The Philosophical Quarterly Vol. 0, No. 0 *2024*
ISSN 0031-8094 https://doi.org/10.1093/pq/pqae092

Playing with labels: Identity terms as tools for building agency

Elisabeth Camp[1] and Carolina Flores[2]

[1] *Rutgers University, New Jersey, USA and* [2] *University of California, Santa Cruz, USA*

Identity labels like "woman", "Black," "mother," and "evangelical" are pervasive in both political and personal life, and in both formal and informal classification and communication. They are also widely thought to undermine agency by essentializing groups, flattening individual distinctiveness, and enforcing discrimination. While we take these worries to be well-founded, we argue that they result from a particular practice of using labels to rigidly label others. We identify an alternative practice of playful self-labelling, and argue that it can function as a tool for combating oppression by expressing and enhancing individual and collective agency.

Keywords: identity labels; essentialism; playfulness; self-construction; identity politics; social categories.

I. Identity labels as shackles for selves

Identity labels—terms like "woman," "Black," "mother," "queer," "LGBTQIA+," "BIPOC," "libertarian," "Christian," and "evangelical"—are pervasive in both political and personal life. They are central tools for bureaucratic management, and also resources for self-understanding and interpersonal communication. We use them personally to introduce ourselves, recount stories, and find friends and colleagues. We use them politically to form coalitions, identify enemies, and predict elections. And we use them bureaucratically to regulate and allocate resources and services. Yet identity labels also elicit widespread negative reactions. At a personal level, many

Correspondence to: Carolina Flores (caro.flores@ucsc.edu)

people chafe at being reduced to a type or classified in terms they find unfitting, while politically, their use by the bureaucratic state is often thought to subserve oppressive projects of surveillance and control (Foucault 1977).

The intuitive unease with identity labels has been backed up by a rich vein of psychological investigation and theorizing.[1] In particular, identity labels have been argued to distort our thinking about groups, by engendering essentialism: the belief "that members of categories share a fundamental nature that grounds a range of common properties" (Leslie 2014: 211). This essentialist thinking about groups has been argued to lead us to ignore significant variation among individuals (e.g. Gelman and Heyman 1999; Gelman 2003; Ritchie 2021a). And these distortions to our thinking about groups and individuals have been argued in turn to distort our interpersonal and political interactions, by creating divisive binaries. For instance, the use of labels has been found to increase preference for in-group members, as well as the generalization of positive information to in-group members and negative information to out-group members (Bigler, Jones, and Lobliner 1997; Baron and Dunham 2015; Liberman, Woodward, and Kinzler 2017 for an overview). More generally, a long-dominant strand in social psychology treats essentialism as a central root of prejudice (Allport 1954). Given this breadth of pernicious effects, some have worried that centering social identities inevitably entrenches and amplifies oppression (Brown 1995, cited in Ritchie 2021c). Perhaps we would be better off eliminating social labels altogether (Haslanger 2011; Leslie 2017; Dembroff and Wodak 2018, 2021).[2]

We take these concerns seriously. We agree that identity labels have deeply pernicious effects, both cognitively and socially. We think it is vitally important to explain these effects theoretically and to contest them practically. However, we have argued elsewhere that blaming those effects on a metaphysical commitment to essentialism mislocates their source (Camp and Flores in press). Instead, these effects arise from identity labels' function as interpretive *frames*, which guide intuitive cognition in profound ways. Labels can perform this function even in the absence of metaphysical commitment; and when such

[1] We are indebted to Ritchie (in press) for this discussion of objections to identity labels and for references to relevant work in psychology. See also Alcoff (2005: chs 1–3), for a critical survey of criticisms of appeals to identity.

[2] Indeed, some have argued that social category *membership* is already in itself "a restriction on human agency" (Bernstein 2024). Along similar lines, Jenkins (2023) considers that an individual can be the victim of ontic injustice simply in virtue of being a member of a social kind when the constraints and enablements which constitute (or partially constitute) the kind, are wrongful to the individual. However, Jenkins allows that identification kinds, which are connected to agential self-identification, need not be oppressive. For other recent discussions of the notion of social identity, see Jenkins (2016), Ásta (2018), Jenkins (2019), and Andler (2022, 2023). Our concern in this paper is with the normative profile of thought and talk about social identities, not of their metaphysics, and there is not a straightforward one-to-one connection between the two (Barnes 2019).

commitment is present, it is better viewed as a consequence than a source of interpretive distortion. A more accurate understanding of identity labels' function better equips us to combat their pernicious effects.

In this paper, we take a more positive tack. Identity labels' pernicious effects are most palpable when they are applied *rigidly*, by others, to members of marginalized and minoritized groups. We focus on an alternative practice: of self-applying identity labels with a playful attitude.

We first explain how identity labels' framing function leads agents to attend, explain, and evaluate themselves and others in terms of social identities (Section II), and how they encode socially constituted scripts for enacting personae and interacting with others (Section III). As tools for guiding intuitive interpretation and action, identity labels can be deployed in multiple ways, each with its own profile of risk and reward. Where rigid other-labelling especially risks flattening and oppressing agency, playful self-labelling has special potential to enhance agency for complex selves (Section IV). We highlight some payoffs of playful self-labelling for agency in Section V and some perils in Section VI, and we offer some tentative recommendations how individuals and collectives should aim to engage with identity labels in Section VII.

In the end, we hope to add playful self-labelling to a more general trend of highlighting positive roles for identity-centric forms of thought and talk.[3] Much of the discussion about social identities over the last decades has rightly focused on the profound harms that are engendered by slurs, prejudice, implicit bias, and epistemic injustice. We believe it is time to deploy identity-centric resources to resist such harms, by focusing and scaffolding liberatory practices.

II. Identity labels as framing devices

Identity labels vary widely, along multiple dimensions. Some denote deeply established categories for broad demographic kinds like gender, race, ethnicity, nationality, and sexuality (e.g. "man," "Asian," "Latino," "BIPOC," "gay," and "LGBTQIA+"). Others denote professional or social roles ("professor," "mother," and "evangelical"); social class or status ("working class," "socialite," "first-gen/FIGLI," and "hoodlum"); or physical or mental ability ("disabled," "Deaf," "autistic," "neurodivergent," "athlete," and "genius"). Some denote fine-grained categories anchored to restricted subcommunities

[3] See, *inter alia*, Saul (2017) and Ritchie (2019) on reasons to use race and gender generics to describe oppression; Hernandez and Crowley (2024) on joyful uses of gendered language in trans communities; Hernandez (2021) on how attending to gender matters for gender affirmation; Bell (2024) on how gender norms can scaffold authenticity; Ritchie (in press) on reasons to introduce identity labels; and Dembroff and Saint-Croix (2019) on agential identities.

("anarchosyndicalist," "twink," and "thrash metal"). Some ("crip" and "libtard") are obviously valenced, while others at least appear neutral. And some ("woman" and "Black") refer to social groups whose members appear to share many deep, perhaps natural properties, while others ("frosh" and "homeless") refer to groups whose members patently share only contingent or fleeting features.

We think it is important to track the variation among such labels while also accounting for their commonalities. As a working umbrella definition, we take identity labels to be compact lexicalized expressions that purport to pick out a social identity or group (Ritchie in press).[4] What do identity labels as a kind *do*, and how might it explain the pernicious effects gestured to in Section I?

Most obviously, identity labels sort individuals into types: labelling someone with "woman" classifies them as a woman. On a simple explanation, their pernicious effects follow directly from this sorting function: identity labels set up boxes that are metaphysically, morally, and/or politically unjustified, and/or force people to be either in or out of some particular box. We agree that identity labels often do these bad things. But we reject it as a general explanation because we don't think it applies to all types or uses of identity labels. Labels like "socialite" or "libtard" don't necessarily entail any robust metaphysical beliefs (Camp and Flores in press), and labels like "homeless" or "autistic" that are not morally harmful still generate many of the objections canvassed above.

In this section, we argue that identity labels are cognitively powerful because they are *frames*: representational vehicles that orient attention, explanation, and evaluation by crystallizing a perspective. In a minimal sense, categorization itself already entails a perspectival frame. Identity labels are embedded within larger presupposed taxonomies (Carnap 1950): for instance, "woman" contrasts with "man"—and also, for many contemporary Americans, with "non-binary person." Employing any one of these gender labels presupposes that the larger taxonomy in which it is embedded tracks stable clusters of lower-level features that are worth monitoring (Boyd 1999). This orients attention to the corresponding identity property and underwrites inferences about further properties that group members are likely to possess (Neufeld 2024, in press).

But identity labels also function as frames in a more robust sense. This is illustrated most clearly by non-literal uses. Thus, we sometimes apply identity labels to people we know don't belong to the relevant category, as when the founder of the Boy Scouts commented on a meeting with African political leaders that "the only man in the room was that woman" (McConnell-Ginet 2002; Leslie 2015); or when someone labels a friend who always

[4]Labels can be used in nominalized or adjectival forms ("Elliott is *a* queer" vs "Elliott is queer"). Nominalized uses promote essentialist patterns of inference more than adjectival uses (Ritchie 2021b).

carries snacks and Band-Aids "the group mom." Other times, we apply labels to category-matching individuals without imputing or endorsing the underlying taxonomy: thus, one of us is fond of sending astrology memes despite vehemently rejecting astrology as tracking well-founded clusters of causal relations. In such cases, *something* is accepted by participants, which is distinct from the claim that the target instantiates the feature(s) constitutive of the corresponding kind. What might this something be?

Classic work on concepts in psychology points the way. Starting with Rosch (1973), it has been abundantly demonstrated that categorization is pervasively facilitated by prototypes or stereotypes.[5] In particular, people are faster and more reliable at judging whether something falls under a concept (e.g. *bird*) when it instantiates many of the corresponding kind's prototypical features (e.g. when it is a robin rather than a penguin). Stereotypes cannot themselves *be* concepts, given that they do not determine category membership: thus, subjects also acknowledge that penguins are birds, and judge non-prototypical odd numbers like 29 to be odd (Armstrong, Gleitman, and Gleitman 1983). Nor do stereotypes systematically compose (Fodor and Lepore 1996). Nonetheless, empirical evidence strongly suggests that using or hearing a label activates a corresponding cognitive structure, which then plays an active role in guiding the rest of one's intuitive thinking about the group and its members.

Moreover, for many social identities these cognitive structures don't just assign putative features (or statistical probabilities of possessing features). They also include a take on which features it is *fitting* for individuals to have, *qua* members of that kind. These normative expectations are revealed especially clearly with dual-character concepts (Knobe, Prasada, and Newman 2013; Del Pinal and Reuter 2017), as reflected in generics and claims about how "real *G*s" behave (Leslie 2015). Thus, we might think that *real* scientists are passionate about increasing our collective understanding of reality even if we know that most people employed in scientific labs don't meet that expectation. More generally, even labels that do not conventionally encode teleological norms, like "Asian," still often activate normatively laden expectations about a paradigmatic way for *G*s to be, where these features may be positively valenced, like *smart*, or negatively valenced, like *greedy*. (Think here of discourse about whether Barack Obama is "really Black.")

These expectations about how it is proper for group members to be locate individuals within a complex social–normative space (Butler 1988), in

[5] Although prototypes and stereotypes are sometimes distinguished, we will use the terms interchangeably, to refer to culturally widespread cognitive structures that attribute many features to a kind probabilistically. See Margolis and Lawrence (1999) and Murphy (2002) for classic discussions of concepts in psychology, and Beeghly (2015) for discussion of the notion of stereotype. See Camp (2015) for discussion of the difference, and relation, between concepts and characterizations (of which stereotypes and prototypes are a special case).

ways that often generate visceral emotional and aesthetic responses. These responses are especially palpable in our reactive attitudes toward nonconforming individuals, ranging from discomfort to disgust. (Think here of many people's response to visible hair on women's armpits and legs, and to the absence of hair on men's.) These expectations are often also encoded in social norms, institutions, and law. (Think here of how our society permits or encourages ridiculing, ostracizing, and sometimes legally punishing men who defy norms of masculinity.)

So, in addition to denoting group categories, identity labels activate intuitive mental representations that attribute both descriptive and normative features. Further, these representations are not just sets of (probabilistic, normed) attributed features: they are internally structured. In particular, identity labels shape patterns of attention and prediction. Different labels render different subsidiary features *salient*: stereotype-conforming features are more likely to be smoothly encoded and recalled, while flagrantly violating features call out for explanation and stereotype-neutral features are likely to be ignored.[6] And in turn, these patterns of attention affect which further features we expect individuals to possess, and how we explain them (Munton 2023).

More specifically, some properties are more *central* than others: they are represented as causing, explaining, and/or normatively grounding many other features (Neufeld 2024). The limit case is metaphysical essentialism, in which group members are taken to share an unobservable essence.[7] However, the properties included within networks of centrality, and the grounds of connection, often extend well beyond anything plausibly attributable to a purported essence. For example, experimental evidence suggests that the prototypes for both female and male professors encode both the properties *smart* and *hard-working*; however, for female professors, *smart* is taken to causally depend on *hard-working*, but not so for male professors (Del Pinal, Madva and Reuter 2017; Del Pinal and Spaulding 2018). Thus, identity labels prioritizes some explanations over others.

Taking stock, identity labels don't just categorize; they also activate complex representations that impute non-extension-determining descriptive and normative features and shape patterns of attention, explanation, and prediction. We can encapsulate these effects by treating identity labels as *frames* (Camp 2019, 2020): representational vehicles that schematize and stabilize cognition by encoding intuitive focal principles that crystallize open-ended

[6]Ella Whiteley (2023) argues that surplus attention can constitute an inherent harm, as reflected in Basquiat's reproaching quip "I am not a Black artist. I am an artist." In other cases, attentional surplus might be a way of challenging marginalization, as with loving attention in contexts of gender affirmation (Hernandez 2021).

[7]See Neufeld (2022) for an overview of recent work on psychological essentialism.

interpretive perspectives.[8] Other familiar species of frame include metaphors, like "Peter is a bulldozer" (Camp 2006, 2017b); telling details, like "He was climbing ladders before he could talk" (Camp 2007), slogans, like "Minds are computers" or "Boys will be boys" (Camp 2019), and narratives (Camp 2024).

So far, in introducing the intuitive cognitive effects of identity labels, we have appealed to stereotypes, understood as culturally entrenched ways of thinking about groups. Attending to these alternative species of frames highlights the need to generalize the relevant cognitive structures beyond stereotypes, to include more bespoke cognitive structures constructed by individuals, and structures that apply to individuals as well as groups. We use *characterization* as a covering term for culture-wide or idiosyncratic cognitive structures, representing any sort of entity. We understand *perspectives* as dispositions to form and update characterizations in application to a certain subject domain; and *frames* as tangible focal principles for expressing and regulating perspectives.

All frames regulate intuitive cognition by crystallizing a perspective, which in turn governs the formation and modification of characterizations by guiding attention, explanation, and response. Different species of frame accomplish this in different ways, with distinct effects. Identity labels work by centering a social identity associated with the label's denoted kind *G*. Thus, deploying the labels "woman," "scientist," or "libtard" reminds or informs an interpreter how they should orient their overall thinking toward the target. It focuses attention by highlighting features that either conform to or markedly violate its associated social identity while backgrounding neutral features. It guides interpretation of attributed traits and behaviors by parsing lower-level features in identity-conforming terms, imputing identity-supporting explanations, and projecting additional identity-conforming features. And it filters emotional, aesthetic, moral, and sometimes legal response by activating corresponding social norms.[9]

Frames in general are useful interpretive tools because they streamline the flux of intuitive cognition. Most perspectives are messy, malleable, amorphous, and tacit. Frames *stabilize* intuitive thinking by offering tangible touchstones that can be accessed and deployed across contexts. They also *schematize* intuitive thinking by encapsulating relatively simple, coherent interpretive principles: heuristics that tell you "all you need to know"—at least to a first approximation—for dealing with the target. As such, they can guide interpretation even if you are largely ignorant about the topic, and even if you encounter new information that conflicts with many of your prior assumptions about it.

[8]Other theorists take the complexity of our thought and talk about social identities to derive from the metaphysics of social identities, claiming that identities are kinds that make ineliminable reference to social norms and agents' relationships to them (e.g. Jenkins 2016; Andler 2022, 2023; Jenkins 2023). Although this is compatible with our view, we remain neutral on the metaphysics of the kinds picked out by identity labels.

[9]For more on identity-centric frames, see Camp and Flores in press.

In the simplest cases of identity-centric framing, a label frames a particular target individual *A* by both sorting *A* as a member of *G* and subsuming *A* under a stereotype for *G*s. But not all cases are this simple. On the one hand, as we've seen, a stereotype can be deployed without commitment to its literal application, as in metaphorical, extended, and non-committal uses. Conversely, identity labels can guide interpretation in the absence of a stereotype. Slurs are an illustrative special case. While many slurs are associated with rich stereotypes backed by an imputed essence, some, like "gaijin" or "midget," lack robust stereotypes (Jeshion 2013). Rather, what slurs in general do is signal the speaker's allegiance to a perspective that centers *G* for thinking about *G*s (Camp 2013). For slurs, that perspective is also derogatory: it treats members of *G* as "others," unworthy of respect. Other identity labels, like "freedom fighter," are laudatory or (more) neutral. But they are all like slurs in treating their associated identity, *G*, as highly central for thinking about *G*s: they impute a high degree of cohesion among members of *G*; parse attributed features in *G*-consonant terms; select *G*-consonant features as mattering; and treat being *G* as explaining and warranting further features and responses.

The combined force of stability and schematization makes frames powerful interpretive tools across domains as various as science, politics, personal relationships, and art. At the same time, these qualities also make frames seductively satisfying and self-reinforcing, by blinding us to all the features we have failed to notice and all the ways we have projected features and explanations beyond our experience. When we don't notice that we are interpretively limited in these ways, the result is *myopic complacency* (Camp in press). This general risk for frames already goes a long way toward explaining identity labels' pernicious effects. Given that identity labels encode a simple, culturally ubiquitous principle for sorting people into coarse-grained groups in terms of an intuitive stereotype, the risks of over-projecting identity-conforming features while ignoring or explaining away non-conforming ones, and of activating ungrounded reactive attitudes, are heightened even further. The likely result is flattening, distortion, and discrimination.

III. Identity labels as social scripts

So far, we've focused on identity labels' role in regulating individual cognition. But identity labels are also essentially *social* tools. They pick out an aspect of the social world (social identities); they are typically governed by culturally dominant characterizations (stereotypes); and they are often applied in social contexts ranging from direct, informal interactions to formal political decisions.

In this section, we argue that identity labels are social tools in a deeper sense: they function to regulate social performance. More specifically, as

identity-centric frames, they sculpt social traits, behaviors, and interactions in ways that profoundly affect who we are.[10]

To understand how identity labels can sculpt action in concert with interpretation, we can turn to Thi Nguyen's (2019, 2020) notion of an *agential mode*. For Nguyen, an agential mode is a focal interest, plus a pattern of attention for recognizing opportunities and obstacles for achieving it, and a set of skills for pursuing it. Like perspectives, agential modes are intuitive, open-ended, complex, structured dispositions that guide attention and response. But where perspectives guide cognition, agential modes guide action. Further, like perspectives, most agential modes are messy, malleable, amorphous, and tacit. Nguyen makes a powerful case that *games* are frames for agential modes: they encode an intuitive way of engaging with the world in a stable, schematized, publicly accessible form. Thus, much as reading fiction can train our actual interpretive habits (Nussbaum 1992; Camp 2017a), so too can playing games train our actual agential habits.

To apply the notion of agential modes to identity labels, we note that many labels are associated with *scripts* (Haslanger 2011): culturally entrenched schemas that present a set of focal interests and ways of achieving them. Thus, labelling someone as a doctor highlights the goal of curing patients and the techniques of modern medicine, along with more fine-grained scripts for diagnosing illnesses, talking with patients, filing paperwork, and so on. Labelling as a runner highlights the goal of running, perhaps along with goals like racing, as well as training practices and more fine-grained scripts for avoiding and managing injuries, eating energy-conducive foods, and so on. "Doctor" and "runner" encode these goals fairly directly, in virtue of the identities they denote. Labelling as a woman accomplishes a similar effect more indirectly, activating a stereotype that highlights culturally entrenched gender goals like raising children, along with skills and scripts for achieving them like getting married, attending to other people's needs, and managing domestic labor. More generally, labels for identities that lack any teleological basis, like ethnic or sexual identities, often present certain interests, skills, and scripts as characteristic and appropriate for that identity.

Treating identity labels as culturally entrenched frames for agential modes forges a connection between intuitive interpretation and action by way of scripts. However, many agential modes and scripts, such as backpacking or playing Sudoku, are not especially social. To connect identity labels to more robustly social performance, we appeal to Erving Goffman's (1956) work on persona construction.

Goffman takes forming and maintaining a stable social identity to be a central activity of our lives. He models everyday interactions as a species of

[10]Identity labels are thus a species of cultural devices that sculpt individual cognition. See Haslanger (2019) for other such devices.

dramatic performance, in which we establish and maintain a persona that enables us to pursue our goals. In order to interact with others in predictable and profitable ways, we must present personae that are intelligible to them, and that we can sustain across widely varying, often unpredictable circumstances. Moreover, a significant part of effective performance involves microbehaviors, like vocal inflection, posture, facial expression, and dress, over which we have little direct voluntary control and which are often affectively charged. Given this, a sustainable persona cannot be put on like a mask or recited like a fixed script. Rather, it must be grounded in an open-ended, embodied, flexible, ability to form nuanced, intuitive interpretations of social situations and to respond with nuanced, intuitive performances in real time. As Judith Butler (1988) emphasizes, over time this continual micro-enactment of personae shapes our actual bodies, and our perceptions of our own and others' bodies, in profound and pervasive ways, creating the appearance and perhaps the reality of a deeply carved nature.

In effect, then, personae are social agential modes: intuitive, open-ended, structured, and deeply embodied dispositions to inhabit the social world in a certain way. While personae, like perspectives and agential modes, are often messy, malleable, amorphous, and tacit, we can think of *characters* on analogy with stereotypes: schematized, stabilized, culturally entrenched frames for performing, as opposed to interpreting, a social identity.

In the simple case, identity labels that encode stereotypes, like "mother," "woman," "doctor," or "Asian," also offer robust scripts for enacting and interacting with their associated characters. However, we saw in Section II that the route from labels to intuitive interpretation does not always trace a straightforward path through stereotypes. So too with personae. On the one hand, not all culturally entrenched characters are encoded in compact lexical expressions: consider Hollywood tropes like *Pretty Woman*'s hooker with a heart of gold. Conversely, not all personae are anchored in culturally ubiquitous characters: think of your family's most beloved (or reviled) grandparent, or your high school's most feared (or adored) teacher. Here again, we take the simple case to offer the key for unlocking a more general phenomenon: how labels can express rich, intuitive, flexible templates for interpreting, enacting, and interacting with social identities.

Explaining the full sense in which identity labels are social tools that regulate social performance requires extending these points at least one more step. Identity labels don't just encode scripts for performing characters, which agents are then free to adopt in pursuing their own, independently constituted interests. Nor do they just encode stereotypes and scripts for others to use in interpreting and interacting with the individuals they label. Because labels play these interpretive, enactive, and interactive roles, they play a crucial causal role in creating and reinforcing those traits and interests within

agents themselves, through the sorts of looping effects described by Ian Hacking (1995).

Hacking argues that the introduction of an identity label for a previously inchoate constellation of features facilitates the creation of formal and informal institutional structures, which constitute a social 'place' for those to whom the label is applied. Hacking was especially interested in the constitution of identities in mental illness, which he illustrated with the emergence of "multiple personality disorder" in the 1970s. The creation of this label provided psychiatrists with a template for diagnosing (and billing for) a previously amorphous constellation of symptoms. This template schematized and stabilized that constellation of symptoms and raised it to salience among psychiatrists and in the general population. Unhappy people who hoped to alleviate their suffering increasingly came to manifest this distinctive constellation of symptoms. And the prevalence of this constellation in turn led psychiatrists to seek a unified underlying explanatory cause. Ultimately, this process culminated in the introduction of a new "way to be a person" into the social terrain (Hacking 2007; cf. Foucault 1977; Mallon 2016).[11]

Hacking's emphasis on expert knowledge and bureaucratic institutions in causal looping effects amplifies two points about the role of identity labels in constructing persons already implicit in Goffman. First, identity labels play an integral role in the construction and perpetuation of *ideology*: what Haslanger (2021: 63) describes as the "cluster of concepts, background assumptions, norms, heuristics, scripts, metaphors (and so on) that enable us to interpret and organize information and coordinate action, thought and affect," and that thereby transform mere material phenomena into meaningful, functional entities. In particular, identity labels implement robust causal nodes within a wider nexus of practices for transitioning among perception, thought, and action. Introducing and deploying a label makes an already nascent causal node available for reinforcement through informal and formal institutional mechanisms, and for extension into a richer network of diagnostic and imaginative entailments (Tirrell 2012).

Secondly, agents are not just slotted into certain social niches because they antecedently instantiate the corresponding set of extension-determining features. Rather, the menu of available identity labels and the reiterated application of labels to individuals sculpts them into becoming the particular agents they become, in a dynamic, self-reinforcing process. Identity labels carve out a social place within a menu of options for how to be. Deploying an identity label at all manifests and reinforces a general commitment to the utility of

[11] Hacking is suspicious of the pseudo-Freudian explanations offered by psychiatrists at the time—suspicions that motivated a shift to the currently preferred label of "dissociative identity disorder." Hacking also discusses the emergence of "autistic," which has endured similar controversies.

carving out that place, as embedded within that menu of options. And applying it to a particular individual locates them in that place, and causes them to change so as to better fit that location's prescriptions.

Sometimes the act of placing is palpably forcible, as with slurring (Kukla 2018) and misgendering (Kukla and Lance 2022). Slurs in particular function to flatten identity, rendering the target invisible except insofar as they instantiate the relevant identity (Camp 2013; Jeshion 2013). But even when identity labels are less overtly coercive, their application still parses and highlights the targeted individual's lower-level features in terms of traits associated with the label's characterization.[12] It licenses others to expect that the target will manifest those traits and behave according to its scripts. And it licenses them to deploy correlative scripts in their interactions with them. Together, these expectations and effects "hail" targeted individuals into enacting and embracing the relevant character for themselves (Kukla 2018; Haslanger 2021; cf. Althusser 1971). As the label becomes increasingly entrenched within society in general and in application to that individual, its interpretive and practical utility are further reinforced. Eventually, it becomes the dominant, maximally effective tool for interpreting, enacting, and interacting with anyone who displays anything in the vicinity of its constellation of traits.

IV. Labels as scaffolds for flexible agency

IV.1 From rigid other-labelling to flexible self-labelling

We believe that our analysis of identity labels as frames explains their pernicious metaphysical, moral, and political effects better than the naïve view of labels as categorization devices. Cognitively, identity labels offer handy heuristics that tell you "all you need to know," at least to a first approximation, about people with that identity. Such heuristics can be useful, insofar as they offer focal principles for regulating attention, explanation, and response that stabilize and schematize otherwise overwhelmingly complex bodies of information. But they thereby risk seductive over-simplification, leading interpreters to ignore individual variation and impute unwarranted metaphysical and normative grounds for gut intuitions. Practically, identity labels offer scripts for enacting and interacting with certain social ways to be. Such scripts can be useful, insofar as they highlight opportunities and obstacles for action and streamline

[12] In particular, it is experimentally challenging to identify independent, objective criteria for assessing stereotypes' accuracy, because both in-group and out-group individuals tend to parse the personal character traits they ascribe to particular people in stereotype-conforming terms (Judd and Park 1993). Similarly for micro-behaviors: the same arm movement is more likely to be parsed as a playful swat or a threatening jab depending on the ascribing agent's stereotype of the race of the person performing it (Duncan 1976).

coordination. But they thereby risk erasing non-conforming opportunities and establishing coercive, self-perpetuating causal loops.

We also think our analysis explains a wider range of identity labels and uses. Not all identity labels traffic in metaphysical essences or encode rich stereotypes, and not all applications of labels are believed by either speaker or hearer to apply literally to their targets. Even interpretively "lean" labels that are taken to be literally false or taxonomically ungrounded still activate identity-centric perspectives and agential modes, which regulate intuitive interpretation and action in ways that are only partly under voluntary control.

However, our primary aim in this chapter is positive: we maintain that identity labels can help to build agency. How is this compatible with acknowledging the interpretive and practical risks rehearsed above?

First, although all frames risk myopic complacency, the risks of misinterpretation, coercion, and discrimination are heightened when identity labels are applied by *others*—especially when they are imposed by users who occupy a privileged social status onto targets who occupy a marginalized status, as with slurs and misgendering. This is a key reason we focus on the potentially ameliorative effects of self-labelling below.

For now, we note that even when a label originates from others, adopting it for oneself enacts and amplifies one's agency in certain ways. Taking up an agential identity by publicly self-applying an identity label "play[s] a critical role in self-determination" (Dembroff and Saint-Croix 2019: 587). Self-applying labels for marginalized identities can improve group members' self-esteem (Ritchie 2021c). Indeed, even self-labelling with negatively valenced labels like slurs can increase a sense of one's own power, others' perceptions of that power, and the power of the slurred group (Galinsky et al. 2013).

Our account allows us to explain and expand on these positive effects. Operating in social space according to a label, with its attendant stereotypes and scripts, transforms abstract opportunities and obstacles for action into intuitive affordances which one can navigate in smooth, flexible ways. In particular, as Goffman emphasizes, a robust intuitive grasp of how you "show up" to others enriches your opportunities for engaging strategically with them. At a minimum, it facilitates "going along to get along." With enough wit, it can equip you to exploit others' expectations and blind spots for your own purposes. Further, as Hacking emphasizes, having a label facilitates recognizing oneself, as well as others who fit the same profile, highlighting opportunities for affiliation.

One might object that these agential advantages are golden handcuffs: benefits purchased at an unacceptably high cost. As critics of ideology emphasize, the same mechanisms that enable a modicum of agency also drastically constrain it (Butler 1988: 521; Haslanger 2021: 57). Identity labels in particular highlight opportunities for action, but from a highly restricted menu of options. They stabilize and amplify traits, but risk turning persons into

caricatures. And they streamline interactions and afford camaraderie, but by exaggerating differences. Given these ways in which labels flatten, coerce, and divide, we might vastly prefer a culture that abolishes them entirely (Dembroff and Wodak 2018, 2021).

While we agree that these dangers are genuine and profound, we think the deeper problem lies in *how* labels are applied, not in the use of labels *per se*. Being labelled by others in a way that one rejects and that fails to accurately track the traits that are central to one's self-conception clearly constitutes a violation of autonomy and a moral harm (Kukla and Lance 2022; Whiteley 2023). Again, this is one reason we focus on self-labelling, as an inherently self-determining act. At the same time, as we've just rehearsed, even self-applied labels can tightly constrain agency. How can this be a good thing?

We think that self-applied labels have these effects insofar they are applied *rigidly*. As we understand it, rigid labelling involves applying one totalizing label (or a small cohesive set of labels) *universally*, across a wide range of contexts, with an *essentializing* attitude that endorses the literal applicability of both its associated concept and characterization, that takes it to reflect a central, immutable aspect of one's self, and that treats its associated norms as *imperatives* for action (Camp and Flores in press).

Not all rigid labelling is bad. Some identities, like "monk", may demand rigidity as a constitutive feature or a matter of respect. Some people may fully and authentically fit within under a totalizing label, such as "mother," at least for a chapter of their lives. Some may need to exploit the stabilizing effects of an identity-centric frame, like "recovering addict" or "Class of 2028," to maintain single-minded focus on potent risks or elusive goals. And some may want to celebrate a marginalized identity, like "autistic" or "Deaf," in order to reclaim pride and build coalitions. We return to some of these cases below.

However, we think that much of the intuitive discomfort and theoretical objections to labelling are driven by the recognition of a mismatch between totalizing, rigidly deployed labels and the vital complexity that many selves exemplify. As Walt Whitman intimates in *Song of Myself* (1891: 51), the "I" of our selfhood is often multitudinous:

> Do I contradict myself?
> Very well then I contradict myself,
> (I am large, I contain multitudes.)[13]

Rigid labelling not only fails to do justice to, but risks eradicating, such multi-dimensionality.

[13]Similarly, Portuguese modernist poet Fernando Pessoa writes:
> Countless lives inhabit us …
> I have more than just one soul.
> There are more I's than I myself. (Pessoa 1998).

It is easy to fall into rigid labelling. As we saw in Section II, frames in general are epistemically and causally self-reinforcing, guiding attention, and amplifying traits in ways that confirm their aptness. Moreover, as we saw in Section III, strategic self-interest and social norms encourage inhabiting a consistent, recognizable social niche. But as a label becomes increasingly entrenched in an agent's internal self-interpretation and external self-presentation, it *ossifies*. Rather than functioning as a rough heuristic for handling a multidimensional social flux, it becomes an absolutist algorithm that admits only label-conforming information and affords only label-conforming action. Cognitively, an ossified frame generates myopic complacency and metaphysical essentialism (Camp 2024, in press; Camp and Flores in press). Agentially, it produces a form of value capture, in which complex, textured values are overwritten with a few highly schematized ones (Nguyen 2021).

However, not all uses of labels are rigid, and ossification and value capture are not inevitable. As tools for interpretation and action, identity labels can be deployed in a multitude of ways, with distinct profiles of risk and reward.

IV.2 Strategic and authentic flexibility

A version of fluid self-labelling is already ubiquitous in both private reflection and public presentation. Different environments frequently elicit different labels for the same individual, framing cognition and action under distinct schemas. In some cases, these labels refer to identities that are orthogonal, such as "professor" (labelling a professional role) and "American" (labelling a nationality). In other cases, they refer to identities that are distinct but overlapping, as with "biracial," and, as it may be, "Black," "white," "Asian," and so on. And in still others, the labels might be (roughly) co-extensive, but differ in their associated characterizations, like "lesbian" and "Sapphic." Fluid labelling of social identities reflects a more general cognitive disposition toward contextual modulation: subjects often display significant diversity in their choice of terms to refer to the very same object, and often innovate with new expressions to meet context-specific needs (Brennan and Clark 1996; Armstrong 2016).[14]

This form of fluid labelling is generally passive, elicited by context in ways that are largely unwitting and sometimes unwilling. The potentially pernicious effects of such passivity for agency are illustrated clearly by stereotype threat (Steele 1997), in which eliciting a self-applied demographic label raises stereotype-conforming traits to salience in ways that negatively affect the agent's performance.[15] One effective antidote to stereotype threat is

[14]Thanks to Josh Armstrong for discussion.

[15]The phenomenon of stereotype threat has been widely discussed and hotly contested. While early, simple characterizations are problematic, more recent meta-analyses suggest that

counter-programming: actively introducing alternative labels or other frames like affirmative mantras, in order to shift the profile of contextually salient features in performance-enhancing ways.[16] We suggest taking the antidote more deeply to heart, and deploying it more widely. In a world that is deeply sculpted by identity labels, intentionally deploying multiple labels for oneself can be a powerful tool for resisting the flattening, coercive effects of other-labelling.

One especially common form of strategically flexible self-labelling is *code-switching*: actively switching labels to match distinct contexts, revising the social identities one makes available to others as one moves between contexts (Dembroff and Saint-Croix 2019). For instance, a working-class person from an immigrant family attending an elite university might label themselves as a cousin from the *barrio* when at home, and as an athlete or engineering major when at school (Morton 2014). Code-switching enables agents to smoothly immerse themselves in a diverse range of contexts in order to unlock the benefits that derive from each of them. A commensurate risk is that one sacrifices one's "home" identity for what one thinks of as merely instrumental gains, ending up adrift, without any comfortable, authentic self.

A related form of strategically flexible self-labelling is *double consciousness*, in which one embraces one label as more or less authentically reflecting one's identity but simultaneously "looks at one's self through the eyes of others" (Du Bois 1903: 11) who are anchored in the dominant culture. Double consciousness enables an agent to smoothly anticipate—and exploit—a culturally dominant perspective on themselves without fully capitulating into embodying the character that is projected onto that identity by the larger culture. A commensurate risk is the increased burden of double bookkeeping: running multiple interpretive and performative ledgers simultaneously, while navigating conflicting, often negative valences for important aspects of one's identity.

Finally, agents may engage in *strategic essentialism* (Irigaray 1985; Spivak 2012): temporarily treating a label, or a taxonomic cluster of labels, as reflecting a real, immutable identity in order to communicate effectively with those who hold essentialist commitments, with the goal of building robust political coalitions with other members of a common group—or of differentiating oneself emphatically from non-members.[17] Strategic essentialism enables agents to leverage the single-minded focus and passion characteristic of rigid labelling. A commensurate risk is that what begins as a temporary, instrumental technique devolves into perpetual practice, carrying with it the malign costs of genuine essentialism.

the effect is genuine (see e.g. Nguyen and Ryan 2008; Pennington et al. 2016). For evidence of the efficacy of counter-programming, see e.g. Martens et al. (2006) and Woolf et al. (2009).

[16] Camp (2017b) diagnoses a similar profile of unwilling perspectival shift in metaphorical insults and advocates cognitive and/or conversational reframing as an antidote.

[17] Thanks to Elli Neufeld for encouraging us to discuss strategic essentialism.

This list of forms of strategic self-labelling highlights the diverse profiles of risk and reward that identity labelling can carry—all of which differ markedly from the pernicious effects that have been the focus of theoretical discussions of identity labels to this point. In principle, all these forms of labelling support a form of flexible agency that is more robust than either the rigid acceptance of a bestowed label or the sort of passive fluidity we find with contextual modulation and stereotype threat. A strategic self-labeler deploys a label intentionally, in pursuit of their own enduring goals and in light of their own assessment of the contexts they find themselves in.

At the same time, though, strategically flexible self-labelling is still fundamentally *reactive*: a compensatory mechanism for navigating a hostile environment. As such, there is an important sense in which a strategic self-labeler remains alienated from at least some of their labels. A more ideal context would obviate the need for such flexibility, allowing one's authentic self to show up more fully within and across contexts.[18]

IV.3 Playfulness

To achieve a more active, authentic, and integrated relationship to identity labels, we need a species of flexibility that goes deeper than the instrumental adoption of a frame in the service of a distinct, stable set of autonomous goals.[19] We point to *playful self-labelling* as evincing this kind of flexibility.

The notion of playfulness we have in mind is anchored in Latina feminist thought (esp. Anzaldúa 1987; Lugones 1987; Ortega 2016), which has emphasized the experiences of those who occupy multiple, marginalized identities at the borders between societies or social milieus.[20] Anzaldúa focuses on Chicanas with indigenous ancestry living near the US/Mexico border, who "continually walk out of one culture and into another" (Anzaldúa 1987: 77). Some[21] Chicanas respond to this situation by developing a distinctive identity: the new *mestiza*, who

> is a mixture of all these identities and has the ability, the flexibility, the malleability, the amorphous quality of being able to stretch, and to go this way and that way, add new labels or names which would mix with the others and they would also be malleable. (Anzaldúa 1987: 79)

Thus, the new *mestiza* shifts among diverse ways of thinking and acting—but not by switching between discrete options in a portfolio of agential modules.

[18] Thanks to Gabby Johnson for discussion.

[19] One can also label others in playful ways; we focus on self-labelling because it raises fewer hermeneutic, epistemic, and moral concerns. We return briefly to other-labelling in Section VI.

[20] Ortega (2016) provides a survey and critical interpretation of this tradition, to which we are indebted.

[21] "Some" because, as standard in standpoint theory, it is an achievement to cultivate an insightful relation to the social world from a marginalized position, not an inevitable consequence of occupying that position. See Wylie (2003) and Dror (2023).

Rather, the identity itself amalgamates dimensions of many different personae into an integrated whole. It also incorporates malleability as a higher-order trait: part of the *mestiza* identity is that one stretches and morphs oneself to fit new labels as they come along, while simultaneously stretching and morphing those labels to fit oneself.[22]

Maria Lugones (1987) identifies *playfulness* as a tool for cultivating this trait of malleability, both within a given identity and across identities. Like Anzaldúa, she emphasizes that the social "worlds" to be navigated are often hostile, imposing frames that are alienating and demeaning, involving caricatured identities that are constructed with fixed stereotypes and scripts. Even so, she "affirm[s] this practice [of world-traveling] as a skillful, creative, rich, enriching and, given certain circumstances, as a loving way of being and living" (Lugones 1987: 3). World-traveling, even to hostile worlds, scaffolds perspectival and agential flexibility: one becomes a different person, occupying a different social niche, with a different profile of traits, interests, and skills. The transformation need not be intentional or conscious; but even when it is, it is not a matter of posing or pretending to be someone *else*. Playfulness is open and genuine, not deceptive: one experiences oneself, directly and intuitively, and is treated by others, as simply *being* that person (Lugones 1987: 11). And by moving between worlds, one doesn't just transform one's own identity; one also gains the capacity to empathize with others who may be very different, by seeing their world through their eyes (Lugones 1987: 8).

As we might put it, playfulness is a higher-order attitude for engaging flexibly and malleably with whatever worlds or identities one encounters; it does not itself commit to or involve any particular content or context. As in Lugones's terms puts it, a playful attitude involves

> openness to surprise, openness to being a fool, openness to self-construction or reconstruction … uncertainty, lack of self-importance, absence of rules or a not taking rules as sacred, a not worrying about competence and a lack of abandonment to a particular construction of oneself, others and one's relation to them. (Lugones 1987: 17)

Lugones contrasts playfulness with the kind of role-playing that characterizes "agonistic play," in which an agent maintains a "fixed conception of him or herself" and aims to win by mastering the game's rules (Lugones 1987: 15). Because agonistic play involves a fixed self-conception and aims at competence, while world-traveling requires abandoning these, agonistic players can never "travel" to other worlds in her sense, but only "conquer" them (Lugones 1987: 16): their goal is to win, not to *be*. Put in our terms, even if the agonistic player acquires intuitive competence with the game's agential mode, they are

[22] In later work, Anzaldúa became less friendly to the use of identity labels, aspiring to "a sense of self (*la nepantlera*) that does not rest on external forms of identification" (Anzaldúa 2009: 302).

still like the strategic self-labeler insofar as they instrumentally adopt a highly scripted agential mode that is dictated by the game context, while keeping their "authentic" self safely quarantined from it.[23] By contrast, an agent who is playful in Lugones's sense engages in a genuinely exploratory act of "trying on" a frame or persona.

Applying playfulness to labelling specifically, we can say that a playful self-labeler feels their way into subsuming their own self under the label's rubric, with no fixed assumptions about what traits the label will entail: what its application will reveal or erase about their self, and no fixed end point for successful or failed application. Playful self-labelling is *open-minded*: the labeler is prepared to discover lurking but heretofore unknown or neglected traits within themselves. Indeed, they are prepared for their traits to be re-parsed at a more fundamental level, and to be highlighted, connected, and evaluated in quite different configurations, as they mold themselves to fit the label's contours. Playful self-labelling is also *open-ended*: the label elicits an indeterminately bounded range of traits, scripts, and skills within any given context; and it invites creative performance of different traits, scripts, and skills across contexts. And playful self-labelling is *humble*: it embraces uncertainty and a lack of sovereignty in the application of any given label. As Lugones puts it, it involves a lack of abandonment to any particular construction of oneself, and an openness about one's own lack of competence and completeness.

Playful labelling thus contrasts markedly with rigid labelling, which treats a single label, or a small fixed set of labels, as literally describing and aptly framing one's "true" self over time and across contexts, in virtue of reflecting a central, immutable essence that dictates how one ought to behave. It also contrasts with strategic labelling, which retains a felt gap between self and label that produces some cloaking or distortion of one's true self. More generally, playfulness helps to combat the myopic complacency that is the besetting sin of framing across the board (Camp 2024). These are all virtues. However, given how easy it is to fall into rigid labelling, achieving playfulness takes considerable effort. In Section IV.4, we explore some ways to facilitate it.

IV.4 Varieties of playful labelling

If one focuses exclusively on the use of identity labels in institutional bureaucracies, political debates, and slurring graffiti, playful labelling might seem like an obscure or *recherché* practice. On the contrary, we take it to be quite common, especially within subcommunities such as borderlands and queer spaces.[24] We also think the practice is manifested in many different guises.

[23] See Camp (2022) for related criticism of Nguyen's notion of striving play in scaffolding a robustly autonomous form of agency.

[24] See Hernandez and Crowley (2024) for discussion of playful labelling in trans and queer communities.

Indeed, given that playfulness as an attitude centrally involves creativity, we should expect playful labelling to take many forms. As we've seen, two key features are non-deceptiveness and multiplicity: being open about the absence of total, authoritative application, and being open to many simultaneous, overlapping labels. However, there is no sure mark of its presence, or absence. Here, we illustrate with three species of playful labelling that we expect to be familiar to most readers.

Many cases of playful labelling involve some measure of non-literality, using linguistic modulation to achieve insightful applicability while flaunting a residual space between language and reality. For instance, a queer woman might call herself "heterito" for insisting on carrying heavy bags and acting in other ways stereotypical of cis men in straight relationships. Similarly, trans and queer people sometimes self-ascribe non-gender categories as their gender: thus, in a 2022 Netflix special, non-binary comedian Mae Martin described their gender as Lumiere, the candlestick in *Beauty and the Beast*.

Literal uses can also be playful, as when the speaker deploys a label to acknowledge a discomfiting degree of fit. Thus, a newly married heterosexual cis woman might call herself "wifey" to mark how far she has shifted toward a stereotype that remains uncomfortably alien. Similarly, when the narrator of Junot Díaz's *The Brief Wondrous Life of Oscar Wao* says "I tried to talk to her, of course—Mr. Community Activist—but she would skitter away from me and my stupid questions" (Díaz 2009: 253), he sincerely avows his do-gooder identity while also acknowledging its limitations and annoyances.

In all of these cases, the non-literality involves *irony*. Irony engages a kind of interpretive doubleness: simultaneously acknowledging and distancing oneself from the applicability of the label's associated expectations and norms. The doubleness is playful rather than strategic, insofar as the acknowledgment and distance are integrated into a single (complex) perspective on a single (complex) target, rather than partitioned into distinct, bistable perspectives.[25] Reclamations are also often at least partly ironic. Thus, when trans YouTuber Contrapoints self-labels as a "biological woman," she simultaneously appropriates and subverts a transphobic frame for gender. Similarly, a wheelchair user who wryly exclaims "It's a miracle! This cripple has risen!" refuses bystanders' pity and highlights their own agency while also acknowledging the challenge, and their genuine accomplishment.[26]

Metaphor engages a different sort of interpretive doubleness: the "twofoldness" of seeing one thing *as* or through the frame of another, in felt awareness of the many remaining features on both sides that fail to fit (Camp 2009).

[25]This integrative model of irony echoes Cleanth Brooks (1947). Sarcasm at least often differs from irony in treating the expectations and norms as uncontroversially presupposed, and the target as having clearly violated them (Camp 2011).

[26]Thanks to Callan Howland for this attested example.

The queer woman's use of "heterito" is plausibly both ironic and metaphorical, but playful metaphors can also be sincere. Thus, in "Metaphors," Sylvia Plath explores her burgeoning identity as a pregnant woman through a rapid-fire kaleidoscope of labels:

> I'm a riddle in nine syllables,
> An elephant, a ponderous house,
> A melon strolling on two tendrils.
> O red fruit, ivory, fine timbers!
> This loaf's big with its yeasty rising.
> Money's new-minted in this fat purse.
> I'm a means, a stage, a cow in calf.
> I've eaten a bag of green apples,
> Boarded the train there's no getting off.
>
> Sylvia Plath (1960)

Here, no single metaphor is taken to be determinative. All are tentative, some at least wry if not ironic, and some at least tentatively sincere. Together, the rapid juxtaposition of labels both amplifies and qualifies the significance of each individual one.

Plath's kaleidoscope of metaphors points toward a third form of playful labelling: a multiplicity of fully literal, sincere, highly fine-grained labels. We noted in Section IV.1 that institutional entrenchment and cultural ubiquity tend to foster rigidity for coarse-grained demographic categories. We also noted in Section IV.2 that agents often shift in a relatively passive way among descriptively accurate labels with orthogonal taxonomies, variously centering professional role, ethnicity, gender, sexuality, familial status, religion, and so on. By contrast, fine-grained categories are more likely to proliferate within subcommunities. One might think here of the nuanced, often highly colorful terms for sexual preferences on some dating sites (e.g. "genderqueer glitterslut") or the bewildering blizzard of insider labels for musical tastes (e.g. "thrash metal," "death metal," and "melo-death") or political orientations (e.g. "anarchosyndicalist," "anarcho-capitalist," "social democrat," and "democratic socialist"). While some members of these subcommunities may treat these labels rigidly, the fact that they are so restricted in scope and dimension encourages partiality: after all, that can't be *all* you really are. The proliferation of labels also fosters a creative attitude toward introducing novel ones to the subterrain, which can support collective playfulness.

V. Building agency through playful labelling

As frames, identity labels function to schematize and stabilize otherwise malleable, amorphous, and tacit perspectives on persons. Stabilization and schematization support agency by offering handy heuristics for navigating the

multidimensional realities of selfhood and social interaction. But especially when embedded into institutional structures, identity labels risk ossifying into rigid templates that flatten individuality, coerce conformity, and amplify divisive boundaries. Playful labelling does the opposite. It expands our range of available actions, in a way that is authentic rather than alienated and integrative rather than modular. The existence of playful self-labelling as a practice demonstrates on a theoretical level that identity labels need not be inherently oppressive. More importantly, on a practical level it can function as an antidote to the pernicious effects of rigid ossification. In this section, we sketch some of the practical payoffs of playful self-labelling, before turning to some of its perils in Section VI.

V.1 Payoffs of playful self-labelling for selfhood

First and most directly, shifting playfully among a variety of labels enriches one's repertoire of agential modes. Deploying a label as a frame involves allowing it to actually guide one's intuitive patterns of interpretation and action. Much as playing a variety of games can be "yoga for your agency" (Nguyen 2019: 458), so does practicing *being* a variety of characters expand one's interpretive and agential skill set, in ways that can only be achieved through intuitive, sustained implementation. However, there is a crucial difference. The game player augments their "Swiss Army knife" (Nguyen 2019: 457) of agential modes by setting up a temporary, nested agent who is immersed within a highly regimented, sealed-off alternative reality; the analogous phenomenon for self-labelling is strategic code-switching as a series of stereotyped, scripted characters. By contrast, playful self-labelling involves open-minded, open-ended, uncertain modulation into an unfamiliar niche. This requires the agent to explore how they, with their own distinctive rich, nuanced, multidimensional, holistically integrated profile of traits, interests and skills, can fruitfully mesh with the label's associated perspective and agential mode, using its stereotypes and scripts more as inspiration than imperative.

To think more concretely about how playful self-labelling scaffolds agency, consider a male first-generation college junior, Jake, who is a philosophy major and a wrestler, and who playfully self-applies the labels "philosophy geek" and "matman." First, playful labelling makes distinct but overlapping profiles of traits available to cultivate, within the scope of each identity and more broadly. For instance, the two labels encourage Jake to re-parse the same lower-level features (e.g. persistent, enjoys tussling, likes structure, and values community) under distinct higher-level traits (e.g. mentally or physically disciplined, active seminar participant, or team member). They highlight distinct profiles of existing but neglected traits (e.g. prone to pedantry, sublimates fear into focus, and super solid core). And they suggest distinct profiles of

skills and identities that he might acquire (e.g. programming, mixed martial arts, teacher, and coach). Each label, and a playful interplay among them, enriches the range of traits actively accessible to Jake, now and as he imagines his future.[27]

Secondly, playful self-labelling fosters a more robustly sustainable self, in the sense emphasized by Goffman. Performing a labelled identity in a playful way requires moving beyond regimented characters and into creative enactment, across a diverse range of circumstances. Thus, Jake might learn that it's at least sometimes OK to be passionate about Kant even if he doesn't know German and also likes David Lewis, and even as he also enjoys (or at least values) doing squats and pinning his opponent. Cultivating open-minded, open-ended, and flexible ways of instantiating each identity discloses new opportunities for implementing those identities and for bridging among them.

Finally, modulating one's self flexibly within and across environments and social identities fosters higher-order flexibility about who one might become. As with the "new *mestiza*," it might become part of Jake's identity that he stretches and morphs to fit new situations, integrating new features (a minor in anthropology, cooking in a restaurant) into a holistic overall coherence. First-order flexibility in traits, bolstered by higher-order identification *as* flexible, mitigates rigid attachment to any particular trait or label. More fundamentally, cultivating comfort with uncertainty and merely partial competence, and success in navigating frightening new contexts in unexpected ways, can instill confidence that one does indeed possess a kind of broad-based, higher-order competence, along with the resilience to build zones of competence when confronted with circumstances that make one's lack of complete competence in a certain domain painfully obvious (Dweck 2006).

V.2 Payoffs of playful self-labelling for collective agency

In Section V.1, we focused on the payoffs of playful self-labelling for building selves. What about its role in collective agency? Most directly, playful self-labelling enables an agent to smoothly inhabit a wider range of social spaces. As Lugones emphasizes, this doesn't just enhance that agent's own personal opportunities; it can also help them to recognize unexpected commonalities with others. Agents who feel at home in multiple communities and who can perform convincingly within those communities are well-placed to build bridges across communities. This is especially important for building solidarity

[27] Along similar but more profound lines, Jones (2022) recounts how shifting among the labels "woman," "mother," "disabled person," and "philosopher" forced her to confront entrenched assumptions about each identity, helping her to subvert those assumptions and to embrace aspects of each identity in a complex whole.

among activist groups organized around distinct dimensions of oppression and pursuing seemingly disparate agendas (Ortega 2016; Flores 2022).

One possibility here is that the various groups continue to be organized around disparate labels, leveraging the focus and passion that have been so useful for identity-based movements as different as labor unions, gay rights (Gamson 1995), and gun rights (Lacombe 2019). While they might employ rigid labelling for strategic purposes, those groups would need to invite and actively embrace non-stereotype-conforming individuals, and at least be open to individual members conjoining their own label with others.

Alternatively, the collective itself might work to employ the label in manifestly playful ways. This may help to welcome those who are identity-curious but not prepared to fully identify. At a deeper level, it can reclaim the ability to frame that identity from the dominant culture and open up possibilities for creative innovation. A more radical possibility is the construction of new labels that more directly acknowledge the complexity of personhood. One might think here of the Gay Rights movement's transformation into queer activism or of fourth-wave feminism, both of which emphasize the oppressive effects of rigid identity labels and embrace a model of selves as radically fluid (Gamson 1995).

Playful self-labelling is especially important for doing justice to intersectionality (Crenshaw 1989), first by subverting the oppressive effects of rigid, stereotyped frames in general, but more importantly by highlighting the ways in which many social identities resist neat analysis as a sum of fixed, discrete types.[28] Thus, Black feminists have noted repeatedly that the dominant frame for womanhood is a stereotype of *white* womanhood, with the result that rigidly self-labelling as a woman risks flattening and distorting the experience of non-white women (as well as of many white women, albeit in a less pernicious way). Individuals who proudly perform the complexity of their social identity can provide a model for others to do the same. And collectives who proudly embrace the deep diversity among their members make space for people of every identity to show up in the world in fuller and more authentic ways.

The most direct payoff from cross-identity coalition building and inclusive collective labelling is an enlarged base of participants, equipped with a wider range of skills, engaged in a shared project that is more resilient against setbacks. Another benefit may be decreased pressure on leaders to perform the movement-defining identity in a paradigmatic manner, with the commensurate risk of disillusionment if they fail to conform to character. Finally, cross-identity coalition-building and inclusion can decrease vulnerability to elite capture, in which elites forestall broader societal disruption by pitting identity-based groups against one another, or by placating specific groups with narrow or superficial changes (Táíwò 2022).

[28]Thanks to Lavi Echeverria for discussion.

VI. Perils of playful labelling

VI.1 Perils of playful labelling for selfhood

We have highlighted the practice of playful self-labelling as a tool for scaffolding more robust, flexible selves in the face of the pernicious effects of flattening, coercion, and division that rigid other-labelling tends to engender. However, like any tool, playful labelling carries risks as well as benefits.

One risk is *habituation*. As we have emphasized, deploying a frame involves actually implementing its corresponding patterns of interpretation and action. This establishes self-reinforcing internal and external looping effects with a gravitational field that can be difficult to escape. This may alter us in subtle ways that we fail to recognize, and/or that we would previously have rejected. It may eradicate significant traits that fall outside of the label even on a playful construal. Worse, it may seduce us into rigidly embracing an identity that is only partly fitting. The meshing of self and frame can be deeply phenomenologically compelling, even beyond revealing latent traits and explaining puzzling ones. This can all seem like evidence that one must have discovered one's true identity, even if most well-positioned observers would take the match to be considerably more awkward.[29]

The most obvious way to sustain playfulness and avoid rigidity with any one label is to try on multiple labels. But this raises the contrary risk, of *superficiality*. Someone in the grip of a rigid label takes it to reflect a deep self that explains a rich range of features and provides general imperatives for action. As we noted in Section IV.1, for some people this may be an accurate perception or an apt aspiration. However, for many people it risks flattening and over-simplifying a more complex reality. At the same time, too much malleability simply abandons the project of authentic selfhood. Even if the idea of a "deep self" is a metaphysical illusion, it is highly attractive (Sripada 2010), and deeply embedded in our social and moral practices. Aiming for robust selves that exhibit a complex, higher-order, flexible species of coherence offers an attractive middle way between rigid essentialism and full dissolution (Camp 2024).

A third risk is *moral harm*. An agent may deploy a frame that augments their own agency, but in the service of a toxic identity that encodes an ethically troubling and/or epistemically inaccurate perspective. Indeed, racist and sexist far-right projects have been at the forefront of recent playful labelling.[30] Appellations like "Proud boy" and "nipster" (for "Nazi hipster") gain allure precisely from their playfulness, enabling self-ascribers to acknowledge those

[29] See Camp (2017a, 2022) for discussion of the risks of habituation, and Nguyen (2021) and Camp (in press) for discussion of the risks of seduction.

[30] Thanks to Gabe Greenberg and Seth Yalcin for discussion.

identities' dark overtones while distancing themselves from a full-throated embrace of their histories and current reality. Irony in particular raises the risk that one pretends to others, and oneself, that one is merely mocking the identity. Even when this pretense is not deceptive, it can still amplify the frame's currency and its attendant pernicious effects.

One form of moral harm is *disrespect*. This is most palpable when those with a privileged social status "go slumming" by self-labelling with minoritized or oppressed identities that do not literally apply. Playful self-labelling and world-traveling are not inherently perks of privilege: on the contrary, as Du Bois, Goffman, Butler, Morton, Anzaldúa, Lugones, and many others have emphasized, oppression imposes strong pressures toward flexible, creative performance of multiple, often boundary-crossing personae. However, those with privilege especially risk enacting moral harm through playful self-labelling. The risk is most obvious when it is performed with bad intent, as when transphobes joke about their gender being an attack helicopter. But sincere self-labelling can still be problematic when it deprives others of economic, political, or cultural resources. Indeed, even sincere, literally applicable but manifestly playful self-labelling performed in contexts where no one objects or is directly oppressed—for instance, a comfortably suburban high schooler trying on the label "Cherokee" to honor a great-great-grandmother she never met—can still disrespect the depth of a target group's marginalization, or promulgate a smug, distorted imagination of what fully and inexorably inhabiting those identities is actually like (Clavel Vázquez and Clavel-Vázquez 2023).

Moving out from an agent's own self-interpretation and self-presentation, we encounter risks in how others interpret playful self-labelling. Most obviously, there is the risk of rejection: blank incomprehension of an attempted non-literal usage, or rigid re-imposition of a more accessible, culturally entrenched identity. But there are also risks of acceptance, such as rigid, literalistic uptake in the form of testimonial report or re-application, and misconstrual by imputing traits that do plausibly fall under the label's frame but that the agent themselves rejects.

Uptake matters because others play such a profound role in constituting our identities. As we saw in Section III, our options for how to be are sculpted by the surrounding social terrain and realized in interaction with others. While rigid uptake can be constraining and coercive, many of the most creative, agency-enhancing cases of playful labelling occur in concert with others, as each party riffs on and responds to the others' performances in nuanced, context-local ways that neither could have anticipated in isolation.

Daniela Dover (2022) argues that an important form of self-construction involves ceding interpretive sovereignty to others in conversation. We hear a vibrant echo here of Lugones's playfulness: openness to surprise and lack of full competence in the absence of a fixed self-conception. Others often have more insight into who we are and who we could become than we ourselves

do. Dover focuses on intimate, sustained, earnest conversations between close friends. We would add that playful experimentation early in a burgeoning relationship, or in a small group, can be at least as revealing and transformative. But these less well-established contexts further heighten the specters of misconstrual and inaccurate extrapolation.

Finally, the potential rewards of playfully co-creating an identity with others foreground the risks of playful other-labelling. Whether we are permitted to label others in ways that diverge from their preferred labels depends on how much interpretive sovereignty they have ceded to us. There might be general ethical duties to grant uptake to others' self-positionings in social space (Kukla and Lance 2022). Indeed, one can wrong others by labelling them in private thought in ways that do not match how they identify (Basu 2023). The force and complexity of these risks is a further reason we have focused on playful self-labelling here.

VI.2 Perils of playful labelling for collective action

In Sections II and III, we argued that identity labels schematize the typically messy flux of intuitive cognition and action into more stable, publicly shareable forms. In Sections IV and V, we argued that schematization and stabilization raise the risk of rigid ossification, and that playful labelling mitigates this risk by harnessing frames' intuitive power with an open-minded, open-ended, and humble attitude. At the same time, the nuance, partiality, and malleability of playful labelling risk compromising the very features that make frames so communicatively effective, in ways that are especially problematic for political action.

First, openly manifesting interpretive humility can lead others to question one's authority more generally. When a leader evinces playfulness in the service of a collective project, this also risks undermining the relevance of the project itself, or the larger collective's competence for pursuing it. This risk is especially pressing for marginalized groups, who are already pervasively subjected to hermeneutic and testimonial injustice. In light of these pressures, it is sometimes wiser for an individual leader to at least temporarily "fake it till they make it," by leaning in to their credibility and playing up identity-centric features. At a collective level, abjuring a simple, unified identity can lead to a loss of focus. In an environment where other groups single-mindedly pursue competing agendas couched in essentialist terms, playful labelling risks having one's own interests drowned out in a sea of seductively sculpted alternatives.

Second, when the social terrain is in fact deeply carved into disparate niches, as in contemporary America, an insistence on ignoring deep commonalities in favor of a kaleidoscope of shifting traits can constitute a form of willful blindness. It is imperative to track, articulate, and intervene on patterns of group-based oppression, and identity labels and generics are effective tools

for doing this (Saul 2017; Ritchie 2019). It is almost impossible to articulate such facts as that Black people in the United States own only 1.5% of the national wealth, not appreciably more than the 0.5% net wealth they owned two years before emancipation, without using identity terms; and these descriptions may need to be rigid in order to reflect and highlight the underlying reality. Abjuring social labels in favor of purportedly individualist "colorblind" or "genderblind" frames does more to relieve the burden of acknowledging social privilege from those who have it than to liberate those who are oppressed by it (Mills 2007).[31]

Third, ceding interpretive authority over a fixed identity in favor of a more open-minded, open-ended, and uncertain construal constitutes a loss of stability and control. One consequent risk is that the label and its frame may evolve in ways that earlier agents would not just have failed to predict but actively rejected. Another risk is a failure of effective coordination, both within the group and with other groups. As we saw in Section III.2, identity labels make social life sustainable by offering scripts for coordinating expectations in the service of smooth interaction. Pervasive playful labelling would almost certainly make social reality more turbulent, both at the level of local micro-interactions and in political life writ large.

VII. Recommendations

We have presented the practice of playful self-labelling as proof of principle that identity labels are not inherently oppressive, and argued that it constitutes a potent tool for mitigating the pernicious effects of rigid other-labelling. Even so, given the risks of rigid labelling sketched in Section IV.1 and the perils of playful self-labelling in Section VI, one might still wonder whether it would be better to eschew identity labels altogether, or at least to work toward a world without them (Dembroff and Wodak 2021).

In an important sense, the suggestion is moot. We have deep reasons to think that even if we can eliminate some particularly toxic labels from circulation, identity labelling in general is unavoidable. Humans are both fundamentally social and cognitively limited. As such, we need tools for managing the multidimensional complexity of our social lives in real time, by tracking and responding to individuals as members of groups. Beyond this, we tend to find labels in general deeply satisfying (Giffin, Wilkenfeld, and Lombrozo 2017).[32]

[31] Ritchie (in press) argues further that introducing labels for unlabelled marginalized groups remedies hermeneutical injustice, and, by increasing the salience of those identities, allows for greater in-group solidarity and power. Moreover, introducing labels for heretofore unlabelled dominant groups can make those identities no longer seem like unmarked defaults relative to which others count as deviations.

[32] Thanks to Josh Armstrong for these points.

More fundamentally, the idea of an autonomous self, prior to and independent of social construction, is itself a pernicious fiction. Social identities are not "artificial, oppressive constraints on the self's natural indeterminacy" or a pre-existing essence (Ortega 2016: 147; cf. Alcoff 2005; Butler 1988, 2023). Rather, selves are essentially constituted through ongoing social interaction.[33] Thus, far from distorting a "pure" antecedent self, identity labelling is a crucial tool for sculpting oneself into a self at all.

Thus, we want to conclude by offering some provisional suggestions for individuals and collectives going forward. Given the nuances and risks of playful self-labelling, and the potential merits of rigid self-labelling in some circumstances, it would be worse than foolish to propose "Self-label playfully!" as a guiding maxim. Playful labelling in hostile environments can be fatal. (This is why, as Lugones notes, she genuinely is not playful within certain worlds, even as she is able to travel to them.) Rather, we suggest that agents should seek out contexts for self-labelling that mitigate these risks: contexts where rules are not sacred, and surprise and foolishness are less likely to be destructive.

The most obvious such contexts are likely to be intimate, or at least in-group, akin to Dover's (2022) contexts of friendship. At the same time, we often find ourselves anew when we mingle with those we take to be other. More importantly, no context is entirely risk-free. Indeed, long-term intimate relationships can be among the most treacherous for trying out new identities. Conversely, when deployed with fierce wit and ingenuity, playful self-labelling can be a powerful political tool for disarming hegemonic rigid labelling even in overtly hostile environments. Thus, there are no hard and fast rules for when to engage in or avoid playful self-labelling.

What advice might we offer to institutions and other collectives interested in scaffolding agency? Certainly not: "Conduct annual mandatory playful-labelling exercises!" Imposed playfulness is oppressive if not outright incoherent. Moreover, as we have noted, rigid labelling does have legitimate uses. In particular, bureaucracies need stable classifications in order to effectively track and control people and resources across times and places (Scott 1999). The individuals and societies that those bureaucracies ultimately serve should be prepared to provide that information when it is relevant. At the same time, they should also work to limit their scope and power as much as possible.

With respect to labelling in particular, we should claw back as much individual and local nuance as possible from rigid institutional abstractions. Practically, this might involve demanding that bureaucracies reduce requirements to provide demographic information when it is not practically relevant, and enrich the range of labels available for specifying that information.[34] More

[33] See Bell (2024) for a recent treatment of self-construction that addresses social embeddedness.

[34] Thanks to Willow Starr for discussion.

specifically, institutions are increasingly developing systems that permit individuals to assign multiple orthogonal identities simultaneously, to opt out of a given taxonomy, or to supply their own label within a taxonomy. For example, the dating app OKCupid includes twenty-two (non-exclusive) gender options and twenty sexual orientation options, in addition to nineteen optional tags related to gender and sexuality.[35] It is easy for this to seem silly or overwhelming. But, if we are right, it is the sort of design choice that increases space for users not just to communicate with others but also to explore and refine their own agency.

VIII. Conclusion

We have argued that in addition to being devices for categorization, identity labels function as frames. This makes them powerful tools for regulating intuitive cognition and social performance. On the one hand, they can oppress, by flattening individual distinctiveness, coercing conformity, and heightening divisive binaries. These perils are most pressing with rigid other-labelling. But identity labels are not in and of themselves obstacles to agency. Applying them flexibly, especially through playful self-labelling, can scaffold agency in ways that support building multitudinous selves, reconfiguring our social world, and engaging in effective collective action against oppression. In the case of identity labels, the masters' tools may indeed be used to dismantle the masters' house—if we expand our repertoire of uses beyond the rigid applications operative in bureaucratic classification and ideological control.

Acknowledgements

This work was produced on a collaborative basis and author names are listed in alphabetical order. We wish to thank audiences at the Social Identities and Cognition Desert Workshop, the UC Irvine Social Unity Research Cluster, the 2023 Southern Society for Philosophy and Psychology, Lafayette College, SLIME 2 at UCLA, the Lisboa Feminist Philosophy of Mind Workshop, the Words Workshop, the 2023 meeting of the Society for Philosophy and Psychology, the University of Lisbon, and the UNED in Madrid, for valuable comments and discussion. We would especially like to thank Kate Ritchie for extensive discussion and Josh Armstrong for comments at SLIME.

Conflict of interest statement

None declared.

[35] https://instinctmagazine.com/you-can-identify-as-a-top-or-bottom-on-okcupid-now/

References

Alcoff, L. M. (2005) *Visible Identities: Race, Gender, and the Self*. Oxford: OUP.
Allport, G. (1954) *The Nature of Prejudice*. Reading, MA: Addison-Wesley.
Althusser, L. (1971) 'Ideology and Ideological State Apparatuses (Notes Towards an Investigation)' (Ben Brewster, Trans.), in L. Althusser (ed.), *Lenin and Philosophy and Other Essays*, pp. 127–86. New York: Monthly Review Press.
Andler, M. (2022) 'Queer and Straight', in Clare Chambers, Brian D. Earp, and Lori Watson (eds), *Routledge Handbook of Philosophy of Sex and Sexuality*. Oxford: Routledge.
Andler, M. (2023) 'What Is Masculinity?', *Synthese*, 202: 1–16.
Anzaldúa, G. (1987) *Borderlands/La Frontera: La Nueva Mestiza*. San Francisco, CA: Aunt Lute.
Anzaldúa, G. (2009) *The Gloria Anzaldúa Reader* (A. L. Keating (ed.)). Durham, NC: Duke University Press.
Armstrong, J. (2016) 'The Problem of Lexical Innovation', *Linguistics and Philosophy*, 39: 87–118.
Armstrong, S., Gleitman, L., and Gleitman, H. (1983) 'What Some Concepts Might Not Be', *Cognition*, 13/3: 263–308.
Ásta (2018) *Categories We Live by: The Construction of Sex, Gender, Race, and Other Social Categories*. New York: Oxford University Press.
Barnes, E. (2019) 'Gender and Gender Terms', *Noûs*, 54: 704–30.
Baron, A. and Dunham, Y. (2015) 'Representing "Us" and "Them": Building Blocks of Intergroup Cognition', *Journal of Cognition and Development*, 16: 780–801.
Basu, R. (2023) 'Morality of Belief II: Three Challenges and an Extension', *Philosophy Compass*, 18: 1–9.
Beeghly, E. (2015) 'What Is a Stereotype? What Is Stereotyping?', *Hypatia*, 30: 675–91.
Bell, R. (2024) 'Being Your Best Self: Authenticity, Morality, and Gender Norms', *Hypatia*, 39: 1–20.
Bernstein, S. (2024) 'Resisting Social Categories', *Oxford Studies in Agency and Responsibility*, 8: 81–102.
Bigler, R., Jones, L., and Lobliner, D. (1997) 'Social Categorization and the Formation of Intergroup Attitudes in Children', *Child Development*, 68: 530–43.
Boyd, R. (1999) 'Homeostasis, Species, and Higher Taxa', in Robert Andrew Wilson (ed), *Species: New Interdisciplinary Essays*, pp. 141–85, Cambridge, MA: MIT Press.
Brennan, S. E. and Clark, H. H. (1996) 'Conceptual Pacts and Lexical Choice in Conversation', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22: 1482–93.
Brooks, C. (1947) 'The Heresy of Paraphrase', in *The Well Wrought Urn: Studies in the Structure of Poetry*, pp. 192–214, New York: Harcourt Brace.
Brown, W. (1995) *States of Injury: Power and Freedom in Late Modernity*. Princeton, NJ: PUP.
Butler, J. (1988) 'Performative Acts and Gender Constitution: An Essay in Phenomenology and Feminist Theory', *Theatre Journal*, 40: 519–31.
Butler, J. (2023) 'Categories by Which We Try to Live', *European Journal of Philosophy*, 31: 283–8.
Camp, E. (2006) 'Metaphor and that Certain "Je Ne Sais Quoi"', *Philosophical Studies*, 129: 1–25.
Camp, E. (2007) 'Showing, Telling, and Seeing: Metaphor and "Poetic" Language', *The Baltic International Yearbook of Cognition, Logic, and Communication*, 3: 1–24.
Camp, E. (2009) 'Two Varieties of Literary Imagination: Metaphor, Fiction, and Thought Experiments', *Midwest Studies in Philosophy*, 33: 107–30.
Camp, E. (2011) 'Sarcasm, Pretense, and the Semantics/Pragmatics Distinction', *Noûs*, 46: 587–634.
Camp, E. (2013) 'Slurring Perspectives', *Analytic Philosophy*, 54: 330–49.
Camp, E. (2015) 'Logical Concepts and Associative Characterizations', in E. Margolis and S. Laurence (eds) *The Conceptual Mind: New Directions in the Study of Concepts*, pp. 591–621. Cambridge, MA: MIT Press.
Camp, E. (2017a) 'Perspectives in Imaginative Engagement with Fiction', *Philosophical Perspectives*, 31: 73–102.
Camp, E. (2017b) 'Why Metaphors Make Good Insults: Perspectives, Presupposition, and Pragmatics', *Philosophical Studies*, 174: 47–64.

Camp, E. (2019) 'Perspectives and Frames in Pursuit of Ultimate Understanding', in S. Grimm (ed.) *Varieties of Understanding: New Perspectives from Philosophy, Psychology, and Theology*, pp. 17–45. Oxford: OUP.
Camp, E. (2020) 'Imaginative Frames for Scientific Inquiry: Metaphors, Telling Facts, and Just-So Stories', in P. Godfrey-Smith and A. Levy (eds) *The Scientific Imagination*, pp. 304–36. Oxford: OUP.
Camp, E. (2022) 'Agency, Stability, and Permeability in "Games"', *Journal of Ethics and Social Philosophy*, 23: 448–76.
Camp, E. (2024) 'Stories and Selves: A Twisted Love Story about the Meaning of Life', *Royal Institute of Philosophy Supplement*, 95: 157–79.
Camp, E. and Flores, C. (in press) '"That's All You Really Are": Centering Identities without Essentialist Beliefs', in S. Haslanger et al. (eds) *Mind, Language, & Social Hierarchy* New York: Oxford University Press.
Camp, E. (in press) 'Perspectival Complacency, Open-Mindedness and Amelioration', in W. Riggs and N. Snow (eds) *Open-Mindedness and Perspective*. Oxford: OUP.
Carnap, R. (1950) 'Empiricism, Semantics and Ontology', *Revue Internationale de Philosophie*, 4/11: 20–40.
Clavel Vázquez, M. J. and Clavel-Vázquez, A. (2023) 'Robustly Embodied Imagination and the Limits of Perspective-Taking'. *Philosophical Studies*, 180: 1395–420.
Crenshaw, K. (1989) 'Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics', *University of Chicago Legal Forum*, 1: 139–67.
Del Pinal, G. and Reuter, K. (2017) 'Dual Character Concepts in Social Cognition: Commitments and the Normative Dimension of Conceptual Representation', *Cognitive Science*, 41: 477–501.
Del Pinal, G. and Spaulding, S. (2018) 'Conceptual Centrality and Implicit Bias', *Mind & Language*, 33: 95–111.
Del Pinal, G., Madva, A., and Reuter, K. (2017) 'Stereotypes, Conceptual Centrality and Gender Bias: An Empirical Investigation', *Ratio*, 30: 384–410.
Dembroff, R. and Saint-Croix, C. (2019) '"Yep, I'm Gay": Understanding Agential Identity', *Ergo: An Open Access Journal of Philosophy*, 6: 571–99.
Dembroff, R. and Wodak, D. (2018) 'He/She/They/Ze', *Ergo*, 5: 371–406.
Dembroff, R. and Wodak, D. (2021) 'How Much Gender Is Too Much Gender?', in R. Mason (ed.) *Hermeneutical Injustice*, pp. 362–76. Oxford: Routledge.
Díaz, J. (2008) *The Brief Wondrous Life of Oscar Wao*. New York: Penguin.
Duncan, B. L. (1976) 'Differential Social Perception and Attribution of Intergroup Violence: Testing the Lower Limits of Stereotyping of Blacks', *Journal of Personality and Social Psychology*, 34/4: 590–8. https://doi.org/10.1037/0022-3514.34.4.590
Dover, D. (2022) 'The Conversational Self', *Mind*, 131: 193–230.
Dror, I. (2023) 'Is There an Epistemic Advantage to Being Oppressed?' *Noûs*, 57: 618–40.
Du Bois, W. E. B. (1903) *The Souls of Black Folk*. Chicago, IL: A.C. McClurg & Co.
Dweck, C. S. (2006) *Mindset: The New Psychology of Success*. New York: Random House.
Flores, C. (2022) 'Keeping Better Company', *The Philosophers' Magazine*, https://www.philosophersmag.com/essays/283-keeping-better-company, accessed 15 May 2024.
Fodor, J. and Lepore, E. (1996) 'The Red Herring and the Pet Fish: Why Concepts Still Can't Be Prototypes', *Cognition*, 58: 253–70.
Foucault, M. (1977) *Discipline and Punish: The Birth of the Prison*. New York: Pantheon Books.
Galinsky, A. D. et al. (2013) 'The Reappropriation of Stigmatizing Labels: The Reciprocal Relationship between Power and Self-Labelling', *Psychological Science*, 24: 2020–9.
Gamson, J. (1995) 'Must Identity Movements Self-Destruct? A Queer Dilemma', *Social Problems*, 42: 390–407.
Gelman, S. (2003) *The Essential Child: Origins of Essentialism in Everyday Thought*. Oxford: OUP.
Gelman, S. and Heyman, G. (1999) 'Carrot-Eaters and Creature-Believers: The Effects of Lexicalization on Children's Inferences about Social Categories', *Psychological Science*, 10: 489–93.
Giffin, C., Wilkenfeld, D., and Lombrozo, T. (2017) 'The Explanatory Effect of a Label: Explanations with Named Categories Are More Satisfying', *Cognition*, 168: 357–69.
Goffman, E. (1956) *The Presentation of Self in Everyday Life*. Garden City, NY: Doubleday.

Hacking, I. (1995) 'The Looping Effect of Human Kinds', in D. Sperber, D. Premack, and A. Premack (eds) *Causal Cognition: A Multidisciplinary Debate*, pp. 351–94. Oxford: Clarendon Press.
Hacking, I. (2007) 'Kinds of People: Moving Targets', *Proceedings of the British Academy*, 151: 285–318.
Haslanger, S. (2011) 'Ideology, Generics, and Common Ground', in C. Witt (ed.) *Feminist Metaphysics*, pp. 179–207. New York: Springer.
Haslanger, S. (2019) 'Cognition as a Social Skill', *Australasian Philosophical Review*, 3: 5–25.
Haslanger, S. (2021) *Ideology in Practice: What Does Ideology Do?* Milwaukee, WI: Marquette University Press.
Hernandez, E. M. (2021) 'Gender-Affirmation and Loving Attention', *Hypatia*, 36: 619–35.
Hernandez, E. M. and Crowley, A. (2024) 'How To Do Things with Gendered Words', in L. Anderson and E. Lepore (eds) *The Oxford Handbook of Applied Philosophy of Language*. New York: Oxford University Press.
Irigaray, L. (1985) *The Sex Which Is Not One*. Ithaca, NY: Cornell University Press.
Jenkins, K. (2016) 'Amelioration and Inclusion: Gender Identity and the Concept of Woman', *Ethics*, 126: 394–421.
Jenkins, K. (2019) 'Toward an Account of Gender Identity', *Ergo*, 5: 713–44.
Jenkins, K. (2023) *Ontology and Oppression: Race, Gender, and Social Reality*. Oxford: OUP.
Jeshion, J. (2013) 'Slurs and Stereotypes', *Analytic Philosophy*, 54: 314–29.
Jones, C. C. (2022) *Easy Beauty: A Memoir*. New York: Simon & Schuster.
Judd, C. M. and Park, B. (1993) 'Definition and Assessment of Accuracy in Social Stereotypes', *Psychological Review*, 100/1: 109–28. https://doi.org/10.1037/0033-295X.100.1.109
Knobe, J., Prasada, S., and Newman, G. E. (2013) 'Dual Character Concepts and the Normative Dimension of Conceptual Representation' *Cognition*, 127: 242–57.
Kukla, Q. (2018) 'Slurs, Interpellation, and Ideology', *The Southern Journal of Philosophy*, 56: 7–32.
Kukla, Q. and Lance, M. (2022) 'Telling Gender: The Pragmatics and Ethics of Gender Ascriptions', *Ergo*, 9: 1–30.
Lacombe, M. J. (2019) 'The Political Weaponization of Gun Owners: The National Rifle Association's Cultivation, Dissemination, and Use of a Group Social Identity', in *The Journal of Politics*, 81: 1342–56.
Leslie, S. J. (2014) 'Carving up the Social World with Generics', *Oxford Studies in Experimental Philosophy*, 1: 208–31.
Leslie, S. J. (2015) '"Hillary Clinton Is the Only Man in the Obama Administration": Dual Character Concepts, Generics, and Gender', *Analytic Philosophy*, 56: 111–41.
Leslie, S. J. (2017) 'The Original Sin of Cognition: Fear, Prejudice, and Generalization' *Journal of Philosophy*, 114: 393–421.
Liberman, Z., Woodward, A. L., and Kinzler, K. D. (2017) 'The Origins of Social Categorization', *Trends in Cognitive Sciences*, 21: 556–69.
Lugones, M. (1987) 'Playfulness, "World"-Travelling, and Loving Perception', *Hypatia*, 2: 3–19.
Mallon, R. (2016) *The Construction of Human Kinds*. Oxford: OUP.
Margolis, E. and Laurence, S. (1999) *Concepts: Core Readings*. Cambridge. MA: MIT Press.
Martens, A. et al. (2006) 'Combating Stereotype Threat: The Effect of Self-Affirmation on Women's Intellectual Performance', *Journal of Experimental Social Psychology*, 42: 236–43.
McConnell-Ginet, S. (2002) '"Queering" Semantics: Definitional Struggles', in K. Campbell-Kibler et al. (eds) *Language and Sexuality: Contesting Meaning in Theory and Practice*, pp. 137–60. Stanford, CA: CSLI Publications.
Mills, C. (2007) 'White Ignorance', in S. Sullivan and N. Tuana (eds), *Race and Epistemologies of Ignorance*, pp. 11–38. Albany, NY: SUNY Press.
Morton, J. (2014) 'Cultural Code-Switching: Straddling the Achievement Gap', *The Journal of Political Philosophy*, 22: 259–81.
Munton, J. (2023) 'Prejudice as the Misattribution of Salience', *Analytic Philosophy*, 64: 1–19.
Murphy, G. L. (2002) *The Big Book of Concepts*. Cambridge, MA: MIT Press.
Neufeld, E. (2022) 'Psychological Essentialism and the Structure of Concepts', *Philosophy Compass*, 17: e12823.
Neufeld, E. (in press) 'Engineering Social Concepts: Labels and the Science of Categorization', in S. Haslanger et al. (eds) *Mind, Language, and Social Hierarchy* New York: Oxford University Press.

Neufeld, E. (2024) 'Engineering Social Concepts: Feasibility and Causal Models', *Philosophy and Phenomenological Research*.
Nguyen, C. T. (2019) 'Games and the Art of Agency', *Philosophical Review*, 128: 423–62.
Nguyen, C. T. (2020) *Games: Agency as Art*. Oxford: OUP.
Nguyen, C. T. (2021) 'The Seductions of Clarity', *Royal Institute of Philosophy Supplement*, 89: 227–55.
Nguyen, H. H. and Ryan, A. M. (2008) 'Does Stereotype Threat Affect Test Performance of Minorities and Women? A Meta-Analysis of Experimental Evidence', *The Journal of Applied Psychology*, 93: 1314–34.
Nussbaum, M. (1992) *Love's Knowledge: Essays on Philosophy and Literature*. Oxford: OUP.
Ortega, M. (2016) *In-Between: Latina Feminist Phenomenology, Multiplicity, and the Self*. Albany, NY: SUNY Press.
Pennington, C. R. et al. (2016) 'Twenty Years of Stereotype Threat Research: a Review of Psychological Mediators', *PLoS ONE*, 11: e0146487.
Pessoa, F. (1998) *Fernando Pessoa & Co. – Selected Poems* (*R. Zenith, Trans.*). New York: Grove Press.
Plath, S. (1960) 'Metaphors', in *The Colossus*. London: Heinemann.
Ritchie, K. (2019) 'Should We Use Racial and Gender Generics?', *Thought: A Journal of Philosophy*, 8: 33–41.
Ritchie, K. (2021a) 'Essentializing Inferences', *Mind & Language*, 36: 570–91.
Ritchie, K. (2021b) 'Essentializing Language and the Prospects for Ameliorative Projects', *Ethics*, 131: 460–88.
Ritchie, K. (2021c) 'Does Identity Politics Reinforce Oppression?', *Philosophers' Imprint*, 21: 1–15.
Ritchie, K. (in press) 'Labelling Unlabelled Identities', in M.G. Isaac, S. Koch, and K. Scharp (eds) *New Perspectives on Conceptual Engineering Vol. 3: Applied Conceptual Engineering*. New York: Springer.
Rosch, E. H. (1973) 'Natural Categories', *Cognitive Psychology*, 4: 328–50.
Saul, J. (2017) 'Are Generics Especially Pernicious?', *Inquiry* 66: 1689–706.
Scott, J. C. (1999) *Seeing like a State: How Certain Schemes to Improve the Human Condition Have Failed*. New Haven, CA: Yale University Press.
Spivak, G. C. (2012) 'Subaltern Studies: Deconstructing Historiography', in G. C. Spivak, *In Other Worlds*, pp. 270–304. Oxford: Routledge.
Sripada, C. S. (2010) 'The Deep Self Model and Asymmetries in Folk Judgments about Intentional Action', *Philosophical Studies*, 151: 159–76.
Steele, C. M. (1997) 'A Threat in the Air: How Stereotypes Shape Intellectual Identity and Performance', *The American Psychologist*, 52: 613–29.
Táíwò, O. O. (2022) *Elite Capture: How the Powerful Took over Identity Politics (and Everything Else)*. Chicago, IL: Haymarket Books.
Tirrell, L. (2012) 'Genocidal Language Games', in I. Maitra and M. K. McGowan (eds) *Speech and Harm: Controversies over Free Speech*, pp. 174–221. Oxford: OUP.
Whiteley, E. (2023) '"A Woman First and a Philosopher Second": Relative Attentional Surplus on the Wrong Property', *Ethics*, 133: 497–528.
Whitman, W. (1891) 'Song of Myself', in *Leaves of Grass*, Deathbed Edition. Philadelphia: David McKay Co.
Woolf, K. et al. (2009) 'The Effect of a Brief Social Intervention on the Examination Results of UK Medical Students: A Cluster Randomised Controlled Trial', *BMC Medical Education*, 9: 1–15.
Wylie, A. (2003) 'Why Standpoint Matters', in R. Figueroa and S. G. Harding (eds) *Science and Other Cultures: Issues in Philosophies of Science and Technology*, pp. 26–48. Oxford: Routledge.