interspeech 2008 incorporating SST 08

Nonverbal Responses to Social Inclusion and Exclusion

Emiel Krahmer[1], *Juliëtte Schaafsma*[1], *Marc Swerts*[1], *Ad Vingerhoets*[2]

[1]Faculty of Humanities, Tilburg University, The Netherlands
[2]Faculty of Social and Behavioral Sciences, Tilburg University, The Netherlands
{E.J.Krahmer, J.Schaafsma, M.G.J.Swerts, A.J.J.M.Vingerhoets}@uvt.nl

Abstract

Applying an elaborate and novel experimental paradigm we collected non-verbal responses to social inclusion and exclusion. A judgement experiment revealed that it was possible to determine whether a person was included or excluded, based solely on non-verbal, behavioral cues, especially if the person was male. A detailed coding using the ECSI scale suggests that included speakers displayed more cues associated with Affiliation and Relaxation, whereas excluded speakers gave more evidence of Flight and Displacement.

Index Terms: paralinguistics, emotion, social rejection, ostracism

1. Introduction

Human beings are deeply motivated to form stable, lasting connections with other people. They strongly desire social attachments and seem inclined to form relationships even in the absence of ulterior motives. Moreover, they are willing to spend considerable time and effort in fostering supportive relationships with others and are generally reluctant to end relationships, even when these relationships have become unnecessary or dysfunctional. This tendency to strive for strong social attachments presumably has an evolutionary basis as it may have promoted human survival and reproduction (Baumeister & Leary, 1995; Leary, 2001).

When people's belonging needs are threatened, they respond in a variety of negative ways. For example, laboratory studies show that being excluded or rejected, even if it is for only a short period of time, is a painful experience (e.g., Eisenberger, Lieberman, & Williams, 2003) that increases self-defeating behaviors (e.g., Twenge, Catanese, & Baumeister, 2002), lowers peoples self-esteem (e.g., Leary, Tambor, Terdal, & Downs, 1995), and leads to aggression toward others (e.g., Twenge, Baumeiser, Tice, & Stucke, 2001; Twenge & Campbell, 2003). At the same time, however, researchers have observed that socially excluded people seem to respond nonverbally in a seemingly detached and emotionally indifferent matter. For example, in experiments in which people were excluded during a ball-tossing game, people have been found to be lethargic and to slowly disengage by slumping down in their chairs and looking downward (Williams, 2001). It has been suggested that people may in this way try to protect themselves or minimize the emotional distress caused by the threats to their inclusionary status (e.g., DeWall & Baumeister, 2006). So far, however, little research has focussed on non-verbal behaviors of excluded persons and on systematically comparing their nonverbal expressions to those of included persons.

The purpose of the current investigation was to explore whether excluded persons display different non-verbal behaviors than included persons. It was hypothesized that judges could determine, only on the basis of non-verbal cues, whether a person is excluded or included. Because excluded people may try to protect themselves from the emotional distress that may arise, it was also expected that excluded people would show non-verbal behaviors that reflect a tendency toward social withdrawal (e.g., a restriction of non-verbal expressiveness), whereas included persons would show higher levels of affiliative non-verbal behaviors. To this end, we first collected data of people who were either included or excluded using a novel experimental paradigm (Section 2). Next we conducted a perception experiment, in which a group of participants had to judge on the basis of video fragments whether a person was included or excluded (Section 3). Finally, the non-verbal behaviors of included and excluded speakers were coded and analyzed using the ECSI non-verbal behavior scale, see Troisi 2002. Troisi & Moles 1999 (Section 4).

2. Audiovisual data collection

2.1. Method

Participants were 59 native speakers of Dutch (37 women), with an average age of 21 years. All were students from Tilburg University, who participated for partial course credit. All had provided written consent for their data to be used for research purposes. Students who had a history of depression or heart complaints were excluded.

The data-collection experiment had a between-subjects design, and participants were randomly assigned to either the Inclusion ($n = 30$) or the Exclusion condition ($n = 29$), where care was taken that sex was balanced for both conditions. The experiment was presented as a study on group decision making under time pressure/stress, and participants were led to believe that they would be engaging in a decision-making discussion with two other participants. In reality, they would communicate with a pair of actors (one male, one female) operating on an elaborate script.

At the start of the experiment, participants were led into a room and told that the other two "participants" were in separate rooms as well. After the global procedure was explained, participants signed a consent form, and six electrodes were applied to measure heart rate. Following the APA guidelines for ethics, participants were informed that they could stop their participation at any moment, without having to give a reason. None of our participants used this right.

Following this introduction, participants were asked to fill in a self-report mood questionnaire ("At this moment, I feel ...") derived from Mackie & Worth (1989) and Krahmer et al. (2004), consisting of six 7-point bipolar semantic differential scales, using the following adjective pairs (English translations of Dutch originals: happy/sad, pleasant/unpleasant, sa-

Accepted after peer review of full paper

Figure 1: *Representative stills of speakers in the Inclusion (top) and Exclusion (bottom) conditions.*

tised/unsatised, content/discontent, cheerful/sullen and in high spirits/low-spirited. The order of the adjectives was randomized; for the analysis negative adjectives were assigned a score of 1 and positive ones a score of 7. After filling in this questionnaire, participants were exposed to a 7-minute film fragment, consisting of underwater scenes filmed in the Red Sea and accompanied with relaxing music to make sure that participants in both conditions were in a comparable state of mind at the start of the discussion. To check whether this was indeed the case, participants were asked to fill in the same mood questionnaire a second time. In addition, saliva samples were collected, to measure participants cortisol levels (cortisol is a stress hormone). Subsequently, participants were accompanied to the discussion room, where they met the other two "participants" (the actors). All three were seated at a hexagonal table, so that each person had one conversation partner on the left-hand and one on the right-hand side, and each had a digital DV camera (25 fps) in front. Both the participant and the actors were recorded, and participants were told that these recordings would be needed to analyze the decision making process afterwards.

At this point, participants received the case to be discussed, in a short text containing the description of a communication problem in a local sport school. Participants were instructed to collectively answer two questions (How did the problems arise? And how could they be solved?), and they were given 4 minutes exactly to answer each one. After this, the three experimenters (one for each conversation participant) left the room. The actual experimental manipulation occurred during the discussion of the second question. In the Inclusion condition, the actors continuously focussed on the contributions of the participant and emphasized how much they appreciated these ("yes, that's an excellent suggestion!"); in the Exclusion condition, the actors discussed the case solely among themselves, ignoring any contributions from the participant.

After 2 x 4 minutes, the experimenters re-entered the discussion room, and each guided one conversation partner (the participant or one of the two confederates) back to one of the individual rooms. Once there, the confederates waited for the next experiment, while the participant was asked to fill in the mood questionnaire once more, to assess the effect of the experimental manipulation on the mood of the participant. After this, participants were shown a second, 7-minute Red Sea underwater scene with soothing music, in an attempt to bring the participants' mood back to more neutral levels (this time interval was also necessary for the cortisol measurements, since it takes some time before an increase in cortisol levels is actually noticeable in saliva samples). Finally, participants filled in the mood questionnaire one last time.

Table 1: *Average mood scores (with standard deviations between brackets) for the four measurements, as a function of social condition (higher values indicate more positive feelings).*

Mood	Exclusion	Inclusion
Mood 1: Initial	5.09 (.81)	5.21 (.76)
Mood 2: After film 1	5.41 (.73)	5.41 (.61)
Mood 3: After manipulation	4.92 (.83)	5.69 (.64)
Mood 4: After film 2	5.37 (.78)	5.56 (.68)

Subsequently, the participants were fully debriefed about the experiment. None of them was suspicious about the experimental set-up; in particular, all believed that they had been interacting with other, "real" participants. Finally, participants signed a non-disclosure agreement, to make sure that future participants were uninformed about the actual nature of the experiment. Overall, the experiment lasted about one hour.

2.2. Results

To find out whether the experimental manipulation worked, we analyzed the self-reported mood scores. Table 1 contains the average scores for the four mood measurements. In all four cases, the internal consistency (Cronbach's α) was high (with α-values between .8 and .9). The average mood scores were submitted to a 2 (Soclal condition: inclusion, exclusion) by 4 (Time of mood measurement: 1, 2, 3, 4) Analysis of Variance (ANOVA). Most relevant for our current purposes is that a significant interaction was found between Condition and Time, $F(1, 57) = 7.69, p < .001$. In particular, as can be seen in Table 1, average mood scores for the two conditions are exactly the same after the first film fragment (as intended), but after the experimental manipulation a clear difference between the conditions can be observed, where the mood increases for included participants and decreases for excluded participants. After watching the second film fragment this effect has disappeared again. Interestingly, there was also a main effect of Time of the mood measurement, and inspection of Table 1 reveals that participants actually feel better *after* than before the experiment (irrespective of the condition they were in), $F(1, 57) = 5.01, p < .01$. There were no further significant effects on the mood scores.

2.3. Conclusion

The results of the mood questionnaires indicate that the elaborate experimental set-up worked exactly as intended. Figure 1 shows some representative stills from participants in the respective conditions. In general, it was felt that clear differences could be observed between participants in the respective conditions. To quantify these differences, we performed a perception experiment and a detailed coding of non-verbal cues, which we describe in the next two sections.

Table 2: *Average perception scores (ranging from -5 "very certainly excluded" to +5 "very certainly included"; with standard deviation and 95% confidence intervals) as function of Social condition and Gender.*

Condition	Gender	Mean	*SD*	95% CI
Inclusion	Male	1.56	.15	(1.24, 1.88)
	Female	1.63	.17	(1.28, 1.98)
Exclusion	Male	-1.68	.20	(-2.10, -1.26)
	Female	-1.22	.19	(-1.61, -0.83)

Table 3: *Average perception scores (ranging from -5 "very certainly excluded" to +5 "very certainly included"; with standard deviation and 95% confidence intervals) as function of Gender and Fragment.*

Gender	Fragment	Mean	*SD*	95% CI
Male	First	-.32	.12	(-.56, -.01)
	Second	.20	.20	(-.21, .62)
Female	First	.53	.19	(.14, .91)
	Second	.12	.18	(-.50, .26)

3. Perception experiment

3.1. Method

For the perception experiment 25 adult native speakers of Dutch participated as judges (8 women), all different from the participants in the data-collection experiment. The stimuli for the perception experiment were prepared as follows: for each of the 59 participants of the data collection phase, two fragments of 8 seconds (200 frames) were selected. One fragment was selected from the beginning of the four minute experimental manipulation (frames 1000 - 1200, i.e., 0.40 - 0.48 minutes), and one from the second half (frames 4000 - 4200, 2.40 - 2.48 minutes). This resulted in 59 x 2 = 118 stimuli. These were presented to individual judges in one of two random orders, to control for potential learning effects. Judges had to indicate by forced choice for each fragment whether they believed the person in the film-clip was included or excluded, and on a five point scale how certain they were of their choice. For data processing perceived inclusion was mapped to "1" and perceived exclusion to "-1", and these scores were multiplied with the certainty score. This resulted in a score ranging from -5 ("very certainly excluded") to +5 ("very certainly included"). The experiment was preceded by a short training session of five stimuli (consisting of random 8 second fragments not used in the actual experiment), to make participants acquainted with the experimental setting. If all was clear, the actual experiment started and there was no further interaction between participant and experimenter. The entire experiment lasted slightly over 25 minuted.

3.2. Results

The data were subjected to a 2 (Social condition: inclusion, exclusion) $\times$ 2 (Fragment: first, second) $\times$ 2 (Speaker gender: female, male) Analysis of Variance (ANOVA). Included participants were indeed perceived as more included ($M = 1.60, SD = .15$) and participants in the excluded condition were overall perceived as more excluded ($M = -1.45, SD = .18$), $F(1, 25) = 367.80, p < .001$. In addition, female speakers were perceived as socially included more often than male speakers (men: $M = -.01, SD = .14$ vs women: $M = .20, SD = .16$), $F(1, 25) = 7.51, p < .05$. Moreover, an interaction between Social condition and Speaker gender was found ($F(1, 21) = 6.92, p < .05$). This interaction can be explained by looking at Table 2; while female speakers were overall perceived as more "included", differences with male speakers are especially large in the Exclusion condition. Also, a significant interaction was found between Fragment and Speaker gender ($F(1, 21) = 26.53, p < .001$). Table 3 shows that, on average, male speakers were perceived as more included in the second fragment, while female speakers appeared more included in the first fragment. No other main or interaction effects reached the significance threshold. In particular, speakers appeared on average equally included or excluded in both the first (early) and the second (late) fragment, $F < 1$.

3.3. Conclusion

The perception experiment revealed that judges can determine, based on the non-verbal cues displayed by speakers, whether or not they are included or excluded from an ongoing conversation, where exclusion is noticeably easier to perceive when the excluded person is male. To get a better insight into which non-verbal signals included and excluded speakers produce (and which cues participants of the perception experiment may have relied on for their classification), we conducted a detailed coding of a selection of included and excluded speakers, which we describe next.

4. Non-verbal behavior coding

4.1. Method

Based on the results of the perception experiment, we selected the 10 speakers for which participants on average were most certain that they were included, and the 10 speakers for which they were most sure they were excluded. For these twenty selected speakers, two extended 30-second fragments were analyzed (0.30 - 1.00 and 2.30 - 3.00). Notice that these fragments properly include the time intervals used for stimuli of the perception test. These selections were coded using the Ethological Coding System for Interviews (ECSI); see e.g., Troisi (2002) or Troisi & Moles (1999). The ECSI is a validated non-verbal behavior scale, consisting of 8 behavioral categories and a total of 37 easy to code nonverbal cues. We selected four behavioral categories for coding, namely "Affiliation" (which is associated with ECSI behaviors 2-6, e.g., smile, head tilt, eyebrow flash), "Flight" (behaviors 10-15, e.g., look away/down, chin to chest), "Displacement" (24-32, e.g., hand-face touching, yawning), and "Relaxation" (33-37, e.g., settle, fold arms, laugh).

Coding was done blind to condition, and without sound (as required by the ECSI guidelines). The coding was generally easy, and difficult cases were resolved after discussion. For data analysis, behaviors associated with each of the four categories were summed per participant and fragment, and submitted to a 2 (Social condition: inclusion, exclusion) $\times$ 2 (Fragment: first, second) Analysis of Variance (ANOVA).

4.2. Results

Table 4 summarizes the results. First, included participants displayed on average more non-verbal behaviors associated with Affiliation ($M = 2.05$) than those who were excluded ($M = 0.85$), $F(1, 36) = 15.52, p < .001$. And in ad-

Table 4: *Average number of non-verbal behaviors (with standard deviation and 95% confidence intervals) in each of the four ECSI categories, as a function of Social Condition.*

Scale	Condition	Mean	*SD*	95% CI
Affiliation	Inclusion	2.05	.21	(1.61, 2.49)
	Exclusion	0.85	.21	(0.41, 1.29)
Flight	Inclusion	1.00	.10	(0.81, 1.19)
	Exclusion	1.40	.10	(1.21, 1.59)
Displacement	Inclusion	1.20	.20	(0.79, 1.61)
	Exclusion	1.90	.20	(1.49, 2.31)
Relaxation	Inclusion	0.70	.16	(0.38, 1.03)
	Exclusion	0.15	.16	(-0.18, 0.48)

dition, included participants showed more signs of relaxation ($M = 0.7$) than did excluded participants ($M = 0.15$), $F(1, 36) = 5.89, p < .05$. In contrast, the excluded participants in this experiment displayed on average more of the non-verbal behaviors associated with Flight ($M = 1.40$) than the included ones ($M = 1.00$), $F(1, 36) = 8.28, p < .001$. The same was true, finally, for Displacement, where excluded participants displayed on average more non-verbal cues in this category ($M = 1.9$) than included participants ($M = 1.2$), $F(1, 36) = 6.25, p < .05$. No significant differences were found between the first and the second fragment.

4.3. Conclusion

The ECSI coding indeed revealed non-verbal differences between included and excluded participants. The former displayed more behaviors associated with Affiliation and Relaxation, while the latter displayed more Flight and Displacement cues. Notice that this pattern can already be seen in Figure 1, where the top stills illustrate head tilt, smile (left) and laughter (right), associated with Affiliation and Relaxation respectively, while the bottom stills illustrate chin to chest (left) and face-touching (right), associated with Flight and Displacement respectively.

5. General conclusion and discussion

Social exclusion is an important problem, and a better understanding of what causes it and of how people respond to it is badly needed. In this paper, we have described a novel way to collect data on social inclusion and exclusion, by letting participants communicate with two other people (who —unbeknownst to the participant— happen to be actors playing an extended script), who either include or exclude the participant during 4 minutes. Self-reported mood questionnaires indicate that the mood of participants drops significantly after exclusion and raises after inclusion. A perception test revealed that judges can reliably distinguish between included and excluded conversation partners, on the basis of their non-verbal behavior. Exclusion was somewhat easier to perceive for men than for women. A detailed coding of a subset of included and excluded participants according to the ECSI scale, revealed that included participants display more behaviors associated with Affiliation and Relaxation, while excluded participants showed more behaviors from the Flight and Displacement categories.

During the data-collection, we not only recorded non-verbal behaviors, but also measured heart rate and cortisol levels. Detailed analyses of these physiological variables failed to yield significant effects of social condition, due in part to large inter-individual differences. We will report on these results in a sequel to this paper. In general, the data-collection method proved to be an efficient way to elicit spontaneous emotional verbal and non-verbal speech data, and offers an interesting alternative to more common emotion induction techniques (cf., Scherer 2003). Various alternatives are worth exploring, for instance varying the lexical content discussed by the actors during the manipulation, which might either strengthen or weaken the effect of the manipulation. This is also something we hope to explore in further work.

6. Acknowledgements

We would like to thank Lennard van de Laar for technical support, and Rian Blankenstein, Lotte Oostrom, Bregje van Rijbroek, Marjolein de Vries for their help with conducting the various experiments.

7. References

[1] Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*, 497-529.

[2] DeWall, C. N., & Baumeister, R. F. (2006). Alone but feeling no pain: Effects of social exclusion on physical pain tolerance and pain threshold, affective forecasting, and interpersonal empathy. *Journal of Personality and Social Psychology, 91,* 1-15.

[3] Eisenberger, N. I., Lieberman, M. D., & Williams, K. D. (2003). Does rejection hurt? An fMRI study of social exclusion. *Science, 302*, 290-292.

[4] Krahmer, E., J. van Dorst and N. Ummelen (2004). Mood, persuasion and information presentation, *Information Design Journal, 12*, 40–52.

[5] Leary, M. R. (2001). *Toward a conceptualization of interpersonal rejection*. In M. R. Leary (Ed.), Interpersonal rejection (pp. 3-20). Oxford: Oxford University Press.

[6] Leary, M. R., Tambor, E. S., Terdal, S. K., & Downs, D. L. (1995). Self-esteem as an interpersonal monitor: The sociometer hypothesis. *Journal of Personality and Social Psychology, 68*, 518-530.

[7] Mackie, D. & L. Worth (1989). Processing decits and the mediation of positive affect in persuasion. *Journal of Personality and Social Psychology, 57,* 2740.

[8] Scherer, K. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication, 40*, 227256.

[9] Troisi, A. (2002), "Displacement activities as a behavioral measure of stress in nonhuman primates and human subjects", *Sress, 5*, 47–54.

[10] Troisi, A. & A. Moles (1999), "Gender differences in depressions: An ethological study of nonverbal behavior during interviews". *Journal of Psychiatric Research, 33*, 243–250.

[11] Twenge, J. M., Baumeister, R. F., Tice, D. M., & Stucke, T. S. (2001). If you cant join them, beat them: Effects of social exclusion on aggressive behavior. *Journal of Personality and Social Psychology, 81*, 1058-1069.

[12] Twenge, J. M., & Campbell, W. K. (2003). Isnt it fun to get the respect that were going to deserve? Narcissism, social rejection, and aggression. *Personality and Social Psychology Bulletin, 29*, 261-272.

[13] Twenge, J. M., Catanese, K. R., & Baumeister, R. F. (2002). Social exclusion causes self-defeating behavior. *Journal of Personality and Social Psychology, 83*, 606-615.

[14] Twenge, J. M., Catanese, K. R., & Baumeister, R. F. (2003). Social exclusion and the deconstructed state: Time perception, meaninglessness, lethargy, lack of emotion, and self-awareness. *Journal of Personality and Social Psychology, 85*, 409-423.

[15] Williams, K. D. (2001). *Ostracism. The power of silence*. New York: The Guilford Press.