram algorithm (Bojanowski et al, 2016).

References

Bojanowski, Piotr, Edouard Grave, Armand Joulin, and Tomas Mikolov. 2017. "Enriching Word Vectors with Subword Information." *Transactions of the Association for Computational Linguistics* 5:135–46. doi: 10.1162/tacl_a_00051.

Sang, Erik Tjong Kim, and Antal Van den Bosch. "Dealing with big data: The case of Twitter." *Computational Linguistics in the Netherlands Journal* 3 (2013): 121-134.