Gender-related Differences in the Production and Perception of Emotion

Marc Swerts and Emiel Krahmer
Communication and Cognition
Tilburg University, The Netherlands
{m.q.j.swerts/e.j.krahmer}@uvt.nl

Abstract
This article discusses a study into gender-related differences in the experience, display and perception of positive and negative emotions. The problem is addressed by a combination of a production and a perception experiment. The production study makes use of a mood induction procedure which aims to elicit positive and negative emotions in subjects by letting them view different kinds of movie clips. This method turned out to be extremely effective, and also revealed that women experience emotions more intensively than men. Video clips collected through the production study were then shown to independent observers, whose task it was to judge whether the recorded speakers were in a negative or positive condition. Overall, judges were very accurate in estimating whether a displayed speaker was positive or negative. Moreover, there were again gender differences in that female speakers were more expressive than male speakers, and female observers were more accurate in their judgments than male observers.

Index Terms: emotion, gender, facial expression, production, perception

1. Introduction
There is a whole body of research which suggests that men differ from women in the way they deal with emotions, where the stereotypical view is that women are more emotional than men. There is an ongoing debate whether such differences are due to sociocultural, cognitive or biological factors (Knapp & Hall 2006). On the one hand, it has been argued that one reason why women have better emotional skills than men is because the former make better use of specific designated brain areas (e.g. Schirmer & Kotz 2006). On the other hand, the extent to which women differ from men in how they deal with emotions could also be biased by a specific culture (Eifenbein et al 2002) or situational context (Ekman 1972), and should be nuanced for specific emotions (e.g. anger is better expressed by men, and sadness by women) (Rotter & Rotter 1988).

Despite such previous work, our knowledge of gender differences in emotion is still incomplete. As a matter of fact, the claim that women are more emotional than men could mean different things: it could refer to the fact that women experience emotions more intensively than men, to the fact that they display such emotions more clearly, or to the fact that they are better in judging emotions in others (Knapp & Hall 2006). As yet, there are virtually no studies which provide an integrated view different kinds of movie clips. This method turned out to be extremely effective, and also revealed that women experience emotions more intensively than men. Video clips collected through the production study were then shown to independent observers, whose task it was to judge whether the recorded speakers were in a negative or positive condition.

2. Production experiment
2.1. Goal
The current experiment aims to induce either “joy” or “sadness” in female and male participants through the use of a film mood induction procedure. The goal is to find out whether there is any evidence that the two genders differ in the extent to which they experience a positive or a negative feeling.

2.2. Method
The experiment consisted of a number of steps, and was conducted on an individual basis. Participants were first informed by written instruction about the fact that they were going to participate in an experiment whose aim it is to investigate how emotions could affect their choice behavior when confronted with certain dilemmas. The reason to give them this instruction with dilemmas was motivated by the claim of Westermans et al (1996) that mood induction procedures are more effective when participants are given a clear goal. Participants were told that they would see a 7-minute fragment of a movie to elicit a certain emotion. After the instruction phase, the participants first had to respond to 3 different dilemmas, such as: “You had in-
positive male  positive female
negative male  negative female

Figure 1: Representative stills from male and female participants in positive and negative conditions (while they watch a movie)

tended to go out for working out, but as soon as you leave your house, it starts raining. What will you do?”, after which they got 2 response options: A “You get on your bike, and you go ahead anyway”; or B “You decide to wait for 15 minutes, and if the weather hasn’t changed after that, you cancel.”

In the next stage, participants had to watch the movie fragments. In the positive condition, participants viewed the first 7 minutes, excluding the begin tune, of episode 5.14 of Friends (The one where everyone finds out). Participants in the negative condition watched a 7-minute fragment (the “liquidation of the ghetto, March 1943”-scene) of Schindler’s list, corresponding to scene 13/14 of Disc 1 of the commercial dvd. After having watched either of these two film clips, participants were asked to complete a short written questionnaire, based on a larger questionnaire adapted from Krahmer, van Dorst and Ummelen (2004). The questionnaire consisted of 6 bipolar 7-point scales (Dutch), the items of which can be translated into English as follows: happy/sad, pleasant/unpleasant, satisfied/unsatisfied, content/discontent, cheerful/sullen and high spirits/low-spirited. For 4 scales, “1” corresponded to a very positive and “7” to a very negative condition, whereas for the other 2 scales, the extremes of the scales had opposite interpretations; the variable interpretation of the extremes of the scale was introduced to keep the participants alert. Finally, the participants were interviewed about the film-fragment they had just seen, after which they were given 3 other dilemmas similar to the ones given before the movie. The whole procedure took about 30 minutes per participant. Note that participants were filmed by a visible camera while they were watching the movie fragments, during the interview session and when responding to the dilemmas given before and after the movie fragment. Figure 1 gives some representative stills of male and female participants in positive and negative conditions (where the participants were watching a happy or sad movie).

2.2.1. Participants

33 participants (17 female, 16 male), all students from the Arts faculty at Tilburg University, between the age of 18 and 30, took part in the experiment on a voluntary basis: 17 of them (8 male, 9 female) were randomly assigned to the positive condition, 16 (8 male, 8 female) to the negative condition. All of them signed a written consent form in which they agree that the video recordings of their reactions during the movie fragments, the interview and the session in which they responded to the dilemmas could be analysed for academic purposes and shown to third parties. After the experiment, all participants received a candy bar to thank them for their participation.

2.2.2. Design

The experiment had a between-subjects design with 2 main factors, i.e., condition (positive, negative) and gender (female, male). The choice to have a between- rather than a within-subject design is that it would seem very difficult to have participants switch between “joy” and “sadness” in the same experiment.

2.3. Results

Before testing the effect of the positive and negative conditions on the experienced moods, the internal consistency between the scores of our 6 scales was tested by means of Cronbach’s $\alpha$. To this end, the scores for 2 out of the 6 scales were reversed to make sure the extremes would have similar interpretations (positive or negative) with 1 being “very positive” and 7 being “very negative”. Cronbach’s $\alpha$ turned out to be very high (0.93) which points out that the internal consistency between the different scales was very good. In the remaining analyses, the scores for the different scales were first pooled, in that we got one average score per participant.

We performed an analysis of variance with condition (negative, positive) and gender (female, male) as independent factors, and the average score per participant as dependent variable. This analysis revealed a significant main effect of condition ($F(1,29) = 58.845, p < .001, \eta^2_p = .662$), whereas the main effect of gender turned out not to be significant. The condition effect was such that the positive condition led to a significantly more positive score of 2.259 (std. error = .182) than the negative score of 4.177 (std. error = .177). In addition to the main effect, there was a significant 2-way interaction between condition and gender ($F(1,29) = 6.527, p < .05, \eta^2_p = .184$). Even though split analyses revealed that the difference between positive and negative scores was significant for both the male ($t = -2.682, df = 14, p < .05$) and the female participants ($t = -11.977, df = 15, p < .001$), it is clear from Table 1 that the difference between the positive and the negative conditions is bigger for female (2.567) than for male participants (1.268).

The production experiment thus revealed that the film mood induction procedure is extremely effective to make participants feel joyful or sad by means of film clips. Interestingly, the

<table>
<thead>
<tr>
<th>Condition</th>
<th>Positive (1)</th>
<th>Negative (2)</th>
<th>Diff (2-1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>2.625 (.258)</td>
<td>3.893 (.258)</td>
<td>1.268</td>
</tr>
<tr>
<td>Female</td>
<td>1.893 (.258)</td>
<td>4.460 (.243)</td>
<td>2.567</td>
</tr>
</tbody>
</table>

Table 1: Average self-ratings, standard deviations and difference scores from male and female participants in positive and negative conditions (1=very positive to 7=very negative)
method also brought to light gender differences in that female participants appeared to produce more extreme scores on the valency scales than the male participants, meaning that the former experienced the different moods more extremely than the latter. Given these production results, it remains to be seen whether the differences are also perceived as such by independent observers, and if so, whether the gender of observers matters there as well.

3. Perception experiment

3.1. Goal

Given the results of the previous experiment about gender differences in the way positive or negative conditions are experienced, the current experiment aims to test whether the different experiences are reflected in gender differences in the expression and perception of these conditions.

3.2. Method

The stimulus materials were selected from the video recordings of the participants of the earlier production study. Using the Adobe Premiere software, we cut 10-second fragments from exactly the middle of the recording of participants watching the 7-minute movie scenes. In addition, we took 10 seconds form the middle of the interview session conducted after the movie session and the completion of the questionnaire. The audio-channel of both kinds of fragments were removed from the clips, so that observers had to base their judgments on the visual information only.

All the fragments (from movie and interview segments) were presented in random order to observers (individually performed experiment), who were asked to judge the emotional content of the presented clip on a 5-point scale (with 1 meaning “very positive” and 5 meaning “very negative”). To compensate for possible learning effects, half of the participants received the clips in one order, whereas the other half received them in the opposite order. The participants saw a clip after which they had 3 seconds to express their judgment on the scale (paper-and-pencil task). To make subjects acquainted with the experimental procedure and the kinds of stimuli, the experiment was preceded by a short practice session where they had to judge 2 clips. After the practice session, there was no further interaction with the experimenters. The whole experiment (including the differences are also perceived as such by independent observers, and if so, whether the gender of observers matters there as well.

3. Perception experiment

3.1. Goal

Given the results of the previous experiment about gender differences in the way positive or negative conditions are experienced, the current experiment aims to test whether the different experiences are reflected in gender differences in the expression and perception of these conditions.

3.2. Method

The stimulus materials were selected from the video recordings of the participants of the earlier production study. Using the Adobe Premiere software, we cut 10-second fragments from exactly the middle of the recording of participants watching the 7-minute movie scenes. In addition, we took 10 seconds form the middle of the interview session conducted after the movie session and the completion of the questionnaire. The audio-channel of both kinds of fragments were removed from the clips, so that observers had to base their judgments on the visual information only.

All the fragments (from movie and interview segments) were presented in random order to observers (individually performed experiment), who were asked to judge the emotional content of the presented clip on a 5-point scale (with 1 meaning “very positive” and 5 meaning “very negative”). To compensate for possible learning effects, half of the participants received the clips in one order, whereas the other half received them in the opposite order. The participants saw a clip after which they had 3 seconds to express their judgment on the scale (paper-and-pencil task). To make subjects acquainted with the experimental procedure and the kinds of stimuli, the experiment was preceded by a short practice session where they had to judge 2 clips. After the practice session, there was no further interaction with the experimenters. The whole experiment (including the differences are also perceived as such by independent observers, and if so, whether the gender of observers matters there as well.

3. Perception experiment

3.1. Goal

Given the results of the previous experiment about gender differences in the way positive or negative conditions are experienced, the current experiment aims to test whether the different experiences are reflected in gender differences in the expression and perception of these conditions.

3.2. Method

The stimulus materials were selected from the video recordings of the participants of the earlier production study. Using the Adobe Premiere software, we cut 10-second fragments from exactly the middle of the recording of participants watching the 7-minute movie scenes. In addition, we took 10 seconds form the middle of the interview session conducted after the movie session and the completion of the questionnaire. The audio-channel of both kinds of fragments were removed from the clips, so that observers had to base their judgments on the visual information only.

All the fragments (from movie and interview segments) were presented in random order to observers (individually performed experiment), who were asked to judge the emotional content of the presented clip on a 5-point scale (with 1 meaning “very positive” and 5 meaning “very negative”). To compensate for possible learning effects, half of the participants received the clips in one order, whereas the other half received them in the opposite order. The participants saw a clip after which they had 3 seconds to express their judgment on the scale (paper-and-pencil task). To make subjects acquainted with the experimental procedure and the kinds of stimuli, the experiment was preceded by a short practice session where they had to judge 2 clips. After the practice session, there was no further interaction with the experimenters. The whole experiment (including the differences are also perceived as such by independent observers, and if so, whether the gender of observers matters there as well.

3. Perception experiment

3.1. Goal

Given the results of the previous experiment about gender differences in the way positive or negative conditions are experienced, the current experiment aims to test whether the different experiences are reflected in gender differences in the expression and perception of these conditions.

3.2. Method

The stimulus materials were selected from the video recordings of the participants of the earlier production study. Using the Adobe Premiere software, we cut 10-second fragments from exactly the middle of the recording of participants watching the 7-minute movie scenes. In addition, we took 10 seconds form the middle of the interview session conducted after the movie session and the completion of the questionnaire. The audio-channel of both kinds of fragments were removed from the clips, so that observers had to base their judgments on the visual information only.

All the fragments (from movie and interview segments) were presented in random order to observers (individually performed experiment), who were asked to judge the emotional content of the presented clip on a 5-point scale (with 1 meaning “very positive” and 5 meaning “very negative”). To compensate for possible learning effects, half of the participants received the clips in one order, whereas the other half received them in the opposite order. The participants saw a clip after which they had 3 seconds to express their judgment on the scale (paper-and-pencil task). To make subjects acquainted with the experimental procedure and the kinds of stimuli, the experiment was preceded by a short practice session where they had to judge 2 clips. After the practice session, there was no further interaction with the experimenters. The whole experiment (including the differences are also perceived as such by independent observers, and if so, whether the gender of observers matters there as well.

Table 2: Average scores (1=very positive to 5=very negative) for different instances of condition, fragment, gender speaker and gender observer

<table>
<thead>
<tr>
<th>Factor</th>
<th>Level</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition</td>
<td>Positive</td>
<td>2.623</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>3.715</td>
</tr>
<tr>
<td>Fragment</td>
<td>Movie</td>
<td>3.468</td>
</tr>
<tr>
<td></td>
<td>Interview</td>
<td>2.870</td>
</tr>
<tr>
<td>Gender speaker</td>
<td>Male</td>
<td>3.302</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>3.036</td>
</tr>
<tr>
<td>Gender observer</td>
<td>Male</td>
<td>3.159</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>3.178</td>
</tr>
</tbody>
</table>

Table 3: Observed scores, standard deviations and difference scores for male and female speakers in positive and negative conditions (1=very positive to 5=very negative)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Speaker</th>
<th>Positive (1)</th>
<th>Negative (2)</th>
<th>Diff. (2-1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>2.941 (.081)</td>
<td>3.662 (.086)</td>
<td>.721</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>2.304 (.090)</td>
<td>3.768 (.101)</td>
<td>1.464</td>
</tr>
</tbody>
</table>

practice session and instructions) lasted about 15 minutes.

3.3. Participants

24 participants, equally balanced across gender and presentation order, all students from the Arts faculty at Tilburg University between the ages of 18 and 30, took part in the perception study on a voluntary basis. None of them had participated as speakers in the production study.

3.4. Design

The design of this experiment was a mixed design with gender of the observer (female, male) as a between-subjects factor and gender of speaker (female, male), film fragment (movie, interview) and condition (positive, negative) as within-subject factors. Since order of presentation did not have a significant effect on the results, this factor will be skipped for the remaining analyses.

3.5. Results

The data were analysed with a 2 × 2 × 2 × 2 repeated measurements anova with gender of observer (male, female) as a between-subjects factor, gender of speaker (male, female), condition (positive, negative) and fragment (movie, interview) as within-subject factors, and the perceived condition (on a scale from 1 to 5) as dependent variable. The analysis revealed main effects of gender of speaker (F(1,22) = 10.985, p < .01, ηp2 = .333), condition (F(1,22) = 271.256, p < .001, ηp2 = .925), and fragment (F(1,22) = 70.564, p < .001, ηp2 = .762), whereas gender of observer turned out not to be significant. From table 2, it can be seen that the significant results can be interpreted as follows: fragments from a negative condition are indeed perceived as being more negative than the positive ones. In addition, it appears that fragments from the movie sequence are perceived on average as being more negative than the interview fragments. And finally, male speakers are on average perceived as being slightly more negative than the female speakers. In addition there turned out to be significant 2-way interactions between mood and gender of speaker (F(1,22) = 52.699, p < .001, ηp2 = .705), between fragment and gender of speaker (F(1,22) = 130.737, p < .001, ηp2 = .856) and between fragment and gender of observer (F(1,22) = 7.694, p < .05, ηp2 = .259), whereas the interaction between condition and gender of observer was marginally significant (F(1,22) = 4.010, p = .058, ηp2 = .154). All other 2-way and higher-level interactions were not significant. Of particular interest for our research are the interactions between condition and gender of speaker, and between condition and gender of observer. Tables 3 and 4 give averages and difference scores for negative and positive conditions, as a function of the gender of the speaker and the ob-
Table 4: Observed scores, standard deviations and difference scores from male and female observers in positive and negative conditions (1=very positive to 5=very negative)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Observer</th>
<th>Positive (1)</th>
<th>Negative (2)</th>
<th>Diff. (2-1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>2.679 (.110)</td>
<td>3.639 (.106)</td>
<td>.960</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>2.566 (.110)</td>
<td>3.791 (.106)</td>
<td>1.225</td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Observed scores, standard deviations and difference scores from negative and positive conditions in interview and movie fragment (1=very positive to 5=very negative)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Fragment</th>
<th>Positive (1)</th>
<th>Negative (2)</th>
<th>Diff. (2-1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Film</td>
<td>3.372 (.090)</td>
<td>4.058 (.090)</td>
<td>.686</td>
<td></td>
</tr>
<tr>
<td>Interview</td>
<td>2.368 (.073)</td>
<td>2.877 (.104)</td>
<td>.509</td>
<td></td>
</tr>
</tbody>
</table>

5. Acknowledgments

This research was conducted in the context of the FOAP project, which is funded by the Netherlands Organisation of Scientific Research (NWO) (see http://foap.uvt.nl). We thank Nadja van Bodegom, Manon de Caluwé, Stefanie Fassaert, Marthy Herckenrath, Eva van de Sande, Annelot Smulders and Sara Theunisse for help with the data collection and the perceptual evaluations, and Lennard van de Laar for technical assistance.

6. References