Association for Information Systems
AIS Electronic Library (AISeL)

ICIS 2022 Proceedings | Human Computer /Robot Interaction

Dec 12th, 12:00 AM

Good Robot, Bad Robot: Customer Responses to Norm-Compliant and Norm-Violating Service Robots

Ruth Maria Stock-Homburg
Technische Universität Darmstadt, rsh@bwl.tu-darmstadt.de

Lea Heitlinger
Technical University Darmstadt, lea.heitlinger@bwl.tu-darmstadt.de

Follow this and additional works at: https://aisel.aisnet.org/icis2022

Recommended Citation

Stock-Homburg, Ruth Maria and Heitlinger, Lea, "Good Robot, Bad Robot: Customer Responses to Norm-Compliant and Norm-Violating Service Robots" (2022). *ICIS 2022 Proceedings*. 8.
https://aisel.aisnet.org/icis2022/hci_robot/hci_robot/8

This material is brought to you by the International Conference on Information Systems (ICIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ICIS 2022 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

Good Robot, Bad Robot: Customer Responses to Norm-Compliant and Norm-Violating Service Robots

Completed Research Paper

Ruth Maria Stock-Homburg
Technical University of Darmstadt
Hochschulstraße 1, 64289 Darmstadt
rsh@bwl.tu-darmstadt.de

Lea Heitlinger
Technical University of Darmstadt
Hochschulstraße 1, 64289 Darmstadt
lea.heitlinger@bwl.tu-darmstadt.de

Abstract

Service robots that interact with customers have penetrated various industries. With a basis in social identity theory, this study examines how customers respond to frontline service robots (FSRs) by investigating norm-compliant versus norm-violating behaviors compared with similar behaviors by human frontline employees (FLEs). In experimental studies, a black sheep effect occurs, such that customers downgrade norm-violating FLE behaviors more than similar behaviors by FSRs. They also upgrade norm-compliant behaviors by human FLEs more than those of FSRs. In service failures, this effect manifests as greater anger and frustration toward the FLE. We establish the underlying mechanism driving the black sheep effect: customers assign FSRs to an outgroup but categorize FLEs to their social ingroup, across different service encounters and independent of interaction frequency.

Keywords: Service robots; customer-robot interaction; social identity theory; black sheep effect; service encounter

Introduction

Customers prefer to be served by frontline employees (FLEs) who are similar to them (Gill et al. 2017), reflecting an ingroup bias, such that people generally favor members of their own social group and rate them as more capable or friendly than members of another group (i.e., outgroup) (Steain et al. 2019). Although they thus approve strongly of norm compliance by ingroup members, they also tend to be more critical of norm violations by those members, because the violations threaten their own social identity (Marques 1990). Prior studies refer collectively to these responses to norm-compliant versus norm-violating behaviors by ingroup members as the black sheep effect (Eidelman and Biernat 2003; Eidelman et al. 2006; Marques et al. 2001; Marques et al. 1988), defined as people's tendency to judge compliant ingroup members more positively but deviant ingroup members more negatively, contrasted with the responses to outgroup members exhibiting the same behaviors. These evaluations can play a role in multiple areas, such as responses to health-related norms (Packer et al. 2021) or group criticism (McCrea et al. 2021). In a service setting, the black sheep effect implies that service encounters involving human FLEs might differ from those conducted by frontline service robots (FSRs), or "system-based autonomous and adaptable interfaces that interact, communicate and deliver service to an organization's customers" (Wirtz et al. 2018, p. 909). FSRs appear in various industry settings, including tourism (Ivanov et al. 2017), shopping malls (Kanda et al. 2009, 2010), hotels (Gockley et al. 2006; Kirby et al. 2010) or restaurants (Mende et al. 2019). As customer–robot service encounters increase in number and spread (Mende et al. 2019), robots are performing functions and roles previously left to human FLEs (Lee et al. 2010). Recent investments also imply that ongoing advancements in artificial intelligence, sensing, actuator, and power technology will soon enable service robots to participate effectively in "dynamic affective exchanges with human partners" (Damiano et al. 2015, p. 1).

New technologies can have an important impact on service firms and as such also on a company's growth (Barrett et al. 2015). Until a few years ago, very little service literature investigated how customers respond

to robots that deliver services (Murphy et al. 2016). Although this number is growing (e.g., Mende et al. 2019), there are still many research gaps due to the growing prevalence of FSRs but also with regard to the likelihood that customers respond to FSRs in unique ways. For example, they might adopt social categorizations and act in accordance with the black sheep effect, though various contingencies may have effects as well (e.g., service failures). We thus aim to test whether the black sheep effect predicts customers' distinct responses to norm-compliant versus norm-violating behaviors by FSRs and FLEs. For this research, we define norm-compliant behaviors as those "geared toward satisfying customers' needs adequately" (Stock and Hoyer 2005, p. 536), whereas norm-violating behaviors include incivility toward customers and inadequate interpersonal treatment during the service encounter (van Jaarsveld et al. 2010). Norms operate in service encounters (Aggarwal 2004), such that service representatives are expected to be sensitive to customers' needs, and if they fail to meet these norms, their behavior appears unacceptable and elicits negative reactions (Wan et al. 2011). During a service encounter, a black sheep effect would imply that customers regard norm-compliant behaviors by human FLEs, as members of the ingroup, more positively than similar behaviors by FSRs, but the FLEs' norm-violating behaviors would be far less acceptable. In turn, we ask, *Do customers upgrade norm-compliant behaviors and downgrade norm-violating behaviors by a human FLE compared with similar behaviors by a FSR?* We further seek to identify some contingencies and moderating effects by addressing a particular type of norm violation by asking, *Does the black sheep effect for FSRs versus FLEs arise after service failures?* We regard incivility as part of norm-violating behaviors expressed by employees (Walker et al. 2017), while the service failure pertains to an actual incident (i.e., booking mistake) with the employee behaving in a neutral way.

For the black sheep effect to arise, customers would have to categorize human FLEs and FSRs as social ingroups and outgroups, respectively. Therefore, we turn to social identity theory (Tajfel and Turner 1986) to establish whether a FSR is perceived as a member of the outgroup while a FLE represents social ingroup (Fraune et al. 2017) as an underlying mechanism of the black sheep effect. That is, *Do customers categorize a FSR and a human FLE into different social groups?* The answers might shift over time, because as customers gain experience interacting with FSRs, their responses might develop (Gockley et al. 2006; Kirby et al. 2010). Thus, *Do customers' social categorizations of FSRs change over time?*

By investigating these research questions, we add to current research in information systems and service research, thus foster interdisciplinarity. As FSRs increase steadily in number (Mende et al. 2019), we need a better understanding of how customers respond (van Doorn et al. 2017; van Pinxteren et al. 2019) and accept these novel technologies (Agarwal and Karahanna 2000). This may be especially important in the light of global pandemics making it necessary to protect FLEs from potentially harmful and risky situations in direct customer contact (Esterwood and Robert 2021). By investigating the black sheep effect, we establish how customers respond to FSRs' norm-compliant versus norm-violating behaviors during a failure-free service encounter, then consider their responses in more detail in a service failure scenario, as a particular form of norm violation. By investigating these nuanced behaviors, we are able to offer guidance for researchers regarding the detailed placement and design of social robots in service contexts. We also consider whether customer responses reflect their social categorization of different service agents as an underlying mechanism of the black sheep effect. In turn, we advance social identity theory in several important ways. That is, we identify the different social categories that customers assign to human FLEs versus FSRs in service settings. Furthermore, we examine how frequency of customer-robot interaction affects a robot's social categorization by customers, relying on the mere exposure phenomenon.

This research provides important insights for managerial practice pointing towards the practical utility of our results. Practitioners can anticipate when and why customers will respond differently to a FSR than a human FLE, by acknowledging the effects of social categorization as potential antecedents of important customer-related outcomes and thus aid service firms. Accordingly, they should take care before placing FSRs in jobs, that were originally reserved for humans, because the robots represent a different social group to customers and might thus evoke less positive customer reactions in regular customer-employee interactions according to the black sheep effect. This notion might be relevant even after several interactions. Yet the evidence we offer for the black sheep effect indicates that customers also might "forgive" a FSR more readily than a FLE for errors, e.g., incivil employee behaviors or even service failures. Given that negative customer reactions often prevail in these interactions (Gelbrich 2010), deploying FSRs might protect FLEs from potential detrimental effects caused by negative customer behaviors. Thus, practitioners can make better informed decisions about when and where to insert FSRs and shape their displayed behaviors accordingly. Finally, this study offers novel evidence about customer responses

following service interactions; the longitudinal effects that we uncover give predictive insights about how customers will respond to regular service interactions with FSRs, such that our study goes beyond one-time effects and offers first insights into long-term acceptance of technologies in service contexts.

Literature Review

An increasing amount of literature has investigated various technologies in the service sector in recent years. According to Heitlinger et al. (2022), technological agents can be characterized along the axes social and physical presence. Robots for instance have both high social and physical presence, while avatars or chatbots have only high social, but low physical presence. We deem this distinction relevant, as research has demonstrated that a lack of embodiment might weaken the FSR's automated social presence (Fasola and Matarić 2011; Lee et al. 2006), whereas a robot's physical embodiment increases its ability to accommodate variations in the environment (Stubbs et al. 2007), making robots unique compared to other technologies (You and Robert 2018). Thus, we focus on research involving actual humanoid robots and exclude literature examining chatbots/avatars (e.g., Crolic et al. 2022; Lv et al. 2021) or machines (such as self-service kiosks; Chen et al. 2021). Research into human–robot interactions largely consists of two streams: robotics research and marketing research. Recent conceptual articles and reviews in both robotics (Belanche et al. 2020a) and services marketing (Čaić et al. 2019; Choi et al. 2021; Ivanov et al. 2017; Kuo et al. 2017; Lu et al. 2020; Merkle 2019; Pitardi et al. 2022; van Doorn et al. 2017; Wirtz et al. 2018; Xiao and Kumar 2019) research indicate that "robots and AI are expected to play an even bigger role in the economy due to what was named 'Fourth Industrial Revolution'" (Ivanov and Webster 2017, p. 1), so "scholars should examine customer acceptance of robots" (Murphy et al. 2016, p. 105). Mende et al. (2019) show in their literature review, that in robotics research, many studies take place within the hospitality sector, such as hotels, retailing, or the health sector. These studies consistently emphasize the importance of human–robot interactions and anticipate the automated social presence of humanoid FSRs, which appear promising for enhancing service encounters, including recovery after a service failure (Choi et al. 2021; Chung and Cakmak 2018b). The research objectives of these studies mainly are to gain insights to optimize robots' technical capabilities by observing their errors. About half of the studies involve actual human–robot interactions, but the rest rely on images or videos of robots. Most studies focus on a single interaction with a robot, reflecting an implicit assumption that humans' emotional reactions do not change over time or through additional interactions (White and Arzi 2005). In a six-month field experiment, Gockley et al. (2006) show that interactions last longer with emotional robots than with neutral robots. A field study in a shopping mall over 26 days (Kanda et al. 2009) reveals that participants who evaluate the robot positively express more interest in the interaction. With a nine-week study, Kirby et al. (2010) determine that the degree of empathy humans offer in response to emotional expressions by robots does not differ from their degree of empathy after verbal expressions.

In marketing research, we find more consideration of FSRs from the perspectives of service representatives or customers. Experimental studies test customer responses to FSRs, mainly in the hospitality sector (Belanche et al. 2020b; Fan et al. 2020; Ho et al. 2020; Jörling et al. 2019; Leo et al. 2020; Mende et al. 2019; Merkle 2019; Yam et al. 2021) but also in laboratory experiments (van Pinxteren et al. 2019; Yam et al. 2021). Jörling et al. (2019) focus on the outcomes of perceived autonomy of the robot and reveal that autonomy decreases perceived behavioral control over the service robot and responsibility for positive but not for negative outcomes. The potential to interrupt the FSRs' autonomy increases perceived behavioral control and responsibility for positive outcomes. Mende et al. (2019) reveal that FSRs elicit greater customer discomfort than humans, which results in enhanced compensatory consumption. Mechanized robots lead to less compensatory consumption than anthropomorphized robots. In a laboratory experiment, van Pinxteren et al. (2019) show that a robot's eye gaze matters for human–robot interactions, by affecting trust, usage intentions, and enjoyment, through the mediation of perceived anthropomorphism. Several studies also address service failures during the customer–robot interaction. Belanche et al. (2020b) reveal that customers make stronger attributions of responsibility for service performance toward humans than toward robots, especially when a service failure occurs; these findings are also mimicked by Merkle (2019) with higher customer satisfaction for FSRs than for FLEs after a service failure. Their findings further indicate that FSRs can provide a valuable contribution, though without considering a black sheep effect explicitly. We regard this point as relevant though, because a FSR's artificial behaviors might determine the effectiveness of customer–robot interactions, as suggested by the caution that marketing "studies have not yet varied the extent of [a robot's] social expression" (Groom et al. 2011, p. 195). Choi et al. (2021) go a step

further and examine FSRs in the context of service recoveries. Further studies examine differences in service recovery following service failure incidents by investigating effects of role theory and perceived sincerity (Ho et al. 2020; Hu et al. 2021). Fan et al. (2020) and Yam et al. (2021) focus on anthropomorphism by including different forms of robots in service failures. An interesting study on attribution of responsibility following a service failure by Leo et al. (2020) found that people attributed less responsibility toward a robot for a service failure compared to a human because the robot was perceived to have less controllability; in case of the robot, more responsibility was attributed toward the service firm. We deem it necessary to explicitly investigate the black sheep effect in service scenarios as it focuses and combines both positive and negative service encounters (including service failures) and thus has additional explanatory value to investigate direct differences between FSRs and FLEs. To contribute to both research streams, we seek insights into customer responses to customer–robot versus customer–human interactions, while also noting potential moderating effects of norm compliance versus violation.

Conceptual Background and Study Framework

Social Identity Theory and the Black Sheep Effect

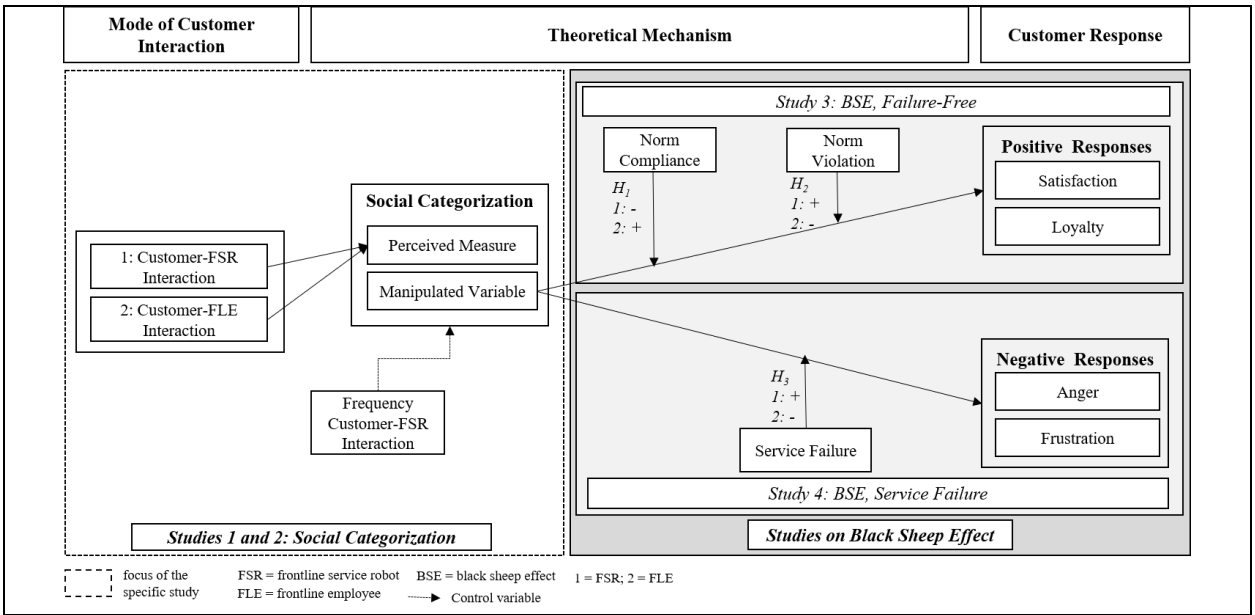

Figure 1. Study Framework

To posit that customers' responses differ when they interact with FSRs versus FLEs, we rely on social identity theory (Tajfel and Turner 1986), which suggests that people define themselves on the basis of both personal and social aspects (Tajfel 1981). The social aspect (i.e., social identification) refers to the perception of belonging to a human group (Ashforth and Mael 1989). Tajfel (1981, p. 255) defines social identity as "that part of an individual's self-concept, which derives from … knowledge of … membership of a social group (or groups) together with the value and emotional significance attached to that membership." Accordingly, "people typically categorize others as ingroup members (i.e., 'like me') or as outgroup members (i.e., 'not like me') when they first see them" (Fraune et al. 2017, p. 1432). Social categorization results in greater favoritism toward ingroup members compared with outgroup members (Ojala and Nesdale 2004). Several studies that apply social identity theory to employee–customer relationships (e.g., Homburg et al. 2009) or human–robot interactions (e.g., Deligianis et al. 2017; Fraune et al. 2017) affirm that people express more ingroup favoritism toward a social robot than toward a computer (Deligianis et al. 2017), though less than their ingroup favoritism toward other humans (Fraune et al. 2017). As differences in social categorization between humans and robots have already been established on a general level (e.g., Fraune et al. 2017), we further investigate this topic in our specific service context. We further involve longitudinal considerations in the form of a mere exposure effect (Zajonc 1968), such that "repeated, unreinforced exposure is sufficient to enhance attitude toward a stimulus" (Bornstein and D'Agostino 1992, p. 545). The mere exposure effect arises in previous human–robot interaction studies (de Graaf et al. 2016; Paetzel et al.

2020). As an extension of social identity theory, the black sheep effect (Marques et al. 1988) describes a psychological process by which "individuals upgrade likable ingroup members and derogate unlikable ingroup members compared with their outgroup counterparts" (Pinto et al. 2010, p. 107). It arises during social interactions (Eidelman and Biernat 2003; Eidelman et al. 2006) and during service employee–customer interactions (Gill et al. 2017).

In the first study, we want to make a deep dive on potential underlying mechanisms: differences in social categorization between FLEs and FSRs in our specific service context (Figure 1, Study 1). Thus, we examine whether customers evaluate a FSR as a member of an outgroup compared with a human FLE as part of the ingroup. Furthermore, we examine the social categorization of robots in a longitudinal online study, to determine possible changes over time. To examine customer responses to FSRs, as members of an outgroup, during the service encounter, we conduct several experiments involving different FSR behaviors (Figure 1). We examine the black sheep effect (Studies 3 and 4) in customer-FSR and customer-FLE interactions. With an experimental laboratory vignette study, we test for it according to whether customers upgrade a human FLE's norm-compliant behavior and derogate the human FLE's norm-violating behaviors, compared with those of a FSR (Study 3), measured in terms of positive customer responses. Thus, the independent variables are the norm-compliant and norm-violating behaviors of the service agent (FSR or FLE). As norm-compliant behavior, we examine customer orientation; norm-violating behaviors are captured by customer-related incivility. The dependent variables are positive customer responses (satisfaction and loyalty) to the service agents' behaviors. Customer satisfaction refers to the extent to which a customer evaluates the interaction positively (Stock and Bednarek 2014). Customer loyalty "is a customer's intention to repeatedly purchase products from the same company" (Stock et al. 2017, p. 225). For Study 4, the independent variable is specifically a service failure, as a norm violation, such that we determine if the black sheep effect occurs during a service failure. If so, the customer would respond more negatively to a human FLE, as a member of the ingroup, than to a FSR, as a member of the social outgroup. Due to the service failure situation, we evaluate negative customer responses: Customer anger is "a powerful negative emotion" (Strizhakova et al. 2012, p. 2), and customer frustration is strongly connected to blame attribution for goal-incongruent incidents (Gelbrich 2010).

Hypotheses Development

Black Sheep Effect in a Failure-Free Service Setting

According to the black sheep effect, a person tends to upgrade ingroup members who behave in a norm-compliant manner, such that those behaviors count more than similar behaviors by outgroup members. Contrarily, they derogate the norm-violating behaviors of ingroup members more than similar behaviors by outgroup members (Pinto et al. 2010). This ingroup bias helps the person to protect a positive image of the ingroup (De Cremer and Vanbeselaere 1999) and safeguard her or his own social identity (Tajfel 1981; Tajfel and Turner 1986). To determine whether customers upgrade norm-compliant behavior by a human FLE, relative to similar norm-compliant behaviors by a FSR, we examine norm-compliant versus norm-violating behaviors in a normal, failure-free encounter (Erasmus et al. 2002). We test whether these behaviors enhance customer satisfaction and loyalty, as critical affective and behavioral customer responses to service encounters (Stock et al. 2017) that can be evoked by pleasant behaviors by service representatives (Homburg et al. 2011; Stock and Hoyer 2005). (User) satisfaction for instance is often studied in information systems contexts (Agarwal and Karahanna 2000; Cenfetelli et al. 2008). Differences in evaluations between robots and humans have been observed in various contexts, e.g., in privacy concerns (Stock-Homburg & Hannig 2020) or decision-making in teams (Gombolay et al. 2015). Thus,

H_1: During the service encounter, customers respond less favorably to pleasant (norm-compliant) behaviors by a FSR, with relatively lower (a) customer satisfaction and (b) customer loyalty, compared with similar behaviors by a human FLE.

Turning to norm-violating predictions (Pinto et al. 2010), we investigate whether customers derogate norm-violating human FLEs as ingroup members more than FSRs as outgroup counterparts when they express incivility (norm violation) toward customers (van Jaarsveld et al. 2010). Consistent with research on customer–FLE interactions, we suggest that customers respond with less satisfaction and loyalty to incivility (Groth and Grandey 2012). If a FSR behaves in a norm-violating manner, it conflicts with customers' predefined scripts, established through previously experienced service encounters (Erasmus et

al. 2002). According to the black sheep effect, customers might respond more negatively to norm-violating human FLEs rather than FSRs, because this norm violation by an ingroup member represents a symbolic threat to human values or distinctiveness (Riek et al. 2008). We hypothesize:

H_2: During the service encounter, customers respond less unfavorably to a FSR's (norm-violating) incivility, with relatively higher (a) customer satisfaction and (b) customer loyalty, compared with similar behaviors by a human FLE.

Black Sheep Effect After a Service Failure

Another norm violation occurs with a service failure, such that the service performance "falls below a customer's expectations" (Hess et al. 2003, p. 129), representing an important construct in information systems research (Tan et al. 2016). An exchange relationship between a customer and a service provider is governed by norms, which lead customers to expect the service they receive will be worth the money (Wan et al. 2011). In the case of a service failure, the service norm is violated, because customers do not receive the expected service (Wan et al. 2011), and customer anger is a particularly striking, negative reaction as an outcome (Laros and Steenkamp 2005). According to the black sheep effect, customers tend to discount a service failure by a FLE more than one by a FSR, so we predict that customer anger and frustration, as negative emotions toward a FSR (Strizhakova et al. 2012) after a service failure, might be stronger towards FLEs than those that emerge in response to failures by FSRs. Thus,

H_3: During the service failure, customers served by a human FLE score higher on (a) anger and (b) frustration than those served by the FSR due to the black sheep effect.

Overview of Empirical Studies

Study	Design & Stimuli	Dependent Variable	Findings	Model (Adjusted Mean and SE)
1	2 (SC) × 4 (service scenarios) BS, online experiment	Social categorization	Social outgroup categorization for FSR and ingroup categorization for FLE are stable across various service scenarios.	FLE = 4.57 (.68) vs. FSR = 3.80 (.68), $F(1, 1140) = 65.23$, $p < .05$
2	7 interactions t_1-t_7 (FSR, service scenarios) WS, online experiment	Social categorization over time	Social outgroup categorization for FSR is stable across multiple interactions.	$F(4.67, 322.43) = .57$, $p = .71$
3	2 (failure-free service SC) × 2 (SB) BS, laboratory experiment	Customer satisfaction and loyalty	The BSE is apparent in the norm-violating condition; the FSR is favored over the FLE. Significant differences in loyalty in the norm-compliant condition.	$F_{SC}(2, 131) = 23.62$, $p < .05$, $F_{SB}(2, 131) = 52.58$, $p < .05$, $F_{IE}(2, 131) = 24.13$, $p < .05$
4	2 (service failure SC) BS, online experiment	Customer anger and frustration	The BSE is apparent, customers are angrier and more frustrated with the human FLE than with the FSR after a service failure.	$t_{Anger}(198) = 1.98$, $p < .05$, $t_{Frustration}(198) = 2.05$, $p < .05$

Table 1. Overview of Studies

Notes: All responses represent actual behavior in the service context. BS = between-subject design; WS = within-subject design; BSE = black sheep effect; SC = service condition (FLE or FSR); SB = service behavior (norm compliance or norm violation); IE = interaction effect.

Table 1 contains an overview of the studies and their objectives. In Studies 1 and 2, we investigate differences in social categorization between FLEs and FSRs as an underlying mechanism of the black sheep effect with two vignette-based online experiments. We test our hypotheses with two experimental studies: Studies 3 and 4 investigate the black sheep effect in human-robot interactions - a vignette-based laboratory experiment and a vignette-based online experiment. In Study 1, to investigate whether customers socially

categorize service robots as members of a social outgroup across multiple service settings, we compare customer responses to a neutral FSR versus a neutral FLE. In this context, neutral means that neither service provider expresses any particular emotions, whether pleasant or unpleasant (Matsumoto and Ekman 2004). Additionally, in Study 2 we examine frequency of interaction and assesses potential changes in the social categorization of FSRs over time. In Study 3, we manipulate the FSR's norm-compliant or norm-violating behaviors. To determine how participants socially categorize a FSR that behaves in either manner, we compare people's responses to the FSR/FLE, measured as positive responses of customer satisfaction and loyalty. In Study 4, we check whether the black sheep effect persists after a service failure on negative customer responses.

The experimental studies all feature experimental vignettes, which ask participants to imagine they are engaged in the detailed scenarios (Collett and Childs 2011). Specifically, participants had to imagine themselves as customers in a hypothetical scenario, involving a service encounter in which they needed help from a service representative. Furthermore, with between-subject research designs, we sought to avoid learning effects, such that each participant was randomly assigned to one condition. We rely on self-reported ratings and previously validated measures when possible. Amazon Mechanical Turk (MTurk) workers participated in the online studies (Studies 1, 2, and 4). Interactions were simulated with videos that mimic real-life service encounters in various settings. We limited the participant pool to MTurk workers with an acceptance rate greater than 95% and residents of the United States. All participants received financial remuneration of $3.50.

Study 1: Differences in Social Categorization Across Service Scenarios (Online Experiment)

Method

To test social categorization of human FLEs across various service settings, we conducted an online experiment with a 2 × 4 between-subjects design, such that the participants were randomly assigned to a FSR or FLE and one of four service scenarios: tour guide, sales assistant at a supermarket, real estate agent, or server at a restaurant. We also checked an online hotel setting, to correspond with our other studies[1]. After being introduced to a scenario, participants watched a series of videos and indicated their social categorization of the service agent with items adapted from Homburg et al. (2009) on a 7-point Likert-type scale; Cronbach's α was greater than .90 for all service scenarios. The overall sample consisted of 1,245 MTurk workers, of whom 637 randomly were assigned to the FSR condition and 608 to the FLE condition. We tested whether they watched the videos attentively and answered carefully with two checks, for which participants had to select a concrete value on a scale (i.e., please pick value 4 on the scale), and excluded any participants who failed the test items. After eliminating 9.6% of participants in the FSR condition and 3.9% in the FLE condition who failed the check, we retained samples of 576 and 572, respectively, for an overall sample size of 1,148 participants. All of them were residents of the United States, with an average age of 39.15 years (SD = 11.61 years). In addition, 56.4% were men. A post-hoc G*Power analysis resulted in a necessary total sample size of 437 participants to detect the effect; with our sample we fulfill this prerequisite. For the calculation, we used the results of the dependent variable. Due to our online setting, we included a large buffer, resulting in the much bigger sample size. For the manipulation check, participants had to rate a series of statements related to the role of the service representative after watching the videos for their randomly assigned service scenario on a 7-point Likert-type scale. The responses confirmed that they recognized the roles of the tour guide in scenario 1 ($t(276) = 24.31$, $p < .05$), sales assistant in scenario 2 ($t(286) = 23.96$, $p < .05$), real estate agent in scenario 3 ($t(305) = 26.96$, $p < .05$), and server at a restaurant in scenario 4 ($t(277) = 29.16$, $p < .05$). In the added scenario, they also recognized the role of the hotel concierge ($t(222) = 19.56$, $p < .05$).

[1] Although the procedure is the same, the data for the hotel setting were collected separately: N = 223, 56.5% men, M_{age} = 40.41 years (SD_{age} = 11.78 years), and differences in social categorization between FLE and FSR: $t(185.12) = -5.24$, $p < .05$.

Findings

We used analysis of variance (ANOVA) to test for a difference in social categorization between the FSR versus FLE experimental conditions, as well as across the service scenarios ($F(7, 1140) = 10.73$, $p < .05$, partial $\eta^2 = .062$). The results indicate that social categorization ($F(1, 1140) = 65.23$, $p < .05$, partial $\eta^2 = .054$) differed significantly for the service representative conditions, with a higher average in the FLE condition ($M = 4.57$, $SD = .68$) than in the FSR condition ($M = 3.80$, $SD = .68$), in line with a social categorization of the human FLE as part of the social ingroup. A post hoc test confirmed that this significant difference in social categorization appeared in all four service scenarios (details in Table 2). Furthermore, neither the main effect of the service scenario ($F(3, 1140) = 1.54$, $p = .204$) nor the interaction between the experimental condition and the service scenario ($F(3, 1140) = 1.28$, $p = .278$) were significant. Thus, participants categorized the human FLE to an ingroup and the FSR to an outgroup in all service scenarios, indicating that the proposed effect of social categorization remains stable across service encounters.

Service Condition					
		FSR	FLE		
		Mean (SD)	Mean (SD)	Mean Difference (SE)	*p*-Value
Service Scenario	1. Tour Guide	3.94 (1.66)	4.79 (1.24)	-.85* (0.19)	.00
	2. Sales Assistant	3.87 (1.88)	4.34 (1.45)	-.48* (0.19)	.01
	3. Real Estate Agent	3.67 (1.86)	4.66 (1.47)	-.99* (0.19)	.00
	4. Service Employee	3.72 (1.80)	4.51 (1.39)	-.78* (0.19)	.00
	5. Hotel Concierge	3.89 (1.83)	5.01 (1.27)	-1.12* (0.21)	.00

Table 2. Study 1, Post Hoc Test of Social Categorization, by Service Scenario and Condition

Notes: Social categorization, 7-point Likert-type scale from 1 ("totally disagree") to 7 ("totally agree"). n(1) = (135|142); n(2) = (160|127); n(3) = (150|156); n(4) = (131|147); n(5) = (106|117). FSR = frontl. service robot; FLE = frontline employee. *$p \leq .05$.

Study 2: Longitudinal Study of Social Categorization of Service Robots (Online Experiment)

Method

To build on the findings about the social categorization of FSRs and FLEs, we sought to test whether it might change over time, due to increased exposure to FSRs through habituation or mere exposure effects. Thus, we conducted a within-subject study over four weeks. The participants completed an online questionnaire twice weekly, for a total of seven waves. We created seven service scenarios, with videos of the robot Pepper: museum guide, store assistant, concierge, real estate agent, information robot at the airport, tour guide and server at a restaurant. The scenario assignments were random in order, but once participants had completed the seventh wave, they all had gone through each service scenario once. In each wave, participants watched a series of videos reflecting the assigned scenario, then provided their social categorization of the FSR, measured with five items on 7-point Likert-type scales adapted from Evers et al. (2008) and Henry et al. (1999). Cronbach's α was again greater than .90 throughout the study. The MTurk workers received increasing incentives over the seven waves, to reduce drop-out rates. We started with an initial sample of 277 participants. After the seven waves, 71 participants remained, each of whom saw each service scenario once. The participants all were residents of the United States, with an average age of 36.55 years (SD = 9.44 years); 56.3% of the sample were men. A post-hoc G*Power analysis resulted in a necessary sample size of 25 participants to detect a medium effect; with our sample we fulfill this prerequisite. Due to our online and longitudinal setting, we included a large buffer, resulting in the much bigger sample size.After watching the videos of each service scenario, participants again rated a series of statements on a 7-point Likert-type scale that checked whether they had watched the videos carefully and understood the content (manipulation check). For all seven waves, they clearly recognized the roles of the FSR as a museum

guide ($t(70) = 5.88$, $p < .05$), store assistant ($t(70) = 6.74$, $p < .05$), concierge ($t(70) = 5.99$, $p < .05$), real estate agent ($t(70) = 5.43$, $p < .05$), information robot at the airport ($t(70) = 6.91$, $p < .05$), tour guide ($t(69) = 6.46$, $p < .05$), or server at a restaurant ($t(70) = 6.85$, $p < .05$).

Findings

To test whether the social categorization of the FSR changed over the course of the seven waves, we conducted a repeated-measures analysis of variance (rmANOVA). In brief, the social categorization did not change as a result of repeated exposures ($F(4.67, 322.43) = .57$, $p = .71$). That is, we do not find support for our prediction that it would increase over time as a result of a mere exposure effect, such that the FSR might come to be considered as part of the social ingroup over the course of the experiment. Instead, our pre-test results indicate that outgroup social categorizations of FSRs are stable, across different service scenarios and over time.

Study 3: Black Sheep Effect, Failure-Free Service (Controlled Real-Life Experiment)

Method

In Study 3, the experimental lab was free of any external sounds, and the room lighting and temperature remained constant. All visual displays and sounds were recorded by high-definition cameras. The experimenter was not visible to the participants but observed video streams from the cameras. Participants read that the FSR acted autonomously, but for consistency, the interaction was being controlled by a hidden programmer. This Wizard of Oz method, first introduced by Gould et al. (1983), is widely employed in human–robot interaction research (see Homburg 2018). It mainly involves "a person (usually the experimenter, or a confederate) remotely operating a robot, controlling any of a number of things, such as its movement, navigation, speech, gestures, etc." (Riek 2012, p. 119). To improve the natural flow of conversation during the experiments, we established natural language processing, so the Pepper robot we used in these studies responded appropriately to comments or questions by participants. Pepper is a 120 cm humanoid robot with 20 degrees of freedom, produced by Softbank (Japan). To standardize the interaction, we adopted standardized service scripts, similar to hotels in practice. The script for Pepper was coded in Python. In Study 3, participants were students, enrolled in psychology, management, or computer science courses. They received financial remuneration of $12. Two trained drama students served as confederates, responsible for transmitting norm-related behaviors. These confederates did not have any knowledge of the hypotheses or specific purpose of the study. To increase the realism of the experiments, they referred to an existing service offer by a hotel. Relevant material was available during the conversation, such as maps and flyers.

Study 3 features a laboratory experiment that mimics a hotel setting, in which participants were served by a FSR or FLE that exhibited either norm compliance or norm violations as a manipulated condition (2×2 between-subjects design, random assignment). The service agent in the norm-compliant condition was very friendly and helpful toward the customer and showed particular interest in the customer's vacation by offering additional sightseeing tips. The norm violation involved incivility toward the customer, such that the service agent explicitly indicated a lack of interest in the customer's trip and indicated a desire not to waste time on the interaction. We obtained responses from 136 participants, of whom 52.2% were men, 42.6% were women, and 5.1% did not specify their gender. Among the 129 participants who indicated their age (7 did not), the mean age was 23.24 years (SD = 6.60 years). A post-hoc G*Power analysis resulted in a necessary total sample size of 32 participants to detect the effect; with our sample we fulfill this prerequisite. For the calculation, we used the results of the interaction effect as a prerequisite for the black sheep effect to arise. Due to our real-life setting, we included a large buffer, resulting in the much bigger sample size. Participants noted how they perceived the service provided by the service agent during the encounter. Those in the norm-violation condition (M = 3.56, SD = 1.69) reported that they perceived significantly lower customer orientation than those in the norm-compliance condition (M = 6.05, SD = 1.11; $t(111.09) = 10.10$, $p < .05$), so the manipulation worked as intended. We first confirmed that the social categorization differed between the FLE and FSR for this sample. Two items (Cronbach's $\alpha = .90$) that measured human likeness on 7-point Likert-type scales affirmed that participants perceived the FLE as significantly more humanlike (M = 5.85, SD = 1.39) than the FSR (M = 3.69, SD = 1.70; $t(99) = -6.34$, $p < .05$).

Findings

Next, we proceeded to the multivariate analysis of variance (MANOVA), with service behavior and service condition as independent variables and customer satisfaction (five items; Stock et al. 2017; Cronbach's α = .95) and loyalty (four items; Stock et al. 2017; Cronbach's α = .92) as outcomes. The results indicated a statistically significant difference between conditions for both service behavior ($F(2, 131) = 52.58$, $p < .05$, partial $\eta^2 = .45$, Wilk's $\lambda = .56$) and the service condition ($F(2, 131) = 23.62$, $p < .05$, partial $\eta^2 = .27$, Wilk's $\lambda = .74$), as well as a statistically significant interaction effect ($F(2, 131) = 24.13$, $p < .05$, partial $\eta^2 = .27$, Wilk's $\lambda = .73$). This effect is a prerequisite for the black sheep effect, in that the outcomes depend on the combination of both independent variables. As we detail in Table 3, the post hoc tests further specified that in the case of norm-violating behaviors by the service employee, customer loyalty was significantly higher in the FSR than the FLE condition ($\Delta M = 1.23$, SE = .33). When the service employee behaved in a norm-compliant manner, participants reported increased loyalty toward the human FLE ($\Delta M = 1.11$, SE = .32). This pattern is in line with H_{1b} and H_{2b}: In the norm violating condition, the social robot was judged more favorably, but in the norm compliant condition, the member of the ingroup (FLE) was favored. The black sheep effect also emerged in relation to customer satisfaction: When the service agent violated norms, participants expressed more satisfaction with the FSR than with the FLE ($\Delta M = 2.39$, SE = .35), in support of H_{2a}. In the norm compliant condition, participants favored the human FLE over the FSR, though this difference was not significant, so we cannot confirm H_{1a}.

		Service Condition							
		FSR		FLE					
		Mean (SD)		Mean (SD)		Mean Difference (SE)		p-Value	
		Sat	Loy	Sat	Loy	Sat	Loy	Sat	Loy
Service Behavior	Norm violation	5.41 (1.48)	5.10 (1.68)	3.01 (1.35)	3.87 (1.31)	2.39* (.35)	1.23* (.33)	.00	.00
	Norm compliance	6.07 (1.06)	5.54 (1.53)	6.46 (.62)	6.65 (.58)	-.39 (.28)	-1.11* (.32)	.16	.00

Table 3. Study 3, Post Hoc Test of Customer Satisfaction and Loyalty, by Service Behavior and Condition (H1 and H2)

Notes: Customer satisfaction (Sat) and loyalty (Loy) were measured on 7-point Likert-type scales from 1 ("totally disagree") to 7 ("totally agree"); n(1) = (33|33); n(2) = (35|35). FSR = frontline service robot; FLE = frontline employee. *$p \leq .05$.

In summary, customers downgraded the human FLE, which was part of their social ingroup, more than the FSR, which was part of a social outgroup, in support of H_{2a} and H_{2b} and the black sheep effect in the norm-violating condition. For the norm-compliant service behavior, the pattern of results indicates ingroup favoritism toward the human FLE, significantly in terms of customer loyalty (H_{1b}) but only directionally with regard to customer satisfaction (non-significant, cf. H_{1a}).

Study 4: Black Sheep Effect, Service Failure (Online Experiment)

Method

To investigate the black sheep effect in a service failure setting, the Study 4 participants interacted with either a human FLE or FSR in a hotel setting, and we used an interactive online study with prerecorded videos that mimic a typical hotel check-in situation. The service failure consisted of a booking mistake that could be blamed on the service agent; an alternative room was offered to the participants. No additional information regarding the circumstances of the service failure was available. The service failure might make people feel angry and frustrated, so we investigate these outcome variables. We excluded participants who failed our attention checks (i.e., picking a certain value on the scale as an item) and did not watch the videos in full (13 in the FSR condition and 26 in the FLE condition). The sample consisted of 200 participants, with 100 in the FSR and 100 in the human FLE condition, of whom 47% were men. They had a mean age of 40.37 years (SD = 11.55 years). A post-hoc G*Power analysis resulted in a necessary total sample size of 150

participants to detect the effect; with our sample we fulfill this prerequisite. For the calculation, we used the results of the t-tests. Due to our online setting, we included a large buffer, resulting in the much bigger sample size. As a manipulation check, we asked participants, after the interaction with the service agent, whether a service failure occurred. On a 7-point Likert-type scale, they clearly recognized the service failure, with a mean value of 5.34 (median = 5.00, SD = 1.39). Furthermore, in response to more detailed questions about the cause of the service failure, they clearly indicated it was attributable to the service agent, as a representative of the hotel, not to other unfavorable conditions (M = 3.16, median = 3.00, SD = 1.26). We assessed participants' perceptions of the realism of the service encounter, on a 7-point Likert-type scale (Cronbach's α = .96). The mean value of 4.70 (SD = 1.73) and median of 5.00 indicated the scenario was generally perceived as realistic. Furthermore, the participants assigned the FLE to their social ingroup (M = 3.63, SD = 1.84) and the FSR to an outgroup (M = 2.59, SD = 1.79; $t(198) = 4.05$, $p < .05$; 7-point Likert-type scales, Cronbach's α = .95, adapted from Evers et al. 2008 and Henry et al. 1999), again confirming the social categorization effect. The outcome variables were customer anger (three items; Gelbrich 2010; Cronbach's α = .94) and frustration (three items; Gelbrich 2010; Cronbach's α = .94) in response to the service failure.

Findings

The results indicated that participants were significantly angrier ($t(198) = 1.98$, $p < .05$) and more frustrated ($t(198) = 2.05$, $p < .05$) with the human FLE than with the FSR ($\Delta M_{Anger} = .56$, SE = .28; $\Delta M_{Frustration} = .58$, SE = .28), as we summarize in Table 4. We thus find evidence of the black sheep effect on customer anger and frustration, such that the FLE was downgraded more, compared with the FSR, for similar behaviors. Thus, we confirmed H_3 and the black sheep effect in a service failure context.

Service Condition							
FSR		FLE					
Mean (SD)		Mean (SD)		Mean Difference (SE)		p-value	
Ang	Frus	Ang	Frus	Ang	Frus	Ang	Frus
3.35 (1.92)	4.16 (1.92)	3.91 (2.10)	4.73 (2.05)	.56 (.28)*	.58 (.28)*	.049	.042

Table 4. Study 4, t-Tests of Customer Anger and Frustration, by Service Condition (H3)

Notes: Anger (Ang) and frustration (Frus) were measured on 7-point Likert-type scales from 1 ("totally disagree") to 7 ("totally agree"); n = (100|100). FSR = frontline service robot; FLE = frontline employee. *$p \leq .05$.

Discussion

Service providers increasingly rely on robots to engage in various service encounters with customers (Mende et al. 2019); some analysts even assert that service robots will be the impetus for the next industrial revolution (Ivanov and Webster 2017). Due to their automated social presence (Chung and Cakmak 2018b), service robots can mimic customer interactions with a real social entity, yet insights into how customers respond to such interactions are growing but remain scarce (Murphy et al. 2016). By investigating customer responses to a FSR compared with a human FLE, we not only shed light on the black sheep effect, by which customers upgrade norm-compliant behaviors and derogate norm-violating behaviors of a FLE as an ingroup member compared with a FSR as an outgroup member (Pinto et al. 2010), but also establish evidence of social categorization in this specific context. That is, in our core study, we identify a black sheep effect during service encounters, such that customers upgrade norm-compliant behaviors and downgrade norm-violating behaviors of the FLE as an ingroup member, but offer contrary assessments of the FSR as an outgroup member. With a pre-study, we demonstrate that customers categorize a FSR as an outgroup member and a FLE as an ingroup member, and this categorization remains stable across multiple service settings and interactions. Together, these studies establish novel insights for research and practice.

Theoretical Implications

This research extends social identity theory (Tajfel and Turner 1986) to customer–robot service interactions as an underlying mechanism for the black sheep effect. Customers define human FLEs as part of their ingroup and FSRs as part of an outgroup, categorizations that remain stable across various service settings (hospitality, retailing, tourism). In line with extant findings, we show that customers prefer service employees who are similar to them (Gill et al. 2017), due to their general favoritism for members they categorize as part of a social ingroup as opposed to an outgroup (Steain et al. 2019). As our contribution, we extend this finding to FSRs and establish that customers consistently and organically consider them an outgroup in comparison to human FLEs. Our findings thus offer crucial insights into the underlying theoretical mechanisms that drive technology acceptance and customer responses to service encounters with FSRs. By investigating longitudinal effects in relation to social categorization, we also add to social identity theory, which does not offer predictions about potentially dynamic changes in social categorization after multiple interactions (Deaux 1993). We find that social categorization of the FSR as outgroup members remained stable across multiple interactions. As service robots continue to spread rapidly into various service settings, more frequent interactions represent a realistic scenario. Our longitudinal design, involving multiple customer–robot interactions across multiple service scenarios, provides insights into the implications of such repeated exposures, in response to calls to examine customer responses to customer–robot interactions (Mende et al. 2019) and service technologies (Bhattacherjee and Premkumar 2004) over time. Our finding of consistent differences in the social categorization of FSRs and human FLEs builds a stable basis for black sheep effects as another extension to social identity theory (Marques et al. 1988).

Regarding the black sheep effect, we argued that a FLE, as a member of the ingroup, earns upgraded assessments when he or she displays norm-compliant behaviors but suffers downgraded ones for norm-violating behaviors, relative to a FSR as an outgroup member (Pinto et al. 2010). This effect has been identified already in customer–FLE service encounters (Gill et al. 2017); with our experimental studies, we extend these insights to include customer–FSR interactions too. That is, the findings, consistent with the black sheep effect, indicate that people express more loyalty toward a norm-compliant FLE compared with a FSR. For customer satisfaction, the difference is not significant but still in the predicted direction, in that participants were more satisfied with the performance of the human FLE. When both service agents violated norms, our study participants responded more positively toward the outgroup FSR, with more satisfaction and loyalty. That is, they "punished" the ingroup human FLE and downgraded this service agent more harshly than the corresponding FSR outgroup members (Lo Monaco et al. 2011; Mendoza et al. 2014). The causes of the black sheep effect are deeply rooted in social identity theory, which predicts that people generally attempt to preserve their own social self-image (Tajfel 1981), a notion that we evaluate with our pre-study regarding social categorization. Upgrading the social ingroup is a protective mechanism, in that it confirms and supports the person's own (positive) self-image. But when confronted by a negatively behaving ingroup member, people try to distance themselves, whereas a social outgroup member, such as the FSR, does not represent a direct threat to their social identity and therefore does not evoke equally negative evaluations (Pinto et al. 2010). By establishing these effects with a comparison of customers' interactions with FLEs versus FSRs, we add to extant research that examines the black sheep effect in either human–human interactions (e.g., Marques et al. 1988) or human–robot interactions (Steain et al. 2019). In addition, we gather new insights into the black sheep effect by investigating service failure situations and negative customer responses (anger, frustration) toward ingroup versus outgroup members. The results clarify that when the service failure can be ascribed to the service agent, the black sheep effect is evident, such that customers express more anger and frustration toward the human FLE than the FSR, signaling "punishment" (Lo Monaco et al. 2011; Mendoza et al. 2014). That is, the black sheep effect occurs in failure-free service encounters but also after a service failure.

Managerial Implications

Social categorization strongly informs human–robot service encounters and directs customers' responses and their acceptance of new technologies, which implies several important implications for information systems and marketing practitioners. Broadly, we show that service employees have tremendous impacts (Mende et al. 2019) and strongly affect customer responses and are relevant to technology acceptance and service outcomes such as customer loyalty, satisfaction, anger, and frustration (Agarwal and Karahanna 2000; Stock et al. 2017; Strizhakova et al. 2012). To create suitable service settings, practitioners should

take these influences closely into consideration when establishing and programming FSRs. With our research, we offer three specific recommendations. First, the insights regarding social categorization indicate that customers will categorize FLEs as part of their ingroup and FSRs as outgroup, and therefore, customer-related responses differ, even if FSRs and FLEs display similar behaviors. These findings are especially relevant for practitioners as we were able to identify the behavior of FSRs and FLEs as important antecedents for customer satisfaction: Due to their automated social presence (Chung and Cakmak 2018b), FSRs are able to create a social experience for customers, however, not to the same extent as FLEs. Thus, firms need to bear in mind that FSRs—in positive service encounters—might lead to decreased customer satisfaction. Our results suggest that these outcomes remain stable over time as outgroup categorization of FSRs does not change after multiple interactions. Second, our findings further extend research on norm-violating (i.e., incivility) behaviors of FLEs. In particular employee-customer interactions it might be natural for FLEs to showcase rejecting behaviors, such as in case of unruly customers (van Jaarsveld et al. 2010). These kinds of situations can have negative psychological effects (e.g., exhaustion; Stock & Bednareck 2014) on FLEs and might potentially even lead to negative exchange spirals (Groth & Grandey 2012). Here, negative emotions (e.g., aggression) continuously increase during the interaction between customer and FLE. According to the black sheep effect, norm violations by a member of a social outgroup are less significant than those by a member of the social ingroup. Our results confirm the black sheep effect after a service failure in that customers respond less negatively to (sometimes inevitable) failures if they encounter a FSR rather than a human FLE. These reactions can effectively buffer negative exchange spirals (Groth and Grandey 2012) in customer-FSR interactions after a service failure. In other words, in case of a service failure it would be beneficial to deploy FSRs, to reduce customers' anger and frustration. Similarly, FLEs will not be confronted with negative customer responses which might have detrimental effects (Walker et al. 2017). Thus, the benefits of FSRs in critical service encounters are twofold: the experience for the customer is more pleasant despite a service failure and negative effects for FLEs can be avoided.

Limitations and Further Research

A lack of control and additional explanatory variables, such as expectations and attitudes, is a limitation of our studies. Due to differences in study designs, only qualitative comparisons between studies are possible. Further research might include other cognitive responses, such as perceived inequity or injustice, that often arise in human–human interactions, to determine their prevalence in human–robot interactions (Grewal 2012; Schoefer and Ennew 2005). In addition, our study design relies mainly on simulated experiments, and the results obtained from these realistic but still artificial conditions, which we control to ensure consistent interactions (Homburg 2018), might not generalize to real-world conditions. Research should continue to examine human–robot interactions in the real world. Finally, extant research addresses various technology-supported interfaces, such as point-of-sale terminals or tablets (Chen et al. 2021; Giebelhausen et al. 2014). Further research should compare responses to terminals versus FSRs, which may evoke differences due to the greater automated social presence of robots (Wirtz et al. 2018).

Conclusion

We provide proof of concept that a black sheep effect occurs in regular and failed service encounters with FLEs and FSRs: Norm-violating FLE behaviors are downgraded to a larger extent than equivalent behaviors by FSRs. In turn, customers upgrade norm-compliant FLE behaviors more than FSR counterparts. This effect is rooted in social identity theory as FSRs represent an outgroup while FLEs social ingroup.

References

Agarwal, R., and Karahanna, E. 2000. "Time Flies When You're Having Fun: Cognitive Absorption and Beliefs about Information Technology Usage," *MIS Quarterly* (24:4), pp. 665-694.

Aggarwal, P. 2004. "The Effects of Brand Relationship Norms on Consumer Attitudes and Behavior," *Journal of Consumer Research* (31:1), pp. 87–101.

Ashforth, B. E., and Mael, F. 1989. "Social Identity Theory and the Organization," *Academy of Management Review* (14:1), pp. 20-39.

Barrett, M., Davidson, E., Prabhu, J., and Vargo, S. L. 2015. "Service Innovation in the Digital Age," *MIS Quarterly* (39:1), pp. 135-154.

Belanche, D., Casaló, L. V., Flavián, C., and Schepers, J. 2020a. "Service Robot Implementation: A Theoretical Framework and Research Agenda," *The Service Industries Journal* (40:3-4), pp. 203-225.

Belanche, D., Luis V. Casaló, L. V., Flavián, C., and Schepers, J. 2020b. "Robots or Frontline Employees? Exploring Customers' Attributions of Responsibility and Stability after Service Failure or Success," *Journal of Service Management* (31:2), pp. 267-289.

Bhattacherjee, A., and Premkumar, G. 2004. "Understanding Changes in Belief and Attitude toward Information Technology Usage: A Theoretical Model and Longitudinal Test," *MIS Quarterly* (28:2), pp. 229-254.

Bornstein, R. F., and D'Agostino, P. R. 1992. "Stimulus Recognition and the Mere Exposure Effect," *Journal of Personality and Social Psychology* (63:4), pp. 545-552.

Čaić, M., Mahr, D., and Oderkerken-Schröder, G. 2019. "Value of Social Robots in Services: Social Cognition Perspective," *Journal of Services Marketing* (33:4), pp. 463-478.

Cenfetelli, R. T., Benbasat, I., and Al-Natour, S. 2008. "Addressing the What and How of Online Services: Positioning Supporting-Services Functionality and Service Quality for Business-to-Consumer Success," *Information Systems Research* (19:2), pp. 161-181.

Chen, N., Mohanty, S., Jiao, J., and Fan, X. 2021. "To Err Is Human: Tolerate Humans instead of Machines in Service Failure," *Journal of Retailing and Consumer Services* (59:C), 102363

Choi, S., Mattila, A.S., and Bolton, L. E. 2021. "To Err Is Human(-oid): How Do Consumers React to Robot Service Failure and Recovery?," *Journal of Service Research* (24:3), pp. 354-371.

Chung, M., and Cakmak, M. 2018b. "How Was Your Stay?: Exploring the Use of Robots for Gathering Customer Feedback in the Hospitality Industry," in *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, China.

Collett, J. L., and Childs, E. 2011. "Minding the Gap: Meaning, Affect, and the Potential Shortcomings of Vignettes," *Social Science Research* (40:2), pp. 513–522.

Crolic, C., Thomaz, F., Hadi, R., and Stephen, A. T. 2022. "Blame the Bot: Anthropomorphism and Anger in Customer–Chatbot Interactions," *Journal of Marketing* (86:1), pp. 132–148.

Damiano, L., Dumouchel, P., and Lehmann, H. 2015. "Artificial Empathy: An Interdisciplinary Investigation," *International Journal of Social Robotics* (7:1), pp. 3–5.

De Cremer, D., and Vanbeselaere, N. 1999. "I Am Deviant, because the Impact of Situational Factors upon the Black Sheep Effect," *Psychologica Belgica* (39:1), pp. 71–79.

De Graaf, M. M., Allouch, S. B., and van Dijk, J. A. 2016. "Long-Term Evaluation of a Social Robot in Real Homes," *Interaction Studies* (17:3), pp. 462-491.

Deaux, K. 1993. "Reconstructing Social Identity," *Personality & Social Psychology Bulletin* (19:1), pp. 4-12.

Deligianis, C., Stanton, C. J., McGarty, C., and Stevens, C. J. 2017. "The Impact of Intergroup Bias on Trust and Approach Behaviour towards a Humanoid Robot," *Journal of Human-Robot Interaction* (6:3), pp. 4–20.

Eidelman, S., and Biernat, M. 2003. "Derogating Black Sheep: Individual or Group Protection?," *Journal of Experimental Social Psychology* (39:6), pp. 602–609.

Eidelman, S., Silvia, P. J., and Biernat, M. 2006. "Responding to Deviance: Target Exclusion and Differential Devaluation," *Personality and Social Psychology Bulletin* (32:9), pp. 1153–1164.

Erasmus, A. C., Bishoff, E., and Rousseau, G. G. 2002. "The Potential of Using Script Theory in Consumer Behaviour Research," *Journal of Consumer Sciences* (30:1), pp. 1–9.

Esterwood, C., and Robert, L. 2021. "Robots and COVID-19: Re-Imagining Human-Robot Collaborative Work in Terms of Reducing Risks to Essential Workers," *ROBONOMICS: The Journal of the Automated Economy* (1:1), pp. 1-5.

Evers, V., Maldonado, H., Brodecki, T., and Hinds, P. 2008. "Relational vs. Group Self-Construal: Untangling the Role of National Culture in HRI," in *2008 3rd ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, Netherlands, pp. 255-262.

Fan, A., Wu, L., Miao, L., and Mattila, A. S. 2020. "When Does Technology Anthropomorphism Help Alleviate Customer Dissatisfaction after a Service Failure?–The Moderating Role of Consumer Technology Self-Efficacy and Interdependent Self-Construal," *Journal of Hospitality Marketing & Management* (29:3), pp. 269-290.

Fasola, J., and Matarić, M. 2011. "Comparing Physical and Virtual Embodiment in a Socially Assistive Robot Exercise Coach for the Elderly," *Center for Robotics and Embedded Systems*, Los Angeles, CA.

Fraune, M. R., Šabanović, S., and Smith, E. R. 2017. "Teammates First: Favoring Ingroup Robots over Outgroup Humans," in *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, Portugal, pp. 1432–1437.

Gelbrich, K. 2010. "Anger, Frustration, and Helplessness after Service Failure: Coping Strategies and Effective Informational Support," *Journal of the Academy of Marketing Science* (38:5), pp. 567–585.

Giebelhausen, M., Robinson, S. G., Siranni, N. J., and Brady, M. K. 2014. "Touch versus Tech: When Technology Functions as a Barrier or a Benefit to Service Encounters," *Journal of Marketing* (78:4), pp. 113–124.

Gill, T., Kim, H. J., and Ranaweera, C. 2017. "Ethnic Stereotyping in Service Provision: When Do Stereotypes Affect the Performance Expectations and Evaluation of Ethnic Service Providers?," *Journal of Service Theory and Practice* (27:3), pp. 520-546.

Gockley, R., Bruce, A., Forlizzi, J., Michalowski, M., Mundell, A., Rosenthal, S., Sellner, B., Simmons, R., Snipes, K., Schultz, A., and Wang, J. 2006. "Designing Robots for Long-Term Social Interaction," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robotics and Systems*, Canada.

Gockley, R., Simmons, R., and Forlizzi, J. 2006. "Modeling Affect in Socially Interactive Robots," in *15th IEEE International Symposium on Robot and Human Interactive Communication*, UK, pp. 558–563.

Gombolay, M. C., Gutierrez, R. A., Clarke, S. G., Sturla, G. F., and Shah, J. A. 2015. "Decision-Making Authority, Team Efficiency and Human Worker Satisfaction in Mixed Human–Robot Teams. *Autonomous Robot* (39), pp. 293–312.

Gould, J. D., Conti, J., and Hovanyecz, T. 1983. "Composing Letters with a Simulated Listening," *Communications of the ACM* (26:4), pp. 295–308.

Groom, V., Srinivasan, V., Bethel, C. L., Murphy, R., Dole, L., and Nass, C. 2011. "Responses to Robot Social Roles and Social Role Framing," in *2011 International Conference on Collaboration Technologies and Systems*, USA, pp. 194–203.

Groth, M., and Grandey, A. 2012. "From Bad to Worse: Negative Exchange Spirals in Employee–Customer Service Interactions," *Organizational Psychology Review* (2:3), pp. 208–233.

Heitlinger, L., Stock-Homburg, R., and Wolf, F. D. 2022. "You Got the Job! Understanding Hiring Decisions for Robots as Organizational Members," in *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction (HRI '22)*, IEEE Press, pp. 530-540.

Henry, K. B., Arrow, H., and Carini, B. 1999. "A Tripartite Model of Group Identification: Theory and Measurement," *Small Group Research* (30:5), pp. 558-581.

Hess Jr., R. L., Ganesan, S., and Klein, N. M. 2003. "Service Failure and Recovery: The Impact of Relationship Factors on Customer Satisfaction," *Journal of the Academy of Marketing Science* (31:2), pp. 127–145.

Ho, T. H., Tojib, D., and Tsarenko, Y. 2020. "Human Staff vs. Service Robot vs. Fellow Customer: Does It Matter Who Helps Your Customer Following a Service Failure Incident?", *International Journal of Hospitality Management* (87), pp. 1-10.

Homburg, C., Mueller, M., and Klarmann, M. 2011. "When Should the Customer Really Be King? On the Optimum Level of Salesperson Customer Orientation in Sales Encounters," *Journal of Marketing* (75:2), pp. 55–74.

Homburg, C., Wieseke, J., and Hoyer, W. D. 2009. "Social Identity and the Service-Profit Chain," *Journal of Marketing* (73:2), pp. 38-54.

Homburg, N. M. 2018. "How to Include Humanoid Robots into Experimental Research: A multi-Step Approach," in *Proceedings of the 51st Hawaii International Conference on System Sciences*, USA.

Hu, Y., Min, H. (K.), and Su, N. 2021. "How Sincere Is an Apology? Recovery Satisfaction in a Robot Service Failure Context," *Journal of Hospitality & Tourism Research* (45:6), pp. 1022–1043.

Ivanov, S. H., and Webster, C. 2017. "The Robot as a Consumer: A Research Agenda," in *Marketing: Experience and Perspectives Conference*, Bulgaria.

Ivanov, S. H., Webster, C., and Berezina, K. 2017. "Adoption of Robots and Service Automation by Tourism and Hospitality Companies," Paper presented at *INVTUR Conference*, Portugal.

Jörling, M., Böhm, R., and Paluch, S. 2019. "Service Robots: Drivers of Perceived Responsibility for Service Outcomes," *Journal of Service Research* (22:4), pp. 404-420.

Kanda, T., Shiomi, M., Miyashita, Z., Ishiguro, H., and Hagita, N. 2009. "An Affective Guide Robot in a Shopping Mall," in *Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction*, USA, pp. 173–180.

Kanda, T., Shiomi, M., Miyashita, Z., Ishiguro, H., and Hagita, N. 2010. "A Communication Robot in a Shopping Mall," *IEEE Transactions on Robotics* (26:5), pp. 897–913.

Kirby, R., Forlizzi, J., and Simmons, R. 2010. "Affective Social Robot," *Robotics and Autonomous Systems* (58:3), pp. 322–332.

Kuo, C.-M., Chen, L.-C., and Tseng, C.-Y. 2017. "Investigating an Innovative Service with Hospitality Robots" *International Journal of Contemporary Hospitality Management* (29:5), pp. 1-17.

Laros, F. J., and Steenkamp, J.-B. E. 2005. "Emotions in Consumer Behavior: A Hierarchical Approach," *Journal of Business Research* (58:10), pp. 1437–1445.

Lee, K. M., Jung, Y., Kim, J., and Kim, S. R. 2006. "Are Physically Embodied Social Agents Better than Disembodied Social Agents? The Effects of Physical Embodiment, Tactile Interaction, and People's Loneliness in Human–Robot Interaction," *International Journal of Human-Computer Studies* (64:10), pp. 962–973.

Lee, M. K., Kiesler, S., and Forlizzi, J. 2010. "Receptionist or Information Kiosk: How Do People Talk With a Robot?," in *Proceedings of the Conference on Computer Supported Cooperative Work*, USA, pp. 31-40.

Leo, X., and Huh, Y. E. 2020. "Who Gets the Blame for Service Failures? Attribution of Responsibility toward Robot versus Human Service Providers and Service Firms," *Computers in Human Behavior* (113), 106520.

Lo Monaco, G., Piermattéo, A., Guimelli, C., and Ernst-Vintila, A. 2011. "Using the Black Sheep Effect to Reveal Normative Stakes: The Example of Alcohol Drinking Contexts," *European Journal of Social Psychology* (41:1), pp. 1–5.

Lu, V. N., Wirtz, J., Kunz, W. H., Paluch, S., Gruber, T., Martins, A., and Patterson, P. G. 2020. "Service Robots, Customers, and Service Employees: What Can We Learn from the Academic Literature and Where Are the Gaps," *Journal of Service Theory and Practice* (30:3), pp. 361-391.

Lv, X., Liu, Y., Luo, J., Liu, Y., and Li, C. (2021). "Does a Cute Artificial Intelligence Assistant Soften the Blow? The Impact of Cuteness on Customer Tolerance of Assistant Service Failure," *Annals of Tourism Research* (87:2), 103114.

Marques, J. M. 1990. "The Black Sheep Effect: Outgroup Homogeneity as a Social Comparison Process," in *Social Identity Theory: Constructive and Critical Advances*, Harvester Wheatsheaf, pp. 131-151.

Marques, J. M., Abrams, D., and Serodio, R. G. 2001. "Being Better by Being Right: Subjective Group Dynamics and Derogation of Ingroup Deviants When Generic Norms Are Undermined," *Journal of Personality and Social Psychology* (81:3), pp. 436–447.

Marques, J. M., and Yzerbyt, V. Y. 1988. "The Black Sheep Effect: Judgmental Extremity towards Ingroup Members in Inter- and Intra-Group Situations," *European Journal of Social Psychology* (18:3), pp. 287–292.

Matsumoto, D., and Ekman, P. 2004. "The Relationship among Expressions, Labels, and Descriptions of Contempt," *Journal of Personality and Social Psychology* (87:4), pp. 529-540.

McCrea, S. M., Erion, C. J., and Thürmer, J. L. 2022. "Why Punish Critical Outgroup Commenters? Social Identity, General Norms, and Retribution," *British Journal of Social Psychology* (61:3), pp. 711-728.

Mende, M., Scott, M. L., van Doorn, J., Grewal, D., and Shanks, I. 2019. "Service Robots Rising: How Humanoid Robots Influence Service Experiences and Elicit Compensatory Consumer Responses," *Journal of Marketing Research* (56:4), pp- 535–556.

Mendoza, S. A., Lane, S. P., and Amodio, D. M. 2014. "For Members Only: Ingroup Punishment of Fairness Norm Violations in the Ultimatum Game," *Social Psychological and Personality Science* (5:6), pp. 662-670.

Merkle, M. 2019. "Customer Responses to Service Robots–Comparing Human-Robot Interaction with Human-Human Interaction," in *Proceedings of the 52nd Hawaii International Conference on System Sciences*, pp. 8-11.

Murphy, J., Hofacker, C., and Gretzel, U. 2016. "Dawning of the Age of Robots in Hospitality and Tourism: Challenges for Teaching and Research," *European Journal of Tourism Research* (15), pp. 104–111.

Ojala, K., and Nesdale, D. 2004. "Bullying and Social Identity: The Effects of Group Norms and Distinctiveness Threat on Attitudes towards Bullying," *British Journal of Developmental Psychology* (22:1), pp. 19–35.

Packer, D. J., Ungson, N. D., and Marsh, J. K. 2021. "Conformity and Reactions to Deviance in the Time of COVID-19," *Group Processes & Intergroup Relations* (24:2), pp. 311-317.

Paetzel, M., Perugia, G., and Castellano, G. 2020. "The Persistence of First Impressions: The Effect of Repeated Interactions on the Perception of a Social Robot," in *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, UK, pp. 73-82.

Pinto, I. R., Marques, J. M., Levine, J. M., and Abrams, D. 2010. "Membership Status and Subjective Group Dynamics: Who Triggers the Black Sheep Effect?," *Journal of Personality and Social Psychology* (99:1), pp. 107-119.

Pitardi, V., Wirtz, J., Paluch, S., and Kunz, W. H. 2022. "Service Robots, Agency and Embarrassing Service Encounters," *Journal of Service Management* (33:2), pp. 389-414.

Riek, L. D. 2012. "Wizard of Oz Studies in HRI: A Systematic Review and New Reporting Guidelines," *Journal of Human-Robot Interaction* (1:1), pp. 119–136.

Riek, B. M., Gaertnert, S. L., Dovidio, J. F., Brewer, M. B., Manla, E. W., and Lamoreaux, M. J. 2008. "A Social-Psychological Approach to Postconflict Reconciliation," in *The Social Psychology of Intergroup Reconciliation*, Oxford: Oxford University Press, pp. 255–274.

Schoefer, K., and Ennew, C. T. 2005. "Emotional Responses to Service Complaint Experiences: The Role of Perceived Justice," *Journal of Services Marketing* (19:5), pp. 261–270.

Steain, A., Stanton, C. J., and Stevens, C. J. 2019. "The Black Sheep Effect: The Case of the Deviant Ingroup Robot," *PloS one* (14:10), e0222975.

Stock, R. M., and Bednarek, M. 2014. "As They Sow, so Shall They Reap: Customers' Influence on Customer Satisfaction at the Customer Interface," *Journal of the Academy of Marketing Science* (42:4), pp. 400-414.

Stock, R. M., de Jong, A., and Zacharias, N. A. 2017. "Frontline Employees' Innovative Service Behavior as Key to Customer Loyalty: Insights into FLEs' Resource Gain Spiral," *Journal of Product Innovation Management* (34:2), pp. 223–245.

Stock, R. M., and Hoyer, W. D. 2005. "An Attitude-Behavior Model of Salespeople's Customer Orientation," *Journal of the Academy of Marketing Science* (33:4), pp. 536–552.

Stock, R., and Hannig, M. 2020. "Is There a Privacy Paradox in the Workplace?," in *Forty-First International Conference on Information Systems (ICIS)*, India.

Strizhakova, Y., Tsarenko, Y., and Ruth, J. A. 2012. "I'm Mad and I Can't Get that Service Failure off my Mind. Coping and Rumination as Mediators of Anger Effects on Customer Intentions," *Journal of Service Research* (15:4), pp. 414–429.

Stubbs, K., Hinds, P. J., and Wettergreen, D. 2007. "Autonomy and Common Ground in Human-Robot Interaction: A Field Study," *IEEE Intelligent Systems* (22:2), pp. 42–50.

Tajfel, H. 1981. *Human Groups and Social Categories,* Cambridge: University Press.

Tajfel, H., and Turner, J. C. 1986. "The Social Identity Theory of Inter-Group Behavior," in *Psychology of Intergroup Relations*, Chicago: Nelson-Hall, pp. 7-24.

Tan, C. W., Benbasat, I., and Cenfetelli, R. T. 2016. "An Exploratory Study of the Formation and Impact of Electronic Service Failures," *MIS Quarterly* (40:1), pp. 1-29.

Van Doorn, J., Mende, M., Noble, S. M., Hulland, J., Ostrom, A. L., Grewal, D., and Petersen, J. A. 2017. "Domo Arigato Mr. Roboto: Emergence of Automated Social Presence in Organizational Frontlines and Customers' Service Experiences," *Journal of Service Research* (20:1), pp. 43–58.

Van Jaarsveld, D. D., Walker, D. D., and Skarlicki, D. P. 2010. "The Role of Job Demands and Emotional Exhaustion in the Relationship between Customer and Employee Incivility," *Journal of Management* (36:6), pp. 1486–1504.

Van Pinxteren, M. M., Wetzels, R. W. H., Rüger, J., Pluymaekers, M., and Wetzels, M. 2019. "Trust in Humanoid Robots: Implications for Services Marketing," *Journal of Services Marketing* (33:4), pp. 507-518.

Walker, D. D., Van Jaarsveld, D. D., and Skarlicki, D. P. 2017. "Sticks and Stones Can Break My Bones but Words Can also Hurt Me: The Relationship between Customer Verbal Aggression and Employee Incivility," *Journal of Applied Psychology* (102:2), pp. 163-179.

Wan, L. C., Hui, M. K., and Wyer Jr., R. S. 2011. "The Role of Relationship Norms in Responses to Service Failures," *Journal of Consumer Research* (38:2), pp. 260-277.

White, R. T., and Arzi, H. J. 2005. "Longitudinal Studies: Designs, Validity, Practicality, and Value," *Research in Science Education* (35), pp. 137–149.

Wirtz, J., Patterson, P. G., Kunz, W. H., Gruber, T., Lu, V. N., Paluch, S., and Martins, A. 2018. "Brave New World: Service Robots in the Frontline," *Journal of Service Management* (29:5), pp. 907–931.

Xiao, L., and Kumar, V. 2019. "Robotics for Customer Service: A Useful Complement or an Ultimate Substitute?," *Journal of Service Research* (24:1), pp. 9-29,.

Yam, K. C., Bigman, Y. E., Tang, P. M., Ilies, R., De Cremer, D., Soh, H., and Gray, K. 2021. "Robots at Work: People Prefer-and Forgive-Service Robots with Perceived Feelings," *Journal of Applied Psychology* (106:10), pp. 1557–1572.

You, S. and Robert, L. P. 2018. "Emotional Attachment, Performance, and Viability in Teams Collaborating with Embodied Physical Action (EPA) Robots," *Journal of the AIS* (19:5), pp. 377-407.

Zajonc, R. 1968. "Attitudinal Effects of Mere Exposure," *Journal of Personality & Social Psychology* (9:2), pp. 1-27.