

Dartmouth College

## Dartmouth Digital Commons

---

Dartmouth College Undergraduate Theses

Theses and Dissertations

---

Spring 5-27-2022

### **Possibility is not a Luxury: The role of identity in representations of possibility**

Hannah M. LeBaron

*Dartmouth College*, [hannah.m.lebaron.22@dartmouth.edu](mailto:hannah.m.lebaron.22@dartmouth.edu)

Follow this and additional works at: [https://digitalcommons.dartmouth.edu/senior\\_theses](https://digitalcommons.dartmouth.edu/senior_theses)



Part of the [Cognitive Science Commons](#)

---

#### **Recommended Citation**

LeBaron, Hannah M., "Possibility is not a Luxury: The role of identity in representations of possibility" (2022). *Dartmouth College Undergraduate Theses*. 266.

[https://digitalcommons.dartmouth.edu/senior\\_theses/266](https://digitalcommons.dartmouth.edu/senior_theses/266)

This Thesis (Undergraduate) is brought to you for free and open access by the Theses and Dissertations at Dartmouth Digital Commons. It has been accepted for inclusion in Dartmouth College Undergraduate Theses by an authorized administrator of Dartmouth Digital Commons. For more information, please contact [dartmouthdigitalcommons@groups.dartmouth.edu](mailto:dartmouthdigitalcommons@groups.dartmouth.edu).

**Possibility is not a Luxury:  
The role of identity in representations of possibility**

*“Feminism loves another science: the sciences and politics of interpretation, translation,  
stuttering, and the partly understood.”*  
– Donna Haraway

Hannah LeBaron  
Advised by Jonathan Phillips and Mingwei Huang  
Cognitive Science Senior Honors Thesis  
Dartmouth College  
27 May 2022

## Acknowledgments

I want to offer special thanks to my two advisors, Jonathan Phillips and Mingwei Huang without whom this project would have been impossible. They supported me through numerous unexpected obstacles and continued to motivate me when I was most challenged. Thank you for being the most exceptional advisors, mentors, and friends.

I would also like to thank Regan Bernhard for her essential mentorship as a research advisor, and the entire philLab for their feedback and support throughout this process. Over the past 3 years it was my work with you that allowed me to feel prepared to attempt such an ambitious independent research project.

Finally, I would like to thank my family and friends at Dartmouth who at various times provided every form of emotional support. I am so grateful to be embraced by a community of overflowing wisdom and compassion.

## **Abstract**

In what ways are our expectations of others and their behaviors limited by the ways they are identified? Broadly, this project asks how identity traits are considered across judgments about which social roles people can assume. This work is an interdisciplinary project that exists within the fields of both Cognitive Science and Women's, Gender, and Sexuality Studies. As such, the first chapter explores the central issues of identity conception from the perspective of feminist theory, including a discussion of category destabilization, intersectionality, and knowledge production in the context of feminist science studies. 4 empirical studies were launched with methods informed by these theory considerations. Study 1 employs a qualitative approach to ask whether there are discernible default identity heuristics being employed for certain social roles. Study 2 develops this question further but switches the focus onto what other judgments, like statistical expectations or prescriptive ideals, might actually be predictive of default identity judgements. Similarly, study 3 considers more broadly whether the associations being made could be represented as a principle, or inherent, aspect of social roles. Finally, study 4 draws upon the work of all the previous studies to investigate how our notions of default identity traits could affect our ability to perceive certain traits as more or less possible. Ultimately, this work concludes by asserting that collectively these studies offer evidence that identity traits could meaningfully limit our scope of possibilities for certain people.

## Contents

<b><i>I.</i></b>	<b><i>Introduction</i></b> .....	<b>4</b>
<b><i>II.</i></b>	<b><i>Background &amp; Theory</i></b> .....	<b>10</b>
1.	Categories and Difference .....	11
2.	Intersectionality, Social Oppression, and the work of Intersectional Analysis.....	16
3.	Epistemology and Feminist Science Studies .....	24
4.	Conclusion .....	31
<b><i>III.</i></b>	<b><i>Methods &amp; Analyses</i></b> .....	<b>33</b>
1.	Representations of Identity .....	34
2.	Study 1a: Qualitative Associations .....	35
2.1	Methods .....	36
2.2	Results .....	36
3.	Study 1b: Quantifying Categories .....	39
3.1	Methods .....	40
3.1	Results .....	41
3.2	Discussion .....	43
4.	Study 2a: What comes to mind? Ideal vs. Average.....	44
4.1	Methods .....	46
4.2	Results .....	47
4.3	Study 2b: Replication without Sampling .....	51
4.4	Results .....	51
4.5	Discussion .....	53
5.	Study 3: Statistical vs. Principled Connections.....	54
5.1	Methods .....	55
5.2	Results .....	55
5.3	Discussion .....	57
6.	Study 4: Constraints of Possibility .....	58
6.1	Methods .....	58
6.2	Results .....	59
6.3	Discussion .....	63
<b><i>IV.</i></b>	<b><i>General Discussion</i></b> .....	<b>64</b>
	<b><i>Sources &amp; References</i></b> .....	<b>76</b>
	<b><i>Supplementary Materials</i></b> .....	<b>80</b>

## I. Introduction

In what ways are our expectations of others and their behaviors limited by the ways we identify them, or the ways they choose to be identified? Broadly, this project asks how identity traits are considered across judgments about which social roles people can assume. How can we theorize the role of cognitive structures in the perpetuation of discriminations? The goal of this series of studies is to examine a critical overlap between the conception of intersectional identity and its potential impact on cognitive representations of possibility. Much of the work of women, gender and sexuality studies focuses on the analysis and deconstruction of oppressive external power hierarchies. By considering a cognitive science approach to this central goal, how can we reconsider the ways these oppressions are represented internally and what the larger impact on ourselves and our society might be? The trajectory of this sequence of empirical studies attempts to analyze a specific area of these broader considerations by exploring and interrogating the ways in which intersectional identity traits are cognitively constructed, conceptualized, and categorized as well as the ways in which identity may be able to affect our expectations of what is possible.

It feels necessary to begin by briefly situating the scope and location of this project as well as myself as its primary author. This project has been designed, implemented, and written as a senior thesis project. Access to numerous resources, invaluable mentorship, and adequate funding have all been made possible by Dartmouth College and my own privilege as an undergraduate student at an ivy league institution. Essentially, my purpose in this investigation is to consider questions about how identity representations may impact our ability to perceive possibility. This question necessarily intersects directly with numerous realities of discrimination, marginalization, and access to privilege. Given the inherent intent of this project to consider these ideas from even the most abstract or objective standpoint, it must be

acknowledged that the multiple realities of these lived experiences will never be captured by a single data point or study. While this project seeks to offer insight into some aspects of identity driven perception which may affect social hierarchies, it does not offer any generalized claims as to what this experience of marginalization can or should be universally understood as.

I want to also situate this work as an interdisciplinary project that exists within the fields of both Cognitive Science and Women's, Gender, and Sexuality Studies. The research pursued throughout asks questions that cannot be relegated to one discipline or the other, and the larger purpose of the project requires the consideration of knowledge from multiple standpoints. One critical issue which arises in this academic intersection is the implementation of methods, particularly in empirical research. In so far as this work seeks to investigate and cast light on the structure of identity and its impact on behavior and cognition, the idea of constructing identity concepts as manipulated variables has implications far beyond the scope of psychology.

This project navigates an intersection between theory and empirical methodology to allow for a uniquely interdisciplinary trajectory. While the chapters are broken down to offer an initial section purely devoted to the further explanation of feminist theory, this should not be interpreted as a suggestion that theory is only relevant to this project as a form of background. In its entirety this project intends to stage a broader conversation between aspects of theory and method and the way these two perspectives interact is essential to the modes of discovery implemented here. The enmeshing of qualitative and quantitative methods as well as the general discussion of results are intended to demonstrate precisely how essential feminist theory is to this form of empirical study which specifically asks questions tied to the conception of identity.

In the following section I wish to explore the central issues of this project specifically from the perspective of feminist theory. I'm choosing to ground the methods and analysis of this



cognitive science research project in feminist theory because I think it necessarily situates the larger societal implications of the construction of identity. While parts of this section may delve into the abstract, I believe these theoretical considerations inform and supplement the more concrete aspects of this project in its entirety. Ultimately, I intend to indulge in this focused analysis of feminist theory to enrich and highlight the interdisciplinary potential of this kind of cognitive science research.

The first chapter of this work explores in depth the ways in which theory not only underlies but also fuels the entirety of this project. By beginning with destabilizing the notion of the “category” specifically as applied to identity, I hope to offer an important rejection of the category as an innately immutable object and offer space to work within this tension throughout. This section continues onto a brief summary of the history and implementation of the concept of intersectionality. The summary specifically highlights intersectionality’s roots in Black Feminism and legal analysis, but also offers suggestions for the ways it can be implemented in future projects without foreclosing on its original intention to acknowledge the unique discrimination Black women are subjected to. Finally, the theory section resolves with a close analysis of knowledge production in the context of feminist science studies. This last theoretical exploration aims to transition the implementation of theory directly onto the design of methods and offer alternative considerations for how we affirm and produce ways of knowing beyond the scope of strictly quantitative data.

The next chapter of this project offers a technical breakdown of the 4 empirical studies implemented throughout this project and each of their specific methods, analyses, and results. Study 1 employs a qualitative approach to ask whether there are discernible default identity heuristics being employed for certain social roles and if these defaults differ meaningfully. Study

2 develops this question further but switches the focus onto what other judgments, like statistical expectations or prescriptive ideals, might be predictive of default identity judgements. Similarly, study 3 considers more broadly what kind of associations are being made when identity traits are considered in the classification of social roles. Specifically, are some traits being consistently represented as a principle, or inherent, aspect of a social role? Finally, study 4 draws upon the work of all the previous studies to investigate how our notions of default identity traits could affect our ability to perceive certain traits as more or less possible when applied to certain social roles.

Ultimately one of the major limitations faced in this project was in the attempt to navigate this negotiation between theory and methods. While the studies that follow explore cognitive representations of possibility and identity, they regretfully do not shed as much light on the intricacies of intersectionality. While the importance of intersectionality is discussed at length from a theoretical standpoint, the scope of these studies was not quite broad enough to do these concepts full justice. While the preliminary studies leave intentional room for a more open-ended examination of identity traits, the final set of studies commit themselves to the examination of a few very specific representations of identity rather than a broad set of overlapping considerations. Ideally, these studies represent merely the tip of the iceberg of a larger investigation and thus could still provide valuable scaffolding for the design of future research that may be more difficult to implement but is ultimately more intentionally intersectional throughout its process.

This project closes with an overarching discussion of results across the four studies and their relevant theoretical implications. Ultimately, the project returns to a consideration of the broader questions surrounding what impact our default representations might have on the world

around us, especially when considering identity. While limitations to these studies are also acknowledged, we conclude by asserting that collectively these studies offer evidence that identity traits could meaningfully limit our scope of possibilities for certain people. This conclusion has implications beyond the purview of this project and offers potentially powerful insight into the way our beliefs have the power to deeply affect each other by prescribing limitations as well as offering opportunities.

## II. Background & Theory

## 1. Categories and Difference

Feminist theory offers powerful insight into the dilemmas that emerge when compartmentalizing aspects of one's identity into categories. The very notion of stable categories that one might employ tactically for research purposes could be considered dangerously reliant on assumptions of universality or coherence. In this section I hope to ground the work of this research in an analysis of the ways feminist theory problematizes the category and its relationship to the production of difference. Ultimately, the construction of identity as a variable begets a deliberate interrogation of the theoretical implications of the "category" itself as a structure which can be harnessed, intentionally or unintentionally, to reinforce oppression. By destabilizing the notion of the category, we seek to mobilize a perception of difference that acknowledges alternatives in a way that critically frames the broader implications of this work.

The "category" has a preexisting and prominent role throughout the history of contemporary Western feminist theory. Notably, the disruption of the idea of a coherent category allows for a re-examination of gender as it is constructed by society and in relationship to other aspects of identity. In Judith Butler's essay *Subjects of Sex/Gender/Desire*, she asserts that "woman" as both a category and a subject can no longer be politically viable for feminist theory. One of her key tactics in this argument is the explicit dismantling of each of these concepts: sex, gender, and desire, in relation to one another. While Butler's work centers around the flawed presupposition of gender as a fixed category, its implications extend beyond this notion and could be applied directly to consideration of other identity concepts including race, class, and ability. Her argument is rooted in the notion that "gender is the cultural meanings that the sexed body assumes," but that there is significant reason to be concerned that the cultural meaning of gender is stable or least of all innate (Butler 2015, 8). Butler concludes that fixing women as the

“self-evident” subject of feminism forecloses possibilities for feminist politics, including coalition building. Ultimately, the paradox of foundational feminist theory is “the assumption that the term women denote a common identity” (Butler 2015, 8). The presumption of universality is fundamentally flawed and becomes a paradox of feminism as it “is effectively undermined by the constraints of the representational discourse in which it functions” (Butler 2015, 7). Butler’s revelations transform the very assumption that feminism is “about women” by deconstructing the restricted positionality of the concept of gendered categorization.

Troubling the notion of women as the privileged subjects of feminism requires the entirety of the feminist movement and theory to reconsider its purpose. The category of “woman” is essentially in flux and Butler urges feminists to instead allow this concept to “serve as a permanently available site of contested meanings. The definitional incompleteness of the category might then serve as a normative ideal relieved of coercive force (Butler 2015, 12). While Butler’s foundational work acknowledges the contradictions of the field, it also offers a trajectory for resistance. By destabilizing identity categories such as gender, and rejecting assumptions of universality, one can “make gender trouble...through the subversive confusion, and proliferation of precisely those constitutive categories that seek to keep gender in its place by posturing as the foundational illusions of identity” (Butler 2015, 16). Thus, the work of category destabilization can be understood as a fundamental part of the active resistance of systematic oppression.

Ultimately, problematizing the “category” as a theoretical framework is only the beginning of the larger process of destabilizing categorical boundaries that articulate identity concepts through assumptions of coherence and universality. Queer theory has offered remarkable examples of what it means exactly to conceptualize this kind of instability through a

lens of shifting relations to heteronormative power. In their discussion of the notion of “Trans” as a concept that intentionally crosses and dissects borders, Susan Stryker, Lisa Jean Moore, and Paisley Currah ask us to “conceptualize gender not as an established territory but rather as a set of practices through which a potential biopower is cultivated, harnessed, and transformed” (Stryker et. al. 14). This fluid understanding of gender situates it within the perpetually evolving context of ‘potential biopower’ rather than establishing it discretely. Further, this new perspective yet again rejects the stable notion of category and offers an alternative wherein “‘Trans-’ thus becomes the capillary space of connection and circulation...[and] ‘gender’ becomes one of several sets of variable techniques or temporal practices (such as race or class) through which bodies are made to live” (Stryker et. al. 14). By “trans-”ing the category of “gender” Stryker et. al, reveals the interconnected space through which all of these terms exist and overlap: a space that is critical to the study of identity.

This notion of queerness disrupting categories extends beyond the abstract and is enacted literally as a form of lived resistance. In the seminal work of *Punks, Bulldaggers, and Welfare Queens* “the assumption that stable collective identities are necessary for collective action is turned on its head by queerness” (Cohen 439). Social change and action require us to “work toward the destabilization and remaking of our identities,” and it is this process which offers liberation from the power invested in the category itself (Cohen 48). Cohen’s work serves as a powerful reminder that coalition building is not grounded in the universality of shared oppression, but rather is a process intrinsically invested in difference and unique experience. There is no universal background of race class or gender, but difference in and of itself is not the problem. Queer collective action against stability asserts the subject beyond the scope of coherence and recognizes the root of the problem is the power invested in the category by

systems of oppression. Queer theory's rejection of the category offers alternative ways to conceptualize difference that don't rely on stable categorization, universal experience, or an assumption of coherence. Identity is understood as a biopolitical experience affected on by systems of difference, but its most natural form exists beyond these artificial restrictions.

An essential aspect of understanding the power invested in the category is the consideration of how categories are mobilized for perceiving difference. Intersectionality examines the way the "category" acts as the social structure for both perceiving and creating difference by highlighting the experiences of those who identify in the "intersection." Much like the spatial metaphor offered by queer theory in the process of "trans-ing," intersectionality considers the alternative realities between and beyond socially constructed expectations.

In *Re-Thinking Intersectionality*, Jennifer Nash probes the ways in which the institutionalization of intersectionality has affected the framework and exposes the limited nature of the theory to describe the ways in which power, privilege, and oppression form intersectional identities. The roots of intersectionality will be discussed at length later, but in this moment, it feels most productive to focus on Nash's critiques of the theory that specifically highlight the way it can be used to dissect category and difference. Specifically, she raises the issue that "intersectionality has yet to contend with whether its theory explains or describes the processes and mechanisms by which subjects mobilize (or choose not to mobilize) particular aspects of their identities in particular circumstances" (Nash 2008, 11). The mobilization of aspects of one's identity reveal the ways in which the intersection of identity can be better understood as a multiplicative rather than merely additive process. The compounding nature of intersecting experiences does not produce a clear summation, but rather elicits entirely new experiences beyond the constraints of any form of "category." For example, the way a queer black woman



experiences each of these traits is relative to every individual trait, the relationship between each pair of traits, and the largest understanding of all the traits. Now combine this with the understanding that each of these traits is affectively subjective rather than essentially categorical, and the multitude of difference is vastly exponential. Nash's reconsideration of the framework of intersectionality establishes how the category is mobilized to perceive difference across institutions of oppression.

Ultimately, a theoretical analysis of the destabilization of identity categories only extends so far. In an experimental setting, the mobilization of categories is a critical part of the empirical research process. It is necessary to consider this implementation of some form of "categorical" analysis in the context of reconciling the category's role in systematic oppression. The role of the category in this project specifically is something I want to situate in the aforementioned work of feminist theory, queer theory, and intersectionality. While this project utilizes quantitative tools to manipulate and measure responses across categories, I don't mean for this process to reaffirm the virtue of those categories as stable or universal. It is important to ground the use of variables in the theoretical engagement of exclusivity that categories can develop. However, ultimately, one can argue that we must utilize some form of analytical categories to understand the discursive conditions of systemic oppression. As Cohen expresses, there is a potential utility to the acknowledgement of categories which is why she advocates for "destabilization" rather than "deconstruction." This also acknowledges the lived reality that traditionally named categories can also foster "community as paths to survival...and shared experiences of oppression" (Cohen 460). While destabilization still advocates for coalition rooted in multiplicity and interconnectedness, identity categories remain prominent forces organizing both individuals and communities.

I would argue that in the direct process of destabilization or indirectly through research, categories must be engaged with rather than ignored. In this project I hope to work within this space of tension and consider multiple questions about the cognitive representations of identity concepts. In doing so however, I do not mean to assert that these concepts are innately immutable. Recognizing the way these categories are presented to intentionally restrict the space of self-expression is a critical step in the process of studying the ways such identity concepts are represented internally and beyond.

Both in theory and implementation the deconstruction of categories highlights the “borders” as spaces of critical overlap. Further consideration of intersectionality as a theory of analysis is an essential aspect of the work grounding this process. By returning to the roots of intersectionality in Black Feminist thought and tracking its evolution and subsequent institutionalization, I next hope to offer further insight into the ways in which feminist theory critically inform the goals of this project.

## 2. Intersectionality, Social Oppression, and the work of Intersectional Analysis

Intersectional analysis and its roots in Black Feminist Thought are essential aspects of the examination of overlapping identities and their role in social oppression. Intersectionality rose to prominence specifically through Kimberle Crenshaw’s employment of the metaphor to consider the systemic oversight of the American legal system when considering situations of both race and gender discrimination. Crenshaw offers the metaphor of intersection to elucidate the experience of living with multiple identities and existing at “a location that resists telling” (Crenshaw 1241). While the concept of intersectional oppression has been present within Black Feminist theory long before the work of Crenshaw, it cannot be denied that her implementation

of the analytical framework sparked a critical evolution of the theory. In this section I wish to introduce not only the key concepts of intersectionality, but also the development of the term as it has been popularized and institutionalized. Ultimately, in applying intersectional analysis to this project and the discipline of cognitive science it is critical to situate the theory within its roots in Black Feminism as well as its present-day politicization.

Intersectionality is a theory and method of analysis with foundations deeply rooted in Black Feminism, but which rose to epistemological prominence through a legal context. Kimberle Crenshaw introduced the term to acknowledge the state's failure to recognize the multiple oppressions experienced specifically by Black women. While Crenshaw deserves significant credit for her contributions to the theory of intersectionality, it is also important to remember the dangerous erasure perpetuated by any narrative which attempts to attribute sole credit for the "discovery" of a theory that has been actively implemented in minoritarian communities and fields well before Crenshaw. This apprehension is more eloquently ascribed by Patricia Hill Collins who asserts, "whether intentional or not, coining narratives perform a form of epistemological gatekeeping that erases and sanitizes the radical potential of intersectionality's more unruly dimensions" (Collins 152). Crenshaw's work doesn't not exist within a vacuum, rather she is building upon preexisting thought, experience, and theory which cannot be cleanly assigned a singular source. Rather than pursue an isolated site of origin, an examination of the history of intersectionality asks us to recognize that "the notion that identity is formed by interlocking and mutually reinforcing vectors of race, gender, class, and sexuality has pervaded black feminist scholarship for decades" (Nash 2008, 3). Indeed the 'unruly' dimensions of original forms of intersectional analysis have explanatory power in understanding how and when this theory was acknowledged by academic scholarship. Documentation of multiple

oppressions existed in many forms, including the political and the personal. Before this epistemological validation, the affect and lived experiences of those with mutually interlocking identities is well documented, but not necessarily as explicitly named. This can be seen through alternative pre-existing terms such as double jeopardy, interlocking oppression, or super exploitation. The specific terminology of “intersectionality” which Crenshaw introduces is thus critical to chart the trajectory of the theory but should by no means mark the ‘beginning’ of the analysis of the experience of multiple oppressions.

Crenshaw’s metaphor of the “intersection” has proved incredibly impactful for the analysis of identity. In many ways, the spatial reimagining of abstract identity concepts makes what is otherwise an embodied affective experience of oppression astonishingly concrete and explicit. While this is what makes the model so powerful, it also suggests the innate dangers that emerge when attempting to make the transcendent endurance of discrimination more theoretically coherent. Crenshaw’s metaphor develops within the context of the legal system’s failure to acknowledge multiple identities and interlocking oppressions. She highlights the well-established legal precedent where minoritarian plaintiffs were forced to argue for only one type of discrimination: race or gender.

While highlighting the intersection of these two “categories” in particular, intersectionality also actively forces us to consider exactly how we negotiate between any set of overlapping identity traits. Race and gender are not the only relevant identifiers, Nash notes “it is incumbent upon both feminist theory and anti-racist work to develop a conceptualization of identity that captures the ways in which race, gender, sexuality, and class, among other categories, are produced through each other, securing both privilege and oppression simultaneously” (Nash 2008, 10). The conceptualization of identity beyond the two traits of race and gender force us to

consider which identity traits enter into which negotiations. To draw upon Crenshaw's visual metaphor once again, which "streets" cross at each intersection? Are there moments when race and ability produce each other more significantly? Or moments when the collision of class and gender deserve special attention? How do we compensate for the fact that each facet of identity is not equal weighted, and these weights might be highly situation dependent?

Crenshaw specifically considers the flawed lens which would consider multiple forms of discrimination as entirely unrelated and parallel rather than overlapping. In her rejection of this parallel formulation, "Crenshaw attacks the assumptions underpinning anti-discrimination law, arguing that black women are compelled to assert either race-based or gender-based discrimination claims instead of causes of action, which reflect their positions as intersectional subjects" (Nash 6). Ultimately the result is targeted directly at Black women whose injuries, "because law recognizes only race- *or* gender-based injuries, ... cannot be wholly addressed by the existing doctrinal structure" (Nash 6). This method further allowed courts to overlook instances of specifically intersectional discrimination against black women by drawing faulty comparisons to either white women or black men. Thus, Crenshaw offers the analogy of the intersection:

Consider an analogy to traffic in an intersection, coming and going in all four directions. Discrimination, like traffic through an intersection, may flow in one direction, and it may flow in another. If an accident happens in an intersection, it can be caused by cars traveling from any number of directions and, sometimes, from all of them. Similarly, if a Black woman is harmed because she is in the intersection, her injury could result from sex discrimination or race discrimination. (Crenshaw 1989, 149)

The intersection analogy dimensionally structures the lived experience of the discrimination faced by Black women. However, this metaphor is intended not only to highlight this erased experience, but also serve as a framework for addressing such harm. Crenshaw's lens of intersectionality can be mobilized as a theoretical tool that helps us observe and directly address issues of systematic oppression. By mapping the marginalized experience, Crenshaw directly emphasizes the institutionalized frameworks that harm those with multiple identities. In her work "intersectionality is a way of responding to doctrinal invisibility with an insistence that law both recognize and redress black women's particular experiences" (Nash 2019, 10). Nash's re-examination of Crenshaw's work focuses on the active nature of intersectionality to actually 'redress' and mobilize to problematize race or gender dominant analysis. While the legal context Crenshaw applies here is particularly potent, the concept of intersectionality has applications beyond the law and has indeed made its mark in many other fields, particularly recently.

Tracking the influence of intersectionality exposes how it has been employed successfully and warped by institutionalization. Addressing this politics of intersectionality brings light to the ways in which intersectionality has evolved to become academically 'mainstream,' but also coopted by neoliberal movements of diversity and inclusivity. The progress narrative here is far from linear. As the concept of intersectionality has swept in to take the focus of women, gender, and sexuality studies it also diverges across other disciplines, and in the words of Nash, "to have intersectionality come to occupy the center of women's studies and to migrate across disciplinary boundaries, to be both filled with promise and emptied of specific meaning" (Nash 2019, 2). The paradox of intersectionality being embraced as a central tenet of feminist theory is the way this allowed it to transition discontinuously into other fields. In this

progression the meaning and application of the theory has dramatically shifted and unfortunately detached itself from some of the central goals envisioned by Black Feminist theorists.

While acknowledgement of the unique experience of overlapping oppression is surely a positive phenomenon across fields, the dilemma it places women of color in is duplicitous. Specifically, the repackaging of intersectionality as merely an initiative of 'diversity' is a dangerous one because intersectionality gets attached to Black Feminism and Black women's bodies, while Black women scholars see no material benefits from this incorporation. The academic institution is one clear example of a place where administrations have coopted intersectional analysis as a weaker form of passive representation. Simply put, the inclusion of more diverse identities, while a noble goal, does nothing to address or hold accountable the structures which oppress minoritarian subjects. Institutional diversity initiatives that do not hold an institution accountable are yet another example of "how women of color are rhetorically summoned as proof of the field's evolution." The tokenization of women of color yet again puts unfair burden on the already oppressed to be the physical markers of advancement. Often even once included, women of color are also disproportionately asked to do the work of accountability from within the system that oppresses them.

While the goal of diversity and inclusivity is often advertised as intersectional, the work these initiatives inspire is essentially rooted in different motivations. From a legal perspective diversity's "value lies not in its capacity to remedy past and ongoing racism and exclusion...The work of diversity, then, is not meant to transform social institutions but to insert bodies into existing structures and even to engage in 'rebranding an organization'"(Nash 2019, 4). The institutionalization of intersectionality as diversity and inclusion initiatives illustrates the drawbacks of oversimplification. Rather than reckoning with preexisting power structures, these

diversity campaigns appease them. The work of intersectionality that scholars like Crenshaw and Nash beseech us to engage in is the radical deconstruction of social institutions which are designed to inflict intersectional harm. In this larger process intersectionality is a tool with which to locate and address the source of structural oppression as a critical step in the process of dismantling the system completely.

As intersectionality is institutionalized and spreads across academic disciplines, the mixed outcomes make it ever more important to consider the roots of this theory in Black Feminism. As new fields attempt to apply and implement intersectional theory it begs the question of not only whose identities are intersectional, but also “who owns intersectionality, and who steals it?”(Nash 2019, 15). Nash asks us to seriously consider interdisciplinary studies as a new frontier in intersectional work, but its roots in Black feminism and specific formulation as a tool to highlight the discrimination faced by Black women must center its migration across fields. Whether being considered in the context of critical race theory, economic disparity, access to healthcare, or psychological experiments, the tool should be recognized as derived from a very specific position and experience of oppression. Intersectionality is often used correctively, to bring an analysis of race and or class where there was only gender, or to gesture diversity and progress. In contrast, the intended purpose of intersectionality, as elucidated by Nash, is “its investment in critiquing and responding to a set of harms that render invisible black women’s injuries” (Nash 2014, 59). In applying intersectionality this intention must be centered to preserve the framework which was specifically designed to highlight erased experiences.

So, for the purpose of this project, what does it mean to incorporate intersectional analysis into cognitive science research? Primarily, I want to at the very least acknowledge the foundational work of Black Feminists (including Crenshaw, Nash, Collins, and all those before



and after) in theorizing a structure through which one can navigate the complex relationship between identity concepts and multiple oppressions. By situating this section in a brief history of intersectionality I hope to utilize the analytical technique in a new discipline without fully unmooring the concept from Black Feminism, but while acknowledging that to apply a theory beyond the scope of the field within it was first developed will result in alterations. However, I want to be wary of the dangerous precedents of institutionalization and instead here engage with this tool as it was intended: in the examination and deconstruction of systematic oppression.

The research presented here asks us to undertake a variety of questions surrounding the role of social markers and categorization. The applied lessons of intersectionality help disentangle the enmeshed relationship these categories have to each other and any given subject. I don't intend to use intersectionality as a reification of categories themselves, but rather an analytical lens through which to examine the way society categorizes identity and systematically perpetuates discrimination in an intentional way. In this project intersectionality serves as an important underlying theory because it offers insight into how society structures identity to perpetuate discrimination. This process seems necessarily related to the way an individual might construct identity concepts in the mind. This relationship between intersectional identity concepts is part of a critical structure through which we perceive ourselves, each other, and interact with the world. Interlocking systems of oppressions shape our experience of identity across domains.

Ultimately, intersectionality is an essential starting point for forming new modes of analysis through this project, but it is not a destination. It is a tool harnessed in the development and analysis of this work but remains an evolving analytical strategy especially in this application. The design of the preliminary studies in this project (studies 1 and 2 specifically)

leave room for identity to be applied and considered not only as a singular category or trait, but as an open ended and overlapping form of classification. Study 1 specifically encourages the description of social prototypes through a multifaceted lens that study 2 employs in asking participants to recognize identity trait associations across 12 different yet clearly overlapping trait categorizations.

Intersectionality also occupies a critical bridge between the practical and theoretical, a space explored more in the final section of this theory review. Moving forward I hope to address the ways in which research methods and epistemological power are outlined by Feminist theory and applied across disciplines in the production of knowledge.

### 3. Epistemology and Feminist Science Studies

Ultimately, this project is a research endeavor in the field of Cognitive Science, however the purpose of this chapter is to ground the work in feminist theories of knowledge production. The pursuit of any form of research or any type of intellectual investigation forces us to critically interrogate what it means to produce knowledge. Fundamentally, we must ask ourselves, what does it mean to “know?” What structures of power restrict our understanding of the boundaries between fact and opinion, or experience and recollection? The production of knowledge is an essentially subjective undertaking, particularly when one is situated within the academic institution at a position of power and restricted access. One specific way processes of knowledge production are both regulated and reaffirmed is through the assumption of a dichotomy between qualitative and quantitative methods. This disjunction is justified by the assumed “objectivity” of quantitative analyses that are considered fundamentally oppositional to qualitative modes of knowing which provide more generous margins for the validation of subjective experience. By arguing for a reconciliation between qualitative and quantitative methods I hope to situate

research as a recursive process of knowing where quantitative and qualitative knowledges are enmeshed. Feminist theory also directly addresses unknowability and the multiplicity of situated knowledges. I want to conclude this chapter by offering a vision for this research project, and the larger project of interdisciplinary research, that embodies the recognition that while we can never truly 'know' anything from an objective standpoint, we can continue to pursue better empirical understandings of the world.

An epistemological analysis is critical to identify false assumptions of objective quantitative methods which must be subverted to reconsider knowledge production from an alternative perspective. As the theory of knowledge production, epistemology centers the investigation of what distinguishes "justified belief" from "opinion." This boundary has been tied directly to the presumed border that divides "objective" quantitative methods from qualitative ones across academic disciplines. Indeed, objectivity is specifically attributed to "to methods or procedures that are fair: statistical, or experimental, or repeated procedures are more objective because they maximize standardization, impersonality or some other quality assumed to contribute to fairness" (Harding 332). A feminist epistemological lens challenges this assumed 'fairness,' as feminist theorist and health behavior scientist Toby Epstein explicitly states that much of this "debate about qualitative and quantitative research has been sterile and based on a false polarization" (Jayaratne 45). Epstein's argument highlights how the both the rejection of qualitative method's validity, and the inability to reimagine quantitative methods being used expansively and ethically highlights an unnecessary polarization of tools that should be implemented in tandem (Jayaratne 47).

The roots of this interdisciplinary debate stem from a perception of subjectivity as the antithesis to knowledge production that must rest on "objective fact." This hierarchy

intentionally and restrictively demands “that neither the individual beliefs or desires of a scientist nor the social values of a scientific community are relevant to the production of knowledge” (Tuana 257). However, the idea of situated knowledge is central to feminist theory which recognizes the inherent positionality of any perspective. By recognizing that there is no such thing as universal objectivity, even within scientific methods, we begin to dismantle the assumed structure that dichotomizes qualitative and quantitative ways of knowing.

Not only is the position of objectivity an impossible one, it is also important to note the ways the pursuit of objectivity can actually act as a hinderance to knowledge production. Objectivity will always be representative of a collection of assumptions that are both constructed and contested. It is not the implementation of perceived ‘neutrality,’ but rather a nuanced stand in for dominant political perspectives and historical trends (Harding 334). Assertions of objectivism are not only dismissive of qualitative ways of knowing, but “objectivism defends and legitimates the institutions and practices through which the distortions and their often-exploitative consequences are generated. It certifies as value-neutral, normal, natural, and therefore not political at all the policies and practices through which powerful groups can gain the information and explanations that they need to advance their priorities” (Harding 337). Harding directly exposes how the assumption of objectivity perpetuates dominant narratives to systemically erase minoritarian experiences. In this sense the pursuit of objectivity is not just misguided, but intentionally harmful to those who do not conform to societal notions of normativity. Rejecting this idea of objectivism requires a re-examination of scientific methods and attributions of validity within both quantitative and qualitative methods for knowledge production.

The intersection of feminist theory and epistemology offers a new lens to consider the arrangement of ideas which is required for any process of knowing. By using feminist analytic

tools, all methods can be re-examined from their relationship to power and their position of privilege. Ultimately, alternative knowledge productions demonstrate precisely how we must seek to redefine knowing. Even down to the hierarchal structures of our citation practices, Black feminist geographer Katherine McKittrick reveals how in the process of research “arranging, rearranging, and collecting ideas outside ourselves are processes that make our ideas our own,” but advocates for resisting a sense of control by thinking “about how our ideas are bound up in stories, research, inquiries, that we do not (or should not claim we) own” (McKittrick 2021, 15). This lesson reaffirms the positionality of all knowledge as subjective, but also digs deeper into the philosophy of citation practices to further question the role of the individual in the imagined (and literal) practice of knowledge discovery. McKittrick clearly rejects the idea of knowledge as the linear pursuit of objective truth, but rather asserts a way of knowing that requires ‘rearrangement’ and acknowledges positionality. A further examination of alternative modes of knowing, specifically those offered through Feminist methodology, will add increased scrutiny to our imposed structures of knowledge production and stretch the perceived limits of our modes of knowing.

The shift of focus back to research methods demonstrates the way feminist science studies can propel a transition in the design of academic structures in the pursuit of knowledge across disciplines. As the relationship between qualitative and quantitative methods is reevaluated, theorists offer a new perspective on the sciences and a rejection of the “neutral” knower. Nancy Tuana advocates for a recognition of the engaged and dynamic position of knowing, a thorough recognition of affective processes, and an examination of embodiment in developing knowledge processes (Tuana 257). These goals are supported by examples of ‘knowing’ which stretch the presumed bounds. Specifically, Tuana offers a metaphor to reify the

notion of qualitative and quantitative embodied knowing. She paints the picture of a health care worker who diagnoses their patients by running diagnostic tests, and viewing their patients' behavior and symptoms, but also by listening empathetically, offering trust and validation, and engaging in dialogue that affirms the patients' needs and experiences. In this metaphor Tuana demonstrates the success of the engaged, 'embodied' scientist who offers care by embracing knowledge holistically rather than selectively or under the guise of "detachment" (Tuana 264).

This descriptive analogy is brought to life in the outline of Allison Jaggar's work melding the roles of qualitative and quantitative methods in research to redesign poverty metrics. In her work Jaggar calls attention the way these metrics can be overly quantitative and thus perpetuate the erasure of lived experience and personal circumstance. However, in this case the intention of quantification is to highlight regions of the world where poverty is a significant issue in a distilled numerical metric; something that cannot be efficiently represented through a purely qualitative expression (Jaggar 500). Jaggar intentionally sits with the tension inherent to designing a metric that cannot ever fully capture experience, and then works deliberately to become the 'embodied' scientist who considers the limitations and compromises implicit in any chosen methods. Ultimately, her redesigned poverty metric does not claim to capture the unknowable, but rather stages a conversation between qualitative information (interviews, discussions, cultural expressions, mapping) and quantitative demographics (age, ability, work, dimensions of hardship) (Jaggar 507). Jaggar's methods here show a contemporary example of the way a reimagined form of knowledge production can be applied and implemented to have a concrete impact. This "conversation" between both qualitative and quantitative methods is something employed directly in the designed study trajectory for this project. One of the intentions of breaking down this progression of studies into so many smaller stages each with

varied designs is to allow for this kind of ‘embodied’ analysis which deliberately considers the limitations of each set up while also intentionally engaging in analysis across these different metrics. The progression of 4 studies implemented in this project demonstrate an attempt to blend with qualitative and quantitative methods in a larger consideration of identity from multiple methodological angles.

Ultimately, the purpose of this interrogation is to reimagine knowledge production through alternative ways of knowing. The work of contemporary researchers across fields demonstrates precisely how powerful this theory becomes when applied in practice. This line of inquiry reveals the necessary dialectic tension between quantitative and qualitative methods while still rejecting the essentialist view that these forms of knowing are oppositional by nature. This vision of knowledge production rejects linearity and objectivity in acceptance of positionality and instability.

While this branch of feminist epistemology offers an opportunity to reconsider the value of quantitative methods even in abstract work, the issue of fundamental unknowability remains. Black feminists reckon with the conception and recapitulation of experiences that exist beyond the scope of knowing or quantification. The fundamental unknowability of Blackness is what McKittrick herself refers to as “the mathematics of unliving, the certification of unfreedom” (McKittrick 2014, 20). McKittrick’s focus on the experiences of enslaved Black people whose ‘unfreedom’ is certified by the quantitative documentation of the Middle Passage, not only exposes the violence of these methods, but also the way that these experiences are truly beyond the scope of documentation. While McKittrick’s work, alongside that of other Black Feminist scholars, tackle the literal nature of unknowability within an incomplete archive, much of

unknowability is also consumed by the more a more abstract notion of the limitations of human position and perspective.

In particular, Feminist scholars Donna Haraway and Sandra Harding expose the complex relationship between objectivity and unknowability, while still advocating for the pursuit of situated knowledge through Feminist science studies. Harding's work directly considers how "Feminist challenges to conventional bodies of knowledge have forced reexamination of empiricist assumptions about the organization of scientific communities, ideals of the knower, the known, and how knowledge should be produced...and scientific method in the sense of 'how to do good research'" (Harding 345). Her argument offers the theorization of "strong" objectivity in order to reconsider a notion of updated scientific methods which at least partially subscribe to standpoint theory. Although notably debated, Harding offers standpoint theory to advocate for a form of scientific method that must start from the perspective of marginal lives to get a better picture of human social order (Harding 341). This commentary on standpoint and perspective converges with the work of Donna Haraway who puts forth situated, embodied, partial knowledge as a more 'objective' way to make responsible claims about reality. In Haraway's work the issue of objectivity is clear: "Feminist objectivity means quite simply situated knowledges" (Haraway 581). She grapples with the misleading illusion of 'the god trick' wherein scientists claim to be able to objectively see and know everything from the disembodied and 'neutral' position of nowhere. In rejecting this form of disembodied knowing, Haraway sits with the reality of unknowability to make her position quite plain:

I am arguing for politics and epistemologies of location, positioning, and situating, where partiality and not universality is the condition of being heard to make rational knowledge claims. These are claims on people's lives. I am arguing for the view from a body, always a



complex, contradictory, structuring, and structured body, versus the view from above, from nowhere, from simplicity (Haraway 589).

The process of knowing, as Haraway sees it, is situated in the body and the structures which scaffold it. Research in cognitive science specifically must compensate for this natural realization of embodiment, as the only way to know someone's mind is inevitably filtered through their own experience of being. Embodiment thus poses a real challenge of any assumption of impartiality when in fact, living is an experience filtered by perspective. However, this does not succumb us to a sense of futility. Rather than flee in the face of unknowability, Feminist theories offer a new trajectory for embodied knowing and methods which embrace situation and perspective to offer knowledge that is not constrained by presumptions of neutrality.

#### 4. Conclusion

The greatest challenge of this project lies in the contradictions that exist at the intersection of Feminist theory and traditional modes of scientific inquiry. In the words of Haraway, "Feminism loves another science: the sciences and politics of interpretation, translation, stuttering, and the partly understood" (Haraway 589). However, throughout the design, implementation, and discussion of this research I hope to find and harness the synergy between feminism and science studies. The two fields inform each other and push the boundaries of knowledge production in unique and yet necessary directions. Feminist science studies reject objectivity, enmeshes quantitative and qualitative methods, and embraces unknowability to produce knowledge that is situated by the perspectives of marginalization. It is for these very reasons that I find this work to be so relevant and imperative. In a time of misinformation and the re-examination of scientific objectivity, how do we continue to pursue knowledge without perpetuating the erasure of those

who are continuously marginalized by research norms? In my mind, feminist science studies offer an alternative trajectory for research methods. The destination may still be unknown, but I attempt to navigate this pathway because I too believe in an alternative notion of knowledge production grounded in notions of embodiment and equality.

### III. Methods & Analyses

## 1. Representations of Identity

The series of studies implemented throughout this project are designed to investigate people's heuristic expectations of identity and what might happen when we are forced to consider identities beyond the scope of the default. We make judgments and decisions every day that require us to make guesses about unmarked details. For example, when considering what to wear every morning one might subconsciously create a hypothetical expectation of the potential weather. This heuristic could be useful in the making of certain decisions like buying clothing or making travel plans. With access to only limited information about what the weather may be like on a given day, we must rely on our cognitive ability to fill in unmarked details to create a detailed enough representation to be helpful when making these judgments. While this entire process may be subconscious, there are clearly certain patterns to the representations being consistently implemented as the default. It seems likely that one's heuristic of "normal" weather is more likely to include possible scenarios that contain a few clouds or falling leaves than lightning strikes or typhoons. Thus, the way the mind fills in default representations is a critical aspect of much of cognition, from small judgments to important decisions.

This same process of filling in the unmarked can be examined through the lens of identity traits. Just as the mind constructs default representations for the world, we also must make underinformed guesses about unmarked facts for other people. In the same way, we can expect that when we rely on default representations of people, we subconsciously consider certain sets of both internal and external traits over others. For example, in order to have a thought about whether a leader can perform a task, one might create a more detailed representation of a "leader" based on their own expectations of what is normal for someone assuming this role. The framework of default identity trait representations serves as a potentially powerful heuristic in

judgments of possibility. We might also consider what the consequences are when these heuristic expectations of identity are violated. How do we react when someone who differs from our assumption of a default “leader” assumes this role? This project empirically investigates what the heuristic expectations for identity are, which identity traits are consistently excluded from the default, and what are potential consequences for violating these “normal” representations?

Study 1a offers an open-ended qualitative examination of the default identity representations for specific social roles to highlight which traits are most referenced in the perception of identity concepts. Study 1b develops a clearer pattern of which trait categories are most significantly referenced in representations by having participants respond directly to the qualitative descriptions from Study 1a. Study 2 extends this consideration of what is recognized as the “default” by having participants evaluate different trait and social role combinations from the perspective of what first comes to mind, what is statistically probable, and what is socially ideal. Study 3 evaluates the same subset of role and trait combinations to scrutinize which traits participants might represent as inherent to filling a social role (principally) rather than merely a likely aspect of filling a social role (statistically). Finally, study 4 has participants judge impossible and possible identity traits in order to analyze relevant patterns of response time and accuracy which highlights one of the direct consequences of these differences in representation.

## 2. Study 1a: Qualitative Associations

The first study was designed to collect qualitative responses about the default identity representations for a set of social roles. While later studies aim to categorize traits more narrowly to elicit quantitative patterns to be analyzed, the aim of this initial study was to broadly and openly ask participants to reflect on their own assumptions of identity. The goal of this study was

to discover which dimensions of identity participants are relying on in a spontaneous generation task. Developmental studies on the cognitive representations of social categories demonstrate that even children of a young age conceptualize intersectional facets of identity (Rhodes et al, 2021). Building off this awareness of the overlapping properties of race and gender, this qualitative experiment seeks to isolate other potential identity traits (height, class, age etc.) that adults are relying on when making social identifications.

## 2.1 Methods

Participants in this study were presented with a series of 15 social roles through a Qualtrics survey and were asked to imagine the physical embodiment of each. The list of 15 roles were selected to cover a wide range of categorizations and encompass both more explicitly occupational roles (officer, teacher, boss, nurse) as well as more abstract social positions (leader, victim, rebel, peacekeeper). The survey asked participants to take a full 2 minutes to type their description as an open-ended text response and encouraged them to include as many details as possible about the physical description of what they imagined. Participants were allowed to proceed to the next prompt when their response met the required 200-character limit.

Data was collected from 50 participants (19 male and 31 female), recruited through Amazon Mechanical Turk and compensated for 20 minutes of their time. Responses from 4 participants were excluded because they were incoherent or clearly directly copied from sources online. The final results considered the responses of the remaining 46 participants.

## 2.2 Results

In order to visualize patterns of descriptions across text responses, word clouds were developed using the full text responses for each social role. All 15 word clouds from Study 1a can be found in the attached supplementary materials, but Figures 1 and 2 exhibit some of the



detects to associate the sampled input with a sentiment output. There is clearly potential for these sentiment outputs to be reflective of not just biased sentiment within participant responses, but also a reflection of bias within the feature detection algorithm itself.

Figure 3 shows the results of this text analysis and general sentiment analysis. Even from these preliminary responses a subtle pattern emerged which revealed a stronger negative sentiment of text analysis correlated directly with roles that were also more frequently associated with female identity traits. To continue the examination of the two traits portrayed in the word clouds above, responses for “leader” had the strongest positive sentiment while the sentiment analysis for “caregiver” skew negative.

social prototype	keywords	sentiment (percent confidence)	social prototype	keywords	sentiment (percent confidence)
parent	parent, child, good parents, children, kids, care, late twenties, role model, mom	Negative 94.9%	nurse	patients, comfortable shoes, nurse, stethoscope, long hours, care of others, blue scrub, middle aged women, white uniform	Negative 77.1%
boss	suit, boss, group of people, employees, white man, many people, in charge of others	Positive 57.9%	assistant	young girls, boss, assistant, google assistant, smart home device, late twenty, central control system, variety of jobs	Negative 98.2%
leader	better vision, leader, big pictures, many people, great leader, attention of everyone, nice dress shoes	Positive 99.4%	student	school, students, young person, educational institutions, desired field, white t-shirt, subject matter, young girls, new information	Positive 95.7%
protector	harm, protector, safety of others, museum guard, valuable artwork	Positive 61.2%	caregiver	family member, caregivers, medical field, lots of love, younger siblings, older women	Negative 54.2%
criminal	orange jumpsuit, jail cell, criminal, illegal things, hard time, physical harm, done wrong, white man	Negative 71.8%	role model	role model, nice car, brown hair, good example, younger people, good person	Positive 98.6%
rebel	young girls, young person, authority figure, rebel, norm of society, leather jacket, dark hair, status quo	Negative 83.6%	victim	victims, taken advantage, need of help, action of others, active social life, many different	Negative 98.7%
officer	blue uniform, officer, bad rap, middle aged man, short hair, sure people, facial hair	Negative 82.5%	peacekeeper	peacekeeper, others, army fatigue, older person, united nations,	Negative 89.5%
teacher	children, students, teachers, knowledge, others, new subject, lot of patience, lot of time, long dress	Negative 89.5%			

Figure 3: Text analysis of qualitative responses across 15 social prototypes from Study 1a



Other less pronounced, but still noticeable, patterns appeared in descriptions of young identities, disabled identities, and racial minorities being among those social prototypes and text descriptions most confidently identified as exhibiting a negative sentiment. While these preliminary results begin to shape expectations about prescriptive associations for both social roles and identity traits, the following study will build off participants' text responses directly to avoid incorporating further sources of external bias.

### 3. Study 1b: Quantifying Categories

Working directly from the original text responses from the 46 participants in Study 1a, Study 1b seeks to isolate data driven categories in adults' spontaneously generated responses to be employed as identity classification categories in Study 2. This approach was motivated by the reality that while many responses seemed to implicitly rely on identity descriptions, not all of these traits were labelled specifically in responses. The task was designed to highlight the implicit biases in descriptions that could not be explicitly withdrawn through a text analysis tool. By passing these responses to be reviewed another set of participants, the aim was to have humans identify implicit identity markers in descriptions without adding the confound of having a single program or researcher make these judgments and introduce their own inherent bias. For example, some participants used coded gendered language in their responses without explicitly describing a prototype as "male" or "female." Instead, they might include descriptions like "has a ponytail," "speaks in a high voice," or "wears high heels" which are physical traits that certainly have socialized connotations with those who identify as female but are of course not explicit markers of gender themselves. The aim of this study was to establish which identity categories are being most heavily relied on in default representations while also accounting for bias that was encoded implicitly in responses from 1a.

### 3.1 Methods

This survey asked participants to directly respond to text responses from Study 1a. Any responses from 1a which did not describe a single individual were excluded for the purpose of this study. The remaining responses were curated and embedded into the prompts of the survey for 1b. In this survey each participant was randomly presented with a response from 1a for each of the 15 roles used in the previous study. The participants were not told explicitly which role the description had been intended to describe, but rather were asked to simply read the description and assess the identity of the person described across 12 potential identity trait classifications.

The possible identity traits were generated directly from the group level analysis of the responses from 1a and included a spectrum of types, everything from race, gender and age to appearance or trustworthiness all of which were frequently referenced across responses. To reduce the task time, “redundant” traits were excluded. For example, for age and height participants only made assessments on whether the person described was “young” and “tall” and were not explicitly asked to consider whether they were “old” or “short.” For each trait participants were asked to use a slider to convey how strongly they agreed that the individual described fit that trait. For example, “How strongly do you agree that the person described is a woman?” Each slider scale started at 0 for “Strongly Disagree” and went up to 100 for “Strongly Agree.” The midpoint of the scale, 50, was labeled as “Neutral.” Each participant recorded their agreement with a slider for all 12 traits for each of the 15 role descriptions they were presented with.

Data was collected from 34 participants (11 female and 23 male) through Amazon Mechanical Turk and participants were compensated for 15 minutes of their time. Individual

responses where subjects did not click on or interact with sliders at all were excluded, but no full subject data was removed with this criterion.

### 3.1 Results

After exclusions were made, the remaining data was organized by both identity trait and social role. As described above, although all 34 responses to descriptions of a “parent” were categorized together into the “parent” condition, each participant was responding to a unique description that was randomly selected from the responses in 1a. After the conditions were divided, the mean, standard deviation, and standard error were calculated for the agreement judgements for each separate trait-role pair. As there were 12 traits and 15 roles this meant there were 180 final trait-role conditions taken into consideration.

Figure 4 depicts these means and standard error calculations. Some preliminary trends emerge from close observation of the plot as traits like gender, height, and trustworthiness are frequently represented with a higher level of confidence and race and sexuality are more consistently met with uncertainty. The initial patterns develop a subtle depiction of which aspects of an individual’s identity are most prominently referenced in default representations as well as which identity traits are consistently recognized or assigned with the most confidence. This task directly considers which types of traits are most strongly relied upon in heuristic expectations for certain roles. The sensitivity to traits like gender suggest that they at least play a prominent role in filling in the “unmarked” details of social roles, but further investigation will consider why prominently represented traits like gender are being included in default judgments.

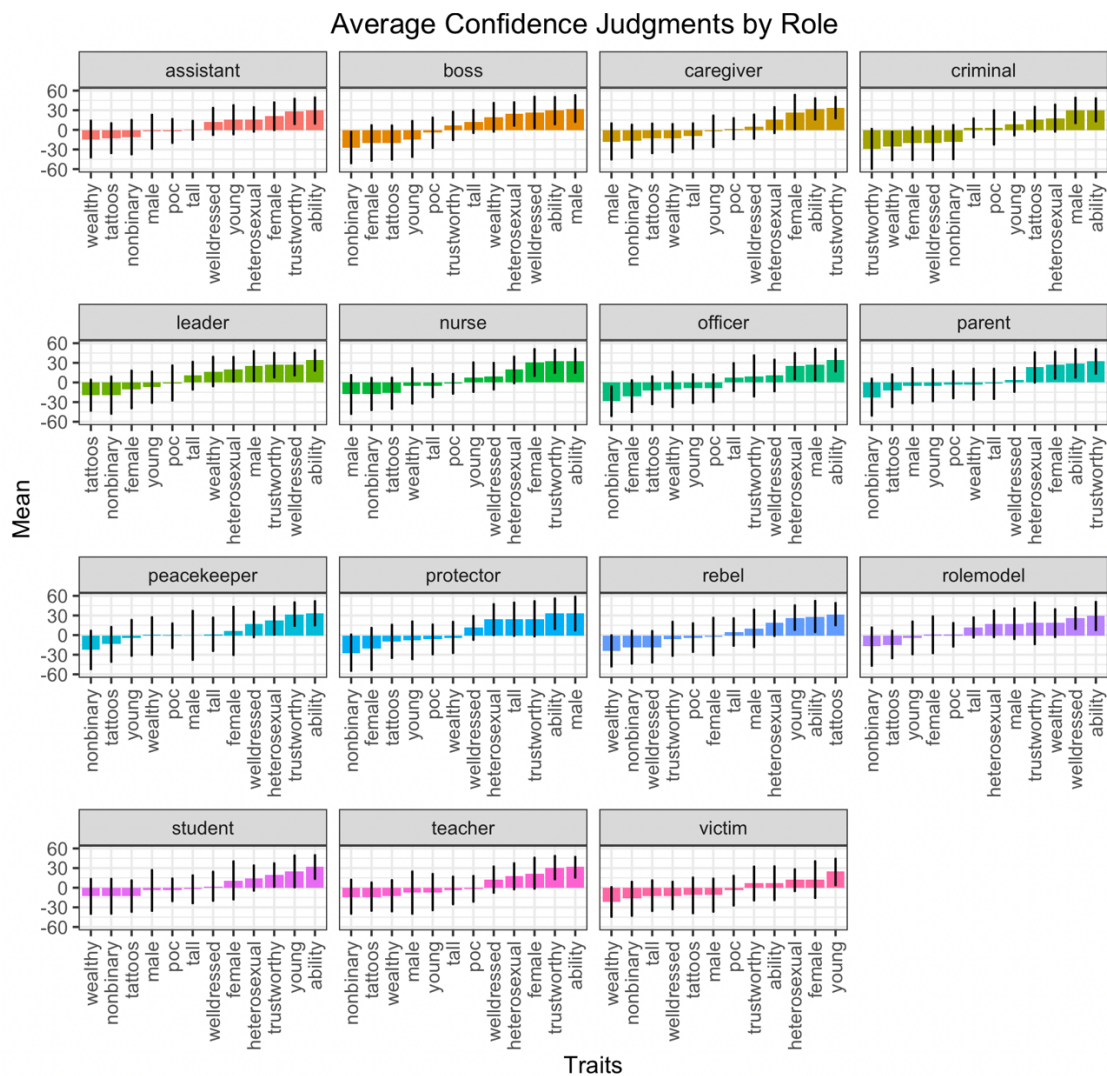


Figure 4: Average confidence agreement judgments for each trait across roles. Adjusted by -50 such that 0 represents “Neutral,” -50 represents “Strongly Disagree” and 50 represents “Strongly Agree”

Figure 5 portrays the same results as Figure 4, but instead groups confidence judgments specifically by trait. This visualization reveals not only the significant reliance on some traits in representations, but also the simultaneous near exclusion of traits like being non-binary or a person of color. Together these figures offer both a snapshot of the way each social role develops its own unique intersection of traits in the formation of a default representation, but also a larger sense of the ways these traits are relied upon across a broad range of social prototypes.

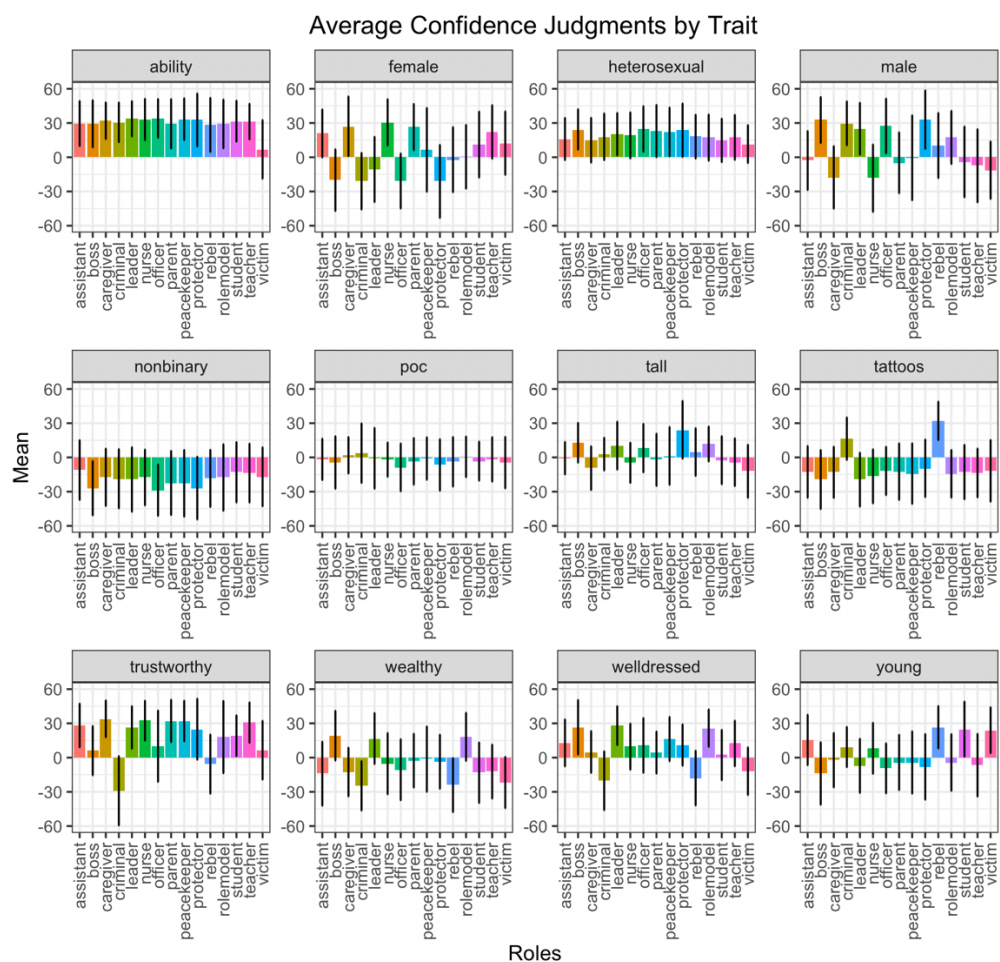


Figure 5: Average confidence agreement judgments grouped by trait. Adjusted by -50 such that 0 represents “Neutral,” -50 represents “Strongly Disagree” and 50 represents “Strongly Agree”

### 3.2 Discussion

To discuss these results from both parts of Study 1 it seems necessary to return to the theoretical context for these studies. In the initial discussion of intersectionality from chapter one, Nash was referenced as a reminder of the multidimensional complexity of Crenshaw’s original “intersection.” While derived from the intersection of race and gender, Nash calls critical attention to the reality that there are many more than just these two traits entering any potential identity intersection, and that these traits might not all enter with equally weighted importance.

I interpret the results from Study 1 as a preliminary qualitative proof of this concept. In this case each social role could be considered a hypothetical “intersection.” When a participant is asked to consider their default representation, the prototype is informed by a unique collision of identity traits and these studies begin to distill these patterns categorically. While results from Study 1a allowed for a broader view of identity traits, Study 1b began to refine these interactions on a narrower scale. For example, while gender, ability, and trustworthiness become quickly visible as traits recognized with high confidence levels (and thus perhaps weighted with significant importance), some trait concepts appear as much more salient in specific roles over others. Having tattoos seems highly associated with default representations of rebels and criminals, but in other contexts representations of this trait seem far less relied upon. These results also leave room for an examination of the interactions between traits, as in the way that protectors are the most strongly associated with being tall and are also most confidently categorized as male.

These studies also highlight certain traits which are consistently left out of the defaults. Being a person of color or identifying as non-binary are both traits that are rarely identified with a confidence above neutral, and in the case of non-binary, is consistently judged to be unlikely to be compatible with default descriptions. This lack of associations probe further consideration of what informs or shapes our capacity to consider certain identities within the default. Are these traits being perceived as unlikely, undesirable, or even less possible? How do these judgments inform our constructions of the default?

#### 4. Study 2a: What comes to mind? Ideal vs. Average

In their 2019 study, “What Comes to Mind?” Bear et al. suggest that default sampling concepts are a compromise of people’s judgments of value and probability (Bear et al. 2019).

Developing from the results from study 1b, study 2 is designed to directly probe whether the default trait and role representations that are first coming to participants' minds are also reflective of prescriptive or statistical perceptions of the norm. While 1b considers what default representations might be, this shift in inquiry reflects an important consideration of why these defaults are constructed in such a way. Are default representations consistently informed by other judgments such as value or probability?

Study 2 replicates the structure of Bear's first study in which participants were asked to respond with an average, ideal, or sampling judgement for 40 different real-world scenarios. For example, "hours of TV watched in a day" or "hours of exercise in a week." Bear predicted that the sampling judgments could be predicted by a multiplicative model of the prescriptive and statistical judgments. In this identity focused variation of Bear's study, instead of prompting participants to respond to everyday scenarios, we ask participants to consider the percentage of individuals of a certain role who identify as a certain trait. For example, "Percentage of leaders who are tall." Following the model of Bear's original study participants were asked to respond to trait-role combinations in a default sampling condition, an ideal condition, and an average condition. These results can then be used to assess the relationship between ideal or average representations of identity and the actual sampling default. For example, if average or ideal judgments strongly predict for the percentages that first "come to mind" in the sampling condition, we may have evidence that default representations are potentially informed by expectations of likelihood or perceptions of prescriptive value.

## 4.1 Methods

Trait-role combinations for study 2 were determined directly from the results of study 1b. We considered the traits that were most strongly associated with each role as well as which role was most strongly associated with each trait. After considering these we narrowed further to include a diverse selection of roles and traits. We wanted the final sample-set to include each role exactly twice and each trait to be represented at least once. Figure 6 depicts the final selection of 30 trait-role combinations which were used as question conditions for study 2.

	female	male	nonbinary	young	poc	wealthy	ability	heterosexual	tall	tattoos	well dressed	trustworthy
parent	■											■
boss								■				■
leader							■		■			
protector				■						■		
criminal		■				■						
rebel			■							■		
officer		■			■							
teacher	■										■	
nurse					■			■				
assistant												■
student				■				■				
caregiver		■					■					
role model									■			■
victim	■				■							
peacekeeper				■		■						

Figure 6: Matrix of trait and role combinations selected as prompts in study 2

All participants in study 2 were first asked to respond to all 30 prompts in a preliminary sampling condition. In the sampling condition participants were asked to merely respond to each phrase with “the first number that comes to mind.” They were explicitly instructed that there were no “correct” answers and were encouraged to respond in under 5 seconds, although responses after 5 seconds were still collected. Every question in Study 2 requested that participants submit their response as a percentage, or in other words a number between 1 and 100. For example, for the young-student trait-role pair participants were presented the prompt:



“Percentage of students who are young.” The decision to measure responses through percentage was made to consistently quantify judgments with a consistent numerical unit.

After submitting responses for all 30 prompts in the sampling condition participants proceeded to complete both the ideal and average conditions in random order. All participants completed all three conditions and began with the sampling condition to attempt to limit potential outside factors which could impact the anchoring of default judgements. In the ideal condition, participants were asked instead to provide judgments on what percentage of individuals of a certain role *should* have a certain trait. For example, “For students, what would be the ideal percentage of them that are young?” Similarly, in the average condition, participants were asked to respond with the percentage they felt was *closest to the actual statistical average*. For example, “On average, what percentage of students are young?”

In the first launch of study 2, study 2a, data was collected from 49 subjects (18 female and 31 male) through Amazon Mechanical Turk and subjects were compensated for 10 minutes of their time. No complete subject data was excluded from analyses.

## 4.2 Results

Responses were grouped by role and trait and mean responses were calculated for the average, sampling, and ideal conditions. A linear regression model revealed that across subject responses average judgements significantly predicted sample judgments with a correlation coefficient of 0.99 ( $p < 0.001$ ) and while ideal judgements were also predictive of sample judgments with a correlation of 0.79, this correlation was significantly mediated by the average

judgments ( $p = 0.963$ ). This suggests that while both ideal and average judgments predict for sampling, average remains a significantly stronger predictor.

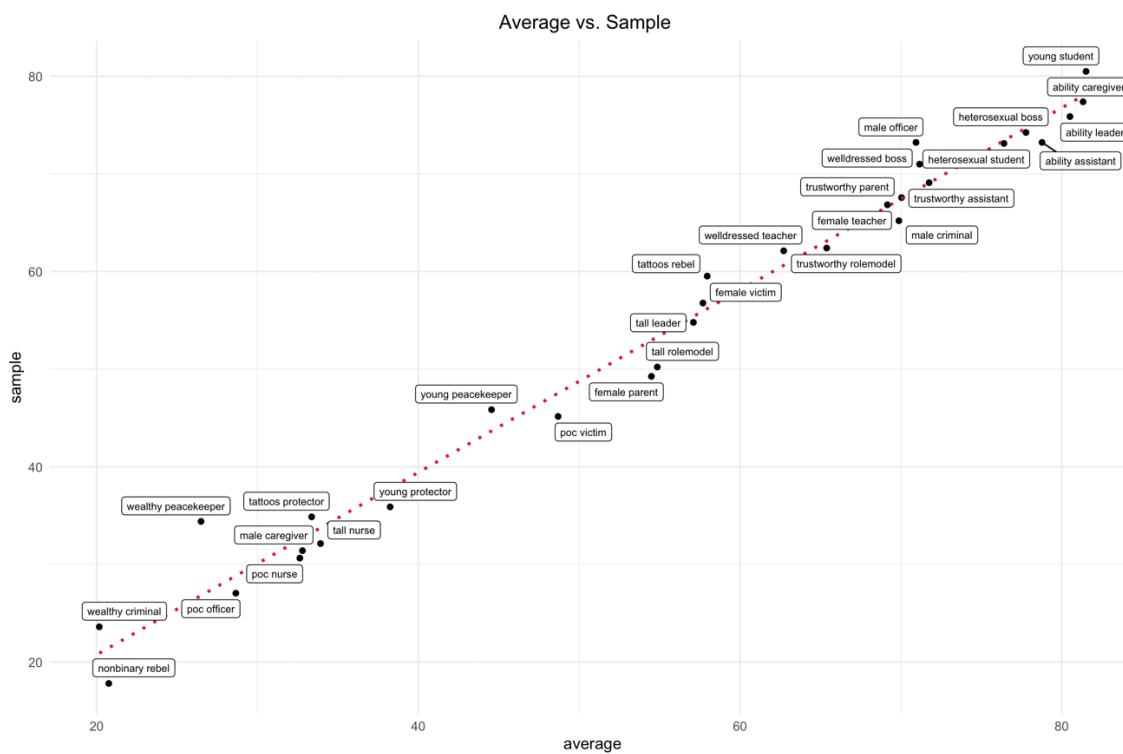


Figure 7: Average judgments predicting for sample judgments in Study 2a

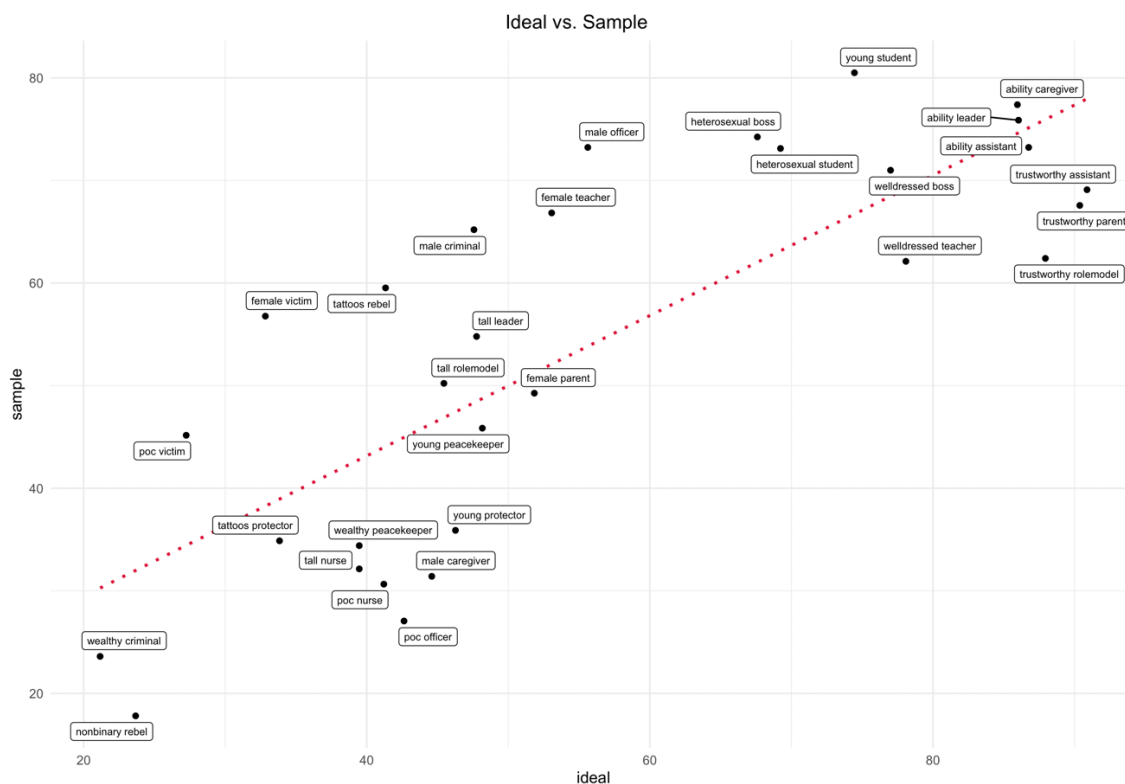


Figure 8: Ideal judgments predicting for sample judgments in Study 2a

Figures 7 and 8 show the distribution and regression lines for ideal vs. sample and average vs. sample. Because the correlation for average and sampling was so strong, we also considered an individual subject response analysis to consider the full distribution of responses. Figure 9 depicts the within subject individual responses of average predicting for sample. These responses had a wider spread as expected representing the variation in responses across subjects, but when considering the individual average and sample judgments for each participant, these still had a significant correlation coefficient of 0.82.

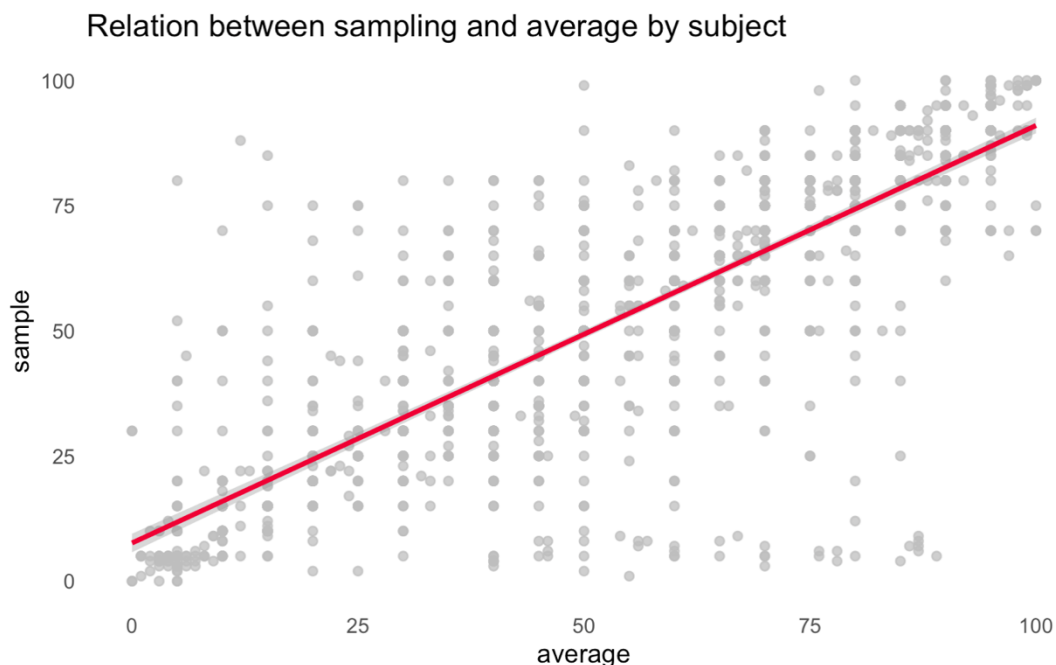


Figure 9: Within subject correlation for Average vs. Sample

Overall, the results from this study demonstrate that people's expectations about what is statistically likely seem to potentially inform their notions of what default identity concepts are. However, one concern about the strong correlation between average and sampling was that, because all participants responded to the sampling condition first, their judgments of average percentages were being strongly impacted by this initial default task. In other words, there was a concern that participants were not actually making judgments about their perception of what was truly most statistically probable, but rather subconsciously reverting to the percentages that first came to mind. Whereas the ideal condition provided participants with a very clearly distinct task (to consider what they believed *should* be the ideal percentage), there was reason for concern that the prompt for average judgments was not interpreted as distinctly different from the sampling task. A replication was run to specifically address this potential confound.

### 4.3 Study 2b: Replication without Sampling

Study 2b was an exact replication of study 2a excluding the sampling condition. 100 participants (42 female, 54 male and 2 nonbinary) were randomly sorted into either the ideal or average condition and completed just the single condition they were sorted into (50 participants for each condition). The purpose of this replication was to investigate whether the initial sampling condition judgements were anchoring the average judgments and causing the significant correlation.

### 4.4 Results

The replication revealed a similarly significant correlation between average responses from 2b and the original sample responses from 2a with a correlation coefficient of 0.988. The replicated ideal responses also similarly predicted for sampling, but with a once again lower coefficient of 0.767. Figures 10 and 11 plot the replications of both the ideal and average condition against the original sampling judgments. The consistency of this replication offers further evidence that the default representation of identity is being informed by what participants judge to be the most statistically probable.

However, it is still important to consider how participants' judgments of what is average is likely far from the real-world statistical average. While many of these "real world" percentages are not logistically measurable, it's worth noting the bias within a perceived average versus true actuality. It's also worth considering the difference in these results from the original Bear et al results which suggested that what comes to mind is informed by both ideal and average judgments. One potential explanation for the extra power of average as a predictor in this study might be the use of percentages as a standard unit for responses across tasks. Using this percentage-based question structure might have significantly altered the task

such that ideal judgments play less of a role and instead average judgments emerge as the more powerful predictor.

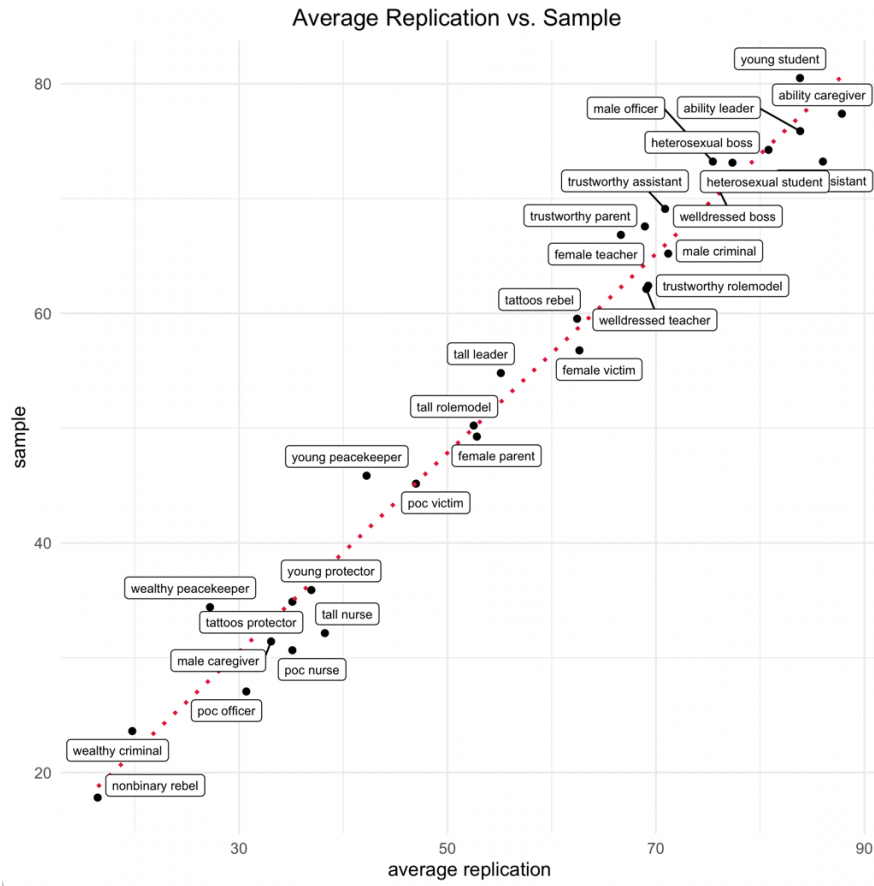


Figure 10: Average judgments from replication predicting for original sample judgments from Study 2a

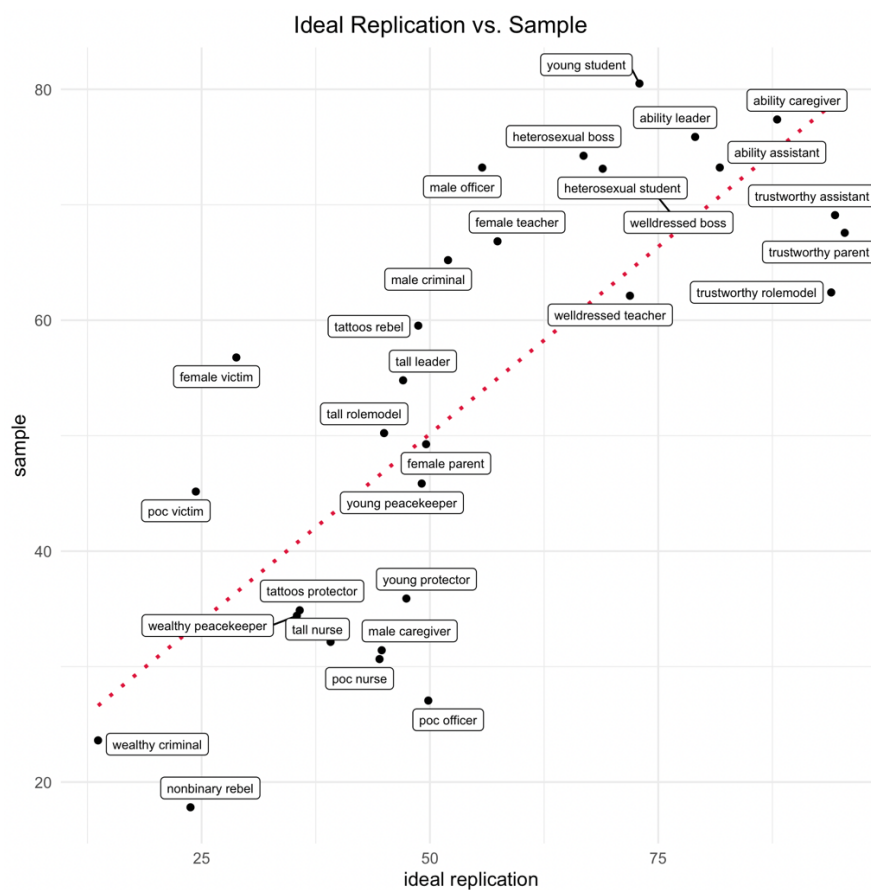


Figure 11: Ideal judgments from replication predicting for original sample judgments from Study 2a

#### 4.5 Discussion

The results from Study 2 show that judgments of what is most “average” strongly predict for sampling judgments across both the different types of identity traits as well as the social roles themselves. This study directly asked whether judgments of what is average or ideal could predict what comes to mind. Ultimately, it seems plausible to include that our ideas of what the default is are at least somewhat informed by both but might be most significantly predicted by our perception of what is statistically probable. These results offer insight to the difference in trait representations across Study 1 by explaining why some traits might be offered a “greater” importance in certain contexts. These most salient default traits may be reflective of expectations

of what is normative, even if these perceptions of normativity do not accurately reflect the state of the world.

## 5. Study 3: Statistical vs. Principled Connections

The goal of study 3 is to investigate which of the traits explored in studies 1 and 2 may exhibit the expected patterns association consistent with those of principled traits and assess the strength of this form of association. While the results from study 2 suggest judgments of what is average might predict default sampling judgments, sampling judgments still correlated significantly with ideal judgements suggesting that perceptions of statistical likelihood are certainly not the only factors informing defaults. Study 3 examines the type of association made between traits and roles to consider whether this form of connection might also be informative in understanding the default.

Starting with the initial category analysis from study 2, study 3 follows the trajectory of the trials from Sandeep Prasada's work on principled and statistical connections in common sense conception (Prasada & Dillingham). Prasada theorizes that the relationships between concepts and their "types" follow two patterns: statistical connections consider the commonality of an associated trait while principled connections are considered a fundamental mode of understanding the 'type' itself.

Specifically, the structure of Study 3 replicates the set-up of Prasada's work in experiment 3a in which participants were presented normative linguistic statements about a concept and its properties and were asked to respond to whether those statements were true or not. For example: "Clocks, by virtue of being clocks, should have alarms." Participants were shown to demonstrate normative expectations about principled concepts (k-properties), but not statistical concepts (t-properties) (Prasada & Dillingham 93). This study sought to



examine whether certain identity traits were more likely to be associated with social roles in a principled rather than statistical way.

## 5.1 Methods

The survey was designed after the structure of Prasada's study using the subset of 30 trait-role combinations selected for use in study 2. Participants responded to all 30 prompts in a random order. Prompts asked whether each trait was an inherent aspect of a person assuming that role. For example, "Leaders, by virtue of being leaders, should be tall." Participants used a slider to indicate how true they believed each sentence to be from a scale of 1 ("definitely not true") to 7 ("definitely true").

Data was collected from 50 subjects (21 females, 28 males, and 1 non-binary person) through Amazon Mechanical Turk and subjects were compensated for 5 minutes of their time. No responses were excluded.

## 5.2 Results

Subject responses revealed a clear pattern of traits that were represented in a way which more closely resembled a principled connection. Trustworthiness, able-bodiedness and being well dressed had consistently high responses with means all  $> 4$ . The table in Figure 12 presents all the mean responses across the 30 trait-role combinations and Figure 13 displays a histogram of the response means.

role	trait	mean	role	trait	mean	role	trait	mean
protector	tattoos	1.42	officer	poc	1.92	officer	male	2.74
nurse	tall	1.46	role model	tall	2.02	student	young	3.38
criminal	wealthy	1.54	peacekeeper	young	2.04	teacher	well dressed	4.08
rebel	nonbinary	1.58	victim	female	2.08	leader	ability	4.14
peacekeeper	wealthy	1.62	protector	young	2.12	assistant	ability	4.24
student	heterosexual	1.64	leader	tall	2.16	boss	well dressed	4.40
victim	poc	1.76	criminal	male	2.26	caregiver	ability	5.42
boss	heterosexual	1.80	parent	female	2.26	parent	trustworthy	5.58
caregiver	male	1.84	teacher	female	2.54	assistant	trustworthy	5.66
nurse	poc	1.86	rebel	tattoos	2.58	role model	trustworthy	6.42

Figure 12: Table of mean judgments of principled connections from Study 3

Overall, most traits were still being represented well below the expected association of principled traits, however there are still certain trends within the distribution of means worth noting. For example, appearance-based traits like being well dressed, tall, having tattoos or even traits like gender or ability which could be visually recognized seem to have some of the strongest associations. There is also a significant cluster of gender trait-based means right above a mean judgment of 2. While this is still significantly below mean judgments for the highest (most principled) cluster, it seems worth considering why gender is still being represented as potentially more strongly associated than other traits like race or sexuality.

Although default heuristics might be formed with both considerations of what is ideal and what is likely, some types of traits seem to be more clearly represented as implicit. Regardless of why precisely traits like being well dressed or able bodied are being employed in some default considerations, their power of being more principally associated with specific roles suggest these traits in particular might guide expectations of what is possible. In order to consider limitations of default representations and the consequences of violating normative expectations, we

specifically want to manipulate the trait being represented most principally and compare this response to the traits that exhibit more statistical patterns of association.

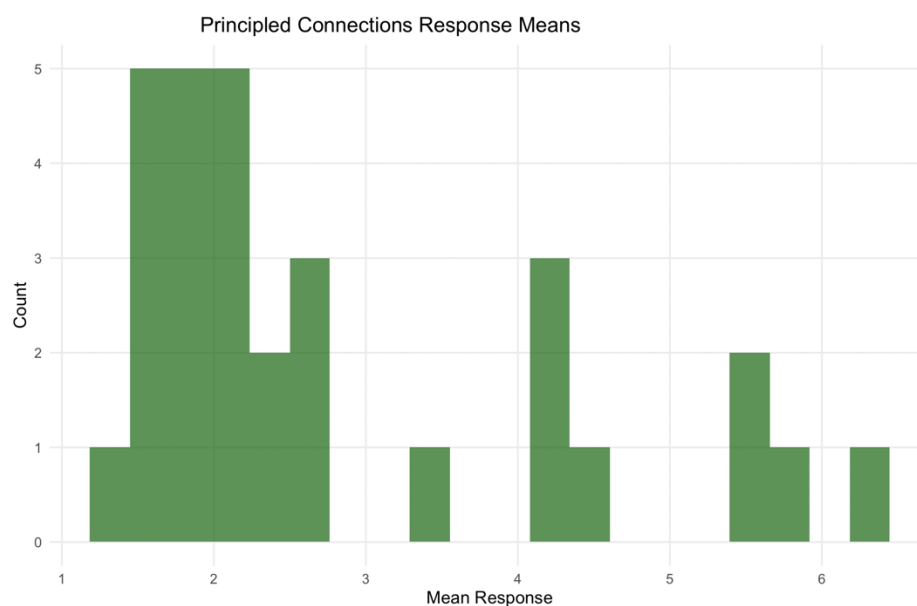


Figure 13: Histogram of mean responses from Study 3

### 5.3 Discussion

While only a few traits tested in Study 3 truly resemble fully principled connections (being trustworthy, able-bodied, or well-dressed), the results still reflect that there may be meaningful differences in the level of principled association between certain trait role pairs. In parallel to the results from Study 1, traits like ability, age, and gender which seemed to emerge with high confidence of association across many different social role contexts, also seem to be closer to resembling principled connections. Once again, traits that were consistently excluded from representations of the default, such as having tattoos or being non-binary, were some of those with the lowest principled connections. These results reflect the conclusion that default representations not only impact on our ability to consider how well people fit into our perceived normal expectations, but also who is unlikely to be represented in our initial judgment of what is possible. The results from Study 3 offer another explanation for how our defaults are shaped, not

by value or probability specifically, but by a perception of connection defined by a sense of innateness.

## 6. Study 4: Constraints of Possibility

The final study implemented builds off the work of all the previous studies to directly address the central question regarding the relationship between identity trait representations and notions of possibility. Using the results from study 3 as a preliminary examination of the potential differences in representation and association for identity traits, this study considers the related consideration of possibility. In essence, rather than considering just implicit association, study 4 inquires how representations of identity may affect the notion of how possible it is for any given person to assume a social role.

### 6.1 Methods

Study 4 was designed and implemented through Testable to allow participants to submit quick keystroke responses to the numerous stimuli and have their response time and accuracy recorded. Participants started by completing a practice condition to train for the task. 20 prompts of either “Yes” or “No” were presented and participants were asked to type “f” for “Yes” and “j” for “No.” If any incorrect responses were submitted in the training condition participants were reminded of the correct key associations and prompted to slow down and be sure to answer correctly.

After the training condition was completed, all participants responded to 75 prompts presented in random order. Each prompt asked participants whether it was possible for a person of a certain role to be a certain trait. For example, “It is possible for a teacher to be transparent.” As in the training task participants were told to type “f” if they believed the

answer to be “Yes” or “j” if they believed the answer to be “No.” In the real testing condition participants were not alerted if they made any mistakes.

All 75 prompts utilized the same 15 social roles that have been tested since study 1, but in different conditions each role was paired with different potential traits to manipulate whether the existence of such a person was possible or not. There were 30 prompts designed for the impossible condition in which the same 15 social roles were each used twice but paired randomly with traits that were realistically impossible for humans to have such as being “invisible,” “inanimate,” or “hollow.” The remaining 45 prompts all described possible traits and included the original 30 trait-role combinations tested in studies 2 and 3 as well as 15 other pairings that included the inverse traits for some descriptions such as “old,” “disabled,” or “short” as the inverses of “young,” “able-bodied,” and “tall.” These inverses were added to the possibility set to expand the set and realistically consider a broader set of traits which might be represented differently than their inverse despite being of the same type.

200 participants (71 females and 129 males) were recruited to complete the full task through Amazon Mechanical Turk, and each was compensated for 5 minutes of their time.

## 6.2 Results

The first variable measured across responses was individuals’ response times. Because response time measurements contain significant noise, response time outliers were removed from the data set prior to the full analyses. Figure 14 depicts the mean response time by traits, and while the standard error is still significant, the trend does seem to reflect the expectation that response times were generally longer for impossible traits than possible ones. A few noteworthy exceptions include untrustworthiness, disability, and being gender nonbinary. It seems worth considering further the way that these traits specifically might be

categorized as less probable or even less ideal such that they constrain greater representations of possibility. On the reverse end, the three pairs with the overall fastest responses times were “white officer,” “male criminal,” and “female parent.”

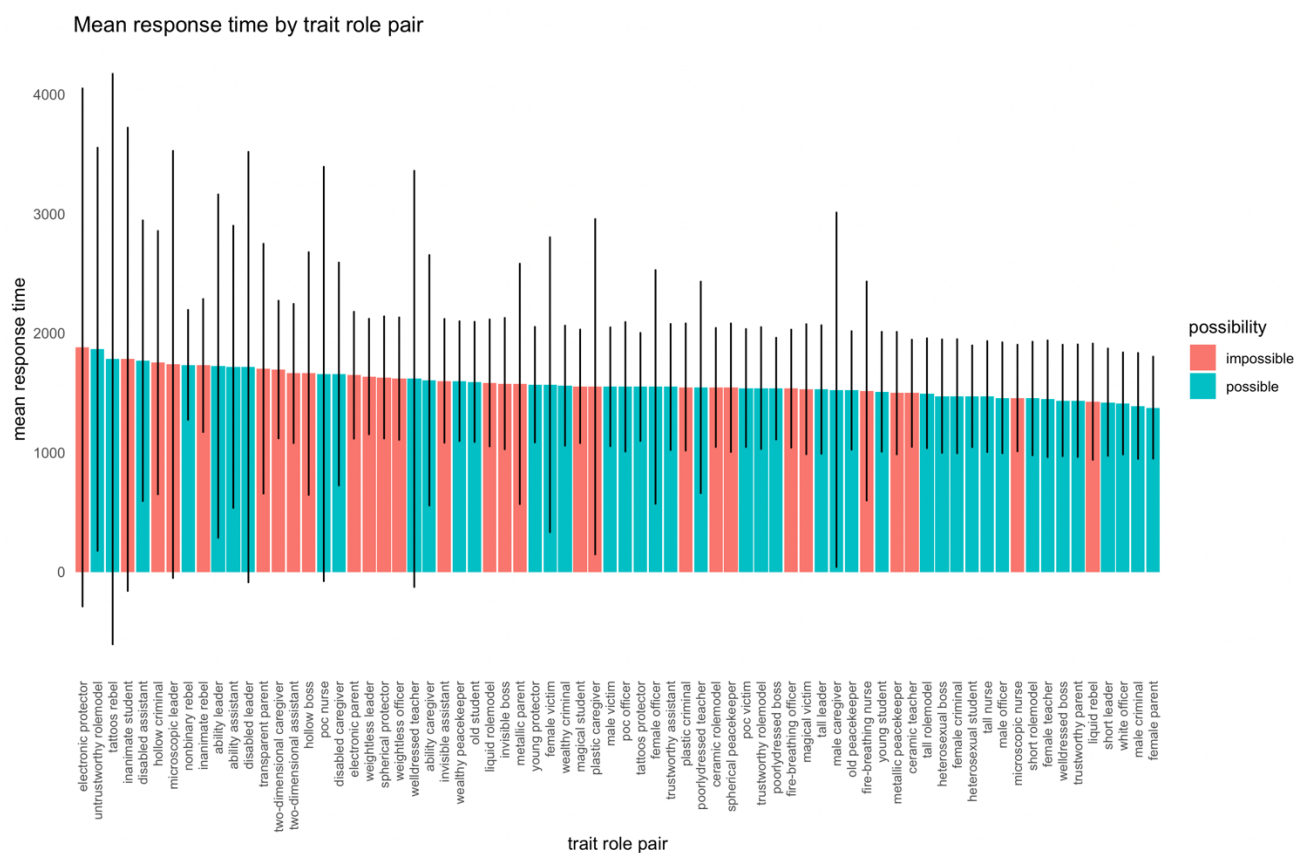


Figure 14: Mean response times by trait role pair for Study 4

In addition, analyses of the accuracy rate portray a noticeable inverse correlation with mean response time where traits that are “most impossible” and have the longest mean response times also frequently have the lowest accuracy rate. Figure 15 portrays this trend, and again the possible traits with the lowest accuracy rates include untrustworthiness, disability, and being gender nonbinary, as well as being poorly dressed or having tattoos. The three least accurately identified possible pairs were “disabled caregiver,” “nonbinary rebel,” and “untrustworthy role model” and were all responded to with an accuracy rate below the average accuracy for

impossible traits (0.901). In contrast the traits with the highest accuracy rate reflect some of the inverse traits including being trustworthy and well dressed as well as being wealthy or heterosexual.

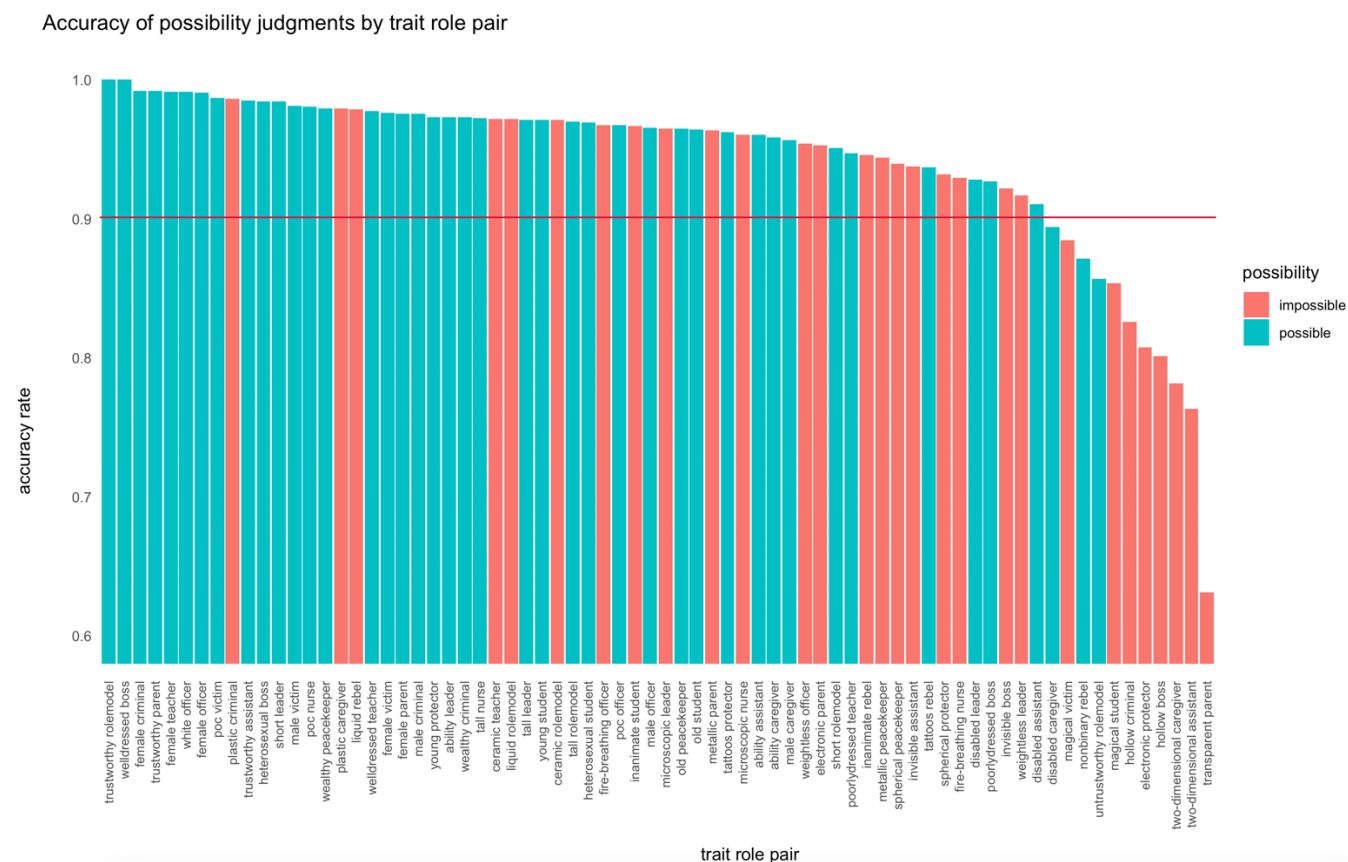


Figure 15: Accuracy rate of responses by trait role pairs for Study 4. Red line marks the average accuracy rate for impossible traits (0.901)

Employing the framework of default representation to examine these results, we can extrapolate that the possible traits which are still leading to the highest response times and the lowest accuracy rate are asking participants to consider an individual who exists beyond the expectations of their own default representation. Whether because they are representing a trait that is conflicting with these or are simply not filling in some of these trait classifications at all, these traits seem to represent the reality that not all identities are represented as equally possible. These results, in conjunction with the findings from the previous studies, begin to

illuminate a pattern of default representations which rely on implicit constructions of normality and thus restrict the expectations of “who” is possible.

Further analyses examined the relationship between these measures of possibility judgments to other related variables from previous studies. For example, if perceptions of what is most average inform our default representations and directly affect the response time and accuracy of possibility judgments, will mean judgments of average from Study 2 correlate with the results from Study 4? Similarly, could trends in possibility judgments be predicted by the general confidence of association (Study 1b), sample judgments of what comes to mind or expectations of what is most ideal (Study 2), or the perceived level or principled association (Study 3). When correlations were run for each of these previous studies in comparison to the possibility results from Study 4 it was found that both judgments of what is most ideal and perceptions of what is most statistically likely strongly predicted for accuracy of possibility judgments. Ideal judgments had a correlation with accuracy of possibility judgments of 0.365 ( $p = 0.051$ ) and average judgments similarly had a correlation coefficient of 0.349 ( $p=0.064$ ). As might be expected, sample judgments of what first comes to mind had the strongest correlation with accuracy of possibility responses with a coefficient of 0.370 ( $p = 0.048$ ). This promisingly suggests an important relationship between our ability to conceive of an individual as possible and our preconceived consideration of what seems both most likely and most valuable to us.

These final analyses offer insight into the complex nature of possibility judgments as compounding observations that reflect not only subconscious biases, but also are predicted by other explicit judgment tasks. The implications of this especially for the case of identity are critical as they suggest that there are other significant factors at play that could be



influencing which people's identities are represented in the default and which are most likely to be restricted.

### 6.3 Discussion

These final results pull together all the findings from the prior studies to assemble a broader examination of how default identity representations directly impact the ability to imagine certain identities as possible within a given role. The patterns of response time and accuracy rate make it clear that certain trait role pairs are clearly represented differently such that our ability to consistently conceive of them as possible is affected. Yet again, traits of disability and being nonbinary are some of the least accurately identified possible traits, so much so that responses to “nonbinary victims” and “untrustworthy rebels” are actually less accurate than the average accuracy for all impossible traits.

Paired with the concept that ideal and average judgments strongly correlate with this pattern of accuracy rate, it seems possible to argue not only that some traits are consistently less marked as possible, but also that these judgments might be directly informed by more inherent perceptions of what is likely or valuable. Ultimately, these results begin to offer a more complete understanding of not only which identity traits are relied upon in default representations, but also a preliminary examination of why.

#### IV. General Discussion

The results from these 4 studies are complementary and interrelated. While each study specifically tackles a smaller sub-question of the larger question regarding how default representations of identity could impact the ability to consider possibilities for people, the results of each study are best discussed in conjunction with one another. This general discussion will attempt to unfold each study in more depth, as well as in relation to the other studies. As these results are considered together, they reveal not only the notable differences in identity heuristics, but also begin to suggest conclusions as to how these differences develop and why they might be so impactful.

The preliminary studies consider a very open-ended and qualitative approach to the consideration of identity heuristics. Studies 1a and 1b jointly ask the most essential question in this project: is there a default identity heuristic for certain social roles? And if patterns can be isolated even from the open-ended descriptors that come to mind, do these patterns reveal meaningful characteristics about what identity traits are most frequently or prominently represented? These studies and their subsequent results serve as the foundation for this project, not only because they ask such a central question, but also because the intentionally qualitative nature of the study design of Study 1a allows for the most organic and intersectional responses.

The default heuristic for any social role is clearly multifaceted and might be structurally different for any given participant, and the unique responses from study 1a demonstrate this explicitly. Rather than diving right into a category driven quantitative model for structured identity types, much of the power of the results from 1a is fueled by its unlimitedness. This study serves as a qualitative proof of concept demonstrating that there clearly are distinct differences in the default identity representations for specific social roles. Descriptions heavily relied on include gendered norms, physical appearance-based traits, and other visual indicators of race or

even class status. As expected, these traits emerged as not only potential indicators of what is expected of certain roles, but also as potential limitations to the conception of possibility.

While still considering the same central questions, the intention of study 1b was to begin to quantitatively analyze the patterns emerging while still intentionally leaving room for the consideration of overlapping traits. While this study, and the ones which follow, explicitly harness the power of categorical labels to classify traits, in study 1b these categories are still represented as an overlapping composite of differences (such as race, class, ability, and gender) allowing for participants to consider not just the relevance or power of a specific trait, but rather the cumulative influence of a multiplicity of categories. Again, the employment of categories is not an attempt to reify the identity category itself (a concept which is explicitly refuted in the theory chapter), but rather offers an acknowledgement of the fact that our thoughts might still utilize a category driven lens when identifying others. In fact, it is precisely this kind of limitation of the category that these studies begin to highlight. Clearly when compared with the open-endedness of 1a, 1b and the following studies quickly begin to foreclose on explicit intersectional representations. While potential changes to study design might have broadened the scope of possible identities considered, and will be discussed further later on, this acknowledgment that at times some categories (such as gender) may be “simpler” or maybe just “more convenient” to track and represent highlights exactly why our default representations have the potential to be both so narrowly dominant and so flawed.

Ultimately, the results from study 1b demonstrate the strength of association between the representation of each trait for each role. By quantifying these associations, we can begin to consider the driving traits more directly for each specific category. For example, the confidence of “female” was among the highest across social roles suggesting that representations of female

were not only frequently included in the default, but also that the depiction of whatever it means to be “female” was confidently recognized and categorized perhaps because it was a trait that was explicitly marked throughout descriptions. On a similar thread, this study also begins to highlight not only the differences in traits that are confidently recognized for specific roles rather than others, but also those identity traits which are rarely associated with any defaults. It is one thing to be cast as within the default for only a specific type of social role, but another situation entirely to be cast out of any form of default representations all together. A variety of traits exhibited patterns of representation most similar to this latter description including being a person of color, being nonbinary, or even having tattoos. This explicit lack of association with any roles serves as a precursor for further investigation of what might be related to our capacity to imagine certain traits as possible rather than others.

The combined results from these initial 2 studies demonstrate not only a qualitative proof of concept that there are differences between the default identity representations for specific social roles, but also begin to quantify these differences in a way that will ultimately come full circle with the final study’s examination of possibility. Studies 2 and 3, however, approach the larger question of identity heuristics from a more predictive and explanatory angle. Rather than asking what exactly these differences in heuristics are, these next studies target questions about what else might predict these effects and what other meaningful relationships between defaults might be isolated to explain why certain trait representations might hold more power than others.

Study 2 directly asks: what predicts the default identity traits that are being represented? Working directly off the traits that seem to be most associated from the results of study 1b, this approach considers whether our default is based on our assumptions of what is most prescriptively valuable or statistically probable or some combination of both. This study takes its

design directly from previous work considering default representations of what comes to mind but narrows the scope specifically to consider the default representations of identity traits. The results show that judgments of what is most “average” strongly predict for sampling judgments across both the different types of identity traits as well as the social roles themselves. While it remains important to consider that sampling judgments may still have been skewed by the study design which had participants submit all their judgments in percentage form and which might not have distinguished average from sampling clearly enough, it still seems plausible to conclude that our ideas of what is “default” is strongly informed by our expectations of what is “most likely.”

However, it is important to remember here that our own representations and expectations of “average” are not necessarily at all reflective of any form of true statistical population average. Rather, we interpret these judgments as a reflection of normative expectations which clearly inform our own associations of what comes to mind. When considering identity specifically it seems fair to conclude that notions of what is assumed to be “statistically typical” inform and structure our heuristic default and thus might even directly impact which identity traits will first come to mind. With these results examined in context with the results from study 1, the differences in trait representations is not only recognized, but also explained on a preliminary level by the notion of expectations of average informing notions of the default.

While average more strongly predicted for sampling judgements in study 2, notions of what would be most “ideal” also correlated with sampling judgements. This relationship was slightly more nuanced, but there was still a high correlation between ideal and sampling judgments and thus there is still reason to believe that expectations of value are in some way integrated into the identity sampling heuristic for social roles. This notion of prescriptive

influences on what comes to mind is more closely examined through study 3. Much like study 2, the aim of this study was not to directly isolate the differences in identity representations across different social categories, but rather to ask what kind of relationship these identity traits might have with the roles they are associated with. While the results of study 2 suggest the power of statistical perception in influencing the default, study 3 acknowledges that these representations are still unlikely to be driven completely by statistical expectations alone. Study 3 demonstrates a spectrum of potential associations spanning from most implicitly associated as a principled aspect of this role, to least associated and merely a statistically represented possibility. Along this spectrum results from study 3 demonstrate not only the different types of associations being made as either more principled or more statistical, but also the notion that these representations should not be distilled to merely a projection of average expectations onto sampling judgments.

While the results from study 3 reveal that only a few traits (being trustworthy, able-bodied, or well-dressed) seem to consistently reflect the association level of principled judgments, there are still some trait and role combinations that seem to have some higher-level principled connections than others. Alongside the results from study 1b, a similar pattern emerges where age and gender seem to be more principally represented than traits like being a person of color or having tattoos. These same traits were recognized with some of the lowest confidence levels across all roles in study 1b and again point to further suggestions that not only are these traits potentially considered less inherent to or less common of certain roles, but also that there may be certain identity categories that are routinely and systematically excluded from nearly every default representation. While this set of studies ultimately was limited in its scope, the representations of these specific traits at least reflect the idea that the reality of default representations is not just their impact on our ability to consider how well people fit into our

perceived normal expectations, but also who will never be represented or considered in our initial judgment of what is possible.

Finally, study 4 builds off the work of each previous study to consider the broadest scope of how default identity perceptions might restrict our considerations of what is possible. Study 1 offers qualitative insight into the default representations of identity as well as draws initial quantitative analysis of the power and erasure of specific identity categories. Studies 2 and 3 consider the way that our expectations of what is statistically average as well as our notions of principled associations could inform these default representations. Study 4 seeks to consolidate these investigations of what default identity representations are and why they follow these patterns to in turn ask how this affects the perception of possibility.

Just as the purpose of study 4 is contextualized by each study the precedes it in this project, the results measured here are also best understood through a broad comparison to the results from each study before. The final analyses run directly compare judgments of ideal and average to the accuracy rate of responses from Study 4. These cross-study correlations further cement that not only are these judgments informing our default representations, but also that these judgments directly impact the perception of possibility itself.

Ultimately these trends suggest some significant and broad conclusions about the overarching inquiry of how default identity representations can actually constrain our ability to imagine certain identities of people as possible. Without overstepping the concrete conclusions, the patterns demonstrated across response time and accuracy rate, as well as the results from the previous studies make it clear that not all identity traits are associated with the same types of roles, represented with the same level of frequency, or judged to have the same implicit value. It seems viable to argue that this project demonstrates that some traits more than others are



consistently under-relied upon in default representations, and this results in a perception of these people's existence as being "less possible." While this conclusion may seem drastic on the surface, studies 2 and 3 offer some preliminary explanations as to why some traits (such as being nonbinary, disabled, or a person of color) may be excluded from the default more than others because of assumptions of normality and the construction of principled associations.

The results from study 4 directly reflect that traits like being non-binary, disabled or having tattoos are consistently among the characteristics that participants need the longest to conceptualize as possible or are most likely to incorrectly classify as impossible. In the opposite vein, the default heuristics of identity traits seem to allow for fast and accurate representations of being white, heterosexual, or trustworthy. While of course this project only had the opportunity to examine a small scope of the infinite combinations of potential identity characteristics, the results from study 4 offer preliminary evidence for a larger take away: we are not all represented by the notion of default and there may be concrete consequences for those forced to persistently exist beyond the scope of possibility. This also begs a return to the idea of intersectionality as a "location that resists" telling. These results push forward the notion that identity traits, including race, gender, class, disability, sexuality and more, do not have equal weights in the conception of identity, and thus cannot be stably isolated. The complexity of these multi-dimensional intersections makes it impossible to locate a single trait (even those like gender which seem to have high primacy over others) as the root of a person's identity or experience of discrimination, and instead force us to consider identity from a broader scope of overlapping interconnectedness.

The general discussion of these results also requires a reflection upon the methods implemented across studies as well as how they might be improved or extended for future work. As mentioned previously in the discussion of 1b, the implementation of categorical identity

frameworks across these studies warrants special consideration. The problematic assumption of any form of stable categorical label for identity is troubled thoroughly in the introductory theory chapter of this project. Nonetheless, the methods of many of these studies implement a category driven form of analysis to isolate and quantify patterned judgements of identity traits. However, this process is not intended to reify any form of sanctity the category might have in conceptualizing identity as a concept itself. Rather the intention of this project is to deconstruct the category both theoretically and empirically to ultimately reconstruct a notion of identity which exposes not only its inability to be captured by a singular statically defined label, but also the ways in which our own internal reliance on categories continues to reveal the inherent limitations of structuring identity in such a way. A purely categorical assumption of identity is fundamentally flawed and limited, however our judgments of others are also frequently flawed and limited. This work suggests that the structure of identity is not merely additive, but rather situational and multiplicative. Perhaps this pattern truly suggests that default representations rely too heavily on categorical assumptions, and we should consider how to advocate for a broader notion of how to acknowledge differences across people.

Considering categories methodologically across this project, one of their major aspects for potential improvement was the incorporation of intersectionality. Study 1a was the only explicitly qualitative study considered and studies 2, 3 and 4 all focused only on responses to a single identity trait rather than any protentional overlaps or intersections. I think a necessary continuation of this work would include an extension beyond category driven analysis through further qualitative work as well redesigned versions of the final studies that are more explicitly intersectional. While these studies were beyond the scope of this project, its intuitively quite simple to consider a continuation of this work that asks not only how possible it is to depict a

woman as a leader or a person of color as a leader, put more specifically a black woman as a leader or perhaps even a disabled woman of color. Identity concepts are intersectional. They overlap in a way that might be at times additive, multiplicative, or wholly unique. What is the consequence of compounding identity traits that are already represented beyond the scope of the default? A likely hypothesis is that these intersections of marginalization are even less represented or associated, and clearly this hypothesis has roots in the theory work of intersectional feminists who offer a qualitative and abstract understanding of what it means to exist within the margins of intersection but beyond the scope of default possibilities.

Another imperative extension of this work is to truly return to the essential and ephemeral aspect of these central questions. If time had allowed, I would have ideally returned to a more methodologically qualitative form of investigation to consider what the true implications of these conclusions are. Living with an identity that is represented beyond the default and attempting to assume a social role you are not associated with according to normative expectations is not a statistical experience. This work utilizes quantitative data to illuminate the reality that these representations exist, but this project stops short of the larger, and perhaps most important question that this conclusion begs us to consider. If some people are not offered the same assumptions of possibility because of the way they identify or are identified (whether this identity is the choice to have a visible tattoo or the un-changeable fact of being marked by a physical disability) what does it mean to experience impossibility? What is the impact on one's ability to live and survive in certain environments that they are not normatively represented in? How do our perceptions of possibility limit the futures of those around us, and what might we be able to do to offer everyone a more equitable access to representation?

While these questions may be far beyond the scope of any singular project, they contextualize the real purpose of this investigation. Regardless of our potential lack of ability to control implicit biases, there is a consequence of the power of expectation. While we might not be able to subvert our own associations, we surely can and should be aware of them. The title of this project is derived from an excerpt of Judith Butler's "Undoing Gender" in which she reflects on being asked "what is the use of increasing possibilities for gender?" Her response is as follows:

Possibility is not a luxury; it is as crucial as bread. I think we should not underestimate what the thought of the possible does for those for whom the very issue of survival is most urgent. If the answer to the question, is life possible, is yes, that is surely something significant. It cannot, however, be taken for granted as the answer. That is a question whose answer is sometimes "no," or one that has no ready answer, or one that bespeaks an ongoing agony. For many who can and do answer the question in the affirmative, that answer is hard won, if won at all, an accomplishment that is fundamentally conditioned by reality being structured or restructured in such a way that the affirmation becomes possible. (Butler 29)

This project offers a new angle on how and why exactly possibility remains so exclusive to some. However, ultimately, I hope that these results continue to enlighten our assumptions of the privilege of possibility to also acknowledge that there is still hope that these conditions are not a stagnant reality. The default heuristics we implement in any cognitive function are not reflective of a mandatory belief, rather it is our own beliefs that collectively impact the reality we choose to live in. I embrace an interpretation of the broader results of this project that suggest not only that our beliefs have the power to limit the possible, but also that our ability to merely

represent something has the chance to make it real. Representations of identity might reflect expectations of the default, but surely this also confirms that the default is not predetermined. While these beliefs may not be mutable, the purpose of this work is to mobilize these results to better understand the underlying judgments of identity and possibility which guide our actions and behaviors. Thus, we can consider our beliefs as having social power beyond the scope of our internal views. In this sense possibility is truly not a luxury to be taken for granted, but perhaps it is an opportunity we have the power to offer one another.

## Sources & References

- Bear, Adam, et al. "What Comes to Mind?" *Cognition*, vol. 194, Jan. 2020, p. 104057. *ScienceDirect*, <https://doi.org/10.1016/j.cognition.2019.104057>.
- Beyond Methodology : Feminist Scholarship as Lived Research*. Bloomington : Indiana University Press, 1991. *Internet Archive*, <http://archive.org/details/beyondmethodolog0000unse>.
- Cohen, Cathy J. "Punks, Bulldaggers, and Welfare Queens: The Radical Potential of Queer Politics?" *GLQ: A Journal of Lesbian and Gay Studies*, vol. 3, no. 4, May 1997, pp. 437–65. *Silverchair*, <https://doi.org/10.1215/10642684-3-4-437>.
- Collins, Patricia Hill. *Intersectionality as Critical Social Theory*. 2019. *read-dukeupress-edu.dartmouth.idm.oclc.org*, <https://doi.org/10.1215/9781478007098>.
- Crenshaw, Kimberle. *Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics*. p. 31.
- . "Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color." *Stanford Law Review*, vol. 43, no. 6, 1991 1990, pp. 1241–300.
- Haraway, Donna. "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective." *Feminist Studies*, vol. 14, no. 3, 1988, pp. 575–99. *JSTOR*, <https://doi.org/10.2307/3178066>.
- Harding, Sandra. "'Strong Objectivity': A Response to the New Objectivity Question." *Synthese*, vol. 104, no. 3, 1995, pp. 331–49.
- Jaggar, Alison M., and Scott Wisor. "Feminist Methodology in Practice: Lessons From a Research Program." *Just Methods: An Interdisciplinary Reader*, edited by Alison Jaggar, Paradigm, 2013.
- Jayarathne, Toby, and Abigail Stewart. "Quantitative and Qualitative Methods in the Social Sciences: Current Feminist Issues and Practical Strategies." *Mixed Methods*, SAGE Publications Ltd, 2006, pp. 267–267. *SAGE Knowledge*, <https://doi.org/10.4135/9781446262566>.

Judith Butler. *Gender Trouble : Feminism and the Subversion of Identity*. Routledge, 2007.

EBSCOhost,

<http://dartmouth.idm.oclc.org/login?url=https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,url,uid&db=e000xna&AN=881409&site=ehost-live&scope=site>.

---. *Undoing Gender*. Routledge, 2004. EBSCOhost,

<http://dartmouth.idm.oclc.org/login?url=https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,url,uid&db=e000xna&AN=110587&site=ehost-live&scope=site>.

Lei, Ryan F., et al. "How Race and Gender Shape the Development of Social Prototypes in the United States." *Journal of Experimental Psychology: General*, Dec. 2021. 2022-15775-001,

EBSCOhost, <https://doi.org/10.1037/xge0001164>.

Lei, Ryan F., and Marjorie Rhodes. "Why Developmental Research on Social Categorization Needs Intersectionality." *Child Development Perspectives*, vol. 15, no. 3, 2021, pp. 143–47. Wiley

Online Library, <https://doi.org/10.1111/cdep.12421>.

Levy, Gary D., and Marianne G. Taylor. "Traditional and Evaluative Aspects of Flexibility in Gender Roles, Social Conventions, Moral..." *Child Development*, vol. 66, no. 2, Apr. 1995, pp. 515–31.

EBSCOhost, <https://doi.org/10.2307/1131594>.

McKittrick, Katherine. *Dear Science and Other Stories*. 2021. read-dukeupress-

[edu.dartmouth.idm.oclc.org](http://edu.dartmouth.idm.oclc.org), <https://doi.org/10.1215/9781478012573>.

---. "Mathematics Black Life." *The Black Scholar*, vol. 44, no. 2, June 2014, pp. 16–28. DOI.org

(Crossref), <https://doi.org/10.1080/00064246.2014.11413684>.

Nash, Jennifer C. *Black Feminism Reimagined: After Intersectionality*. 2019. read-dukeupress-

[edu.dartmouth.idm.oclc.org](http://edu.dartmouth.idm.oclc.org), <https://doi.org/10.1215/9781478002253>.



- . "Institutionalizing the Margins." *Social Text*, vol. 32, no. 1, 2014, pp. 45–65. *DOI.org (Crossref)*, <https://doi.org/10.1215/01642472-2391333>.
- . "Re-Thinking Intersectionality." *Feminist Review*, no. 89, 2008, pp. 1–15. *ProQuest*, <https://doi.org/10.1057/fr.2008.4>.
- Phillips, Jonathan, and Fiery Cushman. "Morality Constrains the Default Representation of What Is Possible." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 114, no. 18, 2017, pp. 4649–54.
- Prasada, Sandeep, and Elaine M. Dillingham. "Principled and Statistical Connections in Common Sense Conception." *Cognition*, vol. 99, no. 1, Feb. 2006, pp. 73–112. *ScienceDirect*, <https://doi.org/10.1016/j.cognition.2005.01.003>.
- Susan Stryker, et al. "Introduction: Trans-, Trans, or Transgender?" *WSQ: Women's Studies Quarterly*, vol. 36, no. 3–4, 2008, pp. 11–22. *DOI.org (Crossref)*, <https://doi.org/10.1353/wsq.0.0112>.
- Tuana, Nancy. "Revaluing Science: Starting From the Practices of Women." *Feminism, Science, and the Philosophy of Science*, edited by Lynn Hankinson Nelson and Jack Nelson, 1996, pp. 17–35.

## Supplementary Materials

















Student



Teacher



Victim