Received: 9 March 2022 | Revised: 27 July 2022 | Accepted: 31 July 2022

DOI: 10.1111/jtsb.12360

ORIGINAL ARTICLE

WILEY

A Motivational Theory of Roles, Rewards, and Institutions

Seth Abrutyn[1] | **Omar Lizardo**[2]

[1]The University of British Columbia, Vancouver, British Columbia, Canada

[2]University of California Los Angeles, Los Angeles, California, USA

Correspondence
Seth Abrutyn, The University of British Columbia, Vancouver, BC V6T 1Z1, Canada.
Email: seth.abrutyn@ubc.ca

Abstract

In this paper, we attempt to rehabilitate the notion of role by linking sociological role theory to recent work on motivational, affective, and cognitive neuroscience specifying the internal mechanisms behind motivated action. We argue that there is nothing inherently problematic or retrogressive in the idea of "role," once its link to a purely normative account of motivated action is severed. Instead, by conceptualizing roles as emerging and persisting around structured *reward systems*, we are able to incorporate contemporary motivational science such that rewards become the proximate causal mechanisms currently missing in role theory. Consequently, a key implication of our argument is that the best way to link role, action, and structures is by reviving the idea of institutions as literal reward systems, which allows us to envision roles as the mechanisms via which the pursuit and delivery of rewards and goal-objects are routinized. Implications for a motivational theory of roles, rewards, and institutions is discussed.

KEYWORDS
institutions, motivation, neurosociology, reward, roles, social action

1 | INTRODUCTION

In an *Annual Review* piece written in 1986, Biddle wrote that "the concept of role is one of the most popular ideas in the social sciences. At least 10% of all articles currently published in sociological journals use the term in a technical sense" (1986, p. 67). Biddle's review article sought to end "confusion and maltintegration" surrounding the concept by synthesizing functionalist, symbolic interactionist, structural, organizational, and cognitive accounts. In an ironic twist, Biddle's review piece

marked the beginning of the end for the concept of role as an analytical category in sociology. Biddle recognized scathing criticism of roles coming from many corners of sociology as new generations swept away Parsonian functionalism's residue. Notably, feminist scholars raised the most blistering criticism of roles, seeing the idea as supporting traditional gender ideologies in families and the workplace (Connell, 1979).

Excepting Jonathan Turner's (2010b) continued emphasis on role dynamics, the role concept survives in truncated form. Symbolic interactionists (Stryker, 1980) speak of "role-identities" while others note people's capacities to "take the role of the other" (Wiley, 1994, p. 86), revealing the diffusion of dramaturgical ideas in sociology but missing the original concept's theoretical bite or general import. For most sociologists, "role" as a sort of sociological signpost survives while role theory (and its analytic utility) is dead. Nevertheless, like so many other ideas tossed aside by association with functionalism, roles remain a promising and empirically powerful tool for understanding a wide range of phenomena, particularly the linkages between structure and agency, culture and action, and structure and culture (Tomasello, 2019). Moreover, it is one of the most mature theories in formalization, development, and empirical testing that sociology has ever produced (Turner, 2001).

In their heyday (see Parsons, 1951b), roles became a vehicle of social structure at the micro-level, albeit predicated on *normative* control, meaning deviation was "abnormal" and even *pathological* (Parsons, 1942, 1951a, 1975). However, like most critiques of Parsons and functionalism, some of it was overdone; like values (Hitlin & Piliavin, 2004) or socialization (Guhin et al., 2021), roles remain a key concept whose widespread currency in sociology offers more opportunities than costs. Indeed, the heavy theoretical lifting moving us beyond Parsonian normativism has already been done (Biddle, 1986; Turner, 2001).

So, what to do? First, the role concept cannot be entirely detached from its link to *social structure*. In a classic statement, Linton (1936) used roles to make sense of the relatively small number of repertoires a foraging society revealed. Small-scale societies present certain observational advantages, as relatively homogeneous populations functioning in small groups afford a view of how people and the collective interface, a point echoed in contemporary sociology by Fine (2012). Second, in this case, kinship subsumes the entire societal apparatus—imposing a kinship 'logic' on economics or politics, or the law (Abrutyn, 2013)—highlighting the importance of possible role-based typifications (parent-child; husband-wife, and the like) and the patterning of behavioral repertoires (Martin, 2009). Third, when a society grows in size, scale, and diversity, roles continue to make sense structurally, helping explain many issues surrounding control, coordination, and integration. When impersonal relationships and encounters outnumber personal ones, local, tangible sites of informal or quasi-formal authority give way to distant, abstract, formalized centers. This ensures trust between strangers, reducing uncertainty across an array of situations while reducing the burden that would emerge absent roles as structural short-cuts. In this way, roles are a mechanism ensuring the large-scale coordination of action and interaction (Berger & Luckmann, 1966). None of this is to say that situations are reproduced perfectly, people are automatons, or that status differences, inequality, regional variation, and the like do not matter. Instead, roles are mechanisms of mimetic isomorphism (DiMaggio & Powell, 1983); in a world of *im*personal and, more expansively, *de*personalized relationships, roles are central tools for understanding why people are motivated to participate in coordinated social action.

2 | ROLES IN FUNCTIONALIST THEORY

Roles have been conceptualized as attached to *status* in structural-functionalist accounts (Linton, 1936; Parsons, 1951b). Status is usually seen as a position in a social structure in which "each actor is an

object of orientation for others" (Parsons, 1951b, p. 25). Relative differences shape orientations and thus access to critical resources: "each actor is oriented to other actors [by] playing a role" (*ibid*). Roles are normative prescriptions and proscriptions conditioning actors' reciprocal orientations. For example, Actor A expects Actor B's response, thereby making coordinating behavior and attitudes easier than in a situation where behavior was random or spontaneous (Parsons & Shils, 1951). Via socialization, people internalize norms, compelling them to enact roles or be subjected to informal or formal sanctions. To borrow a metaphor, for Parsonian functionalism, roles are vehicles of social structure, acting to translate what is macro and external into micro, internal, and interaction-specific roadmaps.

In this way, roles solved a significant analytic problem: dealing with the paradox of social integration and coordination in modernity's wake. As Durkheim (1933) argued, as groups grow in size, density, and diversity, they segment and differentiate but must retain a shared collective consciousness. Consequently, the larger group's cultural system must necessarily become abstract and general to account for horizontal (different clans within a tribe) and vertical (one clan being the clan of chiefs) differentiation, experiential variation, and tensions between super- and subordinates. Parsons saw roles as a solution to this dilemma, being the source of structural and cultural equivalences reducing complexity, and uncertainty, while enhancing trust. Doctors, for example, became a typified social category whose basic repertoires were stereotyped such that a person could find any doctor, in theory, and orient themselves with ease. On the one hand, because some aspects of professional authority are similar (lawyers and doctors draw from comparable beliefs and practices), roles are generalized mechanisms encouraging integration. On the other hand, role performances vary based on the incumbents in a given situation and a host of variables, as symbolic interactionists later emphasized (Turner, 1956, 1976).

Toward the end of Parsons' life, after functionalism had already been mortally wounded, role theory was dealt what appears to be fatal blows by gender scholars who saw it as supportive of existing (traditional) gendered ideologies (Connell, 1979). These scholars criticized role theory for drawing arbitrary boundaries separating the normal from the abnormal, emphasizing stasis over change, assuming smooth and passive socialization, under-emphasizing resistance and lived experience, and under-theorizing conflicts between roles in different arenas (Jackson, 1998, pp. 51-53). Given these critiques, it seems appropriate to abandon roles as explanatory tools. Nevertheless, even as Biddle's paper marks the declining significance of "role," sociology has not been able to quit its oldest manifestation: Mead's idea of *taking the role of the other*. There is something about roles that remains relevant, perhaps even central to the sociological project.

We argue that the overarching issue identified by most critics and responsible for roles' premature death is the *normative commitment* to how things are what many role theorists explicitly or implicitly assume. As Dennis Wrong (1963) pointed out decades ago, this type of motivational mechanism is frequently resorted to in pseudo-explanations for why people act. Roles, for instance, *must be* normatively conformed to so that we can meet the expectations of others. These expectations are consensual because humans *need* ontological security. Are these motivational assumptions correct? So, long as we assume they are, roles have virtually no empirical or theoretical value beyond over-socializing and over-determining behavior. Below, we examine some worthy alternatives to normativist roles, concluding that roles remain a relevant model for envisioning how actors interface with structure and culture. However, to achieve this relevance, we must rebuild the theory of motivation accounting for role performance on new, non-normativist foundations.

3 | TOWARD A NON-NORMATIVIST ROLE THEORY

Almost a decade before Biddle's (1986) review, symbolic interactionists had already begun to pivot away from the role idea, merging it with the concept of *identity* (Stets & Burke, 2000, 2014). Like roles, identities are structured, but rather than study behavioral repertoires, we analyze each role's internalized meaning sets, serving as a "standard" by which external responses to one's performance could be measured against expectations. The focus remains on roles central to contemporary social life (like "mother" or "wife" or "student") but with a focus on the "reflected appraisals" elicited by incongruities between expectations and reality (Burke & Stets, 2009). The move to reflected appraisals gives identities more flexibility, inviting both generalized and idiosyncratic meanings into a given role, but may not represent an apt substitute for the original idea. Two core problems in role-identity research support our assertion.

First, the motivational mechanism in role-identity research relies on outdated homeostatic, drive-reduction models (Kringelbach & Berridge, 2016). Just as functionalists presume roles are governed by normative commitments enhancing conformity to avoid negative sanctions, the identity standard acts like a thermostat with a set point, triggering action only when the external meanings under or over-shoot the preset standard. This is indeed more flexible but still relies on a normativist dynamic—rather than conforming to avoid adverse outcomes, the individual acts deliberately to attempt to produce alignment between expectations and others' responses. A second issue derives from a sole focus on cognition and intrapersonal mechanisms at the expense of behavior motivating us to acquire the meanings in the first place, helping to coordinate with others (Haslanger, 2019). All role theories rely on three core concepts: "patterned and characteristic social behavior, parts, or identities that are assumed by social participants, and scripts of expectations for behaviors that are understood by all and adhered to by performers" (Biddle, 1986, p. 68). Identity theory's focus on meanings removes the individual from social situations and the coordination of bodies and minds in space, losing the scripts of other performers.

An alternative addressing these problems emerged in a theoretical maneuver shifting from role to "category." In many ways, Goffman's conception of roles as embedded in performance allowed gender scholars, for example, to drop them altogether to focus on the "doing" of gender and other categorical distinctions (West & Fenstermaker, 1995; West & Zimmerman, 1987). The focus on role-based *action* is welcome here, but what *drives* people to constantly "do" difference remains unclear. What do people "get" from "doing" categorical roles? These phenomenological approaches are limited concerning the *motivation* for performing roles because they collapse the distinction between action and social action (Campbell, 1996). Roles cease to be characteristics of people, thus subject to affective and cognitive constraints, instead becoming "emergent" features tied to situations (West & Fenstermaker, 1995, p. 9).

3.1 | An Ironic Twist

Nevertheless, the analytic notion of role lurks behind these alternatives, freed from normativist commitments. It is a little ironic, then, that the classic text in medical sociology is Parsons' (1975) writings on the "sick role." While not without its critics (Freidson, 1970), the concept remains relevant to this day for two reasons. First, it is far more flexible than its initial critics presumed, as Parsons left plenty of space for different "forms" of doctor-patient relationships. Second, it is relatively removed from political lines drawn by those who saw roles as supportive of traditional gender ideologies. Undoubtedly, there are doctors and patients throughout the Western world whose relationship, at least

to some degree, remains patterned even as the medical authority of doctors has gradually eroded—albeit never completely replaced (Stivers & Timmermans, 2020). Indeed, we believe the sick role is as exemplary a role as any to raise the question: what would roles look like without a normative motivational mechanism?

Ironically, a (partial) answer may have been stumbled upon by Parsons (1951b) when he began to embrace a naturalistic approach to rewards as primary motivators, going against the sociological tendency to see rewards as indelibly tied to reductive utilitarianism. Parsons notes, for instance, that infants are easily motivated to be socialized, that is, to pay attention to adults, because they need "to secure the specific rewards of being fed, kept warm…and avoid the corresponding deprivations which constitute the first beginnings of his playing a role as distinguished from being merely an object of care" (p. 215-6). Assertions, incidentally, are not at all controversial in contemporary developmental science (Davis & Montag, 2018; Decety & Yoder, 2017). Parsons returns to rewards when talking about social control. Here, "relational rewards," like "alter's attitudes of love, approval, and esteem" (1951b, p. 301), ensure behavior oriented to desired social objects as we seek to secure rewards received in the past. Eventually, these particularized rewards can be institutionalized into generalized role relationships, like a doctor-patient role. The Doctor role "as a technically competent expert" affords occupants the ability to "manipulate the reward system" because their "approval…has meaning to the patient because of his professional authority, which is anchored in the values of the" medical institution (p. 315).

3.2 | The Sociology of Roles and Rewards

Despite its analytic advances, Parsons's theory of rewards relies too heavily on Freudian hydraulics and, even so, remains underdeveloped because rewards never became central to his thinking on roles. Luckily, we can recover insights from another role theorist, Ralph Turner, who went well beyond Parsons' initial insights and which can serve as a bridge to contemporary work in motivational science in building a non-normativist role theory.[1]

Ralph Turner (2001, p. 240ff) specifies commitment to role performance via three mechanisms: functionality, representationality, and tenability. While the former two have been most extensively examined, it is the idea of *tenability* that can best link to a non-normativist account of the motivation for role performance. Turner argues that roles accrue cost/benefit balances as people acquire, practice, and perform them. Even roles assigned at birth are costly endeavors to sustain but are also sources of rewards. Powerful and prestigious roles—Doctors, CEOs, Political Leaders—reveal ledgers in the black, but others, like a classically trained violinist or, hitting more closely to home, a professional sociologist, appear to place the incumbent in the red. Subordinate or less socially valued roles accrue meager amounts of *extrinsic* rewards like money or are predicated on high sunk costs, and intermittent returns do not make as much sense. Instead, some roles inherently or via cultural beliefs provide a suite of *intrinsic* rewards often ignored by accounts focusing solely on the classic triumvirate of wealth, power, and prestige.

In this way, Turner (1978) highlights an important distinction often conflated in sociological approaches to motivation: That between *extrinsic* and *intrinsic* rewards. Extrinsic rewards, like money, may help differentiate roles and induce commitment. Still, they are detachable from the role and situation—portable, if you will. In contrast, intrinsic rewards, "such as admiration or artistic virtuosity or filial devotion," are inseparable from the performance, specific goal objects, and situations in which the role is salient. The power of intrinsic rewards to motivate people to stick to roles more genuinely and fully than extrinsic rewards helps explain why CEOs are typically imagined as

mercenaries with minimal commitments to people or groups while the artist is engrossed in a labor of love (Hahl & Zuckerman, 2014; Ihering, 1968). Indeed, as Turner suggests, the likelihood of low extrinsic rewards—both in the immediate and long-run—and the recouping of intrinsic rewards offset the balance. Consequently, a "meager reward" by utilitarian standards is significant enough to encourage people to keep at it, seeking situations that bring these intrinsic rewards. On the other hand, copious extrinsic rewards may lead people to become less accountable to the role, explore different roles, or even disingenuously play the current role.

In this way, Ralph Turner pulls role theory closer to the contemporary motivational neuroscience of reward (Kim et al., 2016). Thus, Parsons and Turner were right in isolating rewards as crucial to acquiring roles, remaining motivated to play them, and leaving those roles over time. While the reward mechanism is antithetical to standard models of action in sociology, we are proposing neither a crude economic model nor a return to Skinnerian behaviorism. Indeed, when we graft motivational science onto sociological role theory, we find a much more flexible, sociologically satisfying theory of motivation. This approach avoids the oversocializing penchant to rely on external forces like values and norms, which fall flat when scrutinized by modern brain science, as distal causes do not appear to move the dial for mobilizing action (Daw & Shohamy, 2008). Instead, the proximate nature of rewards causes role-based behavior to occur (Miller Tate, 2021). Notably, this approach to role motivation resolves the problem of preferences. Neuroimaging has repeatedly connected value and reward (Di Domenico and Ryan, 2017). As Tamir and Hughes note, "[a] set of dopamine-rich midbrain structures…and areas of cortex…reliably respond both when people anticipate or experience rewarding outcomes…each structure in the circuit computes relevant components of reward processing…" (2018, p. 702).

Notably, the intrinsic rewards Parsons listed as motivating infants to pay attention and learn from adults, like food or comfort, and those more pertinent to sociology, like money, deference, and social approval, trigger the *same* neural structures, which can also respond to extrinsic rewards. fMRI studies, for instance, have shown that money activates the nucleus accumbens/subcallosal region and VTA (Breiter et al., 2001), the same areas activated by food and sweets. Indeed, dopaminergic systems, systems not associated with liking a reward but with seeking or wanting a reward—a distinction discussed in greater detail below, are linked to reward detection and expectation, the representation of goals, and the integration of sensory inputs to prepare for action (Aron et al., 2005).

Significantly, rewards must be detached from pleasure. Indeed, dopaminergic systems are associated with wanting and learning, neurophysiologically dissociable phases from liking. Pleasure is just one phase of the motivational cycle; if it were the only phase, numerous activities associated with roles not "pleasurable" (working on a paper) would be done poorly or avoided, leading to a coordination collapse in an impersonal world (Haslanger, 2019).

Instead, we are suggesting the following:

1. Humans are motivated to act based on the anticipation, consummation, or satiation of rewards. Anticipatory activity can be as rewarding as consummatory activity.
2. Reward objects are shorthand for physical, social, or ideational elements we can pursue, as well as activities or events. Consequently, just about any object can become rewarding via learning and enculturation processes.
3. Rewards are not usually extrinsic but rather *intrinsic.* Intrinsic rewards are first experienced, as Parsons' argued, as neonates. Object-reward in these cases tends to be tightly coupled and often set up developmentally via ontogenetic processes laid down by evolution.
4. Some intrinsic rewards, however, may be psychological or social and thus subject to learning. Regarding the former, we refer to cognitively satisfying states and appraisals, whereas the latter are

rewards derived by interacting with external objects. Despite this important distinction, both types of rewards are intrinsic, built on the fundamental motivational process described below.

5. Finally, some symbolic placeholders of value and media of exchange, like money, can become extrinsically rewarding if they are culturally sanctioned and tied to intrinsic rewards. Extrinsic rewards piggyback on systems designed to guide behavior toward intrinsic reward pursuit.
6. Roles are generalized vehicles not only of social structure or culture but of reward systems. We learn how and why to play roles, internalize their meanings, play them in the future, and, in some cases, merge our self-concept with the role based on the degree to which they are rewarding compared to other roles we have acquired or believe are available.

Role-based action is motivated by the pursuit of specific rewards to which the role serves as a predictably systematic conduit (Lizardo & Stoltz, 2018). Typically learning about reward/object connections involves partaking in role-based action. Nevertheless, where do rewards come from? For an object to become rewarding for the person, it must be constituted as a reward in previous experience. Motivated action implies a history of learning or vicarious attribution that attaining an object leads to a rewarding experience. Note that this applies to all rewards, even tangible rewards we are biologically pre-programmed for. For example, we are motivated to eat and drink because food and water are intrinsically rewarding. Nevertheless, *which* foods we eat or what we prefer to drink to keep hydrated (like sugary soft drinks, fruit juice, water) is based on a learning history, whereby specific foods and drinks became constituted as rewards.

3.3 | The Fundamental Motivational Process

To construct a non-normativist role theory, we must go beyond sociology's limited capacity to conceptualize rewards and incorporate motivational science into action theory. One roadblock is that contemporary work on motivation in sociology has little connection to recent empirical research on motivation in neuroscience, in which rewards and the human reward system are central to the explanatory story of why people do the things they do (Schroeder, 2004), including, we argue, why people take on roles, keep roles, and even drop them. Luckily, contemporary work in affective, cognitive, and motivational neuroscience point to a fundamental reconceptualization of motivation in the broader social and human sciences. Accordingly, instead of listing "fundamental" *motivations* (Turner, 2010a), we propose to focus on fundamental motivational *processes*, understanding there is a massive (perhaps non-enumerable) set of objects that could count as "motivators," which, in turn, provides flexibility to a revamped role theory.

The basis of the motivational process is a set of phases: anticipatory, consummatory, and satiatory (Kringelbach & Berridge, 2016). All three are neurophysiologically dissociable as they involve functionally distinct brain structures and different neurotransmitters. In the *anticipatory* phase, we engage in—flexibly-repeatable or deliberately intentional—activity to attain the potentially rewarding objects. In the *consummatory* phase, we engage with the object, generating hedonic feelings of pleasure if the reward matches or exceeds expectations. Finally, in the satiatory phase, we store linkages between experienced rewards to update the "reward status" of the object. In this phase, we compare what we thought we would get to what we got to set the stage for future reward-pursuit. Using folk psychological labels for these phases of motivation, we can say that the fundamental motivational processes leading objects to be constituted as rewards are *wanting*, *liking*, and *learning*. Thus, pleasure is one aspect or "phase" of motivated action, making the contemporary understanding of reward-based

motivation neither a return to "hedonic" nor operant-conditioning theories. Instead, elements of these theories are incorporated into a more extensive action cycle (Schroeder, 2004).

It is this motivated action cycle, the ability to anticipate—rightly or wrongly—rewarding experiences with an object (or set of similarly classed objects), and the actual reward itself that are the primary drivers of action, *not* distal forces like norms or values (Martin & Lembo, 2020). An entire panoply of potentially rewarding objects has the potential to come to intentionally guide and control motor programs or become a source of habitually motivated activity. Nevertheless, the primary lesson is that only objects constituted as *rewards* have the causal power to energize action by activating the appropriate mental states (Miller Tate, 2021). Abstract "need-states," uncomfortable drives, experiences of "deflection," or "lack of meaning," "ennui," "ontological (in)security," or "loneliness"—or any other "motivation" in a long list of social psychological traditions—are not objects. Therefore, they cannot be constituted as rewards. They can modulate the *incentive salience* of rewarding objects, but need-states are not the objects that motivate action.

Roles *are structured and enculturated shortcuts to different types of rewards*. Previous sociological work has been misleading in thinking roles lead only to a single reward type (typically approval or deference from others). Instead, roles that draw the most commitment lead to multiple rewards (or "reward cocktails") at once so that people can be motivated to engage in a role-relevant activity even if the role stops delivering one type. In some cases, like highly formalized and institutionalized roles (doctor, lawyer, judge), they are well-worn pathways to specific object-reward couplings available only to those occupying the role. In other cases, there is a fuzzier relationship between the objects and rewards one might pursue, leading to greater diversity in behavioral performance. In what follows, we flesh out the motivational process in more detail and then return to the implications for sociological role theory.

3.4 | Wanting/Liking

The first (wanting) phase of the motivated action cycle is characterized by dopamine production associated with expected pleasure, and the extensive (in terms of brain connectivity) system triggering the production and distribution of dopamine in relationship to objects we interface within the environment is associated with wanting (Smith & Berridge, 2007). During the wanting phase, incentives are salient, driving perception, cognition, and action in the situation. Wanting is perhaps the most temporally extended phase of motivated action. Most times, persons are engaged in this phase of activity regarding an assortment of tangible, psychological, or social rewards. Wanting is, in essence, the source of the energizing effect of motivation on behavior, anticipating the use or resultant pleasure of a reward object (or state) (Kringelbach & Berridge, 2016). Many psychological theorists see the study of motivation as equivalent to action driven by want or desire (Schroeder, 2004). Nevertheless, it is only one, albeit important, phase of motivation.

Another phase of motivated action is rooted in *liking* or the set of activities associated with engaging with the goal object or state once it is reached (consumption, interaction, sensory stimulation). Unlike the first phase, in which the pursuit is everything, consumption can be as much about an object's practical, sensual manipulation as it is about its actual consumption. For psychological and social rewards brought about by role performance, liking may resolve itself into passive (covert) activity. This phase is physiologically distinct from wanting, with liking processes linked to more delimited, discrete opioid-producing subcortical areas (pleasure "hotspots") of the brain like the nucleus accumbens and ventral pallidum (Tindell et al., 2005). Sociologists often conflate these two phases of motivation, using terms like preferences, interests, or goals as "motivators" of action, which sometimes

mean people want something and other times they like something. Nevertheless, making a principled distinction between the energized pursuit of reward objects and hedonic experience post-attainment is key to understanding how social environments shape action and how different forms of action like habit, addiction, thrill-seeking, and the like work.

For instance, because of its dissociability from liking, wanting is tied to incentive salience (prominence of rewarding stimuli in the environment or imagination) or goal direction (Di Domenico & Ryan, 2017)—that is, deliberative action. The hedonic response, on the other hand, is tied more closely to learning. That is, wanting is typically based on activating three different energizing mechanisms (initiation, guidance, control). At the same time, liking is tied to the production, maintenance, and modification of internalized schemas (Miller Tate, 2021). *Initiation* is central to motivated action; in the older Parsonian framework, this was theorized as linked "effort" to get action going given the counter-weight of environmental and biological conditions (Silver, 2011). Unfortunately, certain motivational disorders (like major depressive disorder) disable this crucial step so that people have trouble initiating action despite wanting a given reward object (Panksepp & Watt, 2011). *Guidance* refers to mechanisms whereby action, once started, is directed by extant mental models of the environment linking desires to behavioral predictions. Finally, motivated behavior must be *controlled*. This shifts motivation from prediction and mental simulation to practice as we transform desires and predictive models into action. This implies actions must have worked in the past or been vicariously learned such that the hoped-for outcome was not brought about unintentionally (Shepherd, 2017).

Take, for instance, a case falling outside the usual explanatory scope of standard action theory in sociology: "endogenously rewarding" but seemingly non-goal-directed activity such as doing crossword puzzles—or what Cohen (2015) refers to as "engrossing" solitary actions. The reward here, essential to explaining the underlying motivation of solitary activities, comes from the challenge of thinking through a clue instead of finishing the puzzle and not so much the end-game pleasure. Solitary action is sociologically interesting because it routinizes novelty, suggesting a performative context where the expected structures the unexpected—like reading romance novels. In turn, structured surprises activate value-coding dopaminergic neurons associated with "the availability of rewards, evaluating outcomes, and learning" (Di Domenico & Ryan, 2017). The more expected the unexpected, the more engrossing the behavior, and the more likely a second set of salience-coding neurons to provide a further reward. As we shall see, several reward systems institutionalize specific goals, flexibly repeatable behaviors, and expected unexpectedness.

3.5 | Learning, Reinforcing, and Updating

By the time the third phase, satiation, occurs, the pleasure produced by the opioid reward system is already in decline. In this satiated state, the final phase of motivated action, that of strong (reinforcement) learning, occurs, setting up a critical step in the genesis of motivated (future) action taking the form of (a nascent) habit. Note that the (implicit and explicit) "learning" or "typing" the object-reward link that happens at this stage is crucial for conceptualizing the given object, psychological state, or social behavior directed at us as a reward on future occasions.

Neurophysiologically, learning is dissociable from wanting/liking and subserved by distinct midbrain and cortical structures, including the thalamus, the basal ganglia, and the striatum. These are areas that are also associated with skill acquisition and habit formation more generally, speaking to an intimate association, contra the idea of habitual action as "unmotivated" (Giddens, 1979, p. 218), between habitual and motivated action (Yin, 2008). Consequently, the dopaminergic system leading to wanting is closely tied to learning as repeated pursuit and exposure to objects and events allow

people to form mental representations encoding their value salience (Di Domenico & Ryan, 2017), thereby updating or reinforcing existing schema people used during action planning. In this way, past reward-related information helps establish the "anticipatory motivational value of objects and events" (Reeve & Lee, 2019, p. 367), suggesting that prediction lies at the heart of motivation (Miller Tate, 2021), and learning is not confined to a single phase of the motivated action process.

Actions enacted in either phase may become salient and worth repeating due to affective and cognitive experiences during satiation. This is especially the case when the outcomes are more or less rewarding than predicted. As Daw and Shohamy (2008; see also Schroeder, 2004) explain, the dopaminergic reward system in the midbrain (mainly the striatum) is keyed to the receipt of reward when these have a low probability of occurrence or the lack of reward when otherwise expected. In both cases, we see dopamine-releasing neurons in the striatum fire above baseline, whereas the correct prediction of expected rewards produces no activity above baseline. It seems, therefore, that the dopaminergic system implements a reward-prediction error signal (Schultz, 2016), which can then be used to update stored models of the environment in the case rewards are received when they were not expected or fail to come by when they were expected (Miller Tate, 2021). The reward system can thus be seen as implementing a biologically plausible form of "reinforcement learning," allowing people to adapt their behavior to the dynamic reward structure of the environment (Daw & Shohamy, 2008).

Consequently, repeated experiences of liking after attaining the goal object or state are stored in memory and thus drive future episodes of wanting. The recurrent behavior of "binging" (food, TV shows, alcohol) provides an instructive example. Here, the intrinsic rewards felt when (for instance) watching television are ingrained via early and repeated exposure. Binging loops of all three phases, with the appetitive phase becoming highly salient after the reward object (like the series) is exhausted. After consummation, we anticipate the next "big" show, searching for content and, often, willingly binging a TV show that may be less intrinsically rewarding in storytelling or acting. Like the crossword puzzle, the anticipation of the unexpected within a relatively structured activity is rewarding, making habitual motivated action during consummation much more salient and valued, making us crave additional content.

3.6 | Memory and Motivation

At its root, motivation relies on the cognitive capacity for spatial, autobiographical, nondeclarative, and other forms of memory (Stephens, 2019) and on the overall neurocognitive principle of minimizing "over the long term…prediction error signals" (Miller Tate, 2021, p. 18), such that predicted rewards turn into actual rewards received in most cases. As we will see, roles are crucial in establishing predictable links between action and rewards, especially when embedded in institutionalized reward systems. During the learning phase, we internalize both declarative (know-that, know-what) knowledge about the world (of the conceptual and person/situation-specific kinds) and procedural (know-how) knowledge of the way goals, actions, situations, and outcomes fit together (Chi & Ohlsson, 2005; Felix & Stephens, 2020). Some of this knowledge is specific, often related to tangible rewards (refrigerators have food), and some very general, abstract, and flexible behaviors usually pertain to conscious goal-seeking (food is necessary for survival, but only some foods are religiously or morally sanctioned). Based on these models, we generate predictions about the incentive value and reward-salience of objects around us. Predictive models of the environment relevant to motivation are built up from the interconnection of the three energizing mechanisms facilitating motivated action: Initiation, guidance, and control, and the association between wanting and liking built up from

previous experiences. The more tightly coupled and taken for granted the model, the more motivated action is expressed in habitual behavior, no less intentioned or controlled, just less deliberate.

What happens when predicted rewards do not materialize? When learned models of the reward structure of the environment begin to systematically fail us, we have two options. First, via passive inference, we can adjust the model to reduce the amount of error in the future (Pezzulo et al., 2015). Usually, actors are most likely to revise their world models when there are relatively concrete connections between goals, rewards, and situations. This is at the root of flexible repeatability: if action "A" did not work, I may try a variant or attempt action "Z" that I saw a friend use. It is also the root of addiction as people unintentionally reset their "set-point" (Kringelbach & Berridge, 2016): If eating a certain amount does not trigger the same level of reward (due to habituation), one may eat more, or they may seek other goal-objects to substitute or supplement eating. A second solution to prediction error, typically generated when abstract models—or those making predictions that would be most difficult to revise (as in highly institutionalized contexts)—are discrepant, is *active inference*, which involves efforts to alter the sensory signal to align it with what the internalized model predicts. In this case, role-based action becomes a self-fulfilling prophecy, where people predict the rewards their role-based action brings about via routinized mechanisms.

4 | IMPLICATIONS OF THE FUNDAMENTAL MOTIVATIONAL MODEL FOR ROLE THEORY

4.1 | Structured Roles: Institutions as Reward Systems

Given our discussion, one core idea from functionalism should be recovered and put to explanatory work in role theory: Institutions as *socially regulated reward systems* (Crane, 1976; Merton, 1968, 1973; Parsons, 1951b). This powerful suggestion deserves reconsideration given the neuroscientific sea change in our understanding of reward-based motivation already considered. Moreover, a refurbished conception of institutions as reward systems helps solve core analytical problems in institutional theory, which takes a largely cognitivist and phenomenological view of actors and action (Jepperson & Meyer, 2021). This approach focuses on discourses, vocabularies, and other external symbolic elements presumably "enacted" by people at the expense of good old-fashioned motivation (Tolbert & Zucker, 1996).

Merton most fully developed the implications of the idea of institutions as reward systems in the case of science (see also Crane, 1976 for a generalization to other domains). Merton (1973) allowed for the possibility that, under the right conditions, science develops into a self-sustaining, relatively autonomous system productive of sustained motivated action driven by the interplay of intrinsic and extrinsic rewards: "[t]he recognition accorded [to] scientific achievement by the scientist's peers is a reward in the strict sense identified by Parsons" (1968, p. 57, emphasis added). Accordingly, the "socially based psychic income of scientists takes the form of *pellets* of peer recognition that aggregate into reputational wealth" (1988, p. 620). For Merton, the "normative structure" of science emphasizes such informally policed norms as the "communism of ideas," the apportioning of credit via citation, the "naming" (or what Merton (1988, p. 620) called "eponymy") of entire eras, laws, theorems, effects, and the like after the scientists who discover them (Merton, 1957), and the development of an informal reputational system. The normative structure exists because it facilitates the distribution of rewards, thus further "hooking" scientists into the system.

Giving credit for ideas, settling on a given scientist as the "first" to have discovered something, or naming revolutionary effects or techniques after the scientist, all count as intrinsic role-based social

rewards for the scientist (they lead to acclaim, praise, and recognition). These may also lead to other extrinsic rewards (assignment to influential academic, editorial, or administrative posts, grant funding), leading to further psychological and social rewards. The theory of action and motivation underpinning Mertonian sociology of science is consistent with the synthetic approach to reward-based motivated action developed here. Science is *literally* rewarding in the motivational neuroscience sense (Schroeder, 2004, p. 48ff) across multiple reward types. This explains why science is one of the few institutions that can elicit copious amounts of work, commitment, and lifelong devotion from people using purely informal (contra direct domination or coercion) mechanisms. Because scientists (especially those most rewarded) find science rewarding, they are also the most motivated to continuously engage in the patterns of activity that sustain it. The reward system of science veritably "hooks" scientists into the institution via role performance.

Of course, other science-like institutional spheres reveal similar reward systems (see also Crane, 1976). For instance, though we often think of law in terms of courtroom performances, the evolution of Western law was rooted in the development of an intellectual community that endogenously meted out social and psychological rewards. This first happened in Bologna in the 12th century CE and then in other early universities like Salerno and Paris (Berman, 1983). Like science, law begets law because, once an autonomous reward system with norms and rules emerges, the reward system canalizes participants toward specific patterned intellectual pursuits and practices. These include (among others) interpreting, glossing, and amending legal texts (Brundage, 2008). Going further back in time, a significant portion of activities by religious elites, dating back to the earliest periods of writing (Goody, 1986) and onwards to the theological traditions encouraged by the Catholic Church in universities, monasteries, and the like, was predicated on the development of endogenous reward systems among intellectual specialists, rewarding those devoting their time knowledge accumulation and the pursuit of specifically religious intellectual activities (Southern, 1970). More generally, intellectuals organize different reward systems in religious, scientific, artistic, legal, and even medical institutions (Abrutyn, 2013), and the same for artists and cultural producers after the emergence of markets for symbolic goods (Bourdieu, 1993; Crane, 1976). The emergence of "fields," or "ecologies" (Abbott, 1969, 2005; Bourdieu, 1966/1969), settle into periodically stable arrangements dispensing multiple rewards keyed to role performance, keeping players motivated to play the game.

Post-Mertonian debate in the sociology of science (and later science and technology studies) centered on Merton's residual normativism. While much criticism was undoubtedly merited, critics took backward steps in theorizing the link between motivation and institutions. As Bunge (1991, p. 525) notes, for all their bluster, post-Mertonian critics "never tell us what makes scientists tick qua scientists," which was one of the main concerns of Merton's reward systems formulation. One reason for this gap is that most post-Mertonian debate has been driven by the emergence of theories that posit *mono-motivational* action explanations ignoring the linkage between *multiple* rewards in scientific fields. These include approaches like Bourdieu's (1966/1969, Bourdieu and Collier, 1984/1988), making recognition from others (a type of social reward) a master motivation (Peters, 2012), ignoring the complete set of reward cocktails motivating scientists to keep at it systematically dispensed by scientific institutions (an issue also besetting Collins's (2000) account).

We do not advocate a return to Mertonian soft-functionalism or a normativist picture of motivated action in science or any other institution. The problem is that norm-driven approaches conceive deviations from the normative ideals (like meritocracy), such as the cumulative advantage and disproportionate influence and attention wielded and received by the most successful scientists, as problems that need to be papered over. This is unsatisfactory because the "deviations" are the most pervasive empirical features of science and other institutions structured as reward systems. Instead, a successful theory should be able to account for normative and counter-normative behavior, as well as both the

normative and empirical features of institutions, using similar motivational mechanisms at the activity level and its intended and unintended consequences.

4.2 | Institutionalized Roles

If institutions are reward systems, it makes sense to conceptualize roles as *mechanisms for routinizing the pursuit and delivery of rewards and goal objects*, linking roles to a defensible account of reward-driven motivation (Schroeder, 2004). Roles serve primarily as the (predictable) *vehicles* via which individuals traverse the well-worn roads offered by institutions towards various "reward cocktails" they dispense. Just how routinized and predictable reward dispensation is, depends on many factors. These include role formalization (Turner, 2010b), institutional autonomy (Abrutyn, 2013), and role alternatives to similar reward sets.

The most totalizing reward systems, referred to by Coser (1974) as *greedy institutions* and Goffman (1968) as *total institutions*, usually reveal motivational processes most clearly, structured as they are by undivided social relationships or the instrumental, bureaucratic management of "batches" of people respectively. Coser's and Goffman's structural accounts of totalizing role commitment obscure the tight coupling of both formal and informal reward systems and the structural relations induced by undivided loyalty or imposed by singular authoritative fiat. Overemphasizing totalizing structure hides essential inter-individual variation in the coupling of institutionalized roles and rewards. For instance, why some eunuchs or cult members may be more or less committed, dominated, and willingly give everything to the leaders, groups, and communities to which they are attached depends on the number, volume, and frequency of rewards received. Greedy and total institutions couple rewards, objects, behaviors, and swift, severe sanctions in lockstep; in a literal and not metaphorical sense, they are the sociological equivalent of Skinner boxes.

Accordingly, we can account for the rare empirical occurrence of totally "oversocialized" institutions via the difficulty of establishing such a tight linkage of roles and rewards. In the greedy/total case, centralized reward systems can generate lines of motivated action leading people to act in a way benefiting the group even if it leads to self-harm. Greedy/total institutions thus make people utterly dependent on them to receive *all* possible rewards, locking them to their assigned roles. Coser/Goffman institutional dynamics reflect the consequences of purposefully stripping role relationships beyond those demanded by the group, substituting all exogenous reward systems with endogenous (institution-generated) ones. The result is people willing to do just about anything for the group.

This analysis provides a template for theorizing how people can develop emotional and moral attachments to roles even when considering less totalizing relationships, explaining how informal groups or organizations produce commitment and how much collective power they can harness, especially in large, heterogeneous, impersonal, or *de*personalizing arrangements, where roles are crucial to coordinating and controlling people as productive or generalized exchange systems increasingly replace direct/negotiated relational rewards (Lawler, 2001). These systems, like science, are built as the primary means via which smaller units (professional associations or faculties) induce attachments forged via reliable, consistent, and desired access to psychological and social rewards, as well as some tangible rewards (Simpson et al., 2018). For instance, before autonomous scientific institutions consolidated, tangible rewards—particularly patronage—were essential because they gave proto-scientists like Galileo autonomy from the Church's authority and the ability to work (Gaukroger, 2008). Nevertheless, all signs point to the earliest scientists, like early legal or religious experts, being driven by a mixture of intrinsic psychological and social rewards. In this respect, the origins of the institutional reward system of science are illuminating, as it owes its foundation and expansion to a class of

people "who believed in science as an intrinsically valuable preoccupation and who had a reasonable prospect of making their belief generally accepted, even before science proved its economic worth" (Ben-David, 1965, p. 15).

Like a violinist committed to their craft despite meager economic returns, the construction of standardized roles often begins as a labor of love predicated on the cultivation and monopolization of intrinsic rewards first and extrinsic rewards after. Indeed, the continuous search for truth echoed in the practical rules of the scientific method remains at the core of science even with the rise of skepticism about objectivity and the need for reflexive science (Huff, 2017). One need only consider the cautionary tales heaping scorn on scientists seen as primarily motivated by the pursuit of extrinsic rewards (like grant money or prestigious positions) or the suspicion directed at intellectuals oriented towards exogenous social rewards, not under the direct control of the scientific institution (like success as a popular author or accumulating millions of "followers" on social media). The fundamental distinction between the "autonomy" or "heteronomy" of a field is thus dependent on whether actors are primarily oriented toward the institutionally-controlled reward system or to some externally available alternative (Bourdieu, 1993; Crane, 1976).

The link between reward systems and how everyday work is produced is also pivotal to how institutions succeed (or fail) at "making" people committed to their perpetuation (Abrutyn, 2013). Regarding science, the notion of a scientific community, whether defined as a university or other association of scientists devoted to research, a subunit like a college or a faculty, or a discipline and some sort of professional association are collectives built around the purposeful creation of collective goods via productive, direct, or indirect exchange (Lawler, 2001). These exchange systems are ideal for producing affective attachment to the group or institution, generating commitment via a regular, consistent, and predictable cocktail of tangible, psychological, and social rewards (Vanzella-Yang & Abrutyn, 2021). On the one hand, scientists work towards collaborative and solo projects while pursuing knowledge production and dissemination. On the other hand, experienced rewards depend on the system, given an array of personal, impersonal, depersonalized, or indirect exchange relationships, a salient point when socializing incoming scientists.

4.3 | Rethinking Role Exit

More evidence of the relationship between roles and rewards comes from the less-studied process of *role exit*. This dynamic, like commitment, is best observed when people try to break free of total/greedy institutions. As Goffman notes, the longer one's stay in an asylum or prison, the more difficult it is to shake the role "residuals," partly because total commitments dispossess inmates of all other roles (Goffman, 1968, p. 14ff)—and thus pathways to rewards. Goffman's account deals with ritual admission processes designed to break old cultural patterns while ensuing regimentation and tyranny of movement and body—a process also emphasized in Foucault (1975/1995). In our terms, the structural and performative dimensions impose a tight coupling of rewards and the goals of the organization's authority system. Obeying rules means not being punished; the dangling of tangible, psychological, and social rewards (and the exclusion of alternatives) are designed to generate compliance, while informal systems of status and reward are tolerated insofar as they remain embedded and subordinate to the official system. In the total institutional context, members are exposed to a near-continuous Skinnerian box of chains, lights, and levers for reward.

Unsurprisingly, Helen Ebaugh's (1988) account of ex-nuns, transgender individuals, and divorcees draws on a physiological metaphor associated with alcohol abuse to describe the "withdrawal" those attempting to exit long-standing role positions experience. Labeling the role residuals *identity*

hangovers.[2] Ebaugh talks of nostalgia and structural continuities/discontinuities, but surprisingly does not double down on the addiction language she only metaphorically employs (but see Simi et al., 2017). A different way to think of a role hangover is like the plight of an alcoholic endeavoring to become sober. On the one hand, a chemical dependency is a difficult itch to not scratch, which given our theory of motivation, should seem no different from someone playing a role and being routinely, consistently, and adequately rewarded for some indefinite period. Re-wiring the neural circuits creating our sense of self in autobiographical memory and affective signals is not impossible, but it is not as simple as discursively proclaiming a new identity. On the other hand, being forced (like being released from prison), naturally progressing (like aging out of a career), or opting (giving up alcohol for good) out of a routine reward system means not having access to routinized goal objects—many of which are real things, like people or substances—as well as the routines cueing anticipatory rewards. Withdrawal, hangover and relapse are to be expected. Because some role exits may preclude the latter (since most divorces are final) under normal circumstances, we may expect pathological behaviors to trigger rewards by other means.

Thus, it is understandable why an officer in the armed forces or police for most of their adult life may struggle with transitioning to civilian life (Edelmann, 2018); and, worse, why dishonorable discharge or being punished with a desk assignment, respectively, may result in a particularly painful reorientation toward a different set of rewards. There are abundant instances of this process, sometimes described in the literature under the banner of "hysteresis" (Strand & Lizardo, 2017). For instance, athletes are non-totalistic roles deeply entrenched in a particular reward system and subject to "aging" out of their profession regardless of their desire to continue to enact it. It is cliché, but no less accurate, to point to the professional veteran whose skills suddenly elude them or the young athlete who suffers a debilitating injury. In either case, retirement is akin to denying a system of rewards that is impossible to replicate elsewhere. This is similar to Simi et al.'s (2017) white supremacists attempting to shake off the supremacist identity and occasionally "relapsing" into racist outbursts. Ultimately, a general principle derived from the reward-based model of motivation might state that the more a role supplies a significant number, volume, and frequent dispensation of a cocktail of rewards, the more attached the incumbent will be to the role, committed to the institution, and predictable in their behavior. Conversely, the more rewards the role brings, the more painful the exit, especially if forced or unplanned.

4.4 | Roles, Institutions, and Rewards Moving Forward

As we saw in the preceding, the critical problem with functionalist role theory was to make "norms" and "normatively sanctioned" action the cause and the effect of the role structures. How does our reward-based account of role performance improve on this? The first thing to do (as Merton began to do) is to decouple normative (or role-consistent) action from norms as "motivators" of action—see Turner (2010b, pp. 148–150) for alternative ways forward on this. As we noted, people are motivated by *rewards,* not norms. Just because a pattern of action seems to conform to a norm does not mean people are (intrinsically) motivated to sustain the pattern by the internalized norm. Norms are conduits for organizing the reward system but are not motivational in themselves (absent rewards). From the perspective of the contemporary understanding of motivated action, the internalized norm itself lacks causal powers, and it cannot be seen as a direct motivator for action. Instead, an individual could be motivated by the pursuit of several rewards, and norm-following just falls off as a by-product. Norms matter because they serve as *regulators* of the reward system (not of people's actions). Norm-conforming behavior cannot be an explanans or an explanandum simultaneously.

Instead, norm-following and norm-violating motivated action is always an explanandum, and the motivation for following norms must be explained by specifying the rewards people pursue independently of those norms.

For explanatory "symmetry," actions conforming and violating norms are motivated, which means they need to be explained by the same reward-based mechanisms. For instance, take a case of norm-violating action in an institutional field such as science such as Diederik Stapel, a charismatic Dutch psychologist who rose to the top ranks of the field in terms of reputation, influence, and power. Stapel fabricated the data for almost two decades in dozens of studies, including those published in the world's top general science journals (Borsboom & Wagenmakers, 2012), earning more than fifty retractions. This is an extreme case of someone who did not internalize the Mertonian "norms" of the scientific field and whose actions are not motivated by internalized psychological rewards fitting the norms of science (like the dispassionate, self-critical search for truth). Stapel became addicted (in the contemporary motivational science sense) to the role-based rewards that a history of fraudulent actions regularly provided. For instance, in his memoir, Stapel (2012) noted he was driven by "[t]he need to score, ambition, laziness, nihilism, want of power, status anxiety, desire for solutions, unity, pressure to publish, arrogance, emotional detachment, loneliness, disappointment, ADD, addiction to answers" (Stapel, 2012, p. 226).[3]

As such, Stapel was not striving *exclusively* for the usual rewards emphasized by analysts like Bourdieu and Collins (prestige, recognition, vanquishing intellectual opponents in favor of their intellectual in-group), but also a panoply of extrinsic tangible rewards that come with those and the further ratcheting of the social rewards that comes after a person has accumulated all the extrinsic markers and symbols of scientific success and institutional power in their role as a scientist. Once "hooked" to this multiple reward cocktail, Stapel could not stop, even though his behavior was clearly counter-normative. The very Mertonian "normative" system of science (weak disciplinary control, the intellectual "honor system" and the like) *facilitated* his getting away with the fraud for so long. But, and this is the key point, the underlying mechanisms (like being addicted to scientific work and garnering intrinsic and extrinsic rewards from it) are not qualitatively dissimilar from those that sustain motivated action in legitimate scientific work given the same normative system.

The model of reward-driven motivation for role performance offered here can also help make sense of phenomena that were puzzling from the norm-centered conception of science. For example, take the case of institutionalized "over-reward." In Mertonian sociology of science, the most prominent example of such a phenomenon is what Merton came to baptize as the "Matthew Effect." This is the idea that in the realm of science, those that have been most successful in accumulating the most social, tangible, and extrinsic rewards (the most visible signals of recognition, eminence, peer esteem, and reputation, such as the Nobel Prize) have an easier time accumulating those rewards (in terms of receiving recognition and credit for future work and discoveries) in the future, compared to those who have not been rewarded in the past. However, the phenomenon was always an embarrassment to the Mertonian normative model because it could only be conceptualized as a case of "over-reward," thus a distorting force in both the credit-apportionment and proper "communication system" that science should have. In the last case, for instance, the fact that the discoveries and work of the most highly cited people draw "disproportionate" attention from equally (or probably more) worthy work done by less prominent others means that science as an institution for producing and disseminating the "best" knowledge is not as efficient as it should be (Merton, 1968, 1988). Merton, of course, was not shy to postulate several possible "latent functions" that such institutionalized endogenous inequality-generating systems might have.

One possibility that is not usually considered, however, is that such institutionalized "over-reward," may be necessary to sustain motivated actions in fields of organized striving over the long run. Here,

the work in motivational neuroscience can help (Daw & Shohamy, 2008). After receiving a consistent reward history, dopaminergic brain systems learn to *predict* future reward receipt. Accordingly, dopaminergic neurons respond not to receiving rewards when expected, but only to "over" or "under" rewards. This matter for explaining sustained motivation on the part of the most successful players in the context of organized fields of striving. Here, it is precisely a mechanism such as the one described by the Matthew Effect, where the rewards received by the most successful exceed what could be expected given past performance, that will "register" as a reward for a scientist who has had a long history of previous reward receipt. Instead, what counts as an "over-reward" from a normative perspective is motivationally registered as the only reward that can sustain future motivated action in the field. A test implication of this idea is that fields that attempt to restrict mechanisms of cumulative advantage via more democratic reward distribution systems may have the unintended effect of demotivating their most successful members.

5 | CONCLUDING REMARKS

In this paper, we have attempted to rehabilitate the notion of role by linking sociological role theory to recent work on motivational, affective, and cognitive neuroscience specifying the internal mechanisms behind motivated action. We argued that there is nothing inherently problematic or retrogressive in the idea of "role," once its link to a purely normative account of motivated action is severed. In particular, these accounts fail to provide a sociologically plausible account of why people are "impelled" to pursue continuous patterns of role-based action. To rebuild such an account, we brought analytic clarity to the concept of motive, motivation, and motivated action. Regarding the first, we clarified the idea of a motive as an internal cognitive and affective state causally involved in initiating, continuing, and guiding action.

We then engaged in an analytic reconstruction of the motivated action cycle, centered on the notion of reward. Reward disappeared from sociological view but has reemerged with a vengeance in contemporary affective and motivational neuroscience (Kim et al., 2016; Kringelbach & Berridge, 2016; Schroeder, 2004). Classical action-theoretic concepts like value, desire, and want are now conceptualized by motivational science in reward-based terms. Neuroscientists have mapped the structure and function of this reward system, linking it to the various phases of motivated action. Following this work, we distinguished three phases of motivated action: Wanting (striving or desiring), consummation (using and liking), and satiation (reinforcing or updating environmental models). The distinction between the wanting and liking phases is our biggest take home for a sociological theory of motivation: The neurophysiological dissociability between wanting (mobilizing, energizing, anticipating) an object, liking (touching, manipulating consuming) the object, and updating associations between rewards, objects, and actions. A key implication of our argument is that the best way to link role, action, and structures is by reviving the idea of institutions as *literal* reward systems. Consequently, if institutions are reward systems, it makes more sense to conceptualize roles as the mechanisms via which the pursuit and delivery of rewards and goal objects are routinized, with reward-dispensation depending on many field-level factors, like role formalization, institutional autonomy, and role availability.

ORCID

Seth Abrutyn https://orcid.org/0000-0001-6220-3362

Omar Lizardo https://orcid.org/0000-0002-5405-3007

ENDNOTES

[1] Another theorist who has moved far beyond Parsons' limited notion of roles is Jonathan Turner (2010b). We do not leverage his theoretical insights here, as he draws (perhaps too strong) a distinction between motivational and role dynamics. Nonetheless, Turner offers a fascinating alternative to the approach to roles developed here emphasizing their capacity to organize and facilitate social interaction in personal and, especially, impersonal exchanges. In particular, he draws from Ralph Turner's (2001) ideas about role-making and Mead's role-taking to underscore the creative and institutionalized facets of roles. From there, Turner conceptualizes roles vehicles via which individuals connect to culture, norms, social structure, and material ecology. It is in the interstices of these other more macro elements that role-taking and making is experienced and understood. Turner also suggests modern sociology's emphasis on status has trouble explaining how interactional breakdowns are rarely due to status dynamics but actually because of role-playing difficulties—a point first made by Goffman (1967).

[2] It is important to underscore that Ebaugh's discussion of transgender folks is not likely reflective of their experience today (see Schilt, 2011). When she was writing, being gay was extremely sanctioned legally and culturally and thus being trans was not even something imaginable. Consequently, the forces a trans person would face in both transitioning to a different gender (especially when gender was firmly conceived as binary) in a highly hostile time and place provides some context for what might be intensely criticized through gender and sexuality studies today.

[3] Quoted in Borsboom and Wagenmakers (2012, p. 31).

REFERENCES

Abbott, A. (2005). Linked ecologies: States and universities as environments for professions. *Sociological Theory*, *23*(3), 245–274. https://doi.org/10.1111/j.0735-2751.2005.00253.x

Abrutyn, S. (2013). *Revisiting Institutionalism in Sociology: Putting the "Institution" Back in Institutional Analysis*. Routledge.

Aron, A., Fisher, H., Mashek, D. J., Strong, G., Li, H., & Brown, L. L. (2005). Reward, motivation, and emotion systems associated with early-stage intense romantic love. *Journal of Neurophysiology*, *94*(1), 327–337. https://doi.org/10.1152/jn.00838.2004

Ben-David, J. (1965). The scientific role: The conditions of its establishment in Europe. *Minerva*, *4*(1), 15–54. https://doi.org/10.1007/bf01585983

Berger, P. L., & Luckmann, T. (1966). *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. Doubleday.

Berman, H. J. (1983). *Law and Revolution: The Formation of the Western Legal Tradition*. Harvard University Press.

Biddle, B. J. (1986). Recent developments in role theory. *Annual Review of Sociology*, *12*(1), 67–92. https://doi.org/10.1146/annurev.so.12.080186.000435

Borsboom, D., & Wagenmakers, E.-J. (2012). Derailed: The rise and fall of Diederik Stapel. *APS Observer*, *26*(1).

Bourdieu, P. (1969). Intellectual field and creative project (S. France, trans.). *Social Science Information*, *8*(2), 89–119. (Original work published 1966). https://doi.org/10.1177/053901846900800205

Bourdieu, P. (1988). *Homo Academicus*. (P. Collier, trans.). Stanford University Press. (Original work published 1984).

Bourdieu, P. (1993). *The Field of Cultural Production: Essays on Art and Literature*. Columbia University Press.

Breiter, H. C., Aharon, I., Kahneman, D., Dale, A., & Shizgal, P. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron*, *30*(2), 619–639. https://doi.org/10.1016/s0896-6273(01)00303-8

Brundage, J. A. (2008). *The Medieval Origins of the Legal Profession: Canonists, Civilians, and Courts*. University of Chicago Press.

Bunge, M. (1991). A Critical Examination of the New Sociology of Science Part 1. *Philosophy of the Social Sciences*, *21*(4), 524–560. https://doi.org/10.1177/004839319102100406

Burke, P. J., & Stets, J. E. (2009). *Identity theory*. Oxford University Press.

Campbell, C. (1996). On the concept of motive in sociology. *Sociology*, *30*(1), 101–114. https://doi.org/10.1177/0038038596030001007

Chi, M. T. H., & Ohlsson, S. (2005). Complex Declarative Learning. In K. J. Holyoak (Ed.), *The Cambridge handbook of thinking and reasoning* (Vol. 858, pp. 371–399). Cambridge University Press.

Cohen, I. J. (2015). *Solitary Action: Acting on Our Own in Everyday Life*. Oxford University Press.

Collins, R. (2000). The Sociology of Philosophies: A Précis. *Philosophy of the Social Sciences*, *30*(2), 157–201. https://doi.org/10.1177/004839310003000201

Connell, R. W. (1979). The Concept of Role and What to Do With It. *The Australian and New Zealand Journal of Sociology*, *15*(3), 7–17. https://doi.org/10.1177/144078337901500302

Coser, L. A. (1974). *Greedy Institutions; Patterns of Undivided Commitment*. Free Press.

Crane, D. (1976). Reward Systems in Art, Science, and Religion. *American Behavioral Scientist*, *19*(6), 719–734. https://doi.org/10.1177/000276427601900604

Davis, K. L., & Montag, C. (2018). Selected Principles of Pankseppian Affective Neuroscience. *Frontiers in Neuroscience*, *12*, 1025. https://doi.org/10.3389/fnins.2018.01025

Daw, N. D., & Shohamy, D. (2008). The Cognitive Neuroscience of Motivation and Learning. *Social Cognition*, *26*(5), 593–620. https://doi.org/10.1521/soco.2008.26.5.593

Decety, J., & Yoder, K. J. (2017). The Emerging Social Neuroscience of Justice Motivation. *Trends in Cognitive Sciences*, *21*(1), 6–14. https://doi.org/10.1016/j.tics.2016.10.008

Di Domenico, S. I., & Ryan, R. M. (2017). The Emerging Neuroscience of Intrinsic Motivation: A New Frontier in Self-Determination Research. *Frontiers in Human Neuroscience*, *11*, 145. https://doi.org/10.3389/fnhum.2017.00145

DiMaggio, P. J., & Powell, W. W. (1983). The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. *American Sociological Review*, *48*(2), 147–160. https://doi.org/10.2307/2095101

Durkheim, É. (1933). *The Division of Labor in Society*. The Free Press.

Ebaugh, H. R. F. (1988). *Becoming an Ex: The Process of Role Exit*. University of Chicago Press.

Edelmann, A. (2018). Culturally meaningful networks: on the transition from military to civilian life in the United Kingdom. *Theory and Society*, *47*(3), 327–380. https://doi.org/10.1007/s11186-018-9317-7

Felix, C. V., & Stephens, A. (2020). A Naturalistic Perspective on Knowledge How: Grasping Truths in a Practical Way. *Philosophies*, *5*(1), 5. https://doi.org/10.3390/philosophies5010005

Fine, G. A. (2012). Group Culture and the Interaction Order: Local Sociology on the Meso-Level. *Annual Review of Sociology*, *38*(1), 159–179. https://doi.org/10.1146/annurev-soc-071811-145518

Foucault, M. (1995). *Discipline and Punish: The Birth of the Prison*. Vintage Books. (Original work published 1975).

Freidson, E. (1970). *Profession of Medicine: A Study of the Sociology of Applied Knowledge*. Dodd, Mead.

Gaukroger, S. (2008). *The Emergence of a Scientific Culture: Science and the Shaping of Modernity 1210-1685*. Clarendon Press.

Giddens, A. (1979). *Central Problems in Social Theory: Action, Structure, and Contradiction in Social Analysis*. University of California Press.

Goffman, E. (1967). *Interaction Ritual: Essays on Face-to-Face Behavior*. Pantheon Books.

Goffman, E. (1968). *Asylums: Essays on the Social Situation of Mental Patients and Other Inmates*. Aldine Transaction.

Goody, J. (1986). *The Logic of Writing and the Organization of Society*. Cambridge University Press.

Guhin, J., Calarco, J. M., & Miller-Idriss, C. (2021). Whatever Happened to Socialization? *Annual Review of Sociology*, *47*(1), 109–129. https://doi.org/10.1146/annurev-soc-090320-103012

Hahl, O., & Zuckerman, E. W. (2014). The denigration of heroes? How the status attainment process shapes attributions of considerateness and authenticity. *AJS; American Journal of Sociology*, *120*(2), 504–554. https://doi.org/10.1086/678304

Haslanger, S. (2019). Cognition as a Social Skill. *Australasian Philosophical Review*, *3*(1), 5–25. https://doi.org/10.1080/24740500.2019.1705229

Hitlin, S., & Piliavin, J. A. (2004). Values: Reviving a Dormant Concept. *Annual Review of Sociology*, *30*(1), 359–393. https://doi.org/10.1146/annurev.soc.30.012703.110640

Huff, T. E. (2017). *The Rise of Early Modern Science: Islam, China, and the West*. Cambridge University Press.

Jackson, J. (1998). Contemporary criticisms of role theory. *Journal of Occupational Science*, *5*(2), 49–55. https://doi.org/10.1080/14427591.1998.9686433

Jepperson, R. L., & Meyer, J. W. (2021). *Institutional Theory: The Cultural Construction of Organizations, States, and Identities*. Cambridge University Press.

Kim, S.-I., Reeve, J., & Bong, M. (2016).Introduction to Motivational Neuroscience. In *Recent Developments in Neuroscience Research on Human Motivation* (Vol. 39, p. 88). Emerald Group Publishing Limited.

Kringelbach, M. L., & Berridge, K. C. (2016). Neuroscience of Reward, Motivation, and Drive. In *Recent Developments in Neuroscience Research on Human Motivation* (Vol. 19, pp. 23–35). Emerald Group Publishing Limited.

Lawler, E. J. (2001). An Affect Theory of Social Exchange. *American Journal of Sociology*, *107*(2), 321–352. https://doi.org/10.1086/324071

Linton, R. (1936). *The study of man*. Appleton-Century.

Lizardo, O., & Stoltz, D. S. (2018). Max Weber's ideal versus material interest distinction revisited. *European Journal of Social Theory*, *21*(1), 3–21. https://doi.org/10.1177/1368431017710906

Martin, J. L. (2009). *Social Structures*. Princeton University Press.

Martin, J. L., & Lembo, A. (2020). On the Other Side of Values. *American Journal of Sociology*, *126*(1), 52–98. https://doi.org/10.1086/709778

Merton, R. K. (1957). Priorities in Scientific Discovery: A Chapter in the Sociology of Science. *American Sociological Review*, *22*(6), 635–659. https://doi.org/10.2307/2089193

Merton, R. K. (1968). The Matthew effect in science. The reward and communication systems of science are considered. *Science*, *159*(3810), 56–63. https://doi.org/10.1126/science.159.3810.56

Merton, R. K. (1973). *The Sociology of Science: Theoretical and Empirical Investigations*. University of Chicago Press.

Merton, R. K. (1988). The Matthew Effect in Science, II: Cumulative Advantage and the Symbolism of Intellectual Property. *Isis; an International Review Devoted to the History of Science and Its Cultural Influences*, *79*(4), 606–623. https://doi.org/10.1086/354848

Miller Tate, A. J. (2021). A predictive processing theory of motivation. *Synthese*, *198*(5), 4493–4521. https://doi.org/10.1007/s11229-019-02354-y

Panksepp, J., & Watt, D. (2011). Why does depression hurt? Ancestral primary-process separation-distress (PANIC/GRIEF) and diminished brain reward (SEEKING) processes in the genesis of depressive affect. *Psychiatry: Interpersonal and Biological Processes*, *74*(1), 5–13. https://doi.org/10.1521/psyc.2011.74.1.5

Parsons, T. (1942). Age and Sex in the Social Structure of the United States. *American Sociological Review*, *7*(5), 604–616. https://doi.org/10.2307/2085686

Parsons, T. (1951a). Illness and the role of the physician: a sociological perspective. *American Journal of Orthopsychiatry*, *21*(3), 452–460. https://doi.org/10.1111/j.1939-0025.1951.tb00003.x

Parsons, T. (1951b). *The Social System*. The Free Press.

Parsons, T. (1975). The sick role and the role of the physician reconsidered. *Milbank Memorial Fund Quarterly - Health & Society*, *53*(3), 257–278. https://doi.org/10.2307/3349493

Parsons, T., & Shils, E. A. (1951). *Toward a General Theory of Action* (pp. 47–275). Harvard University Press.

Peters, G. (2012). The social as heaven and hell: Pierre Bourdieu's philosophical anthropology: The social as heaven and hell. *Journal for the Theory of Social Behaviour*, *42*(1), 63–86. https://doi.org/10.1111/j.1468-5914.2011.00477.x

Pezzulo, G., Rigoli, F., & Friston, K. (2015). Active Inference, homeostatic regulation and adaptive behavioural control. *Progress in Neurobiology*, *134*, 17–35. https://doi.org/10.1016/j.pneurobio.2015.09.001

Reeve, J., & Lee, W. (2019). Motivational neuroscience. *The Oxford Handbook of Human Motivation*, 355–371.

Schilt, K. (2011). *Just One of the Guys?: Transgender Men and the Persistence of Gender Inequality*. University of Chicago Press.

Schroeder, T. (2004). *Three Faces of Desire*. Oxford University Press.

Schultz, W. (2016). Dopamine reward prediction-error signalling: a two-component response. *Nature Reviews Neuroscience*, *17*(3), 183–195. https://doi.org/10.1038/nrn.2015.26

Shepherd, J. (2017). Halfhearted Action and Control. *Ergo (Ann Arbor, Mich.)*, *4*(9). https://doi.org/10.3998/ergo.12405314.0004.009

Silver, D. (2011). The moodiness of action. *Sociological Theory*, *29*(3), 199–222. https://doi.org/10.1111/j.1467-9558.2011.01394.x

Simi, P., Blee, K., DeMichele, M., & Windisch, S. (2017). Addicted to Hate: Identity Residual among Former White Supremacists. *American Sociological Review*, *82*(6), 1167–1187. https://doi.org/10.1177/0003122417728719

Simpson, B., Harrell, A., Melamed, D., Heiserman, N., & Negraia, D. V. (2018). The Roots of Reciprocity: Gratitude and Reputation in Generalized Exchange Systems. *American Sociological Review*, *83*(1), 88–110. https://doi.org/10.1177/0003122417747290

Smith, K. S., & Berridge, K. C. (2007). Opioid limbic circuit for reward: interaction between hedonic hotspots of nucleus accumbens and ventral pallidum. *Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *27*(7), 1594–1605. https://doi.org/10.1523/jneurosci.4205-06.2007

Southern, R. W. (1970). *Mediaeval humanism and other studies*. Blackwell.

Stapel, D. A. (2012). *Ontsporing*. Prometheus.

Stephens, A. (2019). Three Levels of Naturalistic Knowledge. In M. Kaipainen, F. Zenker, A. Hautamäki, & P. Gärdenfors (Eds.), *Conceptual Spaces: Elaborations and Applications* (pp. 59–75). Springer International Publishing.

Stets, J. E., & Burke, P. J. (2000). Identity Theory and Social Identity Theory. *Social Psychology Quarterly*, *63*(3), 224. https://doi.org/10.2307/2695870

Stets, J. E., & Burke, P. J. (2014). The Development of Identity Theory. In *Advances in Group Processes* (Vol. 39, p. 88). Emerald Group Publishing Limited.

Stivers, T., & Timmermans, S. (2020). Medical Authority under Siege: How Clinicians Transform Patient Resistance into Acceptance. *Journal of Health and Social Behavior*, *61*(1), 60–78. https://doi.org/10.1177/0022146520902740

Strand, M., & Lizardo, O. (2017). The hysteresis effect: theorizing mismatch in action. *Journal for the Theory of Social Behaviour*, *47*(2), 164–194. https://doi.org/10.1111/jtsb.12117

Stryker, S. (1980). *Symbolic interactionism: A social structural version*. Benjamin-Cummings Publishing Company.

Tamir, D. I., & Hughes, B. L. (2018). Social Rewards: From Basic Social Building Blocks to Complex Social Behavior. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *13*(6), 700–717. https://doi.org/10.1177/1745691618776263

Tindell, A. J., Berridge, K. C., Zhang, J., Peciña, S., & Aldridge, J. W. (2005). Ventral pallidal neurons code incentive motivation: amplification by mesolimbic sensitization and amphetamine. *European Journal of Neuroscience*, *22*(10), 2617–2634. https://doi.org/10.1111/j.1460-9568.2005.04411.x

Tolbert, P., & Zucker, L. G. (1996). The Institutionalization of Institutional Theory. In S. Clegg, C. Hardy, & W. R. Nord (Eds.), *The Handbook of Organization Studies* (pp. 175–190). Sage.

Tomasello, M. (2019). The role of roles in uniquely human cognition and sociality. *Journal for the Theory of Social Behaviour*, *50*(1), 2–19. https://doi.org/10.1111/jtsb.12223

Turner, J. H. (2010a). Motivational Dynamics in Encounters. In J. H. Turner (Ed.), *Theoretical Principles of Sociology. Microdynamics* (Vol. 2, pp. 193–235). Springer.

Turner, J. H. (2010b). Role Dynamics in Encounters. In J. H. Turner (Ed.), *Theoretical Principles of Sociology. Microdynamics* (Vol. 2, pp. 133–162). Springer.

Turner, R. H. (1956). Role-Taking, Role Standpoint, and Reference-Group Behavior. *American Journal of Sociology*, *61*(4), 316–328. https://doi.org/10.1086/221763

Turner, R. H. (1976). The real self: from institution to impulse. *AJS; American Journal of Sociology*, *81*(5), 989–1016.

Turner, R. H. (1978). The role and the person. *American Journal of Sociology*, *84*(1), 1–23. https://doi.org/10.1086/226738

Turner, R. H. (2001). Role Theory. In J. H. Turner (Ed.), *Handbook of Sociological Theory* (pp. 233–254). Springer US.

Vanzella-Yang, A., & Abrutyn, S. (2021). Tastes, emotions, and social cohesion: Toward a cultural theory of social exchange. *Journal for the Theory of Social Behaviour*, *52*(2), 315–335. https://doi.org/10.1111/jtsb.12323

von Ihering, R. (1968). *Law as a Means to an End*. A. M. Kelley.

West, C., & Fenstermaker, S. (1995). Doing difference. *Gender & Society: Official Publication of Sociologists for Women in Society*, *9*(1), 8–37. https://doi.org/10.1177/089124395009001002

West, C., & Zimmerman, D. H. (1987). Doing Gender. *Gender & Society*, *1*(2), 125–151. https://doi.org/10.1177/0891243287001002002

Wiley, N. (1994). *The Semiotic Self*. University of Chicago Press.

Wrong, D. H. (1963). Human nature and the perspective of sociology. *Social Research*, *30*(3), 300–318.

Yin, H. H. (2008). From actions to habits: neuroadaptations leading to dependence. *Alcohol Research & Health*, *31*(4), 340–344.

How to cite this article: Abrutyn, S., & Lizardo, O. (2022). A Motivational Theory of Roles, Rewards, and Institutions. *Journal for the Theory of Social Behaviour*, 1–21. https://doi.org/10.1111/jtsb.12360