

## Preference-based escalation: A new interpretation for the responsibility effect in escalating commitment and entrapment

Stefan Schulz-Hardt<sup>a,\*</sup>, Birgit Thurow-Kröning<sup>b</sup>, Dieter Frey<sup>b</sup>

<sup>a</sup> Institute of Psychology, Economic and Social Psychology unit, Georg-August-University Goettingen, Gosslerstrasse 14, D-37073 Goettingen, Germany

<sup>b</sup> Department of Psychology, Social Psychology Unit, Ludwig-Maximilians-University Munich, Leopoldstrasse 13, D-80802 Munich, Germany

### ARTICLE INFO

#### Article history:

Received 28 June 2006

Accepted 4 November 2008

Available online 27 December 2008

Accepted by John Schaubroeck

#### Keywords:

Escalation of commitment

Entrapment

Sunk cost

Self-justification

Responsibility effect

Preference effect

### ABSTRACT

A prominent finding in escalating commitment and entrapment research is the “responsibility effect”: people invest more in a losing course of action or persist with it for longer if they themselves initiated this action (responsibility) as opposed to if it was assigned to them. We argue that this effect is driven by participants’ preferences. Responsible participants usually prefer the chosen alternative since they have chosen it themselves. Non-responsible participants, in contrast, represent a mix of persons who either favor or disfavor the chosen alternative. In two experiments, we demonstrate that responsible participants favor the chosen course of action more strongly than non-responsible participants do, that these preferences facilitate reinvestment in and persistence with the chosen course of action, and that responsibility has no effect over and above this effect of preferences. Non-responsible participants preferring the chosen course of action made similar reinvestments and exhibited similar persistence as responsible participants.

© 2008 Elsevier Inc. All rights reserved.

### Introduction

Research on “escalation of commitment” and “entrapment” deals with the question of why and under what circumstances people persist with losing courses of action (for overviews see Brockner & Rubin, 1985; Staw, 1997). In this research it is usually assumed that people tend to become entangled in losing courses of action and, thus, to “throw good money after bad”. Studies with a more pronounced economic orientation speak of the “sunk cost effect” (Arkes & Blumer, 1985): when deciding among several alternatives – for example, whether to persist with or abandon an action – people not only consider the anticipated future costs and benefits but are also guided by how much they have already invested in a particular alternative. Such processes are used as explanations for loss escalations in diverse contexts, for example in professional sports (Staw & Hoang, 1995), loan decisions in banks (Staw, Barsade, & Koput, 1997), political decisions (Ross & Staw, 1986), and in decisions made by the public administration (Drummond, 1994a; 1994b).

In this paper, we will deal with a particular effect that is highly prominent in research on entrapment and escalating commitment, namely the *responsibility effect* (Staw, 1976). The responsibility effect refers to the phenomenon that when people are confronted

with negative feedback about the success of a decision, they commit more resources to this course of action and persist with it for a longer period of time if they were responsible for making the initial decision than if someone else made this decision. This effect is the best replicated finding in entrapment and escalating commitment research (e.g., Brockner, 1992). The purpose of our paper is to provide a reinterpretation of this central finding: whereas classic research on entrapment and escalation of commitment has interpreted this effect as being due to self-justification concerns among people who feel responsible for having made a decision with negative consequences (e.g., Brockner, 1992; Staw, 1976), according to our reinterpretation this effect stems from the fact that people simply commit more resources to a decision alternative they prefer than to an alternative that they do not prefer – which we will label *preference effect*. In the following, we will briefly describe prior research on the responsibility effect and then outline our reinterpretation.

#### Prior research on the responsibility effect

In his seminal study, Staw (1976) had participants simulate the role of a vice president of a company called “Adams & Smith”. They were faced with the task of determining the allocation of financial resources for research and development. The company was structured into two main divisions, namely “consumer products” and “industrial products”. Half of the participants decided autonomously

\* Corresponding author. Tel.: +49 551 39 13561; fax: +49 551 39 13570.

E-mail address: [schulz-hardt@psych.uni-goettingen.de](mailto:schulz-hardt@psych.uni-goettingen.de) (S. Schulz-Hardt).

mously which division should receive an additional US\$ 10 million (*responsibility* conditions). The other half of the participants were presented with a decision made by the preceding vice president (*no responsibility* conditions). For these participants, a random assignment was realized as to whether the preceding vice president had chosen consumer products or industrial products. Although these participants had also received all the information about the two divisions, they were not asked how they would have decided or which division they preferred.

For all participants, a further manipulation included whether the chosen division performed well or poorly over the next 5 years after receiving the financial boost. Participants were then asked to decide what proportion of a newly available investment pool of US\$ 20 million should be allocated to this particular division. Staw (1976) showed that the initially chosen division received particularly large investments when negative feedback was given and participants were responsible for the initial decision to financially support this division. He subsequently deduced that the increased investments in the unprofitable company division, labeled as *escalating commitment*, represented an attempt to justify prior investments.

In subsequent studies, most researchers focused on the conditions with negative feedback. In several follow-up studies it was shown that people who were responsible for the initial decision made higher resource commitments than participants who had not made the initial decision (Barton, Duchon, & Dunegan, 1989; Bazerman, Giuliano, & Appelman, 1984; Brody & Frank, 2002; Conlon & Parks, 1987; Davis & Bobko, 1986; Kirby & Davis, 1998; Schmidt & Calantone, 2002; Schoorman & Holahan, 1996; Schoorman, Mayer, Douglas, & Hetrick, 1994; Schulz & Cheng, 2002). The responsibility effect was also obtained for the duration of persistence with the losing course of action, which is the central dependent variable in the *entrapment* research tradition: when confronted with negative feedback, persons responsible for the initial decision persisted with the alternative for a longer time than those who had been assigned this alternative (Kameda & Sugimori, 1993). The responsibility effect appeared even if, as in the tradition of sunk cost research, merely a decision on continuation or termination had to be made (Boulding, Morgan, & Staelin, 1997; Brody & Bowman, 1997; Schaubroeck & Williams, 1993; Whyte, 1991). Finally, some studies have also demonstrated the responsibility effect among groups as decision makers (Bazerman et al., 1984; Kameda & Sugimori, 1993).

In some field studies the generalizability of the responsibility effect to real world decision making has been shown. For example, Staw et al. (1997) showed that executive turnover speeds withdrawal from bad bank loans. In other words: the greater the number of current board members who had been involved in the initial decision to provide the loan, the less likely the board was to write it off if it turned out to be a problem loan. In a similar vein, McNamara, Moon, and Bromiley (2002) demonstrated that necessary downgradings in a borrower's credit limit were more frequently made if the lending officer had changed. Thus, prior research has convincingly demonstrated the robustness of the responsibility effect as well as its practical relevance.

From the very beginning, the responsibility effect has been taken as evidence for a *self-justification process* which constitutes one of the major psychological explanations for entrapment and escalating commitment behavior (e.g., Brockner, 1992). According to this view, which is rooted in dissonance theory (Festinger, 1957), people continue to invest further resources in a failing course of action because they do not want to admit that initiating the action was a mistake. Because people who had nothing to do with the initial decision should hardly feel any pressures to justify this decision, they exhibit lower reinvestments and lower persistence (Brockner, 1992; Staw, 1976). In contrast, responsible participants feel responsible for the negative consequences, and in an

attempt to justify the initial decision they show heightened reinvestment and persistence.

Whereas this assumption appears plausible, to our knowledge no studies exist in which feelings of responsibility or self-justification processes have been convincingly shown to mediate the increased investment levels or the longer persistence in the responsibility condition compared to the no responsibility condition. For example, Brody and Bowman (1997) compared information recall among high responsibility vs. low responsibility participants. From a self-justification perspective, the authors expected high responsibility participants to remember more positive information about the chosen alternative compared to low responsibility participants, because positive information about the chosen alternative could be used to justify the initial choice. However, the results revealed no evidence for such a self-justificatory memory bias.

Using the “Adams & Smith” case, Bazerman et al. (1984) showed that, compared to participants in the no responsibility condition, participants in the responsibility condition felt more committed to the initial decision, judged the second investment decision to be more contingent on the first, were more confident about having made a good second investment decision, and were more convinced that additional investments could turn the situation around. However, for the “commitment” variable which seems to have the closest connection to felt responsibility for negative outcomes and to self-justification processes, no significant association with the level of investment could be shown. In contrast, the other three variables were positively correlated with the level of investment, but none of them seems to be very indicative of self-justification. For example, the fact that people who strongly believe that further investments will be successful make higher investments than people who have doubts about the efficacy of further resource allocations is a finding that we would also expect in the case of a purely rational decision making process.

The possibly most sophisticated attempt to show that self-justification drives the escalating commitment among high responsibility participants was made by Conlon and Parks (1987). They replicated Staw's (1976) design and extended it by giving participants the opportunity to search for prospective or retrospective information before making the second investment decision. They found that the only condition where the participants preferred retrospective to prospective information was the “responsibility/negative feedback” condition. They interpreted this finding as the participants' attempt to obtain information that could justify their prior choice. However, Schultze, Pfeiffer, and Schulz-Hardt (submitted for publication) have recently shown that a confound in the original information material might have caused this finding, and they demonstrated that no self-justificatory information search can be observed if the participants are provided with information that unequivocally provides justification vs. criticism regarding the initial decision.

All in all, no clear empirical support exists for the assumption that the responsibility effect rests on the responsibility felt for negative outcomes and subsequent self-justification. In our opinion, the responsibility effect can be parsimoniously explained by (more or less) rational decision-making concerns, with no reference to felt responsibility or self-justification. In the following section, we will outline this alternative explanation.

#### *Preference instead of felt responsibility: reinterpreting the responsibility effect*

If people face negative feedback about the success of an initial decision, they have to reevaluate whether sticking to this decision and, perhaps, increasing one's resource allocations to the chosen alternative has a good chance of finally leading to successful out-

comes, or whether abandoning the chosen course of action and allocating one's resources to a different alternative can be expected to yield a higher return. This evaluation should, at least to some extent, depend on the person's initial belief about how promising the chosen course of action is compared to its competitors.<sup>1</sup> If Person A considers an Alternative X as being superior to its competitors, which means that A has a *preference* for X, whereas Person B considers X to be inferior to some or all of its competitors (i.e., B has a preference for non-X), we should expect A to reinvest more money than B, given that both persons receive the same negative feedback about the success of X.

The crucial point is that these preferences systematically differ between participants in the responsibility and in the no responsibility condition. Consider, for example, the seminal *Staw (1976)* study. Participants in the responsibility condition made the initial decision themselves. If they preferred consumer products, they chose consumer products; vice versa for industrial products. Thus, the chosen alternative should have been the one that they considered most promising. Participants in the no responsibility condition received the same information as participants in the responsibility condition, so again some should have preferred consumer products and some should have preferred industrial products, and some may have judged both alternatives as being equally attractive (which means that they had no preference). However, since the initial decision that had been made by the former vice president was assigned at random, no more than one half of the participants should have considered the chosen alternative as the one with the highest likelihood of success.

As a consequence, the average initial belief about the success of the chosen alternative should have been higher in the responsibility condition than in the no responsibility condition and, thus, participants in the former condition should have considered further resource allocations to the initially chosen alternative to be more likely to succeed in the responsibility than in the no responsibility condition. Because resource allocations are contingent on the decision maker's beliefs about how successful such allocations will be (e.g., *Bateman, 1986; Bateman & Zeithaml, 1989; Bazerman et al., 1984*), participants in the responsibility condition allocate more resources to the initially chosen alternative than participants in the no responsibility condition. Taken together, preferences should mediate the responsibility effect: responsible participants show higher escalation than non-responsible participants because the former – almost inevitably – prefer the chosen course of action more than the latter do.

This parsimonious interpretation of the responsibility effect has not yet been empirically tested, although some previous findings hint at its validity. As mentioned earlier, *Bazerman et al. (1984)* showed that, compared to non-responsible participants, participants in the responsibility condition were more convinced that further investments in the previously chosen alternative would turn the situation around. In a similar vein, *Bazerman, Beekun, and Schoorman (1982)* showed that participants who were responsible for hiring a manager expected the manager to perform better in the future than non-responsible participants did. Both findings are in line with our view that responsible participants usually prefer the chosen alternative and, thus, expect it to perform better compared to non-responsible participants of whom only a subgroup

prefers this alternative. However, since the above-mentioned studies did not measure preferences among non-responsible participants, the data do not directly address our interpretation. *Liao, Finley, and Shafer (2004)* manipulated whether the chosen alternative did or did not match the participants' preferences. In a modified version of *Staw's (1976)* Adams & Smith case, the participants played the role of members of the Financial Management Committee of the company and recommended one of the company's two divisions for the initial investment. However, the final decision about this investment was made by the board of directors, and this board either chose the division proposed by the participants or the other division. (Misleadingly, Liao et al. labeled the former participants "responsible" and the latter ones "non-responsible", although in both cases the participants were not responsible for the final decision.) Following negative feedback, participants who had preferred and recommended the division chosen by the board of directors reinvested more money into the chosen division than participants who had preferred and recommended the non-chosen division. This underlines our argument that preferences in the no responsibility condition have to be taken into account, but due to the fact that the Liao et al. study did not contain a traditional responsibility condition, no real test of our interpretation of the responsibility effect can be derived.

Basically, what one has to do to test our preference interpretation of the responsibility effect is:

- (a) Replicate the classical responsibility effect (on escalation/entrapment).
- (b) Test for the proposed differences in preferences between responsible and non-responsible participants.
- (c) Establish the relation between preferences and subsequent escalation/entrapment.
- (d) Test whether responsibility has an effect on escalation or entrapment over and above the effect that preferences have (i.e., whether preferences fully explain the effect).

It should be emphasized that (d) is the hallmark of a self-justification interpretation of the responsibility effect. According to this interpretation, if you have two persons, A and B, who are similarly in favor of the chosen course of action, but A was responsible for choosing it whereas B has been assigned this course of action, then A should show more escalation than B, because A should feel responsible for the negative consequences whereas B should not. In contrast, according to our approach, they should not differ. Hence, although a study along the lines of the above-mentioned points (a)–(d) would be primarily a test of our preference interpretation, it would also have some implications for the self-justification view of the responsibility effect.

In the following, we report on two experiments that were conducted to provide such a test. In Experiment 1, we used the seminal "Adams & Smith" decision case originally developed by *Staw (1976)*, with *amount of investment* after an initial negative feedback being the primary dependent measure. In Experiment 2, we used a paradigm that involved active role-playing and in which participants' financial rewards were (allegedly) contingent on their behavior. In this experiment, *persistence* with a losing course of action was used as the primary dependent variable. Whereas in Experiment 1 preference vs. non-preference for the chosen course of action was measured, in Experiment 2 it was experimentally manipulated. The general hypothesis tested in these experiments can be summarized as this:

*The effect of responsibility on escalation and entrapment is mediated by differently strong preferences for the chosen course of action among responsible vs. non-responsible participants and is eliminated if these preferences are held constant.*

<sup>1</sup> Studies have shown that base-rate information (as a kind of initial belief) can be completely neglected (e.g., *Bar-Hillel & Fischhoff, 1981*). However, such neglect is restricted to situations where base rates are solely introduced by a condensed statistical format; it does not occur if people generate these base rates from their own experience or from their own evaluations of the alternatives (for a more detailed discussion of this issue see *Betsch, Haberstroh, Glöckner, Haar, & Fiedler, 2001*). Since the latter is the case in the escalating commitment and entrapment situations we refer to, it is reasonable to assume that the person's prior beliefs about the alternatives will, at least to some extent, influence how she evaluates the alternatives after having received negative feedback about the initial choice.

## Experiment 1

In Experiment 1, we replicated the procedure of classical experiments on the responsibility effect, using the seminal “Adams & Smith” case (in a German adaptation). As in most of these experiments, only negative feedback was used, and a responsibility condition was contrasted with a no responsibility condition. The only difference to former experiments was that we measured participants’ preferences by asking them to rate how promising an investment in each of the two divisions was. The division with the higher rating was taken as their preference. Furthermore, the difference between the rating for the chosen alternative and the rating for the non-chosen alternative will be labeled *preference strength*. Large positive numbers indicate that the person clearly favors the chosen over the non-chosen alternative. A value of zero indicates that the person is more or less indifferent to the two alternatives, and the more negative the value becomes, the more strongly the participant *disfavors* the chosen alternative.

Note that the elicitation of preferences in the no responsibility condition did not imply any kind of choice for the participants: they were not asked to state which division they preferred, only to give ratings for each of the two divisions. However, even this slight modification (compared to previous experiments on the responsibility effect) might influence the results. In other words, if we find the predicted effects for preferences, we cannot exclude the possibility that this evidence was produced by the newly introduced ratings, whereas in former experiments that lacked these ratings, no preference effects were present. Therefore, whereas half of the participants (both in the responsibility and in the no responsibility conditions) were asked to give the preference ratings immediately before they had to make a decision for one of the two divisions (responsibility) or were informed about the decision of their predecessor (no responsibility), for the other half these preferences were measured retrospectively, namely *after* the final reinvestment decision had been made, and before participants were fully debriefed. This method has the advantage that the measurement of preferences cannot affect the dependent variable (amount of reinvestment), because the measurement takes place after reinvestment. On the other hand, it opens up the possibility that the stated preferences are not always identical with the true preferences at the moment when the initial decision was made. Thus, the most confident answer to our research question will be possible if the results do not vary dependent on whether preferences are measured prior to or after the dependent variable.

### Hypotheses

The following specific hypotheses were derived for this experiment:

*H1: Responsible participants will make higher reinvestments than non-responsible participants (replication of the classical responsibility effect)*

*H2: Responsible participants’ preferences will be more in favor of the chosen alternative than non-responsible participants’ preferences.*

*H3: The differences specified in H1 will be mediated by the differences specified in H2. When controlling for preferences, effects of responsibility on reinvestment will be eliminated.*

### Method

#### Participants and design

One hundred and thirty eight students from the University of Goettingen participated in the experiment. Of these 138 participants, seven had to be excluded; three due to missing data in the questionnaires, one because of serious problems with compre-

hending German, and three because of, according to the suspicion check, correctly guessing the experimental hypotheses or reporting that bogus feedback had been given.

Of the remaining 131 participants, 60 were female and 71 were male. The participants’ age ranged from 17 to 34 ( $M = 23.51$ ,  $SD = 2.85$ ). The experiment is based on a 2 (*responsibility*: given vs. not given)  $\times$  2 (*assessment of preferences*: initial vs. retrospective) factorial between-subjects design. Participants were randomly assigned to the experimental conditions.

#### Procedure

The experiment took place in a laboratory room. Participants were directly recruited on the university campus. First, the experimenter introduced herself and explained that, as part of a larger project on decision making, we were investigating how people make decisions about financial resource allocations. Participants were told that they would receive an economic case study and should try to make decisions that maximize the respective company’s profit. Participants were led to believe that they would receive performance-based payment, solely dependent on the success of the fictional company in the case study, with an average performance yielding a 3 Euro (about 4.20 US \$) payment. Because negative feedback was given independent of participants’ decisions, all participants received the fixed amount of 3 Euro.

The participants were then handed out a German adaptation of Staw’s “Adams & Smith” case. It describes the situation of a German company called “Ankor AG”. The Ankor AG produces kitchen utensils as well as household appliances and consists of two main divisions, namely “consumer products” and “industrial products”. Both divisions were briefly described to the participants. Furthermore, participants learned that the profitability of both divisions had recently declined. This was illustrated by sales and earnings data for both divisions for the last 10 years; the numbers were identical with those used by Staw (1976). The data showed a significant decrease in earnings, although sales had risen slightly for both divisions. The participants were informed that the managerial board of the Ankor AG believes that deficits in research and development (R&D) in both divisions account for a major part of this decline. Therefore, the managerial board has decided that an additional R&D funding of 10 Million Euro shall be invested in one of the company’s two main divisions.

The subsequent procedure differed slightly between the experimental conditions. In the *responsibility* conditions, the participants were informed that they were the company’s director for research and development. Therefore, they were responsible for making R&D allocation decisions like the one at hand. In the *no responsibility* conditions, participants were informed that the R&D director of the Ankor AG was responsible for making this decision. Half of the participants, namely those in the conditions with retrospective assessment of preferences (taking place at the end of the experiment), now had to make the initial decision about which division should receive the financial support and provide some reasons for their choice (responsibility) or were informed about which division had been chosen by the R&D director (no responsibility). For the other half of the participants, who were in the conditions with initial assessment of preferences, the making (responsibility) or disclosure of this decision (no responsibility) was preceded by the preference assessment. Therefore, participants received a short questionnaire where they were asked to rate how successful an investment in each of the two divisions would be. To avoid priming effects, three distractor items were added, asking participants how motivated they were to obtain good results, how realistic they thought the case study to be, and how well they could concentrate. All five questions used 7-point Likert scales ranging from “not at all” (–3) to “very much” (+3).



All participants were then informed that, 5 years after this investment decision, a new evaluation of the company's R&D activities took place. Participants in the no responsibility conditions should now play the role of the company's new R&D director, whereas participants in the responsibility condition were reminded that they were still acting as the company's R&D director. The participants learned that another 20 Million Euro R&D funding had been made available by the managerial board. This time, participants should decide how much of the available 20 Million Euro the initially funded division should receive. Participants were informed that the remaining funds would be made available to finance other projects within the company.

As a basis for their decisions, the participants received the two divisions' sales and earnings data for the five year period following the initial investment. With regard to this initial decision, the feedback was negative: whereas losses had increased for the division that had received the initial funding, the other division had cut its losses and was now making small profits. The data were identical to the negative feedback in the Staw (1976) study.

The participants were first asked to rate how positive or negative they perceived the two divisions' financial results (both on scales from  $-5$  = very negative to  $5$  = very positive) and then to make their decision about how much of the 20 Million Euro R&D funding should be invested in the initially chosen division. Furthermore, they were asked to write down some reasons for their choice.

For participants in the conditions with initial preference assessment the experiment was over; they were thanked and fully debriefed. For the other half of the participants, this debriefing was preceded by the retrospective preference assessment. We used the same questionnaire as in the conditions with initial preference assessment for this purpose; however, the questions were formulated retrospectively, that is, participants were asked to state how promising each of the two divisions had initially appeared to them (same scales as in the conditions with initial preference assessment).

## Results and discussion

In order to check for possible confounds, the dependence of amount of reinvestment on factors such as age, gender, course of study, or the specific chosen division was examined. No significant effects were found.

### Perception of feedback

A  $2$  (responsibility)  $\times 2$  (measurement of preferences)  $\times 2$  (division rated) analysis of variance (ANOVA) with repeated measurement on the third factor revealed that participants perceived the financial results of the chosen division ( $M = -2.69$ ,  $SD = 1.55$ ) as clearly more negative than the financial results of the non-chosen division ( $M = 2.29$ ,  $SD = 1.56$ ),  $F(1, 126) = 697.46$ ,  $p < .001$ , part.  $\eta^2 = .85$ . Thus, negative feedback has been successfully introduced. No significant interactions with the experimental factors occurred, all  $F$ s  $< 1.35$ , all  $p$ s  $> .25$ .

### Amount of reinvestment – testing for the responsibility effect

The central dependent variable in this experiment was the amount of money that participants reinvested in the initially funded division (chosen by themselves or their predecessor). In a first step, this variable was analyzed using a  $2 \times 2$  ANOVA of the experimental design. This analysis revealed a significant main effect of responsibility,  $F(1, 127) = 7.77$ ,  $p = .006$ , part.  $\eta^2 = .06$ . Participants who were responsible for the initial decision reinvested a higher amount of money in the initially funded division ( $M = 10.52$ ;  $SD = 5.79$ ) than participants without responsibility for the initial decision ( $M = 7.69$ ,  $SD = 5.75$ ). Thus, we successfully rep-

licated the responsibility effect of previous studies on escalating commitment (Brockner, 1992), supporting Hypothesis 1. Measurement of preferences did not significantly affect reinvestment,  $F(1, 129) = 0.09$ ,  $p = .77$ , part.  $\eta^2 = .001$ , and did not significantly interact with responsibility,  $F(1, 127) = 0.44$ ,  $p = .51$ , part.  $\eta^2 = .003$ . In other words, the responsibility effect can be replicated independently, regardless of whether the experimental procedure was identical to previous experiments, with the addition of preferences being measured retrospectively at the end, or whether preferences were measured before the initial funding decision was made (responsibility condition) or disclosed (no responsibility condition).

### Preferences as a mediator for responsibility effects on reinvestment

A central prediction of our approach is that responsible participants should have a clear preference for the chosen division, whereas non-responsible participants should, on average, be more or less indifferent towards that alternative. Therefore, preference strength (i.e., the difference between the ratings of expected success for the chosen alternative and the non-chosen alternative) was analyzed in a  $2 \times 2$  ANOVA of the experimental design. This analysis revealed a significant main effect of responsibility,  $F(1, 127) = 45.12$ ,  $p < .001$ , part.  $\eta^2 = .26$ . As predicted, responsible participants significantly preferred the chosen division ( $M = 1.75$ ;  $SD = 1.50$ ),  $t(63) = 9.33$ ,  $p < .001$  for a test against zero. In contrast, non-responsible participants, on average, marginally disfavored the chosen alternative ( $M = -0.51$ ;  $SD = 2.27$ ),  $t(66) = 1.83$ ,  $p = .071$ .<sup>2</sup> Thus, Hypothesis 2 is also supported by the data.

No significant main effect was obtained for measurement of preferences,  $F(1, 127) = 45.12$ ,  $p < .001$ , part.  $\eta^2 = .26$ , but the interaction of the two factors was marginal,  $F(1, 127) = 2.85$ ,  $p = .094$ , part.  $\eta^2 = .02$ . This was due to the fact that non-responsible participants were, on average, indifferent towards the chosen alternative when preferences were initially measured ( $M = -0.06$ ;  $SD = 2.12$ ),  $t(32) = -0.16$ ,  $p = .871$ , whereas participants with retrospective preference measurement reported that they had, on average, been in disfavor of the chosen alternative ( $M = -0.94$ ;  $SD = 2.35$ ),  $t(33) = -2.34$ ,  $p = .026$  (no such differences were apparent for responsible participants). Thus, the retrospective preference reports might be somewhat biased by having experienced the negative performance of the funded division.

To test whether the differences in preference strength towards the chosen alternative mediate the reinvestment differences between responsible and non-responsible participants, a mediation analysis using the regression approach was conducted. As already outlined in the factorial analysis, responsibility significantly enhances reinvestment,  $\beta = .24$ ,  $t(129) = 2.81$ ,  $p = .006$ , and preference strength,  $\beta = .51$ ,  $t(129) = 6.69$ ,  $p < .001$ . Now, if reinvestment is regressed both on responsibility and on preference strength simultaneously ( $R^2 = .17$ ,  $F[2, 128] = 12.67$ ,  $p < .001$  for the overall regression), then preference strength receives a significant regression weight,  $\beta = .38$ ,  $t(128) = 4.06$ ,  $p < .001$ , indicating that a stronger preference for the chosen alternative is associated with higher reinvestments, when statistically controlling for responsibility. The

<sup>2</sup> From our approach it can be derived that not only the means but also the variances should differ between responsible and non-responsible participants. Responsible participants should more or less differ only in the extent to which they prefer the alternative that they have chosen, whereas non-responsible participants should show a whole range of preferences for and against the chosen alternative. The reported standard deviations support this idea, and the Levene-Test for homogeneity of variances becomes significant,  $F(3, 127) = 5.06$ ,  $p = .002$ . However, because the cell sizes are almost equal, and because in this case the ANOVA is rather robust against violations of variance homogeneity, the results of the ANOVA should still be interpretable. Furthermore, a  $t$ -test with a separate variance estimate for the case of inhomogeneous variances also results in a significant mean difference between responsible and non-responsible participants,  $t(115.14) = 6.75$ ,  $p < .001$ .

**Table 1**

Means for amount of reinvestment in an orthogonal crossing of responsibility and preferences (standard deviations in brackets).

Responsibility	Preference		
	For the chosen alternative	Indifference	For the non-chosen alternative
Responsible	10.95 (6.06) N = 51	9.41 (4.43) N = 11	5.50 (0.71) N = 2
Not responsible	10.86 (5.71) N = 21	4.79 (4.17) N = 14	6.84 (5.55) N = 32

Notes: amount of reinvestment ranges from 0 to 20 (Mio. €). Because preferences are partially dependent on responsibility, the cell sizes necessarily are strongly unequal.

bivariate correlations are  $r = .371$  ( $p = .003$ ) for responsible participants and  $r = .328$  ( $p = .007$ ) for non-responsible participants. Most importantly, in this multiple regression (i.e., controlling for preference) the weight for responsibility is no longer significant,  $\beta = .05$ ,  $t(128) = 0.50$ ,  $p = .620$ . A Sobel test shows that the reduction in the regression weight for responsibility is significant,  $z = 3.47$ ,  $p < .001$ . These results indicate that responsibility has no (additional) impact on reinvestment if preference strength is held constant and that the responsibility effect is fully mediated by preference strength. Similar mediation results are obtained if the participants with initial measurement of preferences and the participants with retrospective preference measurement are analyzed separately. Taken together, Hypothesis 3 has also been shown to be valid.

#### Crossing preferences and responsibility: a categorical analysis

In the foregoing, we have used a continuous preference measure (preference strength) and have tested our central hypotheses using a regression approach. Whereas we believe this to be the most appropriate test (because no experimental manipulation took place for preferences) its implications for our research question can be better illustrated if we additionally use a categorical split into different preference subgroups. Of the 67 non-responsible participants, 21 reported a preference in favor of the alternative chosen by their predecessor (i.e., their preference strength value was positive) whereas 32 preferred the non-chosen alternative (they had a negative preference strength value). The remaining 14 participants were indifferent to the two alternatives (their preference strength value was zero). As expected, most participants in the responsibility conditions – 51 out of 64 – preferred the chosen alternative. However, 11 responsible participants rated the two alternatives equally in expected success (i.e., they were more or less indifferent), and 2 participants even rated the non-chosen alternative slightly higher.

Although the cell sizes are necessarily very different, this pattern allows for a full orthogonal crossing of responsibility and preference (for the means and standard deviations see Table 1). In a 2 (responsible vs. non-responsible)  $\times$  3 (preference for the chosen alternative vs. indifference vs. preference for the non-chosen alternative) ANOVA, only the main effect for the preference factor is significant,  $F(2, 125) = 5.47$ ,  $p = .005$ , part.  $\eta^2 = .080$ . This illustrates the regression finding that preferences have an effect even if responsibility is held constant. In contrast, no significant main effect is found for responsibility,  $F(1, 125) = 0.46$ ,  $p = .50$ , part.  $\eta^2 = .004$ , showing that responsibility does not affect reinvestment if preferences are controlled for. Finally, the interaction is also non-significant,  $F(2, 125) = 1.61$ ,  $p = .21$ , part.  $\eta^2 = .025$ .<sup>3</sup>

<sup>3</sup> Due to the strongly differing cell sizes and the very low number of cases in one of the cells ( $n = 2$ ), we repeated this analysis while collapsing indifferent participants and participants with a preference for the non-chosen alternative into one factor level ("no preference for the chosen alternative" subgroup). The corresponding 2  $\times$  2 ANOVA leads to the same results as the 3  $\times$  2 ANOVA outlined above.

In spite of the non-significant interaction, we tested for responsibility effects within each preference subgroup. The corresponding contrasts are not significant for participants preferring the chosen division,  $t(125) = 0.07$ ,  $p = .95$ , or for participants preferring the non-chosen division,  $t(125) = -0.33$ ,  $p = .74$ , but significant for indifferent participants,  $t(125) = 2.03$ ,  $p = .044$ . This might indicate that responsibility heightens reinvestment selectively among indifferent participants. However, this interpretation would be premature, not only due to the weak statistical basis of the effect, but also due to theoretical considerations: even if participants rate the two alternatives equally regarding their success expectations, they might nevertheless slightly prefer one of them. In the responsibility condition, they will choose the one they slightly prefer, whereas in the no responsibility condition they might end up being assigned the slightly non-preferred option. Therefore, in Experiment 2 we will make sure that both subgroups are treated equally by assigning indifferent non-responsible participants the alternative that they slightly prefer.

#### Summary

In sum, Experiment 1 has supported our preference interpretation of the responsibility effect. Responsible participants have been shown to predominantly hold preferences in favor of the chosen alternative, whereas non-responsible participants were, on average, neutral or slightly in disfavor of this alternative. These differently strong preferences for the chosen alternative mediated the higher reinvestments of responsible as compared to non-responsible participants, and if preferences were held constant, responsibility effects – with the one possible exception mentioned above – were eliminated.

#### Experiment 2

Experiment 2 extends Experiment 1 with regard to three aspects: first, we transferred our analysis from escalation of commitment, which was investigated in Experiment 1, to an entrapment situation. In this situation, time of persistence with a losing course of action was the main dependent variable. Thus, we want to test whether responsibility effects in entrapment (e.g., Kameda & Sugimori, 1993) can also be explained by participants' preferences.

In addition, we wanted to further strengthen the empirical support for our preference interpretation of responsibility effects by experimentally manipulating the correspondence between participants' preferences and the chosen alternative among non-responsible participants. In Experiment 1, this correspondence was "only" measured, which, for example, led to variance in cell sizes if preference subgroups were formed and compared. Thus, in Experiment 2 we measured non-responsible participants' preferences and subsequently assigned their predecessor's choice dependent on their preference, so that this choice either corresponded or did not correspond with their preference. As already mentioned above, in the case of indifference, non-responsible participants indicated the alternative they slightly preferred and were assigned this alternative (whereas indifferent responsible participants simply chose the alternative they slightly preferred). Hence, this leads to three subgroups, namely responsible participants, non-responsible participants who prefer the chosen alternative, and non-responsible participants who prefer a different alternative to the chosen one. These three subgroups allow for an experimental test of our preference interpretation of the responsibility effect, leading to the following hypotheses:

*H1: Non-responsible participants preferring the chosen alternative persist as long as responsible participants.*

*H2: Both responsible participants and non-responsible participants preferring the chosen alternative persist longer with this alternative*

compared to non-responsible participants preferring a non-chosen alternative.

Taken together, this pattern would mean that, when holding preferences constant (namely among participants who prefer the chosen course of action), responsibility does not have an effect. In contrast, when holding responsibility constant (namely among the non-responsible participants), preferences do have an effect. In other words, the escalation effect covaries with preferences, not with responsibility.

Finally, although Experiment 2 contained only conditions with negative feedback (as in Experiment 1), in Experiment 2 the degree of negativity of the feedback was varied. This was done to clarify whether reinvestment or persistence depends solely upon the participants' preferences or whether they still take the specific feedback into account. Furthermore, this allowed us to test a possible limiting condition for preference effects, namely whether preference effects disappear if the feedback becomes extremely negative, because in this situation any preference might be "overruled" by the evidence. With regard to this, our predictions were the following:

*H3: Persistence linearly decreases with feedback negativity.* That is, participants take the specific feedback into account and withdraw earlier from the chosen course of action, the more negative the feedback is.

*H4: Choice status does not interact with feedback.* That is, preference effects are robust across different degrees of feedback negativity.

## Method

### Participants and design

One hundred and seventy undergraduate and graduate students participated in the study. They were recruited through billboard postings and newspaper advertisements. Eight of these 170 participants had to be excluded because in the concluding interview it became apparent that they had been suspicious about our cover story or had been informed about the objectives of the experiment by former participants. Thus, 162 participants remained of whom 98 were female and 64 were male. Ages ranged between 19 and 43 years ( $M = 24.29$ ,  $SD = 5.19$ ).

The experiment is based on a  $3 \times 3$  (choice status: responsibility vs. no responsibility/preferred choice vs. no responsibility/non-preferred choice; feedback: mildly negative vs. moderately negative vs. strongly negative) factorial between-subjects design. Participants were randomly assigned to each of the nine experimental conditions, giving groups of 18 persons. The experiment took place in five separate laboratory rooms.

### Procedure

The paradigm used in Experiment 2 was originally developed by Bobocel and Meyer (1994); several modifications were made for our purposes. Five participants were summoned to an appointment at a time and first convened in a room where they were welcomed by the experimenter and her assistants. They completed a brief questionnaire on demographic data and were assigned a code number on this questionnaire since data sampling was to be carried out anonymously. The experimenter then told the participants that four of them would simulate a department; the fifth person would receive a separate task. The department was to consist of one department head and three coworkers; all of these roles would be assigned by drawing lots. The task of the department consisted in translating enciphered texts. Four strategies of differing effectiveness were available for decryption. The department head would be responsible for selecting the translation strategy to be employed by the coworkers. The target set was to translate at least 45 short texts of standardized length within 1 h. This target was in-

tended to help participants assess feedback appropriately at a later point in time. According to the experimenter, pilot investigations had revealed that the translation of 45 texts represented a manageable task for the department. Fees would be performance-related: they would depend on the number of texts decoded by the department. Additionally, the best two departments would receive a special bonus amounting to 100 Euro (about 90 US \$).

The participants were then asked into separate rooms. The subsequent procedure depended on which condition of the choice status factor was being implemented. In the *responsibility* condition, participants were allotted the role of the department head. They received an instruction sheet for department heads describing their various tasks including the selection of the translation strategy and the monitoring of coworkers' performance. In order to select a strategy, they received an exemplary text (a text on transaction analysis that had been converted to the MS-Word font "Wingdings" and thus contained symbol chains such as "☞■☞●☞♦☞"), a translation table, and a sheet of paper describing the four strategies. For example, one strategy was to translate the text letter by letter, whereas another strategy was to translate one letter, then search for other occurrences of this letter in the text etc. In a pretest we had made sure that, on average, all four strategies were judged as comparably effective. After approximately 10 min deliberation and trial testing, participants estimated the effectiveness of each of the four strategies on a scale from 0 to 10 and decided on one of the strategies. The experimenter then ostensibly informed the (actually non-existent) coworkers of the decision. Furthermore, it was emphasized that experimenters would continually monitor the coworkers and ensure that they actually used the chosen strategy.

In both *no responsibility* conditions, the participants first drew the individual task. They were instructed that their task consisted in making various assessments that would be used in further studies while the other participants would be working as a department. The first assessment concerned the strategies among which the head of department had to decide. To provide department heads with helpful tools in subsequent studies, they should try out the various strategies using the given sample text – for this purpose the same two sheets were distributed as in the responsibility condition – and then assess how effective they judged each of the four strategies to be on scales ranging from 0 to 10. The latter questionnaire differed from the responsibility condition in that a decision was explicitly *not* requested. In the case of two or more strategies being given the same ratings and no alternative strategy being rated higher, experimenters also asked which of the strategies the participant would *most* recommend. Thus, in contrast to Experiment 1, in Experiment 2 participants who had yet not exhibited a preference (and who formed a separate condition in Experiment 1) were forced to do so to enable the correspondence between preference and chosen alternative to be fully experimentally manipulated.

Shortly thereafter a slight confusion was staged in the hallway followed by the sound of a door slamming. The experimenter entered the room appearing surprised and irritated and informed the participant that another participant had forgotten an important appointment and had had to leave the experiment urgently. Unfortunately, the role assigned to the participant who had left was that of the department head. Thus, the remaining participant was asked to stand in as the new department head. Participants always responded positively to this request. They were then given the department head documentation and were able to surmise from it that their predecessor had already chosen a strategy. The coworkers had also already been informed of the choice and had just begun working with it. In the *no responsibility/preferred choice* condition this strategy corresponded to the one judged as the most effective by the participant, in the *no responsibility/non-preferred*

choice condition one of the other three strategies was assigned at random.

From this point onwards the remaining procedure was identical for all conditions. Participants were informed that they would receive feedback every 6 min on the number of texts translated by their coworkers. They were to note down this information on the “performance checklist” specifically devised for this purpose. They were informed that every time they received this feedback they would have the opportunity to change strategy; however, it should be noted that every new strategy involves a certain lead time. To avoid any time being spent idle between the individual feedbacks, the department heads were additionally given a personnel selection task of choosing one candidate for their department from a whole batch of applications.

Participants were now provided with fictional feedback at 6-min intervals. In the *mildly negative* condition, this feedback revealed that the department translated 40 texts within an hour; in the conditions *moderately negative* and *strongly negative* the number of texts translated in an hour fell to 30 and 20 texts, respectively. Special care was taken that no one particular coworker appeared to have a particularly strong or weak performance compared to the others. Once a strategy was changed (or, if no change occurred, after 60 min), the experiment was over. However, before informing the participants about the real purpose of the experiment, the experimenters conducted an interview with each participant to find out whether they harbored any suspicions about the experimental manipulations and to extract their hypotheses on the real nature of the experiment. Subsequently, the five participants again gathered in the welcoming room and were thoroughly debriefed about the experiment. Each participant received a payment of 5 Euro (about 7 US \$) and was dismissed with cordial thanks. The two special bonuses were allocated by drawing lots after the experimental series had been completed.

### Results and discussion

In order to check for possible confounds, the dependence of the duration of strategy persistence on factors such as age, gender, course of study, or the specific chosen strategy was examined. No significant effects were found.

#### Strategy persistence dependent on experimental condition

The means and standard deviations for the duration of strategy persistence (in minutes) in the experimental conditions are shown in Table 2. Hypotheses 1 and 2 were tested using planned pairwise comparisons between the three levels of the choice status factor. As predicted, no significant difference between the responsibility condition ( $M = 34.11$ ,  $SD = 15.67$ ) and the no responsibility/preferred choice condition ( $M = 38.11$ ,  $SD = 17.87$ ) was found,  $t(153) = -1.34$ ,  $p = .169$ . Please note that the mean was even somewhat higher in the latter condition than in the former – so the failure to find a responsibility effect within participants preferring the chosen alternative it is not a matter of insufficient sample size and test power. This finding supports Hypothesis 1. For both conditions, strategy persistence was significantly stronger than in the no responsibility/non-preferred choice condition ( $M = 15.89$ ,

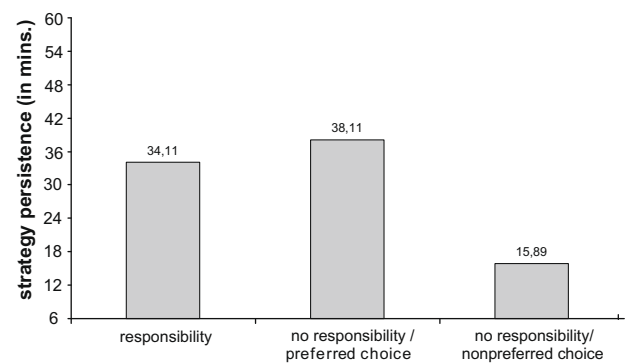


Fig. 1. Average strategy persistence (in minutes) as a function of choice status in Experiment 2.

$SD = 14.33$ ),  $t(153) = 6.29$ ,  $p < .001$  for responsibility,  $t(153) = 7.67$ ,  $p < .001$  for no responsibility/preferred choice. Thus, Hypothesis 2 is also supported by the data. This pattern, which is illustrated in Fig. 1, is fully in line with our preference interpretation of the responsibility effect, because responsibility had no effect over and above the one that preferences had on persistence.

With regard to feedback, we tested for linear effects on persistence by first calculating a main effect in an ANOVA of the full experimental design and then testing for deviations from linearity. The main effect of the feedback factor was significant,  $F(2,153) = 11.00$ ,  $p < .001$ , part.  $\eta^2 = .126$ . Because this factor is measured on an interval-scale (i.e., the difference between mildly negative and moderately negative feedback as well as between moderately negative and strongly negative feedback is 10 texts per hour), we could subsequently test for linear vs. curvilinear feedback effects. Therefore, the linear contrast “–1, 2, –1” for the three factor levels of mildly negative feedback ( $M = 36.67$ ,  $SD = 20.54$ ), moderately negative feedback ( $M = 28.22$ ,  $SD = 18.08$ ), and strongly negative feedback ( $M = 23.22$ ,  $SD = 14.61$ ) was tested for deviations from zero (while controlling for choice status). This contrast was far from significance,  $t(153) = -0.67$ ,  $p = .494$ . Thus, the data can be parsimoniously described with a linear trend, indicating that as fewer texts are translated by their co-workers, participants switch from the chosen strategy earlier. These results support Hypothesis 3.

The interaction between choice status and feedback was insignificant,  $F(4,153) = 1.28$ ,  $p = .280$ , part.  $\eta^2 = .032$ , supporting Hypothesis 4. Taken together, the feedback results indicate that the participants took the feedback into account when deciding upon whether and when to switch their strategy. In other words, they did not blindly adhere to their preference nor did they blindly switch strategy if the chosen strategy was a non-preferred one. Instead, they switched relatively early if the feedback was extremely negative and persisted with the strategy for a relatively long time if the feedback was close to the target. At the same time, in each of the three feedback conditions a preferred strategy had an advantage over a non-preferred one, in that the former was adhered to for longer than the latter. Thus, even extremely negative feedback sets no limiting condition for preference effects.

Table 2

Means for strategy persistence (in minutes) dependent on choice status and feedback in Experiment 2 (standard deviations in brackets).

Choice status	Feedback		
	Mildly negative	Moderately negative	Strongly negative
Responsibility	40.67 (16.79)	35.67 (14.33)	26.00 (12.69)
No responsibility/preferred choice	48.00 (18.18)	37.33 (17.16)	29.00 (13.38)
No responsibility/non-preferred choice	21.33 (17.37)	11.67 (9.08)	14.67 (14.29)

Note:  $N = 18$  in all conditions; the dependent variable can vary between 6 and 60 min.



### Strategy switches

Of the 162 participants, 128 switched the chosen strategy during the course of the experiment. We analyzed the extent to which participants' initial preferences determined the direction of these switches. Therefore, for each participant we rank-ordered the three strategies that had not been initially chosen (either by the participant or by their predecessor) according to that participant's preference ratings. In 13 cases the newly chosen strategy had the same rating as at least one other strategy; these cases are not further considered here. Of the remaining 115 participants, 98 switched to the strategy that had the highest initial preference ranking among the remaining three ones, 13 switched to the strategy initially ranked second-best, and only 4 participants switched to the strategy that initially was their least preferred one. This distribution significantly deviates from equiprobability,  $\chi^2(2, N = 115) = 140.37, p < .001$ . The results show that participants' initial preferences determined the direction of their strategy switches.

### Crossing preferences and responsibility

In Experiment 1 the distribution of preferences allowed for a full orthogonal crossing of responsibility and preference (for the chosen alternative vs. indifference vs. for the non-chosen alternative) although necessarily with very unequal cell sizes. As an inspection of the preference data shows, the same analysis is not possible in Experiment 2, because no participant in the responsibility conditions decided for an alternative that he or she disfavored in the preference ratings. However, four participants in the responsibility conditions and nine participants in the no responsibility/preferred choice conditions rated the chosen strategy and one of its competitors equally high.

With these data, a 2 (responsibility: responsible vs. not responsible)  $\times$  2 (preference status: preference for the chosen strategy vs. indifference) factorial ANOVA was run for persistence. Due to the low cell sizes among the almost indifferent participants, feedback could not be integrated as an additional factor, but was statistically extracted prior to the calculation of the ANOVA (thus, all subsequently reported statistics are corrected for feedback). Preference status had a significant main effect,  $F(1, 104) = 4.05, p = .047$ , part.  $\eta^2 = .035$ . Persistence was higher among participants preferring the chosen strategy ( $M = 37.02, SD = 15.65$ ) than among (almost) indifferent participants ( $M = 29.47, SD = 12.69$ ). In contrast, neither the main effect for responsibility,  $F(1, 104) = 1.57, p = .213$ , part.  $\eta^2 = .015$ , nor the interaction,  $F(1, 104) = 0.17, p = .678$ , part.  $\eta^2 = .002$ , became significant. To allow for comparison with Experiment 1, we also tested for responsibility effects within each of the two preference status levels. Neither among participants preferring the chosen alternative (responsibility:  $M = 35.09, SD = 14.53$ ; no responsibility:  $M = 39.17, SD = 16.71$ ;  $t[104] = -1.30, p = .196$ ) nor among indifferent participants (responsibility:  $M = 21.93, SD = 8.88$ ; no responsibility:  $M = 32.83, SD = 15.81$ ;  $t[104] = -1.19, p = .237$ ) did responsibility have a significant effect on strategy persistence. Please note that the means are even somewhat in the opposite direction, so this lack of responsibility effects is obviously not a problem of sample size and power.

The results of this analysis underline the fact that responsibility no longer has an effect if preferences are held constant. Furthermore, they show that the finding in Experiment 1, where responsibility seemed to selectively heighten escalation among indifferent participants, was most likely due to the fact that indifferent responsible participants chose the alternative they slightly preferred, whereas some of the indifferent non-responsible participants ended up being assigned the slightly non-preferred alternative. In Experiment 2 this difference was controlled for, because all indifferent non-responsible participants were asked to indicate the strategy they slightly preferred and were assigned this

slightly preferred strategy. Under these conditions, no effect of responsibility on persistence was observed anymore.

### Preferences as a mediator for responsibility

The analyses so far have used a categorical preference measure (i.e., a person prefers the chosen alternative, is indifferent, or prefers a different alternative). Although this method of analyzing the data follows from the experimental approach used in Experiment 2, it uses only part of the information that is in the preference data, because differences in preference strength are not considered. Therefore, and to allow for direct comparisons with the results of Experiment 1, we conducted the same mediation analyses as in Experiment 1, thereby testing whether differences between responsible and non-responsible participants in persistence can be explained by the differently strong preferences for the chosen strategy. Thus, in these analyses we did use responsibility instead of choice status as an independent variable. Preference strength for the chosen alternative was measured by taking each participant's rating for the expected success of the chosen strategy and subtracting from it the highest rating for any of the remaining three strategies. As in Experiment 1, high positive values indicate a strong preference for the chosen strategy, zero values indicate (almost) indifference, and the more negative the value becomes, the more the corresponding participant preferred a different strategy than the chosen one. We statistically controlled for feedback effects when conducting the mediation analyses.

If persistence is regressed on responsibility, the weight for responsibility is significant,  $\beta = .19, t(160) = 2.43, p = .016$ , indicating that average persistence was higher among responsible participants ( $M = 34.11, SD = 15.67$ ) than among non-responsible participants ( $M = 27.00, SD = 19.61$ ). This replicates the well-known responsibility effect. Similarly, responsibility receives a significant regression weight when predicting preference strength,  $\beta = .46, t(160) = 4.24, p < .001$ : responsible participants had a clear preference for the chosen strategy ( $M = 2.96, SD = 2.46, t[53] = 8.84, p < .001$  for a test against zero), whereas non-responsible participants were, on average, somewhat in disfavor of the chosen strategy ( $M = -1.28, SD = 4.49, t[107] = -2.96, p = .004$ ). Now, if strategy persistence is regressed on both responsibility and preference strength ( $R^2 = .33, F[2, 159] = 38.93, p < .001$  for the overall regression), the weight for preference strength is significant,  $\beta = .61, t(159) = 8.33, p < .001$ . Thus, even when statistically controlling for responsibility, preference strength is significantly associated with persistence. The bivariate correlations (controlling for feedback) are  $r = .363 (p = .007)$  for responsible participants and  $r = .596 (p < .001)$  for non-responsible participants (because we experimentally manipulated choice status among non-responsible participants, it is not surprising that the correlation is more pronounced in this condition). Most importantly, in this multiple regression the weight for responsibility is no longer significant,  $\beta = -.09, t(159) = -1.20, p = .232$ . A Sobel test indicates that, compared with the first regression, the reduction in the regression weight for responsibility is significant,  $z = 5.10, p < .001$ . Hence, when participants' preferences and the strength of these preferences are controlled, responsibility does not have any additional effect on persistence. Please note that the sign of the regression weight has changed in the second regression – so one cannot argue that a larger sample size and power might still lead to a positive effect of responsibility over and above the effect of preferences.

### Summary

Using a slightly different approach than Experiment 1, the results of Experiment 2 replicate those of Experiment 1 and support our preference interpretation of the responsibility effect. When realizing a no responsibility condition that equals the preferences of the responsible participants (namely preferences for the chosen

alternative) responsibility does not add anything with regard to persistence in comparison with this condition. Only in comparison with a condition of non-responsible participants preferring a non-chosen strategy do we find more persistence among responsible participants and also among non-responsible participants preferring the chosen strategy. An orthogonal analysis of responsibility and preferences (preference vs. [almost] indifference) supported the idea that the action is in the preferences, not in responsibility *per se*. Finally, a mediation analysis showed that responsibility has an effect if preferences are not taken into account, but that this effect is mediated by preference differences between responsible and non-responsible participants, and that it vanishes if preferences are held constant.

## General discussion

The aim of the two experiments presented here was to test a new interpretation of the best-replicated finding in the “escalation of commitment” and “entrapment” literature (e.g., Staw, 1997): the responsibility effect describes the phenomenon that participants who have chosen an alternative themselves invest more resources in that alternative and stick to it for a longer time following negative feedback than participants who have been assigned this alternative do. So far, this effect has predominantly been explained by self-justification (e.g., Brockner, 1992; Staw, 1976). In contrast, we argued that the responsibility effect is the consequence of the participants’ preferences and rational considerations that follow from these preferences. Whereas for responsible participants the chosen alternative is usually the one they prefer, this is only the case for some of the non-responsible participants, namely those whose own preference happens to match that of the decision maker. Since almost any decision theory would suggest that preferring an alternative, that is, considering it more likely to be successful compared to all of its competitors, implies making higher investments in that alternative, the responsibility effect is an almost inevitable consequence of these preferences.

In two experiments with two different paradigms we conducted a competitive test between these two rival explanations. In the first experiment, the participants acted as an R&D director in a fictitious economic decision case, a German adaptation of Staw’s (1976) classical “Adams & Smith” case. In the second experiment, participants played the role of the head of a translation department who had to supervise the performance of her employees and decide upon the translation strategy used by them. In both experiments, we varied whether the initial decision (choice of a division that would gain additional funding in the first experiment, strategy choice in the second experiment) was made by the participants themselves or was assigned to them. Whereas this, so far, is identical to previous research on the responsibility effect, we additionally measured the participants’ preferences for a division or a strategy, respectively, and we measured (Experiment 1) or manipulated (Experiment 2) whether the alternative chosen by their predecessor matched their preferences and how strongly they were in favor or disfavor of the chosen alternative. All participants received negative feedback regarding the success of the initial decision, and we measured how much money they allocated to the previously chosen division (Experiment 1) or how long they persisted with the initial decision (Experiment 2).

Both experiments revealed very consistent results that confirm our preference interpretation: in both experiments, it was shown that responsible participants usually favor the chosen alternative, whereas non-responsible participants, on average, do not. These differences in preference for the chosen course of action mediated the differences in reinvestment or persistence, that is, the more the participants were in favor of the chosen alternative, the more rein-

vestment and persistence occurred, and if preferences were held constant, responsibility did not have an effect on reinvestment and persistence over and above the effect of preferences. Thus, according to our data the responsibility effect is driven by preferences.

To avoid misunderstandings, this does not imply that the responsibility effect does not exist. Its existence has been shown in numerous studies before (e.g., for an overview see, Brockner, 1992), and it has also been significantly shown in both of our experiments: on average, responsible participants made higher reinvestments and showed higher persistence in the face of failure than non-responsible participants did. There is no arguing about that. All we are saying is that there is a very parsimonious explanation for this effect, namely that it is a direct consequence of the fact that responsibility and no responsibility necessarily lead to different preference distributions, and that preferences affect resource reallocations. Simply stated, responsibility fosters escalation and entrapment because it makes you free to choose whatever you prefer.

## Limitations

In both experiments, we only investigated situations with negative feedback. Although this is in line with most studies on escalation of commitment and entrapment, and particularly in line with most replications of the responsibility effect, the original Staw (1976) study contained positive as well as negative feedback. Moreover, the responsibility effect in this study was conceptualized as an interaction between responsibility and feedback, indicating that the tendency of responsible participants to make higher investments than non-responsible participants was more pronounced after negative as compared to positive feedback. Although we did not investigate this interaction, our data exclude the possibility that a responsibility effect could have shown up in the form of such an interaction: because we have shown that following negative feedback the difference between responsibility and no responsibility is completely *eliminated* as soon as preferences are held constant, there is no more room for the possibility that this difference is *smaller* after positive as compared to negative feedback. Nevertheless, it appears promising to extend the present research to situations with positive feedback, because preferences should, at least to some extent, also affect the amount of reinvestment as well as participants’ persistence in situations where things are going well.

Another limitation of our experiments – as well as a limitation of all previous experiments on the responsibility effect that we know of – is that non-responsible participants received some sort of “social cue” by being informed about how another person had decided, whereas responsible participants received no such cue. As a consequence, the objection might be raised that non-responsible participants preferring the selected alternative had a certain “advantage” over responsible decision makers, in that they received validation by their predecessor’s initial decision. Thus, although they might reinvest the same amount of resources as responsible decision makers, they might do this for different reasons. However, Schulz-Hardt, Vogelgesang, Pfeiffer, Mojzisch, and Thürow-Kröning (in press) have recently shown that even receiving validation by *four* persons does not significantly enhance persistence with a failing strategy. And if four validating opinions do not affect reinvestment, it is unlikely that one validating opinion will do so. Furthermore, at least for half of the participants in Experiment 1 and for all participants in Experiment 2 the preference ratings were obtained before information about the predecessor’s choice was given, and these preferences completely mediated the responsibility effect in both experiments. Thus, there hardly seems room left for a social validation effect in this case. Nevertheless, an interesting idea for subsequent experiments could be to in-

form participants in the no responsibility condition that the assigned alternative had been randomly chosen (which, of course, requires a good cover story) and see whether this leads to the same effects as the standard procedure with the (presumed) predecessor.

#### *Implications for the self-justification hypothesis*

As outlined in the introduction, so far the responsibility effect has been predominantly interpreted in terms of self-justification. Obviously, our experiments have not been primarily designed to rule out this self-justification approach to responsibility effects, but rather to test our alternative preference interpretation. Similarly, showing the validity of a particular mediator does not necessarily make other possible mediators invalid. However, in this case we believe that our findings raise at least serious doubts about a self-justification interpretation of the responsibility effect. According to the self-justification view, people who feel responsible for a decision that has led to negative consequences try to justify their initial decision by reallocating more resources to or persisting with the chosen course of action. In contrast, people who were not responsible for choosing this course of action do not have to justify themselves and, hence, make lower reallocations and show less persistence. There is no theoretical reason why this should not work if both groups of people had identical preferences for the chosen course of action and, hence, only differed in the necessity to feel responsible for the consequences. However, as our results show, in this case there is no responsibility effect: non-responsible participants who were assigned a preferred alternative made the same reinvestments and showed the same persistence as participants who had decided for their preferred option and, hence, were responsible.

Of course, a proponent of a self-justification view might argue that those participants who were assigned an alternative they preferred might have felt the need to justify their *preference*, and this pressure might have led to the considerably higher persistence and higher investment in this condition. However, we know from research on dissonance theory that self-justification concerns are stronger after binding decisions than after non-binding preference judgments or recommendations (e.g., Frey, 1981, 1986). Thus, even if self-justification were the driving force among participants in the no responsibility/preferred choice subgroups or conditions in our experiments, we would have expected more persistence and higher reinvestments in the responsibility conditions due to stronger self-justification pressures – which was not the case. Furthermore, in half of the conditions in Experiment 1 the measurement of preferences took place retrospectively, that is, after the dependent variable had been measured. In these conditions, it is very unlikely that participants made higher reinvestments in order to justify a preference which they had not yet been asked about.

Again, it is important to avoid misunderstandings here; even if one would come to the conclusion that the responsibility effect has nothing to do with self-justification, this would not imply that self-justification could not be one of the processes underlying escalating commitment and entrapment. However, our findings imply that it might be fruitful to extend the approach used in this study to other findings that have been interpreted as evidence for self-justification thus far, and similarly test them against alternative, more rational explanations (for an example see Schultze et al., submitted for publication). Studies like this would help to gain more clarity regarding the extent to which self-justification can or cannot explain escalating commitment and entrapment.

#### *The prospect theory interpretation of the responsibility effect*

Self-justification is not the only approach that has offered an explanation for the responsibility effect. The most prominent alter-

native view is the prospect theory perspective on escalating commitment (cf. Whyte, 1986). From this perspective it is argued that responsible participants who made the initial decision themselves experience the subsequent negative feedback as a personal loss, inducing risk-proneness, whereas non-responsible participants should not be in a subjective loss position and should, therefore, be more risk-neutral. It is difficult to judge whether or not our findings are compatible with the prospect theory view. Although the publications so far do not speak to that, one might argue that not only responsibility, but also preferences for a particular course of action might make a person adopt a loss frame if negative feedback occurs. For example, soccer fans often feel very disappointed if their national team loses a match, although they have been “assigned” their nationality. It is, of course, debatable whether this type of disappointment is similar to the loss frame implied by prospect theory (leading to risk-proneness), and it is unclear whether such effects would also be observed in cases of economic investment preferences. Nevertheless, if preferences indeed lead to the adoption of a loss frame in the case of negative feedback (and further studies could test whether this is the case), then prospect theory might be an explanation for why preferences increase reinvestment and persistence in the face of failure.

#### *Implications for the rationality vs. irrationality of escalating commitment and entrapment*

In our experimental tasks, no specific levels of reinvestment or persistence can be considered as the most rational ones. That is why we always spoke of “amount of reinvestment” or “strategy persistence” in our results rather than using terms like “escalation” and “entrapment” that already lead to associations of irrationality. Nevertheless, our results have some implications for the debate on the rationality vs. irrationality of these behaviors (e.g., Bowen, 1987; Brockner, 1992; Northcraft & Wolf, 1984). The escalation of commitment and entrapment literature tends to view persistence and reinvestment following negative feedback as irrational behavior (e.g., Bazerman, 1986; Karlsson, Juliusson, & Gärling, 2005), and the responsibility effect in its self-justification interpretation has always been one of the cornerstones of such a view: if people invest a particularly high amount of money or show strong persistence because they do not want to admit having made a mistake, then they rationalize their decisions instead of deciding rationally.

From the perspective of our preference interpretation, the behavior of these participants seems to be much more rational than has been assumed: if they prefer an alternative, this means that they find this alternative more promising and, therefore, invest more resources in this alternative than those participants who prefer a different alternative and, thus, expect a different alternative to be more successful. This is completely in line with rational decision making based on subjectively expected utility (Edwards, 1954). The fact that they reacted to the levels of feedback in Experiment 2 in a rational manner, that is, they switched strategy earlier the more negative the feedback was, lends additional credit to this interpretation.

Of course, this does not mean that acting on one's preferences always has to be rational. For example, people might selectively expose themselves to information supporting their preference (Frey, 1986; Schulz-Hardt, Frey, Lüthgens, & Moscovici, 2000), or they might fail to attach sufficient significance to evidence showing that their preference is suboptimal and thereby fail to revise such suboptimal preferences (Greitemeyer & Schulz-Hardt, 2003). Nevertheless, in the case of the responsibility effect we do not see sufficient evidence so far to justify classifying it as a violation of rational decision theory.

## Acknowledgements

This research was supported by grants from the German Science Foundation (Deutsche Forschungsgemeinschaft, DFG) to the third author, Project No. Fr. 472/17-1. Experiment 2 was part of Birgit Thurow-Kroening's doctoral dissertation that was supervised by Dieter Frey and Stefan Schulz-Hardt. We would like to gratefully acknowledge the help of Thomas Hacker, Dorothée Haupt, Eva Haussner, Ariane Krogmann, Christoph Meyer zu Kniendorf, Dorothea Schloth, Ute Schlutzen, Julia Schulte-Umberg, Barbara Strauss, and Oliver Walter in carrying out the experiments reported here. We would also like to thank Ramona Bobocel for making available her extensive experimental materials; it provided us with valuable suggestions for the paradigm of Experiment 2. Additionally, we would like to thank Felix Pfeiffer and Thomas Schultze for their support and their helpful comments on an earlier draft of the paper, and the Institute of Psychology of the Christian-Albrechts-University Kiel for making locations and technical support for Experiment 2 available.

## References

- Arkes, H. R., & Blumer, C. (1985). The psychology of sunk costs. *Organizational Behavior and Human Decision Processes*, 35, 124–140.
- Bar-Hillel, M., & Fischhoff, B. (1981). When do base rates affect predictions? *Journal of Personality and Social Psychology*, 41, 671–680.
- Barton, S. L., Duchon, D., & Dunegan, K. J. (1989). An empirical test of Staw and Ross's prescriptions for the management of escalation of commitment behavior in organizations. *Decision Sciences*, 20, 532–544.
- Bateman, T. S. (1986). The escalation of commitment in sequential decision making: Situational and personal moderators and limiting conditions. *Decision Sciences*, 17, 33–49.
- Bateman, T. S., & Zeithaml, C. P. (1989). The psychological context of strategic decisions: A model and convergent empirical findings. *Strategic Management Journal*, 10, 59–74.
- Bazerman, M. H. (1986). *Judgment in managerial decision making*. New York: Wiley.
- Bazerman, M. H., Beekun, R. I., & Schoorman, F. D. (1982). Performance evaluation in a dynamic context: A laboratory study of the impact of prior commitment to the ratee. *Journal of Applied Psychology*, 67, 873–876.
- Bazerman, M. H., Giuliano, T., & Appelman, A. (1984). Escalation of commitment in individual and group decision making. *Organizational Behavior and Human Performance*, 33, 141–152.
- Betsch, T., Haberstroh, S., Glöckner, A., Haar, T., & Fiedler, K. (2001). The effects of routine strength on adaptation and information search in recurrent decision making. *Organizational Behavior and Human Decision Processes*, 84, 23–53.
- Bobocel, D. R., & Meyer, J. P. (1994). Escalating commitment to a failing course of action: Separating the roles of choice and justification. *Journal of Applied Psychology*, 79, 360–363.
- Boulding, W., Morgan, R., & Staelin, R. (1997). Pulling the plug to stop the new product drain. *Journal of Marketing Research*, 34, 164–176.
- Bowen, M. G. (1987). The escalation phenomenon reconsidered: Decision dilemmas or decision errors? *Academy of Management Review*, 12, 548–573.
- Brockner, J. (1992). The escalation of commitment to a failing course of action: Toward theoretical progress. *Academy of Management Review*, 17, 39–61.
- Brockner, J., & Rubin, J. Z. (1985). *Entrapment in escalating conflicts*. New York: Springer.
- Brody, R. G., & Bowman, L. L. (1997). Prior commitment and information recall. *Perceptual and Motor Skills*, 85, 19–27.
- Brody, R. G., & Frank, K. E. (2002). Influence of prior involvement on employees' ratings. *Psychological Reports*, 90, 139–149.
- Conlon, E. J., & Parks, J. M. (1987). Information requests in the context of escalation. *Journal of Applied Psychology*, 72, 344–350.
- Davis, D., & Bobko, P. (1986). Contextual effects on escalation processes in public sector decision making. *Organizational Behavior and Human Decision Processes*, 37, 121–138.
- Drummond, H. (1994a). Escalation in organizational decision making: A case of recruiting an incompetent employee. *Journal of Behavioral Decision Making*, 7, 43–55.
- Drummond, H. (1994b). Too little too late: A case study of escalation in decision making. *Organization Studies*, 15, 591–607.
- Edwards, W. (1954). The theory of decision making. *Psychological Bulletin*, 51, 380–417.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford: Stanford University Press.
- Frey, D. (1981). Reversible and irreversible decisions: Preference for consonant information as a function of attractiveness of decision alternatives. *Personality and Social Psychology Bulletin*, 7, 621–626.
- Frey, D. (1986). Recent research on selective exposure to information. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 19, pp. 41–80). New York: Academic Press.
- Greitemeyer, T., & Schulz-Hardt, S. (2003). Preference-consistent evaluation of information in the hidden profile paradigm: Beyond group-level explanations for the dominance of shared information in group decisions. *Journal of Personality and Social Psychology*, 84, 322–339.
- Kameda, T., & Sugimori, S. (1993). Psychological entrapment in group decision making: An assigned decision rule and a groupthink phenomenon. *Journal of Personality and Social Psychology*, 65, 282–292.
- Karlsson, N., Juliusson, Å., & Gärling, T. (2005). A conceptualisation of task dimensions affecting escalation of commitment. *European Journal of Cognitive Psychology*, 17, 835–858.
- Kirby, S. L., & Davis, M. A. (1998). A study of escalating commitment in principal-agent relationships: Effects of monitoring and personal responsibility. *Journal of Applied Psychology*, 83, 206–217.
- Liao, W. M., Finley, D. R., & Shafer, W. E. (2004). Effects of responsibility and cohesiveness on group escalation decisions. *Advances in Management Accounting*, 13, 245–259.
- McNamara, G., Moon, H., & Bromiley, P. (2002). Banking on commitment: Intended and unintended consequences of an organization's attempt to attenuate escalation of commitment. *Academy of Management Journal*, 45, 443–452.
- Northcraft, G. B., & Wolf, G. (1984). Dollars, sense, and sunk costs: A life cycle model of resource allocation decisions. *Academy of Management Review*, 9, 225–234.
- Ross, J., & Staw, B. M. (1986). Expo 86: An escalation prototype. *Administrative Science Quarterly*, 31, 274–297.
- Schaubroeck, J., & Williams, S. (1993). Behavioral causality orientations and investment decisions following negative feedback. *Journal of Applied Social Psychology*, 23, 1303–1320.
- Schmidt, J. B., & Calantone, R. J. (2002). Escalation of commitment during new product development. *Journal of the Academy of Marketing Science*, 30, 103–118.
- Schoorman, F. D., & Holahan, P. J. (1996). Psychological antecedents of escalation behavior: Effects of choice, responsibility, and decision consequences. *Journal of Applied Psychology*, 81, 786–794.
- Schoorman, F. D., Mayer, R. C., Douglas, C. A., & Hettrick, C. T. (1994). Escalation of commitment and the framing effect: An empirical investigation. *Journal of Applied Social Psychology*, 24, 509–528.
- Schultze, T., Pfeiffer, F., & Schulz-Hardt, S. (submitted for publication). *Information search and self-justification in the escalation paradigm: A critical replication and extension of Conlon and Parks (1987)*.
- Schulz, A. K.-D., & Cheng, M. M. (2002). Persistence in capital budgeting reinvestment decisions – personal responsibility antecedent and information asymmetry moderator: A note. *Accounting and Finance*, 42, 73–86.
- Schulz-Hardt, S., Vogelgesang, F., Pfeiffer, F., Mojzisch, A., & Thurow-Kröning, B. (in press). *Effects of premature praise and forewarning on entrapment in a losing course of action*.
- Schulz-Hardt, S., Frey, D., Lüthgens, C., & Moscovici, S. (2000). Biased information search in group decision making. *Journal of Personality and Social Psychology*, 78, 655–669.
- Staw, B. M. (1976). Knee-deep in the big muddy: A study of escalating commitment to a chosen course of action. *Organizational Behavior and Human Performance*, 16, 27–44.
- Staw, B. M. (1997). The escalation of commitment: An update and appraisal. In Z. Shapira (Ed.), *Organizational decision making* (pp. 191–215). Cambridge, England: Cambridge University Press.
- Staw, B. M., Barsade, S. G., & Koput, K. W. (1997). Escalation at the credit window: A longitudinal study of bank executives' recognition and write-off of problem loans. *Journal of Applied Psychology*, 82, 130–142.
- Staw, B. M., & Hoang, H. (1995). Sunk costs in the NBA: Why draft order affects playing time and survival in professional basketball. *Administrative Science Quarterly*, 40, 474–494.
- Whyte, G. (1986). Escalating commitment to a course of action: A reinterpretation. *Academy of Management Review*, 11, 311–321.
- Whyte, G. (1991). Diffusion of responsibility: Effects on the escalation tendency. *Journal of Applied Psychology*, 76, 408–415.