



Revisiting intersubjective action-effect binding: No evidence for social moderators

Eva Riechelmann¹ · Lisa Weller¹ · Lynn Huestegge¹ · Anne Böckler¹ · Roland Pfister¹

Published online: 22 March 2019
© The Psychonomic Society, Inc. 2019

Abstract

Effect-based accounts of human action control have recently highlighted the possibility of representing one's own actions in terms of anticipated changes in the behavior of social interaction partners. In contrast to action effects that pertain to the agent's body or the agent's physical environment, social action effects have been proposed to come with peculiarities inherent to their social nature. Here, we revisit the currently most prominent demonstration of such a peculiarity: the role of eye contact for action-effect learning in social contexts (Sato & Itakura, 2013, *Cognition*, 127, 383–390). In contrast to the previous demonstration of action-effect learning, a conceptual and a direct replication both yielded evidence for the absence of action-effect learning in the proposed design, irrespective of eye contact. Bayesian statistics supported this claim by demonstrating evidence in favor of the null hypothesis of no effect. These results suggest a limited generalizability of the original findings—for example, due to limitations that are inherent in the proposed study design or due to cultural differences.

Keywords Action control · Social actions · Effect anticipations

Anticipating how other people will react to one's own actions is a fundamental part of action control in social interactions. Another person's predictable behavior can even be used to represent, select, and control one's own movements. Such *sociomotor* actions have recently attracted considerable interest (e.g., Flach, Press, Badets, & Heyes, 2010; Kunde, Lozo, & Neumann, 2011; Müller, 2016; Pfister, Dignath, Hommel, & Kunde, 2013; Weller, Schwarz, Kunde, & Pfister, 2018). The sociomotor framework (Kunde, Weller, & Pfister, 2018) is rooted in ideomotor approaches to action control, which propose that human actions are controlled by anticipating the sensory consequences they typically evoke (Harleß, 1861; Herbart, 1825; James, 1890; see Pfister, 2019; Shin, Proctor, & Capaldi, 2010, for reviews). This idea takes into account that each action inevitably produces a range of sensory effects—for instance, knocking on a table results in a specific sound, a visual image of the respective motion, and proprioceptive sensations

of the movement. Ideomotor theory proposes that agents acquire bidirectional associations between a motor action and these sensory effects (i.e., action–effect associations). These associations can be used for action control: In order to perform a specific action, its sensory effects are anticipated, which in turn trigger the associated motor action. There is considerable empirical support for the assumptions of ideomotor theory—for example, from the manual domain (Elsner & Hommel, 2001; Kunde, 2001, 2006; Pfeuffer, Kiesel, & Huestegge, 2016; Pfister & Kunde, 2013; Wolfensteller & Ruge, 2011) and from the oculomotor domain (Herwig & Horstmann, 2011; Huestegge & Kreutzfeldt, 2012; Riechelmann, Pieczykolan, Horstmann, Herwig, & Huestegge, 2017).

The term *sociomotor actions* specifically refers to situations where the action of an agent not only triggers certain sensory effects in the inanimate environment but also consistently evokes a certain behavior of another person (Kunde et al., 2018). Based on ideomotor theory, it is assumed that the agent can acquire a bidirectional association between his or her action and the sensory effect—that is, the other person's response (also labeled “intersubjective action–effect binding”; Sato & Itakura, 2013). Anticipating the other person's behavior then can reactivate the agent's action. Support for this claim comes from several studies investigating learning and anticipation of social action effects using different experimental designs (e.g., Herwig &

Eva Riechelmann and Lisa Weller contributed equally to this work.

✉ Eva Riechelmann
eva.riechelmann@uni-wuerzburg.de

¹ Department of Psychology, Würzburg University, Röntgenring 11, 97070 Würzburg, Germany

Horstmann, 2011; Kunde et al., 2011; Müller, 2016; Müller & Jung, 2018; Pfister et al., 2013; Pfister, Weller, Dignath, & Kunde, 2017; Sato & Itakura, 2013).

Previous work on sociomotor actions has mainly aimed at showing that anticipations of a partner's behavior are indeed implemented in action control (Kunde et al., 2018; Pfister et al., 2013; see also Wolpert, Doya, & Kawato, 2003, for a related framework). Although the available evidence clearly supports this claim, only few studies have addressed possible peculiarities of social action effects as compared with effects on the agent's body or on the agent's inanimate environment. A notable exception to this rule is a study by Sato and Itakura (2013), who investigated action–effect learning for social action effects as well as social moderator variables in this process. In this study, participants first underwent an acquisition phase in which they repeatedly experienced novel action–effect associations between key presses and a certain mouth gesture of an on-screen face. More precisely, participants could choose between pressing a left or right key on each trial, and their key press consistently triggered a certain change of the on-screen face—for example, lip protrusion or cheek puffing. In a subsequent test phase, participants had to respond with the same left and right key press to novel, arbitrary targets, and the former effect stimuli (i.e., mouth gestures) were presented as primes before the imperative stimuli. These mouth gestures could either be congruent to the to-be-executed response (i.e., the imperative stimulus required the response that had produced the respective mouth gesture in the acquisition phase), incongruent (i.e., to-be-executed and prime-associated responses did not match), or neutral (i.e., inducing a mouth gesture that had not been experienced in the acquisition phase). The authors argued that if participants had acquired bidirectional associations between the key presses and the effect stimuli in the acquisition phase, presenting the effect stimuli as primes in the test phase should activate the associated response. Therefore, responses should be facilitated when the congruent (vs. incongruent) prime is presented.

The results of Experiment 1 in Sato and Itakura (2013) indeed confirmed that participants responded faster in trials involving congruent (vs. incongruent) primes. However, to validate whether the results of Experiment 1 are driven by genuinely social processes as elicited by eye contact (direct gaze of the face stimuli), the authors conducted Experiment 2 where the on-screen face was looking to the left or right instead of directly looking at the participant. Indeed, Experiment 2 revealed no evidence for action–effect learning with averted gaze. These results suggest that action–effect learning and/or retrieval in sociomotor actions can be modulated by genuinely social variables such as eye contact, which is in line with previous research showing that direct eye contact is a powerful moderator of cognitive processes (Senju & Johnson, 2009). The social significance of direct eye contact is further confirmed by studies showing that faces with direct (vs. averted) gaze capture the attention of the perceiver (Böckler,

van der Wel, & Welsh, 2014; Senju & Hasegawa, 2005; Senju, Hasegawa, & Tojo, 2005) and facilitate processing of the observed face (Macrae, Hood, Milne, Rowe, & Mason, 2002; Mason, Hood, & Macrae, 2004).

Although eye contact exerts a strong influence on cognitive processes in a range of different domains, the results of Experiments 1 and 2 in Sato and Itakura's (2013) study can also be explained along different lines. More precisely, the absence of evidence for action–effect learning in case of averted gaze may also be attributed to a lack of attention on the mouth region, where the critical action-contingent changes occurred. Specifically, the face stimuli with averted gaze might have prompted participants to automatically follow this gaze, drawing attention away from the mouth region (see Frischen, Bayliss, & Tipper, 2007, for a review on gaze cueing). Sato and Itakura (2013) designed face stimuli with closed eyes (Experiment 3) to tackle this alternative explanation. Exploring such alternative explanations seems especially warranted in light of studies that showed an impact of social action effects even in the absence of eye contact (e.g., Flach et al., 2010; Pfister et al., 2013; Weller et al., 2018). Because there was still no evidence for action–effect learning in this setting, the authors concluded that the congruency effect as observed in the direct gaze condition (Experiment 1) was indeed due to the eye contact.

However, one might still consider attentional processes as a potential explanation for the null results of their Experiment 3. We argue that visual attention could have been drawn away from the face, causing a lack of attention on the critical mouth region even in Experiment 3, because a nonengaging interaction partner with closed eyes served as stimulus. Several findings—for example, from neonate studies—support this claim by demonstrating that newborns already spend less time looking at a face photograph with the eyes closed compared with the same face photograph with eyes open (Batki, Baron-Cohen, Wheelwright, Connellan, & Ahluwalia, 2000).

Based on this reasoning, the present study was designed to replicate Sato and Itakura's (2013) study and to additionally test whether gaze cueing toward the mouth region (thereby drawing attention to the location of the effect) can reinstantiate a congruency effect even in the absence of direct eye contact. This finding would substantially challenge the idea that eye contact represents a necessary prerequisite for sociomotor learning. In Experiment 1a, we conducted a conceptual replication of Sato and Itakura's (2013) Experiment 1 using photographs with direct gaze as stimuli. Experiment 1b addressed the alternative explanation proposed above by including face stimuli with eyes looking downward in order to guide the participant's attention to the mouth region of the face stimulus (Experiment 1b; for evidence for gaze cuing on the vertical axis, see Langton & Bruce, 1999).

We did not observe evidence for any action–effect learning with either direct or averted gaze in these initial experiments.

Given that our stimulus material was different from the stimuli used in the original study by Sato and Itakura (2013), we then decided to run a direct replication including the original stimulus material (Experiment 2).

Experiment 1a: Conceptual replication

Experiment 1a represents a conceptual replication of Experiment 1 reported by Sato and Itakura (2013). We made every effort to produce stimulus material that matched the stimuli used in the original study, while adopting the stimuli to the different cultural backgrounds of European rather than Asian participants.

As in the original study, participants first experienced novel action–effect associations in an acquisition phase. More precisely, they were instructed to perform left and right key presses that triggered distinct mouth gestures of the face presented on the screen. Participants were asked to spontaneously select each key press while choosing each option about equally often. In a subsequent test phase, corresponding faces were presented as primes shortly before an imperative stimulus, which required a speeded left or right response. If associations between key press and resulting mouth gesture were learned in the acquisition phase, the primes presented in the test phase should activate the prime-associated response, thereby influencing response initiation. Thus, in congruent trials, where the prime-associated response and the to-be-executed response matched, we expected to observe reduced reaction times (RTs) compared with incongruent trials, where prime-associated and to-be-executed response did not match.

Method

Participants

We recruited 32 participants who received either course credits or monetary compensation for participation. Participants were naïve with respect to the purpose of the experiment and gave written informed consent before completing the study.

In contrast to Sato and Itakura's (2013) original study, where no participants were excluded due to an unbalanced proportion of key presses during acquisition, we chose to exclude participants from our initial analyses when the distribution of left and right key presses during acquisition deviated from a balanced distribution at a ratio equal to or exceeding 2:1. Data of eight participants had to be excluded due to this criterion. Data of the remaining 24 participants were analyzed (mean age = 24.6 years, age range: 19–34 years, 20 women, no left-handers). A sample size of 24 participants should ensure a high power of $1 - \beta > .99$ to detect the original effect

size of Cohen's $d_z = t / \sqrt{n} = 4.7 / \sqrt{22} = 1.00$, as reported for Sato and Itakura's (2013) Experiment 1.

Apparatus and stimuli

The experiment was programmed using E-Prime 2.0 (Psychology Software Tools Inc., Sharpsburg, PA, USA), and stimuli were presented on a 23-in. TFT-monitor (refresh rate: 60 Hz, spatial resolution: 1,920 × 1,080 pixels). Participants responded on a standard computer keyboard using a left and right response key with their respective left and right index finger.

Stimuli were designed to be maximally comparable with the stimulus set used in Sato and Itakura (2013). We therefore used four color photographs of one forward-facing Caucasian female face ($7.6^\circ \times 5.7^\circ$, height × width), which differed only with respect to the mouth gestures displayed. The mouth gestures were inserted into the same face pictures (using photo-editing software) to maximize control; all gestures corresponded to the variations used in the original study: mouth closed, lip protrusion, tongue protrusion, and cheek puffing. The faces were cropped to an oval shape (see the Appendix for the complete stimulus set). Stimulus material and experimental program are available on the Open Science Framework (<https://osf.io/z2dw5/>).

Procedure

The procedure of Experiment 1a closely matched the procedure of Sato and Itakura's (2013) Experiment 1, and mainly differed with respect to the stimulus material used (see the Apparatus and Stimuli section for details). As in the original study, the experiment comprised an acquisition phase and a test phase (see Fig. 1).

Acquisition phase Each trial of the acquisition phase started with the central presentation of a fixation cross (1,000 ms), which was then substituted by the female face stimulus shown with the mouth closed. Participants were instructed to respond to this neutral target face with a left or right key press using the left or right index finger, respectively. They were further told to spontaneously select each key press, and to press each key about equally often. Each key press was followed by a black screen (presented for 50 ms; for a similar action–effect delay, see Dignath, Pfister, Eder, Kiesel, & Kunde, 2014; Elsner & Hommel, 2001; Hoffmann, Lenhard, Sebal, & Pfister, 2009). After that, the target face reappeared for 300 ms in the form of an action effect where the mouth gesture had changed to either lip protrusion, tongue protrusion, or cheeks puffing. Importantly, the change in mouth gesture was dependent on the selected key press: the left key press always triggered a specific effect (e.g., lip protrusion) whereas the right key press always triggered a different effect (e.g., cheeks puffing). The

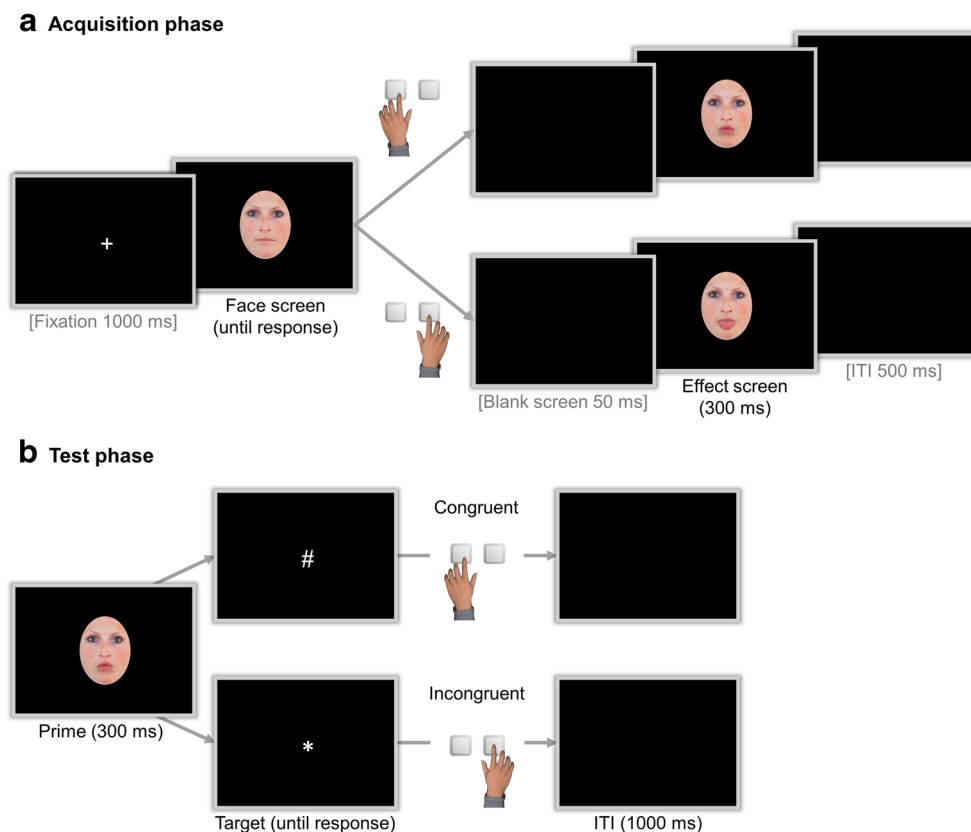


Fig. 1 Schematic depiction of the acquisition (**a**) and test (**b**) phase. **a** After the presentation of a fixation screen, the fixation cross was replaced by a neutral target face requiring the participant to choose between a left or right key press on each trial. After a delay of 50 ms, and contingent upon each key press, one of two target faces with a specific mouth gesture was displayed. The next trial started after an intertrial interval (ITI) of 500 ms. Note that the acquisition phases of Experiments 1a–b and Experiment 2 slightly differed with respect to the timing. There was no fixation screen and ITI in Experiment 2, and the neutral target face was presented on-

screen during the 50-ms delay between key press and effect onset. Thus, the face was constantly presented in Experiment 2 (as in Sato & Itakura, 2013), whereas face presentation was interrupted in Experiment 1a–b. **b** A prime face was followed by a target, which required a left or right key press in accordance with prior instructions. The prime face was either congruent, incongruent, or neutral with respect to the congruency of the prime-associated response (as acquired during acquisition) and to-be-executed response in the test phase. The next trial started after an ITI of 1,000 ms

assignment of mouth gestures to response keys was constant for each participant and counterbalanced across participants. Thus, two out of the three mouth gestures were presented to one participant. The response–effect mapping was not mentioned to the participants, but it was pointed out that the mouth gestures were completely irrelevant for the task. The next trial started 500 ms after effect offset. The acquisition phase consisted of 300 trials in total. After completing the acquisition phase, participants had the opportunity to take a break before the test phase started.

Note that we implemented several minor changes to the original study in the acquisition procedure. In the original study, the key press triggered a change in mouth gesture after a delay of 50 ms without any interruption by a black screen. Further, the trial timing was slightly different. Although there was a fixation interval of 1,000 ms at the beginning and a black screen (presented for 500 ms) at the end of each trial in our study, the original study did not use any such interval, so that the target face was presented on the screen throughout the acquisition phase.

Test phase In the test phase, the previous effect stimuli served as primes. At the beginning of each trial, one effect prime with lip protrusion, tongue protrusion, or cheeks puffing was presented centrally for a duration of 300 ms. Although each participant was familiar with two of the effect primes from the preceding acquisition phase, there was always one unfamiliar prime which had not been presented before. Following the presentation of the prime, one of two target stimuli “*” (0.8×0.8 cm) or “#” (1.0×0.8 cm) was presented at the center of the screen with the instruction to respond to each target with a key press according to a fixed target–response mapping that was instructed at the beginning of the test phase. The target–response mapping was counterbalanced across participants. Participants were instructed to ignore the prime stimulus and to respond as quickly and accurately as possible to the target stimulus. The test phase comprised 120 trials. As in the original study, the test phase comprised 40 trials with congruent and 40 trials with incongruent primes. In another 40 trials, the neutral stimuli with the mouth gesture not presented during acquisition (i.e., without any possibility to acquire

associations to the two response options) served as primes. The next trial started 1,000 ms after the key press.

Design and analysis

The experiment involved the within-subjects factor congruency (congruent vs. incongruent vs. neutral prime) referring to the congruency of the prime-associated response and to-be-executed response in the test phase.

We conducted two types of analyses. First, we performed the exact same analyses as reported for Experiment 1 in Sato and Itakura (2013)—that is, paired-samples t tests to compare error rates and response times between the congruent and incongruent condition while omitting the data of the neutral condition. We report Cohen's d_z as effect sizes for paired-samples t tests (calculated as $d_z = \frac{t}{\sqrt{n}}$). Second, the analyses were extended to include the neutral condition by performing repeated-measures analyses of variances (ANOVAs) with the factor congruency (congruent vs. incongruent vs. neutral) for error rates and RTs. Because the original Sato and Itakura (2013) study did not report any exclusion criteria due to an unbalanced proportion of key presses during acquisition, we additionally report t tests and ANOVAs with all participants included. Error trials were removed prior to analyzing RTs. For violations of the sphericity assumption, we report Greenhouse–Geisser corrected p values along with original degrees of freedom. As in the original study, all following analyses were performed without outlier correction.

Besides traditional null-hypothesis significance testing, we additionally drew on Bayesian statistics for a better interpretation of nonsignificant results. We calculated nondirectional Bayes factors (BF_{01}) using the BayesFactor package Version 0.9.12-2 of the R software environment Version 3.3.2, with a value of 1 as scale parameter for the prior distribution. BF_{01} was computed as $f(\text{data} | H_0) / f(\text{data} | H_1)$, with f denoting marginal likelihoods. We interpreted $BF_{01} > 3$ as evidence for the null hypothesis and $BF_{01} < 1/3$ as evidence for the alternative hypothesis.¹

Results and discussion

Acquisition phase

On average, participants responded 387 ms ($SD = 114$ ms) after presentation of the face stimulus. Descriptively, the distribution of left (52.03%) and right (47.97%) key presses was close to the instructed balanced distribution, even though the statistical comparison indicated a small effect for this comparison, $t(23) = 2.19$, $p = .039$, $d = 0.45$. We ensured that all participants included in the analyses had experienced each

key-effect mapping in sufficient quantity (see Participants section for the exclusion criterion).

Test phase

The mean error rate was 4.38% ($SD = 3.32$) for the congruent and 4.17% ($SD = 3.19$) for the incongruent condition, and a paired-samples t test indicated no significant difference between the two conditions, $t(23) = 0.36$, $p = .723$, $d = 0.07$, $BF_{01} = 5.99$. RTs for correct responses did not differ between the congruent ($M = 448$ ms, $SD = 38.8$) and the incongruent condition ($M = 444$ ms, $SD = 36.5$), $t(23) = 1.00$, $p = .328$, $d = 0.20$, $BF_{01} = 3.97$ (see Fig. 2). Similarly, the repeated-measures ANOVA including the neutral condition was nonsignificant for both error rates and RTs (both F s < 1 ; see Table 1). When all participants were included in the analyses, there were still no significant differences between conditions, neither in error rates or RTs (p s $\geq .165$) nor in the corresponding repeated-measures ANOVAs (both F s < 1).

To sum up, the results of Experiment 1a suggest that participants' behavior was not influenced by the congruency manipulation as implemented in the test phase. Moreover, the analysis of BF_{01} provided clear evidence for the absence of any congruency effect. The observed pattern of results is at odds with the original observations of Sato and Itakura (2013), especially when considering that the present design should come with high power to detect the previously reported effect size. In the light of these results, we ran Experiment 1b, which featured face stimuli with eyes gazing toward the mouth region. This gaze direction might direct visual attention of the participant to the mouth region of the face (see Langton & Bruce, 1999), eventually boosting the build-up of intersubjective action–effect binding.

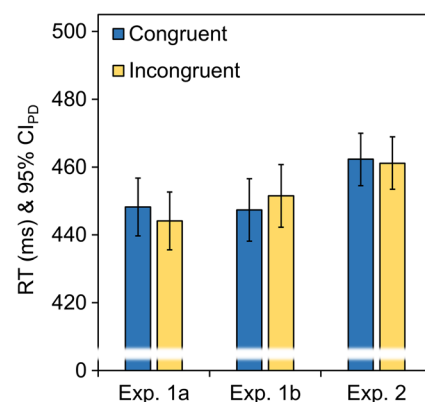


Fig. 2 Mean response times (RT) for congruent and incongruent trials in Experiments 1a, b, and 2. Error bars indicate the 95% confidence interval of paired differences (CI_{pd}) for the comparison of congruent and incongruent RTs, calculated separately for each experiment (cf. Pfister & Janczyk, 2013)

¹ We thank the editor for stimulating these additional analyses.

Table 1 Mean response times (RTs) and error rates for Experiments 1a, b, and 2

Congruency		Experiment 1a		Experiment 1b		Experiment 2	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
RTs (ms)	Congruent	448	38.8	447	32.4	462	56.2
	Incongruent	444	36.5	451	34.5	461	57.0
	Neutral	449	48.3	450	31.7	460	57.0
Error rates (%)	Congruent	4.38	3.32	4.17	3.27	2.92	4.21
	Incongruent	4.17	3.19	3.33	3.51	2.62	3.78
	Neutral	3.85	3.68	3.54	4.42	4.27	3.57

Note. Mean (*M*) and standard deviation (*SD*) are depicted for RTs and error rates of congruent, incongruent, and neutral trials in Experiments 1a, b, and 2

Experiment 1b: Modified conceptual replication

In Experiment 1b, we used the same face photographs of the female individual from Experiment 1a, but now the eyes of the on-screen face were always gazing downward instead of looking directly at the participant (see Appendix). By doing so, visual attention of the participant was directed toward the crucial action effect location, that is the mouth region of the on-screen face (Langton & Bruce, 1999). Again, we expected to observe reduced RTs in congruent trials, where prime-associated and to-be-executed response matched, compared with incongruent trials, where prime-associated and to-be-executed response did not match.

Method

Participants

We tested 27 new participants who received either course credits or monetary compensation for participation. The data of two participants were excluded from analysis due to extreme deviation from the instructed balanced distribution of left and right key presses during acquisition (see Participants section of Experiment 1a for details regarding the exclusion criterion). Another participant was excluded from analysis due to extremely high average RTs ($>$ sample mean + 2 *SD*s). Data of the remaining 24 participants were analyzed (mean age = 25.8 years, age range: 19–49 years, 16 women, three left-handers). Participants were naïve with respect to the purpose of the experiment and gave written informed consent before completing the study.

Apparatus, stimuli, and procedure

Technical equipment and procedure were identical to Experiment 1a. The only difference was the stimulus material used: While eye gaze was always directed toward the

participants in Experiment 1a, the eyes of the face were always looking downward in Experiment 1b (see Appendix).

Design and analysis

Design and analyses were identical to Experiment 1a.

Results and discussion

Acquisition phase

On average, participants responded 325 ms (*SD* = 91 ms) after presentation of the face stimulus. Descriptively, the distribution of left (51.74%) and right (48.26%) key presses was close to the instructed balanced distribution, even though the statistical comparison, $t(23) = 2.39$, $p = .025$, $d = 0.49$, indicated nonequality. However, by excluding participants with extreme deviations from balanced distribution (see Participants section of Experiment 1a for the exclusion criterion), we ensured that each key-effect mapping was experienced sufficiently often.

Test phase

The mean error rate amounted to 4.17% (*SD* = 3.27) for congruent and to 3.33% (*SD* = 3.51) for incongruent conditions, and a t test indicated no significant difference between conditions, $t(23) = 1.07$, $p = .295$, $d = 0.22$, $BF_{01} = 3.71$. Response times did not differ between the congruent ($M = 447$ ms, *SD* = 32.4) and the incongruent condition ($M = 451$ ms, *SD* = 34.5) either, $t(23) = 0.93$, $p = .362$, $d = 0.19$, $BF_{01} = 4.22$ (see Fig. 2). Likewise, the repeated-measures ANOVA did not yield any significant results for error rates or RTs (both F s < 1 ; see Table 1). Including all participants into analyses again did not yield any significant differences between conditions with respect to errors rates and RTs (p s $\geq .192$), as well as with respect to the corresponding repeated-measures ANOVAs (both F s < 1).

These results mirrored the findings of Experiment 1a by showing no evidence for action–effect learning despite using stimuli with the potential to support a shift of attention toward

the mouth region. In light of the procedural differences between the present Experiment 1 and the original design of Sato and Itakura (2013; see Procedure section of Experiment 1a for details regarding the differences), we therefore opted to conduct a direct replication.

As described in the Method section, Experiments 1a–b included a short blank screen between action and effect, whereas the target face was presented on the screen throughout this delay in the original study. At first sight, this aspect of the procedure might suggest that our failure to replicate the original findings was due to change blindness (e.g., Pashler, 1988; Wilford & Wells, 2010). Change blindness occurs when a blank screen separates two flickering images in a change-detection task, and it becomes apparent in terms of reduced change-detection accuracy, especially at short interstimulus intervals. However, some important aspects make it unlikely that change blindness occurred in the present design. Pashler (1988) used pure white or black-and-white checkerboard squares (mask condition) as opposed to a black display (no-mask condition). Given this definition, the blank interval in our design rather resembles the no-mask (instead of the mask) condition of the Pashler (1988) study. In combination with the short interstimulus interval in our design (50 ms), our design is comparable with the experimental condition in which Pashler actually observed best change-detection performance. Furthermore, change blindness predominantly occurs for unexpected changes and is reduced for items that receive preferential attention within a visual composition (see Simons & Rensink, 2005, for a review). Given that the mouth is an important and preferentially attended feature within a face, and that the action-contingent change in our study predictably occurred at the mouth region, we are confident that participants in our study were able to perceive the change and consider it unlikely that change blindness was a confounding factor. Still, in order to parallelize this aspect of the design with the original procedure, we removed the blank interval in Experiment 2 in order to control for a potential confound and to replicate every minute detail of the original Sato and Itakura (2013) study.

Experiment 2: Direct replication

Experiment 2 was a close, preregistered replication of Sato and Itakura's (2013) Experiment 1 that adhered to minute details of the original setup and employed the original stimulus material.²

² We thank Atsushi Sato, who kindly provided the photographs from the original study.

Method

Participants

Another 24 healthy participants were recruited and received monetary compensation for participation (mean age 29.3 years, age range: 20–66 years, 18 women, three left-handers). For one participant, the proportion of left and right key presses (31.44% left vs. 68.56% right) exceeded the range of tolerance as defined in Experiments 1a–b. Because the original study did not report any exclusion criteria due to an unbalanced proportion of key presses during acquisition, we included all participants into the analysis. Participants were naïve with respect to the purpose of the experiment and gave written informed consent before completing the study.

Apparatus and stimuli

Stimuli were presented on a 24-in. monitor (refresh rate: 100 Hz, spatial resolution: 1,920 × 1,080 pixels), and participants responded by using a left and right response key with their left and right index fingers on a standard computer keyboard. The stimulus material was identical to the one used in Sato and Itakura (2013). Stimuli were four photographs (6.9° × 4.5°) of a single, forward-facing female individual with eyes directed at the observer. The faces were cropped with an oval shape, removing all surrounding features. The stimuli only differed with respect to the mouth gestures depicting either a closed mouth, lip protrusion, tongue protrusion, or cheeks puffing. In Experiment 2, the size of the target stimuli amounted to 0.8 cm × 0.8 cm for “*,” and to 1.0 cm × 0.8 cm for “#.”

Procedure, design, and analysis

The procedure of the following experiment matched the procedure of Sato and Itakura's (2013) Experiment 1, with the only exception being that participants underwent a short practice phase consisting of four exemplary trials before the acquisition phase started. Experiment 2 differed from Experiments 1a–b with respect to some minor timing aspects, while closely mirroring the procedure as described in Sato and Itakura (2013): While face presentation was interrupted in the acquisition trials of Experiments 1a–b (see Fig. 1 and the Procedure section of Experiment 1a for details), the face was constantly presented during acquisition in Experiment 2.

To analyze the data, we first conducted the exact same analysis as reported in Sato and Itakura's (2013) study—that is, paired-samples *t* tests for error rates and RTs, with all participants included. In a second step, we then extended the original analyses to match the procedure of Experiments 1a–b. That is, we additionally conducted ANOVAs to also include the neutral condition, and repeated both types of analyses when excluding participants with an unbalanced proportion

of left and right key presses during acquisition. Note that Sato and Itakura's (2013) original study did not apply any outlier corrections. We therefore did not perform any outlier correction for our initial analyses, but validated these results against outlier-corrected analyses of the RTs by excluding RTs that deviated more than 2.5 standard deviations from the corresponding cell mean (computed separately for all participants and conditions).

Results and discussion

Acquisition phase

Mean response time in the acquisition phase amounted to 706 ms ($SD = 275.6$ ms). The distribution of left (48.82%) and right (51.18%) key presses was close to the instructed balanced distribution, $t(23) = 1.38$, $p = .182$, $d = 0.28$.

Test phase

The mean error rate amounted to 2.92% ($SD = 4.21$) for the congruent and to 2.62% ($SD = 3.78$) for the incongruent condition. A paired-samples t test indicated no significant differences between the two conditions, $t(23) = 0.34$, $p = .737$, $d = 0.07$, $BF_{01} = 6.03$. The analysis of RTs for correct responses did not yield any significant differences between the congruent ($M = 462$ ms, $SD = 56.2$) and incongruent condition ($M = 461$ ms, $SD = 57.0$), $t(23) = 0.29$, $p = .775$, $d = 0.06$, $BF_{01} = 6.12$ (see Fig. 2). The repeated-measures ANOVA including the neutral condition as an additional factor level was nonsignificant for both error rates, $F(2, 46) = 1.87$, $p = .166$, $\eta_p^2 = .08$, and RTs ($F < 1$; see Table 1). Note that the pattern of results did not change when omitting the data from the participant with unbalanced left and right key presses in the acquisition phase (see Participants section for details). When excluding this participant, we did not observe significant differences between conditions with respect to error rates, $t(22) = 0.22$, $p = .827$, $d = 0.05$, $BF_{01} = 6.11$, and RTs, $t(22) = 0.57$, $p = .575$, $d = 0.12$, $BF_{01} = 5.35$. The corresponding repeated-measures ANOVAs also did not yield any significant effects for error rates, $F(2, 44) = 1.99$, $p = .149$, $\eta_p^2 = .08$, and RTs ($F < 1$).

Applying an outlier correction to the RT data of correct responses did not change the overall result pattern. The paired-samples t test showed that RTs did not significantly differ between the congruent ($M = 455$ ms, $SD = 53.7$) and incongruent condition ($M = 452$ ms, $SD = 53.7$), $t(23) = 1.02$, $p = .319$, $d = 0.21$, $BF_{01} = 3.89$, and also the repeated-measures ANOVA yielded nonsignificant results ($F < 1$).

Between-experiments analysis

To assess evidence for action–effect learning across the full set of experiments, we conducted a pooled analysis running a two-way split-plot ANOVA, with congruency (congruent vs. incongruent) as a within-subjects factor and experiment (1a vs. 1b vs. 2) as between-subjects factor for RTs of the test phase. On average, participants responded within 452 ms ($SD = 5.1$) in the congruent condition, and within 452 ms ($SD = 5.2$) in the incongruent condition, yielding a nonsignificant effect of congruency ($F < 1$).³ RTs were not significantly different between experiments ($F < 1$). The interaction was not significant, $F(1, 69) = 1.04$, $p = .360$, $\eta_p^2 = .03$. Pooling the data of all three experiments resulted in a congruency effect of 0.35 ms (calculated as the difference between the congruent and the incongruent condition) with a 95% CI of $[-4.32, 5.02]$. This corresponded to a Cohen's d_z of $d_z = 0.02$, with a 95% CI for standardized means of $[-0.21, 0.25]$ and a BF of 10.67. The data of the three experiments combined thus clearly suggest a negligible effect size for the congruency effect as a measure of action–effect retrieval in the present design, whereas the Bayes factor indicates strong support in favor of the null hypothesis.

General discussion

The present experiments revisited action–effect learning for sociomotor actions—that is, actions that aim at triggering predictable responses of a social partner (Kunde et al., 2018). In particular, we employed a study design proposed by Sato and Itakura (2013) that had previously yielded evidence for eye contact as a social moderator of action–effect learning when participants' key-press actions triggered contingent changes of the mouth gesture displayed by a face stimulus.

Whereas the original study had been conducted in Japan, and thus used stimulus pictures of an Asian female face, we used a Caucasian female face as a stimulus along with a German sample for two conceptual replications. Experiment 1a aimed at replicating evidence for intersubjective action–effect learning when eyes were directed at the participant (i.e., Experiment 1 of the original study), whereas Experiment 1b aimed at testing whether action–effect learning for social action effects occurs when the participants' attention is directed toward the location of the action effect by a

³ Note that the pooled analysis comes with a power of $1 - \beta > .95$ even for effects of half the magnitude that was reported in the original study ($d_z = 1.00 \leftrightarrow d_z/2 = 0.50$).

corresponding gaze of the face stimulus. In both experiments, we did not find any evidence for the buildup of action–effect associations as measured in error rates and response times. To rule out that our nonsignificant results can be tracked back to our stimulus material, Experiment 2 featured a direct replication of Sato and Itakura's (2013) Experiment 1, but with Caucasian instead of Asian participants. This experiment also provided no evidence for action–effect learning. Using Bayesian statistics, we found substantial evidence in favor of the null hypothesis. Moreover, pooling the data of all three experiments revealed that the numerical difference between the congruent and incongruent condition was smaller than 1 ms, yielding strong support in favor of the null hypothesis.

Our findings indicate that action–effect learning as studied by Sato and Itakura's (2013) paradigm does not necessarily generalize to other samples of participants. What could be the reasons for the marked differences between the two sets of findings? We suggest that two accounts seem plausible and require further empirical elaboration. First, it could be that both findings—the present null effect and the sizeable effect of the original study—are reliable. According to this view, the difference in results would point to manifest cultural differences between the samples. Second, based on five reported null effects in the literature (the present three experiments as well as Experiments 2 and 3 of the original study), it might be the case that the significant effect observed in Sato and Itakura's (2013) Experiment 1 represents a statistical Type I error. This would suggest that the experimental paradigm as applied in the present experiments and in Sato and Itakura's (2013) experiments might not be suitable to study action–effect learning. In the following, we will discuss both possibilities, starting with the latter proposal.⁴

When assuming that the present experimental design is unsuited to study action–effect learning, the consistent lack of between-condition differences begs the question of whether action–effect learning between key-press actions and following social effects did not take place at all, or whether learning did occur but failed to show in the test phase (Pfister, 2019; Pfister, Kiesel, & Hoffmann, 2011). Critically examining the experimental design seems to locate the issue in the test phase of the original design by Sato and Itakura (2013), especially because of how congruency between target and effect-associated actions was manipulated. More precisely, a task-irrelevant face stimulus preceded the target stimulus

in the test phase, which, according to common ideomotor logic, should prime the associated responses if this facial expression had been triggered by a specific action during the acquisition phase. This within-subjects manipulation is clearly different from many other acquisition-phase/test-phase designs (e.g., Elsner & Hommel, 2001), where congruency is typically manipulated between subjects. Such acquisition-phase/test-phase designs often entail that participants respond either in an acquisition-consistent or in an acquisition-inconsistent mapping to the effect stimuli of the preceding acquisition phase, requiring to attend to the previous effect stimuli (see also Beckers, De Houwer, & Eelen, 2002; Hoffmann et al., 2009). The use of task-irrelevant primes in within designs, by contrast, likely reduces the chance of the prime stimuli to retrieve the associated action, as participants do not necessarily have to attend to the prime. In addition, primes differed only subtly from each other in the stimulus set of the present Experiments 1a–b as well as in the original stimulus set of Sato and Itakura's (2013) study (e.g., compare lip protrusion with cheeks puffing; see Appendix for the stimuli used in Experiment 1a–b, and see Sato & Itakura, 2013, for the stimuli used in Experiment 2). Thus, if participants did not fully attend to the prime stimuli, these subtle differences between primes may not have been recognized, and strong congruency effects may have been precluded. Furthermore, the procedure came with a fixed stimulus-onset asynchrony of 300 ms between prime and target. Even though the time course of response activation due to irrelevant effect primes is not known, it seems possible that activation had already decayed by the time the participants began to plan their response to the target (for positive results with an interval of less than 100 ms between prime and target onset, see Kunde, 2004). Increasing the discriminability of the prime stimuli and presenting primes in closer temporal proximity to the target might thus be a promising way to adjust the proposed within-subjects design (Eder, Rothermund, De Houwer, & Hommel, 2015; Elsner & Hommel, 2004; Müller, 2016; Müller & Jung, 2018; Wolfensteller & Ruge, 2011). But note that the abovementioned limitations are not due to the use of social face stimuli as primes. We rather believe that the failure to find evidence for action–effect learning could represent limitations inherent to the design irrespective of a social (vs. nonsocial) context.

In contrast to the interpretation discussed so far, the difference between the present findings and the findings of Sato and Itakura (2013) could also reflect cultural influences, as predominantly individualistic societies (e.g., German culture) and predominantly collective societies (e.g., Japanese culture) have been shown to differ

⁴ A third tangible interpretation could even suggest that the phenomenon of action–effect learning does not exist for sociomotor actions. We consider this possibility is unlikely due to the robust evidence from different experimental paradigms that has been reported across a range of independent groups (e.g., Flach et al., 2010; Herwig & Horstmann, 2011; Kunde et al., 2011; Müller, 2016; Pfister et al., 2013).

qualitatively in a vast range of cognitive processes and processing styles (Markus & Kitayama, 1991; Nisbett, 2004; Way & Lieberman, 2010). Most relevant for the present discussion, several studies suggest that Japanese culture is different from Western European/North American culture with respect to gaze behavior, gaze processing, and eye contact. For instance, maintaining eye contact during conversation is perceived as attentive and polite in Western culture, while gaze avoidance reflects respectful behavior in Eastern cultures (Argyle, Henderson, Bond, Iizuka, & Contarello, 1986). Additionally, several studies suggest cultural differences in eye movements during the processing of faces whereby Western Caucasian participants predominantly fixated the eye region, together with frequent fixations on the mouth region, showing a triangular fixation pattern. In contrast, the fixation pattern of East Asian participants was biased toward the central face region around the nose, thereby avoiding direct eye contact with the face image (e.g., Blais, Jack, Scheepers, Fiset, & Caldara, 2008). These findings suggest that direct gaze might have a strong and lasting impact on East Asian participants, whereas Western participants might not be as sensitive to this social cue. In line with this speculation, Senju et al. (2013) reported a stronger tendency for gaze following behavior in Japanese participants as compared with British participants. However, Senju et al. (2013) also studied gaze behavior in British and Japanese participants when observing dynamic avatar faces and found Japanese participants to fixate on the eye region more often than British participants, which contradicts the presumed tendency of East Asians to fixate the face center as reported previously (Blais et al., 2008). A similar increased focus on the eye region in East Asian participants was also identified by Jack, Blais, Scheepers, Schyns, and Caldara (2009). In sum, previous studies provided mixed evidence for the hypothesis that Japanese individuals are more sensitive to eye contact than Western individuals are. Based on the heterogeneity of the present database, we believe that further work is needed to corroborate any possible effects of culture on the processing of social action effects. Such further work would be especially informative if it measured eye movements and fixation patterns in the presented experimental design to compare these measures between participants from diverging cultural backgrounds.

It is noteworthy that our samples and the sample of the original study differed with respect to response speed during acquisition. In Experiments 1a and 1b, our participants responded after 356 ms, as opposed to

roughly 1 second in the original study. This difference might be attributed to differences in trial timing, since the acquisition phase in our experiments included a fixation period and an intertrial interval, whereas the sequence of trials was a continuous process in the original study. Thus, the slower RTs in the original study might reflect the time necessary to process the preceding trial. This process has likely taken place during intertrial interval and fixation in our Experiments 1a–b, allowing for faster responses. In Experiment 2, where we implemented the same trial timing as in the original study, we still observed faster response times (706 ms) in the acquisition phase as compared with the original study. Please note, however, that during the test phase, both the German and the Japanese sample responded at nearly the same speed (approx. 450 ms). This is important because meaningful test phase differences in RTs between both samples could have affected the retrieval of action–effect associations differently.

Conclusion

In three experiments, we did not observe any evidence of action–effect learning in a social context using a design that followed the suggestions of Sato and Itakura (2013). The question of whether the diverging results of the present experiments and the original study point towards a Type I error in the original work, or whether this difference could point toward cultural differences, remains to be explored in future work. At the same time, we believe that following the general approach of Sato and Itakura’s work—the dedicated study of peculiarities of social action effects relative to action effects in the inanimate environment—is a highly promising avenue to further inform our understanding of sociomotor actions.

Acknowledgements We thank Atsushi Sato for providing us with the stimulus material of Sato and Itakura (2013), and for stimulating discussions regarding the present findings. We would like to thank Charlotte Erlinghagen and André Michael Interthal for data collection.

Funding This research was funded by grants of the German Research Foundation to A.B. (GZ: BO4962/1-1), L.H. (HU 1847/7-1), and R.P. (PF 853, 2-1).

Compliance with ethical standards

Conflicts of interest The authors declare that there are no conflicts of interest.

Appendix

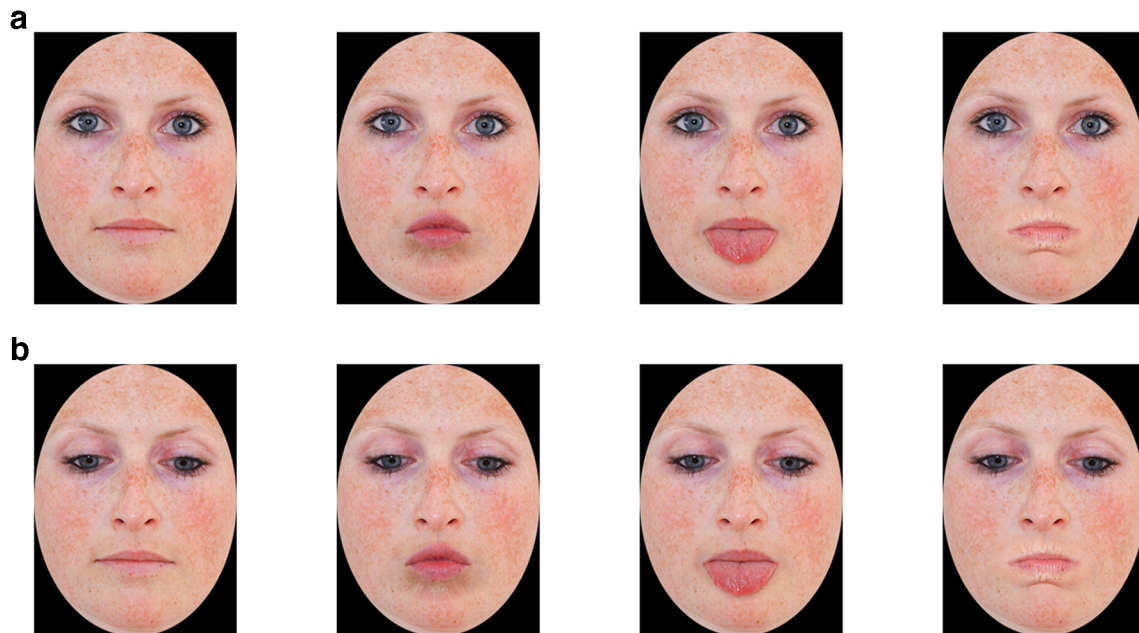


Fig. 3 Stimuli used in (a) Experiment 1a (direct gaze) and (b) Experiment 1b (mouth gaze) showing different mouth gestures (from left to right): mouth closed, lip protrusion, tongue protrusion, and cheeks puffing

References

- Argyle, M., Henderson, M., Bond, M., Iizuka, Y., & Contarello, A. (1986). Cross-cultural variations in relationship rules. *International Journal of Psychology*, 21, 287–315. <https://doi.org/10.1080/00207598608247591>
- Batki, A., Baron-Cohen, S., Wheelwright, S., Connellan, J., & Ahluwalia, J. (2000). Is there an innate gaze module? Evidence from human neonates. *Infant Behavior and Development*, 23, 223–229. [https://doi.org/10.1016/S0163-6383\(01\)00037-6](https://doi.org/10.1016/S0163-6383(01)00037-6)
- Beckers, T., De Houwer, J., & Eelen, P. (2002). Automatic integration of non-perceptual action effect features: the case of the associative affective Simon effect. *Psychological Research*, 66, 166–173. <https://doi.org/10.1007/s00426-002-0090-9>
- Blais, C., Jack, R. E., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture shapes how we look at faces. *PLOS ONE*, 3, e3022. <https://doi.org/10.1371/journal.pone.0003022>
- Böckler, A., van der Wel, R. P. R. D., & Welsh, T. N. (2014). Catching eyes: Effects of social and nonsocial cues on attention capture. *Psychological Science*, 25, 720–727. <https://doi.org/10.1177/0956797613516147>
- Dignath, D., Pfister, R., Eder, A. B., Kiesel, A., & Kunde, W. (2014). Representing the hyphen in action-effect associations: automatic acquisition and bidirectional retrieval of action-effect intervals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40, 1701–1712. <https://doi.org/10.1037/xlm0000022>
- Eder, A. B., Rothermund, K., De Houwer, J., & Hommel, B. (2015). Directive and incentive functions of affective action consequences: An ideomotor approach. *Psychological Research*, 79, 630–649. <https://doi.org/10.1007/s00426-014-0590-4>
- Elsner, B., & Hommel, B. (2001). Effect anticipation and action control. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 229–240. <https://doi.org/10.1037/0096-1523.27.1.229>
- Elsner, B., & Hommel, B. (2004). Contiguity and contingency in action-effect learning. *Psychological Research*, 68, 138–154. <https://doi.org/10.1007/s00426-003-0151-8>
- Flach, R., Press, C., Badets, A., & Heyes, C. (2010). Shaking hands: Priming by social action effects. *British Journal of Psychology*, 101, 739–749. <https://doi.org/10.1348/000712609X484595>
- Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin*, 133, 694–724. <https://doi.org/10.1037/0033-2909.133.4.694>
- Harleß, E. (1861). Der Apparat des Willens [The apparatus of the will]. *Zeitschrift für Philosophie und philosophische Kritik*, 38, 50–73.
- Herbart, J. F. (1825). Psychologie als Wissenschaft neu gegründet auf Erfahrung, Metaphysik, und Mathematik. [Psychology as a science newly founded on experience, metaphysics, and mathematics]. Königsberg, Germany: Unzer.
- Hewig, A., & Horstmann, G. (2011). Action-effect associations revealed by eye movements. *Psychonomic Bulletin & Review*, 18, 531–537. <https://doi.org/10.3758/s13423-011-0063-3>
- Hoffmann, J., Lenhard, A., Sebal, A., & Pfister, R. (2009). Movements or targets: What makes an action in action-effect learning? *Quarterly*

- Journal of Experimental Psychology* (2006), 62, 2433–2449. <https://doi.org/10.1080/17470210902922079>
- Huestegge, L., & Kreutzfeldt, M. (2012). Action effects in saccade control. *Psychonomic Bulletin & Review*, 19, 198–203. <https://doi.org/10.3758/s13423-011-0215-5>
- Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G., & Caldara, R. (2009). Cultural confusions show that facial expressions are not universal. *Current Biology*, 19, 1543–1548. <https://doi.org/10.1016/j.cub.2009.07.051>
- James, W. (1890). *The principles of psychology*. New York, NY: Henry Holt.
- Kunde, W. (2001). Response-effect compatibility in manual choice reaction tasks. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 387–394. <https://doi.org/10.1037/0096-1523.27.2.387>
- Kunde, W. (2004). Response priming by supraliminal and subliminal action effects. *Psychological Research*, 68, 91–96.
- Kunde, W. (2006). Antezedente Effektrepräsentationen in der Verhaltenssteuerung [Antecedent effect representations in behavior control]. *Psychologische Rundschau*, 57, 34–42.
- Kunde, W., Lozo, L., & Neumann, R. (2011). Effect-based control of facial expressions: Evidence from action-effect compatibility. *Psychonomic Bulletin & Review*, 18, 820–826. <https://doi.org/10.3758/s13423-011-0093-x>
- Kunde, W., Weller, L., & Pfister, R. (2018). Sociomotor action control. *Psychonomic Bulletin & Review*, 25(3), 917–931. <https://doi.org/10.3758/s13423-017-1316-6>
- Langton, S. R., & Bruce, V. (1999). Reflexive visual orienting in response to the social attention of others. *Visual Cognition*, 6(5), 541–567. <https://doi.org/10.1080/135062899394939>
- Macrae, C. N., Hood, B. M., Milne, A. B., Rowe, A. C., & Mason, M. F. (2002). Are you looking at me? Eye gaze and person perception. *Psychological Science*, 13, 460–464. <https://doi.org/10.1111/1467-9280.00481>
- Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98, 224–253. <https://doi.org/10.1037/0033-295X.98.2.224>
- Mason, M. F., Hood, B. M., & Macrae, C. N. (2004). Look into my eyes: Gaze direction and person memory. *Memory*, 12, 637–643. <https://doi.org/10.1080/09658210344000152>
- Müller, R. (2016). Does the anticipation of compatible partner reactions facilitate action planning in joint tasks? *Psychological Research*, 80, 464–486. <https://doi.org/10.1007/s00426-015-0670-0>
- Müller, R., & Jung, M. L. (2018). Partner reactions and task set selection: Compatibility is more beneficial in the stronger task. *Acta Psychologica*, 185, 188–202. <https://doi.org/10.1016/j.actpsy.2018.02.012>
- Nisbett, R. (2004). *The geography of thought: How Asians and Westerners think differently... and why*. New York: The Free Press.
- Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, 44, 369–378. <https://doi.org/10.3758/BF03210419>
- Pfeuffer, C. U., Kiesel, A., & Huestegge, L. (2016). A look into the future: Spontaneous anticipatory saccades reflect processes of anticipatory action control. *Journal of Experimental Psychology: General*, 145, 1530–1547. <https://doi.org/10.1037/xge0000224>
- Pfister, R. (2019). Effect-based action control with body-related effects: Implications for empirical approaches to ideomotor action control. *Psychological Review*, 126(1), 153–161. <https://doi.org/10.1037/rev0000140>
- Pfister, R., Dignath, D., Hommel, B., & Kunde, W. (2013). It takes two to imitate: anticipation and imitation in social interaction. *Psychological Science*, 24, 2117–2121. <https://doi.org/10.1177/0956797613489139>
- Pfister, R., & Janczyk, M. (2013). Confidence intervals for two sample means: Calculation, interpretation, and a few simple rules. *Advances in Cognitive Psychology*, 9, 74–80. <https://doi.org/10.2478/v10053-008-0133-x>
- Pfister, R., Kiesel, A., & Hoffmann, J. (2011). Learning at any rate: Action-effect learning for stimulus-based actions. *Psychological Research*, 75(1), 61–65. <https://doi.org/10.1007/s00426-010-0288-1>
- Pfister, R., & Kunde, W. (2013). Dissecting the response in response-effect compatibility. *Experimental Brain Research*, 224, 647–655. <https://doi.org/10.1007/s00221-012-3343-x>
- Pfister, R., Weller, L., Dignath, D., & Kunde, W. (2017). What or when? The impact of anticipated social action effects is driven by action-effect compatibility, not delay. *Attention, Perception & Psychophysics*, 79, 2132–2142. <https://doi.org/10.3758/s13414-017-1371-0>
- Riechelmann, E., Pieczykolan, A., Horstmann, G., Herwig, A., & Huestegge, L. (2017). Spatio-temporal dynamics of action-effect associations in oculomotor control. *Acta Psychologica*, 180, 130–136. <https://doi.org/10.1016/j.actpsy.2017.09.003>
- Sato, A., & Itakura, S. (2013). Intersubjective action-effect binding: Eye contact modulates acquisition of bidirectional association between our and others' actions. *Cognition*, 127, 383–390. <https://doi.org/10.1016/j.cognition.2013.02.010>
- Senju, A., & Hasegawa, T. (2005). Direct gaze captures visuospatial attention. *Visual Cognition*, 12, 127–144. <https://doi.org/10.1080/1350628044000157>
- Senju, A., Hasegawa, T., & Tojo, Y. (2005). Does perceived direct gaze boost detection in adults and children with and without autism?: The stare-in-the-crowd effect revisited. *Visual Cognition*, 12, 1474–1496. <https://doi.org/10.1080/13506280440000797>
- Senju, A., & Johnson, M. H. (2009). The eye contact effect: Mechanisms and development. *Trends in Cognitive Sciences*, 13, 127–134. <https://doi.org/10.1016/j.tics.2008.11.009>
- Senju, A., Vemeti, A., Kikuchi, Y., Akechi, H., Hasegawa, T., & Johnson, M. H. (2013). Cultural background modulates how we look at other persons' gaze. *International Journal of Behavioral Development*, 37, 131–136. <https://doi.org/10.1177/0165025412465360>
- Shin, Y. K., Proctor, R. W., & Capaldi, E. J. (2010). A review of contemporary ideomotor theory. *Psychological Bulletin*, 136, 943–974. <https://doi.org/10.1037/a0020541>
- Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Sciences*, 9, 16–20. <https://doi.org/10.1016/j.tics.2004.11.006>
- Way, B. M., & Lieberman, M. D. (2010). Is there a genetic contribution to cultural differences? Collectivism, individualism and genetic markers of social sensitivity. *Social Cognitive and Affective Neuroscience*, 5, 203–211. <https://doi.org/10.1093/scan/nsq059>
- Weller, L., Schwarz, K. A., Kunde, W., & Pfister, R. (2018). My mistake? Enhanced error processing for commanded compared to passively observed actions. *Psychophysiology*, 55, e13057. <https://doi.org/10.1111/psyp.13057>
- Wilford, M. M., & Wells, G. L. (2010). Does facial processing prioritize change detection? Change blindness illustrates costs and benefits of holistic processing. *Psychological Science*, 21, 1611–1615. <https://doi.org/10.1177/0956797610385952>
- Wolfensteller, U., & Ruge, H. (2011). On the timescale of stimulus-based action-effect learning. *Quarterly Journal of Experimental Psychology* (2006), 64, 1273–1289. <https://doi.org/10.1080/17470218.2010.546417>
- Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, 358, 593–602. <https://doi.org/10.1098/rstb.2002.1238>