

Postprint of a Role-Exchange Explanatory Model Based on Predictive Processing Theory

Authors: Ye Fang, Qiu Huilin, Jiang Ke, Li Changjin

Date: 2023-10-09T00:00:00+00:00

Abstract

Role reversal is one of the fundamental techniques in psychodrama therapy. Based on Moreno's classical theory, a role constitutes the protagonist's self-expression. While contemporary psychodrama therapy practice has accumulated rich clinical experience building upon Moreno's theory, it lacks an explanatory theory that can integratively elucidate the working mechanism of psychodrama therapy. Predictive processing theory, proceeding from the analysis of role reversal experiences, explicates the interactive and constructive relationship between the "protagonist" in the individual's own position and the "counterpart" in the reversed position during the role reversal process. According to predictive processing theory, role reversal is a process of self-construction. Across different levels ranging from a single role reversal therapy session to the entire therapeutic process, this self-construction process manifests respectively as: (1) role reversal produces changes in signal types; (2) role reversal brings about model updating; (3) new interpersonal models are generated through interactive construction. In summary, employing predictive processing theory to explain the structured characteristics of role reversal will facilitate this technique in achieving better therapeutic effects in clinical practice.

Full Text

Preamble

ChinaXiv Cooperative Journal

Advances in Psychological Science 2023, Vol. 31, No. 8, 1504-1516

© 2023 Institute of Psychology, Chinese Academy of Sciences

<https://doi.org/10.3724/SP.J.1042.2023.01504>

An Interpretation Model of Role Reversal Based on Predictive Processing Theory

YE Fang^{1,2,3}, QIU Huilin⁴, JIANG Ke^{1,2,3}, LI Changjin^{1,2,3}

¹ School of Mental Health, Wenzhou Medical University

² Key Laboratory of Alzheimer's Disease of Zhejiang Province, Institute of Aging, Wenzhou Medical University, Wenzhou 325035, China

³ Corning Hospital Affiliated to Wenzhou Medical University, Wenzhou 325000, China

⁴ College of Foreign Languages and Cultures, Sichuan University, Chengdu 610207, China

Abstract

Role reversal is one of the fundamental techniques in psychodrama therapy. Based on Jacob Levy Moreno's classic theory, role constitutes the individual's self-presentation. Contemporary psychodrama practice has accumulated rich clinical experience on the foundation of Moreno's theory, yet lacks an explanatory theory to comprehensively elucidate the therapeutic mechanisms of psychodrama. Predictive processing theory, starting from the analysis of role reversal experiences, explains the interactive and constructive relationship between the "protagonist" in the individual's own position and the "counterpart" in the exchanged position during role reversal. According to predictive processing theory, role reversal is a process of self-construction. Across different levels—from a single role reversal session to the entire therapeutic process—this self-construction manifests as: (1) role reversal produces changes in signal types; (2) role reversal brings about model updating; and (3) new interpersonal models emerge through interactive construction. In summary, applying predictive processing theory to explain the structural characteristics of role reversal can enhance the effectiveness of this technique in clinical practice.

Keywords: psychodrama, role reversal, role construction, self, predictive processing

1. Introduction

Role reversal is a basic technique employed in psychodrama therapy, most commonly applied to address interpersonal relationship difficulties. In interpersonal contexts, the client is referred to as the "protagonist," while the person with whom the client experiences conflict is called the "counterpart." In clinical practice, role reversal proceeds as follows: First, the client is asked to play themselves within a specific scenario. Next, they are required to shift their role position and enact the counterpart's role. Role reversal belongs to the broader category of role-playing techniques, but unlike traditional role-playing, it incorporates a transformative step that requires the performer to adopt the counterpart's perspective in thinking and acting.

Jacob Levy Moreno (1946/1961), the founder of psychodrama, conceptualized "role" as a form of interpersonal experience—understanding relationships through role-playing to ultimately achieve better self-understanding. According to psychodrama role theory, roles are categorized into psychosomatic roles, social

roles, and psychodramatic roles. Psychosomatic roles are the first roles individuals develop after birth, encompassing fundamental physiological needs and bodily experiences, as well as somatic responses triggered by psychological self-experience. Social roles emerge from interactions with others (e.g., mother, friend). Psychodramatic roles refer to objectified individual experiences that become externalized roles relative to the self, such as when abstract experiences like dreams, pain, or pleasure are objectified into roles with which the self can dialogue—for instance, a person might converse with their own “pain.”

Moreno believed that roles contain different action components, including collective action components and personal action components (Moreno, 1961). Collective action components refer to action constraint patterns conferred upon individuals by sociocultural norms, while personal action components refer to action patterns formed from individual experience that possess creative qualities. Moreno argued that the more roles a person possesses, the better their quality of life (Moreno, 1946).

Psychodrama emphasizes three factors: space, time, and action (Blatner, 2000). First, psychodrama must occur in a safe space where individuals attempting to enact life roles cannot be punished (Moreno, 1946; Moreno, 1965). Second, temporally, role reversal can present past scenes, current states, or potential future scenarios on the psychodrama stage according to individual needs. Regardless of which stage the role performance occurs in, individuals may exhibit the same behavioral patterns in the here and now—a phenomenon we might metaphorically call “psychological time travel.” Third, Moreno valued the power of action, famously telling Freud: “You analyze people’s dreams; I teach them the courage to dream again” (Moreno, 1946). Through psychodramatic action coordination, we can bring imagined roles into reality, a process termed “surplus reality” (Moreno, 1965).

Numerous studies have confirmed that psychodrama is closely related to clients’ self-experience, including self-awareness (Boroomandian et al., 2020; Dogan, 2018), self-acceptance (Kaya et al., 2021), inferiority feelings (Tümlü & Şimşek, 2021), and self-compassion (Abeditehrani et al., 2020). Research also indicates that psychodrama relates to emotional competence, such as reducing perceived stress (Boroomandian et al., 2020), enhancing empathy (Kipper & Ritchie, 2003; Soysal, 2021), improving emotional intelligence (Beauvais et al., 2019), and particularly demonstrating positive effects in addressing anxiety (Dorothea, 2016; Rudokaite & Indriuniene, 2019). Psychodrama has also shown symptom-alleviating effects for various mental and physical disorders, including major depressive disorder (Abeditehrani et al., 2020), schizophrenia (Miller et al., 2021), ADHD (Mojahed et al., 2021), phobias (Tarashoeva et al., 2017), substance dependence disorders (Giacomucci & Marquit, 2020), and epilepsy (Chu et al., 2009; Han & Chu, 2010; Pan, 2015).

As an expressive arts psychotherapy technique, psychodrama has recently shown a trend toward integration with other therapeutic approaches such as cognitive-behavioral therapy (CBT) and family therapy (Abeditehrani et al., 2020; Blat-

ner, 1997; Farmer et al., 2018). Beyond clinical applications, psychodrama role theory has also contributed significantly to sociological theory and practice (Moreno, 1961; Žurić Jakovina & Jakovina, 2017). Current research both domestically and internationally primarily treats psychodrama as an intervention method, focusing on pre-post intervention differences through self-reports to obtain evidence supporting its effectiveness (Chen et al., 2019). Since psychodrama practitioners typically combine and repeatedly employ multiple techniques in a single session, research on the mechanism of a single technique remains scarce.

Role reversal, as one of the most important techniques in psychodrama, has been called the “engine of psychodrama,” involving experiences of empathy, compassion, and self-reflection (Karp et al., 1998). Clinical applications have yielded numerous case analyses and studies demonstrating role reversal as a successful intervention technique (Boies, 1972), and it can be applied in individual counseling without requiring auxiliary egos (Blatner, 2007). While various models of role reversal practice exist, current research faces several challenges: First, due to individual differences in personality and experience, researchers struggle to control for similarity of issues across participants (Boies, 1972). Second, researchers have difficulty determining effective evaluation indicators, as the mechanism through which role reversal changes behavioral patterns remains unclear. Third, there is a lack of a unified theory to integrate empirical data and provide scientific explanations for relevant findings (Orkibi & Feniger-Schaal, 2019).

The effectiveness of role reversal stems from the enhanced structuralization of individuals’ psychological experiences regarding self and others during the role-playing process (Carlson-Sabelli, 1989). This paper introduces predictive processing theory to explain the working mechanism of role reversal. By first clarifying the relationship between predictive mind and self-knowledge, and using self-construction as the main thread, we explain through predictive processing model construction how role reversal technology functions in psychological structuration.

2. Role Reversal and Predictive Mind

2.1. Knowing the Self Through Predictive Mind

Extensive research demonstrates that psychodrama is closely related to self-experience, effectively promoting the structural development of individuals’ self-concepts (Butler & Fuhrman, 1980; Rohrbaugh & Bartels, 1975; Yalom, 1995). To explore the internal process through which role reversal in psychodrama affects individual self-experience, we must first briefly clarify the concept of “self-knowledge.”

Human self-knowledge has long been a “hard problem” in philosophy of mind and psychological research (Fei, 2018). The meaning of self and self-experience remains contested. Modern analytic philosophy has argued that individuals know themselves through interactional objects rather than through conscious-

ness or introspection (Gilbert Ryle, 1992). Here, interactional objects refer not only to others but also to the knower themselves. When individuals interact with others or with themselves, they come to know others by observing the performance or experiences of interactional objects, and thereby come to know themselves. Certain types of actions—such as observing, opposing, imitating, and mocking—involve or are superimposed upon other types of actions in some way. These are called “higher-order actions.” Individuals acquire the capacity to take higher-order actions toward the behaviors of interactional objects and apply this capacity in their interactions with themselves, thereby gaining self-knowledge. This is the reflexivity of the self (Mead, 1934).

The reflexivity of the self is a distinctive feature that distinguishes humans from animals (Baumeister, 1998). To date, the multi-faceted hierarchical model of self proposed by Shavelson et al. (1976) remains widely recognized. This model first defines self-concept as a series of organized, hierarchical, multidimensional, and distinguishable self-perceptions, then describes the generative process of self-concept from interpersonal interaction and perceptual experiences of self-attributes and social environments. However, how interpersonal interaction and social awareness generate self-knowledge still requires further theoretical explanation.

Clark proposed predictive processing theory in 2013, which has generated strong responses in philosophy of mind and cognitive science. The theory posits that in an uncertain world, humans need to construct models of the world. These models consist of probability density distributions of world regularities. Individuals use existing models to generate top-down predictive signals, which continuously match actual input signals. The difference between them is called prediction error (PE). Prediction error is fed back to the brain through bottom-up pathways, driving individuals to take action to minimize prediction error. By continuously revising their models to align with the real world, individuals reduce uncertainty about the world (Clark, 2013/2017). Predictive processing theory represents a new paradigm for understanding cognitive processes of mind in recent years (Su & Ye, 2021). This paradigm combines resources from neuroscience, philosophy, biology, and information theory, viewing the mental system as an active “Bayesian inference engine” and has been regarded as the best explanatory inference about mind currently available (Wang, 2021).

The predictive processing approach (hereafter “predictive mind”) provides a possible framework for explaining self-knowledge (Wang, 2021). However, we still need to answer: How does predictive mind function in the process of knowing the self? Predictive processing theory maintains that individuals’ modeling of their own bodies is a natural result of prediction error minimization during environmental interaction (Hohwy & Michael, 2017). A simple example: When walking on a street, individuals must not only predict others’ walking directions and speeds but also have correct knowledge of their own body size to maintain action boundaries with strangers and avoid collisions.

According to predictive processing theory, the self is a set of hierarchical causal

models hidden within the body. Lower-level selves (such as sensory experiences and body cognition) are modulated by higher-level selves (such as personality traits and beliefs) (Hohwy & Michael, 2017). Prediction error feeds back from lower-level structures to higher-level structures, thereby constructing certain stable features of the self. One might say that an individual's predictive signals constitute their self-representation. Higher-level predictive signals depend on stable patterns in lower-level predictive signals (Kwisthout et al., 2017). Based on the free energy principle, individuals achieve self-consistency and continuity through prediction error minimization (Hohwy & Michael, 2017). In short, the mechanism by which predictive mind knows the self is: On the one hand, individuals can represent themselves by emitting predictive signals; on the other hand, the matching of predictive signals and input signals generates new needs for self-representational modeling. These two processes cycle repeatedly until prediction error is minimized.

Self-knowledge is crucial for psychosomatic health. Erikson's theory of ego identity posits that individuals need to continuously integrate experiences of self and object, id and superego, forming internal consistency and continuity to develop adaptive responses to social environments (Erikson, 1959). If individuals' self-knowledge development is incomplete, lacking a clear and continuous self-concept, they are prone to interpersonal conflicts and negative emotions such as anxiety and disappointment, potentially leading to psychological disorders (Han, 2021; Kellermann, 1994). Negative mood states may in turn trigger cognitive immunization, where individuals weaken or deny prediction errors arising from their own unreasonable expectations, refusing to modify their inherent generative models. For example, in specific situations, patients with major depressive disorder who initially hold negative expectations will maintain their original negative expectations when receiving positive signals inconsistent with their expectations through cognitive immunization. Conversely, if patients initially hold positive expectations, they will more easily amplify prediction errors when receiving negative signals, making their expectations negative. This attentional bias reduces the possibility of correct self-knowledge, ultimately leading to a negative cycle (Kube et al., 2020).

Psychodrama's role reversal technique can help individuals accept objectively existing prediction errors through dialogue between the protagonist and counterpart. For patients' stereotypical negative expectations generated in negative mood states, role reversal can help break the negative cycle of cognitive immunization, facilitating timely updating of negative expectations. For patients' pre-existing positive expectations, role reversal, on the one hand, guides individuals to spontaneously generate multiple evaluations of negative events different from before; on the other hand, to some extent, coordinates input signals with the real state of the external world, reducing negative experiences caused by discrepancies between prior positive expectations and posterior negative events. Ultimately, different experiences are integrated into a unified self, thereby enhancing individuals' mental health.

2.2. Role Reversal as Predictive Mind Model Construction

The theoretical foundation of role reversal practice relies on psychodrama role theory. The basic view of self-concept in role theory highly aligns with predictive processing theory (Li, 2021; Hohwy, 2020). Therefore, we can provide a theoretical explanation of psychodrama's mechanism through predictive processing theory.

First, individuals' internal generative models are role-based units. Moreno defined role as the symbolic representation of an individual's functional form—that is, role is the tangible aspect of self, actually and concretely existing within the self (Moreno, 1946/1961; Fürst, 2020). When individuals introduce themselves or learn about others, they tend to report their roles, such as “student” or “mother.” People always know themselves and others through certain roles. Under specific sociocultural norms, each role has universal constraints that constitute the role's collective action components. According to predictive processing theory, roles provide possibilities for model construction in the brain. Such models include prior knowledge about the role; individuals make prior conjectures or predictions about others based on generative models and establish adaptive interpersonal relationships by reducing prediction error in real interactions (Lee et al., 2021). Role stereotypes are a practical example of predictive mind. Insufficient knowledge of roles or a limited role repertoire easily leads to mental health problems. For instance, Levin's (2021) case analysis showed that mothers often experience rapid role and daily life transitions during postpartum childcare, and failure to handle relationships and transitions between roles can trigger postpartum depression. In the workplace, correct self-knowledge and role positioning are essential elements of mental health for professional women (Chen, 2012; Ma & Zhang, 1999). Role reversal enables individuals to try different roles, promoting self-knowledge. When role conflicts arise, individuals can make decisions more calmly based on correct self-knowledge.

Second, the relationship between generative models and the world is essentially constructive. The constructiveness of self already appeared in early psychoanalytic theory. According to Freud's classic psychoanalytic theory, the ego is formed through the interaction between the id and social environment, representing a compromise between id desires and caregiver constraints. In humanistic-oriented ego psychology, concepts such as psychological self and social self also derive from individuals projecting their actions onto others or the environment, then obtaining self-concept from others' feedback. This process also embodies constructivist meaning. Therefore, although different theoretical schools describe and define the self differently, none reject the constructivist presupposition that the self emerges from the interaction between individual action and environment. Moreno explicitly stated that the self develops from socially pre-given roles, not the reverse (Moreno, 1961).

Roles are inherently interdependent, with each role having its counterpart—for example, “husband-wife” or “boss-employee.” Individuals' knowledge of a

particular self-role cannot be separated from interaction with their counterpart. For instance, if a child consistently receives negative evaluations from their mother—“You’re stupid, you can’t do anything right”—the child will likely grow up believing they are incompetent. Similarly, a father is often the first male a daughter encounters; thus, the father-daughter relationship constructs the child’s internal predictive mind model, profoundly influencing her future intimate relationships. The daughter typically expects to find a partner with traits identical to or completely opposite to her father’s.

In summary, individuals connect their internal and external worlds through predictive mind. Role reversal is the model construction process of predictive mind. Role reversal encourages each role to understand the counterpart’s perspective and find ways to coexist peacefully, promoting mental health development. Moreno believed that the capacity for role reversal is crucial for children’s social development, helping them move beyond egocentrism (cited in Kellermann, 1994). Mann and Mann’s (1959) research showed that compared with control groups, role reversal groups demonstrated significant improvements in ratings of “attractiveness as a friend,” “helping the group achieve goals,” and “cooperativeness.” Role reversal facilitates interpersonal relationship regulation. Abeditehrani et al. (2020) applied role reversal to social anxiety disorder populations, targeting their characteristic negative cognitions about others by requiring participants to rate subjective measures before and after role conversion. Results showed that the role reversal group, compared with the role-playing group (which lacked the role conversion step), exhibited significant reductions in “confidence in negative cognitions,” “probability of being negatively evaluated,” and “cost of being negatively evaluated,” supporting the effectiveness of role reversal. The successful application of psychodrama role reversal techniques across different populations demonstrates its therapeutic value (Boies, 1972).

Through role reversal, individuals can better adapt to life. However, a reasonable theoretical explanation for this effect remains lacking. Orkibi and Feniger-Schaal (2019) conducted a systematic review of psychodrama research from 2007–2017, offering critical evaluations of current psychodrama research methods and emphasizing the necessity of theoretical explanations for psychodrama’s specific mechanisms. This paper aims to introduce predictive processing theory to explain the unique therapeutic effects of role reversal technology and provide a referential hypothetical framework for future empirical research on psychodrama.

3. Predictive Processing Theory’s Explanation of Role Reversal

Predictive processing theory posits that the human brain does not passively receive external stimuli but actively generates expectations about upcoming input signals (Clark, 2018; Kwisthout et al., 2017). The comparison between predictive signals and input signals forms prediction error. Following the free energy principle (Friston, 2010; Ransom et al., 2020; Wiese & Metzinger, 2017), indi-

viduals minimize prediction error to enhance their ability to adapt to the environment. When facing prediction error, individuals choose to either take action to satisfy their needs or adjust their expectations to align with the environment; this depends on the precision weighting values assigned to predictive signals and input signals. Higher precision (however encoded) means less uncertainty, which in turn means individuals balance the interaction between top-down and bottom-up processes differently (Clark, 2013/2017).

Role reversal practice occurs within corresponding relationships, including protagonist and counterpart. Each role party possesses models for understanding objective facts, which are often inconsistent or even contradictory. Individuals in their own role generate predictive signals based on their model, receive input signals from the other party, and produce prediction errors through the matching of top-down predictive signals and bottom-up input signals, leading to cognitive dissonance and difficulty in self-integration. The purpose of role reversal is to help individuals structure prediction errors during the “protagonist-counterpart-protagonist” role conversion process, thereby creating opportunities for model reconstruction, promoting self-awareness and growth, and ultimately eliminating prediction errors.

3.1. Role Reversal Induces Signal Changes

There are two types of predictions in daily life (Gilbert Ryle, 1992). The first involves considering explicit propositions, where action results do not differ far from expectations, leading to reduced surprise. The second type of prediction can be expressed as a mood—that is, individuals are always aware of what they are doing without having considered explicit propositions, yet are not surprised by the conclusions reached. When making the second type of prediction, higher-order actions are not always implied, so prediction error does not lead to reduced surprise. When facing common interpersonal situations, predictive moods cause individuals to generate predictive signals without being aware of their content or without generating reflective experiences about their content (Liu, 2021). Due to this lack of reflection, individuals cannot adopt effective strategies to reduce prediction error, ultimately leading to emotional problems.

Because of the elusive nature of the “I” system, I can chase my past self but can only ever grasp the fluttering hem of my coat (Sartre, 1957). The difference between “I” and “you” is that “I” can never grasp my current self but can grasp your (the counterpart’s) current self. Role reversal enacts a specific situation from the client’s life. The signals generated by the client in the protagonist position belong to predictive signals, while signals received from the counterpart belong to input signals. When the client shifts from the protagonist position to the counterpart position, the previously received input signals become predictive signals, and their own previous predictive signals become input signals. The object of examination changes from self to other during this transition, allowing the client to treat the existence and behavior of “I” as an object of knowledge, thereby elevating the level of self-knowledge. This process is illustrated in Figure

1 [Figure 1: see original paper].

Let us use a common parent-child relationship example to demonstrate how the predictive processing mechanism operates. Suppose the client's role is a daughter and the counterpart is her mother. The client's presenting problem is that after leaving home for university for the first time, her mother calls several times daily to inquire about her activities. The client feels her life is being disrupted, which stimulates her to take action: arguing with her mother. In this case, the client holds a model of the parent-child relationship that includes appropriate boundaries for getting along and daily communication frequency with her mother. When the model functions normally, everything is familiar information for the client, and she feels no distress. When the client leaves home for university, the mother's model and the client's model diverge, disrupting the client's normal state and creating prediction error. For example, the client's input signal might be "Mother is disrupting my life; she doesn't care about me or my feelings, only her own needs," while the predictive signal is "Mother should care about me, consider my feelings, and not call me several times daily to disrupt my life" (see Figure 1a). Due to the client's low spontaneity level and limited role repertoire, the input signals received often contain biases. When the client shifts to the counterpart position—that is, the mother's position—the input signals become predictive signals. At this point, "My daughter doesn't understand me; my care for her is instead met with disgust" becomes the input signal, while "My daughter has been physically weak since childhood; I need to remind her to take care of herself, so calling several times is necessary, and my daughter should understand my concern for her" becomes the predictive signal (see Figure 1b). At the conclusion of role reversal, the client shifts back to the protagonist position, and the predictive signal becomes the input signal. This input signal differs from the pre-exchange input signal, transforming from "Mother keeps calling me, disrupting my life; she doesn't care about me or consider my feelings" to "Mother calls multiple times because she is worried about me; she loves me but doesn't know this would disturb me; she does this out of love" (see Figure 1c). The changed input signal aligns with the predictive signal "Mother should care about me, consider my feelings, and not disregard my feelings by calling several times daily to disrupt my life," reducing prediction error and thereby establishing an adaptive relationship with the external world and reducing interpersonal distress.

To facilitate signal transformation more effectively, consultants need to lead clients through adequate warm-up before role conversion. Combining the three stages of role reversal (Kellermann, 1994), the signal change process can also be divided into three steps: First, the client begins by imitating the counterpart's movements and language physically, helping them better inhabit the unfamiliar counterpart role. This stage is empathic role taking and represents the first step in signal change. Next, the consultant collects information related to the client's issue through relational questioning, presenting background information about the counterpart's generation of predictive signals. The final stage is role feedback, where the input signals generated by the client in the counterpart po-

sition include how “I” (the counterpart) perceive “you” (the protagonist), while predictive signals include how “I” (the counterpart) perceive how “you” (the protagonist) perceive “me” (the counterpart). These three steps interact, presenting implicit conflict patterns through the process of signal transformation.

3.2. Role Reversal Leads to Model Updating

Role reversal does not require clients to completely abandon their models but rather, according to the principle of creativity, to integrate adaptive parts from the counterpart’ s model into their own (Kellermann, 1994). Clients need to simultaneously activate representations of self and other. When clients shift to the counterpart’ s position, they are required to temporarily suppress their own feelings and present experiences and actions different from their authentic feelings. Compared with role-playing techniques, this is a creative process requiring high-level volitional control (Yaniv, 2012). Role reversal provides a linking path between the protagonist’ s model and the counterpart’ s model in real scenarios. The subsequent question is: How is prediction error reduced during role reversal?

Rothenberg (1971) proposed the concept of homospatial process to clarify the distinction between creativity and psychopathological symptoms. Homospatial process refers to two concrete entities simultaneously occupying one space, generating a new entity expression (Rothenberg, 1971/1995). For example, German organic chemist Kekulé discovered the benzene molecular structure by superimposing the representation of a snake biting its tail onto the representation of atoms. Yaniv (2012) suggests that homospatial process may be one mechanism of role reversal, where the protagonist’ s model and the counterpart’ s model coexist simultaneously in the individual’ s predictive mind. Moreno believed that individuals have the capacity to create adaptive behavioral patterns at high levels of spontaneity, ultimately enriching their cultural converse. Individuals generate predictive signals based on their models and compare them with input signals from the environment. When predictive and input signals match, individuals consolidate their models; when they mismatch, individuals update their generative models or change the environment. The protagonist has one model that, through position conversion, is brought to the counterpart’ s position. The homospatial process of the two models initiates role reversal. The predictive mechanism in the counterpart position is identical to that in the protagonist position. Since prediction errors generated in the counterpart position and protagonist position occur within the same interactive context, any control of prediction error in either the protagonist or counterpart will bring about changes in the other’ s prediction error, thereby changing the other’ s model (see Figure 2 [Figure 2: see original paper]). In the previous parent-child example, the client matches predictive and input signals in the mother role, reducing prediction error. The homospatial process of mother and daughter models creates momentum for prediction error reduction, so when the client returns to the daughter role, the generative model of the daughter role is altered, helping

the client develop a new cognition: “Mother calls multiple times because she is worried; she loves me.” Ultimately, prediction error in the protagonist position is also controlled. The purpose of role reversal is to help clients step out of their past experiences, form new experiences amid the confusion generated by the coexistence of two models, and escape the constraints of stereotypical patterns.

When do we update our models? Precision weighting determines the extent to which individuals rely on environmental evidence versus their own prior models (Yon & Frith, 2021). When environmental information is ambiguous, we rely more on past experiences and established models to make judgments. When input information has high precision, we lower the weight of prior probabilities in decision-making and trust more in posterior evidence from environmental input. Interpersonal precision weighting works similarly: When interacting with different people, individuals assign different weights to their words. Role reversal can be viewed as a precision weighting process. When the counterpart’s model and the protagonist’s model coexist, the client’s responses to relational questions actually supplement knowledge and details about both models, regulating the precision of each model. Higher detail levels yield higher model precision (Kwisthout et al., 2017). Next, the consultant facilitates direct dialogue between the two roles, promoting weighing and comparison of the two models. Specifically, the consultant invites the client to adjust the distance and orientation of two chairs according to their relationship model with the counterpart to reduce prediction error. However, when discussing the effectiveness of a single role reversal session in changing the client’s real-life predicament, it seems more likely to manifest in helping clients update their original generative models rather than changing the environment.

So how does the mind make predictions, and what is the mechanism of predictive mind throughout the entire psychodrama process? We need to answer this from a dynamic perspective to explain how role reversal technology promotes individual self-healing and growth, helping individuals interact more smoothly with their environment.

Predictive mind makes predictions through paths approximating Bayesian inference (Clark, 2013/2017). The objective world contains numerous discrete random variables, whose stochastic causal relationships constitute causal Bayesian networks (Pearl, 2000). Individuals’ knowledge about world conditions and knowledge about the self are both hierarchical causal models, with each level describable by a causal Bayesian network. Variables in the network can be divided into three types: hypothesis variables, prediction variables, and intermediate variables (Kwisthout et al., 2017). Hypothesis variables are general terms for variables that can generate predictions in the network. Prediction variables are general terms for predictions generated based on hypothesis variables. Logically, hypothesis variables precede prediction variables at the same level, while prediction variables at one level equal hypothesis variables at the next level. Intermediate variables affect the relationship between hypothesis variables and prediction variables (see Figure 3a). The earlier idealized parent-child relation-

ship example helps understand the hierarchical structure of generative models.

Assume the client's issue is at level $n-1$ in the protagonist model. At this level, the prediction variable is "Mother won't keep calling me," and the hypothesis variable is "Mother won't excessively interfere with my life." At level n , the prediction variable is "Mother should trust that I can take care of myself," and the hypothesis variable is "Mother should trust my abilities." At level $n+1$, the prediction variable is "Mother should care about me and consider my feelings," and the hypothesis variable is "Mother loves me." Through role reversal, the counterpart's hierarchical model is constructed in interaction. At level $n-1$, the prediction variable is "Calling several times is necessary," and the hypothesis variable is "My daughter has been physically weak since childhood; I need to remind her to take care of herself." At level n , the prediction variable is "My daughter needs my reminders," and the hypothesis variable is "I worry about her, fearing she can't take care of herself." At level $n+1$, the prediction variable is "I am concerned about my daughter," and the hypothesis variable is "I love my daughter very much." The relationship between predictive and hypothesis variables is complex. Within the psychodrama framework, social systems, cultural contexts, and interactions between different norms all constitute intermediate variables. Every individual exists within culture, which subtly influences their internal conflict models and external manifestations when facing conflicts. Psychodrama practice emphasizes the social and cultural sources behind each role. The collective action components of roles result from generations of interpersonal interaction and significantly influence case conceptualization (Moreno, 1961). Both protagonist and counterpart possess their own multi-level Bayesian models. The interaction between the two models constitutes a larger causal Bayesian network (see Figure 3b [Figure 3: see original paper]). Therefore, the interaction process between protagonist and counterpart is actually an interactive construction process of their predictive processing models within a larger Bayesian network.

3.3. Model Construction Through Interaction

In predictive processing theory, the process of action generates the need for self-construction. The environment contains numerous uncertain factors. To meet survival needs, individuals must construct models of the world and make timely predictions and appropriate responses to upcoming events (Kelly et al., 2019).

Treatment through role reversal is typically not completed in a single session but involves multiple exchanges according to the client's needs within one session. Clients experience a process from role learning to spontaneous development during multiple role reversals, including role taking, role playing, and role creating, thereby expanding their role repertoire and improving quality of life (Moreno, 1961). Specifically, if the counterpart is the client's mother, when the client shifts to the counterpart's position, she will acquire general characteristics of the mother role, such as maternal love and dedication—this is role taking. Next, the client enacts the mother through imitating her speech patterns, tone, and

even expressions—this is higher-level processing of the acquired model. Finally, the client enters the role creating stage, bringing new insights and understandings about the role into the counterpart's model, performing words or actions the mother never said or did at high levels of spontaneity. The role development process follows the rule of model construction from lower to higher levels. While constructing the counterpart's model through role conversion, the client also clarifies the predictive model about the self. The result is that the client's self-knowledge undergoes updating from lower to higher levels. Multiple role reversals align corresponding levels of the protagonist's and counterpart's models, constructing them interactively.

Causal Bayesian networks depend not only on static causal relationships but also on dynamic interactions over time. First, current predictive signals depend not only on current hypotheses but also on predictive signals from a previous time slice (Kwisthout et al., 2017). Second, from a broader temporal perspective, models are updated not only through multiple role exchanges within a single session but also across multiple sessions, as shown in Figure 4 [Figure 4: see original paper]. Goldman and Morrison (1984) described the predictive processing and model transformation process throughout psychodrama therapy through the psychodrama spiral diagram. Psychodrama work begins with the client's current problem, enacting a specific past scene, then connecting to other related scenes and eventually to earlier childhood experiences. The ultimate goal of psychodrama therapy is to achieve integrative catharsis, helping clients better adapt to life through role training. The psychodrama spiral moves through a continuous ascending process from past experiences to future construction, enabling individuals to gain more knowledge about self and others. Based on this structured knowledge, consultants and clients co-create effective strategies to handle prediction errors. Therefore, psychodrama may be a more practical psychotherapeutic technique compared to psychoanalytic schools (Kadyrov, 1997). It should be noted that in actual psychodrama processes, the psychodrama spiral transforms dynamically, with role interactions determined according to the client's current state, reflecting the vitality of psychodrama technology and the integration of individual spontaneity and creative growth.

The mechanism of role reversal raises another question: Does the counterpart that the individual enacts and the counterpart model obtained match the counterpart's actual model? Since "I" cannot output what is not in "I's" own model, the counterpart model expressed by "I" still belongs to part of "I." Regarding this query, this paper argues that this understanding is reasonable to some extent. From the individual's internal perspective, the counterpart model obtained in the counterpart position through role reversal is partly related to the individual's subjective reality—interpersonal conflict is the conflict of the individual's internal models. However, we cannot ignore psychodrama's interpersonal perspective and the connection between intra-subjective and inter-subjective perspectives. Models are constructed from the environment. Both an individual's evaluation of themselves and others' evaluations of them are learned within relationships and culture. That is, others in the environment do

evaluate in certain ways, and the individual's internal world has corresponding parts that either identify with or reject these evaluations, thereby generating internal models. Since there is an "I," there must be a "non-I." "I" obtains construction about "I" within the "non-I."

Spontaneity and creativity are core concepts in Moreno's theory. During the protagonist's self-construction through role reversal technology, conflicts or inconsistencies arise between multiple role models. How does role creation occur in the predictive processing process? Does the degree of correspondence between individuals' recognition of input signals and objective reality affect therapeutic effectiveness? On the one hand, consultants can help clients' input signals approach the real world through guided questioning, relational questions, and facilitating direct dialogue between roles. On the other hand, we need to return to Moreno's view of human nature. He believed that creativity inherent in human nature represents God's infinite creation. Everyone has the capacity to create and change their environment, enabling sustained change and adaptation in their relationship with the environment. Rigid interpersonal interaction patterns can be changed through the power of action, and individuals can gain new ways of coping with old situations or adaptive behaviors for upcoming new situations in psychodrama therapy.

4. Conclusion and Future Directions

Role reversal presents potential conflicts in concrete form, promoting changes in individuals' non-adaptive behavioral patterns. From the perspective of predictive processing theory, role reversal is a process of self-construction. Individuals acquire structured knowledge about self and others through role reversal, reducing prediction error and guiding life toward positive interaction with the real world. The three explanatory models proposed in this paper elaborate the mechanism of role reversal at different levels, with a progressive internal logical relationship: The conversion step first brings about signal change, which further leads to updating of the individual's internal model, and finally involves transformation of the model's hierarchical architecture. Throughout the entire therapeutic process involving multiple role exchanges, predictive processing models undergo dynamic updating through the psychodrama spiral.

As previously mentioned, role reversal can promote self-knowledge and help individuals improve mental health levels. One question requires clarification: As a constructivist framework, predictive processing theory maintains that the relationships between generating predictions based on self-models, matching predictive signals with input signals, and constructing and updating self-models are only logically sequential, not temporally sequential. That is, representation-based top-down predictive processing, experience-dependent bottom-up predictive processing, and model generation, construction, and processing actually occur simultaneously. The interaction construction of protagonist and counterpart models has no sequential order but forms interactively, because the counterpart constitutes part of the protagonist's self-existence.

Applying predictive processing theory to explain role reversal technology can provide a framework for guiding future intervention practice. In work with clients, consultants can concretize the prediction errors generated when clients interact with counterparts, jointly exploring strategies for minimizing prediction error in specific issues. If prediction error stems from the inherent nature of the world and cannot be reduced, clients need to modify their hypothesis variables. If prediction error results from insufficient knowledge of the counterpart's model and can be reduced, clients need to modify random causal relationships in their models or introduce new variables to update the model. Alternatively, other psychodrama intervention techniques (such as mirroring) can be used to resample input signals to achieve prediction error minimization. Psychodrama's theoretical foundation also contains many elements corresponding to traditional Chinese culture. For example, the cultural connotations of "benevolence, righteousness, propriety, wisdom, and harmony" and introspective awareness based on these values, as well as the mind cultivation practices in the School of Heart-Mind, may have internal connections and comparative possibilities with predictive processing theory. How to emphasize the core characteristics of spontaneity and creativity in psychodrama practice while introducing traditional Chinese cultural factors, and explore structural, hierarchical, and development-oriented localized, life-oriented, and effective psychodrama techniques adapted to Chinese characteristics to genuinely enhance psychological healing and national health promotion effects—these should be future exploratory directions guided by predictive processing theory and role reversal technology.

References

- Blatner, A. (1997). Psychodrama: The state of the art. *The Arts in Psychotherapy*, 24(1), 23–30.
- Blatner, A. (2000). *Foundations of psychodrama: History, theory, and practice* (4th ed.). New York, NY: Springer.
- Blatner, A. (2007). Morenean approaches: Recognizing psychodrama's many facets. *Journal of Group Psychotherapy, Psychodrama and Sociometry*, 59(4), 159–170.
- Boies, K. G. (1972). Role playing as a behavior change technique: Review of the empirical literature. *Psychotherapy: Theory, Research and Practice*, 9(2), 185–192.
- Boroomandian, N., Mohammadi, N., & Rahimi Taghanaki, C. (2020). Comparing the effectiveness of group therapy by "interaction and psychodrama" on emotional awareness of adolescent girls with psychological distress. *Journal of Psychological Achievements*, 27(1), 1–22.
- Butler, T., & Fuhriman, A. (1980). Patient perspective on the curative process: A comparison of day treatment and outpatient psychotherapy groups. *Small Group Behavior*, 11(4), 371–388.

- Carlson-Sabelli, L. (1989). Role reversal: A concept analysis and reinterpretation of the research literature. *Journal of Group Psychotherapy, Psychodrama and Sociometry*, 42, 139-152.
- Chen, S. (2012). Research on problems and countermeasures for professional women. *Enterprise Technology Development*, 31(34), 121-122.
- Chen, J., Qu, Y., & Hu, H. (2019). Research progress on the mechanism and effectiveness of psychodrama techniques in depression intervention. *China Contemporary Medicine*, 26(8), 31-34.
- Chu, P., Yan, J., & Li, D. (2009). Effects of psychodrama therapy on negative psychology and coping styles in adolescent epilepsy patients. *International Journal of Nursing*, 28(3), 322-324.
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral & Brain Sciences*, 36(3), 181-204.
- Clark, A. (2017). Predictions, precision, and agentic attention. *Consciousness and Cognition*, 56, 115-119.
- Clark, A. (2018). A nice surprise? Predictive processing and the active pursuit of novelty. *Phenomenology and the Cognitive Sciences*, 17(3), 521-534.
- Dogan, T. (2018). The effects of the psychodrama in instilling empathy and self-awareness: A pilot study. *PsyCh Journal*, 7(4), 227-238.
- Dorothea, G. (2016). Effectiveness of psychodrama group therapy on pupils with mathematics anxiety. In C. Stadler, M. Wieser, & K. Kirk (Eds.), *Psychodrama empirical research and science 2* (pp. 197-215). Wiesbaden: Springer.
- Erikson, E. H. (1959). *Identity and the life cycle: Selected papers*. New York: Norton.
- Farmer, C., Moreno, Z. T., Campbell, D., & Draper, R. (2018). *Psychodrama and systemic therapy*. New York: Routledge.
- Fei, D. (2018). *Research on the mind-body relationship problem*. Beijing: The Commercial Press.
- Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127-138.
- Friston, K. J., Daunizeau, J., Kilner, J., & Kiebel, S. J. (2010). Action and behavior: A free-energy formulation. *Biological Cybernetics*, 102(3), 227-260.
- Fürst, J. (2020). Role development in psychodrama training—Findings and challenges. *Zeitschrift für Psychodrama und Soziometrie*, 19(Suppl. 1), 239-253.
- Giacomucci, S., & Marquit, J. (2020). The effectiveness of trauma-focused psychodrama in the treatment of PTSD in inpatient substance abuse treatment. *Frontiers in Psychology*, 11, 896.

- Gilbert Ryle. (1992). *The concept of mind* (Xu D. J., Trans.). Beijing: The Commercial Press.
- Goldman, E., & Morrison, D. (1984). *Psychodrama: Experience and process*. Dubuque: Kendall Hunt.
- Han, J. (2021). A study on the relationship among social support, self-concept and mental health of college students (Master's thesis). Heilongjiang University, Harbin.
- Han, K., & Chu, P. (2010). Effects of psychodrama therapy on shame and quality of life in children with epilepsy. *Chinese Journal of Practical Nursing*, 26(23), 89-91.
- Hohwy, J. (2020). New directions in predictive processing. *Mind & Language*, 35(2), 209-223.
- Hohwy, J., & Michael, J. (2017). Why should any body have a self? In F. de Vignemont & A. Alsmith (Eds.), *The body and the self, revisited* (pp. 363-391). Cambridge, MA: The MIT Press.
- Kadyrov, I. M. (1997). Psychodrama and psychoanalysis: Two theaters for mental drama. *Journal of Russian & East European Psychology*, 35(6), 55-72.
- Karp, M., Holmes, P., & Tavon, K. B. (Eds.). (1998). *The handbook of psychodrama*. New York: Psychology Press.
- Kaya, Y., Güç, E., Kırlangıç-Şimşek, B., & Cesur-Atınas, M. (2021). The effect of psychodrama group practices on nursing students' self-liking, self-competence, and self-acceptance: A quasi-experimental study. *Perspectives in Psychiatric Care*, 58(4), 1239-1247.
- Kelly, M. P., Kriznik, N. M., Kinmonth, A. L., & Fletcher, P. C. (2019). The brain, self and society: A social-neuroscience model of predictive processing. *Social Neuroscience*, 14(3), 266-276.
- Kellermann, P. F. (1994). Role reversal in psychodrama. In P. Holmes, M. Karp, & M. Watson (Eds.), *Psychodrama since Moreno: Innovations in theory and practice* (pp. 263-279). London: Routledge.
- Kipper, D. A., & Ritchie, T. D. (2003). The effectiveness of psychodramatic techniques: A meta-analysis. *Group Dynamics: Theory, Research, and Practice*, 7(1), 13-25.
- Kube, T., Schwarting, R., Rozenkrantz, L., Glombiewski, J. A., & Rief, W. (2020). Distorted cognitive processes in major depression: A predictive processing perspective. *Biological Psychiatry*, 87(5), 388-398.
- Kwisthout, J., Bekkering, H., & van Rooij, I. (2017). To be precise, the details don't matter: On predictive processing, precision, and level of detail of predictions. *Brain and Cognition*, 112, 84-91.

- Lee, K. M., Ferreira-Santos, F., & Satpute, A. B. (2021). Predictive processing models and affective neuroscience. *Neuroscience & Biobehavioral Reviews*, 131, 211-228.
- Levin, K. R. (2021). *Mommy or me: Perceptions of mothers with postpartum depression on occupational disruption and role conflict* (Unpublished master's thesis). Ithaca College, NY.
- Li, H. (2021). The life view of mind: From pragmatism to predictive mind. *Social Science Front*, (9), 17-25.
- Liu, X. (2021). The effect of mood on predictive sentence processing by older adults. *SAGE Open*, 11(2), 1-10.
- Ma, J., & Zhang, T. (1999). Analysis of the relationship between role cognition, sense of control and work stress. *Chinese Journal of Ergonomics*, (4), 15-19.
- Mann, J. H., & Mann, C. H. (1959). Role playing experience and interpersonal adjustment. *Journal of Counseling Psychology*, 6(2), 148-152.
- Mead, G. H. (1934). *Mind, self, and society*. Chicago: University of Chicago Press.
- Miller, R., Mason, S., & Skolnik, S. (2021). Psychodrama for schizophrenia: Adaptive techniques in mutual aid groups. *Social Work with Groups*, 44(3), 273-287.
- Mojahed, A., Zaheri, Y., & Moqaddam, M. F. (2021). Effectiveness of group psychodrama on aggression and social anxiety of children with attention-deficit/hyperactivity disorder: A randomized clinical trial. *The Arts in Psychotherapy*, 73, 101756.
- Moreno, J. L. (1946). *Psychodrama* (Vol. 1). New York: Beacon House.
- Moreno, J. L. (1961). The role concept: A bridge between psychiatry and sociology. In J. Fox (Ed.), *The essential Moreno* (pp. 60-66). New York, NY: Springer.
- Moreno, Z. T. (1965). Psychodramatic rules, techniques, and adjunctive methods. *Group Psychotherapy*, 18(1-2), 73-86.
- Orkibi, H., & Feniger-Schaal, R. (2019). Integrative systematic review of psychodrama psychotherapy research: Trends and methodological implications. *PloS ONE*, 14(2), e0212575.
- Pan, H. (2015). Effects of psychodrama therapy on psychology and hope in school-age children with epilepsy. *International Journal of Nursing*, (10), 1389-1391.
- Pearl, J. (2000). *Causality: Models, reasoning and inference*. Cambridge, MA: Cambridge University Press.

- Ransom, M., Fazelpour, S., Markovic, J., Kryklywy, J., Thompson, E. T., & Todd, R. M. (2020). Affect-biased attention and predictive processing. *Cognition*, 203, 104370.
- Rohrbaugh, M., & Bartels, B. D. (1975). Participants' perceptions of "curative factors" in therapy and growth groups. *Small Group Behavior*, 6(4), 430-456.
- Rothenberg, A. (1971). The process of Janusian thinking in creativity. *Archives of General Psychiatry*, 24(3), 195-205.
- Rothenberg, A. (1995). Creative cognitive processes in Kekule's discovery of the structure of the benzene molecule. *The American Journal of Psychology*, 108(3), 419-438.
- Rudokaite, D., & Indriuniene, V. (2019). Effectiveness of psychodrama for mitigating school fears among senior secondary school students. *Zeitschrift für Psychodrama und Soziometrie*, 18(2), 369-385.
- Sartre, J. P. (1957). *The transcendence of the ego: An existentialist theory of consciousness*. London: Macmillan.
- Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, 46(3), 407-441.
- Soysal, F. S. Ö. (2021). The effects of psychodrama on emotion regulation skills in emerging adults. *Current Psychology*, 34, 1-14.
- Su, J., & Ye, H. (2021). "Enactive cognition," "interaction theory," and "predictive processing model"—New thinking on the "problem of other minds." *Psychological Science*, 44(6), 1519-1526.
- Tarashoeva, G., Marinova-Djambazova, P., & Kojuharov, H. (2017). Effectiveness of psychodrama therapy in patients with panic disorders: Final results. *International Journal of Psychotherapy*, 21(2), 55-66.
- Tümlü, G. Ü., & Şimşek, B. K. (2021). The effects of psychodrama groups on feelings of inferiority, flourishing, and self-compassion in research assistants. *The Arts in Psychotherapy*, 73, 101763.
- Wang, Q. (2021). The concept of "prediction" in predictive mind. *Fujian Tribune (Humanities & Social Sciences)*, (9), 15-24.
- Wiese, W., & Metzinger, T. (2017). Vanilla PP for philosophers: A primer on predictive processing. In T. Metzinger & W. Wiese (Eds.), *Philosophy and predictive processing* (pp. 1-18). Frankfurt am Main: MIND Group.
- Yalom, I. (1995). *The theory and practice of group psychotherapy*. New York: Basic Books.
- Yaniv, D. (2012). Dynamics of creativity and empathy in role reversal: Contributions from neuroscience. *Review of General Psychology*, 16(1), 70-77.

Yon, D., & Frith, C. D. (2021). Precision and the Bayesian brain. *Current Biology*, 31(17), 1026-1032.

Žurić Jakovina, I., & Jakovina, T. (2017). Role theory and role analysis in psychodrama: A contribution to sociology. *Socijalna ekologija: Journal for environmental thought and sociological research*, 26(3), 151-169.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv –Machine translation. Verify with original.