This project aims to investigate the affective consequences of aggressive behavioral outcomes.

# Study 4

## Introduction

Study 1 failed to show the predicted effect. Explorative analyses and post-experiment discussion led us to the conclusion that various methodical aspects may have contributed to this failure: first, the paradigm failed to establish a strong and experiential link between the colors that were used in the evaluative categorization measure and the actual consequences of behavior. Second, the paradigm confounded hits in the catch-the-box task with experienced losses by the partner, making it difficult to disentangle these two sources of affect. Third, the trial exclusion requirements for the evaluative categorization measure as well as its relative temporal distance from the actual affect caused by reducing the partner’s payout meant it may not have been sufficiently sensitive to capture this affect. Study 2 aimed to address these issues.

Study 2 also failed to show the predicted effect. Explorative analyses indicated the provocation may not have been successful.

For this reason, Study 3 attempted to implement a Taylor-style aggression paradigm using noise blasts instead of money splits to induce aggression. Study 3 showed the expected effect (only) in provoked participants. However, closer analysis of the data led us to the conclusion that some participants may have pressed randomly in the punishment choice task implemented independently from the Taylor task, as they may have had the perception that they could not effectively choose their target. In order to address this and to further more closely illuminate whether emotional responses to punishing aggressors are positive or simply less negative than those to punishing non-aggressors, Study 4 is planned.

To briefly recap the basic theoretical background, although successful goal-directed aggression should by definition lead to satisfaction with having achieved the goal, successful aggression may also lead to regret and negative emotional outcomes for the aggressor. In order to examine which of these two predictions holds true under which circumstances, a more fine-grained analysis of affective responses towards aggressive behavioral outcomes is required.

In particular, we postulate that goal-directed aggressive behavior is satisfying in the moment and should therefore lead to positive affect, though this positive affect may later be modulated or demolished by further self-consideration. If this is so, successfully aggressing should cause a feeling of positive affect as a direct consequence of the behavior when it is goal-directed. Such affective responses should in turn become associated with direct sensory consequences of the aggressive behavior insomuch as these direct sensory consequences are not re-evaluated through further processing.

To illustrate, punching somebody out of anger may cause spontaneous positive affect, which might for example become associated with the sensory perception of ones’ fist connecting. However, in such naturalistic cases, post-aggression rumination may reinterpret these sensory cues – they are, after all, directly connected to an expression of pain by the other person, and considering oneself the source of such pain might be negative. Therefore, the initially positive affect might flip to negative after the fact.

In a setting where the direct sensory consequences of aggressive behavior are not in turn directly connected to expressions of pain, this flipping of their affective meaning should be less likely. Furthermore, measuring this affect soon after the behavior (denying much opportunity for reinterpretation) and using an indirect measure should both also make it less likely that the affect flips. If there was a positive affective response directly after the behavior, such circumstances should be most likely to discover it.

However, the goal behind the aggressive behavior should determine whether the response is positive or not. If the aggressor is achieving a goal with the aggressive behavior (e.g. retaliation after a perceived provocation), the spontaneous response might be expected to be positive; if the aggressor has no such goal, but is aggressing accidentally or solely due to some undesired circumstance (e.g. being instructed to aggress by an experimenter), the spontaneous response should be neutral at best or more likely negative (as causing somebody else harm for no reason is generally aversive to most individuals without specific, rare personality traits).

For this reason, the current study will implement a structured interaction paradigm in which participants compete simultaneously with two fictitious partners, with the winner assigning noise blasts to one of the losers. One partner will be aggressive, often assigning the participant high-intensity noise blasts, while the other will be nonaggressive, assigning low-intensity noise blasts to each partner equally. After each response in this task, a measure of spontaneous evaluative responses will be implemented. Furthermore, participants will alternate this task with tasks in which they are either led to believe they have chosen a partner to assign a medium-intensity noise blast or they actually do assign a partner of their choice a noise blast. After each such response, a measure of spontaneous evaluative responses will be implemented. Note that this study differs from Study 4 only in details of the choice task.

## Method

### Procedure

The experiment follows a 2-cell (partner: aggressive vs. nonaggressive) within-subjects design for manipulation check variables and a 3-cell (punishment target: none vs. aggressive partner vs. nonaggressive partner) within-subjects design for the evaluative misattribution task.

Participants first provide information about themselves which is ostensibly shared with their interaction partners in order to provide an internally consistent justification for the increased aggression by one of their interaction partners. Then, participants alternate between a competitive reaction time task against their two partners in which the winner punishes a partner of their choice with noise blasts with intensity of their choice and a simple partner choice task in which they assign noise blasts with fixed intensity to one or none of their partners, each of which include evaluative misattribution tasks. Next, they answer several questionnaires and are debriefed and paid.

### Self-disclosure task

Participants are informed that some of their personal information will be anonymously shared with their interaction partners and then asked basic demographic questions (sex and age). Then, they are asked about their degree of economic socialism (7-pt Likert scale from 1 – “very liberal” to 2 – “very socialist”) and their conservatism in social matters (7-pt Likert scale from 1 – “very progressive” to 2 – “very conservative”). Finally, participants are informed that their interaction partners will be shown this information, but not a picture of the participant, while they are shown pictures of their interaction partners (two White males selected based on matched attractiveness and aggressiveness ratings in a pretest). This task is included to provide participants with a plausible reason why one of the partners may later focus aggression on them, as this partner may not have liked the responses they gave.

### Competitive reaction time task

Participants are asked to choose which partner shall receive a noise blast if they win, followed by which intensity they desire on a five-point scale. Next, participants see pictures of the interaction partners in the top left and right hand of screen and a box containing the word “ICH” (“me”) in the center top half of the screen. Below each picture and the box is a red circle. After a randomly determined interval of between 500ms and 800ms, the circles turn green and participants must respond as quickly as possible. After a further 1000ms, all circles turn back to red bar the one below the trial winner. After 400ms, a numeric intensity level appears above the selected losing participant as well as a soundwave video superimposed over the picture or box. If the participant was the selected target, a noise blast of corresponding intensity plays as well. The noise and video last for 2000ms, at which point the trial ends. If a participant responds too early or fails to respond during the 1000ms window, a message is displayed that at least one player responded too early or failed to respond respectively for 2000ms before the trial restarts (after partner and intensity choice).

Although participants are instructed that they are competing in a reaction time task against their interaction partners, in each trial, the winner is actually drawn from a randomized list. The lists are constructed in such a way that there are a total of six lists (one for each block of the task), in each of which each player wins at least eight times of the thirty trials per block, and which lead to an even split of winning trials per player over the whole experiment (sixty per player). If the participant wins, the noise blast is shown appearing on the partner of their choice with their chosen intensity. If the nonaggressive partner wins, the noise blast affects either the participant or the aggressive player with an equal distribution and a constant intensity of one (the minimum). If the aggressive partner wins, the noise blast affects the participant in 75% of the cases and the nonaggressive player in 25% of the cases with an intensity of either four or five (the maximum), randomly selected.

The assignment of the two pictures of the interaction partners to the sides of the screen as well as the assignment of the aggressive and nonaggressive roles are counterbalanced across participants.

After each trial of this task, a trial of the evaluative misattribution task follows.

### Evaluative misattribution task

At the beginning of each evaluative misattribution trial, a Chinese ideograph is presented for 100ms, followed by a blank screen for 75ms and then a mask stimulus. Participants respond by rating the masked ideograph as either pleasant or unpleasant using the left and right mouse buttons (assignment of button to evaluation counterbalanced across participants). If participants do not respond within 2000ms, a message instructing them to respond more spontaneously is displayed for 4000ms after their response. Thereafter, the next catch-the-box trial begins.

### Simple partner choice task

After each block of the competitive reaction time task, participants complete a block of the simple partner choice task. Each block of this task consists of six catch trials and twenty-one test trials in randomized order (constrained so that half of the test trials appear in the first half of the block).

In a catch trial, participants see the two pictures of their interaction partners as well as a no-shock symbol with a green frame that jumps between the three options (two partners and no shock) pictures irregularly, remaining on each option for between 550 and 750ms. Participants are instructed to press the space bar when the frame surrounds the partner to which they wish to assign a noise blast of fixed intensity (three, medium), or when it surrounds the no-shock symbol if they wish to refrain from sending a shock. When participants press the space bar, the green frame stops jumping. If a partner was selected, the intensity of the selected tone as well as the soundwave video are displayed over and on top of the selected target’s picture respectively in the same way as in the competitive reaction time task for 2000ms before the trial ends. If the no-shock symbol was selected, the green frame simply remains on the symbol for 2000ms.

Test trials are similar to catch trials, but the interval after which the frame jumps varies between 100ms and 650ms. When participants press the space bar, the frame does not stop where it is, but jumps to the aggressive partner in 33% of the trials, to the nonaggressive partner in the 33% of the trials and to the no-shock symbol in the remaining 33% of the trials. As in the catch trials, the “selected” partner then receives a noise blast for 2000ms or the frame remains on the no-shock symbol for 2000ms.

After each trial of the simple partner choice task, a trial of the evaluative misattribution task follows.

### Questionnaires

After the main study (but before debriefing), participants respond to the long-form scales “BIS II” (sadness), “Anger” and “BAS II” (joy) from the ARES questionnaire (Hartig & Moosbrugger, 2003). These scales are for exploratory analyses with the general expectation that strong trait sensitivity of the BIS or BAS systems on certain dimensions might moderate the expected effect.

### Hypothesis derivation

An evaluation misattribution effect is described by different proportions of positive judgments after different outcomes of the preceding task. Higher proportions of positive judgments are consistent with more implicit liking.

Theoretically, we expect that the aggressive partner will induce a retaliation goal, which in turn should lead to a positive affective evaluation of successfully affecting him with a noise blast. Conversely, for the nonaggressive partner, no retaliation goal should be induced and participants are likely to experience negative affective consequences when successfully affecting him with a noise blast.

Statistically, this pattern can be tested by comparing the positivity proportions after test trials between targets. A significant main effect of target is expected, that is, the positivity score should be higher for the aggressive partner than for the nonaggressive partner under all circumstances. In addition, the positivity score for the aggressive partner should be higher than for the no-shock condition, which in turn should be higher than for the nonaggressive partner.

With regard to manipulation checks, we expect participants to select the aggressive partner in catch trials more often than the nonaggressive partner. We make no predictions about selection frequency for the no-shock symbol. We also expect participants to select higher intensity rates on average for the aggressive partner than for the nonaggressive partner during the competitive reaction time task. Finally, as a positive control for the evaluative misattribution task, we expect participants to show more positivity after competitive reaction time trials in which they win compared to those in which they lose but are not affected by a noise blast and more positivity after the latter compared to trials in which they lose but are affected by a noise blast.

We derive no formal hypotheses with regard to the ARES scales, but will test moderation via these scales in an exploratory fashion.

### Data collection

A power analysis may be conducted to find this effect using a matched-sample t-test (as each individual hypothesis aspect can be reduced to a comparison between two values within-subjects). In order to detect an effect of *dz* = .3 with acceptable power (80%) in a one-sided t-test at a conventional alpha level of 5%, a sample of 71 participants is required.

The planned *N* of collection is 71. Once the criterion of 71 is achieved, data collection will continue for all already scheduled lab slots and then terminate. At that point, the existing datasets will be evaluated for exclusions which require replacement (see below). If the sample falls to below 71 after excluding these, further lab slots will be scheduled, after which the datasets will be evaluated again. This will be repeated until either the final sample is at least 71 or logistical pressure forces data collection to end early (likely to occur if data collection is not completed by the end of May 2020).

This sample size is sensitive to an effect size of *dz* = .300 with 80% power. If 15% of cases are dropped due to non-replaceable exclusions (see below), sensitivity drops to *dz* = .322.

The sample will be recruited from the psychological institute’s mailing list, which comprises a sample of predominantly students with a significant minority of other demographics. The sample is skewed towards higher education and is predominantly white.

## Analysis plan

### Outlier exclusion rules

During data collection, participants will be excluded based on the following criteria (in order):

* Experimenter notes anomalies in behavior as problematic for analysis (e.g. participant is clearly drunk, agitated or otherwise influenced; participant clearly ignores task instructions) at time of data collection

With regard to the number of times participants choose the aggressive partner to receive a noise blast during the competitive reaction time task, we note conditional exclusion rules. For participants who choose the aggressive partner in less than 60% of the competitive reaction time trials (=90 trials), we will proceed as follows:

* Exclude these participants if the final analysis sample comprises at least 71 participants post exclusion.

Participants excluded due to these criteria will be replaced up to a maximum of 120 total participants (economic cutoff).

* If the maximum number of participants is collected, the number of participants choosing the aggressive partner less than 60% of the time is at least 20 and the number of remaining participants is at least 35, we will group the participants according to this variable and include this variable in the analysis. Our hypothesis will be amended to predict a two-way interaction such that the predicted difference occurs more strongly in the group selected the aggressive partner more often than in the other group.
* In all other cases, they will be retained in the analysis and results will be interpreted with caution.

All trials from the first block pair (competitive reaction time task and simple partner choice task) will be discarded as practice trials (but may be included in exploratory analyses).

Participants who respond to 90% or more of the remaining evaluative misattribution trials with the same response will be eliminated from analysis.

### Analyses

All NHST analyses use an existence-of-effect inference threshold of *p* ≤ .05 unless noted otherwise. Effect sizes will be considered in the interpretation of any effect that is found according to this criterion.

Positivity ratios will be calculated for simple partner choice test trials separately for the aggressive partner, the nonaggressive partner, and the no-shock target. These ratios will be subjected to a one-factor within-subjects ANOVA (no-shock vs. aggressive vs. nonaggressive) with planned linear contrasts. The linear contrast will be the critical effect for inference. Individual contrasts will be calculated for descriptive purposes.

If the exclusion rules trigger the new group-level variable, a 2x3 mixed-model ANOVA will be calculated instead with the between factor of the group-level variable and the within factor target using the aforementioned positivity ratios as a dependent measure. The critical effect for inference will be the interaction term’s linear contrast. Individual contrasts will be calculated for descriptive purposes.

The manipulation check of number of choices in catch trials will be evaluated in a one-sample t-test against 12 (=half the number of trials) and average intensity will be evaluated in analogous tests to the (basic) analysis of the index of reduction. The manipulation check for the evaluative misattribution task will be assessed using a within-subjects ANOVA with the factor outcome (participant won vs. participant lost, no blast vs. participant lost, got blasted).

All other analyses are exploratory/supplementary.

**References:**

Hartig, J., & Moosbrugger, H. (2003). Die “ARES-Skalen” zur Erfassung der individuellen BIS- und BAS-Sensitivität. *Zeitschrift für Differentielle und Diagnostische Psychologie*, *24*(4), 293–310. https://doi.org/10.1024/0170-1789.24.4.293