Research Article

Smaller Than Expected

Effects of Imitative Action Regulation After Experiencing Social Exclusion

Carina G. Giesen, Laura Nagel, Matthäus Rudolph, and Klaus Rothermund

Department of Psychology, General Psychology II, Friedrich Schiller University Jena, Germany

Abstract. In two pre-registered studies, we investigated whether processes of imitative action regulation are facilitated after experiencing an episode of social exclusion. We reasoned that imitative action regulation effects should be more pronounced for participants who were socially excluded, providing them with an "automatic means" to socially reconnect with others. Participants played a virtual ball-tossing game to experimentally induce social exclusion or inclusion experiences. Subsequently, pairs of two participants engaged in an observational stimulus–response (SR) binding paradigm modeled after Giesen et al. (2014): Participants observed color categorization responses in their interaction partner ($trial_{n-1}$) and then executed (in)compatible responses in the subsequent trial ($trial_n$), with observation and responding occurring in alternation. Stimulus relation (repetition vs. change) from $trial_{n-1}$ to $trial_n$ was orthogonally manipulated. In both studies, stimulus-based retrieval effects of observationally acquired SR bindings were descriptively larger in socially excluded (compared with socially included) participants. However, none of the effects were statistically significant. Even a joint analysis of both experiments did not show the expected modulation. We discuss the implications of our findings for research on social exclusion effects on imitative action regulation processes.

Keywords: stimulus–response binding, event files, observational learning, social exclusion, action imitation

Did you ever enter the dining hall with a plate full of food, just to find out that your colleagues already had lunch but forgot to ask you to join? Do you recall the feeling of coming to school on Monday, just to hear all your classmates boast about that party on Saturday – the one you were not invited to? As human beings, we have a fundamental need to connect with others and to belong to social groups, which forms a constituent part of our identity (Baumeister & Leary, 1995; Tajfel & Turner, 2004). It thus comes as no surprise that being excluded from social groups is a highly aversive experience, and everybody of us knows situations like the ones described above. But how do we cope with these experiences of social exclusion, and how do we manage to reconnect with others?

Recent findings support the view that adaptive responses to experiences of social exclusion are *two-pronged* (Hess & Pickett, 2010, p. 453): On the one hand, one needs to defend the self against the immediate pain of social exclusion either by externalizing (i.e., showing aggressive/antisocial behavior, Twenge et al., 2007; particularly toward the rejector: Maner et al., 2007) or by internalizing responses (i.e., avoiding self-awareness, Hess & Pickett, 2010; reduced affect, Blackhart et al., 2009; for a review, see Williams, 2007). On the other hand, experiences of social exclusion trigger a set of emotional, cognitive, and behavioral processes that are tuned to rebuild and stabilize social bonds: For instance, Maner et al. (2007) report that socially excluded participants expressed greater interests in making new friends, were more motivated to work with others, and also distributed higher rewards to future interaction partners than control participants. This shows that socially excluded participants adopt *tend and befriend* strategies to reconnect with others. However, more subtle forms of *prosocial tuning* after experiencing social exclusion are also documented. For instance, Wilkowski et al. (2009) report increased sensitivity to gaze cuing in individuals who experienced social exclusion; Hess and Pickett (2010) show that socially excluded participants recalled more social (i.e., other-related) than self-related behaviors in an unannounced memory task. Furthermore, experiences of social exclusion also have impact on automatic affiliative behavior like behavioral mimicry: Lakin et al. (2008) demonstrated that after experiencing social exclusion, participants mimicked their interaction partner more than socially included participants. Increased mimicry is also documented for 5-year-old children who observed third-party exclusion in videos (Over & Carpenter, 2009). Together, these findings point toward a range of subtle behavioral processes that operate automatically to rebuild social bonds with others after experiencing social exclusion.

Approaches to Study Automatic Imitative Behaviors

Research on behavioral mimicry investigates automatic and unintentional copying of manners, gestures, postures, and other motor behaviors between two or more individuals (Chartrand & Bargh, 1999; for an overview, see Chartrand & Lakin, 2013). Usually, the occurrence and frequency of a specific behavior (like face touching) is the dependent variable. Behavior mimicry is often considered as *social glue*, since being mimicked goes along with increased liking, empathy, and helping behavior, and establishes rapport and reduced prejudices for the mimicker (Chartrand & Lakin, 2013). However, the behavior that is mimicked is typically irrelevant for one's own action regulation. This is different, however, for effects of imitative or joint action regulation as studied with stimulus compatibility tasks (like the *automatic imitation* task, Brass et al., 2001; for a discussion, see Chartrand & Lakin, 2013; Heyes, 2011). Here, the copied behavior is a response that is either compatible or incompatible with an observed response; latencies and accuracy of (in)compatible response execution serve as dependent measures. Whereas it is often assumed that studies on behavioral mimicry and compatibility tasks measure similar mechanism(s), this claim is not well-founded empirically (Genschow et al., 2017). The present research was therefore motivated by the following reasons. First, whereas the above discussed studies offer plausible explanations *why* participants show prosocial tuning in their behavior after experiences of social exclusion, they remain silent about *how* such modulation actually takes place. In our view, research on automatic, imitative action regulation and, in particular, transient binding effects between stimuli and observed responses (Giesen et al., 2014) fills this gap and provides new insights into the underlying cognitive mechanisms that might mediate after-effects of social exclusion on social imitation. For instance, it is possible that socially excluded persons give higher priority to the processing and retrieval of social information and are thus more likely to rely on episodic retrieval of stimulus-based bindings with observed responses, which would result in increased imitative responses. Second, it is important to examine whether findings from studies on behavioral mimicry can also be replicated in the realm of imitative action regulation (and vice versa). This is particularly important since paradigms that measure imitative action regulation effects with compatibility tasks may differ in their responsiveness to social modulations. Concerning the "automatic imitation" task (Brass et al., 2001) for instance, there is an ongoing debate on whether the imitation tendencies measured with this task can be socially moderated at all (Cracco & Brass, 2019; Ramsey, 2018). Thus, it is currently unresolved whether performance in this task can be regarded as an indicator of social functioning (see Cracco, Genschow et al., 2018; Cracco, Bardi et al., 2018). This is rather different for the currently endorsed measure of imitative action regulation – the observational stimulus–response (SR) binding paradigm – which so far appears to be sensitive to social modulations (see below). Third, and related to the previous point, the present study can be seen as an important contribution in the process of construct validating the observational SR binding paradigm when evaluating its responsiveness social modulations.

Measuring Imitative Action Regulation: The Observational SR Binding Task

Giesen et al. (2014; 2016) recently showed that stimuli may become bound to responses that are merely observed (but not executed by oneself). Stimulus repetition on a subsequent trial will retrieve the binding between stimulus and observed response from memory. This will systematically bias current behavior. To test this, Giesen et al. (2014) developed the observational SR binding paradigm, which represents a color categorization task shared between a dyad of two coactors. Participants classified the color of word stimuli as red or green via button press in alternating fashion in a sequential design. One participant (the actor) categorized the color of a word stimulus presented in the trial. At the same time, the other participant (the observer) saw only the word stimulus (but no color) and had to observe the action of the actor. To test whether mere observation of the response results in an observationally acquired SR binding, the former observer became actor in the subsequent trial and had to categorize the color of the word stimulus. Stimulus relation from trial$_{n-1}$ to trial$_n$ (stimulus repetition vs. change) and compatibility between observed and to-be-executed responses (compatible vs. incompatible) were manipulated orthogonally. Stimulus repetition from trial$_{n-1}$ to trial$_n$ should retrieve observed responses from memory, producing performance benefits (costs) when retrieved responses are compatible (incompatible) with to-be-executed responses (statistically, this effect pattern is reflected in a stimulus relation × response compatibility interaction).

Indeed, findings by Giesen et al. indicate that people rely on such observationally acquired SR bindings to regulate their own actions. However, retrieval effects only emerged if the observed responses were performed by *socially relevant* others. Specifically, in the study by Giesen et al. (2014), retrieval effects were only obtained when participants had

to interact in a cooperative or competitive way but were absent when participants worked independently of each other. In addition to this situationally induced interdependence between coactors, Giesen et al. (2018) showed that retrieval of observationally acquired SR bindings also occurs for more chronic forms of interdependence, as is constituent for romantic relationships. In their study, retrieval effects for observationally acquired SR bindings emerged only for romantically involved interaction partners, but not for pairs of strangers. This highlights that chronic interdependence will produce automatic retrieval of observational SR bindings independently of the task itself. Together, these findings support the view that retrieval of observationally acquired SR bindings reflect genuinely social effects and indicate that the paradigm is suitable to study effects of social modulatory influences in the context of imitative action regulation.

Building on earlier findings from the influence of social exclusion experiences on automatic affiliative behaviors like behavioral mimicry, we were interested whether similar findings can be obtained for retrieval of observationally acquired SR bindings. In the present study, participants played a virtual ball-tossing game (*Cyberball*, Zadro et al., 2004) to experimentally induce social exclusion or inclusion experiences, before engaging in the observational SR binding task. We reasoned that stimulus-based retrieval effects of observationally acquired SR bindings would be stronger in socially excluded, compared with socially included, participants. Such a finding would attest to the subtle nature of prosocial tuning as an "automatic means" to socially reconnect with others.

To anticipate results, Experiment 1a yielded findings that were in line with our hypothesis but just failed conventional levels of significance. Thus, Experiment 1b was planned as a higher-powered exact replication. According to an a priori calculation with G*Power 3.1 (Faul et al., 2007), n = 156 participants were required (d_z = 0.35, α = .05, $1 - \beta$ = .70). Unfortunately, many data collection appointments had to be canceled due to the nationwide lockdown as a consequence of the corona pandemic in Germany in spring and summer of 2020. The final sample size of Experiment 1b was thus considerably smaller than originally intended. To compensate for this, a joint analysis on data of both experiments was run.

Method

Ethics Vote and Pre-Registration

Both experiments were in accordance with the Declaration of Helsinki. Ethical approval was granted by the Ethical Commission of the FSU Jena (Experiment 1a: FSV 19/15; Experiment 1b: FSV 20/001). Prior to data collection, the exact method, design, hypotheses, data preparation, and planned analyses were pre-registered online at https://www.aspredicted.org.[1]

Participants

In total, 73 psychology students from FSU Jena took part in Experiment 1a. One participant had to be excluded due to excessive error rates (>20% errors in the memory test), meaning that data of 72 students[2] were analyzed (62 female; Mdn_{age} = 20 years, SD_{age} = 3.4 years). Another 68 students from FSU Jena took part in Experiment 1b. However, 10 participants had to be excluded: eight of them because they already took part in Experiment 1a and two additional participants because of excessive error rates (>20%) in the memory test and/or categorization task, meaning that data of 58 students were analyzed (37 female; Mdn_{age} = 20 years, SD_{age} = 7.3 years). All participants were native German speakers. They received partial course credit and sweets or 4.50€ for their participation. Both experiments lasted for 45 min.

Design

Both experiments comprised a 2 × 2 × 2 mixed factors design with the within-subjects factors stimulus relation and response compatibility and the between-subjects factor experimental condition. The between factor experimental condition was manipulated with the *Cyberball* online game: 50% of all participants were socially excluded (Experiment 1a: n = 36; Experiment 1b: n = 29), whereas 50% of all participants were socially included (Experiment 1a: n = 36; Experiment 1b: n = 29; see Procedure for details). Both within-subjects factors were manipulated within the

[1] Experiment 1a: https://aspredicted.org/blind.php?x=2ta2gg, Experiment 1b: http://aspredicted.org/blind.php?x=2yn5ta.

[2] Five participants in Experiment 1a performed the observational SR binding task with a confederate; however, data of confederates were excluded prior to all analyses. Also, we want to correct a misunderstanding that may be derived from the pre-registration text of Experiment 1 regarding the targeted sample size. In the pre-registration, we wrote that "...we will offer the experiment until at least 60 people have agreed to participate or until June 30, 2019 (whichever comes first)." However, we realized that in fact, we adopted a "whichever comes last" strategy: That is, we intended to sample a minimum of 60 participants but would continue with data collection if the laboratories were still available in the booked time period, since a higher N provides more statistical power to the between-subjects factor of experimental condition.

observational SR binding task. Stimulus relation was manipulated by repeating or changing the word stimulus from $trial_{n-1}$ to $trial_n$ for 50% of all trials; response compatibility was manipulated by requiring actors to perform color categorization responses in $trial_n$ that were either compatible or incompatible to observed responses in $trial_{n-1}$ in 50% of all trials (see Procedure for details and examples). Release of rest-state keys in the observational SR binding task served as the dependent variable of interest.

Procedure

Unless mentioned otherwise, Experiments 1a and 1b followed the same procedure: Participants were recruited via e-mail and booked laboratory appointments online. Upon their arrival in the laboratory, participants were informed about the procedure and duration of the experiment. All participants provided written informed consent prior to taking part in the study. Each experiment consisted of three consecutive parts: First, participants played *Cyberball* (Zadro et al., 2004) individually on a computer. Instructions were provided on screen. In the *Cyberball* game, participants were informed that they would play an interactive online ball game together with two other persons; in reality though, both other players were controlled by the computer. When the game started, participants could decide to whom of the two other players they would like to throw a ball. The (programmed) other players kept the ball for a variable interval of 2–5 s to give the impression of a real person contemplating whom to pass the ball next. Participants were randomly assigned to one of the experimental conditions (social exclusion vs. inclusion). Participants in the *social exclusion* condition did no longer receive the ball after four throws; participants in the *social inclusion* condition received the ball as often as the other players. Within each condition, the sequence of throws was fixed, meaning that all participants of the same condition experienced the same sequence of throws. The *Cyberball* game lasted for 45 passes, which corresponded to roughly 6 min.

After the *Cyberball* game was over, participants moved to a different table to perform the observational SR binding task together with an unknown interaction partner. Pairs of two participants were seated opposite to each other at a table, each of them facing a 19″ monitor. Two response pads were fastened to the table on the left and right sides of the monitor (see Figure 1). Each response pad had a huge buzzer button in the middle (either red or green) and two small black keys (rest-state keys) in front of and behind each buzzer button. Response pads were connected to the computer via the parallel port. Participants were asked to place their left and right hands on each rest-state key in front of them and to keep these keys constantly pressed; however, if they had to execute the color categorization response, they were allowed to leave the respective rest-state key to hit either the red or green buzzer button. Participants worked through the observational SR binding task in pairs; however, the computer screens blocked eye contact between participants. Furthermore, participants were not allowed to talk with their interaction partner during the task.

The observational SR binding task was programmed with E-Prime 2. At the start of the experiment, demographic information of both participants (gender, age, handedness, and native language) was collected by the experimenter. Then, task instructions were given on screen. The observational SR binding task started when both participants finished reading the instructions and signaled their readiness by pressing down both rest-state keys.

In the observational SR binding task, both participants (called A and B) had the task to categorize the color of word stimuli in alternating fashion. Twenty-five neutral, monosyllabic, or disyllabic German adjectives serve as word stimuli (e.g., small, quiet, and edgy) in the task. The observational SR binding task works as follows (see Figure 1): Every trial started with a fixation cross presented centrally in white on a black screen (250 ms). Then, a word stimulus appeared in white font (250–400 ms) and then changed its color for participant A (the actor) to red or green; for the other participant B (who only observed the response carried out by participant A), the word remained white. The word stimulus remained on screen until a color categorization response was given by the actor or until 1,500 ms elapsed. Then, the next trial started with the presentation of a fixation cross on each screen (250 ms). Then, another word stimulus appeared in white font on each screen (250–400 ms) and changed its color for participant B; for participant A (who is now observing the response carried out by participant B), the word remained white. The word stimulus remained on screen until a color categorization response was given by the actor or until 1,500 ms elapsed. Thus, each color categorization response serves as a *prime* for the observer, whose color categorization response (speed and accuracy) is measured on the subsequent *probe* trial. After 40 randomly chosen probe trials (25%), a memory test occurred after probe response execution in which actors were asked to indicate which response they observed during the preceding observation trial by pressing the corresponding color key (until response). For the other participant, the screen remained black during the memory test trial.

At first, both participants worked through a practice block of 32 trials. Participants received error and time-out feedback in case of wrong or too slow responses. When more than 20% errors were committed during the practice block, practice was repeated; if the speed and accuracy criteria were not met after three consecutive runs of the practice block, the experiment was terminated (this happened for one pair of participants in Experiment 1b). Upon successful completion of the practice block, the proper experiment started which consisted of two

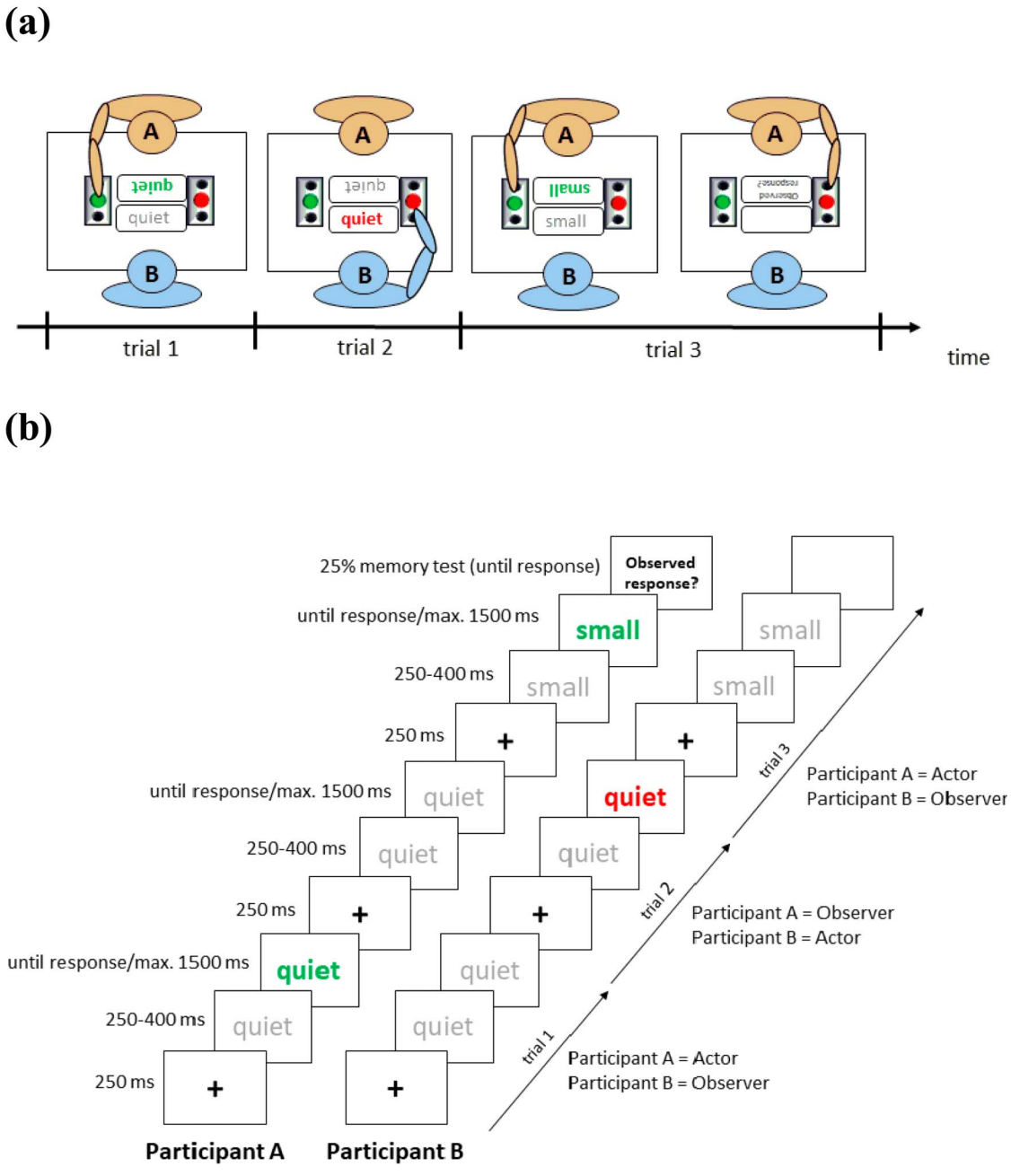

Figure 1. Schematic illustration of experimental setup (A) and trial sequence (B) for participants A and B (for illustrative purposes, background and neutral font colors are inverted). In Trial 1, participant A is actor and categorizes the color of the word stimulus (green); participant B is observer and the word stimulus does not change its color. In Trial 2, participant B is actor and has to categorize the word's color (red), whereas participant A is observer. In Trial 3, participant A is actor and categorizes the word color (green). For each participant, only "actor" trials are analyzed as a function of the immediately preceding (observation) trial. Hence, for participant B, Trial 2 is a stimulus repetition (SR) trial in which the required response is incompatible (IC) with the observed response in Trial 1 (SRIC sequence; cf. Table 1). For participant A, Trial 3 is a stimulus change (SC) trial in which the required response is incompatible with the observed response in Trial 2 (SCIC sequence). After execution of the color response in Trial 3, participant A is prompted with a memory test trial and has to press the response key that corresponds to the response they observed in Trial 2. Word stimuli were presented in red/green font color to actors (depicted in boldface; see online article for colored version of Figure 1). To observers, word stimuli were presented in white font color (depicted in gray). Stimuli are not drawn to scale.

blocks of 160 experimental trials plus two filler trials at the start of each block that were not analyzed (because they had no preceding observation trial). For experimental trials, no error or time-out feedback was provided. Participant A started as actor in the first experimental block, and participant B started as actor in the second experimental block.

Experimental trials were constructed with respect to the experimental design: In 50% of all trials, the word

stimulus was repeated from trial$_{n-1}$ to trial$_n$ (stimulus repetition, e.g., small–small) or changed (stimulus change, e.g., edgy–small). Independent of that, 50% of all required color categorization responses in trial$_n$ were either compatible to observed color categorization responses in trial$_{n-1}$ (e.g., green–green) or incompatible (e.g., red–green). Further counterbalancing factors: 50% of all trials required a red (green) response and 50% of all memory trials required a red (green) response. In total, the observational SR binding task lasted for approximately 30 min.

When the observational SR binding task was completed, participants received a brief paper questionnaire with two sets of questions. The first set of questions measured how participants experienced the experimental situation and interaction partner in the observational SR binding task: On 7-point Likert scales, participants rated the experimental situation (three items: difficult/uncomfortable/negative [1] vs. easy/comfortable/positive [7]; means were averaged to a *situation* score), interaction style (one item: competitive [1] vs. cooperative [7]), and agreeableness of their interaction partner (four items: unsympathetic/insecure/unfriendly/incompetent [1] vs. sympathetic/confident/friendly/competent [7]; averaged to a composite *partner* score). Furthermore, they reported the degree of acquaintance with their interaction partner (1 = *not at all*; 5 = *very well*). A second set of questions measured how participants experienced the *Cyberball* game: On 5-point Likert scales, participants rated the number of passes they received during the game (1 = 0–20%; 5 = 80–100%) and how they felt during the game (six items: sad/uncomfortable/ignored/excluded/dismissed/angry [1] vs. happy/comfortable/noticed/included/accepted/relaxed [5]; averaged to a composite *social inclusion* score). In both experiments, additional questions were assessed for exploratory purpose but are not of interest here. When the questionnaire was completed, participants were rewarded for their participation. At the end of data collection, all participants were fully debriefed.

Results: Pre-Registered Analyses

Raw data and analysis scripts for both experiments and the joint analysis are available online and can be accessed at https://osf.io/vz89f/.

Manipulation Checks

Ratings of Cyberball Experience

We compared mean ratings on the social inclusion scale and received passes for each experimental condition. Participants in the *exclusion* condition felt significantly less included (Experiment 1a: M_{ex} = 2.40; M_{in} = 3.38, $t[69] = 6.27$, $p < .001$; $d = 1.49$; Experiment 1b: $M_{ex} = 2.33$; $M_{in} = 3.24$, $t[55] = 5.25$, $p < .001$; $d = 1.39$) and reported that they received fewer passes (Experiment 1a: M_{ex} = 1.41; $M_{in} = 2.69$, $t[69] = 8.25$, $p < .001$; $d = 1.96$); Experiment 1b: $M_{ex} = 1.66$; $M_{in} = 2.72$, $t[56] = 5.93$, $p < .001$; $d = 1.56$) than participants in the *inclusion* condition. Thus, the manipulation was successful for both experiments.

Ratings of Experimental Situation and Interaction Partner in the Observational SR Binding Task

We computed mean ratings of participants' perception of the experimental situation and their interaction partner for both experimental conditions. Experimental conditions did not differ for ratings of perceived (dis)comfort in the experimental situation (Experiment 1a: M_{ex} = 5.14; M_{in} = 4.79, $t[70] = 1.64$, $p = .105$; $d = 0.39$; Experiment 1b: $M_{ex} = 4.73$; $M_{in} = 4.96$, $|t| < 1$, $p = .390$; $d = 0.23$), perceived agreeableness of interaction partners (Experiment 1a: $M_{ex} = 5.79$; $M_{in} = 5.91$, $|t| < 1$, $p = .507$; $d = 0.16$; Experiment 1b: M_{ex} = 5.60; M_{in} = 5.80, $|t| < 1$, $p = .348$; $d = 0.25$), interaction style (Experiment 1a: M_{ex} = 3.88; M_{in} = 3.53, $|t| < 1$, $p = .400$, $d = 0.20$; Experiment 1b: M_{ex} = 4.21; M_{in} = 3.86, $|t| < 1$, $p = .469$, $d = 0.19$), or degree of acquaintance (Experiment 1a: $M_{ex} = 1.31$; $M_{in} = 1.39$, $|t| < 1$, $p = .658$, $d = 0.10$; Experiment 1b: M_{ex} = 1.41; M_{in} = 1.38, $|t| < 1$, $p = .870$, $d = 0.04$).

Memory Test Performance in the Observational SR Binding Task

We compared memory test performance (mean error rates) between participants of both experimental conditions. Importantly, error rates did not differ between both experimental conditions (Experiment 1a: M_{ex} = 5.5%; $M_{in} = 7.1\%$, $t[70] = 1.29$, $p = .198$; $d = 0.31$; Experiment 1b: $M_{ex} = 5.2\%$; $M_{in} = 4.6\%$, $|t| < 1$, $p = .565$; $d = 0.15$). We are thus safe to conclude that participants of both experimental conditions attended to and memorized observed responses equally well. This finding is important as it argues against the view that socially excluded participants attend less to others (Costantini & Ferri, 2013).

Performance in the Observational SR Binding Task

We only analyzed ($n - 1 \rightarrow n$) trial sequences that were not intervened by a memory test trial (75% of all data). For each participant, only *actor* trials are analyzed as a function of the immediately preceding (observation) trial. Furthermore, only release reaction times (RTs) for correct responses on trial$_n$ that were preceded by observing a correct response on trial$_{n-1}$ were considered. A small percentage (Experiment 1a: 1.1%; Experiment 1b: 1.6%) of all trials were excluded due to

Table 1. *M* (*SD*) Release RT (in ms) for the factorial design of Experiments 1a and 1b, and the joint analysis of data from both experiments

		Experimental condition			
Experiment		Social exclusion *n* = 36 Response compatibility		Social inclusion *n* = 36 Response compatibility	
1a	Stimulus relation	C	IC	C	IC
	SR	508 (61)	513 (56)	516	511
	SC	516 (62)	506 (58)	521	513
	ΔSC − SR	8 [3.9]	−7 [2.4]	5 [4.2]	2 [2.9]
	S × R	15 [4.8]	3 [5.3]		
1b		Social exclusion *n* = 29 Response compatibility		Social inclusion *n* = 29 Response compatibility	
		C	IC	C	IC
	SR	525 (69)	518 (67)	530 (61)	527 (59)
	SC	531 (69)	519 (66)	528 (63)	522 (54)
	ΔSC − SR	6 [3.8]	1 [3.8]	−2 [4.7]	−5 [4.8]
	S × R	5 [4.6]		3 [6.8]	
Joint analysis		Social exclusion *n* = 65 Response compatibility		Social inclusion *n* = 65 Response compatibility	
		C	IC	C	IC
	SR	516 (65)	515 (65)	522 (58)	519 (55)
	SC	523 (66)	512 (61)	524 (56)	517 (56)
	ΔSC − SR	7 [2.7]	−3 [2.2]	2 [3.1]	−2 [2.7]
	S × R	10 [3.7]		4 [4.2]	

Note. C = compatible, IC = incompatible, RT = reaction time, SC = stimulus change, SR = stimulus repetition.
ΔSC − SR = stimulus repetition effects, computed as difference of stimulus repetition minus stimulus change. *S* × R = interaction effect between stimulus relation and response compatibility, computed as the difference between stimulus repetition effects for compatible minus stimulus repetition effects for incompatible response, that is, $(SC\text{-}SR)_C - (SC\text{-}SR)_{IC}$. Positive values indicate interaction effects that conform with expected effects due to retrieval of observational SR bindings (i.e., positive stimulus repetition effects for trial sequences in which required responses are compatible with observed responses; negative stimulus repetition effects for trial sequences in which required responses are incompatible with observed responses). SEs of the means are in square brackets.

errors. Furthermore, trials were also excluded if participants gave a wrong response in the subsequent memory test trial (Experiment 1a: 6.3%; overall: 2.1%; Experiment 1b: 4.9%; overall: 1.6%). Third, outliers[3] on release RTs were excluded (Experiment 1a: 0.9%; Experiment 1b: 0.3%).

Mean release RTs for the factorial design are presented in Table 1. For each experiment, we first computed effect scores for each participant representing the stimulus relation × response compatibility interaction, which indicates stimulus-based retrieval of observed responses (see

[3] Release RT below 300 ms or more than three interquartile ranges above the 75th percentile of the individual RT distribution were regarded as outliers (Tukey, 1977).

Table 1 for details on effect computation). Effect scores were then analyzed as a function of experimental group in one-sided, independent *t*-tests.[4] For Experiment 1a, this difference was not significant, $t(70) = 1.53$, $p = .065$, $d = 0.36$, although the stimulus relation × response compatibility interaction effect scores were descriptively larger for participants in the *social exclusion* condition than for participants in the *social inclusion* condition ($S \times R_{ex} = 15$ ms; $S \times R_{in} = 3$ ms). For Experiment 1b, the difference in interaction effect scores did not differ between experimental groups either, $t(56) = 0.14$, $p = .887$, $d = 0.04$, again although the stimulus relation × response compatibility interaction was descriptively larger for participants in the *social exclusion* condition ($S \times R_{ex} = 5$ ms; $S \times R_{in} = 3$ ms).

Joint Analyses of Experiments 1a and 1b

The previous analyses indicate that both experiments showed data trends in the expected direction, with larger observational SR binding and retrieval effects after social exclusion compared to the inclusion condition. These descriptive patterns, however, failed to reach significance. To exclude the possibility that this was simply a consequence of insufficient power of each single experiment, we ran a joint analysis. $S \times R$ interaction effect scores of both experiments were entered to a 2 (experimental condition: social exclusion vs. inclusion) × Experiment (1a vs. 1b) factorial ANOVA. Although the stimulus relation × response compatibility interaction effect scores were descriptively larger for participants in the *social exclusion* condition than for participants in the *social inclusion* condition also in the joint analysis ($S \times R_{ex} = 10$ ms; $S \times R_{in} = 4$ ms; cf. Table 1), this difference (reflected in the statistical test for the "experimental condition" main effect) was not significant, $F(1,128) = 1.40$, $p = .239$. No other effect was significant (all Fs < 1).

The nonsignificant finding might be a consequence of a surprisingly small effect. To test whether the difference in $S \times R$ interaction effect scores between both experimental conditions was significantly smaller than would be considered meaningful, we first needed to define the smallest ES of interest (SESOI): We performed a mini meta-analysis on all published studies on the observational SR binding paradigm from our laboratory ($k = 3$; Giesen et al., 2014, 2016, 2018), which yielded a mean ES of $d = 0.49$, 95% CI (0.26, 0.73), for social modulatory effects on retrieval of observationally acquired SR bindings. Following the recommendations by Lakens et al. (2018, p. 262), we chose the lower bound of the confidence interval of the meta-analytic ES estimate as SESOI. Given the directional prediction of our study, we ran an inferiority test[5] with the TOSTER package for *R* to test whether the observed group differences in SR retrieval effects obtained in the joint analysis is significantly smaller than the chosen SESOI of $d = 0.26$. The inferiority test was nonsignificant, $t(126.35) = -0.297$, $p = .383$, and is therefore uninformative. In other words, we cannot rule out that the obtained effect is significantly smaller than $d = 0.26$. In addition, we ran a sensitivity analysis with G*Power to determine which ES could be detected with the joint analysis, given equal sample size of $n = 65$ per experimental group, with a statistical power $1 - \beta = .80$, $\alpha = .05$, in a directional (i.e., one-tailed) test. According to this analysis, the experiment was sufficiently powered to detect ESs of $d = 0.43$ and larger but was not sufficiently powered to detect smaller ESs. In concert, results by both inferiority test and sensitivity analysis suggest that the effect of the manipulation was possibly too small to be detected with the present study.

Results: Exploratory Analyses

One explanation for the obtained null-findings is the possibility that the manipulation vanished over time, meaning that the impact of the social exclusion manipulation is only visible for the experimental trials for the first,

[4] Note that the independent *t*-test on $S \times R$ interaction effect scores is mathematically equivalent to the test of a three-way interaction in a 2 (stimulus relation) × 2 (response compatibility) × 2 (experimental condition) mixed models ANOVA, with $t^2 = F$ with df = 1. In addition, we want to inform the reader that for Experiment 1, the one-tailed *t*-test was not explicitly mentioned in the pre-registration of Experiment 1. However, in the descriptions of the hypothesis (see pre-registration, point 2), it says "Triggered by the experience of social exclusion, the participant should show an increased tendency to imitate the behavior of the interacting partner. (. . .) *In contrast, retrieval of observationally acquired S-R bindings should be reduced or even absent for participants in a control condition who were not socially excluded*" (emphasis added). The directional nature of our hypothesis (present SR retrieval after experience of social exclusion, reduced or absent SR retrieval after social inclusion) is thus explicitly stated in the pre-registration. We also want to point out that the pre-registration for Experiment 2 explicitly mentions the use of one-tailed *t*-tests to test the directional hypothesis.

[5] We want to thank an anonymous reviewer for suggesting such an analysis.

Table 2. *M* (*SD*) Release RT (in ms) for block effects for the factorial design on the joint data set

		Experimental condition			
		Social exclusion *n* = 65 Response compatibility		Social inclusion *n* = 65 Response compatibility	
Block	Stimulus relation	C	IC	C	IC
1	SR	518 (66)	521 (66)	528 (64)	527 (61)
	SC	531 (66)	515 (66)	531 (61)	523 (66)
	ΔSC − SR	13 [3.6]	−6 [3.3]	3 [4.3]	−4 [3.9]
	S × R	19 [4.9]		7 [4.5]	
2	SR	514 (68)	509 (59)	516 (56)	511 (53)
	SC	515 (68)	509 (62)	517 (56)	511 (53)
	ΔSC − SR	1 [3.4]	0 [3.1]	1 [3.8]	0 [3.3]
	S × *R*	1 [5.0]		1 [5.2]	

Note. C = compatible, IC = incompatible, RT = reaction time, SC = stimulus change, SR = stimulus repetition.
ΔSC-SR = stimulus repetition effects, computed as difference of stimulus repetition minus stimulus change. *S* × *R* = interaction effect between stimulus relation and response compatibility (see Table 1 for computation details). SEs of the means are in square brackets.

but not the second half of the experiment.[6] To explore this, we ran an exploratory 2 (experimental condition: social exclusion vs. inclusion) × 2 (block: first vs. second half) mixed models ANOVA on interaction effect scores from the joint data set (see Table 2 for means of the factorial design). This analysis showed that interaction effect scores differed significantly from zero (reflected in a significant intercept), $F(1,128) = 5.95$, $p = .016$, $\eta_p^2 = .04$, which reflects the typical SR retrieval pattern (see Table 2). Furthermore, there was a main effect of block, $F(1,128) = 6.72$, $p = .011$, $\eta_p^2 = .05$. Follow-up analyses showed that averaged across the experimental condition factor, $S \times R$ interaction effect scores differed significantly from zero only for the first block ($S \times R_{Block1} = 13$ ms), $t(129) = 3.43$, $p < .001$, $d_z = .30$, but not for the second block ($S \times R_{Block2} = 0$ ms), $|t| < 1$, $p = .830$, $d_z = .02$. Also, stimulus relation × response compatibility interaction effect scores were larger for participants in the *social exclusion* condition than for participants in the *social inclusion* condition in the first block ($S \times R_{ex} = 19$ ms; $S \times R_{in} = 7$ ms), $t(128) = 1.75$, $p = .041$ (one-tailed), $d = .31$, but did not differ statistically in the second block ($S \times R_{ex} = 1$ ms; $S \times R_{in} = 1$ ms; cf. Table 2), $|t| < 1$, $p = .990$, $d = .00$. However, this trend was only descriptive, as neither the main effect of experimental condition nor the experimental condition × block interaction was significant, with $F(1,128) = 1.34$, $p = .250$, $\eta_p^2 = .01$, and $F(1,128) = 1.97$, $p = .163$, $\eta_p^2 = .02$, respectively.

Discussion

In the present study, participants played a virtual ball-tossing game to experimentally induce social exclusion or inclusion experiences and then engaged in the observational SR binding task. We reasoned that stimulus-based retrieval effects of observationally acquired SR bindings would be stronger in socially excluded, compared with socially included, participants. For both experiments, manipulation checks showed that the induction of social exclusion or inclusion experiences was successful. However, the manipulation did not have a robust impact on our dependent measures of interest. Although the descriptive patterns of stimulus-based retrieval effects were in the direction of our prediction, the critical comparison just failed statistical significance in Experiment 1a but was virtually absent in Experiment 1b. Even a joint analysis with both experiments combined did not yield the expected difference in SR retrieval effects as a function of experimental condition.

There are several reasons that might account for the absence of statistical effects. First, although the effect of being socially excluded in the *Cyberball* game was still measurable and led to the expected pattern of significant differences between both experimental conditions when manipulation checks were collected at the end of the experiment, there is reason to believe that the impact of the social exclusion manipulation quickly vanished over time. However, an exploratory analysis of block effects remains inconclusive in this regard, given that the critical interaction with the block factor was not significant. Thus, we can only speculate on possible reasons for the present null findings. Tentatively, assuming that the manipulation did not last throughout the entire experimental session would offer an explanation why SR retrieval is absent during the second half of the experiment, since participants worked independently of each other. Working independently is a condition under which retrieval of observational SR bindings typically does not occur (cf. Giesen et al., 2014; 2018 for similar null findings). Put this way, the experience of social exclusion could have transiently bridged the independence between pairs of participants to satisfy deprived social affiliation needs and to socially reconnect with others. However, future research is needed to follow up on these speculations. In our view, this *post hoc* insight into the transient nature of after-effects of social exclusion in the *Cyberball* game is nevertheless informative for future researchers who aim to use this as a manipulation.

[6] We thank Oliver Genschow for suggesting this explanation and analyses.

Second, the present experiments lack a neutral comparison condition to which one could contrast after-effects of social exclusion versus inclusion experiences on stimulus-based retrieval of observational SR bindings. Such a true baseline would allow us to estimate whether social inclusion already increases stimulus-based retrieval of observational SR bindings. Although this reasoning is somewhat challenged by the fact that $S \times R$ interaction effect scores for socially included participants were only small and seldomly differed from zero (see Table 1), previous studies show that the experience of being socially included makes automatic imitative behavior more likely (e.g., Gleibs et al., 2016; Yabar et al., 2006). Similarly, feelings of rejection or isolation might have counteracted or neutralized the effect of social exclusion on the desire for inclusion, which would lead to reduced observational binding and retrieval effects in the exclusion condition. However, future studies are necessary to address these questions. Ideally, one should then aim to model these neutralizing forces in full extent, for instance, by using separate measures for experienced level of inclusion and the motivation to socially reintegrate with others.

Third, another possible explanation for the absence of effects in the standard analyses is the lack of sufficient statistical power. A sensitivity analysis with G*Power revealed that with the sample size of the joint analysis, our study had sufficient power ($1 - \beta = .80$) to detect ESs of $d \geq .43$, corresponding to medium sized (and larger) effects, but was not sufficiently powered to detect smaller effects. In sum, the post hoc exploratory analysis of block effects, sensitivity analysis, and inferiority test suggest that the effect of the manipulation was possibly too small to be detected with the present study.

Fourth, yet another possibility for the present null findings could be that social exclusion simply does *not* influence imitative action regulation processes as measured with RT-based compatibility tasks. This would imply that predicted effects will not show up, even in larger samples. Indeed, such a finding would represent an important dissociation between imitative effects measured with behavioral mimicry versus stimulus compatibility tasks and would provide further evidence against the idea that both measures tap into the same underlying mechanisms (Genschow et al., 2017).

Implications

The results have important implications for the current debate on whether effects of automatic imitation can be socially modulated. For instance, it is not yet clear whether imitative tendencies as measured with the *automatic imitation* task can be socially modulated at all (Cracco, Genschow et al., 2018; Cracco, Bardi et al., 2018; Cracco & Brass, 2019; Genschow et al., 2017; Ramsey, 2018). However, this task is responsive to manipulations of self versus other focus (Genschow et al., 2019) as well as social content manipulation of imitated gestures (Cracco, Genschow et al., 2018; Cracco, Bardi et al., 2018). Effects were small, though, and thus may emerge only in larger samples. Apart from the present study, the observational SR binding paradigm proved to be responsive to a range of social moderators so far (e.g., situational and chronic interdependency; vicarious reinforcement; Giesen et al., 2014, 2016, 2018). However, it is difficult to draw inferences from the present juxtaposition, particularly because different social moderators were addressed for each paradigm. Some of these moderators may be more easily manipulated and/or have longer-lasting effects, which increases the probability of affecting dependent measures in the expected way. Future research should therefore expose different measures of automatic imitation to the very same social manipulations to draw valid conclusions from the cross-paradigm comparison. After all, the degree of a paradigm's sensitivity to social modulations is not a question that is answered with a simply study, but rather a quest in the service of construct validation.

Of course, our study is not the first to deal with the absence of a social modulation of imitative action regulation effects (e.g., Newey et al., 2019; see also Cracco, Genschow et al., 2018; Cracco, Bardi et al., 2018, for a meta-analysis on social moderators of automatic imitation). Still, we want to emphasize that reporting the present null findings is important in several respects. First, by proving preregistrations and running an exact replication, we put our own research under criteria of strong scientific rigor. Reporting also nonsignificant results will hopefully help to reduce the *file-drawer problem* (Rosenthal, 1979), which – unfortunately – is still an immanent problem in many areas of psychological research. Second, our findings are important for researchers interested in effects of social exclusion on automatic behavior regulation and will provide them with more adequate estimates for to-be-expected ESs that is not conflated as a consequence of publication biases (Ferguson & Heene, 2012). Finally, our results do contain interesting cues for further research in this domain. Specifically, trying to separate and disentangle effects of experienced exclusion from the desire to be included seems to be a promising avenue for investigating and understanding effects of social rejection on automatic tendencies of social imitation.

References

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human

motivation. *Psychological Bulletin*, *117*(3), 497–529. https://doi.org/10.1037/0033-2909.117.3.497
Blackhart, G. C., Nelson, B. C., Knowles, M. L., & Baumeister, R. F. (2009). Rejection elicits emotional reactions but neither causes immediate distress nor lowers self-esteem: A meta-analytic review of 192 studies on social exclusion. *Personality and Social Psychology Review*, *13*(4), 269–309. https://doi.org/10.1177/1088868309346065
Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychologica*, *106*(1–2), 3–22. https://doi.org/10.1016/S0001-6918(00)00024-X
Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*(6), 893–910. https://doi.org/10.1037/0022-3514.76.6.893
Chartrand, T. L., & Lakin, J. L. (2013). The antecedents and consequences of human behavioral mimicry. *Annual Review of Psychology*, *64*(1), 285–308. https://doi.org/10.1146/annurev-psych-113011-143754
Costantini, M., & Ferri, F. (2013). Action co-representation and social exclusion. *Experimental Brain Research*, *227*(1), 85–92. https://doi.org/10.1007/s00221-013-3487-3
Cracco, E., Bardi, L., Desmet, C., Genschow, O., Rigoni, D., De Coster, L., Radkoval., Deschrijver, E., & Brass, M. (2018). Automatic imitation: A meta-analysis. *Psychological Bulletin*, *144*(5), 453–500. https://doi.org/10.1037/bul0000143
Cracco, E., & Brass, M. (2019). Reaction time indices of automatic imitation measure imitative response tendencies. *Consciousness and Cognition: An International Journal*, *68*, 115–118. https://doi.org/10.1016/j.concog.2019.01.001
Cracco, E., Genschow, O., Radkova, I., & Brass, M. (2018). Automatic imitation of pro- and antisocial gestures: Is implicit social behavior censored? *Cognition*, *170*(5), 179–189. https://doi.org/10.1016/j.cognition.2017.09.019
Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146
Ferguson, C. J., & Heene, M. (2012). A vast graveyard of undead theories: Publication bias and psychological science's aversion to the null. *Perspectives on Psychological Science*, *7*(6), 555–561. https://doi.org/10.1177/1745691612459059
Genschow, O., Schuler, J., Cracco, E., Brass, M., & Wänke, M. (2019). The effect of money priming on self-focus in the imitation-inhibition task: A registered report. *Experimental Psychology*, *66*(6), 423–436. https://doi.org/10.1027/1618-3169/a000466
Genschow, O., van Den Bossche, Cracco, E., Bardi, L., Rigoni, D., & Brass, M. (2017). Mimicry and automatic imitation are not correlated. *PLoS One*, *12*(9), e0183784. https://doi.org/10.1037/t51181-000
Giesen, C., Herrmann, J., & Rothermund, K. (2014). Copying competitors? Interdependency modulates stimulus-based retrieval of observed responses. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(5), 1978–1991. https://doi.org/10.1037/a0037614
Giesen, C., Scherdin, K., & Rothermund, K. (2016). Flexible goal imitation: Vicarious feedback influences stimulus-response binding by observation. *Learning & Behavior*, *45*(2), 147–156. https://doi.org/10.3758/s13420-016-0250-1
Giesen, C., Löhl, V., Rothermund, K., & Koranyi, N. (2018). Intimacy effects on action regulation: Retrieval of observationally acquired stimulus-response bindings in romantically involved interaction partners versus strangers. *Frontiers in Psychology*, *9*, 1369. https://doi.org/10.3389/fpsyg.2018.01369
Gleibs, I. H., Wilson, N., Reddy, G., & Catmur, C. (2016). Group dynamics in automatic imitation. *PLoS One*, *11*(9), e0162880. https://doi.org/10.1371/journal.pone.0162880
Hess, Y. D., & Pickett, C. L. (2010). Social rejection and self- versus other-awareness. *Journal of Experimental Social Psychology*, *46*(2), 453–456. https://doi.org/10.1016/j.jesp.2009.12.004
Heyes, C. (2011). Automatic imitation. *Psychological Bulletin*, *137*(3), 463–483. https://doi.org/10.1037/a0022288
Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, *1*(2), 259–269. https://doi.org/10.1177/2515245918770963
Lakin, J. L., Chartrand, T. L., & Arkin, R. M. (2008). I am too just like you: Nonconscious mimicry as an automatic behavioral response to social exclusion. *Psychological Science*, *19*(8), 816–822. https://doi.org/10.1111/j.1467-9280.2008.02162.x
Maner, J. K., DeWall, C. N., Baumeister, R. F., & Schaller, M. (2007). Does social exclusion motivate interpersonal reconnection? Resolving the 'porcupine problem. *Journal of Personality and Social Psychology*, *92*(1), 42–55. https://doi.org/10.1037/0022-3514.92.1.42
Newey, R., Koldewyn, K., & Ramsey, R. (2019). The influence of prosocial priming on visual perspective taking and automatic imitation. *PLoS One*, *14*(1), e0198867. https://doi.org/10.1371/journal.pone.0198867
Over, H., & Carpenter, M. (2009). Priming third-party ostracism increases affiliative imitation in children. *Developmental Science*, *12*(3), F1–F8. https://doi.org/10.1111/j.1467-7687.2008.00820.x
Ramsey, R. (2018). What are reaction time indices of automatic imitation measuring? *Consciousness and Cognition: An International Journal*, *65*, 240–254. https://doi.org/10.1016/j.concog.2018.08.006
Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *86*(3), 638–641. https://doi.org/10.1037/0033-2909.86.3.638
Tajfel, H., & Turner, J. C. (2004). The social identity theory of intergroup behavior. In J. T. Jost & J. Sidanius (Eds.), *Key readings in social psychology. Political psychology: Key readings* (pp. 276–293). Psychology Press. https://doi.org/10.4324/9780203505984-16
Tukey, J. W. (1977). *Exploratory data analysis*. Addison-Wesley.
Twenge, J. M., Baumeister, R. F., DeWall, C. N., Ciarocco, N. J., & Bartels, J. M. (2007). Social exclusion decreases prosocial behavior. *Journal of Personality and Social Psychology*, *92*(1), 56–66. https://doi.org/10.1037/0022-3514.92.1.56
Wilkowski, B. M., Robinson, M. D., & Friesen, C. K. (2009). Gaze-triggered orienting as a tool of the belongingness self-regulation system. *Psychological Science*, *20*(4), 495–501. https://doi.org/10.1111/j.1467-9280.2009.02321.x
Williams, K. D. (2007). Ostracism. *Annual Review of Psychology*, *58*(1), 425–452. https://doi.org/10.1146/annurev.psych.58.110405.085641
Yabar, Y., Johnston, L., Miles, L., & Peace, V. (2006). Implicit behavioral mimicry: Investigating the impact of group membership. *Journal of Nonverbal Behavior*, *30*(3), 97–113. https://doi.org/10.1007/s10919-006-0010-6
Zadro, L., Williams, K. D., & Richardson, R. (2004). How low can you go? Ostracism by a computer is sufficient to lower self-reported levels of belonging, control, self-esteem, and meaningful existence. *Journal of Experimental Social Psychology*, *40*(4), 560–567. https://doi.org/10.1016/j.jesp.2003.11.006

History
Received November 3, 2020
Revision received June 18, 2021
Accepted June 23, 2021
Published online August 10, 2021

Acknowledgments
We thank Nils Meier for his help in programming the experiments. Data were collected as part of Laura Nagel's and Matthäus Rudolphs bachelor theses.

Publication Ethics
Both experiments were in accordance with the Declaration of Helsinki. Ethical approval was granted by the Ethical Commission of the Friedrich Schiller University Jena (Experiment 1a: FSV 19/15; Experiment 1b: FSV 20/001).

Open Data
Raw data and analysis scripts for both experiments and the joint analysis are available online and can be accessed at https://osf.io/vz89f/. The exact method, design, hypotheses, data preparation, and planned analyses were pre-registered online at https://www.aspredicted.org.

Funding
Open Access publication enabled by Friedrich Schiller University Jena.

ORCID
Carina G. Giesen
https://orcid.org/0000-0002-2395-4435

Carina G. Giesen
Department of Psychology, General Psychology II
Friedrich Schiller University Jena
Am Steiger 3, Haus 1
07743 Jena
Germany
carina.giesen@uni-jena.de