INTRODUCTION

Theories of collective intentionality and theories of relational autonomy share a common interest in analyzing the social dynamics of agency. However, whereas theories of collective intentionality conceive of social groups primarily as intentional and voluntarily willed, theories of relational autonomy claim that autonomous agency is both scaffolded and constrained by social forces and structures, including the constraints imposed by nonvoluntary group membership. The question raised by this difference in view is whether social theorizing that overlooks the effects of nonvoluntary social group membership on individual and joint agency overlooks crucial aspects of the social dynamics of agency. To explore this question, this article first evaluates Michael Bratman's planning analysis of individual agency from the perspective of relational autonomy theory and compares it with a narrative self-constitution account of temporally extended agency. It then evaluates Bratman's analysis of shared agency and discusses Shaun Gallagher and Deborah Tollefsen's concept of we-narratives, which extends the notion of narrative construction to shared agency. Overall, the argument aims to show that if we are interested in understanding the social dynamics of agency, it is critical to attend to the way that agents exercise their intentional agency in relation to internalized and external social constraints.

1 INTRODUCTION

Theories of collective intentionality and theories of relational autonomy share a common interest in analyzing the social dynamics of agency. Theories of collective intentionality investigate how, by joining together to form intentions, make plans and decisions, perform actions, undertake commitments, or construct shared narratives, two or more individual agents co-constitute forms of we-agency. Theories of relational autonomy investigate the ways that
personal autonomy is socially scaffolded by interpersonal relationships and by social norms, practices, and institutions. However, despite their common interest in the social dynamics of agency, it is notable that there has been little philosophical dialogue between the respective literatures.

My aim in this article is to engage in such a dialogue and to identify some points of convergence and divergence between these literatures. With respect to convergence, it is evident that various forms of we-agency play a crucial role in scaffolding autonomous agency. In referring to we-agency broadly, I mean to include joint intending, planning, acting, deciding, committing, reasoning, and narrative construction, as well as sharing emotions. Given the centrality and ubiquity of we-agency in our lives, it is surprising that relational autonomy theorists, myself included, have paid so little attention to its role in enabling autonomous agency.

With respect to divergence, while claiming that autonomous agency is only possible by virtue of our sociality and our embeddedness in social relationships and institutions, relational autonomy theorists view the relationship between sociality and autonomy as Janus-faced, as Beate Roessler puts it (2021, p. 8). Some interpersonal relationships, as well as social norms, practices, and institutions structured by unjust social hierarchies thwart the development or impede the exercise of autonomous agency. A central motivation of relational autonomy theory is therefore to explicate the autonomy-impairing effects of social oppression. Social oppression, as defined by Ann Cudd, is “an institutionally structured, unjust harm perpetrated on groups by other groups through direct and indirect material and psychological forces” (2006, p. 25). In focusing attention on the Janus-faced character of sociality, relational theories assume a compatibilist view of the relationship between individual agency and social structures. On this view, while our agency is intentionally guided, our intentions and actions are constrained by social forces and structures, including the constraints imposed by nonvoluntary group membership.

Theories of collective intentionality have not focused similar attention on the relevance of dynamics of social power and oppression for we-agency. This difference points to what seems to be a fundamental difference between how social groups are conceptualized in the two literatures. Influential theories of collective intentionality conceive of social groups primarily as intentional and voluntarily willed in some sense by individual members of those groups. While this view does account for many social groups that play important roles in our lives and in shaping our agency, it excludes groups whose membership is nonvoluntary and based in ascriptive social markers of identity. For relational autonomy theorists, by contrast, our individual identities and agential capacities are constituted in dynamic interaction with these social identity markers, which also shape and constrain the options available to agents. The category of nonvoluntary social groups is therefore critical for explaining how personal autonomy can be undermined by socially oppressive interpersonal relationships and social structures.

The question raised by this difference in view is whether social theorizing that overlooks the effects of nonvoluntary social group membership on individual and joint agency overlooks crucial aspects of the social dynamics of agency. I will argue that it does and that one contribution relational autonomy theory might make to a dialogue between these two literatures is to show why this is the case. The literature on collective intentionality is voluminous. To focus discussion, I will restrict my attention in this article primarily to the work of Michael Bratman, which straddles issues related to both self-governing and shared agency. In Section 2, I evaluate Bratman's planning analysis of individual agency from the perspective of

---

1My discussion here is influenced by Ann Cudd's excellent analysis of the way individual agency is constrained by nonvoluntary social group membership, in Cudd (2006, chap. 2).

relational autonomy theory and compare it with a narrative self-constitution account of temporally extended agency. In Section 3, I evaluate Bratman's analysis of shared agency and then discuss Shaun Gallagher and Deborah Tollefsen's concept of we-narratives, which extends the notion of narrative construction to shared agency. Overall, my argument aims to show that if we are interested in understanding the social dynamics of agency, it is critical to attend to the way that agents exercise their intentional agency in relation to internalized and external social constraints.

2 | AUTONOMOUS AGENCY AND SOCIAL OPPRESSION

Autonomy is a complex concept, which is conceptually connected to a range of other concepts and functions differently in different normative domains. To do justice to this complexity, and to understand the different ways that individual autonomy can be threatened by social oppression, I have proposed elsewhere that autonomy should be understood as a multidimensional concept (Mackenzie, 2014a). According to my analysis, the concept of autonomy comprises three distinct but interacting dimensions: self-determination, self-governance, and self-authorization. To be self-determining is to have the freedom and opportunities required to determine the direction of one's life overall and to exercise practical control over important domains of one's life. The self-determination dimension highlights externalist social and political conditions for autonomy and points to important conceptual linkages between autonomy, freedom, and opportunity. To be self-governing is to exercise the capacities for deliberation and reflection necessary to shape and direct one's life in accordance with one's values, goals, and commitments. The self-governance dimension points to conceptual linkages between autonomy, authenticity, and practical rationality. To be self-authorization is to regard oneself as having the normative authority to take responsibility for one's values and commitments, and this requires holding appropriate attitudes of self-respect, self-trust, and self-esteem. While self-governance and self-authorization highlight internalist psychological conditions for autonomy, these conditions should be understood as socio-relational; that is, constituted in relation to interpersonal relationships as well as broader social forces and structures.

Most of my discussion in this section will focus on the self-governance dimension of autonomy. But I will return at the end of the section to explain the importance of self-authorization and self-determination. In the literature, a distinction is often drawn between two broad kinds of conditions for self-governing agency: competence and authenticity (see e.g., Christman, 2009; Mackenzie, 2008a). Relational theories understand both conditions as socio-relational. Autonomy competence refers to the complex suite of capacities required for critical reflection and deliberation with respect to one's motives, values, and commitments. According to relational theories, this suite includes capacities for reasoning and communication, for imagination and memory, for emotional self-interpretation and regulation, as well as embodied skills. These capacities are socio-relational insofar as they are developed through socialization processes and exercised only with extensive interpersonal, social, and institutional scaffolding (see, e.g., Meyers, 1989). Thus, the extent to which an agent develops these capacities depends on the character of her social environment. Authenticity refers to the idea that choices and actions resulting from competent critical reflection are genuinely the agent's “own,” in the sense that they reflect or are expressive of the agent's considered preferences, values, and commitments. Authenticity is secured if the agent endorses or, more weakly, is not alienated from the motives on which she acts.3

3See especially John Christman (2009) for an influential nonalienation account of authenticity.
Whereas competence conditions explicate the capacities required for governing the self, authenticity conditions demarcate the boundaries of the self that is self-governing. According to relational theories, however, the self is not atomistic. We constitute our self-identities interpersonally and in dynamic interaction with the social contexts in which we are embedded. In my view, this process is best understood as a process of socially scaffolded narrative self-constitution, for reasons I will explain later.

There are extensive debates in the literature concerning what constitutes authentic endorsement and how endorsement secures agential authority. Hierarchical theories, such as those of Frankfurt (1971) and Watson (1975), which understand endorsement in terms of alignment between lower-order and higher-order elements of the agent's motivational structure, are especially vulnerable to the problem of explaining how endorsement secures agential authority. This problem and other difficulties facing hierarchical theories have been extensively discussed in the literature, including by relational autonomy theorists, and I will not revisit these difficulties here. A related debate concerns whether the process that yields authentic endorsement should be conceptualized as synchronic or diachronic. I do not intend to revisit this debate here. Suffice to say that I concur with Michael Bratman (2007) and John Christman (2009) in understanding self-governing agency as a diachronic, temporally extended process. My aim in this section is to discuss two different ways of conceptualizing the diachronic dimensions of self-governing agency: Bratman's planning account and a narrative constitution account. I will argue that from the perspective of relational autonomy theory there are strong reasons to prefer a narrative constitution account.

Bratman's planning account of self-governing agency is motivated by the Lockean intuition that we constitute our identity and agency over time by forging overlapping, cross-temporal first-personal psychological continuities and connections (2007, chap. 2, chap. 10). The Lockean intuition draws attention to two distinctive features of human agency: its reflectiveness and its temporally extended structure. Reflectiveness refers to our capacity to reflect upon and take a stand with respect to our motivations, to endorse or reject them. The temporally extended structure of agency refers to the fact that we conceive of our agency as extending backward into the past and forward into the future. Bratman claims that through backward-looking ties of memory and forward-looking ties of intending and planning, we forge connections between our present, past, and future. The basic claim of the planning account is that we forge these ties and organize our activities over time by forming intentions, planning the means to realize those intentions, and enacting prior intentions. Our intentions, plans, and self-governing policies organize our agency over time by guiding deliberation and practical reasoning.

To effectively guide deliberation, however, intentions and plans must meet certain norms of practical rationality, specifically stability, means-end coherence, and consistency. If I intend to walk every day as part of a plan to walk 10,000 steps a day but then every morning revisit my prior intention to go for a walk, my walking-related intentions are not stable, and this instability will undermine my plan. Stability does not entail rigidity. If there are reasons to revisit the intention, say, because I have sprained my ankle, then I should do so. Nevertheless, prior intentions and plans exert rational pressure not to reconsider them from one day to the next (2007, p. 26). Planning is a complex, nested and evolving activity, involving the development of partial plans and subplans that are filled out over time (Bratman, 1999, p. 3). If I aim to bring about an end, say, to improve my fitness, this will require me to figure out the best means of doing so, given constraints of time and resources, which in turn will require me to develop plans and subplans. The constraint of means-end coherence helps to guide deliberation by focusing my planning activities. If I intend to join
a hiking group as part of the plan to improve my fitness, then I will need to investigate the different groups in my area, work out which is most congenial, find out who to contact to initiate the process of joining, and so on. Finally, the norm of consistency requires that the ends I adopt are mutually consistent and attainable. If my end is to stay fit, then this end will rule out as an option spending six hours on the couch every day bingeing on Netflix. Consistency of ends helps to strengthen diachronic stability because it helps to ensure that the plans and subplans within my overall network of plans and policies are interlocking and mutually supporting rather than clashing.

Self-governing agency, however, involves more than effective planning agency. It also requires strong reflectiveness with respect to which ends the agent should endorse or treat as providing justifying reasons for action. In Bratman's view, higher-order policies, such as a policy to maintain friendships, to spend time with family, to stay fit, not to be dominated by work demands, not to harbor grudges, and so on, function to support self-governing agency by determining where the agent stands with respect to which motives she should endorse or reject. These policies are forms of valuing that reflect the agent's background values, cares and concerns, and enable evaluative rankings of alternative options (1999, p. 61). If I have a self-governing policy not to be dominated by work demands, then this will help guide deliberation about whether to accept an invitation to join a committee that I know will involve a lot of work, by giving weight to the motivation to refuse the invitation. If I refuse the invitation, the self-governing policy gives this decision agential authority because the policy functions as a commitment that determines where I stand. This does not mean that I give no weight to countervailing motivations, but if I am satisfied with this policy, it will function to “crystallize pressures from various elements of one's psychic stew into a more decisive attitude,” (Bratman, 2007, p. 36). Such commitments can of course be reviewed and revised over time, perhaps under pressure from other self-governing policies with which they conflict. However, the more consistency and stability there is within the agent's overall network of self-governing policies and plans, the more effective they will be in coordinating and organizing the agent's life over time (p. 42).

Whether or not Bratman's account of self-governing agency provides a satisfactory response to the problem of agential authority is an issue that I will not engage with here. What is of more interest to my concerns in this article is whether, from the perspective of a relational autonomy theorist, this account adequately conceptualizes both the self of self-governing agency and the agential competences required to govern the self. In my view, it does not, for three connected reasons. First, while Bratman's focus on the temporally extended dimensions of agency and his analysis of planning agency make important contributions to our understanding of self-governing agency, his planning agent is overly thin and exercises an implausible degree of control over her own agency. According to Bratman, the agent just is constituted by her intentions, plans, and self-governing policies. While this analysis avoids the metaphysical difficulties of positing a homuncular agent that is somehow separate from her psychological processes, it relegates into the background important elements of the “psychic stew” that partially constitute the agent's identity over time and in relation to which she exercises her agency.5 It also pays insufficient attention to the role played by the agent's self-concept in “crystallizing . . . a more decisive attitude” out of these elements. Korsgaard's notion of “practical identity,” which she characterizes as the agent's “normative self-conception” provides one way of explicating this self-concept (1996, p. 101). The notion of narrative self-constitution, discussed below, provides another way of explicating the agent's self-concept and its role in guiding agency. According to Korsgaard, the process of constituting one's practical identity is an authorial process involving reflection and deliberation.

5Although Bratman aims to provide a normative theory of agency, in my view an adequate normative theory needs to be based on descriptively adequate assumptions. The analysis of narrative self-constitution developed below is both descriptive and normative.
However, the practical identity that is constituted through this process incorporates myriad nonchosen contingencies over which as agents we have limited control. These include contingencies of social identity, personal history, bodily and intellectual capacity, and relational and social context, all of which define who we are and what matters to us. Bratman acknowledges that self-governing agency requires the self-management of first-order motives, appetites, and desires that conflict with one’s commitments (2007, pp. 217–220). However, self-governing agency involves more than conative self-management because there is more that we need to manage than our desires. We also need to manage, that is, make sense of and constitute, ourselves in relation to the contingencies that shape our practical identities.

Second, while Bratman’s account of shared agency is attuned to certain aspects of the social dimensions of planning agency, his account of self-governing agency is not sufficiently attuned to the role of agents’ social relationships and social environment in shaping their self-concepts and constraining their agency. Bratman’s planning agent is therefore socially decontextualized in important respects. I discuss this issue below and in the following section of the article. Third, the planning agent is overly rationalistic. While Bratman’s account places a great deal of weight on the importance of norms of practical reasoning, and hence on rational competence, it remains silent on the role of affective, imaginative, and embodied competences in self-governing agency. Bratman is unlikely to be fazed by this complaint since he is open that his theory is a theory of rational planning (1999, p. 6). In my view, however, this rationalistic bias results in a somewhat impoverished understanding of the agential competences required for self-governing agency and the extent to which these competences depend on the character of the agent’s social environment. Combining these two criticisms, the theory of planning agency overlooks the extent to which failures of self-governing agency may be due more to the agent’s social environment than to failures of rational planning. I now want to explain why from the perspective of relational autonomy theory a narrative constitution view fares better in these respects than the planning account.

Theories of narrative self-constitution provide an alternative way of conceptualizing the diachronic constitution of identity and self-governing agency, using narrative interpretation as a model. Narratives are holistic, organizing, interpretative structures that explain actions and events by integrating them into meaningful and coherent temporal patterns or sequences. To be meaningful and coherent, a narrative must explain the causal connections between the events and actions it recounts; it must structure event sequences into temporal orderings that need not be chronological but must be intelligible; and it must provide an explanatory and interpretive context within which individual events and their significance can be understood. This interpretive context represents these events from a perspective, or from a variety of perspectives, internal or external to the narrative. A further feature of narratives, highlighted by Goldie, is their power to communicate and arouse evaluative and emotional import (2012, pp. 22–23).

Theories of narrative self-constitution claim that we constitute our identity and agency over time by exercising narrative capacities of understanding and interpretation. Like Korsgaard’s notion of practical identity, a self-narrative is a self-conception by means of which an agent attempts to make sense of who she is and of her temporally extended agency. Talk of “who one is” can be misleading, however, suggesting some deep or true self awaiting discovery. Indeed, the notion of self-governance as “authenticity” has sometimes been interpreted as referring to such a self. The notion of self-narration, as I understand it, requires no such deep self. Who

---

6The discussion that follows refers to some of my earlier work on narrative agency (see especially Mackenzie 2008b, 2009, 2014b; Mackenzie and Poltera 2010), which is influenced by the work of Peter Goldie (2000, 2003, 2012), Marya Schechtman (1996), David Velleman (2006) and Richard Wollheim (1984). My thinking about the explanatory advantages of narrative self-constitution views compared with the planning account has also been influenced by Doug McConnell’s (2014, 2016a, 2016b) work on detrimental self-narratives, discussed below.

7For example, both Marilyn Friedman (1986) and Susan Wolf (1987), mistakenly in my view, interpret the hierarchical views of Frankfurt and Watson as committed to real or deep selves.
one is of course is constrained by features of one's practical identity or self-concept over which one has limited or no control, including contingencies of social identity, biography, and personal history, and embodied and intellectual capacities. Indeed, the process of self-narration partially involves interpreting the significance and meaning of these aspects of one's identity and biography. However, these interpretations are inevitably shifting and open to revision. Events in one's biography take on different practical and emotional saliences at different times of one's life. The significance of various social role identities to our self-conceptions shift over time, and our different social role identities can also create practical conflicts that may put pressure on us to reconfigure our self-conceptions. Self-narration is a dynamic process of creating some kind of provisional coherence and unity among the different strands of one's social identity and self-conception. But these patterns of coherence shift and change over time. As Genevieve Lloyd explains: “The reality of a life lived in time is a perpetual weaving of fresh threads which link events and lives—threads that are crossed and rewound, doubled and redoubled to thicken the web” (1993, p. 144).

Talk of self-narration can also be misleading, in suggesting that we exercise an improbable degree of authorial control over how our lives unfold. However, unlike literary narratives, with carefully structured, teleological plots, human lives are subject to randomness and contingency and unfold and change in haphazard and unpredictable ways. Self-narration, as I understand it, does not require or imply this kind of authorial control over one's life. Narrative understanding is rather a way of responding to and making sense of contingency, flux, and temporal change. Some of the contingencies that may confront us, for example, chronic illness, the experience of trauma, or bereavement, can derail our lives and challenge our self-narratives. To reconstruct our lives and reconstitute our identities in the face of such contingencies, we need to find ways of making sense of them by revising our self-narratives. The extent to which we succeed in doing so will determine the extent to which we succeed in governing ourselves.

Another way in which talk of self-narration can be misleading is that it can imply sole authorship and self-creation. In my view, however, self-narration is deeply relational; hence, it is a coauthorial process. Our self-identities are intersubjectively and socially constituted in relations of dependence and interdependence, beginning with the infant's dependency on and earliest interactions with her caregivers. This primary intersubjectivity, which is rooted in corporeal interactions with caregivers, is subsequently layered by more complex forms of intersubjectivity, which are made possible by cognitive and linguistic development, and by our participation in the social world. This complex intersubjective layering includes the way our self-concepts are shaped and constrained by familial and personal relationships, and by the cultural, religious, historical, and political narratives of the communities into which we were born or to which we now belong. Another way of making this point is to say that while our narrative self-conceptions are first-personal, the first-person perspective is always already social. Our self-narratives are based not only on first-person ascription, but also incorporate second- and third-personal ascriptions. I will come back to this point.

Talk of self-narratives being dynamic and provisional can also be misleading, suggesting that our self-narratives need not be bound by constraints of accuracy, making it difficult to distinguish between authentic and confabulatory or self-deceptive self-narratives. However, the sense in which self-narratives are provisional is not that they are unconstrained by norms

---

8See also David Velleman's analysis of three distinct but interconnected reflexive “guises” of the self: the self of self-image or self-conception; the self of personal identity, which refers to the self of autobiographical memory and future-directed anticipation; and the self of autonomous agency or self-narration (2006, chaps.1, 8, and 9).

9For variants of this criticism, see for example Christman (2004) and Strawson (2004).

10Strawson (2004) criticizes the notion of narrative self-constitution on these and other grounds. I respond to Strawson's critique in Mackenzie and Poltera (2010).
of accuracy and coherence. Rather, our narratives are provisional because our life trajectories are open-ended and at any point in one's life there may be several ways in which one's narrative might unfold in the future, some predictable and foreseeable, some completely unexpected. These unexpected narrative continuations can be the result of attempts to resolve practical conflicts, such as a decision to continue with an unplanned pregnancy. Or they can be the result of self-transformative experiences, such as falling in love, losing one's faith, or coming to terms with the sudden death of a partner. In such situations, envisaging different possible narrative continuations requires envisaging what Jan Bransen refers to as different possible alternatives of oneself (2000, 2002). The process of envisaging different possible alternatives of oneself is a temporally extended, complex imaginative process of narrative projection, in which by imagining different possible futures the agent attempts to figure out which of these possible futures makes most sense. Like planning agency, this process is future-directed and does involve practical reasoning and deliberation. However, since its focus is in some sense on determining who the agent will become, it involves both retrospective and prospective reflection, recruiting narrative capacities for memory and emotion, as well as imagination. It also requires capacities for critically evaluating the plausibility and coherence of the envisaged alternatives, in light of the agent's biography, practical identity, interpersonal relationships, and commitments, as well as the broader social and cultural narratives in which her life is embedded.

Self-governing agency, I contend, involves the exercise of these narrative capacities to develop a coherent, if provisional, self-conception that guides the agent's deliberation, including the activities of planning agency. As such, an agent's narrative self-conception has both motivational and explanatory force. However, there are many ways that self-narration can fail to support self-governing agency. We can fail to be self-governing due to failures of practical rationality, for example, when we fail to form intentions or make plans that align with our self-conceptions, or when we fail to enact our intentions due to weakness of will. Prospective narration can also lead to failures of self-governing agency. For example, self-narration can fail if we do not envisage the alternative possible narrative continuations in sufficient detail; or because no matter how well we envisage them, we cannot control how the future unfolds or how others will act; or because we do not understand our emotions sufficiently well; or because we fall prey to various illusions about ourselves. A lot more could be said about each of these kinds of failure. Here, however, I want to focus on the kinds of failures that interest theorists of relational autonomy; failures that are due to the agent's social environment.

I have claimed that our self-identities are constituted through social interaction and hence that our self-narratives are coauthored. Coauthoring is autonomy-enabling when it supports successful self-narration. However, in social environments where the dominant interpersonal, social, and cultural narratives are oppressive, coauthoring can be autonomy-impairing. The social environment can impede autonomous self-narration either through external or internal constraints. Returning to the multidimensional analysis of autonomy that I outlined at the beginning of this section, by external constraints I mean constraints on self-determination, specifically constraints on agents' social freedoms and opportunities

---

11Schechtman (1996) argues that self-constituting narratives must conform to reality and articulation constraints; Nelson (2001) argues that self-constituting narratives must be credible. I have discussed a range of other constraints on self-constituting narratives in Mackenzie (2008b, 2014b).

12Goldie’s (2012) work on narrative thinking insightfully elaborates on the importance and centrality to our mental lives of narrative capacities for memory, emotion, and imagination.

13For an extended discussion of this process, see Mackenzie (2008b). In that paper, I build on Goldie’s (2003) notion of an “external perspective” and discuss the role that the external perspective plays in evaluating prospective imagination-driven reflection and deliberation.
to determine the direction of their lives. Consider the example of an unplanned pregnancy. Now, the range of possible narrative continuations available to a woman in this situation will depend on her social environment. In a context in which abortion is freely available, the option of terminating the pregnancy may be one possible narrative continuation that she envisages. Assuming the decision that she faces involves conflict among different aspects of her self-conception, envisaging this option is likely to involve an exercise of imaginative projection in which she tries to gauge her possible emotional response to terminating the pregnancy. If she goes ahead and makes the decision to terminate, she will start forming the relevant intentions and plans to enact that decision. In a context in which abortion is restricted, the woman lacks the social freedom even to entertain this option as a possible alternative of herself. The external constraints on her freedom thereby block off the possibility of forming the relevant intentions and plans. Another example is when external constraints in the form of inadequate options create double binds that make it difficult to develop self-narratives that are not riven by internal conflict. A lack of affordable childcare or school hours that do not align with expected working hours create this kind of double bind for many women, who, as a result, experience deep conflict within their self-narratives between the norms of motherhood and the norms of career success, both of which they endorse. This conflict results in ongoing diachronic instability. In Bratman's terms, this instability points to inconsistency in these women's ends. However, this inconsistency is an artifact of social organization, not a failure of planning agency.

Oppressive social environments can also result in the construction of what Doug McConnell refers to as “detrimental self-narratives” (2016a, 2016b). Detrimental self-narratives incorporate content from interpersonal or cultural narrative archetypes that is psychologically damaging to the agent, reinforcing negative self-referring emotions and self-destructive behavior, and thereby undermining self-governance. McConnell discusses the example of a young woman who having been sexually abused by her father and other male relatives, was then not believed by social welfare when she later reported that she had been abused by a foster carer. As a result of this experience of testimonial injustice, she incorporated into her self-narrative the content that she was “dirty and promiscuous,” a “little puppet” for men. This demeaning self-narrative then guided her behavior, stopping her from speaking out any more about the abuse. Drug users' self-narratives of being a “hopeless addict” can similarly guide their behavior in ways that make recovery challenging (2016b).15

Cultural narratives that entrench prejudicial social stereotypes, scripts, and biases, or that rationalize social relations of domination and subordination can be similarly psychologically damaging and self-reinforcing when internalized in agents' self-narratives. As mentioned above, our narrative self-conceptions are based not only on first-personal ascription but incorporate second-personal and third-personal ascriptions, including third-personal ascriptions of social identity. Marina Oshana (2005, p. 90) argues that third-personal ascriptions based on subordinating racialized or gendered scripts or stereotypes are incorporated into our self-conceptions, whether we endorse them or not. Natalie Stoljar develops this idea, arguing that there are three ways that oppressive scripts impair psychological freedom. First, agents “adapt themselves to the script even when they don't endorse or desire it” (2015, p. 118); second, they must “respond to the evaluations” of themselves embedded in the script; third, they are required to “counteract the script” to be treated as social equals (pp. 118–119). Stoljar (p. 119) cites an example from Elizabeth Anderson (2010, p. 53) of an encounter late at night at a gas station with a young Black man who had approached her to

---

14Suzy Killmister (2015) discusses the effect of these kinds of double binds on self-governance, although her analysis of self-governing agency differs in important respects from mine.

15McConnell argues that detrimental self-narratives are self-reinforcing. A key to recovery for drug users and survivors of abuse is therefore to find ways of reinterpreting and revising their self-narratives, especially with respect to their pasts.
offer help with her car troubles. Fully aware of the negative stereotypes associating young Black men with criminality and violence, and the emotional responses typically aroused by these stereotypes, he felt the need to counteract the stereotype by assuring her, ‘‘Don’t worry I am not here to rob you’’... holding his hands up, palms flat at face level, gesturing his innocence.”

The incorporation of damaging narrative content is connected to a further way that detrimental self-narratives can impair autonomy, that is, by impairing an agent's sense of herself as self-authorizing. This is the third dimension of autonomy outlined at the beginning of this section. To be self-authorizing is to regard oneself as having the normative authority to take responsibility for oneself and for one's narrative self-conception. It involves regarding oneself as appropriately positioned and authorized to speak and answer for oneself, and as an equal participant in reciprocal accountability relations—as able to account for oneself to others and also to hold others to account. In my view, thinking of oneself in this way involves holding appropriate attitudes of self-respect, self-trust, and self-esteem. These self-referring psychological attitudes are constitutively social, because they are developed and sustained through intersubjective social relations and normative structures and practices of social recognition. One of the insidious effects of social oppression, discussed by relational theorists, is that experiences of injustice, discrimination, and persistent messages of social inferiority and unworthiness that are embedded in everyday social interactions and conveyed through demeaning stereotypes and scripts can be internalized in feelings of shame and humiliation that erode these attitudes and one's sense of oneself as a self-authorizing agent. Gaslighting and repeated experiences of epistemic injustice can also damage an agent's sense of themselves as self-authorizing, causing crippling self-doubt with respect to one's epistemic capacities, especially one's capacities to deliberate and make up one's own mind.

In this section, I hope to have shown why, from the perspective of relational autonomy theory, a narrative self-constitution account provides a richer analysis of temporally extended agency than Bratman's planning theory, because it provides a conceptual vocabulary for explaining how sociality is integral to the constitution of agency in both enabling and damaging ways. Building on this discussion, in the following section I evaluate Bratman's account of shared agency and Deborah Tollefsen and Shaun Gallagher's concept of we-narratives.

3 | WE-AGENCY AND SOCIAL OPPRESSION

One of the virtues of Bratman's account of planning agency is its attention to the role of joint or shared intentions and shared cooperative activity in temporally extended agency. An agent's

---

16Françoise Baylis (2012) argues that narrative self-interpretation involves “an iterative cycle of ‘self’-perception, ‘self’-projection, ‘other’-perception, and ‘other’-reaction” (p. 128). Achieving some kind of temporary stability or equilibrium among these different perspectives is an ongoing “interpersonal, communicative activity” (p. 123), which can be especially fraught in oppressive social contexts.

17The notion of self-authorization picks up on themes in the work of Charles Taylor (1985) on autonomy as taking responsibility for self; Andrea Westlund (2009a, 2015) on autonomy as a disposition for dialogical answerability; and Paul Benson (2005) on autonomy as having a sense of ownership of one's actions and choices.

18See Joel Anderson and Axel Honneth (2005) and Mackenzie (2008a) for detailed discussions of the importance of these self-referring attitudes for autonomy.

19Cudd (2006, pp. 176–178) claims that shame is one of the indirect psychological harms arising from the assimilation into one's self-concept of socially oppressive beliefs and values.

intentions and plans organize her life over time not only at the intrapersonal level, but also by interlocking at the interpersonal level with the intentions and plans of others with whom she engages in shared activities. Bratman's account of joint intention is modelled and built upon his account of individual intention and planning. His aim is to show how the interpersonal coordination of individual planning agency can yield what he refers to as “modest sociality” (2009, p. 155). This form of sociality is a “we” of “interconnected, interrelated individual planning agents” (p. 163), who temporarily form a group that functions as a causal agent. Joint or shared agency is governed by the same rational norms as individual planning agency: stability, means-end coherence, and consistency. Bratman argues that coordinated interpersonal action and planning that satisfies these rational norms involves the following structure (pp. 155–160):

(i) Each party intends the joint activity. For example, if we plan to hike together on Saturday, I intend that we hike and you intend that we hike. A commitment to our joint activity is embedded in our individual thinking and planning.

(ii) The intentions of each party are interlocked, in the sense that each of our intentions to hike refers to the intentions of the other to hike.

(iii) The intention of each party is that their subplans will mesh and be mutually compatible. Meshing subplans will require that if we have different subplans about which hiking route we should take, we should deliberate, make adjustments to our individual subplans, and reach an agreement about the preferred route for our hike.

(iv) Each party has a disposition to help the other realize the joint plan if needed, for example, I might offer to give you a lift if it would take you too long to get to the start of the route by public transport.

(v) For each party, the persistence of the joint intention to hike together on Saturday depends causally and rationally on her knowledge that the other also intends to hike.

(vi) Each party will be mutually rationally responsive to the other with respect to tracking the joint action. For example, having made the arrangement on Sunday to hike the following Saturday, we might agree to message each other on Friday night to reconfirm the plan to hike together the next day.

(vii) The above conditions are common knowledge among the parties.

Bratman's account of the structure of shared intention allows that each party might have different reasons for engaging in the joint activity. I might want to hike a particular route because I want to see the wildflowers in bloom; you might want to hike it because it involves a challenging climb. Despite these different reasons, our joint intention to hike together provides a common framework of shared valuing with respect to hiking. Nevertheless, it is a not uncommon feature of social coordination that our different reasons might threaten the stability of our joint intention. What if I find out that the wildflowers have stopped blooming and wonder if we should revisit our prior intention to hike this particular route? This is an analogue at the interpersonal level to the issue of stability of prior intentions in individual planning. One response might be to say that if I am sufficiently responsive to your reasons, then this might provide me with a reason not to raise the question of revisiting the route. But given our awareness that we often choose hiking routes for different reasons, we might agree on a shared policy that we take turns in deciding the route. This shared policy functions to mutually reassure each other that once a route has been agreed upon we will not reconsider it. And this mutual reassurance generates an obligation on each of us not to reconsider. 21

21 An important point of difference between Bratman and Gilbert is that whereas Gilbert's account of joint commitments builds obligation into the analysis of joint intention, on Bratman's account obligations arise from and are consequent upon joint intentions. See Bratman (2009) and Gilbert (2009) for discussion of their different views on joint intentions and commitments.
As far as I am aware, Bratman does not investigate how shared agency, as described, supports self-governing agency, except insofar as an individual’s intentions and plans often involve coordination with others. From the perspective of relational autonomy theory, however, shared agency is an important aspect of what I referred to earlier as the “social scaffolding” required for self-governing agency. Joint intentions and the obligations and commitments they generate can function as a way of keeping us on track with our own goals and plans. The commitment to hiking with you on Saturday, for example, helps me to keep on track with my fitness goals; the commitment to coauthoring a paper helps to counter my tendencies toward procrastination. Joint activity also opens a wealth of opportunities of choice and skill development that scaffold autonomous agency. For example, hiking with others expands my ambitions and capacities, enabling me to extend my range of hiking options to include multiday hikes, and to develop the skills and knowledge required to safely undertake these hikes. These kinds of joint activities and the shared experiences and emotions they generate typically also generate shared memories that not only consolidate the sense of “we,” but also, like mutual responsiveness, point to more constitutively dialogical and intersubjective dimensions of autonomous agency. As Andrea Westlund argues (2009b), processes of reasoning, deliberating, and deciding together with others to reach a mutually agreed practical standpoint requires that each party hold herself answerable and accountable to the other(s). Moreover, iterative dialogical deliberation and holding oneself answerable to another often not only requires critical examination of one’s own perspective but can also lead to transformation of that perspective. Mutual responsiveness in situations where the parties are mutually concerned to give due consideration to each other’s perspectives therefore supports not only critical self-reflection, but also one’s sense of oneself as a self-authorizing agent.

However, from the perspective of relational autonomy theory, Bratman’s modest sociality is too modest, for two reasons. First, although shared agency as Bratman characterizes it is central to social life, Bratman’s account of basic shared agency fails to engage with complex forms of shared agency and excludes too many kinds of social groups. This is because it is based on a voluntarist and restricted conception of sociality. Bratman’s account of basic shared agency explicitly excludes any social group that involves relations of asymmetric authority and differential power. So, he says that examples of shared intentional agency include: “our singing a duet together, but not the activities of an orchestra with a conductor; our going to NYC together, or . . . our walking together, but not a school trip led by a teacher; our having a conversation together but not an exchange in which a language teacher guides a novice” (2009, p. 155). These kinds of activities are what he refers to as shared cooperative activities. Now while Bratman does acknowledge that shared cooperative activities “can involve large numbers of participating agents and can take place within a complex institutional framework” (1999, p. 94), giving the example of an orchestra in this context, in much of his work the analysis of joint intention and shared cooperative agency is restricted to simpler two-person cases.

Second, and more importantly for the argument of this article, Bratman’s intentionalist and voluntaristic account of a social group fails to account for nonvoluntary social groups, which are defined by social markers and ascriptions of identity. Nonvoluntary social groups clearly do not meet Bratman’s conditions for either shared agency or shared cooperative activity. Nor do they meet Gilbert’s conditions for joint commitment. So why should their exclusion from a theory of agency and sociality be a concern? I address this question below,

---

22Bratman’s most recent book (Bratman 2022) does engage with more complex forms of shared and institutional agency. A detailed discussion of Bratman’s most recent work is, however, beyond the scope of my concerns in this article.

23See Ann Cudd’s discussion and critique of Gilbert’s account of social groups as constituted by joint commitments (2006, pp. 37–40). My discussion here is influenced by Cudd’s analysis.
via discussion of Deborah Tollefsen and Shaun Gallagher's account of we-narratives. In my view, the concept of a we-narrative is an important advance on Bratman's account, incorporating a richer analysis of shared agency and enabling a more critical perspective on the dynamics of social groups.

Tollefsen and Gallagher (2017; Gallagher & Tollefsen, 2019) work with a much less restrictive conception of social groups than Bratman's, insofar as they include hierarchically structured social organizations and institutions in their analysis of we-narratives. The concept of a we-narrative, which is modelled on the concept of a self-narrative, explains the role that a group's self-conception plays in guiding its activities and deliberations. Although modelled on the concept of a self-narrative, Tollefsen and Gallagher point out that the concept of a we-narrative differs in an important respect. As discussed in the previous section, self-narratives are self-interpretations from the first-person embodied perspective of the agent and incorporate a sense of prereflective experiential self-awareness, as well as a reflective perspective. However, in a we-narrative there is no correlative first-person plural phenomenological sense of “we.” This is because, like Bratman, Tollefsen and Gallagher claim that we as “first-person plural is a grammatical categorization, not an experiential categorization” (2019, p. 217). Even in cases where there seems to be an experiential sense of a fleeting “we,” for example, in spontaneous dancing together, this experience of the “we” is “had by each individual participant, and may be different for each one” (p. 213). While I am sympathetic to this argument, I think there is more to say about how a phenomenological sense of “we” can arise from reciprocal second-personal emotional interaction.24

In the case of groups engaged in longer-term coordinated projects or groups with an institutional structure, the “we” is established reflectively, through communicative practice and intersubjective deliberation relating to shared goals, intentions, plans, and actions. These forms of shared agency constitute a narrative practice, a practice that generates narratives about “what we are doing, have done, and will do, or what we ought to do and want to do” (p. 214). We-narratives are both retrospective and prospective, including reference both to the group's history and its future goals and plans. For example, the we-narrative of a couple will include memories of their history and activities together, as well as reference to their future goals and plans. This narrative reflects their conception of “who” they are as a couple, and in so doing helps to specify the content of their mutual assurances and commitments. Thus, it functions to deepen and stabilize their joint identity and activities. In contexts of disagreement, for example when one member of the couple is tempted to rescind on a joint commitment, the narrative can help to refocus this commitment (Tollefsen & Gallagher, 2017, pp. 103–104). The we-narrative of a club or an institution will include a narrative of its history as well as its goals, mission statement, and future plans, providing a stable identity for the institution, even if the individuals who contributed to it are no longer members. The absence of an experiential first-person plural thus does not preclude reflective first-person plural attributions to the group, either by its participants or by third parties, of shared intentions, agency, identity, rationality, or narrative capacities. The group's we-narrative and activities might be said to constitute its intentional stance, a stance from which individual members can represent or speak, as when a spokesperson for a political party talks about “our policy on climate change” (2017, p. 107; Gallagher & Tollefsen, 2019, p. 217). This stance also plays a regulatory role in putting rational pressure on individual members of the group to cohere with the group's narrative.

The concept of a we-narrative provides a useful explanatory model for understanding how many social groups stabilize their identity and commitments over time. Even so, this conception of a social group is still voluntarist, in the sense that the group's identity is

24For an interesting discussion, see Zahavi (2015).
constructed by individual group members and its we-narrative is agreed upon or at least endorsed by those members. Tollefsen and Gallagher, citing Cassie Striblen (2013), do acknowledge that a group's identity can involve both internal narrative construction by group members, and external construction by non–group members, who can hold the group to account, perhaps by contesting aspects of the group's narrative and values. This idea mirrors the way that members of our social network, for example friends and family, play an important role in holding us to account for our self-narratives, for example, by checking and testing their accuracy. However, there are two issues raised by this account of internal and external construction. First, as Tollefsen and Gallagher acknowledge, it paints a somewhat optimistic picture of the relationship between a group's narrative and its individual members, overlooking the way that “narratives can take on a life of their own, can transcend individuals and come to support an instituted structure, as well as oppress or dominate (in explicit or subtle ways) the intentions of the individuals that belong to the group” (2019, p. 218). In response to this issue, they call for a “critical narrative practice” that involves taking up a critical perspective on the narratives of powerful social, legal, and political institutions and their effects on individual autonomy, interpersonal relationships, and democratic processes.

The second issue is that this account of external construction still remains with a voluntarist framework that does not attend to the social effects of nonvoluntary social group membership. In contrast to internalist and voluntarist analyses of social groups, Ann Cudd provides an externalist definition of social groups that encompasses both voluntary and nonvoluntary social groups. She defines a social group as “a collection of persons who share (or would share under similar circumstances) a set of social constraints on action” (2006, p. 44). These constraints, which shape and guide individual agency, are institutionally structured and include “legal rights, obligations and burdens, stereotypical expectations, wealth, income, social status, conventions, norms and practices” (p. 41). Thus, they include constraints that are the “result of intended collective action,” as well as constraints that are the “unintended consequences of intentional actions” (p. 42). What members of nonvoluntary social groups share in common is the “same social constraints, as a result of others' decisions, while voluntary group members share constraints that result from their decisions and actions to join together, against a background of social constraints consequent on others' decisions” (p. 44). I agree with Cudd that such an externalist account of social groups is needed to explain the effects of social oppression on the agency of both individuals and socially ascribed groups.

Thus, while I support Gallagher and Tollefsen's (2019) call for a critical narrative practice, I think this call puts pressure on the overly voluntarist conception of social groups to which many theories of collective intentionality seem to be committed. A critical narrative practice requires not only taking up a critical perspective on dominant we-narratives. It also requires acknowledging the way dominant we-narratives often involve the construction of “they-narratives” about other social groups that function to maintain oppressive social dynamics and institutional structures. They-narratives are interwoven with prejudicial and often demeaning stereotypes, implicit and explicit biases, false attributions of intentional stances to groups, and distorted histories of intergroup relations that function to justify unjust relations of power. They-narratives also have emotional and evaluative import in that they both arouse and provide justification for negative emotions such as contempt, fear, or hatred directed toward the subjects of the narrative. Colonial narratives about colonized peoples are prime exemplars of they-narratives that have been used to support we-narratives that justify or at least remain silent about histories of dispossession and expropriation. For example, the history of Australia is replete with celebratory we-narratives of nationhood that also involve the construction of false and demeaning they-narratives about the first nations people of Australia. The founding narrative falsehood, that Australia before the British invasion was a “terra nullius,” an uninhabited land, helped
support a false we-narrative that the nation was not founded on violent dispossession and expropriation. Celebratory we-narratives of nationhood that distort the truth of our shared history have shaped social relations and emotional interactions between colonizers and colonized for more than two hundred years.\(^{25}\)

Recalling Stoljar’s (2015) analysis of the way subordinating social scripts constrain individual agency, false and subordinating they-narratives both construct and constrain social group agency, requiring the group to adapt to the narrative, respond to the evaluations of themselves embedded in the narrative, or counteract the narrative. For marginalized social groups, adapting, responding to or counteracting dominant we-narratives can often be experienced as a form of epistemic injustice; an exhausting, autonomy-undermining process of having to repeat the same messages again and again while feeling as though one is never really being heard or understood. At the same time, the process of contesting and countering dominant they-narratives can generate powerful forms of political activism and resistance, and new we-narratives that both reclaim the narrative of “who we are” and function to unify groups who are subject to oppressive they-narratives. For example, first nations people in Australia comprise around two hundred and fifty distinct language and cultural groups. The social group “Aboriginal and Torres Strait Islander people” is thus an artifact of colonization. However, appropriating this social group identity, adopting a flag that represents this identity as distinct from the national identity represented by the Australian flag, and generating new we-narratives about this identity, has been central to Aboriginal political activism and resistance to oppressive colonial they-narratives.

So, my point is not to dispute the explanatory value of the concept of a we-narrative. It is rather to argue that a critical narrative practice points to the need for an expanded analysis of social groups. Voluntary social groups play an important role in social life, and theories of collective intentionality have done much to illuminate this role. However, a critical narrative practice calls for an externalist analysis of social groups, one that explains how nonvoluntary group membership can unjustly constrain both individual autonomy and the agency of groups. A critical narrative practice also needs to attend to the power of we-narratives to arouse shared emotions. While shared emotions enable social group identification and a sense of “we,” these emotions can derive their power to constitute the “we” by fostering hostile and negative emotions toward the subjects of they-narratives.\(^{26}\)

ACKNOWLEDGMENT

Open access publishing facilitated by Macquarie University, as part of the Wiley - Macquarie University agreement via the Council of Australian University Librarians.

REFERENCES


\(^{25}\)Although these we-narratives falsifying the historical record have been debunked by both indigenous and nonindigenous historians, they still retain a grip on the national imagination. The documentary The Australian Wars (2022) by indigenous film-maker Rachel Perkins, is a recent powerful attempt to change the narrative.

\(^{26}\)For helpful comments on an earlier version of this article, thanks to the audience at the 2022 Spindel Conference, and especially to my commentator Sarah Vincent, as well as participants at a colloquium of the Ethics Institute, University of Utrecht.


**AUTHOR BIOGRAPHY**

**Catriona Mackenzie** is Emerita Professor of Philosophy at Macquarie University, Sydney. Her primary areas of research are feminist philosophy, social and political philosophy, philosophical moral psychology, and applied ethics. She has edited six volumes and authored numerous articles and book chapters in these areas. Her most recent edited volumes are *Procedural Justice and Relational Theory: Empirical, Philosophical and Legal Perspectives* (Routledge, 2021), with Denise Meyerson and Therese McDermott, and *Social Dimensions of Moral Responsibility* (Oxford University Press, 2018), with Katrina Hutchison and Marina Oshana.