The Process of Emotion Inference

Matthias Siemer  
University of Miami  

Rainer Reisenzein  
University of Greifswald

Three experiments investigated the process of inferring emotions from brief descriptions of typical eliciting situations, using response time methodology. The initial hypothesis was that emotion inferences are mediated by inferred cognitive appraisals of the eliciting event (concerning e.g., its valence or the responsible agent). This hypothesis was contradicted by the finding of Experiment 1 that emotion judgments are typically made faster than appraisal judgments. To explain this finding, it was hypothesized that emotion judgments are based on automatized (proceduralized) appraisal inferences. This hypothesis was tested in Experiments 2 and 3 using a judgment facilitation paradigm. The results supported the proceduralization hypothesis by demonstrating that appraisal judgments are facilitated by prior emotion judgments.

Keywords: emotions, inference, cognitive processes

Consider the following scenarios: Because of an accident, a friend cannot go on a vacation. You have an appointment with a friend, but she stands you up. Because you are careless at work, a colleague gets hurt. Someone lends you the 50 cents you need to make a call at the pay phone. How would you feel in these situations? More specifically, in which situation would you most likely feel pity, anger, guilt, and gratitude? Numerous studies have shown that people are highly adept at answering such questions. That is, they easily and consensually infer the emotion experienced by a protagonist—who may be the imagined self—from fairly minimal information about the eliciting situation (e.g., Conway & Bekerian, 1987; Reisenzein & Hofmann, 1993; Reisenzein, 1995, Study 3; Siemer, 2001). What is less clear is how people go about performing this task. This question is at the focus of the present article. We report three reaction time experiments that investigated the process of inferring emotions from verbally presented information about eliciting events. The studies were motivated by two considerations. First, the question of how emotions are inferred from information about eliciting events seems highly relevant to several areas of emotion research, but there is little empirical research on it (e.g., Scherer, 1999; Schuster, Rudolph, & Fö rsterling, 1998; van Overwalle, Heylighen, Casaer, & Danicels, 1992). Second, the arguably most straightforward hypothesis about this inference process, and the only one that has so far been empirically tested—namely that people first infer how the eliciting event is appraised by the protagonist and on this basis, infer the protagonist’s emotion—seems to be contradicted by the available data.

We believe that a better understanding of how emotions are inferred from verbally presented, situational information could benefit many fields of emotion research. These include (a) the broader area of research concerned with the inference of emotions from diverse cues (including, apart from situational information, expressive, behavioral, and physiological cues, e.g., Carroll & Russell, 1996; Ekman, Friesen, & Ellsworth, 1972; Scherer, 1986; Scherer, Banse, Wallbott, & Goldbeck, 1991; Schouwstra & Hoogstraten, 1995; Wallbott, 1998); (b) research on the cognitive (i.e., appraisal) processes involved in the generation of emotions (e.g., Frijda, 1993; Lazarus, 1991; Reisenzein, 2001; Smith & Kirby, 2000); (c) research on the influence of anticipated emotions on decision and action (e.g., Bagozzi, Baumgartner, & Pieters, 1998; Baron, 1992; Gilbert, 2000; Mellers, Schwartz, & Ritov, 1999; van der Pligt, Zeelenberg, van Dijk, de Vries, & Richard, 1998); and (d) research on the comprehension of texts (e.g., de Vega, Leon, & Diaz, 1996; Kintsch & van Dijk, 1978; Gernsbacher, Goldsmith, & Robertson, 1992; Gernsbacher & Robertson, 1992; Graesser & Klettke, 2001; Gygax, Oakhill, & Garnham, 2003). Because the present research is primarily rooted in the first two research traditions, we will restrict an elaboration of this point—how the better understanding of emotion inference from situational information could benefit emotion research—to these research areas.

Emotion Inferences From Propositional Information About Eliciting Situations

Research on the problem of emotion inference has traditionally focused on the inference of emotions from “nonpropositional” cues, such as facial expression and prosodic features of verbal expression (e.g., Johnstone, Van Reekum, & Scherer, 2001). However, emotions are frequently also inferred from verbally presented, situational information. Examples are the everyday emotion ascriptions made on the basis of written communications (e.g., letters, books, newspapers) about eliciting situations, as well as emotion inferences from oral communications about eliciting events in interpersonal interaction. Furthermore, it should be noted that, even when information about emotion-eliciting events is obtained through personal observation, this information may ulti-
Emotion Inference and Appraisal Research

Appraisal theories of emotion assume that the emotions elicited by an event depend on how the event is interpreted along a number of appraisal dimensions. These dimensions concern, for example, the goodness-badness of the event, its probability, its focus (whom it primarily concerns), and the responsible agent. For example, one feels pity for another person if one believes that the person is the target of a negative event for which she is not responsible; one feels angry with another person if one believes that this person is responsible for a negative outcome that affects oneself; and one feels grateful to another if one believes that he or she is responsible for a positive outcome that affects oneself (e.g., Ortony, Clare, & Collins, 1988; Weiner, 1986).

Although the assumptions of appraisal theories have been supported in numerous studies, most of these studies investigated remembered emotion-eliciting events (Reisenzin & Hofmann, 1993; Reisenzin & Spielhofer, 1994) or hypothetical scenarios (e.g., Ellsworth & Smith, 1988; Smith & Lazarus, 1993). This research method is based on the assumption that judgments about remembered or hypothetical emotions and appraisals are similar to the corresponding judgments of actual emotions and appraisals. However, this assumption has been questioned (e.g., Parkinson & Manstead, 1992; Parkinson, 1996). One way to evaluate this assumption, and the critique raised against it, is to systematically compare the emotion and appraisal ascriptions made in real and hypothetical situations (Robinson & Clare, 2001).

Another way, pursued in the present article, is to examine in greater depth the process of emotion inference from situational information (cf. also, Smith, 1994). The present research may therefore aid the interpretation of previous work on the link between appraisals and emotion. Beyond that, research on the process of emotion inference from situational information may also shed light on the actual process of emotion generation (cf. Scherer, 1999). However, as this latter point is controversial, we relegate its further discussion to the final section of this article.

The Appraisal Mediation Hypothesis

Appraisal theories of emotion provide not only a plausible account of actual emotion generation. They also suggest a straightforward solution to the problem of emotion inference from information about eliciting events: Emotion inferences are mediated by appraisal inferences. To derive this appraisal mediation hypothesis of emotion inference, only a single additional assumption is needed: namely, that (approximate) knowledge about the relations between emotions and appraisals is a part of folk psychology (see, e.g., Ortony et al., 1988; Weiner, 1995). According to the appraisal mediation hypothesis, the inference of an emotion from situational information thus proceeds in two steps: Presented with a scenario describing an emotion-eliciting event, one first infers the likely appraisals of the event by the protagonist (e.g., by matching the scenario to one or several memory schemas that contain information about the typical appraisals of events of this kind). Second, on the basis of the inferred appraisals, one infers the protagonist’s emotion (by matching the appraisals to emotion concepts that contain inter alia the appraisals typical for the emotion).

To illustrate, consider the following scenario: “Because of an accident, a friend cannot go on a vacation.” Assume that you are asked to decide whether or not you would feel pity in this situation. According to the appraisal mediation hypothesis, this inference requires that you infer your likely appraisal of the situation. Appraisals characteristic of pity include the beliefs that a negative event happened, that this event concerns (primarily) another person, and that the other person is not responsible for his or her fate (e.g., Ortony et al., 1988; Weiner, 1995). In the example scenario, you could infer these appraisals as follows: From your general knowledge about vacations, you infer that not being able to go on a vacation is something negative for your friend. That another person (your friend) is affected by the negative event is directly implied by the description of the situation. And from your knowledge about accidents (which includes that victims of accidents are typically not responsible), you infer that your friend is probably not responsible for his or her fate. Hence, in the first step of the inference process, you infer that the described event is negative, that its focus is the other person, and that the other is not responsible. In the second step, you put the inferred appraisals together and conclude—because the resulting appraisal pattern matches the pattern characteristic for pity—that you would feel pity in this situation.

The appraisal mediation hypothesis of emotion inference has several attractive features. First, it agrees with what is known about the semantics of ordinary language emotion terms: Namely, that their meaning is partly determined by cognitive appraisals (Reisenzin & Hofmann, 1990). For example, “pity” in ordinary language means the feeling typically experienced when a negative event occurs to another person for which he or she is not responsible. Second, the appraisal mediation hypothesis is consistent with the findings of previous studies of emotion inference from verbally presented scenarios; in particular with the finding that, if the participants inferred a particular emotion, they also tended to infer that the pattern characteristic for this emotion was present (e.g., Reisenzin & Hofmann, 1990; Roseman, 1991; Smith & Lazarus, 1993).
Third, the appraisal mediation hypothesis agrees with what is, as far as we can see, the (still) dominant view of the ascription of mental states and traits: Namely, that these ascriptions are the results of theory-based inference processes (e.g., Perner, 1996; Smith, 1994). Finally, of course, the appraisal mediation hypothesis provides a theoretical justification for studying the appraisal-emotion link using hypothetical scenarios (see the General Discussion).

However, the appraisal mediation hypothesis faces a problem: It conflicts with reaction time data on emotion and appraisal inferences. If emotion inferences are indeed made as described by the appraisal mediation hypothesis, then they are necessarily more complex than appraisal inferences in the sense that they require more processing steps. Assuming that complex inferences take more time than the inferences of which they are composed, the appraisal mediation hypothesis therefore predicts that emotion judgments take more time than appraisal judgments. Conversely, if appraisal judgments were found to take more time than emotion judgments, the appraisal mediation hypothesis would be disconfirmed; for “judgments that take longer to arrive at cannot plausibly mediate or come prior to judgments that take a shorter time” (Smith & Miller, 1983, p. 493; Reisenzein, Debler, & Siemer, 1992). Note that these conclusions hold even if appraisal inferences are made in parallel. For even then, the emotion inference cannot be made faster than the slowest mediating appraisal inference (cf. Schweickert, 1983).

Unfortunately, the few studies that tested the described prediction of the appraisal mediation hypothesis (Schuster et al., 1998; van Overwalle et al., 1992) consistently failed to support it. Schuster et al. (1998) found that subjects presented with cause-effect covariation information (cf. Kelley, 1967) needed more time to judge locus of causality and responsibility than to judge emotions such as anger and sympathy that are presumably mediated by causal attributions (cf. Reisenzein, 1986; Weiner, 1986). The same finding was obtained when the participants were additionally given concrete attributional information about degree of effort, ability, and task difficulty. Similarly, van Overwalle et al. (1992) found no significant difference in the time needed to judge shame and guilt given attributional information (concerning locus of causality, stability, globality, and controllability, see Weiner, 1986) and preattributional information (about consensus, consistency, and effectuality). Finally, in two studies on mood influences on appraisal and emotion judgments in hypothetical scenarios, Siemer (2001) found that judgments about anger and sadness were made faster than judgments about responsibility and ability to cope, the appraisals presumably underlying these emotions.

Aims of the Present Research

The aim of Experiment 1 was to replicate and extend the previous findings concerning the relative reaction times of emotion and appraisal inferences in hypothetical scenarios. A reaction time paradigm similar to those used in previous research on attributions and social judgments was used (e.g., Graham & Brown, 1988; Reisenzein et al., 1992; Smith & Miller, 1983). The aim of Experiments 2 and 3 was to investigate an alternative to the appraisal mediation model of emotion inference that is able to explain the RT data. According to this alternative theory, the shorter RTs of emotion judgments, compared with appraisal judgments, reflect the automatization or proceduralization of emotion inferences. This hypothesis, as well as the experimental paradigm used to test it, is described in more detail in the Introduction to Experiment 2.

Experiment 1

Method

Participants

The participants of Experiment 1 were 73 introductory psychology students who participated in partial fulfillment of their study requirements. Fifty-five of the students were female, mean age was 25.1 years.

Stimuli

The emotions pity, anger, guilt, and gratitude were selected for study (e.g., Weiner, 1986). To protect against theoretically irrelevant idiosyncrasies of specific eliciting situations, 14 different scenarios were constructed for each emotion (the exact number was dictated by the logic of the design; see below in the design section). Each scenario consisted of a single sentence that described a typical eliciting event for the given emotion. The descriptions contained no explicit appraisal information, although emotion-relevant appraisals were typically easy to infer. Example scenarios are: “The parents of your friend die in a car accident” (pity); “Part of a manuscript distributed by the lecturer is unreadable” (anger); “You are careless at work; a colleague gets hurt” (guilt); and “A friend listens sympathetically to a problem of yours” (gratitude; the complete list of the scenarios is contained in Appendix A). For each scenario, all four emotion judgments had to be made (e.g., “I feel the following emotion: pity.”). In addition, nine appraisal dimensions, selected on the basis of theoretical proposals as well as empirical results (Reisenzein & Spielhofer, 1994) were assessed: subjective event-evaluation (positive/desirable vs. negative/undesirable), moral event-evaluation (morally right/fair/deserved vs. morally wrong/unfair/undeserved), self-evaluation (self positive vs. self negative), other-evaluation (other positive vs. other negative), importance of the event (important/unimportant), expectedness of the event (expected/unexpected), responsibility for the event (self responsible vs. other responsible), controllability of the event (controllable vs. uncontrollable by self), and focus of the event (concerns primarily myself vs. concerns primarily someone else). For a more detailed explanation of these appraisal dimensions and literature references, readers are referred to Reisenzein and Spielhofer (1994). For each appraisal dimension, two items were formulated that referred to the two opposed dimension values. For example, the two items for the importance dimension were: “That’s important for me” and “That’s unimportant for me” (see Appendix B for the complete item list). Both poles of the appraisal dimensions were used to ensure that about half of the “correct” answers to each scenario were negative, which prevented the adoption of a simple response set (saying “yes” all the time). Finally, for each scenario, a control question was included that referred to the 1st word of the scenario sentence (e.g., “This sentence begins with ‘I’”). Taken together, then, 23 questions had to be answered for each scenario: 4 emotion items, 9 positive and 9 negative appraisal items, and 1 control question.

Design

The design of Experiment 1 was based on a modified Latin square. Each participant received all 4 (emotions: pity, anger, guilt, gratitude) appraisals presumably underlying these emotions.
gratitude) × 23 (questions) = 92 stimulus combinations. However, within each emotion, the 23 questions were balanced across the 14 scenarios for this emotion in such a way that each participant was presented with a given scenario only once (emotion and control questions) or twice (appraisal questions). More precisely speaking, there were 14 different subject groups. These groups were combined with the 14 scenarios for each emotion (e.g., pity) and 14 of the questions (the 4 items asking for the presence of the emotions, the 9 items asking for “positive” appraisals [e.g., “important”, and the control question) according to a 14 (subject group) × 14 (scenarios) × 14 (question) Latin square. In addition, the 9 scenarios combined with “positive” appraisals (e.g., “important”) were presented a second time together with the negative pole of the same dimension (e.g., “unimportant”), making for a total of 23 stimulus combinations for each emotion. This procedure was repeated for each of the four emotions (pity, anger, guilt, and gratitude), resulting in the 92 stimulus combinations presented to each subject. Participants were randomly assigned to one of the 14 groups (scenario-question combinations) with the aim of obtaining about 5 participants per group. Because the randomization procedure was unconstrained, different scenario-question combinations were judged by 3–7 participants. However, pooled across the 14 scenarios for each emotion type, 73 responses were available for each question.

**Procedure**

The participants were tested individually or in small groups in a laboratory equipped with several PCs. The experiment was described as being concerned with judgments of social situations. The participants were first handed a list of the appraisal questions together with an explanation of their meaning (see Reisenzein & Spielhofer, 1994, Appendix B). They were asked to study the list carefully and to indicate if they had difficulties understanding a question (in which case the experimenter gave additional explanations). Subsequently, the participants were instructed about the computerized judgment task. In each trial, a scenario description would be presented on the computer screen; they were asked to carefully read this description and then to press the ENTER key. Following each scenario, a question would be presented which was to be answered by pressing the “1” or “2” key (labeled “yes” and “no,” respectively) on the numerical keypad. The participants were instructed to answer the questions carefully but without undue hesitation (cf. Fazio, 1990); they were not, however, urged to respond as quickly as possible.

Following the instructions, 10 practice trials (using 10 separate scenarios) were run. Subsequently, any remaining questions were answered, and the main experiment was started. Each participant was presented with the 92 scenario-question combinations corresponding to his or her respective experimental condition in an individually determined random order. Each trial started with the presentation of a one-sentence scenario. After the participant confirmed that the sentence had been read and understood, a question was presented. To control for possible effects of different question lengths, the questions were presented letter by letter in average reading speed (50 ms per letter). The computer recorded the question answering time (the time from the beginning of the question presentation to the pressing of the “yes” or “no” keys) in milliseconds. The next scenario was presented with a delay of 1 second. Not counting the instruction and the practice trials, the experiment lasted from 7 to 15 min (M = 10 min).

**Results**

**Judgments (Percentage of Yes-Responses)**

Table 1 shows the percentage of yes-answers for all questions for which more than 50% endorsements were obtained (50% being

<table>
<thead>
<tr>
<th>Judgment</th>
<th>Pity %</th>
<th>Pity ms</th>
<th>Anger %</th>
<th>Anger ms</th>
<th>Guilt %</th>
<th>Guilt ms</th>
<th>Gratitude %</th>
<th>Gratitude ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target emotion</td>
<td>.92</td>
<td>1121</td>
<td>.93</td>
<td>1010</td>
<td>.85a</td>
<td>1061</td>
<td>.95</td>
<td>881</td>
</tr>
<tr>
<td>Event evaluation: negative</td>
<td>.88</td>
<td>1183</td>
<td>.90</td>
<td>1228</td>
<td>.84</td>
<td>1493</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event evaluation: positive</td>
<td>.67</td>
<td>1683</td>
<td>.67</td>
<td>1725</td>
<td>.73</td>
<td>1668</td>
<td>.89</td>
<td>1577</td>
</tr>
<tr>
<td>Moral evaluation: unfair</td>
<td>.64</td>
<td>1890</td>
<td>.60</td>
<td>2021</td>
<td>.59</td>
<td>2200</td>
<td>.51</td>
<td>1602</td>
</tr>
<tr>
<td>Event uncontrollable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Responsibility: other</td>
<td>.66</td>
<td>2658</td>
<td>.52b</td>
<td>1940</td>
<td>.89</td>
<td>937</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Responsibility: self</td>
<td>.68</td>
<td>1603</td>
<td>.70</td>
<td>1547</td>
<td>.63</td>
<td>1585</td>
<td>.71</td>
<td>1576</td>
</tr>
<tr>
<td>Focus: self</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-evaluation: positive</td>
<td>.55</td>
<td>1472</td>
<td></td>
<td></td>
<td>.55</td>
<td>1701</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-evaluation: negative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.79</td>
<td>1419</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other-evaluation: negative</td>
<td>.56</td>
<td>1780</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.89</td>
<td>1103</td>
</tr>
<tr>
<td>Event importance: important</td>
<td>.68</td>
<td>1460</td>
<td>.70</td>
<td>1476</td>
<td>.79</td>
<td>1720</td>
<td>.86</td>
<td>1270</td>
</tr>
<tr>
<td>Expectedness: unexpected</td>
<td>.79</td>
<td>1516</td>
<td>.81</td>
<td>1533</td>
<td>.79</td>
<td>1720</td>
<td>.68</td>
<td>1925</td>
</tr>
</tbody>
</table>

a anger = 53%. b M = .61 without anger-scenarios 1 and 11 (see Appendix A).
the minimum value for regarding an appraisal as “typical” for a given emotion). As can be seen, the percentage of correct attributions of the target emotion in the respective scenarios was generally high: Pity, anger, and gratitude had endorsement percentages > 90%; for guilt, endorsement was 85%. In addition, again with the exception of guilt, the emotion ascriptions were specific to the target emotions, that is, no other emotion was checked with a frequency > 50%. As to the guilt scenarios, the subjects reported in 53% of the cases that they would also feel anger. Examination of the guilt scenarios with particularly high anger endorsements (e.g., guilt-scenario 11, cf. Appendix A) suggests that these anger judgments probably referred to self-directed anger.

Turning next to the appraisal questions, it can be seen from Table 1 that each emotion is characterized by a characteristic appraisal profile. For example, the pity-eliciting events were typically judged as negative (88%), unexpected (79%), and important (68%); as concerning primarily others (focus, 68%), as events for which others or circumstances were responsible (66%), and as uncontrollable by oneself (64%). These findings largely replicate previous results by Reisenzein and Spielhofer (1994; see also Reisenzein & Hofmann, 1990) and are generally in line with the results of other appraisal research (e.g., Roseman, 1991; Roseman, Antoniou, & Jose, 1996; Weiner, 1986). This replication of previous findings was, of course, highly desired for the main aim of the present study, the test of the appraisal mediation hypothesis. In addition, the endorsements of the target emotions in the typical eliciting situations (e.g., pity in pity-situations) were for the most part as high as or higher than the endorsements of the corresponding appraisals (see Table 1). This finding, too, replicates previous results. For instance, in previous appraisal research, participants were often asked to recall autobiographical situations in which they experienced a particular emotion. This instruction implies a (hypothetical) endorsement of 100% for the target emotions. By contrast, the endorsements of appraisals were typically < 100% and in some cases considerably lower (e.g., Reisenzein & Spielhofer, 1994; Roseman, 1991).

**Response Latencies**

An important issue in the analysis of reaction times (RTs) in social judgment tasks is the identification and appropriate handling of outliers (e.g., Fazio, 1990). In the present study, this issue was addressed as follows: the RTs were first standardized within subjects, and the resulting scores were then once more standardized within questions. RTs with resulting z-scores > 3 were replaced by the question mean. This procedure ensured that preexisting differences in response latency, both between subjects and between items (questions) were taken into account in the process of outlier detection. In all, 2.2% of the responses (about 2 per subject) were identified as outliers. They were unsystematically distributed across the experimental conditions. In addition, very fast answers (< 200 ms) were also replaced by the question mean because they were considered chance responses (1.1% of the responses).

Table 1 shows the RTs of the emotion judgments, as well as the RTs of those appraisal judgments for which more than 50% yes-answers were obtained and which therefore can be regarded as at least minimally typical for the respective emotions. Considering first the RTs of the judgments of the target emotions (e.g., pity in the pity scenarios), we found that they differed significantly from each other, $F(3, 216) = 3.2, p = .023$, with gratitude judgments taking the least time, and pity judgments taking the most time. As to the latencies of the appraisal judgments, only responsibility, $F(2, 144) = 18.6, p < .001$, and unexpectedness, $F(2, 144) = 3.9, p = .021$ differed significantly between scenario types. Responsibility judgments were fastest in the guilt scenarios and slowest in the pity scenarios, whereas unexpectedness judgments were slowest in the guilt scenarios and fastest in the pity scenarios (see Table 1). Clearly nonsignificant differences between scenario types were obtained for subjective event-evaluation, moral evaluation, controllability, focus, and self evaluation (all $p$s > .4). To compare the appraisal judgments with the emotion judgments, 2 (emotion vs. appraisal judgment) × 4 (emotion type) repeated-measurement ANOVAs were computed separately for each appraisal judgment. For example, in the ANOVA involving event evaluation, the latency of event evaluation in the pity scenarios was compared with the latency of the pity judgment; the latency of event evaluation in the anger scenarios was compared with the latency of the anger judgment, and so on. The results of these analyses are shown in Table 2. As can be seen, the main effect of judgment type (emotion vs. appraisal) was significant in all ANOVAs (i.e., for all appraisals); the main effect of emotion type (pity, anger, guilt, gratitude) was significant for importance, responsibility, and unexpectedness; and the interaction was significant for responsibility and unexpectedness.

### Table 2

**Two-Way Repeated Measurement ANOVAs for Judgment Latencies, Experiment 1**

<table>
<thead>
<tr>
<th>Appraisal</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Emotion vs. Appraisal (EA)</td>
</tr>
<tr>
<td></td>
<td>Emotion Type (ET)</td>
</tr>
<tr>
<td></td>
<td>EA × ET</td>
</tr>
<tr>
<td>Subjective event evaluation: positive/negative</td>
<td>$F(1, 72) = 7.51, p &lt; .01$</td>
</tr>
<tr>
<td>Moral evaluation: fair/unfair</td>
<td>$F(1, 72) = 22.68, p &lt; .001$</td>
</tr>
<tr>
<td>Responsibility: self/other</td>
<td>$F(1, 72) = 32.09, p &lt; .001$</td>
</tr>
<tr>
<td>Focus: self/other</td>
<td>$F(1, 72) = 46.25, p &lt; .001$</td>
</tr>
<tr>
<td>Self-evaluation: positive/negative</td>
<td>$F(1, 72) = 22.20, p &lt; .001$</td>
</tr>
<tr>
<td>Other evaluation: positive/negative</td>
<td>$F(1, 72) = 5.53, p &lt; .05$</td>
</tr>
<tr>
<td>Event importance: important/unimportant</td>
<td>$F(1, 72) = 37.13, p &lt; .001$</td>
</tr>
<tr>
<td>Expectedness: expected/unexpected</td>
<td>$F(1, 72) = 58.4, p &lt; .001$</td>
</tr>
</tbody>
</table>

**Note.** For guilt scenarios, only judgments of uncontrollability were considered. ANOVAs = analyses of variance.
significant for unexpectedness. Of primary interest is the significant main effect of judgment type. This finding means that the emotion judgments were on average (i.e., across the four emotion/scenario types) significantly faster than each single appraisal judgment (cf. Table 1).²

Discussion

The results of Experiment 1 replicated and extended previous findings which suggested that emotion judgments need less time than appraisal judgments. In the present study, this was found to be true even for appraisals that are consensually regarded as essential for the respective emotions by cognitive emotion theorists, such as a positive event-evaluation in the case of gratitude; other-focus in pity; other-responsibility in (other-directed) anger; or negative self-evaluation in guilt (see, e.g., Ortony et al., 1988; Roseman, 1984; Weiner, 1986). These appraisals were also endorsed, for the emotions in question, by most of our participants (see Table 1). Thus, although we again found, as did previous research, that the situations conducive to pity, anger, guilt and gratitude are appraised in characteristic ways, the RT results of Experiment 1 are again in conflict with the appraisal mediation hypothesis of emotion inference.

This raises the question of how emotion inferences are made, if they are not mediated by inferred appraisals; and how the obtained RT results can be explained. In Experiments 2 and 3, we examined one possible explanation of the findings: Namely, that the inferences assumed by the appraisal mediation hypothesis have become partly automatized or proceduralized. This hypothesis is described next.

The Proceduralization Hypothesis

The proceduralization hypothesis makes two assumptions: (1) Although emotion inferences from situational information are initially mediated by inferred appraisals, as a result of being highly practiced, they have become automatized or “proceduralized” (Anderson, 1987; Branscombe, 1988; Smith, 1994). Anderson assumes that the intermediate steps of a proceduralized inference process (here: the appraisal inferences) are unconscious (not introspectively accessible) and fast. In the extreme case, Anderson (1987) suggests, the intermediate steps can even become entirely omitted through “knowledge compilation,” a special process that operates on the inference procedures (Anderson, 1987).² (2) By contrast, when asked to judge appraisals directly, people rely on a nonautomatic, declarative inference procedure that is accessible to consciousness and comparatively time-consuming. The proposed model of emotion inference is based on the general model of the automatization of cognitive skills developed by Anderson (1983). For this reason, the model is necessarily similar to other domain-specific models of cognitive processes and their automatization that are based on Anderson’s theory. Although the present model of emotion inference thus claims no originality with respect to the general processes of automatization that it posits, it contributes to current theorizing on emotion (and appraisal) inferences in at least two ways: (a) inspired by Anderson’s theory, the present model features mechanisms of automatization, in particular the mechanism of knowledge compilation, that have not been previously considered in the literature on emotion and appraisal inferences, and (b) in contrast to dual-process models of cognition (e.g., Smith & Neumann, 2005) that view automatic and controlled processes as independent processes with fundamentally different properties, the present theory allows to model the transition from deliberate to automatic inferences as a continuous (and possibly reversible) process.

It also has to be noted that some recent models of automatic appraisals in emotion were independently developed (e.g., Clore & Ortony, 2000; Ellsworth & Scherer, 2003; Smith & Kirby, 2001; Smith & Neumann, 2005). Inasmuch as these authors attribute properties to automatic appraisals (e.g., that automatic appraisals are fast and unconscious) that are based on, or in line with, general theories of the automatization of cognitive processes, our model of emotion inference is necessarily also similar, in these respects, to theories of automatic appraisals. However, note that the present model deals with emotion inference, not with emotion generation. Even if both processes are subject to the same general mechanisms of automatization, this does not mean that they are also identical in other respects. Therefore, to keep our presentation straightforward, we will for now focus exclusively on the process of emotion inference and relegate a more systematic comparison of the present model of emotion inference with models of emotion generation to the General Discussion.

To see specifically how knowledge compilation can lead to the proceduralization of originally appraisal-mediated emotion inferences, it is helpful to describe the process of emotion inference in the language of a production system (cf. e.g., Anderson, 1983; Smith, 1994). Within this theoretical framework, the inference of an emotion from situational information would be regarded as a goal-directed process whose aim is to determine—in inference situations like those used in Experiment 1—whether or not the protagonist of a given scenario experiences a given emotion (for short, whether the emotion fits the scenario). As mentioned, we assume that the appraisal mediation hypothesis correctly describes how emotions are initially (i.e., prior to proceduralization) inferred from situational cues. In the language of a production system, this original two-step inference process can be described in general terms as follows (cf. Smith, 1994):

I₁ (Inference of emotions from appraisals): IF the goal is to determine whether <person P> experiences <emotion E> in <situation S>, THEN conclude <person P> has <emotion E> in <situation S>

If the appraisals of the situation are known, the production I₁ initiates a subproduction I₂ whose aim is to determine the appraisals a₁ to aₙ. This latter inference procedure can be described as follows:

I₂ (Inference of appraisals from situational cues): IF the goal is to determine whether <person P> makes appraisals a₁ to aₙ of <situation S>, THEN conclude <person P> makes appraisals a₁ to aₙ of <situation S>

² To guard against the danger that the results were biased by differing proportions of yes-responses obtained for the compared judgments, the analyses were repeated for yes-responses only. The results were largely the same. Details of these analyses can be obtained from the first author on request.

³ Since we cannot distinguish between these two possibilities on the basis of our data, we will in the following allow for both mechanisms of automatization as accounts for our findings.
and \(<\text{situation } S>\) has features \(<f_1>, <f_2>, \ldots, <f_n>\)

THEN conclude \(<\text{person } P>\) makes appraisal \(a_1, a_2, \ldots, a_n\)

\(<a_n>\) of \(<\text{situation } S>\)^4

The inferred appraisals are then returned to procedure II, where they are used to infer the emotion \(E\) from the appraisals.

Anderson (1983; 1987) proposed that if a sequence of productions (here: the sequence \(I1\rightarrow I2\)) that is in the service of the same processing goal (here: emotion inference) is frequently repeated, a new production is created that directly connects the conditions of the first with the outcomes of the second production. In the present case, a new production \(I3\) would be created that combines \(I1\) and \(I2\) as follows:

\(I3:\) IF the goal is to determine whether \(<\text{person } P>\) has \(<\text{emotion } E>\) in \(<\text{situation } S>\)

and \(<\text{situation } S>\) has features \(<f_1>, <f_2>, \ldots, <f_n>\)

THEN conclude \(<\text{person } P>\) has \(<\text{emotion } E>\) in \(<\text{situation } S>\)

Hence, as a consequence of knowledge compilation, emotion inferences are based directly on the situational features that allow the inference of emotion-specific appraisals, and are no longer causally mediated by inferred appraisals.

Finally, automatic processes are usually less error prone than controlled processes (e.g., Bargh, 1994). This property of automatic processes could in part explain the different endorsement percentages of the appraisal and emotion judgments obtained in Experiment 1.

The proceduralization hypothesis is able to account for all of the existing data on emotion inferences from situational information. (1) At minimum, the proceduralization hypothesis explains why emotion inferences do not need more time than appraisal inferences: This is so because in proceduralized emotion inferences, the mediating appraisal inferences occur very quickly or are even completely omitted (cf. our first assumption). (2) If one also accepts our second assumption that when asked to make appraisal judgments, people rely on a more time-consuming, nonautomatic inference procedure, then one can also explain why emotion judgments can even be faster than appraisal judgments. (3) Finally—like the appraisal mediation hypothesis—the proceduralization hypothesis explains why scenarios are consensually judged to elicit a particular emotion are also consensually judged as being appraised in a particular way: This is so because the situational schemas that are used to infer particular emotions in a direct, appraisal-unmediated way still contain the features that allow the inference of the appraisals characteristic for these emotions.

Experiment 2

The proceduralization hypothesis assumes that the inference of an emotion from information about an eliciting event involves the matching of the eliciting event to situational schemas that also allow the inference of the appraisals characteristic for the emotion.

This assumption was tested in Experiment 2 using a variant of the so-called judgment facilitation paradigm, originally proposed by Bassili and Racine (1990) for the investigation of the temporal sequence of person versus situation (internal-external) attributions of behavior (cf. Smith & Miller, 1983). The judgment facilitation paradigm rests on the following assumption (cf. Bassili & Racine, 1990): If, in the process of making a judgment, information relevant to a second, subsequent judgment is activated in memory, less time is needed for the second judgment. In other words, if the first judgment activates information relevant to the second, then the second judgment is facilitated by the first. For example, if the information activated in the process of making an emotion judgment is the same or partly the same as that accessed when making an appraisal judgment, then the appraisal judgment will be facilitated if it is preceded by the emotion judgment.5

If one combines the logic of the judgment facilitation paradigm with the proceduralization hypothesis of emotion inference, two testable predictions can be derived:

1. Inferences of the appraisals of a situation are facilitated by the prior inference of an emotion for which the appraisals are characteristic. For short, appraisal judgments (AJs) are facilitated by corresponding emotion judgments (EJs). The reason is that the situational features matched to an emotion schema during emotion inferences are identical to the features accessed for the corresponding appraisal inferences.

2. EJs are also facilitated by corresponding AJs, but to a lesser degree. The reason is that each emotion is characterized by a pattern of appraisals, and that a single appraisal judgment activates only a subset of the situational features that are accessed in the process of emotion inference.

Method

Participants

The participants were 48 introductory psychology students who participated in partial fulfillment of their study requirements. Thirty-two of the participants were female; mean age was 24.2.

---

4 Production rule I2 combines the inference rules for single appraisals of the form: IF the goal is to determine whether \(<\text{person } P>\) has \(<\text{emotion } E>\) of \(<\text{situation } S>\) and \(<\text{situation } S>\) has feature \(<f_j>\), THEN conclude \(<\text{person } P>\) makes appraisal \(<a_i>\) of \(<\text{situation } S>\) and so forth. Note that \(f_j\), is typically a combination of more elementary features \(<f_{i_1}, f_{i_2}, \ldots, f_{i_n}>\).

5 The judgment-facilitation paradigm is also well suited to rule out a possible alternative account of the results of Experiment 1. Before a person can make an emotion or appraisal inference from situational information, he or she must first translate the question into some internal representational format. It might be possible that, compared with the emotion questions, the participants of Study 1 needed more time to translate the appraisal questions into internal representations. However, a significant advantage of the judgment facilitation paradigm is that facilitation effects are by nature relative effects. That is, absolute RT is irrelevant, as only RTs to the same questions under different conditions (following or not following a potentially relevant first question) are compared. This leaves the time to translate the questions into some internal representational format constant. In other words, the time to grasp the meaning of the questions is effectively subtracted out of the relevant comparisons.
Materials

The emotions of anger, guilt, and gratitude were selected for study. For each emotion, 24 one-sentence scenarios were constructed; hence, a total of 72 scenarios were used. They included most of the scenarios for these emotions from Experiment 1 (see Appendix A). For each scenario, judgments of the target emotion and judgments of the eliciting event on four appraisal dimensions (event-evaluation, responsibility, fairness/deservedness/moral evaluation, and focus) were made. All appraisal questions asked for that pole of the respective appraisal dimension for which the theoretically expected answer was positive. In addition, to prevent the use of a simple acquiescence response strategy, two control questions were included.

Procedure

Participants were tested individually or in small groups. The instructions mirrored those used in Experiment 1 with 2 exceptions. (a) Different from those used in Experiment 1, Experiment 2 used a cue-word technique to control for possible effects of question length. That is, the participants were asked to memorize brief cue-words that represented the questions. POSITIVE versus NEGATIVE served as cues for positive versus negative event evaluation; SELF versus OTHER/SITUATION as cues for the responsibility appraisal; FAIR versus UNFAIR as cues for the fairness appraisal, CONCERNS ME versus CONCERNS OTHER/S as cues for the focus appraisal; and ANGER, GUILT, and GRATITUDE as cues for the emotion questions. (b) Also different from Experiment 1, the instructions emphasized accuracy and speed of the answers. That is, the participants were asked to respond as fast as they could without making mistakes. We thought that this instruction would reduce the danger that the subjects used different speed-accuracy trade-offs for appraisal and emotion judgments (e.g., slowed down their appraisal judgments to reach higher accuracy).6

After the instructions and five practice trials, the experiment was started. In each trial, a scenario was first presented in the same manner as in Study 1. Participants pressed the return key to indicate that the scenario had been read and understood. Subsequently, the cue-word for the first question was presented and the participants responded by pressing the “yes” or “no” key. The cue-word for the second question was shown 300 ms later. If the first question had been an emotion question, the second question was an appraisal or control question (i.e., a question referring to a nontarget emotion or an oppositely poled appraisal dimension) and vice versa. Which of the four appraisal questions and the two nontarget emotion questions served as the control item for a given scenario was randomly determined. The next scenario was presented 300 ms after the second response. This procedure continued until all scenarios had been presented. The order of the scenarios was individually randomized. The entire scenario-rating task lasted from 5 to 9 minutes.

Design

Four factors were varied, all of them within-subjects. The first factor, emotion type (with levels anger, guilt, and gratitude) represents the type of emotion suggested by a scenario. As demanded by the logic of the judgment facilitation paradigm, each scenario was followed by two questions. The second factor, question type (with levels emotion and appraisal) codes whether the question was an emotion or an appraisal question. The third factor, judgment position, represents the position in which the emotion or appraisal question was asked after the scenario (first position vs. second position). The fourth factor, appraisal type (with levels event-evaluation, fairness, controllability, and focus) refers to the kind of appraisal question that was presented together with (i.e., before or after) an emotion question.

The combination of these factors resulted in five different questions (one emotion question and four appraisal questions), each of which could occur either in the first or in the second position after a scenario. If an emotion question occurred in the first position, the second question was an appraisal question, and vice versa. In addition, to prevent adoption of a simple acquiescence response style, two control questions were included for each scenario, one asking for an unpredicted pole of an appraisal dimension and the other for an unpredicted emotion. Which of the four appraisal dimensions and which of the two unpredicted emotions served in the control questions for each scenario was randomly determined. Because the two control questions were never paired with each other, this resulted in 12 question combinations.

The 12 question conditions for each emotion type were then combined with the $2 \times 12 = 24$ scenarios for this emotion by means of 2 Latin squares of size 12, stacked on top of each other. Each of the 12 rows of this stacked Latin square represents a different experimental group that received the 12 question combinations paired with a different combination of (2) scenarios. This procedure was repeated for each emotion. As a consequence, there were 12 experimental groups, each of which received the 36 combinations of the 12 question conditions and the 3 emotion types with a different selection of scenarios, 2 for each combination (hence, each participant judged 72 scenarios). Participants were randomly assigned to the 12 groups, with 4 participants per group.

Results

Prior to the subsequent analyses, the answers and response times of each participant within experimental conditions (12 question conditions $\times 3$ emotion types) were aggregated. With the exception of the 1st-placed emotion questions, all aggregated answers and reaction times were based on $N = 2$ responses (scenarios). As concerns the 1st-placed emotion questions, the answers and RTs were collapsed across all appraisal conditions ($N = 12$).

---

6 This is a deviation from the original judgment facilitation paradigm. Bassili and Racine (1990) did not inform subjects that the responses were timed, arguing that because their judgments (situation and person attributions) did not have “correct” answers, they did not want to pressure subjects into giving quick, unreflective answers. However, this reasoning does not apply to the present experiment because the judgments studied here do have a “correct” (at least in the sense of consensually shared and theoretically predicted) answer. Furthermore, the proportions of yes-responses obtained in Experiment 2 were very similar to those obtained in Experiment 1, indicating that quality of judgments was not affected by the additional speed instruction in Experiment 2.
Judgments

The proportion of yes-answers obtained for the emotion and appraisal questions, and the statistical comparisons for first-versus second-place judgments are summarized in Tables 3 and 4, respectively. As can be seen from Table 3, emotion ascriptions were high in all conditions. The position of the emotion judgments (first or second) had no effect on the endorsement percentages with the exception of the guilt judgment, for which endorsement was lower if it followed the responsibility judgment than if it preceded the latter judgment or followed other appraisal judgments.

The proportions of yes-answers obtained for the appraisal questions were generally also high (the exception is the responsibility judgment in gratitude scenarios) and as high as or higher than those obtained in Study 1. Statistical comparisons revealed that question position had a significant influence only on event evaluation in guilt scenarios: Guilt scenarios were more often evaluated as negative when the evaluation followed a guilt judgment than when it was made before this judgment.

In sum, with the exception of the guilt judgments and the evaluation judgments in the guilt scenarios, question position had no statistically reliable effect on the proportion of yes-answers. Even for the two exceptions, the effect of question position was very small in absolute terms.

Reaction Times

Prior to further analysis, outliers were identified and replaced using the same procedure as in Experiment 1. However, because we expected the judgment facilitation effects to be of smaller magnitude than the unconditional RT effects that were at the focus of Experiment 1, we used a more restrictive criterion for outliers, namely a z-value > 2.5. Using this criterion, 170 responses (1.2%) were identified as outliers and replaced.

Replication of Previous Findings

The mean first-position RTs and the statistical comparisons of the appraisal and emotion judgments are summarized in Table 5. As can be seen, with one exception, all appraisal judgments took significantly longer than the respective emotion judgments. The exception concerned the evaluation judgment in the guilt scenarios, which did not take significantly longer than the guilt judgment (although it was still longer in absolute terms). Hence, the RT results of Study 1 were replicated.

Table 3
Proportion of Yes-Answers for Emotion Judgments and Statistical Comparisons Within Scenarios, Experiment 2

<table>
<thead>
<tr>
<th>Emotion type</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anger</td>
<td>Guilt</td>
</tr>
<tr>
<td>First Place</td>
<td>.96</td>
</tr>
<tr>
<td>Post evaluation</td>
<td>.95</td>
</tr>
<tr>
<td>Post responsibility</td>
<td>.92</td>
</tr>
<tr>
<td>Post focus</td>
<td>.92</td>
</tr>
<tr>
<td>Post fairness</td>
<td>.97</td>
</tr>
<tr>
<td>Friedman-ANOVA</td>
<td>ns</td>
</tr>
</tbody>
</table>

Note. First-place judgments are based on N = 12 judgments. ANOVAs = analyses of variance.

Table 4
Proportion of Yes-Answers of Appraisal Judgments and Statistical Comparisons for First- Versus Second-Position Judgments, Experiment 2

<table>
<thead>
<tr>
<th>Appraisal</th>
<th>Position</th>
<th>Anger</th>
<th>Guilt</th>
<th>Gratitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluation</td>
<td>First</td>
<td>.87</td>
<td>.81</td>
<td>.95</td>
</tr>
<tr>
<td></td>
<td>Second</td>
<td>.91</td>
<td>.92*</td>
<td>.96</td>
</tr>
<tr>
<td>Responsibility</td>
<td>First</td>
<td>.78</td>
<td>.91</td>
<td>.44</td>
</tr>
<tr>
<td></td>
<td>Second</td>
<td>.76</td>
<td>.92</td>
<td>.45</td>
</tr>
<tr>
<td>Focus</td>
<td>First</td>
<td>.86</td>
<td>.79</td>
<td>.90</td>
</tr>
<tr>
<td></td>
<td>Second</td>
<td>.81</td>
<td>.80</td>
<td>.95</td>
</tr>
<tr>
<td>Fairness</td>
<td>First</td>
<td>.81</td>
<td>.61</td>
<td>.96</td>
</tr>
<tr>
<td></td>
<td>Second</td>
<td>.83</td>
<td>.64</td>
<td>.93</td>
</tr>
</tbody>
</table>

* Significantly different from first-position judgment-proportion, Wilcoxon test, p < .05.

Judgment Facilitation Effects

The facilitation of emotion judgments by appraisal judgments is shown in Figure 1 and that of appraisal judgments by emotion judgments in Figure 2.

Facilitation of emotion judgments by preceding appraisal judgments. A three-factorial repeated measurements ANOVA of the EJ reaction times with the factors emotion type, appraisal, and position yielded a significant effect for the three-way interaction, F(6, 282) = 2.20, p < .05. Therefore, we decided to conduct separate two-way (emotion X position) ANOVAs for the different preceding appraisal (see Table 6). These ANOVAs revealed significant position effects indicative of judgment facilitation for all appraisals. In addition, there was a significant interaction of position and emotion type in the case of the evaluation, focus, and fairness judgments, and a marginally significant interaction for the responsibility judgments (p < .07).

Figure 1 also summarizes the results of t tests comparing the first-position emotion judgments with the emotion judgments following the four appraisal judgments, separately for each of the three emotions. As can be seen, the emotion judgments were significantly facilitated by preceding appraisal judgments in 9 of
the 12 cases. The exceptions were anger judgments were not significantly facilitated by prior judgments of responsibility and focus, and gratitude judgments were not significantly facilitated by prior fairness judgments. However, all obtained RT differences were positive, indicating facilitation.

Facilitation of appraisal judgments by preceding emotion judgments. A three-factorial repeated-measurement ANOVA of the AJ reaction times revealed a marginally significant three-way interaction, $F(6, 282) = 1.93, p < .08$. Therefore, we conducted three separate two-way repeated measurement ANOVAs, one for each emotion type (see Table 7). Significant position effects indicative of judgment facilitation effects were obtained for all emotion types; in the case of the gratitude scenarios, they were moderated by a significant position X appraisal interaction.

Figure 2 also summarizes the results of $t$ tests comparing the first-position appraisal judgments and the appraisal judgments following the respective emotion judgment, separately for each of the three emotions. In 9 of the 12 cases, the appraisal judgments were significantly facilitated by the preceding emotion judgments. The exceptions were evaluation appraisals were not facilitated by prior anger judgments, and fairness and responsibility appraisals were not speeded up by preceding gratitude judgments. However, only in the last-mentioned case was the RT difference in a direction opposite to a facilitation effect. In sum, most AJs facilitated the corresponding EJs, and vice versa.

Comparison of Facilitation Effects

Having established both EJ and AJ facilitation effects, we next asked which of these effects was stronger. To answer this question, we conducted a three-way repeated measurement ANOVA on the facilitation scores (i.e., the RT differences between first and second-position judgments) with the factors emotion type, appraisal type and facilitation type (facilitation of appraisal judg-
Results of the Two-Way Repeated Measurements ANOVAs for Emotion Judgments, Experiment 2

<table>
<thead>
<tr>
<th>Effect</th>
<th>Evaluation</th>
<th>Responsibility</th>
<th>Focus</th>
<th>Fairness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position (P)</td>
<td>$F(1, 47) = 19.07, p &lt; .001$</td>
<td>$F(1, 47) = 8.48, p &lt; .01$</td>
<td>$F(1, 47) = 31.65, p &lt; .001$</td>
<td>$F(1, 47) = 51.37, p &lt; .001$</td>
</tr>
<tr>
<td>Emotion Type (ET)</td>
<td>$F(2, 94) = 17.68, p &lt; .001$</td>
<td>$F(2, 94) = 22.41, p &lt; .001$</td>
<td>$F(2, 94) = 21.17, p &lt; .001$</td>
<td>$F(2, 94) = 11.77, p &lt; .001$</td>
</tr>
<tr>
<td>P × ET</td>
<td>$F(2, 94) = 4.35, p &lt; .05$</td>
<td>$F(2, 94) = 2.86, p &lt; .07$</td>
<td>$F(2, 94) = 8.11, p &lt; .01$</td>
<td>$F(2, 94) = 4.84, p &lt; .05$</td>
</tr>
</tbody>
</table>

Note. ANOVAs = analyses of variance.

Discussion

With respect to the first-position judgments, Experiment 2 replicated the main result of Study 1 that appraisal judgments need more time than emotion judgments. This result was replicated despite several procedural variations: A cue-word approach and a combined speed-and-accuracy instruction were used, and the responses were embedded into a judgment facilitation paradigm. Hence, the findings of Study 1 are not the result of idiosyncrasies of question presentation or of the instructions used in Experiment 1.

More important, however, are the results of Experiment 2 concerning judgment facilitation effects. As predicted, we found both facilitation of AJs by prior EJs and facilitation of EJs by prior AJs, as well as evidence that the former effect is stronger. If one accepts the logic of the judgment facilitation paradigm, the facilitation of AJs by prior EJs means that during the process of emotion inference, information relevant to appraisals was activated. Hence, as predicted by the proceduralization hypothesis, the information (situational features or schemas) activated during emotion inference is, at least in part, the same as that activated during the inference of the corresponding appraisals.

The facilitation of EJs by preceding AJs supports the assumption that the information activated when making appraisal judgments is relevant for emotion inferences. As mentioned in the introduction to Experiment 2, this facilitation effect can also be explained by the proceduralization hypothesis. However, it needs to be emphasized that this effect only means that information activated during appraisal judgments can be used and was indeed used by our participants to speed up emotion judgments. The EJ facilitation effect does not mean that unconditional emotion judgments (i.e., those not preceded by AJs) are based on such information; indeed, according to the “knowledge compilation” version of the proceduralization hypothesis, this is not the case (cf. the introduction to Experiment 2). In this respect, the observed facilitation of AJs by prior EJs is more conclusive because it demonstrates that unconditional emotion judgments activate information that overlaps at least partly with information that is activated during the inference of the corresponding appraisals.

Results of the Two-Way Repeated Measurements ANOVAs for Appraisal Judgments, Experiment 2

<table>
<thead>
<tr>
<th>Effect</th>
<th>Anger</th>
<th>Guilt</th>
<th>Gratitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position (P)</td>
<td>$F(1, 47) = 13.52, p &lt; .001$</td>
<td>$F(1, 47) = 26.64, p &lt; .001$</td>
<td>$F(1, 47) = 13.16, p &lt; .001$</td>
</tr>
<tr>
<td>Appraisal (A)</td>
<td>$F(3, 141) = 19.2, p &lt; .001$</td>
<td>$F(3, 141) = 8.72, p &lt; .001$</td>
<td>$F(3, 141) = 36.95, p &lt; .001$</td>
</tr>
<tr>
<td>P × A</td>
<td>$F(3, 141) = 1, ns$</td>
<td>$F(3, 141) = &lt; 1, ns$</td>
<td>$F(3, 141) = 3.01, p &lt; .05$</td>
</tr>
</tbody>
</table>

Note. ANOVAs = analyses of variance.
Finally, again in accord with predictions, we obtained evidence for a differential facilitation effect: Overall, the facilitation of AJs by preceding EJs was stronger than the facilitation of EJs by preceding AJs. This finding supports the assumption that the memory information activated during the inference of appraisals makes up only part of the situational schemas that is matched during an emotion inference. For example, the inference of responsibility activates only those schema components that are relevant for determining whether the person described in the scenario is accountable for the event. By contrast, the inference of anger also requires the activation of the situational features or schemas needed for the inference of evaluation, fairness, and focus.  

A statistically significant differential facilitation effect was however not obtained for all appraisal-emotion combinations. There are 2 possible explanations for this finding. First, it could simply be because of insufficient statistical power. As mentioned, in addition to the 4 cases where a significantly stronger facilitation effect was obtained for AJs than for EJs, there were five other cases in which the RT differences were in the same direction; they may have failed to reach significance because of insufficient statistical power. Therefore, it would be premature to conclude that differential facilitation effects exist only for some emotions or appraisals.

Second, the differential facilitation effect could indeed be restricted to certain emotions or appraisals. If so, this could in principle be explained by assuming that some appraisals are inferentially connected. For example, some appraisals may spontaneously trigger other appraisals or may require the prior inference of other appraisals (cf. Reisenzein, 2001). A closely related hypothesis was proposed by van Overwalle et al. (1992) to explain the finding that the presentation of preattributinal information speeded up emotion judgments to the same degree as attributional information. The authors speculated that both preattributional and attributional information are stored in the same schema, and that the activation of one component of this schema tends to activate its remaining components. In the present case, one could similarly assume that there is, for each emotion, an abstract appraisal schema that contains the appraisals characteristic for this emotion, and that the activation of one of the appraisals activates the remaining appraisals. Alternatively, the activation of an appraisal component could prime all emotion schemas that contain this appraisal component, and thereby constrain emotion inferences to a subset of emotions.

Finally, the findings of Experiment 2 suggest that, in the present experimental paradigm, emotion judgments are not made spontaneously during scenario comprehension. For in this case, they should not have been facilitated by preceding appraisal judgments. This conclusion is in line with recent results by Gygax et al. (2003) and Gygax, Garnham, and Oakhill (2004), who found that specific emotions are not spontaneously inferred during text comprehension. More generally, this finding underscores the conclusion of previous studies that different facets of automaticity—such as efficiency and spontaneity—can be dissociated (e.g., Bargh, 1994).

Experiment 3

The aim of Experiment 3 was to replicate the central findings of Experiment 2 and to further support their interpretation by ruling out two possible alternative hypotheses.

First, at least as far as the observed facilitation of AJs by preceding EJs is concerned, one could object that the crucial factor was not the inference that an emotion was present in a given scenario, but simply the activation of the emotion concept by the question. Second, because only first-versus second-position judgments were compared in Experiment 2, one could object that facilitation effects were confounded with position effects. That is, it could be argued that facilitated judgments were faster simply because they were made in the second rather than in the first position. For example, although the first question appeared on the screen only after the participants confirmed that they had read and understood the scenario, they may have still been engaged in the process of story comprehension when the first question was presented. If so, the faster responses to second-position judgments observed in Experiment 2 could have been because of a delay of the first-position judgments, rather than to a facilitation of the second-position judgments.

Method

Participants

The participants of Experiment 3 were 48 introductory psychology students who participated in partial fulfillment of their study requirements (38 women, mean age = 27.3).

Design and Materials

The design of Experiment 3 was a modification of the design of Experiment 2. The same emotions as in Experiment 2 were studied (anger, guilt, and gratitude), but there were three differences: (a) the number of appraisal dimensions was reduced to three (event-evaluation, fairness, and responsibility). (b) An additional position-control question similar to the control item used in Study 1 was included (participants were asked whether a particular word was contained in the scenario text). This allowed us to test for position effects by comparing judgments following the position-control question with judgments following an appraisal or emotion question. (c) An additional experimental factor, emotion concept priming (priming vs. no priming) was included. Because of this additional factor, the emotion scenarios used in Experiment 2 had to be supplemented by 32 additional scenarios for each emotion, resulting in 56 scenarios for each emotion.

Procedure

The procedure of Experiment 3 was identical to that of Study 2 except for the greater number of scenarios and the addition of the emotion concept priming. In half of the trials (priming condition), the name of the emotion suggested by the scenario was presented immediately before the presentation of the first question. The priming was realized as follows: Immediately after the participants...
confirmed that they understood the scenario sentence, a fixation point appeared at the center of the screen below this sentence. After a delay of 400 ms, the “fitting” emotion word was presented in upper-case letters for 150 ms. Again 150 ms later, the first question was shown. The SOA between prime and question was thus 300 ms. In the other half of the trials (no priming), the emotion-word prime was replaced by a string of “X”s.

Hypotheses

Two main predictions were made: (1) the findings of Experiment 2 are replicated if the position of the judgments is held constant, and (2) the facilitation of emotion-related appraisal judgments by preceding emotion judgments is not simply the result of the priming of the emotion concepts. In Experiment 3, this means that the facilitating effect of emotion judgments on appraisal judgments is not eliminated if the emotion concept is primed (by presenting the emotion word) before the emotion judgment is made.

Results

Prior to the subsequent analyses, the answers and reaction times of each participant were aggregated within the 14 question conditions and the three emotion types. Each aggregated answer and RT is based on N = 2 responses for the priming and the no-priming conditions, with the exception of the first-placed emotion judgments that were collapsed across all other conditions (N = 10).

Judgments

Preliminary data analyses revealed that neither the position of the judgments nor the priming condition had significant effects on the proportion of yes-answers, and that these proportions were largely identical to those obtained in Experiment 2. Therefore, we do not present these findings in detail.

Reaction Times

Prior to the RT analyses, outliers were identified and replaced using the same procedure as in Experiment 2. Two hundred seventy-eight responses (1.7%) were identified as outliers and replaced.

Effectiveness of the priming manipulation. To check the effectiveness of the priming manipulation, we compared the first place emotion judgments preceded by an emotion word prime with those preceded by the control prime (the string of “X”). These comparisons showed that the judgments in the priming condition (M = 715 ms) were significantly faster than those in the no-priming condition (M = 826 ms), F(1, 47) = 17.25, p < .001, η² = .27. Hence, the priming manipulation was effective in activating the emotion concepts. This priming effect did not interact with type of emotion scenario, F(2, 94) < 1, ns.

Facilitation of emotion judgments by preceding appraisal judgments. To test whether the facilitation of EJs by prior AJs obtained in Experiment 2 was replicated, the no-priming condition (which can be directly compared to the experimental conditions in Exp. 2) was analyzed. Two-way repeated measurements ANOVAs with the factors position and emotion type were conducted for each of the three appraisal dimensions. These analyses revealed significant main effects of position for all three appraisal judgments (all ps < .05). Hence, the facilitating effect of AJs on subsequent EJs found in Experiment 2 was replicated.

Facilitation of appraisal judgments by preceding emotion judgments. For each emotion, separate three-factorial repeated measurement ANOVAs were conducted with the factors appraisal, position, and emotion concept priming. The results of these analyses are summarized in Table 8. As can be seen, there were significant main effects of appraisal and position for all scenarios, as well as an additional, significant interaction of the two factors for the anger scenarios. Hence, the facilitation of AJs by preceding EJs obtained in Experiment 2 was also replicated.

Looking next at the effects that involved the factor emotion concept priming, we found a significant three-way interaction of priming, appraisal, and position for the anger scenarios, and a significant interaction of priming and position for the gratitude scenarios.

T tests comparing the judgment conditions (see Figure 3) revealed, for the no-priming conditions, significant facilitation effects in all cases with the exception of a nonsignificant effect of gratitude judgments on judgments of responsibility (replicating the corresponding finding of Experiment 2). For the priming conditions, six out of nine facilitation effects were significant. The exceptions were nonsignificant effects of anger and gratitude judgments on judgments of responsibility, and of gratitude on fairness. Hence, a similar pattern of AJ facilitation effects was obtained in the no-priming and the priming conditions: six of the eight facilitation effects obtained in the no-priming condition were replicated in the priming condition. Taken together, significant facilitation effects were obtained in 14 of the 18 cases.

Table 8

Results of the Two-Way Repeated Measurement ANOVAs for Appraisal Judgments, Experiment 3

<table>
<thead>
<tr>
<th>Emotion type</th>
<th>Position (P)</th>
<th>Appraisal (A)</th>
<th>Priming (PR)</th>
<th>P×A</th>
<th>P×PR</th>
<th>A×P×PR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anger</td>
<td>F(1, 47) = 18.84, p &lt; .001</td>
<td>F(1, 47) = 51.22, p &lt; .001</td>
<td>F(1, 47) &lt; 1, ns</td>
<td>F(1, 47) = 3.20, p &lt; .05</td>
<td>F(1, 47) = 2.40, ns</td>
<td>F(2, 94) = 2.01, ns</td>
</tr>
<tr>
<td>Guilt</td>
<td>F(1, 47) = 51.22, p &lt; .001</td>
<td>F(1, 47) = 34.14, p &lt; .001</td>
<td>F(1, 47) = 1.38, ns</td>
<td>F(1, 47) = 1.42, ns</td>
<td>F(1, 47) = 1.41, ns</td>
<td>F(2, 94) = 1.54, ns</td>
</tr>
<tr>
<td>Gratitude</td>
<td>F(1, 47) = 9.07, p &lt; .01</td>
<td>F(2, 94) = 34.13, p &lt; .001</td>
<td>F(1, 47) &lt; 1, ns</td>
<td>F(2, 94) = 1.87, p &lt; .10</td>
<td>F(2, 94) = 1.41, ns</td>
<td>F(2, 94) &lt; 1, ns</td>
</tr>
</tbody>
</table>

Note. ANOVAs = analyses of variance.
Discussion

Experiment 3 yielded two main findings. First, the main results of Experiment 2 were replicated when the position of the judgments was held constant. This was true both for the appraisal and the emotion facilitation effects.

Second, with two exceptions, the priming of the emotion concept did not eliminate the facilitating effects of EJs on subsequent AJs. Therefore, the hypothesis that the mere activation of the emotion concept is responsible for the observed EJ facilitation effects can be rejected. The two exceptions were judgments of responsibility following judgments of anger and gratitude, which were not facilitated further when the emotion concept had been primed. Conceivably, attributions of responsibility are so strongly connected to the concepts of anger and gratitude that these attributions are already activated by the mere presentation of these emotion concepts.

Finally, the obtained emotion-concept priming effect (the speeding up of the emotion judgments by the matching emotion word prime) further supports the conclusion drawn in the discussion of Experiment 2, that the emotions were not spontaneously inferred during scenario comprehension (see also, Gygax et al., 2003, 2004). For if the emotions had already been spontaneously inferred during scenario comprehension, it seems unlikely that the presentation of the matching emotion concept would have further speeded up emotion judgments.

General Discussion

The studies reported in this article investigated the process of emotion inference from situational information (brief scenario descriptions). The starting point of our inquiry was the appraisal mediation hypothesis of emotion inference, which posits that emotion judgments are mediated by inferred appraisals. As explained in the introduction, this hypothesis predicts that emotion judgments should take at least as much time as the slowest mediating appraisal judgment. The results of Experiment 1, which were replicated in Experiments 2 and 3, did not confirm this prediction.

On the contrary, similar to the findings of previous studies (Schuster et al., 1998; Siemer, 2001; van Overwalle et al., 1992) appraisal judgments took on average more time than the corresponding emotion judgments. This was true even for appraisal dimensions that are widely regarded as central to the corresponding emotions.

To explain these findings, we proposed a modified version of the appraisal mediation hypothesis of emotion inference, the proceduralization or automatization hypothesis. The proceduralization hypothesis assumes that as a result of extensive practice the initially mediating appraisal inferences have become fast and inaccessible to consciousness or have even come to be entirely omitted through the process of knowledge compilation. This hypothesis implies that the situation schemas that are activated during emotion inferences contain the same information that is also accessed during appraisal inferences. This implication of the proceduralization hypothesis was tested in Experiments 2 and 3 using a judgment facilitation paradigm. Supporting the proceduralization hypothesis, both experiments found that (a) emotion judgments facilitated subsequent appraisal judgments and (b) appraisal judgments also facilitated subsequent emotion judgments. Experiment 3 further corroborated these conclusions by demonstrating that the obtained judgment facilitation effects (a) were not an artifact of the position of the judgments and (b) were not because of the mere activation of the emotion concept, as the priming of the emotion concepts did not facilitate AJs and did not eliminate facilitation of AJs by preceding EJs.

In addition, Experiment 2 found that EJs facilitated subsequent AJs to a higher degree than vice versa. This result supports the assumption that each appraisal judgment activated only part of the situational schema that is accessed during an emotion judgment. Although this differential facilitation effect was not statistically reliable for all appraisal-emotion combinations, this finding may have been because of insufficient statistical power. Alternatively, these asymmetric facilitation effects were also found in Experiment 3, but possibly because of the differences in procedure, were less pronounced.
some appraisals may be connected to other appraisals in such a way that the activation of one appraisal activates other appraisals, and in some cases even the complete schema that is accessed in the course of an emotion inference. Different appraisals are probably differentially central to and thus differentially able to activate other appraisal dimensions of a complete emotion schema (Reisenzein & Spielhofer, 1994).

In sum, Experiments 2 and 3 provided experimental support for a modified version of the appraisal mediation theory of emotion inference, the proceduralization hypothesis. This hypothesis preserves the strengths of the appraisal mediation hypothesis, but other than the appraisal mediation hypothesis, it can also explain the problematic reaction time data reported in Experiment 1.

Is Simulation an Alternative?

A basic assumption of the present research is that the process of emotion inference from situational information is a process in which conclusions about another’s emotional state (or one’s own emotional state in hypothetical or future situations) are derived from knowledge about the concrete eliciting situation and more general folk-psychological assumptions (e.g., Heider, 1958; Perner, 1996; Weiner, 1986). However, an alternative possibility must at least briefly be considered. This possibility is that, to infer their likely emotional reactions to the scenarios, our participants simulated the actual process of emotion generation (e.g., Goldman, 1993; Gordon, 1986; Davies & Stone, 1996; Scherer, 1999). According to this simulation theory of emotion inference, people infer their own or other’s likely emotional reaction to a given event by imagining that the event happens to them. This imagery presumably elicits a facsimile (i.e., a “fantasy” version) of the emotion that they would experience if the described situation actually happened to them. The simulated emotion is then introspectively “read off” and attributed to the actual protagonist of the story.

The question to which degree people use a simulationist versus a theory-based strategy for predicting their own or others’ mental states has been the subject of substantial debate during the past years (e.g., Davies & Stone, 1996; Perner, 1996; Saxe, 2005; Stitch & Nichols, 1997). Although the issues are far from decided, there seems to be increasing agreement among the participants of this debate that both strategies—simulation and theory-guided inference—play a role in mental state ascriptions. If one accepts this conclusion, the main question becomes when and under what circumstances which of the two strategies is used. Specifically, is simulation a likely alternative explanation for the findings of our studies? We believe that it is not, for two reasons. First, there is evidence that verbally presented hypothetical scenarios, and even recalled emotional experiences, elicit emotional reactions only under special circumstances; in particular, if the situations are imagined or recalled, respectively, vividly and in substantial detail (Ayduk, Mischel, & Downey, 2002; Keltner, Ellsworth, & Edwards, 1993; Siemer, 2005; Strack, Schwarz, & Gschneidinger, 1985). The emotion inferences required in our studies certainly did not encourage such processes.

Second, the simulation theory cannot easily explain the RT effects obtained in our experiments. At least, it seems that to explain these findings, the simulation theory would have to be amended by additional assumptions that are quite similar to those made in the proceduralization theory. For example, to acquire substance, the simulationist account needs to explicate how the emotional reactions were simulated in the first place. But if appraisal theories are right that it is not (real or imagined) situations per se that cause emotions, but rather the appraisals of these situations, it would seem that the simulationist theory needs to assume that the participants simulated emotions by first simulating appraisals. Consequently, to explain that appraisal judgments take more time than emotion judgments, the simulation theory would have to assume that the simulated appraisals were automatic and unconscious and had to be “reconstructed” by means of a more time-consuming process. And to explain the facilitation of emotion judgments by preceding appraisal judgments, simulation theory would have to assume, for example, that the appraisal judgments primed the emotion concepts, as a result of which the simulated emotion could be more quickly verbalized. As documented by the proceduralization hypothesis, suitably modified versions of these additional assumptions (about automatic appraisal inferences and priming processes) are alone sufficient to explain the RT findings; no process of mental simulation needs to be postulated. Nonetheless, further empirical investigation of the issue (e.g., by using potential concurrent indicators of simulated emotions, such as facial EMG) would be an interesting task for future research.

Implications for Appraisal Research

To the degree that our results support the hypothesis that emotion inferences from situational information are (typically) automatized or proceduralized, they vindicate previous studies of the appraisal-emotion link, specifically studies that investigated this link by presenting participants with hypothetical emotion-eliciting scenarios. As noted in the introduction, this research has been based on the assumption that judgments about hypothetical appraisals and emotions are similar to the corresponding judgments of actual appraisals and emotions. As mentioned in the introduction, the appraisal mediation hypothesis of emotion inference provides a straightforward justification of this assumption: If emotion ascriptions in hypothetical scenarios are based on inferred appraisals, emotion inferences reflect at least approximately (i.e., to the degree that the underlying, folk psychological theory of appraisal-emotion links is accurate) the appraisals and emotions that occur in real situations. Whereas the original appraisal mediation hypothesis was put into question by the finding that emotion judgments require less time than appraisal judgments, the present results suggest that a modified version of this hypothesis can be upheld. For although emotion judgments are usually not based on conscious and controlled appraisal judgments, one can still say that they are based on implicit or proceduralized appraisal judgments in the sense described earlier.

In the introduction, we also raised the possibility that research on emotion inferences from situational information may shed light on the actual process of emotion generation. The remainder of the discussion will be devoted to a discussion of this issue.

Generally speaking, drawing conclusions from the process of emotion inference to the process of emotion generation would be warranted if the two processes are similar in relevant respects. This may be the case even if the process of inferring emotions from situational information does not involve a simulation of the process of emotion generation (see above). For given that the two processes solve similar tasks—bridging the gap between situation
and emotion—it is reasonable to assume they involve partly the same or similar subprocesses. To the degree that this is the case, the process of emotion inference from situational information may provide suggestions about how the process of emotion generation looks like.

In particular, the present theory of emotion inferences suggests a new view of how automatic appraisals might work. Many appraisal theorists assume that emotion-eliciting appraisals are often automatic in the sense of unconscious and efficient (e.g., Lazarus, 1991; Reisenzein, 2001; Scherer, 2001; Smith & Kirby, 2001). However, there is so far no agreement precisely on how automatic appraisal operates and only little empirical research has been devoted to this question. For example, Leventhal and Scherer (1987; see also, Scherer, 2001) proposed that appraisals can occur on different levels of processing, distinguishing between a hard-wired, a schematic, and a conceptual level. Lazarus (1991) linked automatic appraisals more specifically to Leventhal’s (1982; see also Leventhal & Scherer, 1987) notion of schematic processing. He proposed that schematic processing results when a situation is repeatedly appraised in a particular way and eventually causes the automatic occurrence of the emotional reaction. Smith and Kirby (2001) distinguish between conscious reasoning processes and automatic associative processes. The latter processes, they propose, are capable of associatively activating appraisals. The respective associations are thought to be the result of the cooccurrence of specific patterns of sensory information and specific appraisal meanings. Many models of automatic appraisals in emotion generation are dual-process models (Smith & Neumann, 2005). In these models, automatic and deliberate appraisals are seen as distinct processes that exist independently of each other. Often, these distinct appraisal processes are also seen as implementing appraisals of different complexity and are attributed to different neural systems (Ochsner & Barrett, 2001). Reisenzein (2001) suggests that schematic appraisal consists of the retrieval of an appraisal from a schema for this event in which the appraisal is stored together with other information of this event. Thus, appraisals become immediately activated once it is recognized that an event fits a particular schema.

The present research, which suggests that emotion inferences are based on proceduralized appraisal inferences, points to a related possibility of how actual, automatic appraisal could work: automatic appraisal could be based on proceduralized appraisals. That is, as a result of repeated appraisals of the same or similar events, appraisals as mediating steps between a situation schema and the emotion have become fast and unconscious, and may even have come to be entirely omitted. This hypothesis could account for several properties of seemingly automatic emotion generation, in particular: (a) once a situation matches the antecedent of an emotion generating procedure, the emotion is triggered automatically, without any deliberate and effortful intermediate steps and (b) people are often unaware of the appraisals that triggered an emotion (cf. Frijda, 1986).

Although this proceduralization hypothesis of emotion generation is generally in line with previous accounts of emotion generation (in the sense of being a possible elaboration of certain assumptions of these theories), there are some noteworthy differences. First, several of the above-described appraisal theories assume that the schematic situation representations that trigger automatic appraisals may also contain nonpropositional representations of sensory stimuli. By contrast, the proceduralization hypothesis focuses exclusively on propositional representations of situations. Second, the proceduralization hypothesis does not cover appraisals that are presumably nonpropositional in nature, such as appraisals of intrinsic (sensory) valence (cf. Scherer, 2001). Third, in contrast to what at least some authors seem to assume, the proceduralization hypothesis of emotion generation entails that complex, propositional appraisals can become automatized, reflecting Lazarus’ (1995) claim that in order to be adaptive automatic appraisals must often be quite complex.

In line with previous suggestions (e.g., Lazarus, 1991), the proceduralization hypothesis of emotion generation assumes that “automatic emotions” occur as responses to events that have been repeatedly associated with particular appraisals and resulting emotions. That is, automatic emotions are most likely to occur as responses to familiar events and situations; whereas in unfamiliar situations, appraisals are likely to be more nonautomatic (i.e., conscious and deliberate, cf. Ellsworth & Scherer, 2003).

Accordingly, one important research topic for the future is the question which factors may cause proceduralized appraisal judgments to become controlled and conscious again. For example, the regulation of emotions by the deliberate reappraisal of a situation (Gross, 1998) may be difficult if the initial appraisal is automatic and unconscious. The reappraisal of such situations may therefore require to 1st de-automatize the automatic appraisals. Although it may be possible in principle to intentionally interrupt automatic appraisal processes and to shift into a controlled processing mode, in many cases it may be easier to rely on “natural de-automatizers.” We believe that two such natural de-automatizers are unfamiliarity and unexpectedness. Hence, one strategy to accomplish the de-automatization of appraisals may consist in the attempt to make the situation to some degree unfamiliar again, for example by introducing novel elements into the situation, or by changing the perspective or the frame of reference. A second related factor that may cause the de-automatization of appraisals is unexpectedness. Indeed, unexpectedness (or perhaps the ensuing surprise) may be a primary factor that disrupts automatic processing and triggers controlled inference strategies (Meyer, Reisenzein, & Schützwohl, 1997; Pyszczynski & Greenberg, 1987; Stiensmeier-Pelster, Martini, & Reisenzein, 1995). However, unlike change of perspective, unexpectedness of a situation is probably not subject to deliberate influences. Although, unlike a change of perspective, the unexpectedness of a situation seems largely beyond the deliberate influence of the person concerned, it may be possible to introduce unexpected elements in schematic situations (e.g., by changing familiar contexts, etc.). In any case, a better understanding of the process of de-automatization or de-proceduralization of appraisals may be highly relevant for questions of emotion regulation research.

These considerations also suggest two extensions of the research presented in this article. The first extension concerns a closer exploration of the effects of familiarity and unexpectedness on the automaticity of emotion inferences. We would predict that RT differences between emotion and appraisal judgments should disappear or be reversed when people judge scenarios describing unexpected or unfamiliar events. The second extension of the present research concerns our suggestions concerning the actual processes of emotion generation. Here, two interesting hypotheses that should be tested in future research are whether emotions based
on deliberate and conscious appraisals are indeed, as we predict, more easily controllable than emotions based on automatic appraisals, and whether “natural de-automatizers” (adding unfamilial and unexpected elements to familiar and expected situations) can indeed be utilized to render appraisal processes, and thus emotions, more controllable by de-automatizing them.

References
Appendix A

Emotion Scenarios Used in Experiment 1

**Pity**

1. A friend fails an important examination.
2. Your father loses his job because of downsizing.
3. The doctor tells your friend that she cannot have children.
4. At the railway station, you see an elderly woman crying.
5. A friend breaks his leg just before a sporting competition.
6. Because of an accident, a friend cannot go on a vacation.
7. The lease of a friend is revoked; she is without a home.
8. You learn that a female friend was attacked during the night.
9. Your travel companion's money and luggage are stolen.
10. The parents of your friend die in a car accident.
11. A friend loses his lover to someone else.
12. Because of financial problems, a friend has to quit his studies.
13. A friend is desperate because his lover left him.
14. Your colleague loses his job because of illness.

**Anger**

1. While entering the bus, your shopping bags tear apart.
2. The telephone rings all night long, but when you pick it up, nobody answers.
3. You have an appointment with a friend but he stands you up.
4. Twenty kilometers away from home your bicycle breaks.
5. A classmate leaves your lecture notes on the subway.
6. Part of an article your lecturer distributed is unreadable.
7. Your best friend discloses a secret entrusted to her.
8. You cannot concentrate because the people around you are noisy.
9. The money your parents loaned you has still not arrived.
10. You lend your car to someone; it is returned battered and with an empty tank.
11. One of your dinner guests comments negatively on your cooking abilities.
12. You learn that your colleague has slandered you to the boss.
13. You find out that your friend has lied to you.
14. Your sister ruins the stereo set you lent her.

**Guilt**

1. Because you did not keep an appointment, your colleague could not finish his work.
2. Turning left in your car, you fail to see a pedestrian and run her down.
3. You are careless at work; a colleague gets hurt.
4. You forgot to refuel your car; it breaks down when your friend uses it.
5. You didn't feel like helping your friend; she fails the exam.
6. You tell people your friend was responsible for a prank; he is punished for it.
7. You forgot to return the urgently needed money to your friend.
8. You lied to your friend about an important matter.
9. Your partner is deeply hurt when learning that you lied to him.
10. You make fun of a friend's looks in his presence.
11. You accidentally leave a colleague's lecture notes in the subway.
12. You ridicule a friend in his presence and in front of other people.
13. You forget a study group appointment; the others have to do the work themselves.
14. Someone slips on the banana peel that you threw away.

**Gratitude**

1. Hitchhiking: After a long wait, you are finally picked up by a car.
2. Someone lends you the 50 cents you need to make a call at the pay phone.
3. When you leave the crowded store, someone opens the door for you.
4. Your neighbor helps you with the repair of your broken bike.
5. You have to lie in bed because you have the flu; a friend does the shopping for you.
6. A friend helps you translate a difficult text.
7. You have an acute sneezing attack; someone hands you a tissue.
8. Your colleague is willing to change work shifts with you.
9. You swam out too far into the sea; a boat saves you.
10. A friend helps you find an affordable apartment.
11. Until you have found something else, a friend lets you stay at his place.
12. In a difficult financial situation, a friend lends you money.
13. A friend listens sympathetically to you when you tell him a problem.
Appendix B

Emotion and Appraisal Questions Used in Experiment 1

Emotions
I experience the following feeling: Pity.
I experience the following feeling: Anger.
I experience the following feeling: Guilt.
I experience the following feeling: Gratitude.

Subjective Event Evaluation—positive/negative
I experience this as positive/desirable.
I experience this as negative/undesirable.

Moral Event Evaluation—positive/negative
This is morally right/fair/deserved.
This is morally wrong/unfair/undeserved.

Self-Evaluation—positive/negative
I evaluate myself positively.
I evaluate myself negatively.

Other-Evaluation—positive/negative
I evaluate someone else positively.
I evaluate someone else negatively.

Event Importance (important/unimportant)
This is important for me.
This is unimportant for me.

Expectedness of the Event (expected/unexpected)
I expected this.
This was unexpected.

Responsibility for the Event (self/other)
I am responsible for this.
Someone else/circumstances are responsible.

Controllability of the Event (controllable/uncontrollable)
I can still influence/change this.
I can no longer influence/change this.

Focus of the Event (self/other)
This concerns primarily myself.
This concerns primarily someone else.