Role-Taking Theory and its Application to Interpersonal Conflicts

Inaugural-Dissertation

zur Erlangung des Doktorgrades

der Philosophie an der Ludwig-Maximilians-Universität

München

vorgelegt von

Johannes Schwabe

aus Detmold (Lippe)

München 2019

Erstgutachter: Prof. Dr. Mario Gollwitzer

Zweitgutachter: Prof. Dr. Dieter Frey

Datum der mündlichen Prüfung: 09.07.2019

Table of Contents

1 Introduction and Summaries of the Manuscripts

1.1 Content and Contributions

How do people behave in interpersonal interactions? And why do they behave the way they do? In this thesis, I present a novel social-cognitive account of role-taking explaining how fixed positions affect our emotions, thoughts, and behavior in interactions. The basic notion of Role-Taking Theory (RTT) is that actors represent roles in form of mental schemata, and that role-taking activates the role-specific mental schema. In its application, Role-Taking Theory (RTT) was used to explain third-party reactions to conflicts. As a result, this thesis entails three manuscripts (MS) currently under review for publication in international peer-reviewed journals: one theoretical (MS 1), and two empirical (MS 2&3). In the theoretical article, we explain the core premises of RTT, discuss RTT's contributions over and above existing theory, apply RTT to interpersonal conflicts, and outline future directions. In the two empirical articles, we report first empirical tests of hypotheses deriving from RTT's assumptions. In sum, the manuscripts provide answers to (1) how the number of constituting roles can be identified in any interpersonal interaction (MS 1&2), (2) how features of identified roles can be derived (MS 1–3), (3) how features of roles affect actors due to role-taking (MS 1–3), (4) why some actors take over certain roles in conflicts more frequently than others (MS 2), (5) why some actors benefit from taking over certain roles in conflicts more than others (MS 2), (6) why actors sometimes stick to roles even when this comes with costs (MS 1&3), and – ultimately – (7) how all these mechanisms tied to role-taking explain when and why interactions stabilize or destabilize. In a final chapter at the end of this thesis, the incremental contributions of these manuscripts over and above existing literature are discussed both in its novel theoretical rationale as well as in empirical insights. In sum, I argue that RTT's universality and precision are apt for sustained empirical and theoretical work in the future. However, RTT is still in an early developmental stage, and thus, I prepare the next steps by analyzing the present state, outlining the most critical challenges in theory development, and identify the most promising directions for future research.

1.2 Chronology of the Project

Throughout this PhD project, RTT underwent an ongoing process of improvement fueled both by an ever-repeating critical assessment of its theoretical core – ultimately resulting in MS 1 – as well as by empirical tests of deriving hypotheses – resulting in MS 2 and 3. The theoretical work led to an article of its own, and thus, the theoretical introduction often found in the beginning of dissertations is already covered by MS 1. Therefore, this introduction should be read as a general foreword including summaries of the manuscripts. The reason why we tested the theory in the application context at hand – and not in any other interpersonal interaction – lies in the chronology of this PhD project. To make this chronology more comprehensible and to give readers more context for evaluating the merits of the present work, I briefly sketch the project's developmental road in this introduction.

1.2.1 Starting Point

This thesis represents a milestone on an ongoing road of developments that originally started with the following research questions: How do third-party actors react when witnessing injustice? How can their reactions be classified and explained? And, given that perceptions of injustice often lie at the heart of conflicts (Montada, 2013), and that third-parties' reactions are often decisive for the course and outcome of conflicts (Skarlicki & Kulik, 2004), how can insights about reactions to injustice help to explain the stabilization and escalation of conflicts? Prior research has mainly focused on the central protagonists of conflicts succeeding norm violations (i.e., "party A" and "party B", or "victim" and "perpetrator", Shnabel & Nadler, 2008), and thus, studies on third-party reactions are comparably sparse. Third-party studies mainly focused on specific behaviors (e.g., punishment, retribution, forgiveness, or compensation) and did not offer an integrative framework. RTT, in the first draft, has been developed with the intention to fill this gap in the third-party literature.

That is, in the beginning of the project, we had no intention to develop a generalized role-theoretical approach applicable to all interpersonal interactions. However, the idea to pursue a role-theoretical approach to answer the above questions was already incorporated in the very first project outline. For example, building on previous research investigating group phenomena stabilizing bullying behavior in school classes (i.e., the participant role approach, Salmivalli, Lagerspetz, Björkqvist, Österman, & Kaukiainen, 1995), we already assumed that third-party reactions to interpersonal conflicts could be categorized in a finite set of roles and that actors could only take one role at a time.

1.2.2 Preliminary Qualitative Study

The very first step I took to approach the above research questions was a qualitative interview study (N = 28), aimed at achieving a categorical characterization of the behavior associated with the assumed third-party roles. At the time, we assumed that third-parties could react in one of the following four ways: support one of the involved parties (i.e., victim or perpetrator), try to reconcile the conflict, or remain neutral and passive. Using a semi-standardized questionnaire, I asked participants to recall events in which they witnessed injustice (i.e., either a norm violation regarding conduct, distribution, or law), and used qualitative content analysis (Mayring, 2010) to achieve categories of behavior distinguishing the corresponding four roles victim supporter, perpetrator supporter, arbitrator, and bystander. However, although this preliminary qualitative step was of great practical value for many later stages of the project (e.g., when creating ecologically valid vignettes), its empirical value is more of descriptive character and in hindsight did not appear substantial enough to legitimate publication. Nevertheless, it is mentioned here for chronological completeness and interested readers should contact me for further information on results.

1.2.3 Early Theoretical Developments

Sometime around this point, we sensed that none of the candidates we had in mind as potential overarching role-theoretical frameworks for our project actually had the social-

cognitive answers we were looking for. That is, although the role-concept had been widely used in social psychological and sociological theory (e.g., Mead, 1934; Parsons, 1951; Salmivalli et al., 1995; Stryker & Burke, 2000), none of the existing approaches actually explained the social-cognitive process underlying role-taking. In other words, all these approaches failed to recognize that role-taking actually *is* the phenomenon which requires an explanation from a social-psychological standpoint. Which social-cognitive processes affect actors once they identify themselves with a social role? RTT delivers a comprehensive framework for all roles answering this question, and this is the central theoretical contribution of this PhD project. We began to develop terms through which the role-taking process could be explained from a social-cognitive perspective. We defined a social role as the mental schema containing all associations an actor ties to a position in an interaction, and role-taking as the activation of this mental schema (see MS 1). However, at that point in time, the scope of the theory was still limited to third-party reactions to injustice, and the approach was termed “Moral Roles Theory”.

1.2.4 First Set of Studies: Explaining Third-Party Reactions to Conflicts

The first line of experimental studies testing empirical hypotheses deriving from the framework was carried out in form of quantitative and (quasi-)experimental online studies. At that point, we went beyond mere classification of third-party reactions and asked ourselves how role-taking as a social-cognitive process would affect third-parties, for example, whether role-taking affects the situational self-concept (or “working self-concept”, Markus & Kunda, 1986), and which mechanisms specifically related to role-taking could explain why some third-party actors take over pro-social conflict roles more often than others. The corresponding studies found their way into this thesis in form of MS 2. Thus, MS 2 entails the chronologically oldest studies, which are based on the earliest versions of RTT.

In the two studies, we tested a central assumption derived from our approach – that role-taking in conflicts is intricately intertwined with the moral self-concept. We hypothesized

that (1) pro-social role-taking (i.e., as an arbitrator or victim supporter) is predicted by individual differences in general moral self-concept relevance; (2) taking over a pro-social role increases the situational moral self-concept; and (3) the more relevant the general moral self-concept for an actor, the higher the situational moral self-concept increase after pro-social role-taking. The results from both studies (total N = 731) support these hypotheses. In sum, we interpret the results as indicative for a motivational mechanism: pro-social role-taking causes an increase in situational moral self-concept and the more central the moral facet for the general self, the more actors are motivated to reap this situational benefit. Theoretically, these studies already incorporated an important conceptual separation: RTT – in contrast to its closest role-theoretical relative Identity Theory (IT, e.g., Stryker & Burke, 2000), extends the role-concept to roles in short-termed interactions and links role-taking effects to the situational self-concept and not the general identity of actors (see MS 1 for a more detailed separation of the two approaches).

1.2.5 Extending Universality: Developing Role-Taking Theory

At that time, I became increasingly drawn into the development of our role-theoretical approach, as we faced many challenges on the way to a first article version of the theory. Most importantly, we wanted to make sure that what we had in mind was not already said and done elsewhere. For instance, we wondered whether role-taking was already covered under the umbrella of self-categorization theory (SCT, Turner, Hogg, Oakes, Reicher, & Wetherell, 1987). SCT uses identification processes related to groups such as self-stereotyping or depersonalization to explain intergroup behavior (Turner et al., 1987). However, although RTT and SCT share a central feature in that they both are concerned with identification processes related to social categories, the group-level processes assumed by SCT do not fit the individual character of social roles (see MS 1). This pattern, that is, some similarity with existing theory but also a clear-cut feature separating RTT from it, was the result of many literature reviews during that time.

Although the concern to have overlooked something stayed with me for quite a long time, we felt confident enough to present a first version of the theory at conferences and also reached out to a journal with a first pre-submission inquiry, offering a stand-alone theoretical article. In response, we received very reinforcing feedback from the community. It was even suggested to widen the scope of our theory to such a general extent, that it would not only incorporate role-taking in interpersonal conflicts, but in all interpersonal interactions. We took this as a great opportunity and went with it. That being said, MS 1 and the final discussion (see Chapter 5) also demonstrate the challenges that come with the ambitions to specify such a universal theory. From that point on forward, the project was a balancing act: to publish the role-theoretical approach in universal form as a stand-alone theoretical article, and to continue testing central derivations empirically in parallel.

1.2.6 Second Set of Studies: When Arbitrators Prolong Conflicts

The second line of experimental studies reported in MS 3 investigated a paradox hypothesis that follows from the results reported in MS 2: if arbitrators (and victim supporters) are motivated to experience themselves as moral persons during the conflict, and if arbitrators lose this ongoing benefit of their role once the conflict is over (because the roles are embedded in the conflict situation), this should motivate arbitrators to keep conflicts going. That is, the increase in moral situational self-concept due to pro-social role-taking can undermine the ostensible role goal of arbitrators – to solve the conflict – because arbitrators do not want to lose this role-specific benefit. Data from two studies – a correlational online study (N = 170) and an experimental lab study (N = 107) – support the hypothesized indirect effect where pro-social role-taking motivates third-parties to prolong conflicts through benefitting the moral self-concept. This effect derived from the RTT framework helps to explain why conflicts can be hard to solve even when third-parties in pro-social roles are involved in conflicts. On a broader conceptual level, it demonstrates how actors can be

motivated to keep interactions going because they benefit from the role they took in the interaction.

1.2.7 Role-Taking Theory: Present State and Future Directions

Together with MS 2, the second line of empirical research reported in MS 3 demonstrates parts of RTT's explanatory power. However, in an applied sense, there are still many questions left open with regard to third-party reactions to conflicts (see MS 2 & 3). More importantly, on a theoretical level, RTT entails many more hypotheses and presumptions yet to be tested empirically (see MS 1 and Chapter 5). This temporary imbalance presumably is a natural phenomenon accompanying all focused theoretical efforts leading to novel frameworks. Nevertheless, the crucial factor deciding over RTT's future impact is whether this imbalance will be equalized by well-designed empirical studies both rigidly testing the central assumptions of RTT as well as firmly demonstrating its explanatory power. For this cause, the most pressing lines of future empirical work are discussed in the Outlook Section of MS 1 and in Chapter 5.

References

Markus, H., & Kunda, Z. (1986). Stability and malleability of the self-concept. *Journal of Personality and Social Psychology*, *51*, 858-866. doi: 10.1037/0022-3514.51.4.858

Mayring, P. (2010). Qualitative Inhaltsanalyse. In G. Mey & K. Mruck (Eds.), *Handbuch Qualitative Forschung in der Psychologie* (pp. 601-613). Wiesbaden, Germany: VS Verlag für Sozialwissenschaften.

Mead, G. H. (1934). *Mind, self and society*. Chicago: University of Chicago Press.

Montada, L. (2013). Gerechtigkeitskonflikte und Möglichkeiten ihrer Lösung. In M. Gollwitzer, S. Lotz, T. Schlösser, & B. Streicher (Eds.), *Soziale Gerechtigkeit. Was unsere Gesellschaft aus den Erkenntnissen der Gerechtigkeitspsychologie lernen kann* (pp. 35-54). Goettingen: Hogrefe.

Parsons, T. (1951). *The Social System.* London: Routledge.

Salmivalli, C. (1999). Participant role approach to school bullying: Implications for interventions. *Journal of Adolescence*, *22*, 453-459. doi: 10.1006/jado.1999.0239

Salmivalli, C., Lagerspetz, K., Björkqvist, K., Österman, K., & Kaukiainen, A. (1996). Bullying as a group process: Participant roles and their relations to social status within the group. *Aggressive Behavior*, *22*, 1-15. doi: 10.1002/(SICI)1098-2337(1996)22:1<1::AID-AB1>3.0.CO;2-T

Shnabel, N., & Nadler, A. (2008). A needs-based model of reconciliation: satisfying the differential emotional needs of victim and perpetrator as a key to promoting reconciliation. *Journal of Personality and Social Psychology*, *94*, 116-132. doi: 10.1037/0022-3514.94.1.116

Skarlicki, D. P., & Kulik, C. T. (2004). Third-party reactions to employee (mis) treatment: A justice perspective. *Research in organizational behavior*, *26*, 183-229. doi: 10.1016/S0191-3085(04)26005-1

Stryker, S., & Burke, P. J. (2000). The past, present, and future of an Identity Theory. *Social*

Psychology Quarterly, *63,* 284-297. doi: 10.2307/2695840

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Cambridge, MA, US: Basil Blackwell.

2 Manuscript 1 – Role-Taking Theory

Schwabe, J., & Gollwitzer, M. (2019). *Role-Taking Theory*. Manuscript under review.

Role-Taking Theory

Johannes Schwabe & Mario Gollwitzer

Department of Psychology, Ludwig-Maximilians-Universität München, Germany

Acknowledgments

We want to thank Nadja Born and Clarissa Zwarg for proof-reading the final manuscript.

Conflict of Interest Statement

The authors declare that there is no conflict of interest in any form.

Abstract

In this article, we present Role-Taking Theory (RTT) – a novel theoretical approach shedding light on the social-cognitive effects of role-taking. RTT starts from the premise that all social interactions have a prototypical social structure with a finite set of social roles. Every actor present in a social interaction necessarily takes over one of these roles. RTT states that prototypical role-specific schemata (including role-specific goals, expectations, self-perceptions, and relational patterns to other roles) shape actors' feelings, thoughts, and behaviors due to role-taking. We argue that role-taking can be meaningfully differentiated from processes related to larger social categories (e.g., social groups), and we highlight RTT's theoretical increment over and above existing role-theory. To exemplify its explanatory power, we demonstrate RTT's applicability to behavior in interpersonal conflicts succeeding norm violations. We end by discussing practical implications that address RTT's relevance for understanding the stabilization and escalation of interpersonal conflicts.

Keywords: Social Roles, Role-Taking, Social Cognition, Social Categorization, Interpersonal Conflicts

Role-Taking Theory

How do people behave in social interactions? And why do they behave the way they do? Social psychological answers to these questions often rely on two principles. First, the social context (e.g., the salience of social categories like social groups or social roles) shapes social behavior. Second, this influence of the social context is intricately linked to a person's self-concept (Hogg, Terry, & White, 1995). That is, actors identify with social entities like groups or roles and act accordingly. Role-Taking Theory (RTT) – the theoretical approach we will describe in this article – expands prior theorizing with regard to both principles. A role is defined as the mental schema an actor ties to a position in an interaction (e.g., a negotiation, a conflict, the trading of goods, etc.). Role-taking is the activation of the mental schema: all associations tied to the role are activated once the role is taken over (i.e., role-specific goals, expectations, self-perceptions, etc.). RTT comprehensively explains a number of phenomena that only have been loosely described in the literature but have never been theoretically explicated by previous approaches. These phenomena include, for instance, (1) the observation that individuals who have taken over a particular social role behave stereotypically in accordance with role-specific expectations, (2) the observation that actors often are reluctant to change or exit their role, or (3) the observation that actors' perceptions and representations of social interactions (as well as the perceptions of other involved actors) become increasingly biased due to role-taking.

RTT explains these phenomena by capturing the multifaceted social-cognitive effects of role-taking through the activation of the mental schema associated with a given role. It is applicable to any interpersonal interaction. The goal of the article is to expand theory about identification processes related to social entities by (1) delivering a precise definition of roles, (2) specifying a process which can explain the effects related to role-taking, and (3) differentiating this process from other related processes that have been described by Social Identity Theory (SIT; Tajfel & Turner, 1979, 1986), Self-Categorization Theory (SCT;

Turner, Hogg, Oakes, Reicher, & Wetherell, 1987), and Identity Theory (IT; Stets & Serpe, 2013; Stryker, 1980; Stryker & Burke, 2000).

The present article consists of four sections. In Section 1, we give a brief overview of role theories in psychology and sociology. In Section 2, we present the premises of RTT, explain how RTT defines "social role" and "role-taking", and differentiate RTT from other theoretical approaches. In Section 3, we demonstrate how RTT can be applied to a given interaction on the example of social conflict situations, derive empirically testable hypotheses in this context, and elucidate how RTT can explain typical conflict phenomena. In Section 4, we critically assess RTT's contribution to the existing literature, discuss future directions, and end with concluding remarks.

Section 1: Social Structures, Social Roles, and Role-Taking

RTT is rooted in sociological and social psychological theory on social roles, most notably, symbolic interactionism (Blumer, 1969; Mead, 1934), and structural functionalism (Parsons, 1951; for an overview of classic role-theoretical approaches and an attempt to integrate them, see Biddle, 1986, 2013; Biddle & Thomas, 1966). In structural functionalism, a "role" is the procedural aspect of an actor's participation in a patterned interactive relationship: "what the actor does in his relations with others seen in the context of its functional significance for the social system" (Parsons, 1951, p. 16). This procedural aspect is accompanied by the positional aspect of the actor's participation in the social system (termed "status" by Parsons, 1951), that is, "where the actor in question is 'located' in the social system relative to other actors" (Parsons, 1951, p. 15). Thus, Parsons (1951) differentiates the participation of an actor in a social structure into two facets, "role" and "status", which enables him to distinctly analyze social systems in his further theory building.

By contrast, in the early versions of symbolic interactionism (e.g., Mead, 1934), a "role" was less formalized, less fixed, and rather an emergent product of an ongoing negotiation between the involved actors. Actors understand the meaning of their own and

others' social conduct through the lenses of their role schemata. The role concept is used in the description of individual processes (e.g., "role-taking" and "role-playing"), that – according to symbolic interactionism – foster social-cognitive development (e.g., the development of a self-concept). Thus, both approaches use the term "role" (or "statusrole"; Parsons, 1951) to describe a particular unit of a social system, and imply or directly argue that many social systems and social situations come with a set of social roles. However, functionalism uses the term "role" to describe a mechanistic perspective on social systems, while symbolic interactionism stresses the importance of role-taking for the individual actor and his or her development of a relation towards the social world (see Figure 1).

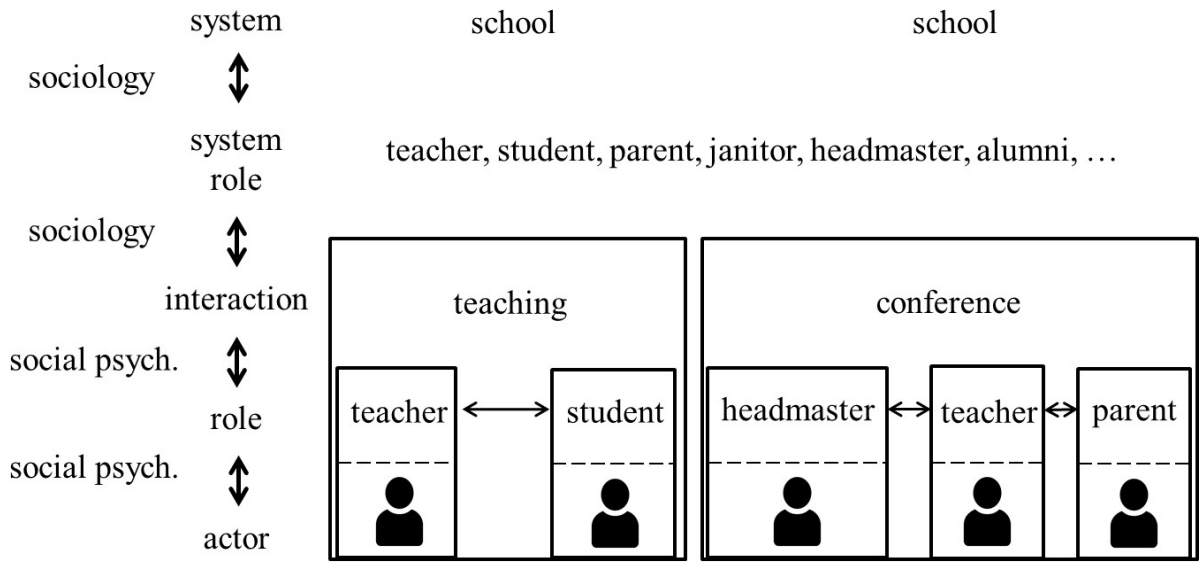

Figure 1. Typical levels of role-analysis: sociology versus social psychology.

Although heavily criticized (Biddle, 1986), both classic role theories were very influential, also in social psychology. Many social psychological approaches are rooted in one of the approaches mentioned above. For instance, the structural functionalist idea of roles as components of self-perpetuating social systems is mirrored in the "participant role approach" (Salmivalli, 1999; Salmivalli, 2010; Salmivalli, Lagerspetz, Björkqvist, Österman, & Kaukiainen, 1996). This approach aims to explain the stability of bullying behavior in school classes as a group phenomenon based on a stable system of roles and their proponents. The

symbolic interactionist view of roles as important vehicles for the development of a self-concept and further social-cognitive abilities was pursued in traditional symbolic interactionism (Blumer, 1969), and in structural symbolic interactionism (Kuhn, 1964), on which Identity Theory is mainly based (Stets & Serpe, 2013; Stryker, 1980; Stryker & Burke, 2000). Identity Theory (IT) is concerned with the relation of long-term societal roles (e.g., father, tenant, social worker) to the general self-concept: in IT terms, an actor's "identity" is the part of the self defined by the societal roles played by this actor (Stryker & Burke, 2000; Swann, Rentfrow, & Guinn, 2003). IT refers to the process connecting general identity and role identities as "self-verification:" actors try to verify certain aspects of their general identity (e.g., the moral aspect, the competent aspect, etc.) through their performance in corresponding roles (Burke, 1991; Stryker & Burke, 2000; Swann et al., 2003). Central ideas of symbolic interactionism are also relevant for social psychological research on perspective-taking, theory of mind, frame analysis, impression management, and the inclusion of others in the self (e.g., Ames, Jenkins, Banaji, & Mitchell, 2008; Aron, Aron, & Smollan, 1992; Davis, Conklin, Smith, & Luce, 1996; Goffman, 1974; Schlenker, 1980).

However, the crucial difference between sociological and social psychological role concepts is that sociology mainly focuses on the interrelation between a social role and the respective social system (i.e., a role is understood as a component of and in its relevance for a social system), while social psychology is concerned with the interrelation between a social role and the individual actor (i.e., a role is understood as an attribute of and in its relevance for an individual, see Figure 1).[1] The social-cognitive processes specifying the latter have still not been comprehensively explained. RTT is designed to fill this rather blank spot on the social psychological side. One notably extension of existing theory by RTT is that roles do not only exist as relatively long-term and situation-unspecific positions in stable societal systems (e.g., a manager, an organ donor, a voter), that may be treated as rather stable attributes of an actor,

[1] In regard of this criterion, symbolic interactionism and its theoretical heritage can also be understood as social psychological theory, while functionalism is classic sociological theory.

but as short-term-oriented, situation-specific positions in constantly and spontaneously emerging interpersonal systems (i.e., in interactions). That is, in RTT, roles are rooted in interactions. Independent of actors and situational instances, a position exists as a generalized feature of an interaction. In a given situational instance of an interaction, a role is an actor's mental representation of a position in this instance. That is, RTT captures the social-cognitive effects tied to positions in everyday interactions to contribute to a better understanding of psychologically relevant phenomena in these interactions.

Five Basic Features of Social Roles

Despite their differences, sociological and social psychological role concepts often share a number of common features. These features are also integrated in the core of RTT's social-cognitive perspective on social roles.

Self-relevance. First, roles are linked to an actor's self-concept (McConnell, 2011; Mead, 1934; Stryker & Burke, 2000). For example, being a father, a student, an employee, and so forth, is – at least usually – an important part of how people see themselves. That said, the extent to which people identify with roles varies (between people, between roles, and between situations), and roles taken over in brief social interactions (e.g., paying at a cashier while shopping as a client) are usually not self-defining parts of one's generalized or chronic self-concept. However, even short-term roles are self-relevant as they are associated with the active self-concept (Wheeler, DeMarree, & Petty, 2007; also see "working self-concept;" cf. Markus & Kunda, 1986; McConnell, 2011).

Expectations. A second feature of a "role" is that it comes with a set of socially shared normative expectations. A "father" is expected to care for his child; an "employee" is expected to do his or her work properly; an "arbitrator" is expected to be unbiased, and so forth. In fact, following a classic sociological definition (Gecas, 1982), a role is nothing else

than the sum of all expectations subsumed under its label.[2] Violating these expectations often leads to social sanctions, which can be more or less formal (e.g., formal written warnings, or informal social exclusion). Beyond external sanctions, violating role-specific expectations can also be accompanied by internal sanctions, like self-directed negative emotions such as guilt or shame (Heise, 1977, 1979; Smith-Lovin & Heise, 2016).

Individual Entity. A third feature of a social role is that this social category describes one single individual. That is, unlike social groups, this entity can meaningfully exist even if taken over by only one actor, and without further proponents of one and the same role. Although at one given point in time, many actors can take over the same role in different instances of an interaction (i.e., there are many judges in different trials, and many mothers at different family dinner tables at one given point in time), or the same role in the same instance (i.e., two friends on the phone, many bystanders of an accident), a social role is a category relating to one single individual. That is, where SIA processes always presume at least two members of a group, roles come without this assumption. Thus, actors only need the respective prototype of the role in their mind to identify with it, and no real proponents of that category. The "individual" character of social roles – in comparison to groups – also translates to the explanatory scope: while the SIA is mainly concerned with group processes and intergroup behavior, RTT focuses on individual processes and interpersonal behavior. Likewise, while SIA processes are typically associated with the collective self (Hogg, Abrams, & Brewer, 2017), role-taking in the RTT sense is concerned with the personal and relational self (Brewer & Gardner, 1996; Sedikides, & Brewer, 2001), and does not assume a shift from the "I"- to the "We"- level of self-abstraction due to identification. However, if a role becomes salient that is embedded in a group, there might be spillover effects, and vice versa. We further elaborate on potential consequences of this interesting nesting phenomenon of roles and groups in the outlook in Section 4.

[2] In an attempt to integrate sociological role concepts, Gecas (1982, p. 14) has defined a role as "the behavioral expectations associated with a position or status (either formal or informal) in a social system."

Complementarity. A fourth feature of roles is that they are typically complementary, for example: mother/child, leader/follower, student/teacher, and prisoner/guard. The complementarity of dyadic roles is systematically mutual: the leader role can only meaningfully exist in combination with the follower role and vice versa. However, when social interactions become more complex than a dyad, roles do not necessarily exist in mutual complementarity: a bystander role – for example a fan observing a fight between two boxers – cannot meaningfully exist without the boxers, but the boxer roles can meaningfully exist without the fan. Note that this necessary role feature of complementarity also does not apply to social groups. The complementarity of social roles also implies that roles are disjunctive categories: one actor can only take over one of several complementary roles in a given situation at a given time.

Mutual Reinforcement. A fifth feature is that roles often add to their mutual maintenance. In other words, roles reinforce each other, and some role-structures tend to reinforce themselves (and the interaction in which they are embedded) once they are in place. There are several mechanisms through which reinforcement can occur. One mechanism is connected to the complementarity of roles. For example, if one of two actors claims a role or does not want to exit a role, the other actor is forced into taking the complementary role for as long as the interaction lasts. A second mechanism is that a network of social roles and their proponents developed in a given interaction can become a stable feature of that group of interaction partners, that is, to retain one's role can become an over-arching group norm for every member (Salmivalli, 1999). A third mechanism is rooted in the interdependence of roles (i.e., in regard of associated goals). For positively interdependent roles (i.e., roles where the goal attainment of one role is positively related to the goal attainment of the other role, e.g., friends), one role proponent might have to compensate for the other role proponent (e.g., if one of the friends is notoriously bad in keeping in touch, the other one has to do more to reach the shared goal – to keep the friendship alive). For negatively interdependent roles (e.g., two

athletes fighting for the same award), a role proponent has to keep up with the other contestant to not lose the struggle. Thus, for example, if one of the two athletes puts more effort into the role (i.e., identifies more strongly with the role, engages into more role-consistent behavior, etc.), the other athlete is pressured to do the same. Thus, even when there is negative interdependence between complementary roles, the interaction that contains these different roles typically strives for a kind of equilibrium (regarding identification and engagement with the roles, but also resources shared between the roles like power, interests, etc.). Arguably, there are interactions marked by equilibria that are stiff, but stable, cementing ever-repeating behavioral patterns, such as the infamous fights over the same issues by old married couples.

Summary

To sum up, social roles have been a key concept in both sociological and social psychological theory to describe stable positions in social systems and interactions, and to elucidate how these positions feed back onto individual actors. However, up until now, no theory has comprehensively explained the social-cognitive effects tied to social roles. Social-psychological approaches were concerned with identification processes related to larger social categories (such as the Social Identity Approach), or the relation of long-term societal roles to the general self-concept (such as Identity Theory). Thus, it is unclear how the five features typically associated with social roles (i.e., self-relevance, expectations, individual entity, complementarity, and mutual reinforcement) play out in everyday, sometimes short-lived instances of social interactions, that is, how they affect individual actors in everyday interactions. RTT provides the answer to these questions.

Section 2: Premises of Role-Taking Theory (RTT)

RTT is built on six premises, which we will present and critically assess in the following.

Premise 1: All social interactions have a prototypical social structure consisting of a finite

and disjoint set of positions that can be adopted by individuals.

We argue that all interactions (including all interpersonal interactions that have been studied by social psychologists), such as conflict situations, exchange situations, negotiations, interpersonal trust dilemmas, friendships, and so forth, consist of a finite set of positions. This set – unique for every type of interaction – is the prototypical social structure of an interaction. At a minimum, interactions consist of two primary positions – in conflict situations, victim and perpetrator; in a trust dilemma, trustor and trustee; in an exchange situation, buyer and seller, and so forth. But many situations are more complex and consist of more than two primary positions: in negotiation situations, there may be more than two stakeholders, and their negotiation may be supervised by a neutral arbitrator; a conflict situation can include third parties beyond victim and perpetrator, and so on. The secondary positions completing the prototypical social structure of any given interaction can be derived top-down by asking how further positions can relate to the goals of the primary roles, and the interaction in total (see Section 3 for an application example). The point we want to make in this regard is that the number of these positions is always bounded (i.e., finite) and the positions are disjunctive and exhaustive: each person involved in the situation has to take a position, and he or she can take not more than one position at a given time. Together, these positions form the "prototypical social structure" of the respective interaction.

Premise 2: From repeating encounters with a social interaction, individuals abstract a mental template of the prototypical social structure of that interaction.

Premise 3: This mental template contains a generalized mental schema of the positions in the respective social structure.

The representation of the prototypical social structure of an interaction also incorporates a mental schema of all positions involved. A schema is a cognitive structure that represents information about a concept, its attributes, and its relationship to other concepts (Brewer & Treyens, 1981; Rumelhart, 1980; Zwebner, Sellier, Rosenfeld, Goldenberg, &

Mayo, 2017). For instance, a "self-schema" contains all cognitive and affective associations tied to one's identity (Markus, 1977; Markus & Wurf, 1987).

In our conceptualization, a "role" is an actor's mental schema of a position in an interaction in its entirety, that is, with all (stereotypical) expectations, self-relevant attributes, and further associations tied to this position. To exemplify, all associations tied to the "teacher" role, that is, its goals (i.e., to increase students' knowledge and skills), its relations to its complementary role (the "student"), the expectations tied to it, and the stereotypes and self-relevant attributes of this role (e.g., a teacher is a competent, knowledgeable, responsible, fair, … person) can and some of them have to be part of the mental schema belonging to the "teacher" role. It is important to note that certain associations are necessarily part of any given actor's representation of a given position (e.g., a teacher should be competent – an *injunctive norm* connected to the teacher role), while others are more open to individual interpretations of a role (e.g., teachers often wear turtleneck sweaters – a *descriptive norm* or stereotype connected to the teacher role), which can, but do not have to be, socially shared. The necessary associations of a given role and their evaluations can be derived from features of the interaction in which it is embedded. For example, interactions where the primary roles have negatively interdependent goals (e.g., conflicts, negotiations, or fights) are typically more aversive for at least one of the involved actors than interactions with positive interdependence. Thus, roles defined by the goal to end the interactions resulting from such negative interdependence (e.g., arbitrators), are typically evaluated positively on a broader level.

Furthermore, the possible actions (i.e., actors' decisions) are subsumed under a role in form of expectations (Gecas, 1982; Schank & Abelson, 1977). For instance, a typical two-person trust dilemma consists of two actors (trustor and trustee) and a relatively narrow set of decision possibilities: the trustor has something that he or she can entrust; entrusting creates an added value, and the trustee can give something back so that both the trustor and trustee

have increased their respective outcomes (cf. Berg, Dickhaut, & McCabe, 1995). The decision tree that describes the potential decisions and their consequences for all actors – the behavioral options of a trust dilemma and the script on which the dilemma is based – is easy to understand. And the more often one has encountered a similar situation in the past, the more easily understandable it becomes and the more routinely actors behave in such a situation. Very often, people have no difficulties understanding the decision tree and abstracting the resulting positions that form the prototypical social structure of a situation. It is as if there is an intuitive understanding for many social interactions even without knowing much about the specific details. This intuitive understanding of a social interaction develops into what we call a *mental template*: an abstract and prototypical version of the interactional social structure and its potential consequences that likely fits many specific situations. However, the degree to which this mental template and the contents of the embedded role schemata are explicit or implicit knowledge (or both) arguably varies between content, actors, over time, and between social interactions.

Premise 4: The mental schema of a given position consists of all associations tied to this position (e.g., self-perceptions, goals, general evaluations, behavioral options and expectations, symbols, etc.).

Stereotypical knowledge about a position does not only consist of normative expectations (i.e., how a role-holder typically behaves and how they should behave), but also of other position-specific associations that are not necessarily "behavioral" in nature. Position-specific associations can also be evaluative (e.g., victims are weak; bystanders are egoistic; sellers are untrustworthy; etc.) or resemble "mini-theories of mind" (e.g., trustors feel anxious about being exploited; arbitrators feel morally elevated by solving a conflict; buyers are eager to make a "hot deal"; etc.). Furthermore, people gain knowledge about the symbols serving as signals for a position (e.g., brides wear white dresses on their wedding day; arbitrators stand or sit between the conflict parties; etc.). These stereotypical associations

are activated when one thinks of a specific role. This notion is substantiated by a large body of empirical evidence, for example, research on automatic "moral agency" ascriptions to perpetrators and "moral patiency" ascriptions to victims of norm violations (Gray & Wegner, 2009).

Premise 5: When an actor adopts a given position, this actor's role-specific mental schema is activated. This process is defined as role-taking.

We have discussed and explained Premises 1 to 4 mainly from a "third-party" perspective, that is, from the perspective of someone who imagines, observes, or hears about a social interaction (such as a trust dilemma or a social conflict situation), but who is not directly involved in it. Premise 5 now gives an answer to what happens when a person gets involved in an interaction. Given that a social interaction consists of a finite set of positions, and based on the notion that each individual involved *has to* take a position, "getting involved" means that a person automatically adopts a certain position. According to Premise 5, taking a position activates all the role-specific associations that we talked about in Premise 4. Regarding the self-relevant content of the role schema role-taking means linking a role-specific self-schema to one's active self-concept (Davis et al., 1996). More specifically, the more salient a taken role in a given situation, the higher the salience of role-specific dimensions of self-evaluation in the active self-concept (cf. Markus & Kunda, 1986). Thus, we argue that role-taking always triggers certain role-specific self-directed cognitions and emotions, no matter how well the actor plays the role. The activation of the role-specific self-schema should always make role-specific self-dimensions more accessible (e.g., taking the teacher role should always lead to a higher accessibility of the "academic" or "knowledgeable" dimension in the active self-concept). However, the extent to which an actor perceives to perform well in a role ("role-performance", Stets & Burke, 2000) influences the self-evaluation on this dimension.

As noted above, a role can be embedded in complex long-term interactions (e.g.,

families, romantic relationships) as well as in simple short-term interactions (e.g., playing Frisbee, shopping), and important long-term roles might be relevant for the general self-concept. However, RTT argues that the self-related facet of the role-taking process in a given interactional context is identical for long- and short-term roles: role-specific mental constructs become more accessible and shape actors' goals, expectations, social perceptions, active self-concepts and – ultimately – interpersonal behavior.

Premise 6: Role-taking shapes how actors feel, think, and behave.

Finally, a sixth premise of RTT is that taking a role shapes how individuals feel, think, and behave in a given situation. Given what we know about the effects of activated mental constructs on psychological states and behaviors (e.g., Bargh, 1990; Bargh, Raymond, Pryor, & Strack, 1995; Galinsky, Gruenfeld, & Magee, 2003; McConnell, 2011; but see also Doyen, Klein, Pichon, & Cleeremans, 2012; Harris, Coburn, Rohrer, & Pashler, 2013, for critical comments and failed replications), this is rather a corollary than a premise: role-taking makes actors feel, think, and behave in accordance with (and sometimes even in contrast with) role-specific goals and other role-related attributes.

Just as attitudes, stereotypes, or norms, role-specific attributes are learned either explicitly (e.g., features of the "teacher" role are at least partly defined by laws) or implicitly (e.g., via observational, implicit, incidental, associative learning or behavioral reinforcement). Thus, some role-specific features can be easily verbalized in a propositional form (e.g., "teachers should support students"), while other features are mere associations, which are harder to verbalize, to justify, and to question. Nevertheless, building on associative network models (e.g., Anderson, 1983), these associations are automatically activated by taking over a role.

The extent to which associations prime an actor's feelings, thoughts, and behaviors can vary between individual actors, associations, and interactions. Moreover, the extent to which the influence of role-specific attributes occurs via deliberative/thoughtful vs.

associative processes (cf. Strack & Deutsch, 2004) depends on the specific circumstances. First, actors are arguably more aware of explicitly codified role expectations (e.g., how to dress as a lawyer at a trial) than of implicit attributes connected to a role (e.g., how a father should talk to his daughter in a supermarket); explicit attributes are, thus, more likely to be processed deliberately, while implicit attributes are more likely to be processed automatically. Second, role attributes that are particularly salient to a specific person in a specific situation are more likely to be processed deliberately (e.g., a person on a diet is more sensitive to and reflects more on associations concerning food, cf. Strack & Deutsch, 2004). Third, as interactions require or allow more vs. less time for deliberation and reflection, the extent to which role attributes shape behavior consciously vs. automatically depends on constraints like time pressure (e.g., a team of cooks in a busy restaurant kitchen vs. pen pals) and other features of interactions that put actors under very high or very low stress or cognitive load (cf. Strack & Deutsch, 2004).

Furthermore, role-taking does not only affect self-representation, but also other-representation. Roles always have counterparts, and they always draw their meaning from them (see feature “mutual complementarity”). Thus, taking a role always implies ascribing the complementary role(s) to others. In that regard, we assume that the same basic cognitive principles (e.g., accentuation and contrast effects, Eiser, 1971; Krueger, Rothbart, & Sriram, 1989; Tajfel & Wilkes, 1963) underlying many effects of identification with larger social categories on the self (e.g., self-stereotyping, Hogg & Turner, 1987), also apply to role-taking.

To illustrate the role-taking process, categorizing oneself as a “teacher” (and thus, the other interaction partner as a “student”) can influence the interpretation of ambiguous stimuli in a conversation. Imagine a group of doctoral students and their advisor sitting in a bar, having a drink. In many previous interactions between the individuals, the advisor typically had a “teacher” role, and the doctoral students took a “student” role, and thus, it is hard for them to shrug these roles off, even in a bar. However, they know each other for many years

now, and with an increasing number of drinks, the topics become more private. This might lead some of the doctoral students to rather represent the interaction in line with the current more private context than in line with previous interactions, and maybe even represent themselves and everybody else, including the advisor, as having the same role right now: "friend".[3] One of the doctoral students then may say something like: "Don't know much about history." If the advisor still is in the "teacher" role, he or she might perceive this statement in line with the "student's" role: as a call for a sermon about historical events. However, if the advisor also already switched into the "friend" role, he or she might rather interpret the line as a random contribution to small talk and answer with the second line of "What a wonderful world this could be."

Summary

To sum up, we conceptualize a social role as the mental schema containing all associations an actor ties to a given position in a given prototypical social structure. Furthermore, role-taking is defined as the activation of the mental schema associated to the specific role an actor takes. Thus, the role-taking process can be related to the five features of social roles mentioned before. *First*, taking a social role – both short-term and long-term roles – is associated with the actor's self-concept. Long-term roles are associated with the generalized or chronic self-concept as well as with the active self-concept: by choosing to accept a position in a company, this position (and its features) will become part of one's general self-concept in one way or another (Stryker & Burke, 2000), and by getting involved in a concrete interaction in which a long-term role is relevant, role-specific cognitions and emotions will become more accessible in the active self-concept. Short-term roles are usually not relevant for the chronic self-concept, but RTT argues that the role-taking effect in a given situation is the same for short-term and long-term roles: both affect the active self-concept. The extent to which role-taking positively reflects on the self – for example, increases state

[3] Arguably, a typical effect after a number of drinks.

self-esteem – depends on how well the expectations associated with a role are fulfilled (role-performance, Stets & Burke, 2000). *Second*, taking a social role means subjecting oneself – that is, one's actions, one's attitudes, one's goals, etc. – to role-specific expectations. Role-taking means that violating such expectations will likely be sanctioned or will result in losing the role. *Third*, role-taking affects the *personal* and *relational* (active) self-concept, and – in distinction to depersonalization (SCT, Turner et al., 1987) – does not cause a shift in self-construal from the personal 'I' to a social 'We' identity. *Fourth*, one actor can only meaningfully take over or keep a role when the complementary role is occupied as well. *Fifth*, taking a social role affects the social equilibrium between roles. That is, a once stable interaction can become unstable with the appearance of an additional actor, and vice versa.

Summing up the differences and similarities to Social Identity Theory (SIT), Self-Categorization Theory (SCT), and Identity Theory (IT; cf. Hogg et al., 1995; Stryker & Burke, 2000; Turner et al., 1987), RTT is concerned with sometimes short-lived instances of social categories anchored in interpersonal interactions, not necessarily relevant on a societal level (IT), while explaining individual processes and interpersonal behavior rather than intergroup behavior (SIA). In sum, RTT explains the individual social-cognitive processes underlying role-taking in given instances of an interaction, while IT describes the ongoing engagements of actors in long-term roles and their interrelations with the general self-concept, and the SIA describes effects of identification with a social group. Regarding the relation to the self-concept, RTT relates roles to the individual and the relational self (Sedikides & Brewer, 2001), instead of the collective self (cf. SIA).

Section 3: Application Blueprint – Role-taking in Interpersonal Conflicts

In the following, we apply RTT to the domain of interpersonal conflicts. We derive the prototypical social structure (i.e., the finite set of roles in this interaction), the core content of the role-specific schemata in this context, and exemplify how role-taking effects related to these schemata can be specified.

How to derive the Prototypical Social Structure

Picture the following situation in your team at work: two of your colleagues, David and Marie, do not talk to each other anymore, because Marie accuses David of always leaving a mess in the shared kitchen. For you, Marie (the accuser) takes the victim role, and David (the accused) takes the perpetrator role in a conflict evolving out of a perceived norm violation. Indeed, the mental template through which humans intuitively understand and represent norm violations is a prototypical perpetrator-victim dyad (Gray, Waytz, & Young, 2012). In line with that reasoning, people automatically infer the existence of a victim, if there is a perpetrator and vice versa (i.e., "dyadic completion;" Gray, Schein, & Ward, 2014). For example, participants who suffer from unfair distributions in economic games (e.g., the dictator game) are more likely to infer an intentional agent behind the distribution (Morewedge, 2009). Furthermore, the "victim" role is automatically associated with inferiority, powerlessness, and emotional sensitivity ("moral patiency"), while the "perpetrator" role is automatically associated with superiority, power, and insensitivity ("moral agency;" Gray & Wegner, 2009). Thus, people follow the principle of mutual complementarity when representing norm violations and they automatically activate role-specific schemata when confronted with proponents of the two primary roles of this interaction. However, victim and perpetrator are not the only roles that exist in social conflicts. Whenever third parties are present in a conflict situation, there are more roles to be filled.

Scholars have used the term "conflict" in a variety of meanings (Bar-Tal, 2000; Rubin, Pruitt, & Kim, 1994), and draw upon several classification features which differentiate one conflict from another (e.g., type of social units that are in conflict, conflict matter, etc.). RTT deals with interpersonal encounters, and we thus chose an interpersonal definition for the present article. In the following, a "conflict" is any manifest negative interdependence of goals or interests between at least two individuals (Deutsch, 2006), resulting from a perceived

norm violation (cf. Montada, 2013). The interactional starting point of such a conflict typically consists of one actor (e.g., the "victim") accusing another actor (the "perpetrator") of a norm violation (e.g., a norm of justice, fairness, distribution, conduct, etc.), and voicing some kind of demand (e.g., "I want you to stop doing this;" "I want an apology;" "I want a compensation," etc.), which the other party (e.g., the perpetrator) may be reluctant to grant. Conceptually, perceived norm violations are thus regarded as a key feature of the representation of this interaction.[4] Using this definition leaves us with two primary roles in the representation of the interaction: victim and perpetrator.

The next step is to derive the set of further positions, that is, of third parties potentially getting involved in the interaction, and associations tied to these positions. Independent of the type of interaction, this can be done by mapping the possible relations between a new role and (1) the roles that already exist in the situation (i.e., victim and perpetrator), and (2) the interaction itself (i.e., the conflict interaction constituted by the primary roles). Mapping all possible goals that new roles may want to achieve in the interaction results in a two-dimensional system in which new roles can be located, with the dimensions (1) *partiality* with three manifestations (i.e., neutral, victim side, perpetrator side) and (2) *maintenance* with three manifestations (i.e., keep the interaction going, end the interaction, or neither of the two). Thus, conflicts consist of a finite set of seven roles: two primary roles (i.e., victim and perpetrator), and five third-party roles (i.e., victim supporter, perpetrator supporter, arbitrator, escalator, and bystander, see Figure 2).

This finite set implies that each individual present has to take one role in a conflict situation: as soon as a third-party actor enters a conflict, every action (or non-action) is an act of positioning in the two-dimensional space defined by "partiality" and "maintenance" (see Premise 1). In the case of Marie's and David's conflict about the kitchen mess you – as a

[4] We are aware that many scholars would defy this notion and argue that this is not a necessary feature in representing a conflict (Deutsch, 2006). Therefore, it is important to note, that we do not claim validity of the following for all interactions with a negative interdependence, but for all interactions in which this negative interdependence is a result of a perceived norm violation in the eyes of the individual of interest.

third-party – cannot *not* take a role:[5] you are forced to position yourself as supporter (of either Marie or David), arbitrator, escalator, or bystander. Although each third-party actor has to take a role, some of these roles are more and some are less likely to be chosen and some may even appear implausible at first glance (e.g., the "escalator" role). That said, the seven roles described here map the entire domain of roles in conflict situations, and RTT assumes that people engage in the same kind of mental abstraction that we have described here (i.e., mapping relations and goals). This abstraction process takes place intuitively (schematically) for any relevant social interaction.

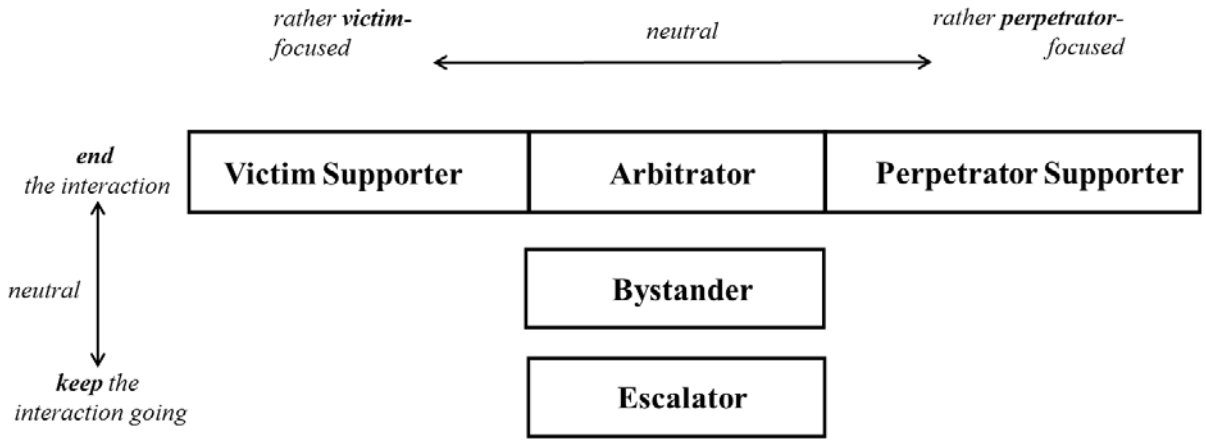

Figure 2. Schemata of third-party roles in conflicts following RTT.

The Core Content of Role-Specific Mental Schemata

As noted above, people build their role-schemata bottom-up (see Premise 2). That is, the associations tied to a given position can be manifold, and some may differ between actors due to different learning experiences with a position. However, certain associations are – from RTT's view – necessarily tied to certain positions by all actors, because they directly follow from the interactional context in which these positions are embedded. More specifically, the defining aspects of a "conflict" interaction (e.g., a perceived norm violation, primary roles perceive to have negative interdependent goals) shape role schemata (i.e., translate into general evaluations, expectations, and self-schemata tied to the positions), because they are

[5] This alludes to Paul Watzlawick's famous first axiom of communication: "You cannot not communicate." (Watzlawick, Beavin, & Jackson, 2011, p. 49).

universal features of that interaction. In the following, we will focus on these “shared” associations (e.g., “injunctive norms”) and their consequences, and will not go into detail about effects of any “secondary” associations (e.g., “descriptive norms”).

Arguably, interactions defined by negative goal interdependence are perceived as more aversive than interactions defined by positive goal interdependence. In addition, in interactions with negative goal interdependence, interaction partners perceive and evaluate each other more strongly on the “competence” than on the “warmth” dimension, while the opposite is true in interactions with positive goal interdependence (e.g., friendships, romantic relationships, families, see Fiske, Cuddy, & Glick, 2007). For the two primary roles in conflicts (i.e., victim and perpetrator), this has already been empirically established: victims are seen as non-agentic and sensitive, and perpetrators are seen as agentic and insensitive (“moral typecasting;” Gray & Wegner, 2009). The moral typecasting effect does not only exist among uninvolved observers, but also among victims and perpetrators themselves (“moral transformation”, Gray, 2010). Expanding this notion, RTT assumes that the same kind of “typecasting” also occurs for the other roles involved in a conflict. The main RTT hypothesis in this regard is that third-party roles in conflicts are always morally charged (i.e., they have associations tapping into the moral domain), because they draw their meaning from how they relate to a perpetrator-victim dyad. In a given situational instance, the more the victim’s demands are morally justified, (1) the more moral the victim supporter is perceived to be, (2) the more pressure for justification and potential moral condemnation on the perpetrator supporter, and (3) the more morally relevant is what arbitrators, escalators, and bystanders do – with the arbitrator role perceived as most prosocial, the escalator role perceived as most antisocial, and the bystander role perceived as antisocial or neutral, depending on the conflict’s destructive potential.

Implications and Empirical Hypotheses

RTT can explain (1) why people in conflicts often stop communicating with each

other, (2) why people often perceive proponents of other roles as “closed-minded” (and why this is often true), (3) why all actors (even arbitrators) have a motivation to keep a conflict going, (4) why third-parties’ perceptions of a conflict tend to be biased after role-taking, and (5) why conflicts sometimes become “intractable”, that is, why they tend to stabilize and escalate so easily. In this section, we will explain each of these five phenomena against the conceptual background of RTT.

How role-taking explains (1) diminished and dysfunctional interaction and (2) reduced perspective-taking

Two typical phenomena in destructive conflicts are (1) diminished and/or dysfunctional interaction (e.g., communication) between the conflict parties, and (2) a lack of perspective-taking. When you as a freshly involved third-party ask one of the proponents of the primary roles why the conflict has not been solved yet, they typically answer something like: “The other side is stubborn. There is no sense in talking to them about it. I really do not understand what they are thinking.” This perception is often caused by a typical effect of role-taking: role-taking accentuates perceptions of self and others in line with role-contingent categories, especially in negatively interdependent interactions. This leads actors to perceive differences in an exaggerated manner (Krueger et al., 1989). Since perceived similarity covaries with perspective-taking (Cialdini, Brown, Lewis, Luce, & Neuberg, 1997; Davis et al., 1996), role-taking tends to make perspective-taking harder. In conflicts, this effect is especially severe, because, as noted above, roles in conflicts are necessarily associated with moral evaluations (DeScioli & Kurzban, 2013; Gray & Wegner, 2009; Schwabe & Gollwitzer, 2019a). And since actors are highly motivated to avoid threats to their self-evaluation on this dimension (Brambilla & Leach, 2014) and due to the fact that moral convictions are perceived as non-negotiable (Skitka, Baumann, & Sargis, 2005), actors’ behavior in conflict situations often escalates the conflict rather than creating consensus.

How role-taking explains (3) the motivation of actors to keep a conflict going

As noted above, role-taking in conflicts allows actors to perceive themselves as moral or prosocial persons. This is especially true for the roles of victim, victim supporter, and arbitrator. Also, actors in sided roles typically perceive themselves to be on the victim side (victim or victim supporter), and neutral actors typically perceive themselves as doing the best to solve the situation (as an arbitrator, Schwabe & Gollwitzer, 2019a). That is, the large majority of involved actors subjectively perceives themselves to be in a moral or prosocial role in conflicts (Schwabe & Gollwitzer, 2019a). As noted above, people are highly motivated to perceive themselves as moral and prosocial. Thus, actors have a motivation to keep their roles. Unfortunately, since the roles lose their meaning once the interaction is over, a solution of the conflict threatens the beneficial effect of role-taking. That is, although the interaction ("conflict") itself is typically evaluated negatively, the embedded roles are typically associated with positive effects for the individual actors. First data collected are compatible with this hypothesis (Schwabe & Gollwitzer, 2019b).

How role-taking explains (4) unfavorable biases in conflict perceptions

As noted above, the effects of role-taking are moderated by how an actor subjectively perceives certain situational characteristics, for example, the distribution of proponents of other roles. In addition, actors evaluate the extent to which other actors match their respective role prototype and fulfill role-specific expectations. This evaluation of actors' role-fit is an important feature of role-taking. The extent to which others fulfill their role expectations directly reinforces, corroborates, and stabilizes one's own role and vice versa. Perceiving others to play their role badly can threaten one's own role and, thus, diminish the beneficial effects tied to one's role. Thus, the more an actor is motivated to profit from role-taking (e.g., with regard to one's moral self-concept), the more he or she prefers others to act in accordance with their role prototypes. These motivated perceptions are likely to be shared among all proponents of the same role (and among proponents of one's supporter roles), which creates an "echo chamber" that reinforces and stabilizes such biased perceptions of

other role proponents. This, in turn, makes conflicts hard to be solved.

How third-party role-taking explains (5) the stabilization of conflicts

Individuals balance two opposing needs when "managing" their self-concept: a need for inclusion (i.e., belonging, assimilation, inclusion, immersion) and a need for differentiation from others (e.g., Brewer, 1991; Fromkin & Snyder, 1980). Since role-taking affects the self-concept, these two needs and the resulting consequences for "identity management" should apply to role-taking effects as well. Interestingly, in the context of roles, third-parties can differentiate themselves best from others by taking a role that nobody else has taken before. That is, the degree to which the need for differentiation can be satisfied depends on the distribution of other role-proponents, and results in a "take-the-free-seat"-motivation. For example, the morality-boosting effect tied to taking a moral role (such as victim supporter; Schwabe & Gollwitzer, 2019a, 2019b) is likely to be higher (1) the more perpetrator supporters are already present, and (2) the fewer victim supporters are already present (see features "complementarity" and "mutual reinforcement" in Section 2). This motivation, in turn, explains why conflicts develop a stable structure: actors tend to "fill up" the prototypical social structure. In other words, the effect of role-taking on the active self-concept is moderated by the existing structure of role proponents; and this can contribute to an equilibrium of forces in a conflict, creating a stable social structure. Whenever a new actor joins a conflict, the combination of the need for differentiation and the existing role-structure "pulls" them towards vacant roles.

Interestingly, this mechanism is incompatible with what we know about the psychological processes underlying the "bystander phenomenon" (i.e., the more bystanders are present in a helping situation, the less likely a given bystander will help; Latané & Darley, 1968), and it apparently contradicts other research showing that people tend to align their actions more with a majority than with a minority (Asch, 1951, 1955; Thompson & Fine, 1999). Hundreds of studies in social psychology show how strong the pressure to conform

with the majority opinion and the reluctance to join a minority perspective are, and RTT does not deny the validity of these effects. However, RTT argues that if actors think they can reap the benefits of taking the free seat at low (social) cost, then they will be more likely to take it even though the majority pressures them to do otherwise. Future research should delineate the specific predictors for a "take-the-free-seat" motivation vis-à-vis a conformity motivation in the context of multiple-role settings (e.g., conflicts) in more detail. Building on an identity management perspective, one hypothesis in that regard is that the more one is motivated to differentiate oneself from others, the stronger is the "pull"-effect of an open role.

Summary

Summing up the consequences of role-taking in conflicts, the process offers explanations for diminished and destructive communication, reduced perspective taking, motives to keep a conflict going, motivated biases in conflict perceptions, and the stabilization of conflicts. More specifically, RTT elucidates why interpersonal conflicts are often hard to reconcile, even – and sometimes particularly – when third-parties are involved. Importantly, some of these effects occur spontaneously and without explicit self-regulation, while other processes are more deliberate and willful. As a consequence, stereotypical conflict perceptions can cause a negative conflict dynamic, because actors are more strongly perceived as proponents of their respective roles and the conflict is perceived to have a solid social structure. As outlined above, this biased perception can create a vicious cycle of mutual reinforcement of role-proponents between the two conflict sides, recreating and cementing the prototypical social structure. Ultimately, these processes of social exchange affect actors in their conflict behavior and explain why reconciliation in conflicts is often so difficult.

Section 4: General Discussion

Critical Summary

Social roles have always been a key concept for social psychologists and sociologists (Mead, 1934; Parsons, 1951; Kuhn, 1964; Stets & Serpe, 2013; Stryker, 2008; Stryker &

Burke, 2000). However, “roles” as a relevant category for social cognition have been somewhat overlooked by social-cognitive psychology so far, resulting in a lack of empirical research. RTT is aimed to remedy this issue. In the novel social-cognitive approach outlined here, we build on the idea that actors use social roles not only descriptively, but that role schemata affect our everyday behavior through a role-taking process. In RTT terms, a role is the mental representation an actor has about a fixed position in an interaction. Role-taking is the activation of a mental schema containing all associations tied to a given position in a given prototypical social structure. RTT explains how roles affect an actor’s cognitions, emotions, and behaviors. Applied to norm conflicts, we have demonstrated how actors in conflicts necessarily adopt exactly one out of seven roles in their representation of the interaction: victim, perpetrator, victim supporter, perpetrator supporter, arbitrator, escalator, and bystander. Underlining the five features of social roles (i.e., identity-relevance, expectations, individual entity, complementarity, and mutual reinforcement), the process of role-taking can harden the social structure of conflicts, explaining stabilization and escalation.

To critically assess RTT’s contribution over and above existing theory, it first needs to be evaluated against previous role-theoretical approaches. To our knowledge, RTT is the first approach designed to capture social-cognitive processes that occur when individuals take over a specific role in an interaction (see also Hogg et al.’s critique of IT in this regard; Hogg, et al., 1995). These processes are similar but have to be separated from those investigated in the domain of social identity and self-categorization theory. Second, RTT is an attempt to integrate classic sociological role theories (e.g., symbolic interactionism, Mead, 1934; structural functionalism, Parsons, 1951) with modern social-cognitive theorizing. Doing so, RTT acknowledges both the stable nature of interactional structures as well as the importance of subjective representations of these structures when it comes to individual behavior. Third, RTT is broadly applicable and can (hopefully) contribute meaningfully to a better understanding of complex and intricate social phenomena in a variety of contexts. Thus, RTT

has a high level of universality (Popper, 1934/2005).

Having said that, the high degree of universality also creates challenges for RTT, for example, for the degree of precision in regard of the role-taking process. For instance, regarding the RTT assumption that role-taking activates *all* associations tied to a position, a critical mind could demand RTT to exactly postulate which associations *are* and which *are not* involved in role-taking processes (i.e., demand a finite set of all types of associations possibly tied to positions in interactions). A further, even more challenging demand could be to rank them in importance, and to specify how exactly an accessible association from a certain area (e.g., self-relevant cognitions and emotions, behavioral expectations, symbols) affects behavior. In other words, the current state of our theorizing does neither deliver a finite framework of all areas of associations possibly involved in role-taking nor a model specifying how associations from different areas possibly interact once activated through role-taking. Future research on RTT may suggest that the theory requires more refined concepts and stronger assumptions in that regard.

In sum, RTT's ambitious effort to clarify the role-concept as well as the aligning process in an over-arching manner for all interpersonal interactions comes with optimistic promises, but also with some challenges. Nevertheless, at least in the domain of interpersonal conflict situations, RTT has been shown to be helpful in order to understand why conflicts tend to become hard to solve. More precisely, RTT allows deriving empirical hypotheses that are novel and partly incompatible with other theoretical approaches. Although empirical evidence is currently sparse, first data are in line with these empirical hypotheses (Schwabe & Gollwitzer, 2019a, 2019b).

Outlook

We hope that this article motivates social psychologists to take a closer look at role-taking. In the following, we describe what we think are the most pressing questions and the most promising avenues for further research in this regard.

"Pull" Effects. RTT argues that actors can – under certain circumstances – be "pulled" into roles. Pull effects can take place both with regard to role choice (i.e., actors can be pulled into choosing a role, see for example "take-the-free-seat" motivation in Section 3) as well as to consistency of feelings, thoughts, and behavior with a role after role choice (i.e., actors' experiences and behaviors can become more and more consistent with the role schema over time). In the most extreme case, both effects take place simultaneously: actors internally adapt more and more to an involuntarily chosen role until – at one point – they start playing this role out of an internal motivation. In other words, role-taking can occur automatically, involuntarily, sometimes even against one's better interests. A good example for being involuntarily pulled into a role has been given in Berne's (1964, 2011) influential book "Games people play", in which he describes transactional analysis as a psychotherapeutic approach. Here, he describes "ego states" (e.g., child, parent, and adult), which are mind-sets that are often tied to roles in a "family" interaction. It is important to note that an ego state is not the same as a role: ego states can be part of any role schema (i.e., the "child" ego state can also be part of the "lover" role). However, the psychoanalytic aura of transactional analysis notwithstanding, Berne (2011) convincingly describes how quickly and automatically people activate a role schema (incorporating an ego state), how immediately the ego state shapes thinking, feeling, and acting in a situation, and how strongly people can be pulled into a role even if taking it causes aversive consequences. For instance, Berne (2011) describes how hard it can be to not activate the "child" ego state when communicating with one's parents: in a role not chosen freely, actors are "pulled" into a certain mindset because of role-specific associations. RTT argues that the same dynamics can operate in all social systems. Under certain circumstances, actors are pulled towards a specific role and activate role-specific associations, even if they do not necessarily want to. These "pull" effects of roles have hardly been understood so far, and we believe that RTT provides a theoretical framework to explain and analyze such effects much better than before.

Role-Making. In most of the examples we used here, we deliberately chose roles for which the behavioral expectations that are tied to it are clearly specified. This is true for the teacher role, the trader role, the arbitrator role, et cetera. But it is a truism that roles differ in the extent to which they provide space for individual arrangements. The teacher role can be played in many nuances – the "stern teacher," the "supportive teacher," the "cool teacher," and so on. In general, roles are not only being taken and then played uniformly, roles are also filled with individuality. In other words, there are "strong" and "weak" roles similar to "strong" and "weak" situations (Mischel, 1977): weak roles allow more degrees of freedom (i.e., allow more "role-making", Mead, 1934) than strong ones. Arguably, the more actors engage in role-making, the more likely they identify with their role. This suggests that there are two ways of reaping self-related benefits from playing a role: one is the extent to which an actor fulfills the expectations that are defined by their role (i.e. the quality of role-taking), and the other is the extent to which one can fill a role with individual features (i.e., role-making). Future research may want investigate which of these two ways is more relevant for reaping the benefits of role-taking.

Long-Term Effects of Short-Term Roles. Identity Theory proposes that long-term roles shape the chronic self-concept (Stryker & Burke, 2000). Now, RTT anchors roles in interactions of which some are short-lived. One question that follows: can short-term roles affect the chronic self-concept as well? And, if yes, what exactly is the psychological process operating here? One very simple assumption could be that playing a short-term role repeatedly (because it is part of one's daily routine) can affect the chronic self-concept even though taking it in one of these instances is not very meaningful.

Roles and Groups. Now that we can clearly separate roles from groups, it should be possible to elucidate the added value of the "role-taking approach" for a "social identity approach" and vice versa in more detail. The central conceptual question is: to what extent do role-taking and SIA processes (e.g., depersonalization, Turner et al., 1987) differ and interact

with each other? More specifically, role-taking effects might moderate group-identification effects, and vice versa: the more one is motivated to profit from a role embedded in a group, the more one is motivated to identify with a group, and vice versa. However, there might be boundary conditions for this effect. For example, if one of the two categories (role or group) bears negative ramifications, how do actors respond to that? For instance, if an actor has a very positive role in a very negative group or interaction (e.g., you are an honorable soldier of a nation with reprehensible interests, or you are a sincere manager in a dubious company, etc.), he or she may identify more strongly with the role than with the group. Furthermore, effects distinctly related to roles and to groups arguably not only mutually moderate the initial identification processes, but also the way how people deal with downstream consequences of identification. For example, social mobility (i.e., trying to cope with negative ramifications of group-identification by trying to switch to a group with higher status, SIT, Tajfel & Turner, 1986) arguably becomes more unlikely the more beneficial the role in that group is.

Going beyond individual-level processes, a final interesting question is how groups are affected by the positions they take in (intergroup) interactions, that is, whether and how RTT's principles also apply on an intergroup level. Many of the interactions we find on an interpersonal level also exist on an intergroup level, and importantly, they have the same prototypical structure (i.e., the positions in intergroup and interpersonal conflicts, negotiations, and so on are congruent). It would be interesting to compare the effects of role-taking for individuals and groups, and investigate, for example, whether group role-taking shapes collective identity and how group role-taking trickles down to individual members, for example, through self-stereotyping processes (Hogg & Turner, 1987). Sustaining these hypotheses with empirical results would allow us to understand the relations of social structures and the self-concept as well as their downstream effects on social behavior more comprehensively.

Concluding Remarks

Supplying precise and testable hypotheses, the RTT framework constitutes a foundation for sustained empirical and theoretical work in the future. This work will shed further light on how, when, and why actors recreate prototypical social structures in interactions through role-taking, and which consequences and byproducts are caused by role-taking – not only in norm conflicts, but in every type of interpersonal interaction. These insights are of crucial importance for interventions aiming to resolve dysfunctional and aversive interactions, but will also prove beneficial to foster the cohesion of rather desirable social structures like organizational teams, romantic relationships, and families, by bolstering the underlying social interactions. Furthermore, they will help actors to engage in effective and successful interactions. We thus hope that the framework outlined here is only the first stepping stone for a surge of new research and theory, creating a new, modern, and truly social-cognitive home for a long forlorn social category: social roles.

References

Ames, D. L., Jenkins, A. C., Banaji, M. R., & Mitchell, J. P. (2008). Taking another person's perspective increases self-referential neural processing. *Psychological Science, 19,* 642-644. doi: 10.1111/j.1467-9280.2008.02135.x

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior, 22,* 261-295. doi: 10.1016/S0022-5371(83)90201-3

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology, 63,* 596-612. doi: 10.1037/0022-3514.63.4.596

Asch, S. E. (1951). Effects of group pressure on the modification and distortion of judgments. In H. Guetzkow (Ed.), *Groups, leadership, and men.* Pittsburgh, Pa.: Carnegie Press.

Asch, S. E. (1955). Opinions and social pressures. *Scientific American, 193,* 31-35. doi: 10.1038/scientificamerican1155-31

Bar-Tal, D. (2000). From intractable conflict through conflict resolution to reconciliation: Psychological analysis. *Political Psychology, 21,* 351-365. doi: 10.1111/0162-895X.00192

Bargh, J. A. (1990). Goal and intent: Goal-directed thought and behavior are often unintentional. *Psychological Inquiry, 1,* 248-251. doi: 10.1207/s15327965pli0103_14

Bargh, J. A., Raymond, P., Pryor, J. B., & Strack, F. (1995). Attractiveness of the underling: An automatic power→ sex association and its consequences for sexual harassment and aggression. *Journal of Personality and Social Psychology, 68,* 768-781. doi: 10.1037/0022-3514.68.5.768

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior, 10,* 122-142. doi: 10.1006/game.1995.1027

Berne, E. (1964). *Games people play: The basic handbook of transactional analysis.* New York: Ballantine Books.

Berne, E. (2011). *Games people play: The basic handbook of transactional analysis.* Tantor eBooks.

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

Biddle, B. J. (2013). *Role theory: Expectations, identities, and behaviors*. London: Academic Press.

Biddle, B. J., & Thomas, E. J. (1966). *Role theory: Concepts and research*. New York: Wiley.

Blumer, H. (1969). *Symbolic Interaction: Perspective and Method.* Englewood Cliffs, NJ: Prentice-Hall.

Brambilla, M., & Leach, C. W. (2014). On the importance of being moral: The distinctive role of morality in social judgment. *Social Cognition, 32,* 397-408. doi: 10.1521/soco.2014.32.4.397

Brewer, M. B. (1991). The social self: On being the same and different at the same time. *Personality and Social Psychology Bulletin, 17,* 475-482. doi: 10.1177/0146167291175001

Brewer, M. B., & Gardner, W. (1996). Who is this" We"? Levels of collective identity and self representations. *Journal of Personality and Social Psychology, 71,* 83-93. doi: 10.1037/0022-3514.71.1.83

Brewer, W. F., & Treyens, J. C. (1981). Role of schemata in memory for places. *Cognitive Psychology, 13*, 207-230. doi: 10.1016/0010-0285(81)90008-6

Burke, P. J. (1991). Identity processes and social stress. *American Sociological Review, 56,* 836-849. doi: 10.2307/2096259

Cialdini, R. B., Brown, S. L., Lewis, B. P., Luce, C., & Neuberg, S. L. (1997). Reinterpreting the empathy–altruism relationship: When one into one equals oneness. *Journal of Personality and Social Psychology*, *73*, 481-494. doi: 10.1037/0022-3514.73.3.481

Davis, M. H., Conklin, L., Smith, A., & Luce, C. (1996). Effect of perspective taking on the

cognitive representation of persons: A merging of self and other. *Journal of Personality and Social Psychology*, *70*, 713-726. doi: 10.1016/0010-0285(81)90008-6

DeScioli, P., & Kurzban, R. (2013). A solution to the mysteries of morality. *Psychological Bulletin*, *139*, 477–496. doi: 10.1037/a0029065

Deutsch, M. (2006). Cooperation and competition. In M. Deutsch, P.T. Coleman, & E C. Marcus (Eds.), *The Handbook of Conflict Resolution: Theory and Practice* (23-42). San Francisco: Jossey-Bass.

Doyen, S., Klein, O., Pichon, C.-L., & Cleeremans, A. (2012). Behavioral priming: It's all in the mind, but whose mind? *PloS one*, *7*(1), e29081. Retrieved from https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0029081

Eiser, J. R. (1971). Enhancement of contrast in the absolute judgment of attitude statements. *Journal of Personality and Social Psychology*, *17*, 1-10. doi: 10.1037/h0030455

Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*, 77-83. doi: 10.1016/j.tics.2006.11.005

Fromkin, H. L., & Snyder, C. R. (1980). The search for uniqueness and valuation of scarcity. In K. J. Gergen, M. S. Greenberg, & R. H. Willis (Eds.), *Social Exchange: Advances in theory and research* (pp. 57-75). New York: Plenum Press.

Galinsky, A. D., Gruenfeld, D. H., & Magee, J. C. (2003). From power to action. *Journal of Personality and Social Psychology*, *85*, 453-466. doi: 10.1037/0022-3514.85.3.453

Gecas, V. (1982). The self-concept. *Annual Review of Sociology*, *8*, 1-33. doi: 10.1146/annurev.so.08.080182.000245

Goffman, E. (1974). *Frame analysis: An Essay on the Organization of Experience*. New York, NY: Harper Colophon.

Gray, K. (2010). Moral transformation: Good and evil turn the weak into the mighty. *Social Psychological and Personality Science, 1*, 253-258. doi: 10.1177/1948550610367686

Gray, K., Schein, C., & Ward, A. F. (2014). The myth of harmless wrongs in moral cognition: Automatic dyadic completion from sin to suffering. *Journal of Experimental Psychology: General*, *143*, 1600-1615. doi: 10.1037/a0036149

Gray, K., Waytz, A., & Young, L. (2012). The moral dyad: A fundamental template unifying moral judgment. *Psychological Inquiry*, *23*, 206-215. doi: 10.1080/1047840X.2012.686247

Gray, K., & Wegner, D. M. (2009). Moral typecasting: divergent perceptions of moral agents and moral patients. *Journal of Personality and Social Psychology*, *96*, 505-520. doi: 10.1037/a0013748

Harris, C. R., Coburn, N., Rohrer, D., & Pashler, H. (2013). Two failures to replicate high-performance-goal priming effects. *PloS one*, *8*(8), e72467. Retrieved from https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0072467

Heise, D. R. (1977). Social action as the control of affect. *Systems Research and Behavioral Science*, *22*, 163-177. doi: 10.1002/bs.3830220303

Heise, D. R. (1979). *Understanding events: Affect and the construction of social action*. Cambridge, UK: Cambridge University Press.

Hogg, M. A., Abrams, D., & Brewer, M. B. (2017). Social identity: The role of self in group processes and intergroup relations. *Group Processes & Intergroup Relations*, *20*, 570-581. doi: 10.1177/1368430217690909

Hogg, M. A., Terry, D. J., & White, K. M. (1995). A tale of two theories: A critical comparison of Identity Theory with Social Identity Theory. *Social Psychology Quarterly*, 255-269. doi: 10.2307/2787127

Hogg, M. A., & Turner, J. C. (1987). Intergroup behaviour, self-stereotyping and the salience of social categories. *British Journal of Social Psychology*, *26*, 325-340. doi: 10.1111/j.2044-8309.1987.tb00795.x

Krueger, J., Rothbart, M., Sriram, N. (1989). Category learning and change: Differences in

sensitivity to information that enhances or reduces intercategory distinctions. *Journal of Personality and Social Psychology*, *56*, 866-875. doi: 10.1037/0022-3514.56.6.866

Kuhn, M. H. (1964). Major trends in symbolic interaction theory in the past twenty-five years. *The Sociological Quarterly*, *5*, 61-84. doi: 10.1111/j.1533-8525.1964.tb02256.x

Latané, B., & Darley, J. M. (1968). Group inhibition of bystander intervention in emergencies. *Journal of Personality and Social Psychology*, *10*, 215-221. doi: 10.1037/h0026570

Markus, H. (1977). Self-schemata and processing information about the self. *Journal of Personality and Social Psychology*, *35*, 63-78. doi: 10.1037/0022-3514.35.2.63

Markus, H., & Kunda, Z. (1986). Stability and malleability of the self-concept. *Journal of Personality and Social Psychology*, *51*, 858-866. doi: 10.1037/0022-3514.51.4.858

Markus, H., & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology*, *38*, 299-337. doi: 10.1146/annurev.ps.38.020187.001503

McConnell, A. R. (2011). The multiple self-aspects framework: Self-concept representation and its implications. *Personality and Social Psychology Review*, *15*, 3-27. doi: 10.1177/1088868310371101

Mead, G. H. (1934). *Mind, self and society*. Chicago: University of Chicago Press.

Mischel, W. (1977). On the future of personality measurement. *American Psychologist*, *32*, 246-254. doi: 10.1037/0003-066X.32.4.246

Montada, L. (2013). Gerechtigkeitskonflikte und Möglichkeiten ihrer Lösung. In M. Gollwitzer, S. Lotz, T. Schlösser, & B. Streicher (Eds.), *Soziale Gerechtigkeit. Was unsere Gesellschaft aus den Erkenntnissen der Gerechtigkeitspsychologie lernen kann* (pp. 35-54). Goettingen: Hogrefe.

Morewedge, C. K. (2009). Negativity bias in attribution of external agency. *Journal of Experimental Psychology: General*, *138*, 535-545. doi: 10.1037/a0016796

Parsons, T. (1951). *The Social System.* London: Routledge.

Popper, K. (2005). *The logic of scientific discovery*. London: Routledge.

Rubin, J. Z., Pruitt, D. G., & Kim, S. H. (1994). *Social conflict: Escalation, stalemate, and settlement*. New York: McGraw-Hill.

Rumelhart, D. E. (1980). Schemata: the building blocks of cognition. In R. J. Spiro, B. Bruce & W. Brewer (Eds.), *Theoretical issues in reading comprehension*. Hillsdale, NJ: Lawrence Erlbaum.

Salmivalli, C. (1999). Participant role approach to school bullying: Implications for interventions. *Journal of Adolescence, 22*, 453-459. doi: 10.1006/jado.1999.0239

Salmivalli, C. (2010). Bullying and the peer group: A review. *Aggression and Violent Behavior, 15*, 112-120. doi: 10.1016/j.avb.2009.08.007

Salmivalli, C., Lagerspetz, K., Björkqvist, K., Österman, K., & Kaukiainen, A. (1996). Bullying as a group process: Participant roles and their relations to social status within the group. *Aggressive Behavior*, *22*, 1-15. doi: 10.1002/(SICI)1098-2337(1996)22:1<1::AID-AB1>3.0.CO;2-T

Schank, R. C., & Abelson, R.P. (1977). *Scripts, plans, goals and understanding: An inquiry into human knowledge structures.* Mahwah, NJ: Lawrence Erlbaum.

Schlenker, B. R. (1980). *Impression management: The self-concept, social identity, and interpersonal relations.* Monterey, CA: Brooks/Cole Publishing Company.

Schwabe, J., & Gollwitzer, M. (2019a). *Explaining third-party behavior in norm conflicts: A role-taking approach*. Manuscript under review.

Schwabe, J., & Gollwitzer, M. (2019b). *When arbitrators prolong conflicts: Beneficial role-taking effects motivate resistance against conflict resolution*. Manuscript under review.

Sedikides, C., & Brewer, M.B. (Eds.). (2001). *Individual self, relational self, collective self.* Philadelphia: Psychology Press.

Skitka, L. J., Bauman, C. W., & Sargis, E. G. (2005). Moral conviction: Another contributor

to attitude strength or something more? *Journal of Personality and Social Psychology*, *88*, 895-917. doi: 10.1037/0022-3514.88.6.895

Smith-Lovin, L., & Heise, D. R. (Eds.). (2016). *Analyzing Social Interact: Advances in Affect Control Theory*. Abingdon-on-Thames: Routledge.

Stets, J. E., & Burke, P. J. (2000). Identity Theory and Social Identity Theory. *Social Psychology Quarterly*, 224-237. doi: 10.2307/2695870

Stets, J. E., & Serpe, R. T. (2013). Identity Theory. In J. DeLamater & A. Ward (Eds.), *Handbook of Social Psychology* (pp. 31-60). Dordrecht: Springer.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*, 220-247. doi: 10.1207/s15327957pspr0803_1

Stryker, S. (1980). *Symbolic interactionism: A social structural version*. Menlo Park, CA: Benjamin/Cummings.

Stryker, S. (2008). From Mead to a structural symbolic interactionism and beyond. *Annual Review of Sociology*, *34*, 15-31. doi: 10.1146/annurev.soc.34.040507.134649

Stryker, S., & Burke, P. J. (2000). The past, present, and future of an Identity Theory. *Social Psychology Quarterly*, *63*, 284-297. doi: 10.2307/2695840

Swann, W.B., Rentfrow, P.J., & Guinn, J.S. (2003). Self-verification: The search for coherence. In M. R. Leary & J. P. Tangney (Eds.), *Hand**b**ook of Self and Identity* (pp. 367-383). New York, NY: Guilford.

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W.G. Austin & S. Worchel (Eds.), *The Social Psychology of Intergroup Relations* (pp. 33-47). Monterey, CA: Brooks-Cole.

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *The Social Psychology of Intergroup Relations* (pp. 7-24). Chicago: Nelson-Hall.

Tajfel, H., & Wilkes, A. L. (1963). Classification and quantitative judgement. *British Journal of Psychology*, *54*, 101-114. doi: 10.1111/j.2044-8295.1963.tb00865.x

Thompson, L., & Fine, G. A. (1999). Socially shared cognition, affect, and behavior: A review and integration. *Personality and Social Psychology Review*, *3*, 278-302. doi: 10.1207/s15327957pspr0304_1

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Cambridge, MA, US: Basil Blackwell.

Watzlawick, P., Beavin, J. H., & Jackson, D. D. (2011). *Pragmatics of human communication: A study of interactional patterns, pathologies and paradoxes*. New York: WW Norton & Company.

Wheeler, S. C., DeMarree, K. G., & Petty, R. E. (2007). Understanding the role of the self in prime-to-behavior effects: The active-self account. *Personality and Social Psychology Review*, *11*, 234-261. doi: 10.1177/1088868307302223

Zwebner, Y., Sellier, A. L., Rosenfeld, N., Goldenberg, J., & Mayo, R. (2017). We look like our names: The manifestation of name stereotypes in facial appearance. *Journal of Personality and Social Psychology*, *112*, 527-554. doi: 10.1037/pspa0000076

3 Manuscript 2 – Explaining Third-Party Behavior in Norm Conflicts

Schwabe, J., & Gollwitzer, M. (2019). *Explaining third-party behavior in norm conflicts: A role-taking approach*. Manuscript under review.

Explaining Third-Party Behavior in Norm Conflicts:

A Role-Taking Approach

Keywords: interpersonal conflicts, role-taking, social roles, morality, self-concept

Funding

This research received no specific grant from any funding agency in the public, commercial, or non-profit sectors.

Abstract

When people witness conflicts succeeding moral transgressions in their group, they can react in one of the following ways: (1) support one of the involved parties (i.e., perpetrator or victim), (2) try to reconcile the conflict, or (3) remain neutral and passive. Building upon Role-Taking Theory (RTT), we conceptualize these reactions as social roles. A central assumption derived from RTT – that role-taking in norm conflicts is intricately intertwined with the moral self-concept – is translated into three hypotheses: (1) taking over a pro-social role is predicted by individual differences in general moral self-concept relevance; (2) taking over a pro-social role increases the situational moral self-concept; (3) the more relevant the general moral self-concept, the higher the situational moral self-concept increase after pro-social role-taking. Results from two studies using both experimental and correlational designs (total $N = 731$) support these hypotheses and are discussed against the background of Role-Taking Theory.

"One does what one is; one becomes what one does." (Musil, 1930)

Introduction

Imagine Bob and Rob, two members of a working group. They created a new webpage together. However, when Bob presents the new design, he does not mention Rob's contribution and takes all the praise for himself. When Rob becomes aware of Bob's behavior, he is furious, accuses Bob of taking false credit, and demands an apology. Bob, however, does not show any kind of remorse and is not willing to make concessions. A conflict emerges, and bit by bit, the entire work group becomes involved in it, forcing every member to react in one way or another. How will the other group members react? And which processes can explain why a given actor reacts in a given way? In this article, two empirical studies shed light on these questions by making use of a novel role-theoretical approach, Role-Taking Theory (RTT).

Norm Conflicts

Perceived norm violations, like the false credit-taking in the scenario described above, lie at the heart of interpersonal conflicts (Montada, 2013). In a typical course of events, the victim accuses the perpetrator of having violated a norm, and demands some kind of restitution or, at least, an acknowledgment. This demand creates a pressure to react, not only for the accused perpetrator, but also for third-parties, who might consider this demand more or less legitimate. Thus, third-parties become inadvertently involved in a conflict. Their actions can have decisive consequences for the further course of events (Skarlicki & Kulik, 2004). For instance, depending on third parties' reactions, Bob might get away with his egoistic behavior or might be forced to apologize and make good for it. If third parties do not act or even defend Bob, Rob's feelings of being mistreated may have downstream consequences, such as counterproductive work behaviors or layoff intentions (Colquitt & Zipay, 2015).

Social psychological research on team conflicts – such as the one presented here – has mainly focused on the central protagonists' behaviors (i.e., victim and transgressor; see Gray & Wegner, 2009; Shnabel & Nadler, 2008). Studies on third-party responses in normative conflicts is comparably sparse and has mainly focused on specific behaviors, such as third-party punishment (e.g., Lotz, Baumert, Schlösser, Gresser, & Fetchenhauer, 2011), retribution (e.g., Skarlicki & Rupp, 2010), forgiveness (e.g., Green, Burnette, & Davis, 2008), or compensation (e.g., Lotz, Okimoto, Schlösser, & Fetchenhauer, 2011). What is missing, however, is an integrative framework explaining inter-individual differences in third-party reactions in norm conflicts more generally. Role-Taking Theory (RTT) has been developed with the intention to fill this gap (Schwabe & Gollwitzer, 2018).

Social Roles

Role-Taking Theory starts from the assumption that reactions to norm conflicts can be conceptualized as social roles. Besides the "victim" and the "perpetrator" role, third parties involved in the conflict situation can take one (and only one) of five roles (see Figure 1): "victim supporter," "perpetrator supporter," "arbitrator," "escalator", and "bystander." In the example of Rob and Bob, the other members of the team could support Rob's claims for an apology ("victim supporter"), argue in favor of infamous Bob ("perpetrator supporter"), try to mediate between Rob and Bob (e.g., by trying to find a solution that is acceptable for both; "arbitrator"), try to stir up the conflict ("escalator"), or try to stay out of the conflict ("bystander").

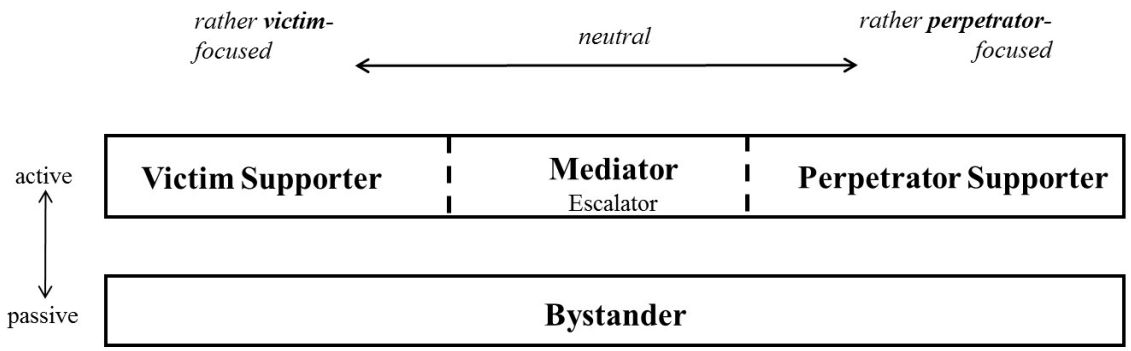

Figure 1. Schematic of the third-party roles' behavior in norm conflicts resulting out of

activity and sidedness as classification dimensions.

The notion that third-party reactions to social conflicts can be conceptualized as "social roles" is not entirely new: research on bullying in schools has made a similar argument. For instance, according to the "participant role approach" (Salmivalli, Lagerspetz, Björkqvist, Österman, & Kaukiainen, 1996), all parties directly or indirectly involved in a bullying episode are forced to take one of several roles, and that the role-taking process, in turn, creates a stable social system of bullying. According to RTT, the roles that third parties can take in a social conflict can be located in a two-dimensional space; the two partly dependent dimensions are (1) partiality (victim-sided, neutral, perpetrator-sided), and (2) interaction maintenance (end the interaction, neutral, prolong the interaction; cf. Schwabe & Gollwitzer, 2018). More specifically, "victim supporters" and "perpetrator supporters" are sided (i.e., not neutral) and active in a conflict, "arbitrators" follow the goal to solve the conflict and are usually active, "escalators" are usually active while trying to prolong the conflict, and "bystanders" are neutral and passive while trying to stay out of the interaction.

Notably, the goals also translate into behavioral expectations towards role-proponents, and are mutually exclusive. In other words, actors can only occupy one role at a time, although they may change these roles over the course of a conflict. Thus, from a social systems viewpoint, social roles can be used to describe reactions in conflicts succeeding norm violations. However, from a social psychological viewpoint, the individual-level processes explaining behavioral phenomena connected to role-taking are more interesting: which processes can explain why a particular actor takes over a given role? What do actors gain from taking over a certain role? And which effects does role-taking have on individual actors and on the course and outcome of conflicts?

Role-Taking as a Social-Cognitive Process

Our theoretical approach – Role-Taking Theory – does more than just describing which roles exist in interpersonal conflict situations; the theory also specifies the social-

cognitive mechanisms involved in the role-taking process. RTT assumes that each of the roles mentioned above – victim supporter, perpetrator supporter, arbitrator, escalator, and bystander – is accompanied by a specific mental schema. This mental schema contains all associations tied to a given position, that is, how actors typically behave in this position, the extent to which the position is associated with positive vs. negative stereotypes, whether relations to other positions are more or less benevolent vs. conflicting, et cetera. For example, the arbitrator role is associated with activity, neutrality, and an orientation towards maximizing the joint outcome of all parties involved; the bystander role, on the other hand, is associated with passivity, neutrality, and an orientation towards maximizing one's own individual outcome.

In RTT terms, role-taking is defined as the activation of a role-specific mental schema: whenever an actor takes over a role, the role-specific mental schema is activated. And, even more importantly, the role-specific mental schema becomes associated with the situational self-concept (or "working self-concept," see Markus & Kunda, 1987). The situational self-concept consists of all self-referent cognitions and emotions in a particular situation, referring to, for instance, (1) *morality* (i.e., the extent to which one perceives her-/himself to be a moral person), (2) *self-esteem* (i.e., a general evaluation of oneself), (3) *identity expression* (i.e., the extent to which one's actions express "who I am"), (4) *meaning* (i.e., the extent to which one's actions are experienced as meaningful), and (5) *power* (i.e., the extent to which one's actions provide one with a sense of power).

A particularly relevant facet of the situational self-concept is morality, the extent to which one is able to perceive oneself as a moral (or immoral) person (Brambilla & Leach, 2014). In norm conflicts, which are inherently charged with moral issues (DeScioli & Kurzban, 2013), the moral situational self-concept is even more relevant than in other situations: everything an actor does and says in a conflict episode can, in principal, be evaluated in morality-related terms, that is, perceived as "right" or "wrong." More

specifically, taking over a pro-social role in a norm conflict (i.e., the arbitrator or victim supporter role), allows the actor both to express and to experience him- or herself as a moral person, at least temporarily. Based on this theoretical argument, we hypothesize that taking over a pro-social conflict role (i.e., arbitrator or victim supporter) is correlated with a more positive moral situational self-concept than taking over the bystander role.

However, a morality-boosting effect of role-taking does not explain *why* a given actor takes a given role. Why do some actors take over "moral" roles more often than others? Do some actors profit from "moral" role-taking more than others? To shed light on these questions, we will briefly sum up individual-level factors explaining reactions in norm conflicts in the following.

Individual Differences in Role-Taking

Based on RTT's notion that pro-social role-taking both expresses and nourishes one's moral identity, we argue that role-taking is uniquely predicted by personality traits reflecting the centrality of one's moral self. Two constructs are relevant in this regard: sensitivity to injustice from an observer's perspective – in short, "Observer Sensitivity" (OS) – and "Moral Identity" (MI). People scoring high on OS are more sensitive to cues of observed injustice than people scoring low on OS (Baumert, Gollwitzer, Staubach, & Schmitt, 2011). Furthermore, they experience strong negative emotions in the face of observed injustice, and they tend to ruminate longer about observed or alleged injustice. In line with these findings, OS is related to pro-social behavioral dispositions such as modesty, agreeableness, or perspective taking (Schmitt, Gollwitzer, Maes, & Arbach, 2005), cooperative choices in social dilemmas (e.g., Fetchenhauer & Huang, 2004; Gollwitzer, Rothmund, Pfeiffer, & Ensenbach, 2009; Thomas, Baumert, & Schmitt, 2011), solidarity with the disadvantaged (Gollwitzer, Schmitt, Schalke, Maes, & Baer, 2005), and a willingness to engage in costly sanctioning of rule-breakers ("altruistic punishment;" see Lotz et al., 2011). Thus, OS can be related to role-taking in interpersonal conflicts in a very straightforward manner: OS is expected to

positively predict the likelihood of taking over a pro-social conflict role, that is, arbitrator or victim supporter.

Moral Identity (Aquino & Reed, 2002), henceforth abbreviated "MI," is defined as "a self-concept organized around a set of moral traits" (p. 1424). Importantly, MI consists of two dimensions: Moral Internalization is the degree to which moral traits (e.g., caring, honest, and fair) are central to one's general self-concept. Moral Symbolization is the degree to which the traits are reflected in one's actions — stated differently, the degree to which a person wants to communicate the possession of moral traits to others. According to RTT, people with a high MI should be more likely to take over a pro-social role in a conflict; this applies both to "high internalizers" as well as to "high symbolizers." However, only "high internalizers" should actually feel more moral after taking over a pro-social role because doing so provides actors with cues for their moral situational self-concept, and "high internalizers" react more readily towards such cues (Aquino & Freeman, 2009). "High symbolizers", on the other hand, are more concerned about *appearing* moral to others instead of nourishing their moral self. Thus, if taking a pro-social role in interpersonal conflicts fosters one's moral situational self-concept, as RTT proposes, it follows that Moral Internalization, but not Moral Symbolization, should amplify (i.e., positively moderate) this effect. Empirical evidence for this mechanism would corroborate the theoretical rationale of RTT and shed light on how role-taking affects the moral dimension of the situational self-concept in dependence of the general importance of a moral identity.

The Present Research

The present paper describes two studies in which the following hypotheses – derived from the RTT framework – are empirically tested:

1) Observer Sensitivity uniquely predicts pro-social role-taking (i.e., over and above Moral Identity and more general morality-related personality traits). That is, the higher an actor scores on Observer Sensitivity, the higher the likelihood of taking

over a pro-social role compared to a passive role ('bystander').

2) Moral Internalization and Symbolization both uniquely predict pro-social role-taking (i.e., over and above Observer Sensitivity and more general morality-related personality traits). The higher an actor scores on Moral Internalization and Moral Symbolization, respectively, the higher the likelihood of taking over a pro-social role compared to a passive role ('bystander').

3a) Taking over a pro-social role in a conflict elicits a higher moral situational self-concept during the conflict than taking over a passive role ('bystander').

3b) This effect is amplified by Moral Internalization, but not by Moral Symbolization.

In Study 1, personality traits and (self-reported) role choice in a conflict scenario were measured online. After taking over a role in this conflict, participants rated five facets of their situational self-concept during the conflict (i.e., morality, self-esteem, identity, meaning, and power) with morality as the central DV. In Study 2, we replicated Study 1 with an experimental design: Instead of measuring participants' role choice via self-reports, participants were randomly assigned to one particular role in a conflict. Again, the moral situational self-concept during the conflict was the central DV in this study.

STUDY 1

Method

Sample

Participants were recruited via a university-wide mailing list including students and university staff members. All participants completed the study online using the survey platform SoSci Survey (Leiner, 2014). As a reward, a Tablet PC was raffled among all participants. The final sample consisted of n = 659 participants (65% women) between 18-70 years (M = 29, SD = 10.4). This large sample size was the result of our efforts to ensure statistical power for each of the logistic regressions in the multinomial regression model of role-choice (see below): we aimed for at least 100 participants per role. A post-hoc power

analysis for the logistic regression with the two roles chosen by the fewest participants as outcome variables (bystander and victim supporter, n = 110 each), and a given Odds Ratio of 2 for a continuous predictor (corresponding to a small to medium effect size; Chen, Cohen, & Chen, 2010), revealed a power of 86% to detect a significant effect. While the majority of the sample consisted of students (66%, n = 432), we also recruited a substantial amount of participants in their working life (30%, n = 197).

Independent Variables

Moral Identity (MI). Participants indicated their age and sex and then responded to the ten items of the MI scale (Aquino & Reed, 2002). The MI scale measures the centrality of a set of moral traits (e.g., caring, fair, helpful) to a person's self-concept with two subscales. The Moral Internalization subscale measures the degree to which persons internalize the set of moral traits into their personal self-concept (e.g., "Being someone who has these characteristics is an important part of who I am;" 5 items, α = .74). The Moral Symbolization subscale measures the effort persons spend on communicating to others, that they have these moral traits (e.g., "I often wear clothes that identify me as having these characteristics;" 5 items, α = .77). Response scales ranged between 1 ("completely disagree") and 6 ("completely agree"), with higher values indicating a higher centrality of morality to a person's self-concept.

Observer Sensitivity (OS). Next, participants completed the ten items of the OS Scale (e.g., "I am upset when someone is undeservingly worse off than others;" α = .87; Schmitt, Baumert, Gollwitzer, & Maes, 2010; Schmitt et al., 2005). Again, response scales ranged between 1 and 6.

Honesty-Humility and Agreeableness. Two traits from the HEXACO model of personality (Ashton et al., 2004) were included as covariates in the present study: Honesty-Humility (e.g., "I would never accept a bribe, even if it were very large;" 10 items, α = .67), and Agreeableness (e.g., "I rarely hold a grudge, even against people who have badly

wronged me;" 10 items, $\alpha = .76$). Response scales ranged between 1 and 5. By controlling for Agreeableness and Honesty-Humility, we were able to scrutinize the hypothesized effects of Moral Identity and Observer Sensitivity over and above broader personality traits, which are relevant in norm conflicts.

Dependent Variables

Role-Taking. Participants were told to imagine being a member of an organizational working group. They then read a short description of a norm conflict evolving in this team (see Supplementary Material A, for all vignettes). Each participant read one vignette, which was chosen randomly out of a set of three vignettes, each containing a different kind of norm violation (e.g., violation of fairness norms, personal harm) to extend external validity. To ensure that the described situations were plausible, immersive, and free of undesired biases (e.g., strongly favoring a certain role), a set of five vignettes was pretested. In a between-subjects design, 88 participants read one of the five vignettes and rated it with regard to immersion ("I can picture myself in that situation very well."), general plausibility ("What happened in the described situation appears plausible to me."), plausibility of the conflict ("It appears plausible to me that a conflict developed out of that situation."), plausibility of third-party reactions ("How plausible would it be for third parties in this situation to support [the perpetrator] / support [the victim] / do nothing / try to reconcile?"), responsibility for causing the conflict ("Which of the two actors caused the conflict?"), and severity of the norm violation ("How much damage did [the perpetrator] cause with his behavior?"). Two vignettes were discarded because of the pretest results. One was discarded because the norm violation was rated as very mild in the pretest ($M = 1.81$, $SD = 1.11$, on a scale from 0 = "no damage was caused by the behavior" to 5 = "a very big damage was caused by the behavior"), and the evolving conflict did not appear very plausible to participants ($M = 2.94$, $SD = 1.35$ on a scale from 0 = "not at all plausible" to 5 = "very plausible"). Another vignette was discarded because the supposed perpetrator of the norm violation was not clearly identifiable as the

perpetrator (44% of all participants in the pretest sample wrongly identified the victim as being responsible for the conflict).

The situation described at the end of all three vignettes portrayed a team meeting in which a conflict between two team members is debated (see Supplementary Material A). While the perpetrator is neglecting his wrongdoing, the victim wants everybody in the work group to recognize the wrongdoing and to support the claim for consequences for the perpetrator. Next, the emotional, cognitive, and behavioral reactions from participants during that meeting were assessed in open format. Thus, participants were asked how they would feel, think, and behave during this conflict meeting. Participants were then asked to subsume their respective reaction under a particular role (i.e., victim supporter, perpetrator supporter, bystander, or arbitrator[6]), which fitted their reaction best. The supporter roles were labeled as 'supporter of [*name*]' to avoid biases through the labels of victim and perpetrator. By assessing the reaction in an open format first and anchoring participants' role choice this way, we prevented participants from choosing their role based on the label. That said, some participants confused the names of victim and perpetrator or chose a role label which was completely incongruent with their reaction described beforehand (e.g., indicating 'arbitrator' as a role, but describing their behavior as clearly sided). To correct for this kind of mislabeling, two independent raters, who were blind to the hypotheses, checked the congruence between reaction and label for all cases. Only obvious cases of mislabeling, indicated by both raters, were corrected (3.6% of all cases, $n = 24$)[7].

Situational Self-Concept (SSC). After role-taking, participants were asked to imagine acting as described before (i.e., during the meeting). In other words, they were told to imagine themselves in their specific role in the actual conflict situation. The five facets of the

[6] The escalator role was not included in designs of Study 1 and Study 2, because it is – similar to the perpetrator supporter role, but to an even stronger extent – a role which is very implausible in a stripped down experimental vignette with no further context factors (e.g., relationships to victim or perpetrator, status in the work team, monetary incentives, etc.).

[7] We also ran all tests with non-corrected labels, to check if this correction of mislabeling changed any results. No differences emerged between the two samples.

situational self-concept as described in the Introduction were assessed by asking participants how acting out their role in the conflict would make them think and feel about themselves (17 items, see Supplementary Material B). Items were developed in accordance with the definition of a situational self-concept as described above. The facet "moral situational self-concept" was most relevant for the present purposes. Since we were unable to find a suitable self-report measure of the moral situational self-concept in the literature, we developed three items based on the definition given above ("What I said and did in the situation shows that I am a moral person;" "What I said and did in the situation shows that I am a good person;" "What I said and did in the situation is an expression of my fundamental convictions and values"). Item- and scale-level analyses (including all 17 items) showed that these three items loaded on a common factor (with loadings ≥ .3), which explained 5.2% of the observed variance (s. Supplementary Material B). Given a sufficient internal consistency (Cronbach's α=.78), we aggregated the three items into a "situational moral self-concept" scale.

The other facets of the situational self-concept were *self-esteem* (3 items, α = .74, e.g., "Because of the things I said and did in the situation, I am satisfied with myself", adapted from the Rosenberg Self-Esteem Scale, Rosenberg, 1979), *identity* (4 items, α = .85, e.g., "What I said and did in the situation is an expression of my personality", self-developed), *meaning* (2 items, r = .74, p < .01, e.g., "What I said and did in the situation was meaningful", self-developed), and *power* (5 items, α = .83, e.g., "I think I had some power in the situation," adapted from Anderson & Galinsky, 2006). Response scales ranged between 1 and 6. A list of all items can also be found in the Supplementary Material B.

Results

Correlations and descriptive statistics for all measured variables are reported in Table 1.

Table 1

Correlations Among and Descriptive Statistics For All Measured Variables

Measure	*M*	*SD*	1	2	3	4	5	6	7	8	9	10
1. MI-Internalization	5.16	0.69	-									
2. MI-Symbolization	3.17	0.98	.38*	-								
3. Observer Sensitivity	4.51	0.81	.38*	.28*	-							
4. Honesty-Humility	3.61	0.61	.17*	.11*	.10*	-						
5. Agreeableness	3.19	0.62	.16*	.14*	-.06	.14*	-					
6. SSC: Self-Esteem	4.60	0.88	.12*	.17*	.10*	-.03	.05	-				
7. SSC: Meaning	3.80	1.15	.18*	.21*	.16*	-.08	.03	.57*	-			
8. SSC: Expression	4.58	0.98	.23*	.18*	.15*	.03	.07	.34*	.37*	-		
9. SSC: Moral	4.21	1.03	.25*	.23*	.28*	-.06	.09*	.59*	.56*	.50*	-	
10. SSC: Power	3.23	1.11	.08*	.22*	.12*	-.14*	-.04	.50*	.62*	.30*	.47*	-

Note. $N = 659$. MI-I = Moral Identity. SSC = Situational Self-Concept. All scales ranged from one to six, except for HH and AGR, which ranged from one to five.

* $p < .05$.

To test Hypotheses 1 and 2 regarding predictors of role-taking, a multinomial regression model was specified. The outcome was the (self-selected) role-choice with the roles of (1) victim supporter (n = 110), (2) arbitrator (n = 439), and (3) bystander (n = 110), with the bystander role as the reference category (coded with 0). The fourth role, perpetrator supporter, was not chosen by any participant. Obviously, this role was not sufficiently plausible or attractive enough in the presently described situation. The results of the multinomial regressions are reported in Table 2.

Table 2

Multinomial Logistic Regression Predicting Choice of Role with Bystander as Reference Category

Model	Predictor	*B*	*SE*	*p*	*OR*	95% CI
Arbitrator	(Intercept)	1.47	0.12	<.01		
	Internalization	0.29	0.12	.02	1.33*	[1.06, 1.67]
	Symbolization	0.10	0.12	.39	1.11	[0.87, 1.41]
	OS	0.34	0.12	<.01	1.40*	[1.11, 1.76]
	HH	0.13	0.11	.26	1.14	[0.91, 1.41]
	AGR	0.07	0.12	.55	1.07	[0.86, 1.34]
VS	(Intercept)	0.02	0.15	.87		
	Internalization	0.36	0.15	.02	1.43*	[1.06, 1.93]
	Symbolization	-0.04	0.15	.78	0.96	[0.71, 1.29]
	OS	0.26	0.15	.07	1.30	[0.98, 1.73]
	HH	-0.28	0.14	.04	0.76*	[0.58, 0.99]
	AGR	-0.19	0.14	.19	0.83	[0.63, 1.10]

Note. N = 659. All predictors were z-standardized. VS = Victim Supporter. OS = Observer Sensitivity. HH = Honesty Humility. AGR = Agreeableness. CI = confidence interval for odds ratio (*OR*).
* $p < .05$.

The five predictors, OS, Moral Internalization, Moral Symbolization, Honesty-Humility, and Agreeableness, were simultaneously entered into the regression model. As predicted, Observer Sensitivity (B = .34, p < .01) and Moral Internalization (B = .29, p = .02) had significant positive effects on the arbitrator vs. bystander contrast. In other words, the

higher a person's scores on Observer Sensitivity or Moral Internalization, the higher the likelihood that he/she chooses the arbitrator over the bystander role – above and beyond more general personality traits (Honesty-Humility and Agreeableness), and – which, in contrast to our hypothesis, did not show the same predictive effect – Moral Symbolization. Regarding the victim supporter role, Moral Internalization (B = .36, p = .02) had a significant positive effect on the victim supporter vs. bystander contrast. Although Observer Sensitivity had a similar effect in size on the victim supporter vs. bystander contrast (B = .26) in comparison to its effect on the arbitrator vs. bystander contrast (B = .34), this effect was not significant on a 5% level (p = .07). Honesty-Humility had a significant negative effect on the victim supporter vs. bystander contrast (B = -.28, p = .04). Taken together, the results show that the likelihood of choosing a pro-social role as an arbitrator increases with higher scores on Observer Sensitivity and Moral Internalization, while the likelihood of choosing a pro-social as a victim supporter increases with higher scores on Moral Internalization, and decreases with higher scores on Honesty-Humility.

Table 3

Mean Scores, Standard Deviations, and Univariate Analyses of Variance for Facets of the Situational Self-Concept as a Function of Role

	Moral		Self-Esteem		Power		Meaning		Expression	
Role	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Bystander	3.25[a]	1.20	3.78[a]	1.17	2.22[a]	1.01	2.70[a]	1.32	4.02[a]	1.30
Victim Supporter	4.47[b]	0.90	4.62[b]	0.75	3.41[b]	0.99	3.93[b]	0.99	4.72[b]	0.85
Arbitrator	4.39[b]	0.86	4.80[b]	0.70	3.44[b]	1.02	4.04[b]	0.96	4.66[b]	0.87
ANOVA (F, η_p^2)	69.76*, .18		66.04*, .17		64.22*, .17		72.84*, .19		24.94*, .07	

Note. N = 659. All scales range from one to six.

* p < .001.

Hypothesis 3a regarding the effects of role-taking on the moral situational self-concept was tested with a one-factorial ANOVA. The mean scores and standard deviations for the moral situational self-concept as a function of role-choice and the corresponding results of the ANOVA are reported in Table 3. Participants who took over pro-social roles as an arbitrator or victim supporter profited more from their roles in regard of their moral situational self-concept in comparison to participants who took over the bystander role, $F(2,656) = 69.76$, $p < .01$, $\eta_p^2 = .18$. That is, pro-social roles – in comparison to the bystander role – provided them with a stronger sense of their moral character.

To test Hypothesis 3b regarding the amplifying role of Moral Internalization (but not Symbolization), we specified multivariate linear regression models with the moral situational self-concept as dependent variable. The results for the Internalization model are displayed in Table 4. In the first step, two dummy-coded variables were included to contrast the effects of taking the arbitrator or the victim supporter role (coded 1, respectively) versus the bystander role (coded 0). Additionally, Moral Internalization (z-standardized to facilitate the interpretation of regression weights) was included as predictor. In the second step, the interaction terms of Internalization × arbitrator (vs. bystander; dummy 1), and Internalization × victim supporter (vs. bystander; dummy 2) were added to the model. Doing so did not decrease the unexplained variance significantly ($\Delta R^2 < .01$, $p = .43$). That is, in a quasi-experimental design with a self-selection to the experimental groups (i.e., to the roles), Moral Internalization (which was used in the same study to predict role-taking and is therefore confounded with the role-choice) did not amplify the effects of pro-social role taking on the moral situational self-concept. We specified an analogous two-step model with Moral Symbolization replacing Internalization, yielding the same pattern of results ($\Delta R^2 < .01$, $p = .87$)[8].

[8] Given the hypothesized pattern of interactions would have emerged, it would have been necessary to further scrutinize the effect by also running the models including Moral Internalization and Symbolization as

Table 4

Hierarchical Regression Analysis Summary for Role-Taking and Moral Internalization Predicting Participants Moral Situational Self-Concept

Step and predictor variable	*B*	*SE B*	β	R^2	ΔR^2
Step 1:				.21*	
Dummy 1 (Arbitrator)	1.04	0.10	.48*		
Dummy 2 (VS)	1.15	0.12	.42*		
Internalization			.19*		
Step 2:				.21*	<.01
Dummy 1 (Arbitrator)	1.05	0.10	.48*		
Dummy 2 (VS)	1.16	0.13	.42*		
Internalization			.15*		
Dummy 1 x Internalization			.05		
Dummy 2 x Internalization			-.03		

Note. $N = 659$. Reference Category for the Role Dummies = Bystander. Internalization was z-standardized.
* $p < .05$.

Discussion

The results of Study 1 support the hypotheses that Moral Internalization and Observer Sensitivity predict role-taking choices in interpersonal conflicts succeeding norm violations, but do not support the hypothesis that Symbolization also predicts pro-social role-taking. People with a strong Moral Internalization disposition were more likely to take over a pro-social role as an arbitrator or as a victim supporter in comparison to the bystander role – over and above more general personality traits, that is, Agreeableness and Honesty-Humility. People scoring high on Observer Sensitivity were more likely to take over a pro-social role as an arbitrator in comparison to the bystander role. Interestingly, Moral Symbolization did not predict pro-social role-taking. This might be a result of the fact that this was an online study:

moderators of pro-social role-taking effects four more times, one for each of the other four assessed facets of the situational-self-concept. Although the hypothesized effect was not found in Study 1, we still ran these models. No interaction effects of pro-social role-taking and Internalization or Symbolization emerged regarding the other four facets.

Moral Symbolization might show a predictive effect in settings where the presence of others – to whom one can communicate one's moral traits through role-taking – is more salient than in an online imagination task.

The results also support Hypothesis 3a that pro-social role-taking elicits a moral situational self-concept. Participants taking over a pro-social role (i.e., arbitrator or victim supporter) perceived a significantly higher moral situational self-concept during the conflict. These effects are in line with our reasoning that taking over pro-social roles in conflicts boosts actors' moral situational self-concept more than taking over a passive bystander role. Regarding Hypothesis 3b, the data from Study 1 do not support the notion that actors scoring high on Moral Internalization profit even more strongly in their moral situational self-concept from pro-social role-taking in comparison to taking over a bystander role. However, in the present study, the free choice of roles – which was necessary to test the predictive effects of the traits on role-taking – only offered a quasi-experimental design to investigate the role-taking effects on the situational self-concept. This design is not optimal to test Hypothesis 3 (a and b) due to the non-independence between MI and role choice. To test Hypotheses 3a and 3b more rigidly, we conducted a second experimental study with a random assignment of roles.

STUDY 2

In Study 2, we tested Hypotheses 3a and 3b in an experimental design. In addition, we also wanted to avoid a priming of moral idealism through the assessment of Moral Identity directly prior to role-taking. Therefore, we assessed Moral Identity independently at a measurement occasion four months prior to the actual study.

Sample. Participants were first-year psychology students. In the first week of the semester, they completed a battery of paper-pencil questionnaires that – among other trait scales – also included the same OS and MI scales that had been used in Study 1. However, the OS scale is not relevant for the proposed motivational mechanism; thus, we only focus on the

MI scale here (5 items for Internalization, α = .78; 5 items for Symbolization α = .78). Four months later, students were invited to take part in an online study on "behavior in conflict situations." Seventy-two students completed the online study, and 55 (i.e., 76%) of them could be unambiguously matched with their data from the questionnaire battery on the basis of a personalized code. A post-hoc power analysis in regard of a R^2-increase corresponding to a medium-sized effect of the proposed interaction in Hypothesis 3b (f^2 = .15) yielded a power of 70%.

Materials and Procedure. The materials and procedure were the same as in Study 1, except that (1) no traits were assessed directly prior to reading the vignettes, and (2) roles in the conflict situation (victim supporter, arbitrator, bystander) were randomly assigned instead of chosen. Furthermore, to avoid biases created by the role labels, a schematic depiction was used to assign the roles (see Figure 2). As a manipulation check, participants were asked how they would behave when taking over their respective role in the conflict, and answers were probed with regard to their role-consistency by two independent raters (e.g., no active behavior as a bystander, no side-taking by arbitrators, active support of the victim as a victim supporter). Participants would have been excluded if both raters indicated role-inconsistent behavior. However, no data had to be excluded on the basis of this rule. Afterwards, like in Study 1, the same self-referent cognitions and emotions during the conflict session were assessed (all Cronbach's $\alpha > .75$)**.**

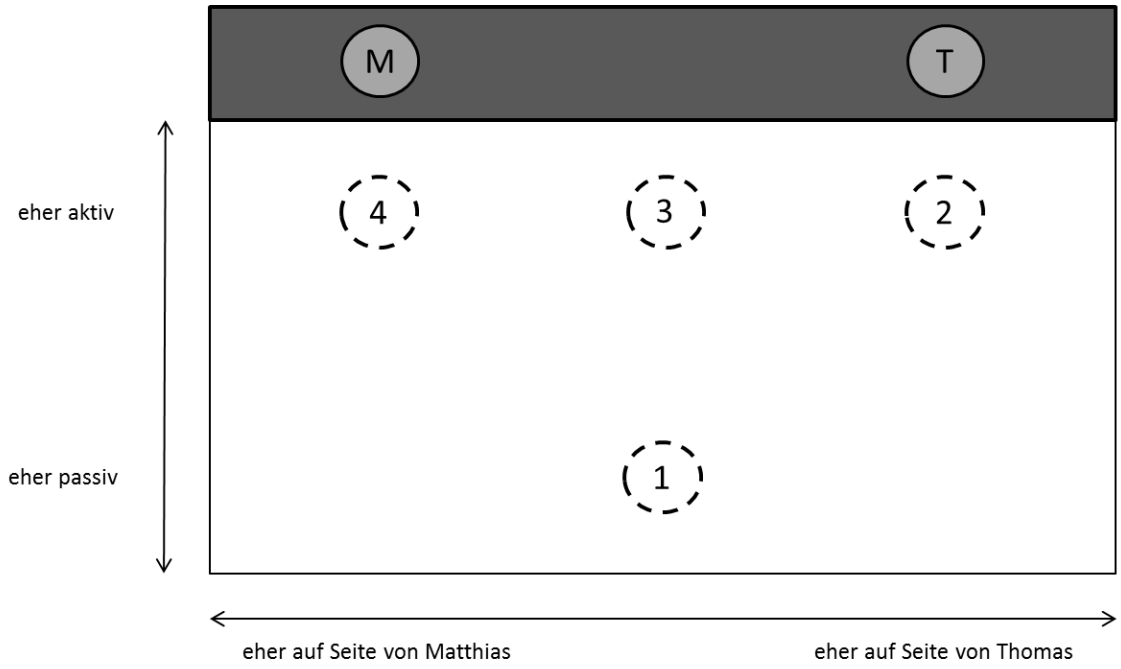

Figure 2. Schematic used for the role assignment of bystander (1), victim supporter (4), and arbitrator (3) in Study 2.

Results

Correlations and descriptive statistics for all measured variables are reported in Table 5.

Table 5

Correlations Among and Descriptive Statistics For All Measured Variables

Measure	*M*	*SD*	1	2	3	4	5	6	7
1. MI-Internalization	4.27	0.63	-						
2. MI-Symbolization	2.33	0.99	.34*	-					
3. SSC: Self-Esteem	4.70	0.94	.06	-.02	-				
4. SSC: Meaning	4.21	1.10	.04	.01	.71*	-			
5. SSC: Expression	4.45	1.09	.25	.20	.56*	.65*	-		
6. SSC: Moral	4.13	0.96	.12	.09	.70*	.80*	.74*	-	
7. SSC: Power	3.53	1.08	-.20	.02	.48*	.58*	.48*	.60*	-

Note. $N = 55$. MI-I = Moral Identity. SSC = Situational Self-Concept. Internalization and Symbolization ranged from zero to five. All SSC scales ranged from one to six.

* $p < .05$.

Hypothesis 3a stating that pro-social role-taking is associated with a higher moral

situational self-concept than taking over a bystander role was tested by an ANOVA. The results are displayed in Table 6 together with Means, Standard Deviations, and ANOVAs for all measured facets of the situational self-concept. The results replicate the effect from Study 1: participants who were assigned to a pro-social role (arbitrator or victim supporter) reported a significantly higher moral situational self-concept during the conflict meeting than proponents of the bystander role, $F(2, 52) = 4.50$, $p = .02$, $\eta_p^2 = .15$.

Table 6

Mean Scores, Standard Deviations, and Univariate Analyses of Variance for Facets of the Situational Self-Concept as a Function of Role in Study 2

	Moral		Self-Esteem		Power		Meaning		Expression	
Role	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Bystander	3.67[a]	1.09	4.13[a]	1.05	3.11[a]	1.30	3.57[a]	1.31	3.92[a]	1.31
Victim Supporter	4.42[b]	0.90	5.02[b]	0.60	3.95[b]	0.72	4.40[b]	0.57	4.77[b]	1.02
Arbitrator	4.42[b]	0.65	5.09[b]	0.74	3.67[b]	0.92	4.76[b]	0.79	4.79[b]	0.71
ANOVA (*F*, η_p^2)	4.50*, .15		7.98*, .24		3.21*, .11		7.77*, .23		4.83*, .15	

Note. $N = 55$. All scales range from one to six.

* $p < .05$.

To test Hypothesis 3b stating that this effect is amplified by Moral Internalization, but not by Moral Symbolization, we specified multivariate linear regression models with the moral situational self-concept as dependent variable. The results for the model with Internalization as moderator are displayed in Table 7. In the first step, two dummy-coded variables were included to contrast the effects of taking the arbitrator or the victim supporter role (coded 1, respectively) versus the bystander role (coded 0). Additionally, Moral Internalization (z-standardized to facilitate the interpretation of regression weights) was included as predictor. In the second model, the two interaction terms of Internalization × arbitrator (vs. bystander; dummy 1), and Internalization × victim supporter (vs. bystander;

dummy 2) were added to the model.

Table 7

Hierarchical Regression Analysis Summary for Role-Taking and Moral Internalization Predicting Participants Moral Situational Self-Concept in Study 2

Step and predictor variable	*B*	*SE B*	β	R^2	ΔR^2
Step 1:				.16*	
Dummy 1 (Arbitrator)	0.74	0.30	.37*		
Dummy 2 (VS)	0.75	0.31	.35*		
Internalization			.09		
Step 2:				.29*	.13*
Dummy 1 (Arbitrator)	0.78	0.27	.39*		
Dummy 2 (VS)	0.78	0.29	.36*		
Internalization			-.40		
Dummy 1 x Internalization			.27		
Dummy 2 x Internalization			.54*		

Note. $N = 55$. Reference Category for the Role Dummies = Bystander. Internalization was z-standardized.
* $p < .05$.

Doing so decreased the unexplained variance significantly ($\Delta R^2 = .13$, $p = .02$), supporting the hypothesis that Moral Internalization amplifies the effects of pro-social role-taking on the moral situational self-concept. For people scoring high in Internalization (+1 SD above the mean), pro-social role-taking as an arbitrator in comparison to taking a bystander role had a significant positive effect on the predicted moral situational self-concept ($B = .88$, $p < .01$), as well as pro-social role-taking as victim supporter ($B = 1.22$, $p < .01$). In contrast, for people scoring one standard deviation below the mean in Internalization, taking the arbitrator role ($B = 0.67$, $p = .02$) still led to a significant increase in the predicted moral situational self-concept, while taking the victim supporter role ($B = 0.34$, $p = .12$) did not. The interaction is graphically displayed in Figure 3. Note that the Internalization × arbitrator interaction coefficient in Table 7 is not significant according to conventional levels ($p = .11$). However, the pattern of the interaction and the simple effects supports the notion that Internalization

moderates role-taking effects on the moral situational self-concept. We ran an analogous model for Moral Symbolization (see Table 8). Although a similar pattern resulted, using Symbolization as a moderator of role-taking effects did not decrease the unexplained variance significantly ($\Delta R^2 = .05$, $p = .21$). This lends support to our hypothesis that Moral Internalization, but not Moral Symbolization, amplifies the positive effect of pro-social role-taking on the moral situational self-concept[9].

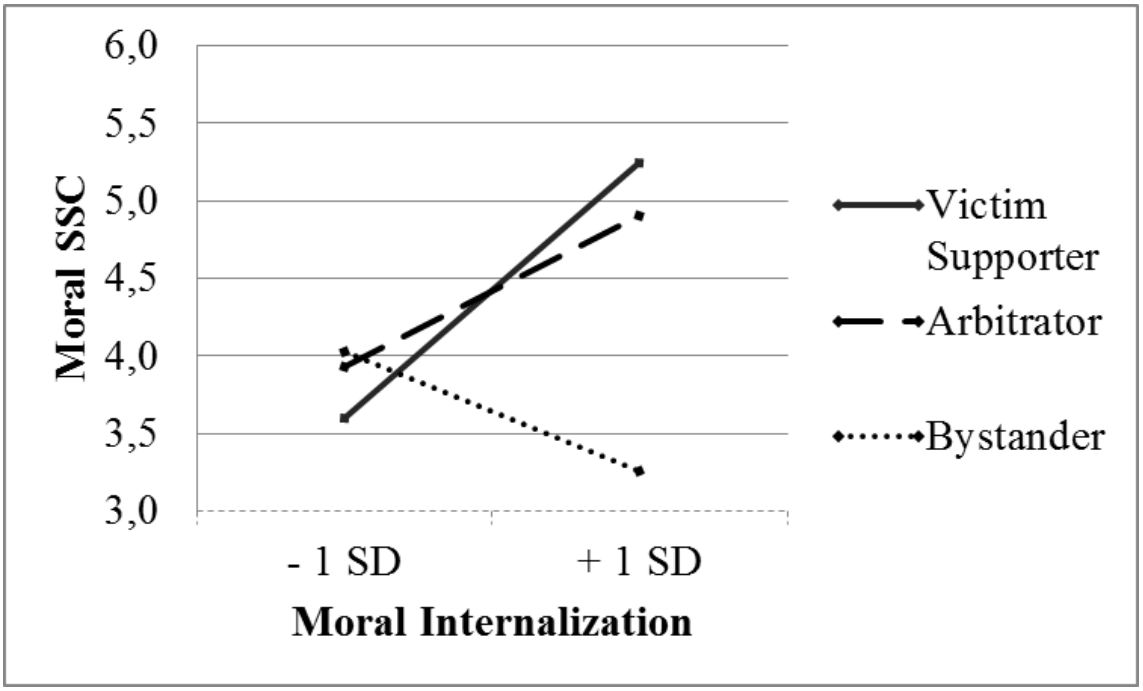

Figure 3. Relationships between Moral Internalization and the Moral Situational Self-Concept in Study 2 for each role.

[9] To test the interaction of Internalization and role-taking over and above effects of Symbolization and its interaction with role-taking, we also ran a full model (see Supplementary Material B, Table 10), yielding the same pattern of results. To even further scrutinize the hypothesized effect, we ran the model including Moral Internalization as a moderator of pro-social role-taking effects four more times, one for each of the other four facets of the situational-self-concept experienced in the conflict situation. Although one could expect similar patterns at least for *expression of identity* ($\Delta R^2 = .09$, $p = .06$) and for *meaningfulness of behavior* ($\Delta R^2 = .08$, $p = .06$), no further significant moderating effects were found. In conclusion, a high Moral Internalization (but not Moral Symbolization) shows the expected moderating effect in regard of the moral facet of the self-referent cognitions and emotions (but not in regard of the other facets) supplied by pro-social role-taking.

Table 8

Hierarchical Regression Analysis Summary for Role-Taking and Moral Symbolization Predicting Participants Moral Situational Self-Concept in Study 2

Step and predictor variable	*B*	*SE B*	β	R^2	ΔR^2
Step 1:				.15*	
Dummy 1 (Arbitrator)	0.74	0.30	.37*		
Dummy 2 (VS)	0.75	0.31	.35*		
Symbolization			.05		
Step 2:				.20*	.05
Dummy 1 (Arbitrator)	0.75	0.29	.37*		
Dummy 2 (VS)	0.82	0.31	.38*		
Symbolization			-.31		
Dummy 1 x Symbolization			.25		
Dummy 2 x Symbolization			.34		

Note. $N = 55$. Reference Category for the Role Dummies = Bystander. Symbolization was z-standardized.
* $p < .05$.

Discussion

The results of Study 2 further support the hypothesis that pro-social role-taking in norm conflicts reflects positively on the moral situational self-concept, and that people with a high Moral Internalization profit even more from the beneficial effect of pro-social role-taking. The latter effect supports the proposed motivational mechanism: people who have strongly internalized moral aspects into their personal self-concept are drawn towards roles that elicit a moral situational self-concept because they feel even more moral in pro-social roles than people who do not value the moral aspect of their self as highly. In this sense, Study 2 sheds light on how stable features of one's self-concept (e.g., a high centrality of moral aspects for the general self) interact with situational factors (e.g., pro-social role-taking in conflicts) on the moral situational self-concept.

General Discussion

In this article, we aimed to test a novel role-theoretical approach (RTT) by shedding light on (1) individual-level mechanisms explaining inter-individual differences in third-party reactions in norm conflicts, and (2) the effects of role-taking on the situational moral self-concept. Building on a definition of social roles as the mental schemata actors tie to social positions, we tested whether pro-social role-taking (arbitrator or victim supporter) can be uniquely predicted by personality traits reflecting the centrality of a moral self-concept (i.e., Moral Internalization, Moral Symbolization, Observer Sensitivity). In line with Hypothesis 1, Moral Internalization and Observer Sensitivity predicted pro-social role-taking over and above other pro-social traits (i.e., Honesty/Humility and Agreeableness). However, Moral Symbolization did not show the hypothesized predictive effect.

Second, we tested whether pro-social role-taking leads to a more positive moral situational self-concept than taking over the bystander role – this follows from RTT's central claim that taking over a prosocial role positively reflects on one's moral self-concept. Results were consistent with this hypothesis in two studies using both correlational and experimental designs: participants who selected (Study 1) or were assigned to (Study 2) a pro-social role experienced themselves as more moral persons during the conflict in comparison to participants in the bystander role. In addition, we investigated a potential motivational mechanism underlying the predictive effect of Moral Internalization for pro-social role-taking. According to RTT, taking over a pro-social role nourishes the moral situational self-concept. What follows from this argument is that the effect of taking over a pro-social role on moral self-referent moral cognitions and emotions should be particularly pronounced among "high internalizers" (i.e., people with high values on Moral Internalization), but not among "high symbolizers" (i.e., people with high values on Moral Symbolization), because only Moral Internalization reflects the centrality of being moral for one's self-concept. This hypothesis (3b) was tested in both studies. While the effect did not occur in Study 1 (in which role-taking was self-selected and, thus, quasi-experimental), the results of the more rigid test in Study 2

(in which role-taking was randomly assigned) were in line with this hypothesis: people with relatively high values on Moral Internalization (but not those with relatively high values on Moral Symbolization) felt even more moral after taking over the arbitrator or the victim supporter role in a conflict.

The difference in results between Study 1 and Study 2 needs to be discussed. The quasi-experimental design of Study 1 made it more difficult to test the hypothesized moderation effect rigidly, because Internalization and role-taking were confounded. One potential alternative explanation of the null effect in Study 1 caused by the quasi-experimental design lies in a post-hoc legitimization of a bystander role-choice in Study 1: actors scoring high on Moral Internalization were, arguably, more prone to justify their inactive role-choice post hoc by reporting a more moral situational self. That could be the reason why one obtains similar positive regression coefficients of Internalization and the moral situational self-concept in all roles (s. Step 2, Table 4). In Study 2, post-hoc justifications are obsolete, because roles were assigned, not selected. This might have been the decisive factor to reveal an amplifying effect of Internalization on the moral situational self-concept boost after prosocial role-taking (s. Step 2, Table 7). In conclusion, the results lend support to our theoretical argument that pro-social role-taking in norm conflicts is related to a process of moral self-regulation, rather than, for example, impression management concerns. In sum, when explaining third-party reactions, the results highlight the importance of how these reactions feedback on the actors' moral self-concept *during* the conflict, and how inter-individual differences moderate this feedback, reinforcing the reaction in this way.

Limitations

The present research faced three methodological challenges: demand effects, selectivity of the samples in both studies, and sample size issues in Study 2. Demand effects are relevant because participants might have taken pro-social roles more often or reported more positive self-referent cognitions and emotions after pro-social role-taking because they

inferred that this were the effects we were looking for (e.g., from the role labels). We tried to minimize such artifacts by (1) capturing the reactions in Study 1 in open format first, (2) avoiding the labels ‘victim’ and ‘perpetrator’ in general, and (3) by not using any role labels in Study 2 at all (instead, we used a schematic description of the various roles, see Figure 2). Of course, we cannot rule out that participants guessed our hypothesis despite these demand-reducing strategies.

Second, the samples in both studies were dominated by students, while the second sample consisted of first-year psychology students only. This raises the question of generalizability. Especially in regard of interpersonal conflict behavior, effects might differ between psychology students and the general population. For example, compared to the general population, psychology students might be prone to prefer choosing a pro-social role over the bystander role. What needs to be said in this regard is that role-choice was only self-selected in Study 1, and that the sample of Study 1 also comprised a substantial amount of students of other departments, as well as non-students (30%). In Study 2, where we only sampled psychology students, roles were randomly assigned.

Third, the sample size of Study 2 was relatively small. This raises issues of statistical power and the question of reliability. However, a fluke in regard of the central result (i.e., the interaction of pro-social role-taking and Moral Internalization on the moral situational self-concept) seems to be quite implausible, taking the consistency of the overall data pattern into account (see Table 7). Nevertheless, the insignificant interactions of pro-social role-taking and Moral Symbolization on the moral situational self-concept, and the insignificant interactions of pro-social role-taking and Moral Internalization on the other facets of the situational self-concept might be due to a lack of statistical power in Study 2. Given the results of this study, these effects might be a bit smaller than the interaction of Internalization and pro-social role-taking (i.e., around increases in R^2 of .05). To test this effect size in further studies with a power of at least 80%, a total sample size of $N = 187$ is required. Thus, a replication with a

potentially larger (and preferably more diverse) sample is desirable.

Outlook and Conclusion

The present findings can only be the first step in a potential research program on third-party roles and role-taking in conflicts. Many research questions are still left open. Most importantly, future research needs to investigate how, when, and why third-party role-taking tends to escalate or stabilize the conflict situation, versus how, when, and why it leads to de-escalating or constructive conflict behavior. That is, it has to shed light on how, when and why the third-party roles and the role-taking effects tied to them foster unfavorable versus favorable outcomes of conflicts for all involved actors. More specifically, future studies should investigate (1) paradoxical phenomena resulting from role-taking, and (2) the basic social-cognitive and motivational processes underlying role-taking and its behavioral consequences.

Paradoxical Phenomena of Role-Taking

The present research looked at one specific psychological process involved in role-taking: the morality-boosting effect of taking over a prosocial role. Now, since this effect is empirically established, one paradox follow-up hypothesis directly follows from RTT: if arbitrators (and victim supporters) are indeed motivated to experience themselves as moral during the conflict, and if arbitrators lose this ongoing benefit of their role once the conflict is over (because the roles are embedded in the conflict situation), this should motivate arbitrators to keep conflicts going (Schwabe & Gollwitzer, 2018). That is, the effect found in this study could motivate arbitrators in a way that is undermining their ostensible role goal. This might explain why sometimes conflicts remain unsolved, even though third parties are involved.

Furthermore, roles also differ in many other aspects, for example, in behavioral expectations, symbols, and relationships to other roles, which – following RTT – has substantial consequences for the matters discussed above. For example, roles can only

meaningfully exist in a relationship with at least one other role (e.g., no victim can exist without a perpetrator, no victim supporter without a victim, etc.). If a person is motivated to benefit from a victim supporter role in some way, they first need to find a victim. That is, people scoring high in Internalization and low in Honesty-Humility (see Table 2), might be generally more prone to categorize actors as victims in ambiguous cases. Furthermore, this effect is likely to be stronger the more an actor is motivated to profit from the benefits the victim supporter role has to offer (e.g., in regard of the moral SSC). This would result in a biased black-and-white representation of the conflict, which, in turn, is an obstacle to conflict resolution. Thus, the more victim supporters are motivated to benefit from their role-taking in regard of their situational self-concept, the more their conflict representation is likely to become biased and one-sided. This is another example how stable role-related features (inter-role relationships and motivations for role-taking) can help to explain unfavorable third-party phenomena in norm conflicts.

Social-Cognitive Processes

Second, theoretical and empirical work has to investigate the social-cognitive processes underlying role-taking and differentiate it from similar processes that have been described in other theories, for example, the depersonalization process described in Self-Categorization Theory (Turner, Hogg, Oakes, Reicher, & Wetherell, 1987). Although role-taking and depersonalization can be theoretically separated very clearly (e.g., in regard of the self-related part of the process[10]), the processes are congeneric on a basic theoretical level (i.e., both deal with a certain kind of identification with a social category). Thus, over-arching theoretical principles incorporated in SCT (e.g., accentuation, the meta-contrast principle) are likely to apply to the basic cognitive underpinnings of role-taking as well, potentially causing similar effects. One example for a deriving empirical research question: does role-taking moderate social projection, just like social categorization in groups does (Clement & Krueger,

[10] For example, depersonalization includes a shift in the level of self-construal abstractness from an individual 'I' identity to a social 'We' identity, while RTT's role-taking does not.

2002)? That is, does role-taking, for example as a victim supporter, cause people to assume that the victim or another victim supporter (but not the perpetrator supporter) is similar to them, thinks, feels, and will behave the same like they do? Effects like these could further explain biased and non-cooperative third-party behavior.

In sum, research following these and similar directions will create stepping stones on the road to a comprehensive role-theoretical understanding of third-party behavior in interpersonal conflicts. Besides paving the way for role-theoretical research in other domains (e.g., negotiations, solutions to social dilemmas, etc.) this understanding is also relevant for developing effective and efficient interventions (i.e., coachings, training programs) aimed at preventing situations like the one described at the beginning of this article (i.e., Bob and Rob's conflict over taking credit for ideas), or, at least, at helping third-parties to deal with them efficiently. Thus, we hope this first empirical study investigating third-party role-taking in norm conflicts will serve as a blueprint for further research investigating social roles and the social-psychological processes involved in role-taking in conflict situations.

Declaration of Conflicting Interests

The Authors declare that there is no conflict of interest.

References

Anderson, C., & Galinsky, A. D. (2006). Power, optimism, and risk-taking. *European Journal of Social Psychology, 36*, 511-536. doi: 10.1002/ejsp.324

Ashton, M. C., Lee, K., Perugini, M., Szarota, P., De Vries, R. E., Di Blas, L., Boies, K., & De Raad, B. (2004). A six-factor structure of personality-descriptive adjectives: solutions from psycholexical studies in seven languages. *Journal of Personality and Social Psychology, 86*, 356-366. doi: 10.1037/0022-3514.86.2.356

Aquino, K., & Freeman, D. (2009). Moral identity in Business Situations: A Social-Cognitive Framework for Understanding Moral Functioning. In D. Narvaez & D. Lapsley (Eds.), *Personality, Identity, and Character: Explorations in Moral Psychology* (pp. 375-395). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511627125.018

Aquino, K., & Reed, I. I. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology, 83*, 1423-1440. doi:10.1037/0022-3514.83.6.1423

Baumert, A., Gollwitzer, M., Staubach, M., & Schmitt, M. (2011). Justice sensitivity and the processing of justice-related information. *European Journal of Personality, 25*, 386-397. doi: 10.1002/per.800

Brambilla, M., & Leach, C. W. (2014). On the importance of being moral: The distinctive role of morality in social judgment. *Social Cognition, 32*, 397-408. doi: 10.1521/soco.2014.32.4.397

Chen, H., Cohen, P., & Chen, S. (2010). How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. *Communications in Statistics—Simulation and Computation, 39*, 860-864. doi: 10.1080/03610911003650383

Clement, R. W., & Krueger, J. (2002). Social categorization moderates social projection. *Journal of Experimental Social Psychology, 38*, 219–231. doi:10.1006/jesp.2001.1503

Colquitt, J. A., & Zipay, K. P. (2015). Justice, fairness, and employee reactions. *Annual Review Organizational Psychology and Organizational Behavior*, *2*, 75–99. doi: 10.1146/annurev-orgpsych-032414-111457

DeScioli, P., & Kurzban, R. (2013). A solution to the mysteries of morality. *Psychological Bulletin*, *139*, 477–496. doi: 10.1037/a0029065

Fetchenhauer, D., & Huang, X. (2004). Justice sensitivity and distributive decisions in experimental games. *Personality and Individual Differences, 36,* 1015-1029. doi: 10.1016/S0191-8869(03)00197-1

Gollwitzer, M., Rothmund, T., Pfeiffer, A., & Ensenbach, C. (2009). Why and when justice sensitivity leads to pro-and antisocial behavior. *Journal of Research in Personality, 43,* 999-1005. doi: 10.1016/j.jrp.2009.07.003

Gollwitzer, M., Schmitt, M., Schalke, R., Maes, J., & Baer, A. (2005). Asymmetrical effects of justice sensitivity perspectives on prosocial and antisocial behavior. *Social Justice Research*, *18*, 183-201. doi: 10.1007/s11211-005-7368-1

Gray, K., & Wegner, D. M. (2009). Moral typecasting: divergent perceptions of moral agents and moral patients. *Journal of Personality and Social Psychology*, *96*, 505-520. doi: 10.1037/a0013748

Green, J. D., Burnette, J. L., & Davis, J. L. (2008). Third-party forgiveness: (Not) forgiving your close other's betrayer. *Personality and Social Psychology Bulletin*, *34*, 407-418. doi: 10.1177/0146167207311534

Leiner, D. J. (2014). SoSci survey. Retrieved from https://www.soscisurvey.de.

Lotz, S., Baumert, A., Schlösser, T., Gresser, F., & Fetchenhauer, D. (2011). Individual differences in third-party interventions: How justice sensitivity shapes altruistic punishment. *Negotiation and Conflict Management Research, 4,* 297-313. doi: 10.1111/j.1750-4716.2011.00084.x

Lotz, S., Okimoto, T. G., Schlösser, T., & Fetchenhauer, D. (2011). Punitive versus compensatory reactions to injustice: Emotional antecedents to third-party interventions. *Journal of Experimental Social Psychology, 47*, 477-480. doi: 10.1016/j.jesp.2010.10.004

Montada, L. (2013). Gerechtigkeitskonflikte und Möglichkeiten ihrer Lösung. In M. Gollwitzer, S. Lotz, T. Schlösser, & B. Streicher (Eds.), *Soziale Gerechtigkeit. Was unsere Gesellschaft aus den Erkenntnissen der Gerechtigkeitspsychologie lernen kann* (pp. 35-54). Göttingen: Hogrefe.

Musil, R. (1930). *The man without qualities.* London: Picador.

Rosenberg, M. (1979). *Conceiving the self.* New York: Basic Books.

Salmivalli, C., Lagerspetz, K., Björkqvist, K., Österman, K., & Kaukiainen, A. (1996). Bullying as a group process: Participant roles and their relations to social status within the group. *Aggressive Behavior*, *22,* 1-15. doi: 10.1002/(SICI)1098-2337(1996)22:1<1::AID-AB1>3.0.CO;2-T

Schwabe, J., & Gollwitzer, M. (2018). *How stable positions in interpersonal interactions shape social cognition and behavior: A Role-Taking Theory*. Manuscript under review.

Schmitt, M., Baumert, A., Gollwitzer, M., & Maes, J. (2010). The justice sensitivity inventory: Factorial validity, location in the personality facet space, demographic pattern, and normative data. *Social Justice Research*, *23*, 211-238. doi: 10.1007/s11211-010-0115-2

Schmitt, M., Gollwitzer, M., Maes, J., & Arbach, D. (2005). Justice sensitivity. *European Journal of Psychological Assessment, 21*, 202-211. doi: 10.1027/1015-5759.21.3.202

Shnabel, N., & Nadler, A. (2008). A needs-based model of reconciliation: satisfying the differential emotional needs of victim and perpetrator as a key to promoting reconciliation. *Journal of Personality and Social Psychology*, *94*, 116-132. doi: 10.1037/0022-3514.94.1.116

Skarlicki, D. P., & Kulik, C. T. (2004). Third-party reactions to employee (mis) treatment: A justice perspective. *Research in organizational behavior*, *26*, 183-229. doi: 10.1016/S0191-3085(04)26005-1

Skarlicki, D. P., & Rupp, D. E. (2010). Dual processing and organizational justice: The role of rational versus experiential processing in third-party reactions to workplace mistreatment. *Journal of Applied Psychology*, *95*, 944-952. doi: 10.1037/a0020468

Thomas, N., Baumert, A., & Schmitt, M. (2011). Justice sensitivity as a risk and protective factor in social conflicts. In E. Kals & J. Maes (Eds.), *Justice and Conflicts* (pp. 107-120). Berlin: Springer.

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Cambridge, MA, US: Basil Blackwell.

Supplementary Material

Supplementary Material A

Vignette A
There is a conflict in an organizational team. At the core of the conflict are Matthias and Thomas. They both have been in the team for a long time and do not differ from each other regarding work experience or skills level. Both have strengths and weaknesses - each of them in their own area.
Following situation led to the conflict:
Matthias got instructed to create a new database system, which should be used in the whole company. Because of Matthias´s lack of experience regarding such systems Thomas supported him. Together they created a very efficient system.
When Matthias presented the new system to his supervisors he earned much praise for his good work. Thomas´s assistance was never mentioned.
Thomas thereupon felt treated unfairly.
He accuses Matthias of passing him over to take all the credit for their cooperation alone.
Matthias tries to reject these allegations.
A conflict emerged from this situation affecting all team members. Thomas expects the other team members to also condemn Matthias´s behavior. Matthias expects the other team members to calm Thomas down. In the meantime their relationship has suffered a lot.
Obviously this alos has consequences for the working environment of the team.
The other team members take different positions regarding the situation.

Vignette B
There is a conflict in an organizational team. At the core of the conflict are Matthias and Thomas. They both have been in the team for a long time and do not differ from each other regarding work experience or skills level. Both have strengths and weaknesses - each of them in their own area.
Following situation led to the conflict:
Thomas and Matthias applied for a new attractive job in the working group. They are now engaged in an ongoing competition for the job.
Both candidates have an important project deadline scheduled shortly before the decision is made. The assessment of the project will play an important role in the decision to fill the position.
There´s always an opportunity in the team to get feedback from an „expert“ before handing in an important project. The expert has excellent knowledge in one field. His task is to examine the project and make suggestions for improvements. Thereby many mistakes and flaws can be fixed.
Thomas´s project had to be checked by someone with expert knowledge in IT. Matthias is the expert of the team regarding IT.
Despite asking for feedback on his project, Thomas did not receive an answer from Matthias.
Thomas thereupon feels treated unfairly. He takes Matthias behavior as an act of sabotage.
Thomas accuses Matthias of behaving unfairly and harming him. Matthias tries to reject these allegations.
A conflict emerged from this situation affecting all team members. Thomas expects the other team members to also condemn Matthias´s behavior. Matthias expects the other team members to calm Thomas down. In the meantime their relationship has suffered a lot.
Obviously this alos has consequences for the working environment of the team.
The other team members take different positions regarding the situation.

Vignette C

There is a conflict in a team of a scientific institute. At the core of the conflict are Matthias and Thomas. They both have been in the team for a long time and do not differ from each other regarding work experience or skills level.

Following situation led to the conflict:

Thomas and Matthias applied for a new attractive job in their working group. They are now engaged in an ongoing competition for the job.

Both candidates conduct an important study which will be finished shortly before the decision is made. The assessment of the studies will play an important role in the decision to fill the position. Both studies are demanding and both are following a tight schedule. Due to the time pressure the studies are equally stressful and burdensome for both candidates.

Thomas and Matthias have to work at the same laboratory to conduct their studies. They have to share the lab. The lab should be left as clean and tidy as possible to prevent a loss of time for the following occupant. However, sometimes the schedules are too tight which forces the following occupant to finish tidying up the lab.

Lately, Thomas feels treated unfairly by Matthias. Increasingly frequent, the lab was in a mess when he arrived. Thomas increasingly suspects Matthias of purposely leaving the lab without tidying up to harm him and safe time. As a consequence, Thomas has serious time pressure which also causes psychological strain. He takes Matthias's behavior as an act of sabotage.

Thomas accuses Matthias of behaving unfairly and harming him. Matthias tries to reject these allegations.

A conflict emerged from this situation affecting all team members. Thomas expects the other team members to also condemn Matthias´s behavior. Matthias expects the other team members to calm Thomas down. In the meantime their relationship has suffered a lot. Obviously this alos has consequences for the working environment of the team.

The other team members take different positions regarding the situation.

Supplementary Material B

Table 9

Summary of Items and Factor Loadings for Promax Five-Factor Solution for the Situational Self-Concept Items (N =659).

Item	Factor loading					Communality
	1	2	3	4	5	
1. What I said and did in the situation shows that I am a moral person.	-.04	-.06	**.97**	-.08	.01	.75
2. What I said and did in the situation is an expression of my fundamental values and convictions.	.28	-.08	**.31**	.24	.07	.47
3. What I said and did in the situation shows that I am a good person.	-.01	.03	**.82**	-.08	-.01	.62
4. Because of the things I said and did in the situation, I am satisfied with myself.	.02	.07	.22	**.59**	-.07	.58
5. I have the impression that people perceived me as a valuable person because of the things I said and did in the situation.	-.04	.14	**.35**	.23	.13	.50
6. Because of the things I said and did in the situation, I feel like a loser. (r)	-.05	-.07	-.17	**.96**	.01	.69
7. What I said and did in the situation carried important meaning.	.01	.05	.06	-.13	**.89**	.78
8. What I said and did in the situation was meaningful and important.	.02	-.01	.01	.12	**.76**	.72
9. What I said and did in the situation gives me a sense of what kind of person I am.	**.66**	.01	.11	-.13	.08	.50
10. What I said and did in the situation has nothing to do with me as a person. (r)	**.66**	-.03	-.07	.06	.02	.42
11. What I said and did in the situation is an expression of my personality.	**.91**	-.01	-.04	.01	-.04	.75
12. My personality is reflected in what I said and did in the situation.	**.92**	.03	-.03	-.03	-.03	.78
13. Through what I said and did in the situation, I was able to show how assertive I am.	.07	**.85**	-.05	.01	-.05	.67
14. From what I said and did in the situation, one can tell that I have a strong willpower.	-.06	**.81**	.01	-.06	-.08	.53
15. Through what I said and did in the situation I was able to convince others of my point of view.	.04	**.62**	.14	.09	.01	.61
16. What I said and did in the situation probably did not have much effect. (r)	.01	.21	-.10	.14	**.44**	.41

17. I think, I had some power in the situation.	-.04	**.58**	-.09	-.08	.25	.45

Note. Boldface indicates highest factor loadings. r = reversed.

Table 10

Hierarchical Regression Analysis Summary for Role-Taking and Moral Symbolization and Internalization Predicting Participants Moral Situational Self-Concept in Study 2

Step and predictor variable	*B*	*SE B*	β	R^2	ΔR^2
Step 1:				.16*	
Dummy 1 (Arbitrator)	0.74	0.30	.37*		
Dummy 2 (VS)	0.75	0.31	.35*		
Symbolization			.02		
Internalization			.09		
Step 2:				.32*	.16*
Dummy 1 (Arbitrator)	0.77	0.28	.38*		
Dummy 2 (VS)	0.80	0.29	.37*		
Symbolization			-.23		
Internalization			-.36		
Dummy 1 x Symbolization			.29		
Dummy 2 x Symbolization			.05		
Dummy 1 x Internalization			.20		
Dummy 2 x Internalization			.55*		

Note. $N = 55$. Reference Category for the Role Dummies = Bystander. Internalization and Symbolization were z-standardized.
* $p < .05$.

4 Manuscript 3 – When Arbitrators Prolong Conflicts

Schwabe, J., & Gollwitzer, M. (2019). *When arbitrators prolong conflicts: Beneficial role-taking effects motivate resistance against conflict resolution*. Manuscript under review.

Abstract

When people witness conflicts succeeding moral transgressions in their group, their reactions can be conceptualized as social roles. Previous studies show that taking a prosocial role fosters actors' moral situational self-concepts. However, since the roles cease to exist as soon as the situation resolves, this positive effect can only be sustained as long as the conflict remains unsolved. Based on this reasoning, we tested the following hypotheses in two studies (N=170 and N=107): (1) arbitrators, compared to bystanders, have a higher motivation to keep their role, even if that means to prolong the conflict; (2) this effect is mediated by positive effects of prosocial role-taking on the moral situational self-concept. Data from both studies support the hypothesized indirect effect where prosocial role-taking fosters the motivation to prolong the conflict through benefitting the moral self-concept. The results are discussed against their theoretical background: Role-Taking Theory.

Keywords: interpersonal conflict, role-taking, social roles, morality, self-concept

Introduction

Imagine you are leading an organizational team and a conflict developed between two members. Paul and John created an awesome product together, but Paul feels as if John was taking all the praise for himself. The issue appears to be intractable, keeps the team members from their work, and endangers the team climate. Thus, you, as the leader of the group, suggest a way to solve the matter. Would not everybody in the team be happy about your move, especially those team members who made themselves available as arbitrators of the conflict? Although the most reasonable answer to this question may be "yes," it may sometimes in fact be "no." In the present article, we explain this phenomenon on the basis of a novel role-theoretical approach to interpersonal conflicts: Role-Taking Theory (RTT).

Social Roles in Norm Conflicts

The conflict between John and Paul is a typical example of a norm conflict. Although John and Paul are the core conflict parties in this situation, other team members become necessarily involved in it. Notably, the behavior of such third-parties can be decisive for the course and outcome of the conflict (Skarlicki & Kulik, 2004). While prior research on norm conflicts has mainly focused on the victim and the perpetrator (Shnabel & Nadler, 2008), third-party reactions in conflicts has largely escaped scholarly attention so far. Role-Taking Theory (Schwabe & Gollwitzer, 2019a) is a social-cognitive framework explaining third-party behavior in conflicts, but also in other social situations.

In RTT terms, third-party reactions to conflicts are conceptualized as social roles. Specifically, five roles can be distinguished: victim supporter, perpetrator supporter, arbitrator, escalator, and bystander (see Figure 1). The roles can be derived from (1) the partiality of a role (victim side, neutral, perpetrator side), and (2) the role's motivation to solve the conflict (Schwabe & Gollwitzer, 2019a). Supporters of one side (i.e., victim or perpetrator) typically have no motivation to solve the conflict; bystanders are neutral on both dimensions, escalators are motivated to keep the conflict going, and arbitrators should be

motivated to end the interaction by mediating between the two parties. However, RTT does not only provide a classification of the positions in conflicts, but also describes a process specifying how taking over a specific role is intricately connected to actors' self-concepts. This process explains why actors sometimes "stick" to their roles, even when this means to prolong the conflict.

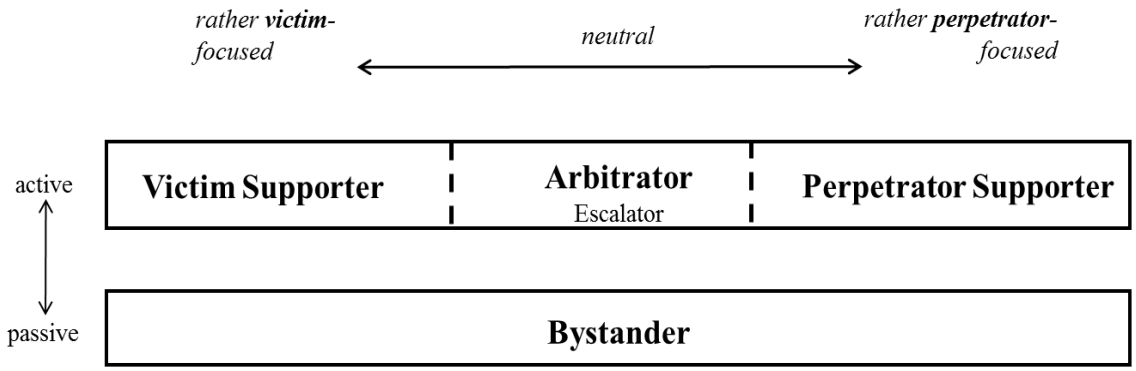

Figure 1. Schemata of third-party roles in conflicts following RTT.

Role-Taking as a Social-Cognitive Process

As explained above, the arbitrator role is an active role with the goal to maximize the joint outcome of all parties. Thus, the arbitrator role is a moral role: performing well in that role reflects positively on the (situational) moral self-concept. In line with this reasoning, experimental data shows that the role-taking process leads people to experience themselves as more moral when they take over the arbitrator role than when they take over the bystander role (Schwabe & Gollwitzer, 2019b). However, the moral self-boosting effect only lasts as long as the conflict keeps going. Building on the premise that people – and particularly arbitrators – are motivated to perceive themselves as moral (Brambilla & Leach, 2014; Schwabe & Gollwitzer, 2019b), a paradoxical hypothesis emerges: arbitrators experience themselves as more moral than bystanders in norm conflicts, and this, in turn, motivates them to keep their role, even when that means to keep the conflict going. This motivation, however, is incompatible with the central role goal of the arbitrator, that is, to solve the conflict. If this hypothesis were substantiated empirically, it would deliver an explanation why norm conflicts

are sometimes so hard to solve: even third-parties in moral roles may sometimes have a motivation to keep conflicts going, because they profit from their roles in regard of their moral situational self-concept.

The Present Research

The present paper describes two studies empirically testing the following hypotheses:

3) Arbitrators express more resistance against a conflict resolution that would make their role unnecessary, even if this alternative conflict resolution would be more likely to solve the conflict.

4) This effect is mediated by the moral situational self-concept: during the conflict, arbitrators experience themselves as more moral than bystanders, and this motivates resistance against an alternative conflict resolution.

In Study 1, participants' role choices in a conflict scenario were measured online. After taking over a role, participants rated five facets of their situational self-concept during the conflict, including the moral facet. After a turn of events in the conflict, their resistance against the solution of the conflict was measured via self-reports. In Study 2, we replicated Study 1 with an experimental lab study design: instead of measuring students' role choice, they were randomly assigned to either the bystander or the arbitrator role in a conflict acted by confederates. Again, the moral situational self-concept during the conflict and the resistance against a solution were measured.

STUDY 1

Method

Sample

Participants were recruited via a university-wide mailing list including students and university staff members. All participants completed the study online using the survey platform SoSci Survey (Leiner, 2014). As a reward, a Tablet PC was raffled. In total, 222 participants completed the online study, out of which 171 chose the bystander or the mediator

role. One participant had missing data on the dependent variable. The final sample consisted of n=170 (65% women) between 20 and 61 years (M=38, SD=11.1). The sample size was the result of our efforts to ensure statistical power for test of the indirect effect postulated in H2. Simulation analyses suggested that 148 participants are needed to detect a small- to medium-size effect with a power of .80 (Fritz & MacKinnon, 2007).

Measures and Procedure

Participants first completed two personality scales[11]. Next, demographics were measured (age, gender, level of education, native language, and occupation).

Role-Taking. Participants imagined being a member of an organizational working group, in which a norm conflict evolved (see Supplementary Material A, for all vignettes). The vignettes had already been used in a previous study (Schwabe & Gollwitzer, 2019b). Participants were told that the described conflict was being debated in team meetings. In these meetings, the perpetrator denies his or her wrongdoing, while the victim demands some kind of acknowledgment and restitution from the accused team member.

To ensure attentiveness, we included two items asking for the names of the perpetrator and the victim of the original norm violation. Participants re-read the vignette if they answered incorrectly. Next, participants were asked how they would (1) feel, (2) think, and (3) behave during the meetings in open format. Participants then chose a label for their role (i.e., victim supporter, perpetrator supporter, bystander, or arbitrator[12]). The supporter roles were labeled 'supporter of [*name*]' to avoid biases. By assessing role-choice in an open format first and then asking them to choose a respective label for their role, we prevented

[11] The scales were a 10-item Moral Identity scale (Aquino & Reed, 2002), measuring the centrality of a set of moral traits (e.g., caring, fair, helpful) to a person's self-concept, and the 10-item Observer Sensitivity scale (Schmitt, Gollwitzer, Maes, & Arbach, 2005; Schmitt, Baumert, Gollwitzer, & Maes, 2010), measuring the extent to which people react sensitively to observed injustice. Both measures were included to provide us with the opportunity to replicate other findings regarding the correlation between personality and role choice; however, they are not relevant for testing the present hypotheses.

[12] The escalator role was not included in Studies 1 and 2, because we considered it not very plausible that anybody would take that role in the present context.

participants from choosing their role based on the label[13].

Dependent Variables

Situational self-concept (SSC). After role-taking, participants were asked to imagine themselves in their specific role in the meetings. Participants were asked how acting in their role would make them think and feel about themselves (8 items, Supplementary Material B). The items were selected from a larger pool piloted elsewhere (Schwabe & Gollwitzer, 2019b) on the basis of a factor analysis. The (1) "moral situational self-concept" was assessed with one item that had the highest factor loading in the pilot study ("What I said and did in the situation shows that I am a moral person.")[14]. Two further facets of the SSC were (2) general self-esteem (two items, r=.58, e.g., "Because of the things I said and did in the situation, I am satisfied with myself," adapted from Rosenberg, 1979), and (3) sense of power (two items, r=.40, e.g., "I think I had some power in the situation," adapted from Anderson & Galinsky, 2006).[15] Two self-developed secondary facets of the SSC were (4) meaningfulness (one item: "What I said and did in the situation was meaningful and important"), and (5) expression of identity (two items, r=.38, e.g., "What I said and did in the situation is an expression of my personality"). Response scales ranged between 1 ("not at all") and 6 ("absolutely").

Resistance against conflict solution. Participants read that the boss had suggested a solution to solve the conflict (see Supplementary Material A). Resistance was measured with five items (α=.88, e.g., "I think there are good arguments against the solution suggested by the boss;" "I am going to discuss the idea of continuing the weekly meetings with the boss;" full list in Supplementary Material C).

---

[13] Two independent raters blind to the hypotheses checked the congruence between reaction and label and were instructed to flag any cases in which reaction and label were incongruent. We decided a priori to remove cases in which both raters flagged such incongruence. This was, however, not the case.

[14] Item- and scale-level analysis (including all 8 items, principal axis factor analysis with Oblimin-direct rotation, Supplementary Material B), assuming a 5-factor model, showed that this item represented a unique factor (loading=.65), which explained 37.8% of the observed variance.

[15] Because participants in previous studies criticized a forced choice regarding the power items (e.g., because they were somewhat difficult to answer without information about the reaction of the other actors involved in the conflict), participants were allowed to have missing answers on the two power items. Nineteen participants used this option (16 arbitrators, 3 bystanders). All other items had to be answered.

Results and Discussion

Correlations and descriptive statistics can be found in Table 1. To test Hypothesis 1, mean values on "resistance" were compared between arbitrators (M=3.17, SD=1.36) and bystanders (M=2.82, SD=1.38). Despite a tendency in the hypothesized direction, arbitrators and bystanders did not differ significantly in their resistance, $t(168)=1.23$, $p=.22$, $d=0.26$, 95% CI for d [-0.25, 0.76]. Thus, Hypothesis 1 was not supported. We will come back to this in the General Discussion.

Notably, a non-significant total effect of role choice on resistance does not imply that testing Hypothesis 2 – an indirect effect via moral SSC – is impossible (MacKinnon & Fairchild, 2009). Thus, we tested Hypothesis 2 despite the absence of the total effect. Looking at the effect of role choice on the moral SSC ("a-path"), we found that – as predicted – arbitrators experienced themselves as more moral than bystanders during the meetings, b=0.45, 95% CI for b [0.20, 0.69]. Looking at the (direct) effect of the moral SSC on resistance ("b-path"), we also found a positive correlation, suggesting that the more moral participants felt in that situation, the more they expressed resistance, b=0.25, 95% CI for b [0.08, 0.41]. Finally, we looked at the indirect effect of role choice on resistance via moral SSC ($N_{Bootstrap}$=5000, using PROCESS, v2.16.3, Hayes, 2013). This effect was also significant, b_{indirect}=0.11, 95% CI for b [0.03, 0.26], leaving a small and non-significant direct effect of role choice on resistance ("c'-path"), b=0.06, 95% CI for b [-0.23, 0.35]. To further scrutinize this result, we specified a multiple mediation using the three facets of the SSC (i.e., morality, power, self-esteem) simultaneously (using M*plus* 6, Muthén & Muthén, 1998-2011). In this model (see Figure 2), the moral SSC uniquely emerged as the only relevant mediator with a significant specific indirect effect, b_{indirect}=0.23, 95% CI for b [0.03, 0.43]. The direct effect of role choice on resistance was small and non-significant, b=0.13, 95% CI for b [-0.49, 0.75]. This pattern of results supports the assumed motivational mechanism through which role-taking as an arbitrator can lead to the stabilization of a conflict. This

paradoxical effect was small to medium in size. Furthermore, Study 1 was quasi-experimental, rendering alternative interpretations possible. In addition, role-taking was operationalized “as if”, not in real interaction. To test our hypotheses more rigidly, we conducted an experimental lab study.

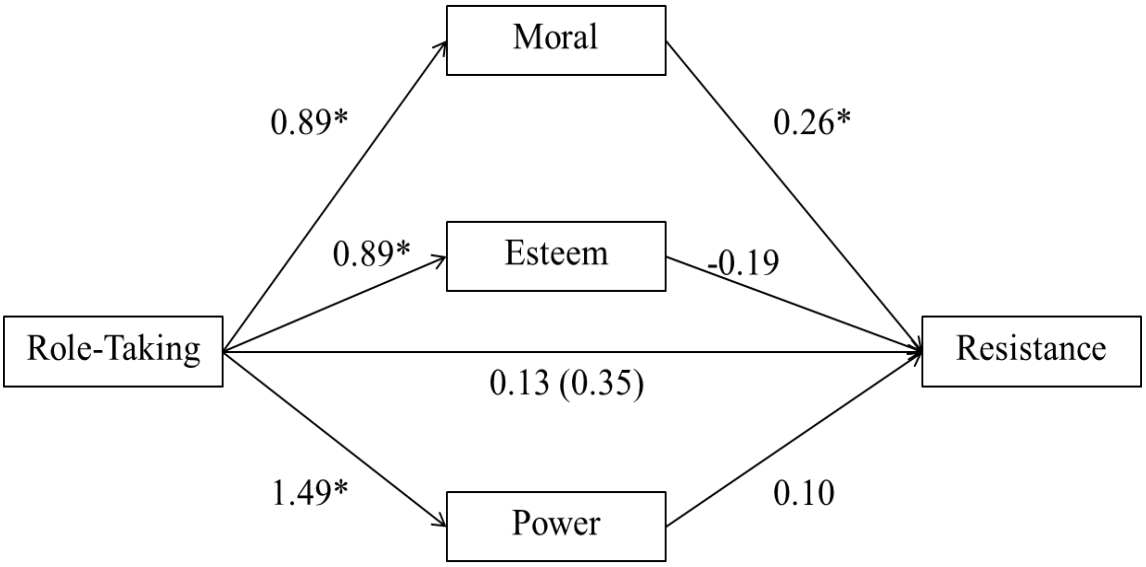

Figure 2. Multiple mediation path model with the situational self-concept facets self-esteem, morality, and power as mediators between role-taking (Dummy, arbitrator=“1”, versus bystander=“0”) and the resistance against the conflict solution in Study 1.

* $p < .05$

Table 1

Correlations among and Descriptive Statistics for all Measured Variables in Study 1

Measure	*M*	*SD*	1	2	3	4	5	6	7	8	9
1. Resistance against Solution	3.11	1.36	-								
2. Dummy 1: Arbitrator vs. Bystander	.84	0.37	.09	-							
3. SSC: Moral	3.96	1.26	.24*	.26*	-						
4. SSC: Self-Esteem	5.03	0.93	.07	.36*	.34*	-					
5. SSC: Power	3.44	1.13	.13	.50*	.37*	.52*	-				
6. SSC: Expression	5.11	0.87	.06	.21*	.13	.42*	.22*	-			
7. SSC: Meaning	4.48	1.13	.11	.40*	.43*	.63*	.52*	.33*	-		
8. MI-Internalization	5.02	0.78	.18*	.13	.40*	-.05	.11	-.02	.11	-	
9. MI-Symbolization	2.98	1.06	.23*	.16*	.36*	.14	.25*	.01	.22*	.49*	-
10. Observer Sensitivity	4.35	0.86	.19*	.09*	.40*	-.05	.06	.05	.12	.37*	.31*

Note. N=170, except for the correlations including the SSC: Power scale, where 19 participants had missing values. MI=Moral Identity. SSC=Situational Self-Concept. All scales ranged from one to six.

* $p < .05$.

STUDY 2

Method

Sample

One-hundred and seven students ($M_{Age=}$24, n=83 female) were recruited. They were compensated either in cash or by certifying course credit. Given the effect size found in Study 1, a sample size of 140 would have been required to detect this effect with a power of .80 (Fritz & MacKinnon, 2007). However, this ambitious goal could not be achieved because of sampling constraints towards the end of the academic semester. Eventually, we were able to recruit 107 participants by the end of the academic semester.

Cover Story

Students were told that the department had recently been discussing the issue of free-riders in student groups, that is, students who did not contribute adequately to a group task. Allegedly, there had been a number of conflicts among students about this issue. The department has now decided to invite the conflict parties and discuss their issue with a third person (i.e., the real participant).

Measures and Procedure

Role-Taking. Participants took part in a fake event scheduling, where they had to appoint a date to meet the students in conflict. In the lab, participants read a standardized instruction repeating the cover story and their role-assignment of either bystander or arbitrator (see Supplementary Material D). Bystanders were instructed to stay passive in the conflict, to solely observe and to refrain from taking anyone's side. Arbitrators were told to actively mediate between the conflict parties with the ultimate goal to solve the conflict. To standardize their behavior, arbitrators were told to collect and write down the arguments of both actors on a sheet, and to sum them up in paraphrases for the conflict parties in a break. Furthermore, arbitrators had to keep track of actors' speaking times to give both actors a chance to voice their view.

Dependent Variables

Situational Self-Concept (SSC). In the meeting room, after a brief introduction, the experimenter started an audio recording and the confederates started their conflict enactment using a standardized script. After approximately 15 minutes, the experimenter announced the break. During the break, the arbitrators summarized the exchanged arguments; afterwards, the moral SSC was measured with three items (one item of Study 1 plus two additional items, α=.90, e.g., "Based on how I played my role one can tell that I am a moral person").

The other four facets of the SSC were assessed with the same items as in Study 1: (1) general self-esteem (2 items, r=.54, e.g., "Because of the things I said and did in the situation, I am satisfied with myself;" adapted from Rosenberg, 1979), (2) sense of power (2 items, r=.39, e.g., "I think I had some power in the situation," adapted from Anderson & Galinsky, 2006), (3) meaningfulness of their reaction (1 item: "What I said and did in the situation was meaningful and important;" self-developed), and (4) expression of identity (2 items, r=.62, e.g., "What I said and did in the situation is an expression of my personality;" self-developed). Response scales ranged between 1 ("not at all") and 6 ("absolutely").

Ecological validity and evaluation of the enacted situation. During the break, participants answered 16 items to check if participants experienced the situation in the intended manner[16].

Resistance against conflict solution. Our dependent variable was assessed similar to Study 1. After the break, the conflict discussion went on for approximately ten minutes. The conflict remained unsolved. The confederates ended their "play" by asking the experimenter what to do now. The experimenter mentioned the opportunity of a second meeting, but this

[16] We assessed this ecological validity by measuring participants perception of the conflict (4 items, e.g., "It is not surprising that this situation resulted in a conflict"), the atmosphere during the conflict meeting (3 items, e.g., "The atmosphere in this situation is relaxed"), the relation between victim and perpetrator (3 items, e.g., "The two conflict parties are considerate with each other"), and the general valence of the conflict situation (6 items, e.g., "I do not perceive the conflict and the situation to be particularly bad").

would depend on the willingness of everybody involved[17]. Participants were then told that the department already has a plan for solving conflicts like this one quickly and reliably. In order to put this plan into action, the participant's involvement would no longer be required. However, the department would like to hear the participant's perspective about this plan. Six items measured the degree to which participants resisted against this solution (α=.84, e.g., "I think we should continue the process together," or "I would be glad if the conflict was over for me now;" reversed, see Supplementary Material E). Response scales ranged from 1 ("not at all") to 6 ("absolutely").

Manipulation check. Participants answered six items about the difficulty to become involved into the situation (1 item), to act in the assigned role (1 item), their performance in their role (2 items), and how much their role contributed to the conflict solution (2 items). To ensure that participants did not guess our hypotheses, we conducted a funneled debriefing. Finally, participants' demographics (i.e., age, gender, occupation, language skills) were assessed. Participants were then rewarded, carefully debriefed, and dismissed.

Results and Discussion

Correlations and descriptive statistics can be found in Table 2. Study 2 replicated the results from Study 1. While the hypothesized total difference in resistance between arbitrators (M=4.23, SD=0.98) and bystanders (M=4.15, SD=1.01, cf. Hypothesis 1) was not statistically significant, $t(105)=0.43$, $p=.67$, $d=0.04$, 95% CI for d [-0.34, 0.42], the specific indirect effect formulated in Hypothesis 2 was replicated in the experimental lab study. Looking at the effect of role on the moral SSC ("a-path"), we found that arbitrators felt more moral during the

[17] At this point of the experiment, we tried to assess participants' motivation to keep their role (and to keep the conflict situation going) with a behavioral measure: participants were asked whether they would voluntarily take part in a potential second conflict meeting. However, they would not be rewarded in cash or course credit for this second meeting. If agreed, the experimenter asked the participant for how long the room should be blocked for the next meeting, stating that any time between fifteen minutes and three hours was possible. Although participants' decisions to have another meeting and the time they would want to invest in such a meeting would have been a very face-valid behavioral measure, we could not use the responses in the present analyses because they did not sufficiently vary across participants: only 5 out of 107 participants said they were unwilling to meet a second time, and only 10 participants answered the question how much time they were willing to invest in a straightforward manner. The other participants shrugged or gave no specific answer at all or tried to discuss this issue with the two conflict parties or the experimenter.

conflict than bystanders, *b*=0.77, 95% CI for *b* [0.28, 1.25], and looking at the direct effect of the moral SSC on resistance ("b-path"), we also found a significant positive correlation, *b*=0.23, 95% CI for *b* [0.09, 0.38]. Finally, we found a significant indirect effect, $b_{indirect}$=0.18, 95% CI for *b* [0.05, 0.40], leaving a small and non-significant direct effect of role on resistance ("c'-path"), *b*=-0.09, 95% CI for *b* [-0.48, 0.29]. This suggests that arbitrators (vs. bystanders) are more likely to resist against the proposed alternative conflict solution because this would prevent them from reaping the moral benefits of taking the arbitrator role. Again, we investigated whether the effect was unique for the moral SSC (see Figure 3). The significant total indirect effect of *b*=0.29, *p* < .01, 95% CI for *b* [0.09, 0.49] consisted of two significant specific indirect effects, the "moral" facet, *b*=0.17, *p*=.04, 95% CI for *b* [0.01, 0.31], and the "powerful" facet, *b*=0.19, *p*=.04, 95% CI for *b* [0.01, 0.32]. Arbitrators felt more moral and more powerful than bystanders, which, in turn, motivated arbitrators to resist against the conflict solution. These specific indirect effects were small to medium in size (Fritz & MacKinnon, 2007). The direct effect of role on resistance over and above indirect effects was negative, but non-significant, *b*=-0.21, *p*=.27, 95% CI for *b* [-0.58, 0.16].

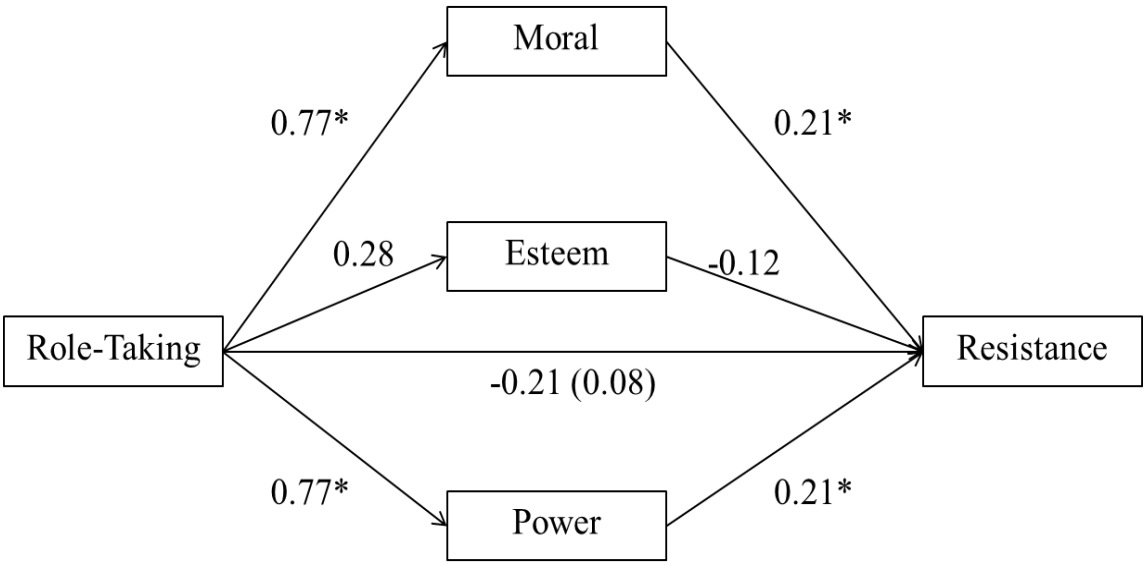

Figure 3. Multiple mediation path model with the situational self-concept facets self-esteem, morality, and power as mediators between role-taking (Dummy, arbitrator="1", versus bystander="0") and the resistance against the conflict solution in Study 2.

* *p* < .05.

Table 2

Correlations among and Descriptive Statistics for the central Measured Variables in Study 2

Measure	*M*	*SD*	1	2	3	4	5	6
1. Resistance against solution	4.19	0.99	-					
2. Dummy 1: Arbitrator vs. Bystander	.49	0.50	.04	-				
3. SSC: Moral	3.21	1.32	.29*	.29*	-			
4. SSC: Self-Esteem	4.83	0.98	.04	.14	.37*	-		
5. SSC: Power	2.68	1.20	.29*	.32*	.34*	.26*	-	
6. SSC: Expression	3.47	1.44	.06	.25*	.48*	.21*	.30*	-
7. SSC: Meaning	3.47	1.50	.27*	.38*	.65*	.49*	.48*	.44*

Note. *N*=107. SSC=Situational Self-Concept. All scales and items ranged from one to six.

* $p < .05$.

General Discussion

The present research tested two paradoxical hypotheses resulting out of a novel role-theoretical approach (RTT, Schwabe & Gollwitzer, 2019a). The hypotheses deducted from RTT and tested here shed light on why norm conflicts sometimes become "intractable" and remain unsolved, even though neutral and active third-parties with the goal to solve the conflict are involved (i.e., "arbitrators"). Data from a quasi-experimental online study and an experimental lab study support the hypothesis that arbitrators experience themselves as more moral during a conflict than bystanders, and that this, in turn, motivates their resistance against ending the conflict. This suggests that arbitrators are motivated to keep their role – even when that means to prolong the conflict – in order to continue experiencing themselves as a moral person.

On a broader level, this substantiates RTT's claim that third-party role-taking in norm conflicts is intricately intertwined with the situational self-concept, especially with the moral dimension. Resistance against a conflict solution was related uniquely to role-taking effects on the *moral* (in Studies 1 and 2) and the *powerful* dimensions (in Study 2). The latter effect, though not explicitly expected, is consistent with what has been discussed in the power literature. Power-Approach Theory (Keltner, Gruenfeld, & Anderson, 2003) states that power increases goal-directed behavior, but also increases social distance and hypocrisy, and reduces perspective-taking and external influence on behavior (Galinsky, Magee, Gruenfeld, Whitson, & Liljenquist, 2008; Galinsky, Magee, Inesi, & Gruenfeld, 2006; Lammers, Galinsky, Gordijn, & Otten, 2012; Lammers, Stapel, & Galinsky, 2010). Thus, our findings suggest that taking the arbitrator role has made actors feel more powerful, and this role-specific sense of power decreased their motivation to resolve the conflict. Notably, the indirect effect via sense of power only emerged in Study 2, but not in Study 1, so future research may want to scrutinize the replicability of this specific effect. Moreover, and more importantly, the indirect effect via sense of power in Study 2 existed above and beyond the hypothesized indirect effect

via the moral self-concept. Thus, the indirect effect via the moral self-concept is indeed unique – a finding which even strengthens the tenability of RTT's central assumption that arbitrators want to continue reaping the moral benefits of being in a pro-social role.

The fact that the data support the hypothesized indirect effect (i.e., Hypothesis 2), but not the total effect (Hypothesis 1) speaks against RTT at first glance. However, a non-significant total effect in the presence of a significant indirect effect does not necessarily mean that no "unmediated" effect exists (MacKinnon & Fairchild, 2009). Instead, it is possible that such an effect is suppressed by another (negative) effect. Importantly, such a negative effect is plausible, and, in retrospect, it makes sense: assuming that the arbitrator's goal is to solve the conflict, s/he should endorse any means to achieve this goal. In the situation described (Study 1) and enacted (Study 2) in the present research, this means to endorse the conflict resolution, and not oppose it. To put it more bluntly: two hearts seem to be beating in an arbitrator's chest. One is to fulfill his or her duties and solve the conflict, and the other may be to reap the benefits of having taken a pro-social role. Looking at the pattern of results in Study 2, this interpretation is plausible given the negative (albeit non-significant) direct effect of role on resistance. The pattern of results in Study 1, however, is not consistent with this interpretation. This suggests that further research should investigate the potential negative suppression effect discussed here more directly.

Limitations

We faced five methodological challenges in the present research: demand effects in both studies, selectivity of the sample in Study 2, sample size issues in Study 2, a potential experimenter artifact in Study 2, and difficulties in the assessment of behavioral dependent measures in Study 2. First, participants in both studies could have guessed our research hypotheses and tried to answer accordingly, for example, by reporting a lower moral SSC and a lower resistance as a bystander, but not as an arbitrator. However, the role was freely chosen in Study 1, making a demand artifact less plausible, and we checked for potential hypothesis-

guessing in Study 2: although 17 participants (18.2 %) reported feeling uncertain whether the conflict actually was real or acted, none of these sceptics guessed the true research question.

Second, regarding the selectivity of the sample in Study 2, the rather young, female, and academic student sample (M_{age}=24, n=83 female) raises questions of generalizability. For example, the indirect effect could be potentially caused by a stronger need for a moral SSC of young female students. However, previous research suggests that the need for a moral identity and the high importance of the moral domain in self-representation are universal and not restricted to a specific age or gender group (Aquino & Reed, 2002; Brambilla & Leach, 2014). Nevertheless, a replication with a more diverse sample is desirable. Taking the observed effect sizes into account, future studies should try to recruit at least 148 participants to reach a .80 power (Fritz & MacKinnon, 2007). We were not able to reach this number in Study 2 due to feasibility constraints. However, the morality-related indirect effect specified in H2 was also found in Study 1, where we reached a power well above .80.

As a further limitation, the team of experimenter and confederates in Study 2 was neither blind to the hypotheses nor to the experimental condition of the participants. That is, experimenter bias (Rosenthal & Fode, 1963) might have been an issue in Study 2. Potentially, participants could have been influenced by subtle cues of experimenter and confederates. However, the team members had no motivation to produce a hypothesis-consistent result, as the study outcome was irrelevant for them. Furthermore, we took several measures to prevent subtle experimenter cues in the procedure: (1) actors were instructed to follow the standardized script as closely as possible, (2) actors were thoroughly trained and supervised by the first author of the present paper, and (3) audio recordings were checked for deviations from the script after each session. Although experimenter bias can never be ruled out completely, it is only plausible in Study 2 – and, importantly, we also found the effect stated in H2 in Study 1.

As a final limitation, the behavioral measure of participants' motivation to keep their

roles (see Footnote 7) could not be used. Thus, it remains an open question when and how resistance translates to a behavioral level. A study which gives arbitrators and bystanders an opportunity to communicate with the conflict parties more freely and the assessment of subtle behaviors fostering vs. impeding a conflict solution appears to be a desirable next step.

Outlook and Conclusion

The present research delivered evidence for an effect that appears rather trivial at first glance: people can be motivated to keep an interaction going because they profit from keeping the role they took in the interaction. However, for arbitrators in a conflict, this effect becomes paradoxical, because an arbitrator *ought to* solve the conflict. Notably, the effect we found in our studies cannot be explained by other theoretical approaches (e.g., moral licensing, costly signaling, or behavioral consistency): resistance was neither clearly moral nor immoral, and thus, moral licensing is irrelevant here (Merritt, Effron, & Monin, 2010). Costly signaling theory (Gintis, Smith, & Bowles, 2001) does not apply because participants rated their resistance anonymously. And, finally, behavioral consistency (Festinger, 1957) cannot be reasonably applied here because – at least in Study 2 – arbitrators were explicitly instructed to solve the conflict. Thus, Role-Taking Theory appears to be best equipped to explain the effects reported here.

Further (role-theoretical) questions emerge from our research. There is no reason to believe that this paradoxical role-taking effect is restricted to the arbitrator role and the moral domain of the situational self-concept. Theoretically, it can be transferred to other roles situationally prompting other domains of self-related cognition and emotion. Analogous self-related mechanisms might motivate therapists, teachers, or CEOs to keep interactions going longer than necessary. Future studies answering these questions can build on RTT's role-theoretical framework and the present research to further investigate how role-taking effects on the situational self-concept motivate actors to keep their role and, therefore, the over-arching interaction.

Acknowledgements

For Study 2, Miriam Hausner, Sören Lauff, Leonora Schaefer, and Marisa Timm helped putting the experimental material together and conducted the lab study.

References

Anderson, C., & Galinsky, A. D. (2006). Power, optimism, and risk-taking. European Journal of Social Psychology, 36, 511-536. doi: 10.1002/ejsp.324

Aquino, K., & Reed, I. I. (2002). The self-importance of moral identity. Journal of Personality and Social Psychology, 83, 1423-1440. doi: 10.1037/0022-3514.83.6.1423

Brambilla, M., & Leach, C. W. (2014). On the importance of being moral: The distinctive role of morality in social judgment. Social Cognition, 32, 397-408. doi: 10.1521/soco.2014.32.4.397

DeCelles, K. A., DeRue, D. S., Margolis, J. D., & Ceranic, T. L. (2012). Does power corrupt or enable? When and why power facilitates self-interested behavior. Journal of Applied Psychology, 97, 681-689. doi: 10.1307/a0026811

Festinger, L. (1957). A theory of cognitive dissonance. Evanston, IL: Row Peterson.

Fritz, M. S., & MacKinnon, D. P. (2007). Required sample size to detect the mediated effect. Psychological Science, 18, 233-239. doi: 10.1111/j.1467-9280.2007.01882.x

Galinsky, A. D., Magee, J. C., Gruenfeld, D. H., Whitson, J. A., & Liljenquist, K. A. (2008). Power reduces the press of the situation: implications for creativity, conformity, and dissonance. Journal of Personality and Social Psychology, 95, 1450-1466. doi: 10.1037/a0012633

Galinsky, A. D., Magee, J. C., Inesi, M. E., & Gruenfeld, D. H. (2006). Power and perspectives not taken. Psychological Science, 17, 1068-1074. doi: 10.1111/j.1467-9280.2006.01824.x

Gintis, H., Smith, E. A., & Bowles, S. (2001). Costly signaling and cooperation. Journal of Theoretical Biology, 213, 103-119. doi: 10.1006/jtbi.2001.2406

Hayes, A.F. (2013). An introduction to mediation, moderation, and conditional process analysis. New York: The Guilford Press.

Keltner, D., Gruenfeld, D. H., & Anderson, C. (2003). Power, approach, and inhibition.

Psychological Review, 110, 265-284. doi: 10.1037/0033-295X.110.2.265

Lammers, J., Galinsky, A. D., Gordijn, E. H., & Otten, S. (2012). Power increases social distance. Social Psychological and Personality Science, 3, 282-290. doi: 10.1177/1948550611418679

Lammers, J., Stapel, D. A., & Galinsky, A. D. (2010). Power increases hypocrisy: Moralizing in reasoning, immorality in behavior. Psychological Science, 21, 737-744. doi: 10.1177/0956797610368810

Leiner, D. J. (2014). SoSci survey. Retrieved from https://www.soscisurvey.de.

MacKinnon, D. P., & Fairchild, A. J. (2009). Current directions in mediation analysis. Current Directions in Psychological Science, 18, 16-20. doi: 10.1111/j.1467-8721.2009.01598.x

Merritt, A. C., Effron, D. A., & Monin, B. (2010). Moral self-licensing: When being good frees us to be bad. Social and Personality Psychology Compass, 4, 344-357. doi: 10.1111/j.1467-8721.2009.01598.x

Muthén, L. K., & Muthén, B. O. (1998-2011). Mplus User´s Guide. Sixth Edition. Los Angeles, CA: Muthén & Muthén.

Rosenberg, M. (1979). Conceiving the Self. New York: Basic Books.

Rosenthal, R., & Fode, K. L. (1963). The effect of experimenter bias on the performance of the albino rat. Behavioral Science, 8, 183-189. doi: 10.1002/bs.3830080302

Schwabe, J., & Gollwitzer, M. (2019a). How stable positions in interpersonal interactions shape social cognition and behavior: A Role-Taking Theory. Manuscript under review.

Schwabe, J., & Gollwitzer, M. (2019b). Explaining third-party behavior in norm conflicts: A role-taking approach. Manuscript under review.

Schmitt, M., Baumert, A., Gollwitzer, M., & Maes, J. (2010). The justice sensitivity inventory: Factorial validity, location in the personality facet space, demographic pattern, and normative data. Social Justice Research, 23, 211-238. doi:

10.1007/s11211-010-0115-2

Schmitt, M., Gollwitzer, M., Maes, J., & Arbach, D. (2005). Justice sensitivity. European Journal of Psychological Assessment, 21, 202-211. doi: 10.1027/1015-5759.21.3.202

Shnabel, N., & Nadler, A. (2008). A needs-based model of reconciliation: satisfying the differential emotional needs of victim and perpetrator as a key to promoting reconciliation. Journal of Personality and Social Psychology, 94, 116-132. doi: 10.1037/0022-3514.94.1.116

Skarlicki, D. P., & Kulik, C. T. (2004). Third-party reactions to employee (mis) treatment: A justice perspective. Research in Organizational Behavior, 26, 183-229. doi: 10.1016/S0191-3085(04)26005-1

Supplementary Material

Supplementary Material A

Vignettes in Study 1

Vignette A
In the following, you will read about a conflict in an organizational team. Please read through the conflict situation carefully. While doing so, imagine being part of the team. Try to picture yourself in the situation as well as in any way possible.
At the core of the conflict are Matthias and Thomas. They both have been in the team for a long time and do not differ from each other regarding work experience or skills level. Both have strengths and weaknesses - each of them in their own area.
Matthias got instructed to create a new database system, which should be used in the whole company. Because of Matthias´s lack of experience regarding such systems, Thomas supported him. Together they created a very efficient system.
When Matthias presented the new system to his supervisors he earned much praise for his good work. Thomas´s assistance was never mentioned.
Thomas thereupon felt treated unfairly.
He accuses Matthias of passing him over to take all the credit for their cooperation alone.
Matthias tries to reject these allegations.
A conflict emerged from this situation affecting all team members. Thomas expects the other team members to also condemn Matthias´s behavior. Matthias expects the other team members to calm Thomas down.
… [Manipulation Check items, Next page]
Continue to imagine being part of this team. The conflict develops in the following way:
Meanwhile, the relationship between the two has suffered. Of course, this also has repercussions on the team climate. The conflict developed into a constant topic of debate. The company rules say that such cases should be dealt with inside of teams. Therefore, there are weekly team-meetings, where the conflict gets debated.
These meetings are obligatory for everyone until the conflict is solved.
The other team members take different positions regarding the situation:

1. Some try to actively mediate the conflict from a neutral position.
2. Some try to keep out of the conflict.
3. Some try to support Matthias.
4. Some try to support Thomas.

… [Role choice, SSC items, Next page]
Imagine the situation develops in the way described on the next page.
[Next page]
The weekly meetings have taken place three times now. A satisfying solution could not be reached yet. The negative influence on the work climate continues. The head of department wants to prevent the conflict to keep the team from work. The head of department therefore proposes to stop the weekly meetings and to solve the conflict in a different way.

Vignette B
In Vignette B, “Matthias” and “Thomas” were changed to “Sarah” and “Julia”.

Supplementary Material B

Table 3

Summary of Items and Factor Loadings for Promax Five-Factor Solution for the Situational Self-Concept Items (N =222).

	Factor loading					
Item	1	2	3	4	5	Communality
1. What I said and did in the situation shows that I am a moral person.	**.65**	.03	.01	-.04	-.01	.44
2. Because of the things I said and did in the situation, I am satisfied with myself.	.15	-.15	.04	**-.67**	-.25	.84
3. Because of the things I said and did in the situation, I feel like a loser. (r)	.01	.10	.03	**-.71**	.07	.54
4. What I said and did in the situation was meaningful and important.	.14	-.22	.24	-.32	**-.37**	.69
5. What I said and did in the situation has nothing to do with me as a person. (r)	.03	**.67**	.05	-.05	-.12	.54
6. What I said and did in the situation is an expression of my personality.	.04	.16	-.03	.01	**-.74**	.66
7. What I said and did in the situation probably did not have much effect. (r)	-.09	.10	**.69**	-.10	.04	.51
8. I think I had some power in the situation.	.18	-.07	**.58**	.08	-.04	.44

Note. Boldface indicates highest factor loadings. r = reversed.

Supplementary Material C

List of Items for the Resistance against the Conflict Solution in Study 1.

1. I think there are good arguments against the solution suggested by the boss.
2. I am going to discuss the idea of continuing the weekly meetings with the boss.
3. I think the boss's proposal makes sense. (r)
4. In my opinion, the chances to solve the conflict are better when continuing the path we started.
5. I am happy when these weekly meetings are over now. (r)

Supplementary Material D

Instructions for arbitrators and bystanders in Study 2

Instruction A (Bystander)

Problem description. This experiment should help with a problem that occurs in almost every department. It`s about problems in groups who prepare graded presentations together. The problem arises when individual group members do not fulfil their duty and do not contribute the part of the work they should be contributing. However, the same grade is often given to everyone because it is difficult for lecturers to know exactly who contributed how much to a presentation. The "free-riders" therefore take advantage of the commitment of other students, from which the others suffer. Such situations can lead to disputes between students and then generally burden the study situation.

How such disputes develop often depends on how other students react to the situation. With this experiment, we would therefore like to examine how the conflict parties react when an uninvolved person assumes a certain role in the conflict. We would also like to examine the impact of participation in such conflicts on those not initially involved, in order to estimate the burden of participation in such a dispute. The results will be used to develop a strategy for preventing such problems from burdening the study situation.

In order to be able to answer the question as realistically as possible, we conducted a short survey in the higher semesters to determine in which presentation groups there were conflicts. We asked the arguing members whether they would be willing to come to us for a short interview. In the interview, they shall describe the current state of the debate and their position. Your role as a third person is to adopt a certain role in the conversation. The students do not know that you as an initially uninvolved third person will be sent into the interview with a certain role.

Your role: **bystander / observer**

Role instruction. The behavior of third parties should be as clearly defined as possible to ensure experimental control. Therefore, please follow these instructions as close as possible during the interview.

Stay out of the discussion as far as possible. Listen carefully to the conflict parties, but do not intervene in the conversation. Do not let the conflicting parties lead you to take sides and please do not try to mediate the conflict. Therefore, be as neutral and passive as possible.

After a few minutes, we will have a break in the discussion in order to understand the state of mind of all participants. Please also do not interfere in the conversation afterwards. After the interview, we conduct a short follow-up interview with all participants.

Instruction B (Arbitrator)

Problem description. The description of the problem was the same for arbitrators and bystanders (see above).

Your role: **Mediator**

Role instruction. In your role, you should actively contribute to solve the conflict between the conflicting parties. However, as we have to ensure that the behavior of the mediators is as clearly defined as possible in terms of experimental control, you will mainly use two classic mediation strategies. Please follow these strategies:

1. *Objectification*

The view on the conflict is usually very subjective and different between the parties to the conflict. In order to soften the one-sided view of the two parties and to come to an objective discussion, the mediator often tries to work out objective points of view which are then confirmed by the conflict parties ("validated"). Therefore, please collect the parties' arguments on the note pad. You will read these to the parties once the examiner asks you to do so and ask the conflict parties if they can confirm these objective positions.

2. *Fair communication*

It is important in a dispute that the conflict parties feel treated fairly and able to present their views adequately. Therefore, during the conflict, please ensure that none of the parties exceeds the maximum speaking time of 30 seconds per comment with the stopwatch. If one party exceeds their speaking time, please point out how important it is to give the other party time for their comments.

Please limit yourself as far as possible to these two strategies and do not bring own arguments into the discussion in order not to risk experimental comparison.

After a few minutes, we will have a break in the discussion in order to understand the state of mind of all participants. Please do not interfere in the conversation afterwards. After the interview, we conduct a short follow-up interview with all participants.

Supplementary Material E

List of Items for Resistance against the Conflict Solution in Study 2.

1. I think there are good arguments to integrate me (and participants in general) in the further process.
2. I would not like it if the conflict parties would have to sort the conflict out between each other.
3. I think we should keep continuing the process together.
4. I am glad when I do not have anything further to do with the conflict. (r)
5. I am glad when the conflict is over for me now. (r)
6. I think it would be wrong to leave the conflict parties alone with the conflict.

5 Final Discussion – State and Future of Role-Taking Theory

5.1 Introduction

In this chapter, I provide a transparent and self-critical evaluation of the present state of Role-Taking Theory (RTT). Throughout, I follow recent suggestions aimed at (re-)establishing scientific practices that foster efficiency and reliability in scientific progress (Glöckner, Fiedler, & Renkewitz, 2018). The publication of RTT in form of a theoretical stand-alone article (Manuscript 1) and the publication of the first two empirical articles (Manuscript 2 and 3) are important milestones on RTT's developmental road: from a first three-page outline to – hopefully – a textbook state of theory. However, many further steps need to be taken before RTT can be enshrined in the textbooks of social psychology with its own chapter. In this final discussion, I will lay ground for the next steps by evaluating the present state of the theory and by outlining the most critical challenges to be overcome in the future. How substantial is the contribution of RTT in the present state over and above existing theory? Do we really need the role-concept or can we explain the same phenomena more parsimoniously by focusing on the associated behavior? Where are the most pressing areas for improvements in the near future? Which directions of theory development are most promising? To which extent is RTT empirically corroborated, for example, through the results presented in Manuscript 2 and 3? And which parts of the theory should have the highest priority for further empirical testing? Although Manuscript (MS) 1 already contains some answers to these questions, the scope of a journal article limits their comprehensiveness, and thus, they deserve some afterword. While debating these questions, I will alternate my stance between a critical "reviewer" and an enthusiastic "advocate" and try to come to valid conclusions by integrating both perspectives. In general, this evaluation is also aimed to offer support for anyone continuing to deepen the role-theoretical understanding of social behavior as well as to underline the contribution of this dissertation over and above existing theory.

5.2 Assessment of Empirical Content

Empirical content. In his seminal epistemological work (1934/2005), Popper argued

that a theory – even before it is entered into the reiterative cycle of empirical testing and reformulation to assess its *degree of corroboration* (Bewährung) – should be evaluated according to its *empirical content* (empirischer Gehalt) over and above competing theories. The empirical content of a theory increases with its degree of falsifiability: the stronger the incorporated laws and the more prohibitive the derivations, the higher the potential to be falsified and thus, the higher the empirical content of a theory. The two criteria deciding over the degree of falsifiability are a theory's *level of universality* (Allgemeinheit) and its *degree of precision* or *specificity* (Bestimmtheit, Popper, 1934/2005). A useful theory needs both: a high level of universality means that it claims validity in many situations, and a high degree of precision means that it does not allow many "subclasses" in realizations of predictions (Glöckner & Betsch, 2011). A theory that is superior to another with respect to one of the two criteria has "unique" empirical content and constitutes a scientific advantage (if corroborated empirically). Before comparing RTT with competing theories, I assess it independently by discussing its universality and precision including potential improvements.

Universality. To judge universality, the most exclusive antecedent condition of a theory is decisive. To be able to compare antecedent conditions (and for other reasons, see Chapter 5.9), I formalized the present state of RTT in Appendix A. As of the still very universal character of RTT's most exclusive antecedent condition (see statement 5, Appendix A), RTT has a high level of universality: It claims validity for any interpersonal interaction and all persons who abstracted a mental schema about their position in an interaction, and thus, arguably, for a large majority of people in interactions. However, it can be debated how large the proportion of "schemata-abstracting people" really is. On the one hand, taking the critical reviewer stance, it appears obvious that statement 3 is over-inclusive: not *all* people but only the ones fulfilling certain preconditions might be able to abstract a mental schema from repetitive encounters with an interpersonal interaction (e.g., people who are able to form second-order representations as in Theory of Mind paradigms, Premack & Woodruff, 1978;

Schaafsma, Pfaff, Spunt, & Adolphs, 2015). On the other hand, it is difficult to draw clear lines in that regard. Even people with somewhat impaired cognitive abilities might gather *some* knowledge tied to positions in interactions (Premack & Woodruff, 1978). However, an extension of the theory would be needed to obtain precise predictions about the role schemata in these cases, because they are arguably less influenced by objective features of an interaction (i.e., contain less of the *fixed* associations), and more influenced by *random* processes (see below). In the most extreme case, people with certain impairments might develop role schemata completely independent of any objective features of interactions. Applying RTT to such cases would take away parts of its utility, because deriving hypotheses from universal features of interactions would be impossible. In conclusion, RTT would profit from more refined inclusion criteria, somewhat reducing its universality. Nevertheless, even taking a reduction as mentioned above into account (i.e., to all people who are able to form second-order representations), RTT still would be a very universal theory with regard to the proportion of individual actors it claims validity for. Furthermore, it is not plausible to assume that its validity is restricted to a certain cultural context or period of time because the *fixed* components derive from stable universal features and the *random* components are explicitly conceptualized as malleable. To expand universality even further, the theory could easily be extended to also incorporate groups as actors (see Section 4 in MS 1). An open question regarding universality is whether RTT also applies when one of the interactions partner is non-human or merely imagined, for instance, in human-machine interactions or in spiritual interactions like praying.

Precision. RTT can and needs to be improved in its degree of precision. There are three main areas that bear potential for improvement. A first area is concerned with how components of role-schemata get processed in the course of role-taking and whether the present state of RTT is precise enough in its definition of the role-taking process. In a dedicated chapter (5.3), I argue that RTT would profit from focusing more on the basal

cognitive processes involved in role-taking and on empirical corroboration in corresponding paradigms, for example, regarding memory encoding and retrieval effects. Two further highly pressing areas that directly fall into RTT's field of responsibility are a comprehensive set of rules on how to precisely obtain hypotheses about both the *fixed* and the *random* components of role schemata, and a better specification of the nature of role-schemata, for example, by offering a finite categorization of elementary types constituting the slots of role-schemata. After all, the precision of RTT role-taking predictions will increase with knowledge about the role schemata: the more complete the set of known slots (i.e., types of components such as behavioral expectations, self-relevant associations, symbols, etc.) and known components of these slots, the more precise the predictions. That is, I argue that RTT needs to provide more guidance to enable its users to specify role-taking hypotheses in a clear-cut and standardized manner. Potential issues and refinements in these two areas are discussed in two designated chapter below (5.4 and 5.5).

The more refined and precise RTT will become in these three areas (i.e., in shedding light on the role-taking process, in guidance on how to derive specific schemata-components, and in specifying the nature of role-schemata), the more impact it will have in the long run. Imagine Tom, a researcher using RTT to predict role-taking effects in the classroom. The precision of Tom's predictions about effects of taking the "teacher" role depends on how precisely he can obtain the role-specific schema using RTT's rules. He thus needs to know which elementary types and which exact components are supposed to be part of the teacher role-schema according to RTT. Afterwards, he needs to know as precisely as possible when which component will be *activated* and *used* in response to which stimuli through or during role-taking. In the following, I argue that the present state of the theory is not comprehensive enough to reach an acceptable degree of precision in all interactions it claims to be applicable on, and thus still requires sustained theoretical and empirical work to fulfill its promises. However, I also argue that RTT has enough potentially unique empirical content to make this

future work worthwhile.

5.3 Improving Precision: The Cognitive Basis of Role-Taking

First issue: more guidance on activation and use of knowledge due to role-taking. In the present state, RTT gives away some of its potentially unique precision because there is no explicit differentiation between *activation* and *use* of role-specific schema-components, and no further guidance on questions surrounding these two topics. RTT states that role-taking *activates all associations* tied to a position, which is a general and not very refined definition of a presumably quite complex cognitive process, especially if one takes into account that the accessibility of role-specific schema-components most likely changes over the course of one role-taking instance, for example, in dependence of the occurrence of additional role-related stimuli (e.g., a student asking a question would be a role-related stimuli for a teacher). In the worst case, readers could even wrongfully interpret that RTT claims that role-specific components would be *inevitably* used in response to any subsequent stimuli in the interaction. To be very clear about this, activation in the sense of Higgins (1996) means that some piece of knowledge (e.g., a proposition in a role-schema) is *temporarily increased in accessibility,* that is, that the likelihood to be applied to some object in response to subsequent stimuli is increased. However, this likelihood to be used also depends on *applicability,* that is, whether a given piece of activated knowledge fits attended features of a stimulus and *judged usability,* that is, whether applying knowledge to some stimulus is judged as appropriate (Higgins, 1996). RTT presumably has quite some things to say about these matters, as a comprehensive analysis of an interaction and an embedded role-schema would not only imply which components are generally available and then generally activated through role-taking, but also which situational triggers most likely activate which specific components the most, and which components of the schema are applicable to which specific objects in the interaction.

In this line of thought, RTT's potential shows nicely: RTT assumes that roles are the

"boxes" in which interpersonal social cognition is sorted and packaged. This implies that a role-related stimulus does not only trigger one specific item out of the box but always "opens the whole box:" role-taking always triggers a multifaceted set of effects that is unique for each role. If future empirical work could empirically elucidate this "packaging"-character of role-schemata, that is, how a role-related stimulus triggers not only stimulus-specific components but also additional components usually unrelated to the stimulus, and RTT could comprehensively predict these additional components, then RTT's impact would benefit strongly.

Thus, future iterations of RTT need to satisfy the need for more precision in this regard, increasing its falsifiability as well as its potential impact. However, as a prerequisite, more knowledge about role-specific schema-components seems to be necessary (see Chapter 5.4). Once RTT offers safe and sound methods to derive schemata-components, RTT's impact would benefit strongly from testing whether role-schemata have similar functions in directing attention and processing information as other more basal schemata (e.g., in paradigms similar to Brewer & Treyens, 1981).

Second issue: role-taking and dual systems models. Readers might ask themselves how role-taking relates to the "reflective" system and the "impulsive" system in dual systems terms (e.g., Strack & Deutsch, 2004; Gawronski & Bodenhausen, 2006, 2011). Relating role-taking to such models in a comprehensive manner would make for a theoretical paper of its own, but a little afterword to the respective section in MS 1 (see Section 2, Premise 6) seems appropriate. To elaborate, dual system models are a group of very influential theoretical models sharing the same basic assumption that there are two cognitive systems with different principles of representation and information processing. Both systems work simultaneously and compete for behavioral responses. The two systems are the "impulsive" (or "associative") system, in which cognitive elements are stored in form of an associative network (Smith, 1998), and the "reflective" (or "propositional") system, in which elements are connected

through semantic relations to which a truth value is assigned (Strack & Deutsch, 2004). In regard to storing cognitive elements, Strack & Deutsch (2004) assume – along with many others (see Smith, 1998) – that the associative network underlying the impulsive system is relatively stable and can be thought of as a long-term memory system, while the reflective system is more flexible but can only contain a small number of propositions simultaneously which fade away quickly without rehearsal (i.e., the reflective system as a temporary storage can be thought of similar to the working memory, Baddeley, 1986).

First of all, RTT implies that role-taking affects both systems. That is, with respect to the impulsive system, role-taking should have long-term effects by shaping the associative network underlying the impulsive system, as well as short-term effects by temporarily increasing the accessibility of the role-specific schema. One implication that RTT carries for the content of the associative network underlying the impulsive system (Strack & Deutsch, 2004): since the associative strength between the elements in this network is contingent on the frequency of their co-occurrence, and given that roles have the socio-structural functions proposed by RTT, then the associative strengths in this network of a given person should at least partly develop as a function of the roles this person takes over. In other words, each time a role gets taken it leaves its unique footprints in the associative strengths because each role comes with its unique but universal package of social cognition. Therefore, role-taking should be an important factor in the development of the associative network.

With respect to the reflective system, RTT's role-taking concept is in line with the idea that the relations between elements of the reflective system cannot only be simple logical relations such as *is a, is not,* or *implies,* but also social relations "… such as friend, enemy, spouse or partner" (Strack & Deutsch, 2004, p. 225). Thus, if RTT provides the exact properties of role-schemata, it informs both the impulsive and the reflective pathway. The most interesting question in that regard is whether the different types of schema-components differ in the degree to which they play out not only in the impulsive but also in the reflective

system. For example, the results of MS 3 could also be read as a conflict between an impulsive response motivated by self-related associations and cognitions and a reflective response motivated by behavioral expectations: arbitrators feel more moral during conflicts and the motivation to keep this effect might create an impulsive resistance against the conflict solution. To respond in line with the behavioral expectations tied to the arbitrator role (i.e., to foster or at least not withstand a solution) arguably requires more cognitive effort, and thus, this second process most likely takes the reflective route. Being able to theoretically and empirically separate such processes more clearly would increase RTT's precision. Thus, the features of dual systems frameworks should be kept in mind when deriving role-taking hypotheses and empirical studies should investigate when which schema-components lead to impulsive versus reflective reactions. In experimental studies, cognitive load paradigms could be used to distinguish between reflective versus impulsive pathways (e.g., as in Gollwitzer, Braun, Funk, & Süssenbach, 2016).

One minor issue related to RTT from a dual systems perspective is that role-schemata can contain quite a high number of propositions. For example, Tom might derive quite a long list of propositions that logically co-vary with pursuing the goal to teach (e.g., behavioral expectations like teachers are expected to answer students' questions). Combined with the fact that role-taking is said to *activate all* of them, and without further guidance on the issues of *applicability* and *use* described in issue one, the "critical reviewer" could ask whether the present state of RTT does not contradict an assumption of dual system models, namely that only a few propositions can be used in the reflective system at the same time (Strack & Deutsch, 2004). However, as noted above, activating a high number of propositions of a role-schema in the role-taking process does not mean that all of them are simultaneously used in each and every subsequent instance of response. Still, this is another reason why this topic needs to be investigated.

Summary. To sum up, the role-taking process as the activation of all associations tied

to a position-specific schema fits well into the discussed classic and more recent (social-) cognitive theorizing. However, as of now, RTT has not been rigidly tested in one of the classic paradigms originally demonstrating the importance of cognitive schemata, for example, focusing on information processing or memory retrieval (e.g., Brewer & Treyens, 1981). Since the idea of roles as schemata is the cognitive foundation on which RTT relies, RTT would highly benefit from a line of studies illuminating how role-taking influences basic cognitive processes. This line of studies should be implemented with high priority in future empirical work. Once this foundation is laid, RTT needs to concern itself with more precise guidance on how specific components of a role-schema are not only increased in accessibility due to role-taking but actually *applied* to which objects in an interaction under which conditions. In this line of studies, RTT would benefit from studies illuminating the "packaging"-character of role-schemata, that is, whether stimuli can trigger usually unrelated components via the activation of role-schemata.

5.4 Improving Precision: Deriving Specific Schema-Components

Present state. RTT's present state provides one very general rule on how to derive the components of the role-schemata as a basis to predict role-taking effects: components have to be derived deductively from the positions' goals, because the goals (and their interrelations) are universal features of the prototypical social structure (see Section 3 in MS 1). That is, when Tom tries to theoretically derive the role-schemata for "teacher" and "student", he would deduct components that co-vary with pursuing the goals to teach and to learn. Until now, RTT would supply Tom mainly with the slots "behavioral expectations" and "self-related cognitions and emotions", and components thereof. For example, Tom might arrive at the prediction, that teachers are associated with being knowledgeable persons, and this self-related cognition might affect the situational self-concept of proponents through role-taking. How would Tom derive the "knowledgeable" component? Tom would argue that teachers have to know more than their students as a precondition to fulfill the role's goal. Thus,

proponents of the role are – or at least appear to be – more knowledgeable than proponents of their complementary role in nearly every instance of interaction, and thus, this component should be learned as a *fixed* part of the teacher role. This deductive path would also be the first answer RTT gives on how to derive the elementary *types* of components tied to roles: a role schema contains all the elementary types that *logically* co-vary with the role-specific goal. Unfortunately, this deductive path cannot be applied to *random* components, which is an issue in itself (see second issue below), and although the *fixed* components should find their way into role schemata as byproducts of goals, the deductive path is often more complicated than in this example, and it only produces valid assumptions under certain preconditions. These might not always be fully met, as I discuss in the following. As this might prove as more of an issue for components relevant for reflective role-taking processes (e.g., the role-components feeding into syllogistic reasoning), Tom might need more guidance especially in this regard.

First issue: consistency of co-variation between schema-components and roles and its effect on representation. Some components *always* co-vary with pursuing a given goal, but others only co-vary under certain conditions, and thus, elements of the role-specific schema can be assumed to differ in associative strength (Strack & Deutsch, 2004). This also feeds back into specifying the role-taking process. Differences in associative strength of schema-components might lead to a hierarchy of role-taking effects: the higher the default associative strength of a component within the role-schema, the higher the increase in accessibility due to role-taking. Although this would be important for precise role-taking predictions, Tom might not have been able to grasp this from MS 1, and he might wonder how to derive the associative strength of a component within the role-schema. Assuming that the relations between elements in an associative-network perspective on role-schemata are just contingent on the frequency of co-occurrence (Smith, 1998), one simple preliminary answer would be that Tom needs to derive or estimate the percentage of all instances of the

interaction in which the conditions for co-variation of the component and the position are met and take this estimation as a prediction for associative strength.

Moreover, it is questionable to which degree the propositions people associate with the roles are valid representations of the world, that is, whether they truly mirror the ones Tom – our RTT-in-the-present-state scientist – would deduct in a top-down manner (cf. Strack & Deutsch, 2004). After all, the present state of RTT assumes that role-schemata can also be learned bottom-up (see statement 3, Appendix A). Thus, Tom would have to include all the biases that make learning processes so delicate into his reasoning. That is, if Tom would want to achieve perfectly precise predictions about role-schemata, he would not only need to be able to estimate the frequency to which the conditions are met under which certain components co-vary with the pursuit of a given goal, he would also need to have quite extensive psychological knowledge about how people generally learn propositional contingencies, including, for example, whether this process differs when acting in a position as compared to observing others acting in it. For instance, although the core associations with the "teacher"-role and the "student"-role are independent of whether an actor already acted in both roles or not, many other associations arguably differ quite substantially between actual teachers and the rest of the world. RTT, thus, puts quite some trust into its user's competence by not explicitly incorporating any further guidance in these matters, which threatens its precision. At the very least, this threat should be reduced by specifying the "repeated encounters" of statement 3 (Appendix A), for example, by inserting a statement if and how often people have to be proponents of a position, proponents of the complementary position, or observers of a position to abstract the role-schemata assumed by RTT.

Second issue: *random* components. The need for more guidance on how to uncover specific role-components becomes even more obvious for the *random* associations, that is, the ones which cannot be derived deductively from the prototypical social structure at all. For example, while some behavioral expectations follow quite naturally from goals and many

self-related associations follow right after an intermediate deductive step (see example above), others are of rather symbolic nature and some symbols tied to roles are completely irrelevant to role-goals and have no important interactional function whatsoever. For example, it is not a necessary precondition for a bride to wear white to get married. In fact, it is not related to the role-goal or any other universal feature of the interaction at all. Such associations thus cannot be deducted, but they still might be important to precisely predict role-taking effects in certain contexts. For example, although this was not the original reason for the white color of wedding dresses[18], it is often associated with purity and virginity, and thus, one could predict corresponding role-taking effects on the situational self-concept and on behavior: brides feel more "pure" and "maidenlike" after role-taking because they associate "white" and "purity" with the role[19]. Given a cultural context in which such associations are explicitly shared, RTT would predict that brides do not even have to wear the white dress for this role-taking effect, because the associations are generally tied to the role schema and a self-categorization as "bride" should be sufficient to activate them[20]. However, as long as RTT only provides rules on how to obtain schemata deductively, it cannot provide hypotheses about such *random* and sometimes context-sensitive features of roles.

Third issue: emergent features of schemata. Thirdly, RTT might have to provide rules how the integration of several components of different elementary types in one role-schema affects the schema as a whole to achieve precise predictions. In the most extreme case, the integration of a variety of different components in one and the same schema might not only lead to meta-representations (e.g., in some roles, actors might learn that the self-related associations tied to a role are to be ignored, for example, as a "selfless" monk), but lead to emergent phenomena so that the schema as a whole cannot be understood just by the

[18] White wedding dresses were presumably introduced to the world by rich aristocrats because the apparent impracticability of this color demonstrated wealth and luxury.

[19] Of course, the "purity" association could also be alternatively explained through religious norms tied to the "bride", e.g., abstinence before marriage.

[20] However, for some, this self-categorization might be very hard or not possible at all without using the symbols tied to the role.

sum of its parts. For example, coming back to the *Games people play* (Berne, 1964/2011), Berne argues that people tie very complex clusters of associations to roles in families. An advocate of transactional analysis would even argue that these so-called *ego states* (Berne, 2011) can hardly be dissociated into discrete components, and their effects, once activated, could not be fully understood by analyzing the psychological processes tied to each of these parts in an isolated fashion. Even with a finite set of rules through which all discrete parts of this complex cluster could be derived deductively, RTT could fail to grasp far enough in such cases, as long as it does not deliver rules about how these discrete parts interact in forming emergent features of schemata. Given this was true, and in the face of the apparently enormous complexity that would be added to this issue through the diversity of roles, the critical reviewer would conclude at this point that RTT is in far over its head by solely relying on a deductive approach, that is, that the present state of the theory is not even close to provide an acceptable degree of precision when it comes to comprehensive predictions of the mental schemata tied to all the roles it claims to be applicable on. Arguing as an advocate, although it might be theoretically plausible to assume emergent features of schemata including very complex clusters of components, this level of complexity is presumably only reached by a very narrow set of long-term roles. In fact, the roles in families might even be the only example where schemata develop such emergent features. Thus, as RTT's main explanatory focus in the present state lies on phenomena tied to short-termed roles, excluding the family roles from RTT's universe of applicability might be a bearable (temporary) compromise to deal with this third issue.

Fourth issue: transfer of components between roles. However, the critical reviewer would comment that transactional analysis even assumes that *ego-states* – although learned in family interactions – can also be tied to more short-termed roles as well (e.g., the *lover* role, see MS 1). Although this effect – and the transactional perspective as a whole – has not undergone rigid empirical testing, this would raise an additional theoretical point of concern.

This kind of transfer of components from one role-schema to another would imply an even stronger call for more comprehensive rules on component-derivation, because such transferred components would not be inherent to the interaction, but "imported byproducts", and thus, could not be obtained by merely analyzing one single interaction deductively. In this case, RTT would need to provide extensive guidance on which degree of similarity between interactions is sufficient to cause such transfer-effects, which features of an interaction are decisive for similarity, and whether this transfer is mutual or if there is some kind of hierarchy causing one-way transfers from more important roles to others (e.g., transactional analysis would argue that there is a hierarchy in transfers: *ego-states* learned in family roles become part of other roles more often than vice versa, Berne, 2011). To illustrate, Tom could potentially ask himself whether the typical asymmetries in age, status, knowledge and power between teachers and students are sufficiently similar with the ones between parents and children to trigger such transfer-effects. If so, Tom's hypotheses about the role-schemata would contain components that are not necessarily *random* but would only follow from the deductive path after adding quite an extensive set of further assumptions.

Summary. To sum up, the issues related to deriving precise hypotheses about specific components of role-schemata appear to be so substantial that the deductive pathway incorporated in the present state of RTT is not sufficient to cope with them. Fortunately, there is at least a partial remedy to these issues: some of the *random* associations are explicitly shared, at least in a given cultural context, just as the bride's white dress. Although one cannot derive them deductively, they are known to everyone who has at least some rudimentary knowledge about the cultural context in which RTT is to be applied. Thus, they can be obtained inductively. However, to use inductive reasoning or inductive empirical approaches to obtain the role-schemata also bears dangers and limitations, for example, overgeneralization and circular reasoning. Still, it might be the only way to cope with the issue of *random* associations as well as with the challenges posed in the next chapter.

5.5 Improving Precision: Nature of Role-Schemata

First issue: complexity of component types. First, although discrete components of certain types might be meaningful parts of role schemata, role-taking activates *all* components tied to a position, and thus, knowledge about one discrete component of one type only allows a prediction of a tiny mosaic stone in a most likely very complex interplay. Although it is plausible to assume direct effects for many discrete components of a given type (e.g., of a specific expectation on corresponding behavior), there is also evidence for indirect and interaction effects between the types in the model linking role-taking and observable outcomes (Burke & Reitzes, 1981; Gecas, 1982; Jussim, 1989). For example, the degree to which actors fulfill behavioral expectations activated through role-taking might be moderated by downstream effects of simultaneously activated self-relevant associations (e.g., teachers only fulfill the expectation to behave respectful towards their students as long as they experience themselves as persons to be respected), or the effects of self-relevant associations on the situational self-concept might be moderated by the degree to which behavioral expectations are fulfilled (e.g., teachers only experience themselves as persons to be respected as long as they can answer students' questions, cf. MS 3, also see *role-performance*, Burke & Reitzes, 1981). Thus, for precise predictions of measurable outcomes, it is crucial to know as many of the elementary types constituting the slots of a role-schema as possible and to have a model about their interrelations with regard to specific outcomes.

This reasoning also already implies RTT's integrative impetus: if roles are the categories under which social cognitions are packaged in interpersonal interactions and role-taking activates the whole package, then all the *types* making up their own slot in a role-schema should each have their own spot in conceptual models explaining specific outcomes (e.g., a given behavior). Although not every type might be equally relevant for every outcome-specific model, their potentially simultaneous activation and use would at least require the consideration of potential direct effects and interaction effects for all types. Thus,

RTT implies that models focusing on effects of one type only, for example, just on expectations, or just on self-relevant associations, fail to reach far enough in interpersonal interactions (also see MS 3).

Second issue: *random* types of components. Making things even more difficult, the set of types constituting role-schemata arguably differs between roles. For example, the "teacher" in contrast to the "judge" is not associated with a formalized symbol of any kind. Tom could not have known that, using RTT in the state of MS 1. Tom could not have known, because such symbols are often *random components* of a schema, that is, they cannot be derived deductively (see second issue on deriving specific components). Therefore, equally important as developing more refined rules on how to obtain knowledge about specific components, it appears necessary to provide a rationale which elementary *types* are incorporated in a role-schema under which conditions. Although the deductive path gives some answers on this, they do not seem to be sufficient.

Summary. In sum, I argue that more knowledge about the nature of role-schemata is necessary, especially about the types of components that constitute the slots of role-schemata. When predicting role-taking effects on specific outcomes, direct and indirect effects of all types of components should be considered.

5.6 A Potential Remedy: Inductive Derivation of Role-Schemata

The only remedy especially for *random* components would be accounting for them by adding an inductive approach: deriving the components on which role-taking hypotheses are based not only top-down from universal features of the interaction, but also bottom-up through some sort of measurement or inductive reasoning. Inductive reasoning would be the simplest of all methods to capture explicitly shared *random* components of role-schemata, just as the white wedding dress. Another blunt method to identify *random* associations would be to gather large samples of associative lists people produce in response to the role in question and to integrate them in an associative model. Of course, both methods are also severely

limited (e.g., by social desirability). *Random* components of which actors are less aware are even harder to identify. For example, there might be symbols tied to roles that are rather subliminal (e.g., victim supporters might engage into more eye-contact with victims than perpetrator supporters), and actors might store corresponding implicit associations in role-schemata. Although theoretically, Tom could obtain knowledge about the implicit relation of a known semantic component to the role-schema using implicit measures (e.g., the IRAP, Implicit Relational Assessment Procedure, Barnes-Holmes et al., 2006), Tom could not identify a previously unknown component using such a method. Furthermore, the debate about the reliability of measures and validity of results of the social priming literature illustrates the meticulous character surrounding this topic (e.g., Doyen, Klein, Pichon, & Cleeremans, 2012; Harris, Coburn, Rohrer, & Pashler, 2013).

In sum, inductive approaches offer easy solutions to obtain explicitly shared *random* role-components but *random* components tied implicitly to roles would constitute a challenge which presumably can only be overcome with yet to be invented methods, for example, in the realm of neuro-imaging. Thus, for now, RTT should only incorporate an additional statement guiding researchers to also include explicitly shared *random* components when specifying hypotheses about role-schemata and corresponding role-taking effects.

5.7 Comparative Assessment of Empirical Content

Now, that we are thoroughly aware of RTT's present state and potential improvements, it is time to assess its empirical content over and above competing approaches. In MS 1, we already discussed how RTT relates to Identity Theory (e.g., Stryker & Burke, 2000) and the Social Identity Approach (i.e., Social Identity Theory, Tajfel & Turner, 1986; and Self-Categorization Theory, Turner, Hogg, Oakes, Reicher, & Wetherell, 1987). We concluded that RTT does not compete with but rather complements them, because RTT differs in assumed processes and explanatory focus.

Still, the critical reviewer might wonder whether RTT's role-concept is truly necessary

to explain the phenomena it claims to be applicable on. When debating such questions, *parsimony* or *simplicity* is very commonly referred to as an important criterion: everything else being equal, the simpler theory should be preferred (e.g., Brandstätter, Gigerenzer, & Hertwig, 2006; Glöckner et al., 2018). Of course, "everything else" is typically not equal: theoretical approaches usually differ in universality and precision (Glöckner et al., 2018). Still, the critical reviewer might suggest to drop the role-concept from the theory and to focus only on the behavior acted out in roles. To substantiate the point, he or she might argue that such a reductionist approach would be more parsimonious than RTT while being at least as precise, and point to evidence that the characterization of a particular behavior may also already activate categories corresponding with personality traits (e.g., Uleman, 1999). Furthermore, he or she might note that the empirical results reported in MS 2 and 3 could potentially be explained just as well simply by the behavior that was carried out in the roles (in sensu).

That being said, advocates have arguments on their side as well. The true symbolic interactionist would take a radical stance to counter this critique (i.e., argue that social behavior only carries meaning in the first place because it is attached to it through an ongoing process of interactive negotiation, Mead, 1934). However, an empiristic advocate would point to all the studies demonstrating that context matters in categorization processes, for example, because the context implies alternative causes of the same behavior (e.g., Trope, 1986). Crucially, the actors' positions matter: third-parties react differently to perpetrator's behavior than victims (Skarlicki & Kulik, 2004), victims have different needs than perpetrators after transgressions and thus respond differently to reconciliatory interventions (Shnabel & Nadler, 2008), and when arbitrators resist against a conflict resolution, they are motivated by different processes than bystanders (MS 3).

Thus, in conclusion, the question is not whether it is helpful for successful interactions to gather knowledge about positions as an actor (it most definitely is, see above) or whether

people are influenced by this knowledge in *some* roles (they obviously are, e.g., Berne, 1951; Gray & Wegner, 2009; Salmivalli, 1999). The question is how strong the influence of role-schemata over and above behavioral schemata in the universe of all interactions really is, in other words, whether the resolution of specific role-schemata in comparison to more general behavioral schemata really is more fine-grained in all interactions. One corresponding line of studies should investigate whether role-taking of very similar roles (i.e., roles where the behavior is comparable to the utmost degree) still causes meaningful differences in actors' experience and behavior that can only be predicted using RTT, for example, by analyzing further universal features of the interaction (e.g., the role's goal, interdependence of goals, etc.). In sum, the size of RTT's increment in explanatory power over and above all other categorization processes (e.g., on the basis of group-membership or plain behavior) has to be demonstrated empirically and thus can be debated at the present state, but if the critical reviewer would want to reject RTT's theoretical foundation entirely, this could only be done by simultaneously dismissing large parts of social-cognitive theory and corresponding empirical evidence.

5.8 Assessment of the Degree of Corroboration

Now, if RTT constitutes a scientific advantage as of its empirical content, it should be evaluated to which degree RTT is corroborated through empirical results. First of all, as outlined above, RTT shares some basic principles with other role-theoretical and social-cognitive approaches that are corroborated empirically already (e.g., IT, Stets, Burke, & Savage, 2018; participant role approach, Salmivalli, 2010). That is, the discussion of the degree of corroboration has to be focused on the novel contributions of RTT only. In that regard, the empirical studies reported in MS 2 and 3 are examples for RTT's explanatory power. By suggesting that third-party role-taking in conflicts is intertwined with the situational and chronic moral self-concept and that an increase in moral situational self-concept due to role-taking can motivate arbitrators to resist against a conflict resolution, the

two manuscripts empirically demonstrate that RTT has an increment over IT (Stryker & Burke, 2000) by extending the role-concept to positions in short-termed interactions. Especially the results reported in MS 3 are a point in case for RTT. The hypothesis that arbitrators would be motivated to keep the conflict going followed from the general RTT principle that roles are embedded within interactions and thus lose their meaning once the interaction is over. This principle is not shared by the sociological IT view on roles. IT (Stets & Burke, 2000, p. 225) conceptualizes roles as "… the relatively stable, morphological components of social structure that are termed roles. Thus, like social identity theory, identity theory deals principally with the components of a structured society:" Assuming a relatively stable and more "macro"-character of roles, IT could not have been used to derive the hypotheses that led to manuscript 3.

That being said, both manuscripts only entail one set of roles in one interaction (i.e., third-party roles in conflicts), raising questions of generalizability, and the theoretical reasoning was mainly based on one of the five basic features of roles assumed by RTT: identity-relevance. In conclusion, empirical evidence corroborating the novel theoretical contributions of RTT is yet spare. For example, future empirical work still has to proof that RTT's novel perspective also has incremental explanatory value in explaining the relationships between the other four role-features and role-taking effects: expectations, individual entity, complementarity, and mutual reinforcement (see Outlook Section of MS 1 for potential studies).

5.9 Formalization

One practice recently suggested regarding theory development is the formalization of theories (Glöckner & Betsch, 2011; Glöckner et al., 2018). To formalize a theory, all if-then statements constituting the theory are listed in an explicit and formal way (Glöckner & Betsch, 2011). If the theory entails corresponding statements, formalization would also mean to translate the theory into mathematical formulae. The idea is to prevent insufficient

specification of antecedents and consequences of a theory, thereby reducing the danger of a theory being unfalsifiable. Furthermore, in comparison to the prose form, this practice is supposed to facilitate (1) the assessment of the *empirical content* of a theory, (2) the identification of theories applicable to the same phenomenon, (3) the confident use and updating of a theory by researchers other than the original authors, and (4) the easy storage of a theory in open data-bases to facilitate public access to the most recent version together with empirical results indicative of the theory's degree of corroboration (Glöckner et al., 2018). Although RTT does not offer statements that can be translated into mathematical formulae as of yet, I want to contribute to these ends, and thus, one can find the basic formalization of RTT at the end of this final discussion (Appendix A). The formalization is not included in the main text because semantically, it does not contain anything that is not already included in the prose form of the theory (MS 1).

5.10 Conclusion

RTT's present state is a much needed first step opening many doors for future role-theoretical research. The number and quality of the above issues in precision should not intimidate future efforts: according to RTT, roles are – next to groups – the most important social categories affecting actors in their everyday life. Thus, naturally, the questions and issues surrounding roles and the role-taking process are almost as complex as social life itself: How can we describe the structures of interpersonal social life? How do these structures affect actors on the most basal cognitive level? And can we derive explanations for feelings, thoughts, behavior, and the development of interactions from insights on these matters? RTT supplies convincing answers by clearly defining the category "social roles", by specifying roles through their positional anchors in interpersonal interactions, and by delivering guidance on how to derive role-taking hypotheses from these anchors. In face of the challenges brought about by the ambitious degree of universality, RTT's degree of precision is acceptable. Still, as discussed above, RTT needs to improve in specifying the role-taking process, the nature of

role-schemata, and the rules for component-derivation. The inclusion of an inductive approach to derive hypotheses about role-schemata seems to be inevitable in face of the issues surrounding *random* components..

In many ways, RTT's potential is comparable to the one of the SIA approaches: just as the SIA provides rationales about how a super-ordinate and universal social category (the "group") affects individual proponents, RTT provides a rationale about how the super-ordinate category "role" affects proponents. Although not every individual part of SIT or SCT has been novel just by itself upon publication (Tajfel & Turner, 1986), the integrative combination undoubtedly was novel and substantial enough to spark sustained empirical and theoretical work investigating social groups. Considering the apparent lack of a social-cognitive role-theory that is accepted within social psychology (Hogg, Terry, & White, 1995), the same – hopefully – applies to RTT. Still, large parts of RTT have yet to be corroborated empirically, and more impactful empirical results need to demonstrate its explanatory power in the future. In this regard, the most promising lines of future research surround the role-features of complementarity and mutual reinforcement (e.g., "pull-effects", see MS1), and the cognitive basis of role-taking (see above).

To conclude, the work carried out in this PhD project dealt with many initial difficulties sparked by the discovery of a surprising theoretical and empirical gap in social psychology. Rediscovering social roles as a social category and applying modern social-cognitive theorizing to explain associated processes arguably was a long due step for a comprehensive understanding of social structure and how it affects involved actors. Going forward, the outlined future directions will illuminate how important this step was to understand actors' feelings, thoughts, and behaviors in interpersonal interactions.

References

Baddeley, A. D. (1986). *Working memory.* New York: Oxford University Press.

Barnes-Holmes, D., Barnes-Holmes, Y., Power, P., Hayden, E., Milne, R., & Stewart, I. (2006). Do you really know what you believe? Developing the Implicit Relational Assessment Procedure (IRAP) as a direct measure of implicit beliefs. *The Irish Psychologist, 32*, 169-177. doi: 10.1037/h0100577

Berne, E. (2011). *Games people play: The basic handbook of transactional analysis.* Tantor eBooks.

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Making choices without trade-offs. *Psychological Review, 113*, 409–432. doi: 10.1037/0033-295X.113.2.409

Brewer, W. F., & Treyens, J. C. (1981). Role of schemata in memory for places. *Cognitive Psychology*, *13*, 207-230. doi: 10.1016/0010-0285(81)90008-6

Burke, P. J., & Reitzes, D. C. (1981). The link between identity and role performance. *Social Psychology Quarterly*, *44*, 83-92. doi: 10.2307/3033704

Doyen, S., Klein, O., Pichon, C.-L., & Cleeremans, A. (2012). Behavioral priming: It's all in the mind, but whose mind? *PloS one*, *7*(1), e29081. Retrieved from https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0029081

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: an integrative review of implicit and explicit attitude change. *Psychological Bulletin*, *132*, 692-731. doi: 10.1037/0033-2909.132.5.692

Gawronski, B., & Bodenhausen, G. V. (2011). The associative-propositional evaluation model: Theory, evidence, and open questions. *Advances in Experimental Social Psychology*, 44, 59-127. doi: 10.1016/B978-0-12-385522-0.00002-0

Gecas, V. (1982). The self-concept. *Annual Review of Sociology*, *8*, 1-33. doi: 10.1146/annurev.so.08.080182.000245

Glöckner, A., & Betsch, T. (2011). The empirical content of theories in judgment and decision making: Shortcomings and remedies. *Judgment & Decision Making, 6,* 711-721.

Glöckner, A., Fiedler, S., & Renkewitz, F. (2018). Belastbare und effiziente Wissenschaft. *Psychologische Rundschau, 69,* 22-36. doi: 10.1026/0033-3042/a000384

Gollwitzer, M., Braun, J., Funk, F., & Süssenbach, P. (2016). People as intuitive retaliators: Spontaneous and deliberate reactions to observed retaliation. *Social Psychological and Personality Science, 7,* 521-529. doi: 10.1177/1948550616644300

Gray, K., & Wegner, D. M. (2009). Moral typecasting: divergent perceptions of moral agents and moral patients. *Journal of Personality and Social Psychology, 96,* 505-520. doi: 10.1037/a0013748

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological Review, 108,* 814-834. doi: 10.1037/0033-295X.108.4.814

Harris, C. R., Coburn, N., Rohrer, D., & Pashler, H. (2013). Two failures to replicate high-performance-goal priming effects. *PloS one, 8*(8), e72467. Retrieved from https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0072467

Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A.W. Kruglanski (Eds.), *Social Psychology: Handbook of Basic Principles* (pp. 133–168). New York: Guilford.

Jussim, L. (1989). Teacher expectations: Self-fulfilling prophecies, perceptual biases, and accuracy. *Journal of Personality and Social Psychology, 57,* 469-480. doi: 10.1037/0022-3514.57.3.469

Mead, G. H. (1934). *Mind, self and society*. Chicago: University of Chicago Press.

Popper, K. R. (1934/2005). *Logik der Forschung* (11th ed.). Tübingen: Mohr Siebeck.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind?. *Behavioral and Brain Sciences, 1,* 515-526. doi: 10.1017/S0140525X00076512

Salmivalli, C. (1999). Participant role approach to school bullying: Implications for interventions. *Journal of Adolescence*, *22*, 453-459. doi: 10.1006/jado.1999.0239

Schaafsma, S. M., Pfaff, D. W., Spunt, R. P., & Adolphs, R. (2015). Deconstructing and reconstructing theory of mind. *Trends in Cognitive Sciences*, *19*, 65-72. doi: 10.1016/j.tics.2014.11.007

Shnabel, N., & Nadler, A. (2008). A needs-based model of reconciliation: satisfying the differential emotional needs of victim and perpetrator as a key to promoting reconciliation. *Journal of Personality and Social Psychology*, *94*, 116-132. doi: 10.1037/0022-3514.94.1.116

Skarlicki, D. P., & Kulik, C. T. (2004). Third-party reactions to employee (mis) treatment: A justice perspective. *Research in organizational behavior*, *26*, 183-229. doi: 10.1016/S0191-3085(04)26005-1

Smith, E. R. (1998). Mental representation and memory. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of Social Psychology*, Vol. 1 (4th ed., pp. 391–445). New York: Oxford University Press.

Stets, J. E., & Burke, P. J. (2000). Identity theory and social identity theory. *Social Psychology Quarterly*, 224-237. doi: 10.2307/2695870

Stets, J. E., Burke, P. J., & Savage, S. V. (2018). Exchange, Identity Verification, and Social Bonds. *Social Psychology Quarterly*, *81*, 207-227. doi: 10.1177/0190272518785866

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*, 220-247. doi: 10.1207/s15327957pspr0803_1

Stryker, S., & Burke, P. J. (2000). The past, present, and future of an Identity Theory. *Social Psychology Quarterly*, *63*, 284-297. doi: 10.2307/2695840

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *The Social Psychology of Intergroup Relations* (pp. 7-

24). Chicago: Nelson-Hall.

Trope, Y. (1986). Identification and inferential processes in dispositional attribution. *Psychological Review, 93,* 239-257. doi: 10.1037/0033-295X.93.3.239

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the Social Group: A Self-Categorization Theory*. Cambridge, MA, US: Basil Blackwell.

Uleman, J. S. (1999). Spontaneous versus intentional inferences in impression formation. In S. Chaiken & Y. Trope (Eds.), *Dual-process Theories in Social Psychology* (pp. 141-160). New York: Guilford Press.

5.i Appendix A

(1)

IF: for all social interactions

THEN: the interaction consists of a finite set of positions involved actors can have.

Positions are social categories distinguished by their goal in the interaction.

(2)

IF: for all social interactions AND all persons

THEN: persons can only take one position at a time.

Taking a position is the self-categorization as a proponent of a position.

(3)

IF: for all persons AND repeated encounters with a social interaction

THEN: people abstract mental schemata of the positions in the interaction.

(4)

IF: all associations constituting a position-specific mental schema AND for all persons AND all social interactions

THEN: associations are learned either explicitly or implicitly.

(5)

IF: any kind of participation in an interpersonal interaction AND for all persons who tie a mental schema to their respective position in the interaction

THEN: the accessibility of the mental schema tied to the position the person took in the interaction increases.

Accessibility is the likelihood that a concept will be used when responding to subsequent

stimuli (Higgins, 1996).

The mental schema consists of all associations tied to the position by the respective actor.

[GOALS]

The goals of primary positions are either positive interdependent, negative interdependent, or neutral.

The goals of third party positions in interactions are located in a two-dimensional space.

One goal dimension of third party positions is the form of interdependence of the third party goal with the goals of the primary positions (positive, negative, or neutral).

One goal dimension of third party positions is the maintenance of the interaction (end, sustain, or neutral).

6 Deutsche Zusammenfassung

Inhalt. Wie verhalten sich Personen in interpersonalen Interaktionen? Und warum verhalten sie sich so wie sie es tun? In dieser Dissertation präsentiere ich in drei Manuskripten und einer finalen Diskussion einen neuen theoretischen Ansatz, der zur Beantwortung dieser Fragen einen sozial-kognitiven Prozess der Rollenübernahme spezifiziert. Der theoretische Ansatz, die Role-Taking Theory (RTT), wird in Manuskript 1 geschildert und erklärt wie Akteure durch Eigenschaften ihrer jeweiligen Position in Interaktionen beeinflusst werden. Den Kern der Theorie bilden drei Sätze: Eine soziale Rolle ist das mentale Schema, das ein Akteur an eine Position in einer Interaktion knüpft. Die Rollenübernahme („role-taking“) ist die Aktivierung des Rollenschemas bei Übernahme der Position. Das Rollenschema formt durch die Aktivierung Fühlen, Denken und Verhalten. In ihrer empirischen Anwendung wurde die RTT in vier empirischen Studien dazu benutzt, die Reaktionen von Drittparteien in Konflikten zu erklären (*Ns* = 659, 55, 170, 107). Die Ergebnisse dieser Studien stützen empirische Hypothesen, die sich aus zwei Annahmen der RTT ableiten. Erstens, die Rollenübernahme in Konflikten ist auf komplexe Art und Weise mit dem moralischen Selbstkonzept verwickelt. Zweitens, Akteure in prosozialen Rollen (wie Schlichter oder Unterstützer des Opfers) erfahren durch die Rollenübernahme eine Steigerung im situativen moralischen Selbstkonzept und können dadurch motiviert sein, ihre Rolle zu behalten, selbst wenn das bedeutet, den Konflikt zu verlängern. Die empirischen Studien werden in Form von zwei weiteren zur Veröffentlichung eingereichten Manuskripten in Kapitel 3 und 4 berichtet. Zudem diskutiere ich in einem finalen fünften Kapitel selbstkritisch den aktuellen Stand der RTT, weise auf mögliche Verbesserungen in theoretischer Präzision hin und beleuchte sowohl die Herausforderungen als auch die Chancen in zukünftiger Forschung. Zusammenfassend liefert diese Dissertation mit der RTT einen umfassenden theoretischen Rahmen zur Erforschung interpersonaler Interaktionen sowie erste empirische Beiträge, die die Annahmen der RTT stützen und ihre Nützlichkeit verdeutlichen.

Manuskript 1 – Role-Taking Theory. Beim ersten Manuskript (Kapitel 2) handelt es sich um einen theoretischen Artikel in welchem wir die Prämissen der RTT erläutern, diskutieren worin die RTT über bisherige Ansätze hinausgeht, die RTT beispielhaft auf interpersonale Konflikte anwenden und die zentralen zukünftigen Forschungsfragen besprechen. Dazu wird nach einem kurzen Abriss der theoretischen Wurzeln des Rollenbegriffs vor allem verdeutlicht, dass bisherige Ansätze zur Rollenübernahme hauptsächlich auf langfristig übernommene Rollen (wie „Vater“, „Arbeitnehmer“, oder „Linker“) und ihren Zusammenhang zum chronischen Selbstkonzept fokussiert waren. Problematisch aus sozialpsychologischer Perspektive ist dabei, dass (1) eine große Anzahl von Rollen ausgespart wurde (Rollen in kurzfristigen Interaktionen), (2) die eher soziologische Definition der Kategorie „soziale Rolle“ keine klare Trennung zu sozialen Gruppen zulässt und vor allem, dass (3) der sozial-kognitive Prozess der Rollenübernahme bisher nicht zufriedenstellend beschrieben wurde.

Die RTT stützt sich in ihrer sozial-kognitiven Ausrichtung des Weiteren vor allem auf fünf grundlegende Aspekte, die auch in bisherigen Ansätzen immer wieder mit Rollen in Bezug gesetzt wurden: Identitätsrelevanz, Erwartungen, Einzelkategorie, Komplementarität und wechselseitige Verstärkung. Erstens sind Rollen identitätsrelevant, da die Übernahme einer Rolle (egal ob kurz- oder langfristig) das situationale Selbstkonzept und – bei bedeutsamen Rollen – das chronische Selbstkonzept beeinflusst. Zweitens sind Rollen mit Erwartungen verknüpft: Das Verletzen von Rollen-Erwartungen kann zu internalen Sanktionen (wie Schuld oder Scham) aber auch zu externalen (wie Bestrafung) führen. Drittens beschreiben Rollen als soziale Kategorien anders als Gruppen ein einzelnes Individuum: Der Prozess der Rollenübernahme setzt daher keine bereits vorhandenen Vertreter dieser Kategorie voraus. Zudem unterscheidet sich die Rollenübernahme durch den individuellen Charakter von Rollen von Prozessen die die Identifikation mit Gruppen beschreiben: Obwohl auch die Rollenübernahme die Identifikation mit einer sozialen

Kategorie beinhaltet, findet bei der Rollenübernahme kein Wechsel von einer „Ich“-Ebene der Selbstkategorisierung zu einer „Wir“-Ebene statt. Die RTT wird zudem in ihrem Fokus auf interpersonales Verhalten klar von Theorien abgegrenzt, die sich mit der Erklärung von Intergruppenverhalten beschäftigen. Viertens kann eine Rolle nur sinnvoll eingenommen werden, wenn auch die komplementäre Rolle besetzt ist: Ein Opfer braucht einen Täter, ein Schüler einen Lehrer, ein Führer braucht Folgende und vice versa. Fünftens führen die vielfältigen Wechselbeziehungen zwischen komplementären Rollen, wie zum Beispiel in der Zielerreichung (aber auch in der Verteilung anderer Ressourcen wie Macht, Anerkennung oder Identifikation mit den Rollen) häufig dazu, dass sich Rollenvertreter wechselseitig in ihren Rollen verstärken. Diese beiden letzten Eigenschaften (Komplementarität und wechselseitige Verstärkung) können erklären unter welchen Umständen Akteure in ihre Rollen hineingezogen und in ihnen „gebunden“ werden können. Dabei impliziert die RTT, dass in den multivariaten Verteilungen der oben genannten Ressourcen Equilibria entstehen können, die „fest gefahrene“ Interaktionen erzeugen. Solche Interaktionen sind durch sich immer wiederholende Handlungsmuster der Akteure gekennzeichnet (wie der immer gleiche Streit des alten Ehepaars).

Neben den fünf Eigenschaften von Rollen besteht der Kern der RTT aus sechs Prämissen, die in Manuskript 1 kritisch entwickelt werden. Erstens: Alle sozialen Interaktionen haben eine prototypische soziale Struktur, die aus einem finiten und disjunkten Set von Rollen besteht. Zweitens: Von wiederholten Erfahrungen mit einer sozialen Interaktion abstrahieren Individuen eine mentale Schablone der prototypischen sozialen Struktur der Interaktion. Drittens: Diese mentale Schablone enthält ein generalisiertes mentales Schema zu jeder Position in der sozialen Struktur. Viertens: Das mentale Schema einer spezifischen Position besteht aus allen Assoziationen die an diese Position geknüpft sind, wie zum Beispiel Ziele, Selbst-Wahrnehmungen, Erwartungen, allgemeine Bewertungen, Verhaltensoptionen und Symbole. Fünftens: Wenn ein Akteur eine Position

einnimmt, wird das rollen-spezifische mentale Schema aktiviert. Dieser Prozess wird als Rollenübernahme definiert. Sechstens: Die Rollenübernahme formt das Fühlen, Denken und Verhalten von Akteuren.

Im weiteren Verlauf des Artikels wird erläutert, wie die RTT benutzt werden kann um Hypothesen über den Inhalt von spezifischen Rollenschemata aufzustellen. Dabei steht vor allem die Analyse des Ziels einer Rolle im Vordergrund. Laut RTT ist jede Rolle von einem universalen Rollenziel gekennzeichnet ist, von dem sich weitere *fixe* Schemakomponenten logisch ableiten lassen (zum Beispiel selbst-relevante Assoziationen oder Verhaltenserwartungen). Neben den fixen Komponenten, die sich von solchen universalen Eigenschaften von Rollen ableiten lassen, besitzen Rollenschemata laut der RTT aber auch *zufällige* Komponenten, so zum Beispiel manche Symbole (zum Beispiel die weiße Farbe von Brautkleidern). In Bezug auf die Rollenübernahme wird unabhängig vom fixen oder zufälligen Charakter von Komponenten angenommen, dass die erhöhte kognitive Verfügbarkeit rollenspezifischer semantischer Konzepte sowohl über reflektierte als auch über impulsive Routen kognitiver Verarbeitung Effekte auf Fühlen, Denken und Verhalten hat.

Im dritten Abschnitt wird die RTT auf interpersonale Norm-Konflikte angewendet. Von der RTT lässt sich ableiten, dass involvierte Akteure genau eine von sieben Rollen übernehmen können: Opfer, Täter, Schlichter, Eskalateur, Unterstützer Opfer, Unterstützer Täter, oder Bystander. Zudem erläutern wir in diesem Abschnitt, welche Erklärungsansätze die RTT für Phänomene in Konflikten bietet, wie zum Beispiel für verminderte und destruktive Kommunikation, reduzierte Perspektivübernahme, Motive zur Aufrechterhaltung von Konflikten, motivierte Verzerrungen in der Konfliktwahrnehmung und die Stabilisierung von Konflikten. Dabei liegt ein besonderer Fokus auf der Rolle von Drittparteien in Konflikten.

Im letzten Abschnitt argumentieren wir, dass weitere Forschung zu den Grundlagen der RTT vor allem beleuchten sollte, durch welche Mechanismen Akteure in Rollen hineingezogen werden (siehe oben zu Komplementarität und wechselseitiger Verstärkung), welche Wechselwirkungen es zwischen Prozessen der Rollenübernahme und der Gruppenidentifikation gibt, ob kurzfristig übernommene Rollen auch langfristige Effekte haben können und welche Bedeutung individuelle Stile in der Ausgestaltung von Rollen für die Konsequenzen der Rollenübernahme haben. Zusammenfassend argumentiert die RTT, dass ein bedeutender Anteil sozialer Kognition an mentalen Repräsentationen fixer und universaler Positionen in Interaktionen gebunden ist und daher die Analyse von Rollenschemata und ihrer Effekte für die Erklärung interpersonaler Phänomene sehr nützlich ist.

Manuskript 2 – Drittparteien in Konflikten. In MS 2 und MS 3 werden die Resultate erster empirischer Tests von Hypothesen berichtet, die sich im Kontext interpersonaler Konflikte aus Annahmen der RTT ableiten lassen. Die Fragestellungen in diesem Anwendungsbereich: Wie reagieren Dritte, wenn sie Zeuge von Ungerechtigkeit werden? Wie lassen sich ihre Reaktionen erklären? Und unter der Annahme, dass Normverletzungen und damit wahrgenommene Ungerechtigkeit häufig Ausgangspunkt von Konflikten sind und Drittparteien häufig eine wichtige Rolle für den Verlauf und den Ausgang von Konflikten spielen, was kann man aus Erkenntnissen über die Reaktionen von Drittparteien auf Ungerechtigkeit über die Stabilisierung oder Eskalation von Konflikten lernen?

Bisherige Forschung hatte sich vor allem auf die ursprünglich beteiligten Akteure fokussiert (Partei A und Partei B, oder „Opfer“ und „Täter“ nach einer Normverletzung). Forschung zu Drittparteien untersuchte hauptsächlich einzelne Verhaltensweisen Dritter wie Bestrafung, Vergeltung, Vergebung, oder Kompensation. Ein integrativer Rahmen zur Erklärung der Reaktionen von Drittparteien und inter-individueller Unterschiede in diesen

Reaktionen fehlte. Der ursprünglich erste Entwurf der RTT war dazu gedacht, diese Lücke zu schließen.

Zunächst folgt aus der RTT die Klassifikation der Reaktionen Dritter durch die fünf Rollen Unterstützer Opfer, Unterstützer Täter, Schlichter, Eskalateur und Bystander. Zudem lässt sich aus der Identitätsrelevanz von Rollen und der moralischen Bedeutsamkeit von Rollen in Konflikten die Hypothese ableiten, dass die Übernahme einer prosozialen Rolle (Schlichter oder Unterstützer Opfer) im Vergleich zu einer neutralen Rolle (Bystander) durch Persönlichkeitseigenschaften vorhergesagt werden kann, die die Zentralität der moralischen Komponente für das Selbst erfassen. Die Rollenübernahme einer prosozialen Rolle im Vergleich zu einer passiven Rolle sollte des Weiteren zu einem höheren moralischen situativen Selbstkonzept führen. Dieser Effekt sollte – falls die Rollenübernahme tatsächlich durch Selbstregulation und weniger durch Impression Management motiviert ist – bei Internalisierern (Personen, die moralische Eigenschaften wie fair, ehrlich und gerecht in ihr Selbstkonzept internalisiert haben) aber nicht bei Symbolisierern (Personen, die gerne nach außen kommunizieren, dass sie moralische Eigenschaften besitzen) verstärkt auftreten.

In Studie 1 (N = 659) wurden die Persönlichkeitseigenschaften und die Rollenwahl der Teilnehmenden in einem Konfliktszenario online gemessen. Nach der Rollenübernahme schätzten die Teilnehmenden ihr situatives Selbstkonzept während des Konflikts in fünf Facetten ein (Moral, Selbstwert, Identitätsausdruck, Bedeutung und Macht), wobei die moralische Facette die zentrale abhängige Variable war. Studie 2 (N = 55) replizierte Studie 1 mit einem experimentellem Design: Die Rollen wurden randomisiert zugewiesen. Die Ergebnisse beider Studien unterstützen die Hypothesen und sind ein Indiz für einen potentiellen motivationalen Mechanismus: Personen, die moralische Eigenschaften stark internalisiert haben, wählen deshalb prosoziale Rollen, weil sie sich durch die Rollenübernahme noch stärker als moralische Personen erleben als Personen, die moralische Eigenschaften nicht so stark internalisiert haben.

Manuskript 3 – Wenn Schlichter den Konflikt verlängern. Die Hypothesen dieser Studie folgten aus den Ergebnissen aus Manuskript 2 und einer Annahme der RTT: Wenn Akteure in prosozialen Rollen sich als moralische Personen erleben und dieser belohnende Effekt der Rollenübernahme verschwindet, sobald der Konflikt gelöst wird (weil die Rollen ihre Bedeutung verlieren sobald die Interaktion endet), sollte sich für Schlichter daraus eine Motivation ergeben, den Konflikt aufrecht zu erhalten. Das heißt, die prosoziale Rollenübernahme als Schlichter im Vergleich zur neutralen Bystander-Rolle sollte das moralische situative Selbstkonzept steigern und diese Steigerung sollte positiv mit dem Widerstand gegen eine Konfliktlösung korrelieren. Dieser indirekte Effekt wurde in zwei Studien überprüft. Studie 1 (N = 170) ist eine quasi-experimentelle Online-Studie (freie Rollenwahl) und Studie 2 (N = 107) eine experimentelle Laborstudie mit randomisiert zugewiesenen Rollen. Der indirekte Effekt zeigt sich in beiden Studien: Schlichter fühlen sich während des Konflikts moralischer als Bystander und je moralischer sie sich fühlen, umso mehr Widerstand zeigen sie gegen eine Konfliktlösung. Dieser Effekt erklärt, warum es manchmal schwierig sein kann, Konflikte zu lösen, selbst oder gerade wenn Drittparteien in prosozialen Rollen beteiligt sind.

Finale Diskussion. In der finalen Diskussion evaluiere ich selbstkritisch und transparent den aktuellen Stand der Role-Taking Theory und bereite so die nächsten Schritte der Weiterentwicklung vor. Ich zeige vor allem drei Bereiche auf, in denen die RTT in ihrer theoretischen Präzision verbessert werden kann. Erstens sollte zukünftige empirische und theoretische Arbeit die basalen kognitiven Prinzipien beleuchten, die der Rollenübernahme zugrunde liegen. Die Konzeptualisierung von Rollen als mentale Schemata und der Rollenübernahme als die Aktivierung derselben ist das Fundament der RTT. Daher sollte zukünftige Forschung dieses Fundament streng untersuchen, um davon ausgehend weiterreichende Hypothesen zu Effekten der Rollenübernahme abzuleiten. Ursprünglich wurde die Relevanz mentaler Schemata in der Informationsverarbeitung und

Gedächtnisenkodierung in experimentellen Paradigmen demonstriert, in denen ein Schema (z.B. „Büro“) aktiviert wurde und danach gemessen wurde, ob schema-konsistente Stimuli (z.B. Bürostuhl, Kalender, Computer) besser erinnert wurden als inkonsistente. Ähnliche Paradigmen, die die basalen Eigenschaften von Rollen als soziale Schemata testen, würden uns dabei helfen, die kognitiven Grundlagen der Rollenübernahme besser zu verstehen.

Der zweite Bereich dreht sich um die Eigenschaften von Rollenschemata: Welche Typen von Komponenten (Selbstwahrnehmungen, Erwartungen, etc.) können in Rollenschemata gespeichert sein, welche nicht und unterscheidet sich der Prozess der Rollenübernahme je nach Komponententyp in Art (z.B. reflektiert versus impulsiv) oder in Relevanz für verschiedene kognitive und behaviorale Konsequenzen? Hier argumentiere ich, dass die RTT viel Potenzial besitzt, weil sie eine umfassende Analyse universaler Eigenschaften von Interaktionen anregt, die genau für diese Fragen informativ sein sollte. Zum Ausschöpfen dieses Potenzials muss diese Analyse in Zukunft aber stärker theoretisch ausgearbeitet werden.

Drittens muss die RTT ihren Nutzern mehr Unterstützung dabei bieten, präzise Hypothesen über die spezifischen Komponenten von Rollenschemata zu bilden. Dazu muss neben mehr Hilfestellungen zum deduktiven Weg (dem Schließen auf spezifische Rollenkomponenten ausgehend von universalen Eigenschaften wie den Rollenzielen) wohl auch ein induktiver Weg spezifiziert werden, um auch *zufällige* Komponenten wie zum Beispiel Symbole mit zu erfassen (zum Beispiel das weiße Kleid der Braut).

Zusammenfassend argumentiere ich, dass die RTT zwar bisher nur partiell empirisch getestet wurde, aber im Vergleich zu bestehenden Theorien zusätzlichen empirischen Gehalt (Popper, 1934/2005) besitzt und damit ein vielversprechender theoretischer Ansatz ist. Ob sich die RTT in Zukunft behaupten und nachhaltigen Einfluss in der Sozialpsychologie haben wird, hängt davon ab, ob sich ihre Versprechungen in methodologisch rigiden empirischen Studien einlösen lassen. Die zukünftigen theoretischen und empirischen

Entwicklungsbereiche, die im Ausblick in Manuskript 1 (z.B. „Pull-“Effekte von sozialen Rollen) und in der finalen Diskussion besprochen werden, weisen diesbezüglich den Weg.

I Danksagung

Dank gebührt zuvorderst Prof. Mario Gollwitzer, für die Idee zur RTT, das Vertrauen, die Geduld, die positive Einstellung und das Verständnis.

Dank gebührt auch Prof. Dieter Frey, für die Begutachtung.

Danke auch an Prof. Thomas Eckert, für den Beisitz.

Danke an Peter, Philipp und Gunnar, für die Freundschaft und die Kollegialität.

Danke liebe Freunde, für die Ablenkung.

Danke liebe Anna, für das Zuhören, das Verstehen und das Lachen.

Danke lieber Andreas und liebe Franzi, die ihr immer da seid.

Für Anne.

II Veröffentlichungen

Schwabe, J., & Gollwitzer, M. (2019a). *Role-Taking Theory*. Manuscript under review.*

Schwabe, J., & Gollwitzer, M. (2019b). *Explaining third-party behavior in norm conflicts: A role-taking approach*. Manuscript under review.*

Schwabe, J., & Gollwitzer, M. (2019c). *When arbitrators prolong conflicts: Beneficial role-taking effects motivate resistance against conflict resolution*. Manuscript under review.*

* = relevant für die Dissertation.