

The Induction of Mood via the WWW

Anja S. Göritz

Published online: 22 November 2006
© Springer Science+Business Media, LLC 2006

Abstract Five experiments explored the methodology of Web-based mood induction and yielded empirical evidence for its feasibility and its limits. Study 1 examined the suitability of the Velten method and mood-suggestive photographs to induce both positive and negative moods. Negative mood was successfully induced, while the positive mood induction failed. In Study 2, cartoons successfully ameliorated mood, whereas jokes were ineffective. In Study 3, associations with affectively valenced words were not suitable in producing a positive or a negative mood. In Study 4, picture-illustrated emotive texts successfully induced a positive and a negative mood. In Study 5, valence and arousal ratings were obtained for the stimuli used in Study 1 and 3. Based on the results, recommendations for the application of online mood induction procedures are given.

Keywords Mood induction · Effectiveness · WWW · Internet · Online

Experimental mood research relies on mood induction procedures (MIPs) to elicit the desired moods. Although Web-based experimentation has many advantages and is now established (Birnbaum, 2004; Gosling, Vazire, Srivastava, & John, 2004; McGraw, Tew, & Williams, 2000), experimental mood research via the World Wide Web (WWW) is still in its infancy. In a pilot study, Göritz and Moser (2006) have demonstrated the effectiveness of the Velten MIP (Velten, 1968) and mood-suggestive photographs in inducing a neg-

ative mood. However, these two methods were ineffective in generating a positive mood. As a third MIP, autobiographical recall was tested. It failed to induce positive or negative mood. In re-testing the Velten MIP and mood-suggestive photographs and in testing four more MIPs, this article examines the methodology of Web-based mood manipulation and presents more evidence for both its feasibility and limits.

Because in Web-experiments there is no physical contact between participants and researcher, some of the conventional MIPs are ineligible to be used online such as hypnosis, drugs, and manipulation of body language. Films, music, social interaction via videoconferencing, and other MIPs that require sophisticated IT and bandwidth on the part of the participants cannot (yet) be used with general samples. Other MIPs such as gift-based procedures can be employed online if the gift comes in electronic form (e.g., a redeemable voucher by e-mail). Some procedures, however, can be applied via the WWW with only minor modifications. These are, for example, the Velten method, autobiographical recall, aptitude tests with or without false feedback, mood-suggestive photographs, jokes, stories, cartoons, and associations with affectively valenced words. They all share the quality of solely relying on text or still pictures and on standard HTML input fields. Moreover, they make no special demands on participants' hard- and software.

Online MIPs need to meet the same quality criteria as offline procedures. However, because there is less control over the participants and the participation process on the WWW, Web-based mood induction faces the potential problem of systematic noncompliance. If participants find an online MIP unacceptable, which might predominantly be the case when a negative mood is induced, they might abandon the experiment or incompletely or untruthfully fill in the questionnaire more often than participants who are induced into a positive mood. For example, a negative MIP

A. S. Göritz (✉)
Department of Organizational and Social Psychology, University
of Erlangen-Nürnberg,
Lange Gasse 20,
90403 Nuremberg, Germany
e-mail: anja.goeritz@wiso.uni-erlangen.de

might appear ineffective because the data for the participants who dropped out precisely because they had experienced mood deterioration are lost. Therefore, a variety of collateral data on the participation process, such as response latitudes and the extent of item-nonresponse, should be drawn upon to identify noncompliant participants. It should then be tested whether the proportion of noncompliers differs across experimental mood conditions.

The following sections describe the test of six MIPs in five studies (1) the Velten technique, (2) emotion-laden photographs, (3) jokes, (4) cartoons, (5) associations with affectively valenced words, and (6) picture-illustrated emotive texts.

Study 1

This experiment examined the effectiveness of the Velten technique and the presentation of emotionally charged photographs in arousing positive and negative moods. In the standard Velten procedure (Velten, 1968), participants read at their own speed 60 positively or negatively tinged self-referent statements. The Velten procedure is generally effective: In Larsen and Sinnett's (1991) meta-analysis, participants in Velten-type inductions scored .76 *SD* higher than controls. Westermann, Spies, Stahl, & Hesse (1996) found in their meta-analysis that the Velten method more effectively deteriorates (mean weighted effect size based on 72 studies, $r = .52$) than improves mood (mean weighted effect size based on 46 studies, $r = .38$). The r statistic used throughout this paper is Pearson's product-moment correlation and a measure of effect size (.10 is a small effect, .30 is a medium effect, and .50 and greater is a large effect). The r statistic can be converted to a Cohen's $d = [2r]/[(1-r^2)^{0.5}]$.

The display of mood-suggestive photographs belongs to a group of procedures that present emotive material without explicit instructions for getting into the target mood. Westermann et al. (1996) assessed the effectiveness of a group of procedures based on material similar to photographs. The effectiveness was high in inducing both a positive mood (mean weighted effect size from 13 studies, $r = .53$) and a negative mood (mean weighted effect size from 16 studies, $r = .50$). However, only one article integrated in this meta-analysis relied on actual pictorial stimuli (i.e., McArthur, Solomon, & Jaffe, 1980). In their Studies 1 and 2, the effects of the negative mood inductions were $r = .69$ and $r = .59$. The effects of the positive mood inductions were $r = .71$ in Study 1 and $r = .25$ in Study 2.

Because the negative Velten and the negative Photo MIP had effectively induced a negative mood in Göritz and Moser (2006), it was expected that in the current within-subjects experiment participants' post-experimental mood would be more negative than their pre-experimental mood. The Velten

and the Photo MIP were used in a short and a long version. No hypotheses were stated concerning effectiveness differences between the long and the short version of each procedure. Because in Göritz and Moser (2006) the positive mood induction had not succeeded, no directional hypotheses were formulated on the impact of the positive Velten and Photo MIP.

Method

Participants received an e-mail invitation asking them if they would like to take part in a *study concerned with evaluating media contents*. Participants' pre-induction mood was measured on *Scale B - Elated Mood* of Hampel's (1977) *Adjective Scales for the Assessment of Mood*. The instructions for filling out the scale said: *Ask yourself with each one of the following adjectives 'How do I feel at the moment?'* The scale comprises the seven adjectives *hilarious, cheery, joyful, humorous, sporting, relaxed, and comfortable* (Cronbach's $\alpha = .84$) with seven anchors ranging from *completely applicable to not at all applicable*. Hampel (1977) has reported several studies that demonstrate the validity of this scale. The participants were assigned to one of eight ($2 \times 2 \times 2$) experimental conditions (Velten or Photo procedure, positive or negative version, long or short form).

In the long Velten MIP, 60 self-referent statements appeared on six pages. These statements were a free translation of the original statements by Velten (1968). Positive statements suggested cheerfulness, optimism, and liveliness; for example, *I've certainly got energy and self-confidence to spare*. Negative statements reflected pessimism, dissatisfaction, and lethargy; for example, *It's so discouraging the way people don't really listen to me*. The short Velten MIP was identical to the long version, except that from each page of the long version a randomly selected set of five statements was discarded. Thus, there were 30 statements left that appeared on six pages.

In the long Photo MIP participants viewed on ten pages either 58 positive or 57 negative pre-tested¹ pictures that had been culled from the WWW and printed media. Pleasant photographs covered the themes "contented families," "desirable women," "comfortable home," "happy children," "fulfilled partnership," "success at sports," "unspoiled nature," "cuddly animals," "holidays," and "children and animals." Unpleasant photographs covered "starvation," "train

¹In a pre-test, four male and two female judges rated the photographs on a 6-point scale with anchors in the following order: *clearly mood-lifting, slightly mood-lifting, mood-neutral, slightly depressing, clearly depressing, and shocking*. For ethical reasons, 12 photographs were excluded which at least three judges found shocking. From the remaining photographs, the 58 pictures with the lowest mean score were used for the positive and the 57 with the highest score for the negative mood induction.

crash,” “unpleasant feelings,” “illness,” “war,” “concentration camp,” “sports accidents,” “environmental destruction,” “industrial disasters,” and “animal experiments.” Instructions on each page, varying slightly from page to page, read: *Let all photographs work on you for a while.* A radio button was located under each photograph. At the end of the positive pages was: *Mark the photograph which you find the best (nicest, most appealing, etc.),* whereas the negative pages said: *Mark the photograph which you like the least (find saddest, most moving, etc.).* This rating was intended to obscure the study’s true purpose. The short Photo MIP contained a selection of 29 positive and 29 negative photographs from the long versions, spread over ten pages. The duration of each MIP is given in Table 1.

After the induction, participants rated their mood on the *SES Scale A – Elated Mood*. This scale is a parallel scale to *Scale B - Elated Mood*. It comprises the seven adjectives *merry, lighthearted, enthusiastic, frolicsome, elated, lively, and free* (Cronbach’s $\alpha = .94$). In the norm sample, the two parallel scales *A* and *B* correlated at $r = .92$ ($N = 846$; Hampel, 1977). The same page included some distraction items which, on account of the cover story, were related to mass media; for example, to choose one out of three magazines that the participants would like to read. In the end, participants went through three pages of remedial cartoons.

Participants

The study was carried out in a new university-based online panel. An online panel is a pool of people who have agreed to occasionally take part in Web-based studies (cf. Göritz, in press). The panelists (i.e., members of the online panel) were volunteers from all walks of life. In part, the panelists had selected themselves for the panel via search engines, links, mailing lists, and fliers. In part, the panelists had been probabilistically sampled through faxes, letters, and e-mails (cf. Göritz, 2004). A total of 317 German-speaking panelists over 16 years old were invited. Of these, 221 (70%) accessed the study’s first page. There resulted 203 (92%) valid and 18 invalid cases: One person quit before the mood induction, 3 people quit during the mood induction, 2 people did not answer two or more mood scale items, 3 people were deemed to have interrupted the experiment because they took longer than 20 min for viewing the mood inducing stimuli, and 9 people produced suspicious response effects (i.e., with each mood scale item they clicked the same answer).² The valid

²Several measures were taken to ensure that participants only took part once. In Studies 1, 3, and 4, and in part in Study 2 participants were recruited from an online panel: They could access the pages with a personal password only. In Study 2 there were also ad hoc recruited participants. All datasets were sifted according to IP-number

Table 1 Study 1: Self-reported pre- and post-induction mood, invalid datasets, and duration for the velten procedure and emotive photographs, each in a long and a short version

	Velten positive		Velten negative		Velten negative		Photos positive		Photos negative		Photos negative	
	long	short	long	short	long	short	long	short	long	short	long	short
<i>n</i>	26	23	27	26	26	26	26	26	26	26	26	23
Pre-induction mood (<i>SD</i>)	29.5 (7.2)	29.0 (8.3)	30.7 (7.6)	28.1 (6.4)	29.4 (8.4)	28.9 (8.3)	30.2 (7.2)	28.9 (8.3)	28.4 (9.3)	28.4 (9.3)	28.9 (8.3)	28.4 (9.3)
Post-induction mood (<i>SD</i>)	29.5 (8.5)	28.4 (9.7)	20.7 (10.1)	21.2 (8.1)	28.2 (8.6)	27.8 (9.5)	27.8 (9.5)	16.6 (8.3)	19.6 (9.7)	16.6 (8.3)	16.6 (8.3)	19.6 (9.7)
Invalid datasets (%)	1 of 27 (3.7)	3 of 26 (11.5)	5 of 32 (15.6)	2 of 28 (7.1)	2 of 28 (7.1)	2 of 28 (7.1)	1 of 27 (3.7)	2 of 28 (7.1)	2 of 25 (8.0)	1 of 27 (3.7)	1 of 27 (3.7)	2 of 25 (8.0)
Duration [min:sec] (<i>SD</i>)	4:23 (2:24)	3:38 (1:30)	4:48 (2:23)	3:38 (2:02)	8:03 (3:28)	8:03 (3:28)	5:03 (1:31)	9:14 (3:38)	7:28 (3:49)	9:14 (3:38)	9:14 (3:38)	7:28 (3:49)

Note. The mood measure ranged from 7 (negative) to 49 (positive).

sample consisted of 63 women (31%) and 140 men. The average age was 34 years ($SD = 10$). Two people had not (yet) completed school, 10 had graduated after 9 years of school, 50 had graduated after 10 years of school, and 141 had graduated from high school.

Dependent variables

The post-pre difference on the parallel *SES* scales A and B *Elated Mood* reflected the mood induction's effectiveness. Contrasting the proportion of invalid cases across experimental conditions reflected compliance.

Results

The participants in the eight conditions did not differ in their pre-experimental mood, $F(7,195) = .33$, *ns* (cf. Table 1). Thus, the random assignment of the participants to the conditions was successful.

Both the long and the short positive Velten MIP were ineffective because the post-pre mood difference was negligible, $t(25) = .03$, *ns* and $t(22) = .48$, *ns*, respectively. Both the long and the short negative Velten MIP were effective in deteriorating mood, $t(26) = 5.62$, $p < .001$, one-tailed, $r = .49$ and $t(25) = 5.39$, $p < .001$, one-tailed, $r = .43$, respectively. The long positive Photo MIP was ineffective, $t(25) = 1.44$, *ns*. Contrary to intuition, with the short positive Photo MIP the initial mood was more positive than the post-induction mood, $t(25) = 3.25$, $p = .01$. Both the long and the short negative Photo MIP effectively deteriorated the mood, $t(25) = 7.51$, $p < .001$, one-tailed, $r = .60$ and $t(22) = 6.49$, $p < .001$, one-tailed, $r = .42$, respectively. The proportion of invalid datasets did not differ among the eight conditions, Fisher's Exact Test = 4.55, $N = 220$, *ns*. None of the MIPs were significantly more or less effective with men than with women.

Discussion

Study 1 examined the Velten technique and mood-suggestive photographs, each in a long and a short version. The positive Velten and Photo MIP failed to improve participants' mood. Surprisingly, the short positive Photo MIP even deteriorated participants' mood. Perhaps this was due to boredom. Or perhaps this was due to positive photographs not having a sufficiently strong positive valence or portraying enough arousal. Study 5 examines this possibility.

and timestamps. Therefore, it can be assumed that very few, if any, multiple participations remained undetected.

As expected, both versions of the negative Velten and Photo MIP led to significant mood deterioration. A comparison of this experiment's results with meta-analytic effects yielded $r = .46$ for the negative Velten MIP averaged across the short and the long version, against $r = .43$ in Westermann et al. (1996)³ and $r = .25$ in Larsen and Sinnett (1991).⁴ The mean effect of the negative Photo MIP amounted to $r = .53$. In Westermann et al. (1996) a similar technique yielded $r = .47$. Thus, both the negative Velten and Photo online MIP yielded similar results to equivalent offline procedures. The procedures' application was uncompromised by mood-related noncompliance.

The present study as well as the pilot study by G6ritz and Moser (2006) have not satisfactorily achieved an amelioration of mood. That is why the next study focused on the induction of positive mood.

Study 2

In this study it was examined whether cartoons and jokes successfully induced a positive mood. Both cartoons and jokes were tested in a short and a long version. The idea to utilize cartoons arose when some participants of Study 1 reported that the remedial cartoons had exhilarated them. Surprisingly, cartoons and jokes as mood lifters have scarcely been used. Some of Linden and Frankish's (1988) participants rested quietly for 20 min, whereas other participants read a book with cartoons. The cartoon group's mood improved relative to the group that rested quietly at $r = .34$. Carnevale and Isen (1986) asked a group of participants to sort 30 cartoons into piles. Participants who had been exposed to the cartoons reported their mood was more positive ($r = .37$) and more pleasant ($r = .30$) than did control participants.

Reading the jokes and cartoons was expected to improve participants' initial mood. However, no directional hypotheses were stated about possible effectiveness differences between jokes and cartoons and between the short and the long versions.

Method

The study was announced as *an investigation into humor appreciation*. The procedure corresponded to that of Study 1, except that each person was assigned to one of

³Mean effect sizes are based on my own re-analysis of Westermann et al.'s (1996) data. To obtain mean effects, each individual effect size was Fisher-transformed and weighted by sample size (Hedges & Olkin, 1985, pp. 230–231).

⁴Larsen and Sinnett (1991) did not differentiate between negative and positive induction.

Table 2 Study 2: Self-reported pre- and post-induction mood, invalid datasets, and duration for cartoons and jokes, each in a long and a short version

	Cartoons long	Cartoons short	Jokes long	Jokes short
<i>n</i>	131	117	117	98
Pre-induction mood (<i>SD</i>)	29.2 (8.4)	29.4 (8.0)	29.3 (8.0)	29.2 (8.2)
Post-induction mood (<i>SD</i>)	30.1 (8.5)	30.3 (8.4)	29.6 (8.7)	28.7 (8.5)
Invalid datasets (%)	50 of 181 (27.6)	62 of 179 (34.6)	50 of 167 (29.9)	44 of 142 (31.0)
Duration [min:sec] (<i>SD</i>)	5:41 (3:16)	4:16 (2:15)	7:25 (4:04)	4:57 (2:37)

Note. The mood measure ranged from 7 (negative) to 49 (positive).

four (2×2) experimental conditions: cartoons long, cartoons short, jokes long, and jokes short. Prior to each MIP, an instruction page appeared. For example, the text for the long version of the cartoons said: *On the following pages you will see a number of cartoons. Please look at them one after the other. After a set of 8 cartoons, please rate the previous pictures' funniness and then have a look at the next 8 cartoons.* A similar text introduced the other three conditions. The long version of the cartoons consisted of 32 cartoons, each on a separate page. After a set of eight cartoons, to render the cover story more credible, a page with a nine-point face scale appeared. The short version of the cartoons consisted of 21 cartoons, complemented by three face scales. The long version of the jokes featured 28 jokes and four face scales, whereas in the short version, 15 jokes and three scales appeared. Cronbach's α with the pre-induction mood measurement on *Scale B – Elated Mood* (Hampel, 1977) was .89, and it was .94 with the post-induction mood measurement on *Scale A*.

Participants

There were two types of participants in this study: (1) different panelists from the same online panel that had been used for Study 1 and (2) non-panelists who had found the study through links, newsletters, newsgroup postings, and word of mouth. In total, 354 panelists were invited. Of these, 193 (55%) proceeded to the first study page. In addition, 706 non-panelists loaded the first page. Of all 899 persons who had accessed the first page, 463 (52%) produced valid and 436 invalid datasets: 230 people quit the study on the welcome page, 123 discontinued their participation in between the second and the last page, 16 people did not respond to two or more mood scale items, 11 people were deemed to have interrupted the experiment because they took longer than 25 min for the mood induction, 19 people were considered click-throughs because of their extremely short visit (i.e., they took less than 2:11 min for the mood induction), and 37 people clicked the same answer with each mood scale item. The sample that produced valid data consisted of 180 women (39%) and 283 men. The average age was 32 years ($SD = 10$). Fourteen people had not (yet) completed school,

33 had graduated after 9 years of school, 100 had graduated after 10 years of school, and 316 had graduated from high school.

Dependent variables

The post-pre difference on the parallel *SES* scales A and B *Elated Mood* reflected the mood induction's effectiveness. Contrasting the proportion of invalid cases across the four conditions reflected compliance.

Results

The participants in the four experimental conditions did not differ in their initial mood, $F(3,462) = .01$, *ns* (cf. Table 2).

The long version of the cartoons improved mood significantly, $t(131) = 1.70$, $p = .045$, one-tailed, $r = .11$. The slight mood improvement achieved by the short version of the cartoons bordered on significance, $t(117) = 1.56$, $p = .06$, one-tailed, $r = .06$. The long and short versions of the jokes, however, did not improve participants' initial mood, $t(117) = .57$, *ns*, one-tailed, and $t(98) = 1.00$, respectively. Compliance did not differ among the four conditions, $\chi^2(3, N = 669) = 2.15$, *ns*. Neither of the two MIPs was significantly more or less effective with men than with women.

Discussion

Study 2 demonstrated that cartoons improved mood significantly, whereas jokes did not. No meta-analytic effect sizes from previous research are available for comparison. It is unlikely that experimental demand accounts for the effects because (1) the experiments' purpose was concealed and (2) there were differences between cartoons and jokes, albeit experimental demand was similar in both MIPs. Yet, because in Studies 1 and 2 mood was assessed on the basis of self-reports, experimental demand cannot be fully ruled out as an explanation for the effects found. To ensure that observed mood effects are solely the result of a genuine mood change, mood needs to be measured through non-self-report measures. Moreover, because emotion words used in

self-reports can encode many features of an emotion episode other than internal experience (Sabini & Silver, 2005), it is useful to complementarily measure participants' experienced emotional states by other means than emotion words. Besides testing a new MIP, the next study focused on eliminating possible demand effects by using a supplementary non-self-report mood measure.

Study 3

This experiment examined participants' associations to emotion-laden words as a method of mood induction (cf. Isen, Johnson, Mertz, & Robinson, 1985). A word-association task can easily be implemented in a Web form and – properly framed – rouses little or no experimental demand. Participants who brought up associations with positively valenced words were expected to be in a better mood, and those who brought up associations with negatively valenced words to be in a more negative mood than participants who associated with neutral words.

Pleasantness ratings of unfamiliar neutral words were used as a non-self-report measure of mood (cf. Isen et al., 1985). Forgas' (2002) Affect Infusion Model assumes that a person's mood is transferred onto an evaluated stimulus more strongly the less familiar and ordinary that stimulus is. Thus, a person in a positive mood is expected to judge an unfamiliar word more positively than a person in a neutral mood. By contrast, a person in a negative mood is expected to evaluate an unfamiliar neutral word more negatively than a person in a neutral mood.

Method

To conceal the experiment's purpose, the task was announced as *an association test to assess creativity*. It came in two versions, namely as a free and as a forced-choice word-association task. In the free association task, participants were instructed to put down in an open-ended fashion the first association that would come to their mind to every cue word presented. In the forced-choice task, participants were instructed to choose one among four provided response options that came closest to their own associations to the cue word. For example, possible answers to the cue word *sun* were *feeling the warmth on the skin*, *summer*, *enjoying the day*, and *growing and prospering*.

Participants were randomly assigned to one of six treatments: free association or forced-choice association to positive, neutral, or negative cue words. In each condition, ten pre-tested cue words appeared on three pages. The positive words were *bliss*, *sun*, *smile*, *peace*, *love*, *honesty*, *friends*, *family*, *friendship*, and *vacation*. The neutral words were *table*, *lamp*, *parliament*, *town*, *pencil*, *stone*, *person*, *pic-*

ture, *radio*, and *river*. The negative words were *war*, *violence*, *rape*, *destruction*, *kidnapping*, *terrorism*, *hostage*, *bomb*, *suppression*, and *hatred*. As a non-self-report manipulation check, participants rated the pleasantness of four unfamiliar neutral words (i.e., *perlator*, *aviatics*, *brevarium*, and *soubrette*) on a 9-point scale ranging from 1 = *very positive* to 9 = *very negative*. These words and their frequency of occurrence were culled from the linguistic word pool at <http://wortschatz.informatik.uni-leipzig.de>. As a self-report manipulation check, participants rated their mood on a 9-point scale ranging from 1 = *very positive* to 9 = *very negative*. The self-reported mood was measured with only one item to keep experimental demand low.

Participants

A sample of 1122 people who were over 16 years old was drawn from the same online panel used for Studies 1 and 2. Of those, 444 (40%) called up the first study page. Of those, 380 (86%) produced valid and 64 invalid datasets: Of the invalid datasets, 22 people abandoned the study prematurely, four people did not rate the pleasantness of two or more of the infrequent words, one person did not self-report her mood, five people ignored at least two cue words in the association task, 14 people participated for an unusually long time (i.e., they took more than 4 min for the forced-choice association task or more than 6 min for the free association task), and 18 people took less than 15 sec for the mood induction. The sample that produced valid data consisted of 181 women (48%) and 199 men. The average age was 31 years ($SD = 10$). Five people had not (yet) completed school, 16 had graduated after 9 years of school, 71 had graduated after 10 years of school, and 288 had earned a high school degree.

Dependent variables

Participants rating of four unfamiliar neutral words was used as a non-self-report measure of mood and their rating of their current mood was used as a self-report measure of mood.

Results

As shown in Table 3, in the free association task the planned contrast of the pleasantness rating of the word *perlator* among the positive, neutral, and negative mood groups (with contrast coefficients $-1, 0, \text{ and } +1$) was significant, but not in line with the stimulus material, $t(195) = 2.01, p = .05$. The contrasts pertaining to the words *aviatics* and *brevarium* were not significant, $t(196) = 1.28, ns$ and $t(195) = 1.61, ns$, respectively. The contrast for the word *soubrette* was significant, but again not in line with the stimulus material,

Table 3 Study 3: Non-self-reported mood (four word ratings), self-reported post-induction mood, invalid datasets, and duration for forced-choice and free associations to emotive words

	Free association		Free association		Forced association		Forced association	
	positive	neutral	negative	neutral	positive	neutral	negative	neutral
<i>n</i>	74	56	69	61	65	61	55	
Word rating <i>perlator</i> (SD)	5.69 (1.31)	4.70 (1.86)	5.13 (1.80)	4.92 (1.59)	4.92 (1.59)	5.16 (1.77)	5.35 (1.43)	
Word rating <i>aviatics</i> (SD)	4.47 (1.62)	4.07 (1.75)	4.13 (1.43)	4.48 (1.72)	4.48 (1.72)	4.10 (1.46)	4.11 (1.65)	
Word rating <i>brevarium</i> (SD)	5.11 (1.63)	4.89 (1.81)	4.65 (1.66)	5.02 (1.57)	5.02 (1.57)	4.62 (1.28)	4.47 (1.60)	
Word rating <i>soubrette</i> (SD)	4.65 (1.85)	4.21 (2.24)	3.93 (1.81)	4.83 (1.61)	4.83 (1.61)	4.36 (1.65)	3.91 (1.72)	
Post-induction mood (SD)	3.81 (2.09)	3.63 (1.36)	3.57 (1.59)	4.03 (1.53)	4.03 (1.53)	4.07 (1.59)	4.02 (1.71)	
Invalid datasets (%)	13 of 87 (14.9)	7 of 63 (11.1)	14 of 83 (16.9)	8 of 73 (11.0)	8 of 73 (11.0)	10 of 71 (14.1)	11 of 66 (16.7)	
Duration [min:sec] (SD)	5:05 (3:30)	4:05 (2:43)	5:57 (5:45)	2:15 (1:13)	2:15 (1:13)	2:13 (0:53)	2:04 (0:49)	

Note. The mood measure and the word ratings ranged from 1 (positive) to 9 (negative).

$t(195) = 2.20, p = .03$. With regard to the self-reported mood, the contrast among the positive, neutral, and negative mood groups (with coefficients $-1, 0,$ and $+1$) was not significant, $t(196) = .80, ns$. Neither of the two MIPs was significantly more or less effective with men than with women.

In the forced-choice association task, the contrasts for the pleasantness rating of the words *perlator* and *aviatics* among the positive, neutral, and negative mood groups were not significant, $t(178) = 1.43, ns$ and $t(178) = 1.25, ns$, respectively. The contrasts for the words *brevarium* and *soubrette* were significant, but not in line with the stimulus material, $t(178) = 1.99, p = .05$ and $t(178) = 3.03, p = .01$, respectively. With regard to the self-reported mood, the contrast among the positive, neutral, and negative mood groups (with coefficients $-1, 0,$ and $+1$) was not significant, $t(178) = 0.04, ns$. Compliance did not differ among the six conditions, $\chi^2(5, N = 443) = 1.98, ns$.

Discussion

Study 3 tested whether producing associations with emotive words is a useful method to elicit a positive, negative, or neutral mood. Neither the non-self-report word ratings nor the self-report mood measure evidenced any mood effects in line with the stimuli. By contrast, with several word ratings, significant effects appeared that were not in line with the stimuli. One reason for the lack of effectiveness might be that the positive and negative cue words that were used as mood-inducing stimuli might not have significantly differed from the neutral cue words in valence and arousal. Study 5 examines this possibility.

Study 4

This experiment examined picture-illustrated emotive texts as a method of inducing a positive, neutral, and negative mood. This MIP is a mixture between photographs and story MIPs. Stories and photographs have separately been used for mood induction in offline experiments, but not combined. Verheyen and Göritz (2003) identified 13 studies that have used texts for inducing positive or negative mood (i.e., Erber, 1985; Forgas, 1998; Johnson & Tversky, 1983; Knapp, 1986; Popp, 1988). Extracting the effect sizes from these studies and weighting them by their sample size (cf. Hedges & Olkin, 1985, pp. 230–231) yielded a mean effectiveness of $r = .57$ for negative mood inductions and $r = .35$ for positive mood inductions.

Similar to Study 3, the focus in this study was on avoiding demand characteristics to possibly account for any mood difference observed. Therefore, the two non-self-report measures of mood word pleasantness ratings (cf. Study 3) and

subjective probabilities of negative and positive life events were used along with a one-item self-report measure of mood. Johnson and Tversky (1983) have shown that negative mood results in higher estimates of the frequency of many personal risks and other undesirable events. Conversely, positive mood leads to a decrease in the judged frequency of risks. These findings have been endorsed by Wright & Bower (1992), who found happy people to report higher probabilities for positive events and lower probabilities for negative events and sad people to report lower probabilities for positive and higher probabilities for negative events.

Thus, expectations were that participants who had read the positive text rated two unknown neutral words more positively, whereas participants who had read the negative text rated the words more negatively than participants who had read the neutral text. Furthermore, participants who read the positive text were expected to estimate the likelihood of experiencing three positive life events as higher and the likelihood of experiencing three negative events as lower than control participants. Conversely, participants who read the negative text were expected to estimate the likelihood of experiencing three positive events as lower and the likelihood of experiencing three negative events as higher than control participants. Finally, participants in the positive MIP were expected to rate their current mood more positively than participants in the neutral MIP, and participants who underwent the negative MIP to rate their mood more negatively than control participants.

Method

The topic of the study was *context-dependence of the perception of Website contents*. Participants were randomly assigned to one of three conditions, namely to read and evaluate a neutral, funny, or sad text. Each text spanned two pages and featured two images at the end. The neutral text was an excerpt from the first chapter of Hawking (1997). It describes humankind's evolving image of the universe from ancient times to the 17th century. The pictures were two illustrations of the solar system. The funny text consisted of 31 howlers allegedly found by teachers in children's essays. The howlers had been compiled from several humorous Web sites. Two cartoons featuring children accompanied the text. The sad text introduced the controversy surrounding the death penalty and described two electric chair executions. It was based on Wefing (2002) and Kiesow (2002). Two photographs at the end depicted the execution of Allen Lee Davis in 1999. For ethical reasons, people assigned to the negative text were informed that the text-to-come was possibly upsetting, and that they may skip it. Of 154 people who made this choice, 18 (12%) chose to skip the text. They were excluded from further analyses.

After they had finished reading the text, participants rated the pleasantness of the two unfamiliar neutral words *saté* and *soubrette* on a 7-point scale ranging from 1 = *very negative* to 7 = *very positive*. Subsequently participants rated six probabilities that they were to experience (again) six positive or negative life events (i.e., divorce, unexpected fortune, robbery, a long life, accidental fall resulting in a fracture, happiness in old age) on a 9-point scale ranging from 1 = *very likely* to 7 = *very unlikely*. Finally, participants rated their current mood on a one-item scale ranging from 1 = *very negative* to 7 = *very positive*.

Participants

A sample of 1242 panelists was drawn from the online panel that was used for the other studies. Of those, 512 (41%) called up the study's first page. Of the 494 people who did not skip the negative text, 294 (60%) produced valid and 200 invalid datasets: Of the invalid datasets, 77 people abandoned the study prematurely, eight people did not rate the pleasantness of one or both infrequent words, two people did not report their mood, five people took more than 10 min to read one page of the two-page text, and 108 people viewed at least one page of the two-page text for less than 30 sec.⁵ The sub-sample that produced valid data consisted of 143 women (49%) and 151 men. The average age was 27 years ($SD = 10$). Two people had not (yet) completed school, 13 had graduated after 9 years of school, 40 had graduated after 10 years of school, and 239 had earned a high school degree.

Dependent variables

The rating of two unfamiliar neutral words and estimations of the likelihood of experiencing three positive and three negative life events were used as non-self-report measures of mood. Participants' rating of their current mood was used as a self-report measure of mood.

Results

As shown in Table 4, the planned contrast of the rated pleasantness of the word *saté* among the positive, neutral, and negative mood group (with contrast coefficients +1, 0, and -1) was significant, $t(290) = 2.74, p = .01$. The contrast pertaining to *soubrette* was not significant, $t(291) = 1.21, ns$. The contrasts of the negative life events *divorce*, *robbery*, and *accidental fall* among the positive, neutral, and negative mood groups (with coefficients +1, 0, and -1) were not significant, $t(290) = .13, ns$, $t(289) = .18, ns$, and $t(291) = 1.13, ns$, respectively. The contrasts of the positive

⁵ A sensitivity analysis showed that the effects were invariant no matter how strictly the criteria for inclusion were set.

Table 4 Study 4: Non-self-reported mood (two word ratings and six subjective probabilities), self-reported post-induction mood, invalid datasets, and duration for emotive texts

	Positive text	Neutral text	Negative text
<i>n</i>	118	76	100
Word rating <i>saté</i> (<i>SD</i>)	4.56 (1.24)	4.59 (1.25)	4.07 (1.44)
Word rating <i>soubrette</i> (<i>SD</i>)	3.83 (1.09)	4.28 (1.21)	4.02 (1.19)
Probability <i>divorce</i> (<i>SD</i>)	6.56 (2.02)	6.57 (2.25)	6.60 (2.03)
Probability <i>fortune</i> (<i>SD</i>)	6.16 (2.05)	5.76 (2.19)	6.32 (1.95)
Probability <i>robbery</i> (<i>SD</i>)	6.50 (1.57)	6.16 (1.66)	6.54 (1.60)
Probability <i>long life</i> (<i>SD</i>)	4.39 (1.67)	4.26 (1.81)	4.43 (1.64)
Probability <i>accident</i> (<i>SD</i>)	4.81 (1.90)	4.76 (1.92)	4.51 (1.95)
Probability <i>happiness</i> (<i>SD</i>)	3.50 (1.66)	3.67 (1.87)	3.76 (1.74)
Post-induction mood (<i>SD</i>)	4.97 (1.34)	4.66 (1.28)	4.38 (1.48)
Invalid datasets (%)	50 of 168 (29.8)	96 of 172 (55.8)	54 of 154 (35.1)
Duration [min:sec] (<i>SD</i>)	4:10 (1:43)	3:51 (2:00)	3:42 (1:26)

Note. The mood measure and the word ratings ranged from 1 (negative) to 7 (positive). The probability ratings ranged from 1 (likely) to 9 (unlikely).

life events *unexpected fortune*, *long life*, and *happiness in old age* among the positive, neutral, and negative mood group (with coefficients -1 , 0 , and $+1$) were not significant, $t(291) = .57$, *ns*, $t(291) = .17$, *ns*, and $t(290) = 1.08$, *ns*, respectively. The contrast in self-reported mood among the positive, neutral, and negative mood group (with coefficients $+1$, 0 , and -1) was significant, $t(291) = 3.09$, $p = .002$. The effect size for the induction of a positive mood relative to a neutral mood was $r = .12$, and the effect size for the induction of a negative mood relative to a neutral mood was $r = .10$. Compliance differed among the three conditions, $\chi^2(2, N = 494) = 26.67$, $p < .001$. The effectiveness of the MIP did not differ with participants' sex.

Discussion

Study 4 tested whether picture-illustrated emotive texts successfully induce a positive, neutral, and negative mood. One of the pleasantness ratings and the self-reported mood evidenced a significant mood shift in the anticipated direction. However, subjective probabilities of negative and positive life events yielded no evidence of a mood change. The effect sizes found in this study are much smaller than effect sizes gained from the meta-analysis by Verheyen and Göritz (2003). The induced shift towards a positive mood in this study was $r = .12$ compared to $r = .35$, and the shift towards a negative mood in this study was $r = .10$ compared to $r = .57$ in Verheyen and Göritz (2003).

Despite the success of the mood induction with picture-illustrated text, this online MIP is to be used with caution. The percentage of invalid datasets was not evenly distributed among the three mood conditions. Compliance was highest in the positive mood condition and lowest in the neutral mood condition. Many volunteer participants might find it unacceptable to read a lengthy and possibly boring text on the screen.

Study 5

In Study 1 and in Study 3, mood-suggestive photographs and emotion-laden cue words were used as mood-inducing material. Before their use in Study 1 and 3, these researcher-developed stimuli had not undergone a systematic rating for valence and arousal by a sufficiently large and heterogeneous group of people. The lack of effectiveness of the positive mood induction with mood-suggestive photographs in Study 1 and Göritz and Moser (2006) might be due to having used positive and negative photographs that largely differed in arousal and absolute valence. The lack of effectiveness of both the positive and negative mood induction with cue words in Study 3 might be due to having used stimuli that did not have a significant valence in the expected direction or might have been very low in arousal. To shed more light on the ineffectiveness of the positive mood induction with regard to the stimuli, Study 5 obtained ratings for valence and arousal for all photographs used in Study 1 and for the words used in Study 3.

Method

Participants were randomly assigned to one of five conditions; namely, they either rated the 57 negative photographs or the 58 positive photographs used in Study 1 and in Göritz and Moser (2006), or they rated the 10 negative, 10 neutral, or 10 positive cue words used in Study 3. People rated each stimulus' valence on a 9-point scale ranging from $-4 = \textit{extremely negative}$ to $+4 = \textit{extremely positive}$ and arousal on a 9-point scale ranging from $0 = \textit{not at all aroused}$ to $8 = \textit{extremely aroused}$. Each stimulus appeared on a separate page. The order of the stimuli was the same as in Study 1 and 3, respectively. For ethical reasons, people assigned to rate the negative photographs were informed beforehand that some of the photographs were graphic, and if they did not want to see the negative photographs they may shift to

positive pictures. Of 185 people who made this choice, 8 (4%) switched to the positive photographs. They were excluded from further analyses.

Participants

A sample of 2456 panelists, who were at least 18 years of age, was randomly drawn from the online panel that was used for the other studies. Of those, 1025 (42%) called up the study’s first page. Of the 1017 people who did not switch to the positive photographs, 845 (83%) produced valid and 172 invalid datasets: Of the invalid datasets, 163 people abandoned the study prematurely, two people did not rate six or more photographs, and seven people did not rate two or more cue words. Missing stimulus ratings by participants who skipped less than six photographs or one word were replaced by the mean rating of that stimulus. The sub-sample that produced valid data consisted of 525 women (62%) and 320 men. The average age was 36 years ($SD = 12$). Six people had not (yet) completed school, 93 had graduated after 9 years of school, 256 had graduated after 10 years of school, and 490 had earned a high school degree.

Variables of interest

It was examined whether (1) arousal and absolute valence ratings of the 58 positive photographs were comparable to those of the 57 negative photographs used in the long Photo MIP in Study 1, (2) arousal and absolute valence ratings of the 29 positive photographs were comparable to those of the 29 negative photographs used in the short Photo MIP in Study 1, and (3) valence and arousal of both the ten negative and the ten positive words differed from those of the ten neutral cue words used in Study 3 and whether arousal and absolute valence of the 10 positive words were comparable to those of the 10 negative words.

Results

As shown in Table 5, the absolute value of the average valence per photograph was weaker with the 58 positive photographs than with the 57 negative photographs that were used in the long Photo MIP in Study 1, $t(303) = 8.80$, $p < .001$. The size of this effect is $r = .44$. Similarly, the mean arousal per photograph was weaker with the 58 positive photographs than with the 57 negative photographs, $t(303) = 7.85$, $p < .001$, $r = .41$. With the subset of the 29 photographs that were used in the short Photo MIP, the absolute value of the valence per photograph was weaker with the positive photographs ($M = 1.3$, $SD = 0.9$) than with the negative photographs ($M = -1.8$, $SD = 0.7$), $t(303) = 5.69$, $p < .001$, $r = .30$. The mean arousal per photograph was weaker with the 29 positive photographs

Table 5 Study 5: Average valence and arousal per stimulus

	Photographs positive	Photographs negative	Cue words positive	Cue words neutral	Cue words negative
<i>n</i>	174	131	175	177	188
Mean valence per stimulus (<i>SD</i>)	1.1 (0.8)	-1.9 (0.7)	2.7 (0.8)	0.5 (0.6)	-3.1 (0.7)
Mean arousal per stimulus (<i>SD</i>)	2.1 (1.4)	3.3 (1.2)	3.6 (1.5)	1.7 (1.3)	4.8 (2.0)

Note. The valence ratings ranged from -4 (negative) to +4 (positive). The arousal ratings ranged from 0 (not aroused) to 8 (aroused).

($M = 2.3$, $SD = 1.5$) than with the 29 negative photographs ($M = 3.2$, $SD = 1.2$), $t(303) = 5.91$, $p < .001$, $r = .32$.

As also shown in Table 5, the mean valence per word was different between the ten negative and neutral words, $t(363) = 49.75$, $p < .001$, $r = .93$, and between the neutral and positive words, $t(350) = 28.09$, $p < .001$, $r = .83$. The absolute value of the average valence per stimulus was weaker with the positive words than with the negative words, $t(361) = 5.10$, $p < .001$, $r = .26$. The mean arousal per word was different between the negative and neutral words, $t(363) = 17.68$, $p < .001$, $r = .68$, and between the neutral and positive words, $t(350) = 12.24$, $p < .001$, $r = .55$. The mean arousal per stimulus was weaker with the positive words than with the negative words, $t(361) = 6.50$, $p < .001$, $r = .32$.

Discussion

Study 5 revealed that the negative and positive photographs used in Study 1 and in Göritz and Moser (2006) and the negative and positive cue words used in Study 3 were not comparable with regard to absolute valence and arousal. The negative stimuli were consistently stronger in absolute valence and arousal than the positive stimuli.

Consequently, in future studies — especially if one is interested in the comparative effectiveness of positive and negative mood conditions within the same MIP — stimuli should be used that are equated for valence and arousal. Instead of home-developing mood-inducing materials — as was done in these studies due to finding out too late that standardized stimuli existed — it is recommended to use existing material that is calibrated, such as the *International Affective Picture System* (Lang, Bradley, & Cuthbert, 2005) or the *Affective Norms for English Words* (Bradley & Lang, 1999). Moreover, using ready-made materials facilitates the comparison of results across studies.

In view of the positive valence of the positive photographs used in Study 1 it would have been expected that undergoing the positive Photo MIP would have mildly improved participants' mood. Instead, the positive Photo MIP in part even led to a deterioration of mood. This finding becomes plausible when assuming that the dose-response-relationship of stimulus-induced mood change is not linear or that this relationship is asymmetric for positive and negative mood, or both.

Asymmetries in the effectiveness of positive and negative mood-inducing stimuli have also been found in Study 1 and Göritz and Moser (2006) with the Velten MIP: The negative Velten MIP was effective, whereas the positive Velten MIP was ineffective. Moreover, in their meta-analysis Westermann et al. (1996) calculated a lower mean weighted effect size for the induction of positive ($r = .41$) than negative mood ($r = .53$). Although in some of the summarized

studies in this meta-analysis the stimuli may not have been equated for valence and arousal, thereby accounting for the difference in overall effectiveness, it is perhaps no accident that positive stimuli regularly seem to be weaker in absolute valence and arousal than negative stimuli. For future studies it therefore might be expected that within the same MIP — even if positive and negative stimuli are matched for valence and arousal — the induction of positive mood will in general be less effective than the induction of negative mood.

The lack of effectiveness of both the positive and negative mood induction with emotion-laden words in Study 3 does not seem to be due to having used unsuitable stimuli. Study 5 revealed that although positive cue words were weaker in valence and arousal than negative words, both positive and negative words were significantly different from the neutral cue words in expected valence and arousal. Thus, there are other reasons why this MIP failed (see Study 3 for a discussion).

General discussion

Six mood induction procedures were developed or adapted for use on the WWW: the Velten procedure, mood-suggestive photographs, jokes, cartoons, associations to affectively valenced cue words, and picture-illustrated emotive texts. The Velten method and mood-suggestive photographs successfully induced a negative mood, but failed to induce a positive mood. Cartoons successfully induced a positive mood, whereas jokes were ineffective. Associations to affectively valenced words were not suitable in calling forth a positive or negative mood. Finally, picture-illustrated emotive texts successfully induced a positive and negative mood. The effect sizes for inducing a positive mood achieved with cartoons and emotive texts were small.

The lacking effectiveness of the positive Velten procedure, pleasant photographs, associations to affectively positive words, and jokes as well as only modest effects of cartoons and emotive texts in inducing a positive mood might at least in part be explained by online participants' positive baseline mood, rendering any further mood lifting difficult: The norm sample of the mood scale *SES* (Hampel, 1977) recorded an average positive mood score of 22.0 ($SD = 10.7$; $N = 846$), whereas participants' pre-induction mood was as positive as 29.3 ($SD = 7.8$; $N = 203$) in Study 1 and 29.3 ($SD = 8.1$; $N = 463$) in Study 2. This is a difference of $r = .27$ and $r = .33$, respectively.

From that, it seems that participants' baseline mood on the WWW is generally more positive than that of participants in the laboratory. Either the online participants' mood had been improved through their prior surfing on the Web, or a selection effect was at play, in the sense that people in a positive mood are more likely to take part in online

studies. Using the WWW for leisure might have a similar effect as undergoing a positive MIP because the Internet is a tool for engaging in several types of rewarding behavior such as being in a relationship (Griffiths, 1999). Indeed, Novak, Hoffman, and Yung (2000) have extended the Flow construct (cf. Csikszentmihalyi & LeFevre, 1989) to encompass navigation behavior on the WWW. An additional reason for the better baseline mood of WWW participants could be that taking part in online studies might on average be more pleasant than taking part in offline studies because it is often more voluntary and less tedious. In contrast, offline participants frequently take part for course credit, and they need to spend more time physically getting to the laboratory. Moreover, if people take part in online studies for little or no remuneration they might boost their mood by attributing their efforts to their helpfulness (Harris, 1978; Williamson & Clark, 1989) or by perceiving their participation in such a study as self-rewarding (Festinger, 1962). Moreover, participants in Studies 1 and 2 assessing their initial mood on a mood scale consisting of seven positively tinged adjectives might have exerted an auto-suggestive effect on their mood. If it proves to be a robust effect that online participants are in a more positive baseline mood than offline participants, the WWW could be of limited use for positive mood manipulations on account of the ceiling effect which would make it difficult to further improve their already positive mood. However, more experiments using different MIPs and mood measures are needed to examine this conjecture further.

Contrary to initial concerns, participants did not produce invalid data more often if undergoing a negative mood induction. Instead, participants tended to be less compliant if they underwent a neutral MIP. It is a challenge for the future to create experimental materials that are affectively neutral but at the same time involving enough to dissuade volunteer respondents from prematurely abandoning or flippantly clicking through a study.

The focus of the present studies was to test the feasibility of Web-based mood induction. The optimal procedural variants of these MIPs, in terms of length, instructions, and stimuli, were probably not yet chanced upon in the present series of studies. Honing these MIPs remains a goal for future studies. In doing so, it is desirable to use mood-inducing stimuli that have been standardized for valence and arousal.

Moreover, new MIPs (e.g., based on Virtual Reality) or established ones such as films and music (Westermann et al., 1996) might be used on the WWW. With the wide penetration of high-bandwidth connections and state-of-the-art computers, such MIPs might be used with special samples even today (e.g., volunteers from an online panel who might get screened for compliance with technological requirements). With general samples, it is a question of time when these types of MIPs can be used. Another task is to search for reliable and sensitive non-self-report measures of mood that

lend themselves to be used in the framework of Web-based mood research.

Recommendations for inducing mood via the WWW

The outcome of the present studies allows giving preliminary suggestions for both Web-based mood induction in general and for the examined procedures' practical application in particular: Both the negative Velten and Photo technique qualify to be employed as negative mood inducers. For the induction of a positive mood one might opt for cartoons or an emotive text. If one wishes to elicit both a positive and a negative mood with comparable materials, the only validated web-based MIP to-date relies on picture-illustrated emotive texts.

Besides its effectiveness, the extent to which a technique meshes with both the cover story and the ensuing task should influence the decision of its use. Furthermore, knowing the duration of the altered mood state helps in selecting a procedure. For example, Frost and Green (1982) found that after a 10 min waiting period, negative mood induced with the offline Velten MIP was significantly diminished, and positive mood had completely dissipated relative to a neutral mood condition. Using the offline Velten MIP, Isen and Gorgoglione (1983) found no difference in self-reports between participants in positive and dysphoric mood conditions after participating in a 4 min post-induction word categorization task. These results point to the need to examine the duration of altered mood states when mood is manipulated with the help of Web-based MIPs.

On account of operational particularities of collecting data via the WWW, the mean effects found with online MIPs are likely to be smaller than those found for offline MIPs. Random error tends to be greater on the WWW than in comparable laboratory experiments due to participants' diverse Internet connections and IT equipment, variation in the time of day they take place, and less control over factors influencing their mental presence while participating (McGraw et al., 2000; Shavit, Sonsino, & Benzion, 2001).

Acknowledgments This work has been supported by a University of Erlangen-Nürnberg post-doctoral scholarship (HWP). I thank Klaus Moser for helpful comments on two of the studies presented in this paper and for providing the institutional backdrop for all five studies.

References

- Birnbaum, M. H. (2004). Human research and data collection via the internet. *Annual Review of Psychology*, 55, 803–832.
- Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Stimuli, instruction manual and affective ratings*.

- Technical report C-1, Gainesville, FL. The Center for Research in Psychophysiology, University of Florida.
- Carnevale, P. J. D., & Isen, A. M. (1986). The influence of positive affect and visual access on the discovery of integrative solutions in bilateral negotiation. *Organizational Behavior and Human Decision Processes*, *37*, 1–13.
- Csikszentmihalyi, M., & LeFevre, J. (1989). Optimal experience in work and leisure. *Journal of Personality and Social Psychology*, *56*, 815–822.
- Erber, R. (1985). *Choosing among multiple categories: The effect of mood on category accessibility, inference, and affect*. Unpublished doctoral dissertation, Pittsburg, PA: Carnegie-Mellon University.
- Festinger, L. (1962). Cognitive dissonance. *Scientific American*, *207*, 93–107.
- Forgas, J. P. (1998). Asking nicely? The effects of mood on responding to more or less polite requests. *Personality and Social Psychology Bulletin*, *24*, 173–185.
- Forgas, J. P. (2002). Feeling and doing: Affective influences on interpersonal behaviour. *Psychological Inquiry*, *13*(1), 1–28.
- Frost, R. O., & Green, M. L. (1982). Velten mood induction procedure effects: Duration and postexperimental removal. *Personality and Social Psychology Bulletin*, *8*, 341–347.
- Göritz, A. S. (2004). Recruitment for online access panels. *International Journal of Market Research*, *46*, 411–425.
- Göritz, A. S. (in press). Using online panels in psychological research. In A. Joinson, K. McKenna, T. Postmes, & U.-D. Reips (Eds.), *Oxford handbook of internet psychology*. Oxford: Oxford University Press.
- Göritz, A. S., & Moser, K. (2006). Web-based mood induction. *Cognition and Emotion*, *20*, 887–896.
- Gosling S. D., Vazire S., Srivastava S., & John, O. P. (2004). Should We Trust Web-Based Studies? A Comparative Analysis of Six Preconceptions About Internet Questionnaires. *American Psychologist*, *59*, 93–104.
- Griffiths, M. (1999). Internet addiction: Fact or fiction? *Psychologist*, *12*, 246–250.
- Hampel, R. (1977). Adjektiv-Skalen zur Einschätzung der Stimmung (SES) [Adjective scales to assess mood]. *Diagnostica*, *23*, 43–60.
- Harris, M. B. (1978). Effects of altruism on mood. *Journal of Social Psychology*, *102*, 197–208.
- Hawking, S. (1997). *Die Illustrierte Kurze Geschichte der Zeit* [The illustrated A brief history of time]. Reinbek, Germany: Rowohlt.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. New York: Academic Press.
- Isen, A. M., & Gorgoglione, J. M. (1983). Some specific effects of four affect-induction procedures. *Personality and Social Psychology Bulletin*, *9*, 136–143.
- Isen, A. M., Johnson, M. M. S., Mertz, E., & Robinson, G. F. (1985). The influence of positive affect on the unusualness of word associations. *Journal of Personality and Social Psychology*, *48*, 1413–1426.
- Johnson, E. J., & Tversky, A. (1983). Affect, generalization, and the perception of risk. *Journal of Personality and Social Psychology*, *45*, 20–31.
- Kiesow, R. M. (2002). Überwachen und Rächen. Todesstrafe live [Surveilling and avenging. Death penalty live]. *Kursbuch*, *147*, 30–38.
- Knapp, A. (1986). Die Auswirkungen emotionaler Zustände auf das Lösen eines sozialen Dilemmas [The effects of emotional states on the solving of a social dilemma]. *Zeitschrift für Sozialpsychologie*, *17*, 160–172.
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*. Technical Report A-6. Gainesville, FL: University of Florida.
- Larsen, R. J., & Sinnett, L. M. (1991). Meta-analysis of experimental manipulations: Some factors affecting the Velten mood induction procedure. *Personality and Social Psychology Bulletin*, *17*, 323–334.
- Linden, W., & Frankish, J. (1988). Expectancy and type of activity: Effects on pre-stress cardiovascular adaptation. *Biological Psychology*, *27*, 227–235.
- McArthur, L. Z., Solomon, M. R., & Jaffe, R. H. (1980). Weight differences in emotional responsiveness to proprioceptive and pictorial stimuli. *Journal of Personality and Social Psychology*, *39*, 308–319.
- McGraw, K. O., Tew, M. D., & Williams, J. E. (2000). The integrity of Web-delivered experiments: Can you trust the data? *Psychological Science*, *11*, 502–506.
- Novak, T. D., Hoffman, D. L., & Yung, Y. F. (2000). Measuring the customer experience in online environments: A structural modeling approach. *Marketing Science*, *19*, 22–42.
- Popp, M. (1988). Wirkung von Informationsbedingung, emotionalem Zustand und externer Speichermöglichkeit auf das Lösen sozialer Fallen [Effect of information condition, emotional state, and possibility of external storage on the solving of social traps]. *Archiv für Psychologie*, *140*, 33–51.
- Sabini, J., & Silver, M. (2005). Why emotion names and experiences don't neatly pair. *Psychological Inquiry*, *16*, 1–10.
- Shavit, T., Sonsino, D., & Benzion, U. (2001). A comparative study of lotteries-evaluation in class and on the Web. *Journal of Economic Psychology*, *22*, 483–491.
- Velten, E. (1968). A laboratory task for induction of mood states. *Behavior Research and Therapy*, *6*, 473–482.
- Verheyen, C., & Göritz, A. S. (2003). *Text-Based Online Mood Induction*. Final thesis by first author, supervised by second author. University of Erlangen-Nuremberg, Department of Organizational and Social Psychology.
- Wefing, H. (2002). Die Lebenskandidaten. Wendet Amerika sich wirklich von der Todesstrafe ab? [Candidates for life. Does America really turn away from capital punishment?] *Frankfurter Allgemeine Zeitung*, *26.6.2002*, *145*, p. 45.
- Westermann, R., Spies, K., Stahl, G., & Hesse, F. W. (1996). Relative effectiveness and validity of mood induction procedures: A meta-analysis. *European Journal of Social Psychology*, *26*, 557–580.
- Williamson, G. M., & Clark, M. S. (1989). Providing help and desired relationship type as determinants of changes in moods and self-evaluations. *Journal of Personality and Social Psychology*, *56*, 722–734.
- Wright, W. F., & Bower, G. H. (1992). Mood effects on subjective probability assessment. *Organizational Behavior and Human Decision Processes*, *52*, 276–291.