LUNDS UNIVERSITET
INSTITUTIONEN FÖR PSYKOLOGI

Electrophysiology of stereotypes: N400 as a measure of the beautiful is good stereotype

ANDERS RASMUSSEN

Kandidatuppsats ht 2006

Handledare: Mikael Johansson

Abstract

Our tendency to associate attractive people with positive traits and unattractive people with negative traits is well documented. Stereotypes have traditionally been measured using either explicit measures such as questionnaires, or implicit measures such as the Implicit Association Test. In the present experiment it was tested whether the beautiful is good stereotype could also be measured using the N400 event-related potential (ERP) component which is sensitive to semantic and pragmatic rule violations. While their EEG was measured participants were presented with faces rated as attractive or unattractive, followed by words rated as positive or negative. As predicted, the results showed that the N400 was larger for unattractive-positive compared to unattractive-negative face-word pairs, thus confirming that the N400 is sensitive to stereotype violations. However, there was no difference between attractive-negative and attractive-positive pairs. A main effect of attractiveness on EEG was also found.

Key words: EEG, N400, Stereotypes, Beautiful is good, Attractiveness

Electrophysiology of stereotypes: N400 as a measure of the beautiful is good stereotype

Beautiful is good

Being beautiful is associated with many advantages as well as some disadvantages. If you are a criminal offender you will receive a lower sentence if you are also attractive, unless you were using your attractiveness to commit the crime (Sigall & Ostrove, 1997). If you choose an honest living instead, you will still be better off in terms of income (Frieze, Olson & Russel, 1991). If you are an attractive professor, you are likely to be seen as a better lecturer compared to your unattractive counterpart (Riniolo, Johnson, Sherman & Misso 2006). As an attractive individual you will also be treated better by other people, receive more help and co-operation, and have a more favorable self-perception (Langlois, Kalakanis, Rubenstein, Larson, Hallam, & Smoot, 2000). Even two month old infants will prefer you to your not so attractive counterpart. When shown two faces simultaneously, one rated by adults as attractive and the other as unattractive, infants spend more time looking at the attractive face (Langlois & Roggman, 1987).

At the core of all these advantages that attractive people so unjustly are enjoying lays a tendency in people to associate attractiveness with positive traits and unattractiveness with negative traits. In the first study to investigate such attribution differences participants were asked to pair personality traits and faces. The hypothesis which was confirmed in the study was that attractive individuals would be associated with more favorable personality traits (Dion, Berscheid & Walster, 1972). The authors coined the name “What is beautiful is good” to describe this stereotype, however, subsequent studies have shown that this name may not be so accurate. In one recent study neutral faces were used as a baseline to which attractive and unattractive faces could be compared, thus allowing an absolute measure of how much we associate negative and positive traits with attractive and unattractive faces respectively. The authors concluded that it would perhaps be more accurate to refer to the stereotype as “what is ugly is bad”, reflecting the fact their participants had a stronger association between unattractive faces and negative traits than between attractive faces and positive traits (Griffin & Langlois, 2006). Similarly, another study showed that the beautiful is good stereotype works both ways. If we are given a positive description of a person we will automatically see that person as more attractive, suggesting that the stereotype could also be called “what is good is beautiful” (Gross & Crofton, 1977; Little, Burt, & Perrett, 2006). In any case, our tendency to associate attractive people with positive traits and unattractive people with negative traits is by now well established from a large number of empirical studies, most of which have been summarized in two meta-analysis studies by Eagly, Ashmore, Makhijani & Longo (1991) and Langlois et al. (2000).

So why would people hold this stereotype? What purpose does it serve? How did it get started? One frequent argument is that our assessments of what is considered beautiful is largely determined by media images, but this explanation fails to explain the cross-cultural agreement on evaluations of facial attractiveness (Cunningham, Roberts, Barbee & Druen, 1995), and it also fails to explain why infants, too young to be influenced by our medias, also prefer attractive faces rather than unattractive faces (Langlois & Roggman, 1987). An evolutionary approach, on the other hand, could explain these findings. In our evolutionary history, those individuals, who have been attracted to other individuals of good phenotypic quality i.e. those with higher fertility and better health for instance, would have had, on average, higher fitness than those who were biased towards partners with bad phenotypic quality. Facial characteristics that people tend to find attractive such as symmetry and averageness have been linked to better physical health and thus higher phenotypic quality (Henderson & Anglin, 2003; Kanazawa & Kovar, 2004; Thornhill & Gangestad, 1999). For this reason, natural selection is likely to have favored those with a preference for symmetry and averageness and thus also facial attractiveness. When Shackelford and Larsen (1999) tested the link between physical attractiveness and health directly by correlating the physical attractiveness with different measures of physical health, they found that those rated as more attractive were indeed healthier. In sum these studies suggests that we have an evolved preference for attractive faces because they signal good phenotypic quality. For a good review and meta-analysis on these issues see Rhodes (2006).

Stereotypes

Stereotypes, in general, can be defined as beliefs, knowledge or expectations tied to a particular social group (Hamilton & Sherman, 1994). It is the prototypical mental picture that appears in the mind when thinking about a doctor, a cleaner, or a beautiful person (Banaji, 2001). These pictures may be either accurate (e.g. most professors are men), or inaccurate. Stereotypes are distinguished from the intuitively proximate concepts of attitudes and prejudice in that the two latter have an emotional component whereas stereotypes is a cognitive structure (Banaji, 2001). As unfair and untrue as stereotypes may be, they are probably very important for us to be able to function properly. We use stereotypes to categorize people and when we meet a member of a category we use our stereotype to make hypothesizes about the individual (Banaji, 2001). Stereotypes, including the beautiful is good stereotype can thus be viewed as a knowledge structure (Hamilton & Sherman, 1994). When we see a person belonging to a specific category other congruent concepts that we associate with the category are primed. This is the logic behind Greenwald's Implicit Association Test (IAT; see below) and it is the logic in this experiment as well.

The simplest way to measure stereotypes is to ask people, yet doing so is associated with problems such as social desirability effects and experimenter expectancy effects (Shaughnessy, Zechmeister & Zechmeister, 2005). Social desirability effects refer to when a participant answers questions with a conscious or unconscious motive to seem like a good person. Since stereotypes can be very politically incorrect, social desirability effects might become particularly pronounced. For example, when asked "do you think this unattractive person is stupid" or "do you think this black person is violent", not many people would openly admit that they think so. Experimenter expectancy effects occur when the participant makes hypothesizes about what kind of results the experimenter is seeking. The expectancies might influence the results in any direction, perhaps the participant tries to "help" the experimenter by acting in accordance with whatever he or she thinks is the hypothesis being tested, or maybe the participant wants to disprove the hypothesis. In any case it is a problem for the experimenter.

Because of these limitations implicit measures, which do not to the same extent rely on the honesty of the participants, have an important role, especially when it comes to measuring stereotypes which people generally do not like to reveal. One of the most famous implicit measures of stereotypes is the implicit association test (IAT). This test employs reaction time measures to indirectly measure the strength of associations among different concepts (Greenwald, McGhee & Schwartz, 1998). Participants are asked to categorize different stimuli that appear on a screen by pressing one of two category buttons. When two associated concepts (e.g. "beautiful" and "good") shares the same response key, reaction times are short, however, when two unassociated concepts (e.g. “ugly” and “good”) share response key, reaction times are long (Greenwald et al., 1998). Though this test is not as vulnerable to social desirability effects and experimenter expectancy effects there are still problems associated with it. For example, the IAT is dependent on the reaction time of the participants, which varies between different people (Nosek, Greenwald & Banaji, in press). Another related problem with the IAT is the effect of age. Older participants have longer reaction times and therefore whatever stereotypes they have will become more pronounced. A perhaps more serious problem is that many IAT experiments have not properly controlled for salience. Coming back to the theoretical example just mentioned, perhaps ugly faces and negative words are more conspicuous or salient which in turn gives rise to faster reaction times when these concepts share the same response key. In one experiment where the effect of salience was subtracted from the effect association the effect vanished completely, thus suggesting that it was merely the salience of the words that mattered, not how much they are associated (Rothermund & Wentura, 2004). Implicit measures in general also have a lack of construct validity. A person who gets a high score on one measure may get a low score on another measure even if both claim to measure the same thing.

By using a more direct, physiological measure it is perhaps possible to circumvent these issues. Whether it will ever be possible to explain social cognitive phenomenon in terms of neurophysiological processes is at best uncertain, however, EEG studies such as this one could well be the first step in this direction.

Electroencephalography (EEG) and N400

Electroencephalography or EEG is a noninvasive measure of excitatory post-synaptic potentials generated by a population of neurons. Measures are made with small electrodes attached to the scalp of the participant (Luck, 2005). Even though it is a crude measure in the sense that it cannot detect activity of individual neurons, EEG have impressive temporal resolution, and it can reveal different states, such as sleep or seizures, as well as the occurrence of different events such as the perception of something unexpected. This later category of waves is referred to as Event-Related Potentials or ERPs and they occur in response to many different cognitive, sensory, and motor events (Luck, 2005; Rugg & Coles, 1998).

The N400 ERP component which, as the name implies, is a negative deflection taking place approximately 400 milliseconds post-stimulus, is a measure of the semantic relatedness between a stimulus and the context in which it was presented. In the original paradigm in which the N400 was measured participants were presented with sentences, some of which ended in unexpected way e.g. “he took a sip from the transmitter” instead of “he took a sip from the glass”. The N400 component was elicited following the presentation of incongruous sentences but not after the presentation of congruous, normal sentences (Kutas & Hillyard, 1980). In the authors own words “The N400 wave may be an electrophysiological sign of the reprocessing of semantically anomalous information” (p. 203). This view has since been supported by evidence from numerous studies (see Kutas & Federmeier, 2000; Luck, 2005; Rugg & Coles, 1995, for reviews). The N400 has already found some practical applications. Among other things the N400 component has shown potential as a clinical tool for diagnosing schizophrenia (Kumar & Debruille, 2004), and as a lie detector tool (Boaz, Perry, Raney, Fischler & Shuman, 1991).

Although the N400 component has traditionally been demonstrated in experiments utilizing written stimuli the N400 component reflects deeper forms of processing and therefore it is modality independent. For example, if participants are sequentially presented with two pictures depicting famous individuals, and then asked to judge whether they belong to the same or to different occupational categories, the N400 will appear if the faces belong to different occupational categories i.e. if they are incongruent, but not if they belong to the same category (Barrett & Rugg, 1989). Similarly, the N400 has been demonstrated in a study using auditory instead of written material (Faustman, Murdoch, Finnigan & Copland, 2005).

Recently, Hagoort, Hald, Bastiaansen & Petersson (2004) demonstrated that the N400 component is sensitive not only to violations of language rules as in Kutas and Hillyard’s original study, but also to violations of our knowledge about the world. The authors were not surprised to find that the N400 was elicited following presentation of the sentence “Dutch trains are *sour* and very crowded”. This just confirms what has been found in previous studies. However, the ERP component was also elicited following the sentence “Dutch trains are *white* and very crowded”. This latter sentence is not a semantic violation i.e. it follows the rules of language, but it is a violation of knowledge since in the Netherlands it is well known that trains are yellow, not white. As mentioned before, stereotypes can also be seen as knowledge structures (Hamilton & Sherman, 1994). Hence if stereotypes are knowledge structures and if the N400 component is sensitive to violations of knowledge structures, it follows that the N400 component should be sensitive to violations of stereotypes.

Because the N400 component is automatic and therefore independent of whether the participant makes a response, it has some potential advantages as a measure of stereotypes. For example, the N400 component is unlikely to be affected by social desirability effects and experimenter effects. Likewise, reaction times which, as mentioned above, can influence the results in an implicit association test should not affect the outcome in the present experiment.

Thus in the present experiment it is predicted that the N400 component will be larger when serially presented face-word pairs violates the “beautiful is good” stereotype, compared to when the pair is in agreement with the content of the stereotype. More specifically we predict that the N400 component will be larger following the presentation of an attractive face followed by a negative word or an unattractive face followed by a positive word (stereotype incongruent pairs), than when an unattractive face is followed by a negative word or when an attractive face is followed by a positive word (stereotype congruent pairs). It will also be assessed whether this N400 is larger for either attractive or unattractive faces or whether the effect is the same.

A side-track in the present experiment was to test whether there where any differences in EEG solely depending on whether an attractive or unattractive face was shown. A few fMRI studies have revealed that several different brain regions are sensitive to the attractiveness of faces (O´Doherty, Winston, Critchley, Perrett, Burt & Dolan, 2003; Winston, O´Doherty, Kilner, Perrett & Dolan, 2007). It is conceivable therefore that EEG activity would also be different for attractive and unattractive faces.

Method

Participants

Subjects were 22 Swedish university students aged between 18-40, with a median of 22. To make recruiting easier an ipod was raffled out to one of the participants. All subjects were right handed as determined by tooth-brushing hand.

Material

Two types of stimuli were needed to conduct the present experiment, words rated for valence, and faces rated for attractiveness. A long list of personality attributes were generated by translating English words used in a previous study on stereotypes (Rosenberg, 1977), into Swedish, and then finding synonyms to these words as well as synonyms to the synonyms. This resulted in a list of 331 words describing personality attributes.

A total of eleven judges were given this list and for each word they were asked whether they thought that the words were positive, negative, or neutral. The word got +1 for every judge that rated the word as positive, -1 when negative, and 0 when the judge thought the word was neutral or when the judge did not know what the word meant. Based on this quantification every word received a valence score between +11 (if everyone thought the word was positive) and -11 (if everyone thought the word was negative). Eighty words which were rated as positive by everyone or nearly everyone, and 80 words rated as negative by everyone or nearly everyone were selected to be used in the experiment. Tests were done to check that the positive and negative words did not differ significantly in length, in absolute valence (how far away the words were from zero valence), or in how frequently they are used in the Swedish language as measured by how many times the word appear in Swedish newspapers for every one million words (Språkbanken, Press 98 corpus; http://spraakbanken.gu.se). The results of these tests are shown in Table 1.

A total of 40 color photographs, 20 male and 20 female, depicting only the face were used in the experiment (see Figure 1). These pictures which were morphs of two or more faces had been rated for attractiveness in a previous study (Braun, Gruendl, Marberger & Scherber, 2001). Based on these previous ratings, the faces were categorized as either attractive or unattractive. T-tests were done to check the validity of these categorizations. All words and faces were used to create face-word pairs that were congruent, agreeing with the stereotype, or incongruent, thus violating the stereotype (see Figure 1). All stimuli were presented on a computer monitor, using the E-prime software (Psychology Software Tools, Inc.; http://www.pstnet.com/products/e%2Dprime/). All words used as stimuli in this experiment can be found in Appendix A and all faces that were used can be found in Appendix B.

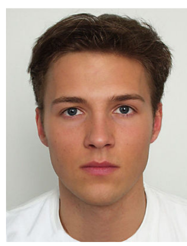
Stimulus type	Face	Word
Unattractive incongruent		Kompetent (Competent)
Attractive congruent		Kompetent (Competent)
Attractive incongruent		Rädd (Afraid)
Unattractive congruent		Rädd (Afraid)
Time »	0 milliseconds »	1000 milliseconds

Figure 1. Examples of the four different types of face-word pairs used in the experiment. Based on whether the face is attractive or unattractive and whether the word is positive or negative the different trials are categorized into one of four categories. From top to bottom the four examples shown here represent a (1) unattractive-incongruent, (2) attractive-congruent, (3) attractive-incongruent, and a (4) unattractive –congruent trial.

While the positive and negative words did differ in valence, $t(158) = 210.8$, $p < .001$, there were no significant differences in length $t(158) = 1.310$, *ns*, absolute valence $t(158) = 1.368$, *ns*, or in frequency of use $t(158) = .445$, ns.

The t-test testing whether the faces in the attractive category were more attractive than the faces in the unattractive category revealed a highly significant difference between the two categories $t(38) = 17.95$, $p < .001$. Means and standard deviations for words as well as faces are shown in Table 1.

Table 1. Means and standard deviations for the words and faces used in the experiment.

Variable	Mean	Standard deviation
Valence, positive words	10.66	.476
Valence, negative words	-10.53	.763
Length, positive words	8.36	2.522
Length, negative words	7.85	2.424
Frequency, positive words	35.79	32.74
Frequency, negative words	32.83	49.77
Attractiveness, attractive faces	5.47	.437
Attractiveness, unattractive faces	2.80	.502

Procedure

The experiment was conducted after the participants had already taken part in another experiment which lasted about an hour. Due to the long duration of the setup required prior to an EEG experiment and the relatively short duration of the two experiments this solution was accepted as the most effective. Even though this probably resulted in more tired subjects it was constant across subjects and experimental conditions. It is also possible that this prior experiment was beneficial to the present experiment in that EEG data tends to "calm down" over time. Before the start of the experiment the participants were given written instructions on the screen. To make sure that they had understood they were asked to explain themselves what they were supposed to do. If something was unclear further instructions were given verbally.

The N400 component as previously mentioned is elicited automatically as long as the participants are paying attention to the stimuli. In this experiment they were given an orienting task simply to make sure that they paid attention to the stimuli. The participants' task was to judge whether personality characteristics fitted a face presented just before. They were asked to focus on a cross in the middle of the screen, and after a brief moment one of the 40 faces appeared on the screen for 1000 milliseconds, immediately after this, one of the 160 words describing a personality characteristic appeared for 1000 milliseconds. It was during this time that the participants had to press one button if they thought that the word fitted the face or another button if they thought the word did not fit the face. After this presentation a cross appeared again for 2500 milliseconds, after which another face and then another word appeared. Participants were instructed to try and time their blinks so that they occurred in this time window. This procedure was repeated 160 times so that each face was presented four times and each word once. Based on whether the face was attractive or unattractive and whether the word was positive or negative, each trial was categorized as "attractive congruent" (ac), "attractive incongruent" (ai), "unattractive congruent" (uc) or "unattractive incongruent" (ui).

ERP recording

The EEG was continuously recorded from 40 silver-silver chloride electrodes mounted in an elastic cap, "Quick cap" (Neuromedical Supplies; www.neuro.com/neuromed/quickcap.htm) and labeled according to the extended 10-20 system (Jasper, 1958). Prior to the experiment impedances were lowered to about 5 kΩ. The EEG from all electrode sites was recorded with reference to the left mastoid electrode, and re-referenced off-line to the average of the left and right mastoids. In order to monitor eye movements and blinks, electrodes were placed above and below the left eye and at the outer canthi. The EEG was amplified with a NeuroScan Nuamp amplifier (Neuroscan; www.neuro.com) with bandpass cutoffs of 0.5–20 Hz (12 dB) and digitized on-line with a sampling rate of 500 Hz. Triggers with information about whether the face was attractive or unattractive and whether the following word was positive or negative, were sent from the stimulus presentation computer to the amplifier when the face appeared on the screen. The epochs used for analyses had a duration of 2200 ms, starting 200 ms prior to the face presentation and ending 1000 ms after the word presentation. The pre-stimulus sampling period was used for baseline correction. Epochs containing artefacts greater than ±50 µV were rejected before averaging. Average ERPs were formed for the following six conditions:

attractive face, unattractive face, attractive-congruent, attractive-incongruent, unattractive-congruent, unattractive-incongruent.

Statistical tests

To test the hypothesis that attractive and unattractive faces would produce different EEG activity patterns a repeated measures analysis of variance (ANOVA) was carried out using the following factors: Face (attractive, unattractive), Anterior/posterior (frontal, central, parietal, occipital), and Laterality (left, mid, right). Mean amplitudes from electrodes F3, FZ, F4, C3, CZ, C4, P3, PZ, P4, O1, OZ, and O2, were extracted from a time window between 250 and 700 milliseconds following the face presentation. The hypothesis that incongruent face-word pairs would be associated with a larger N400 deflection compared to congruent pairs was tested with ANOVAs employing the factors of Congruency (congruent, incongruent), Anterior/posterior and Laterality on data measured 300-600 milliseconds after the presentation of the word. This hypothesis was tested separately for attractive and unattractive faces thus resulting in two different ANOVAs.

Results

As can be seen in Figure 2 the ERP waveforms for attractive and unattractive faces started to deviate after about 250 ms following presentation of the face. The difference, which lasted until about 700ms after the face had been presented, consisted of a more positive-going wave for the attractive faces compared to the unattractive faces. Figure 3 shows that the effect occurred almost all over the scalp though it was more pronounced over central and parietal areas and less pronounced over frontal and temporal areas. The ANOVA revealed a main effect of attractiveness on EEG between 250 and 700 millisecond following presentation of the face, $F(1,21) = 8.790$, $p = 0.007$, with no significant interactions involving electrode site (i.e. Anterior/posterior, Laterality).

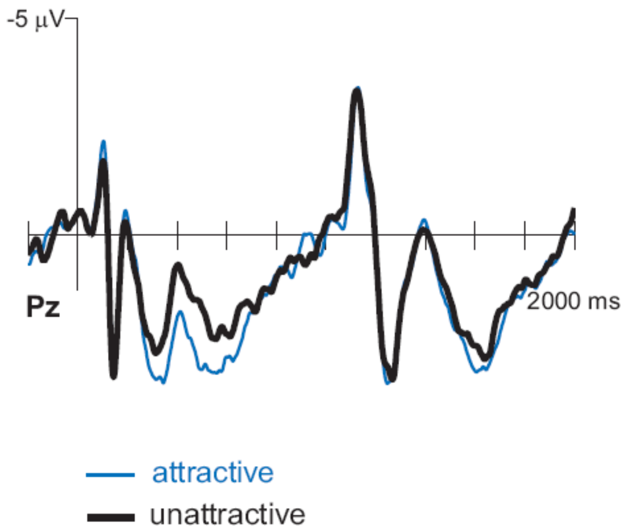


Figure 2. Average ERP waveforms for attractive and unattractive faces.

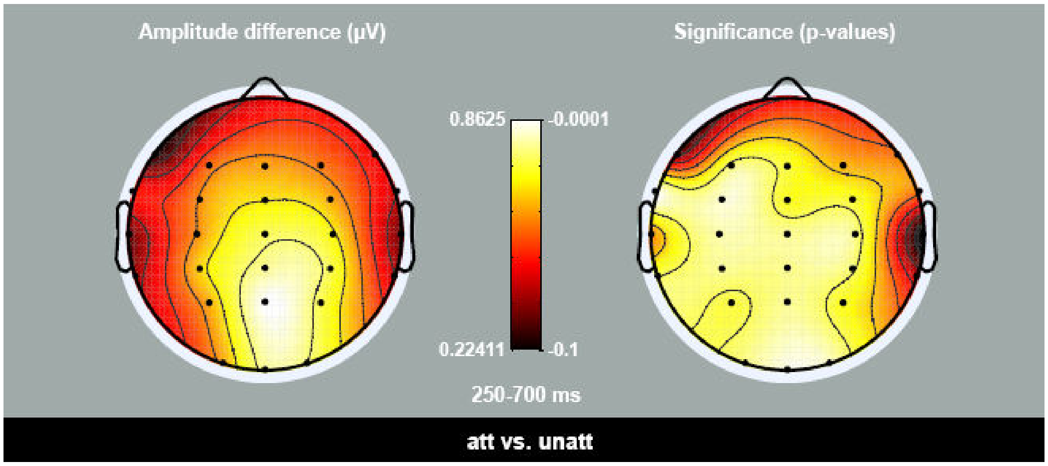


Figure 3. Differences in amplitude and significance levels for presentations of attractive and unattractive faces.

Figure 4 shows the waveforms that appeared following congruent and incongruent trials for attractive and unattractive faces respectively. Visual inspection of the waveforms for the attractive faces suggests that there are no differences as a result of congruence. For the unattractive faces, on the other hand, there is a clear deviation between the two lines in the time window 1300-1600ms, which is 300-600ms after presentation of the word. The difference consisted of a more negative going wave for the incongruous trials.

The ANOVAs testing whether there was a difference in the N400 resulting from a face-word mismatch revealed one significant effect and one non-significant effect. There was a significant difference in voltage in the expected time window between unattractive-congruent and unattractive-incongruent face-word pairs, $F(1,21) = 6.668$, $p = .017$. As can be seen in Figure 5, the effect showed the typical N400 distribution, maximal over central-parietal recording sites. There was no corresponding effect for attractive faces. That is, there was no significant ERP difference following presentations of attractive-congruent and attractive-incongruent face-word pairs, $F(1,21) = .041$, ns. Figure 6 confirms this conclusion.

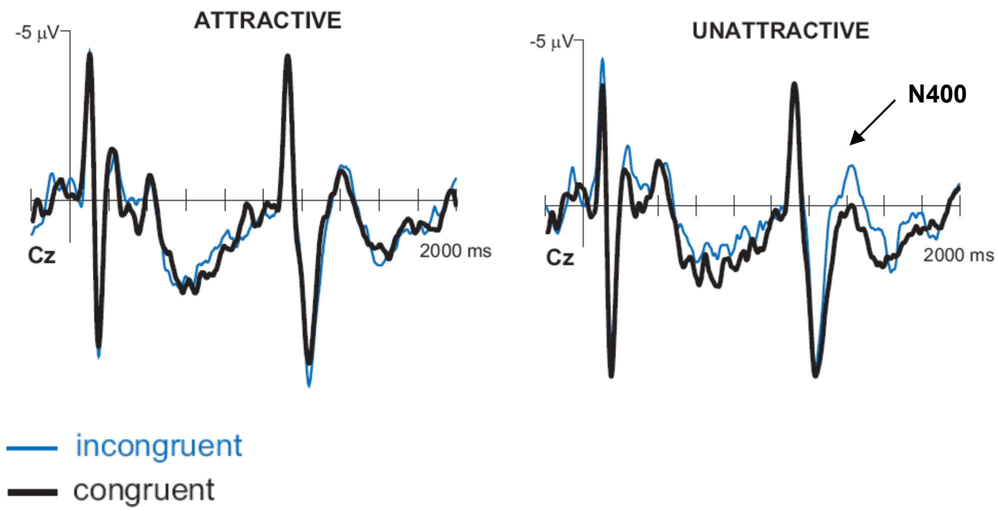


Figure 4. Average ERP waveforms following incongruous and congruous trials. The waveforms are shown separately for attractive and unattractive faces.

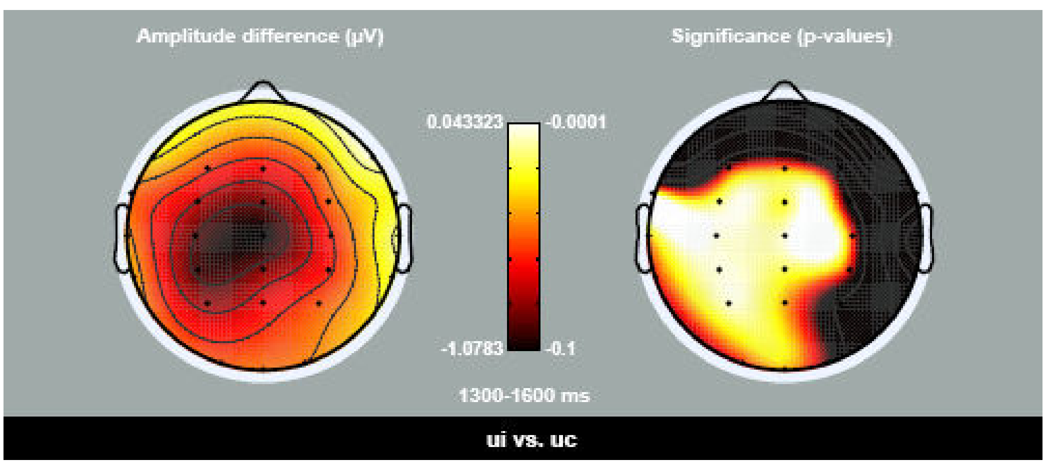


Figure 5. Differences in amplitude and significance levels for unattractive congruent (uc) and unattractive incongruent (ui) trials. There where significant differences between congruent and incongruent trials.

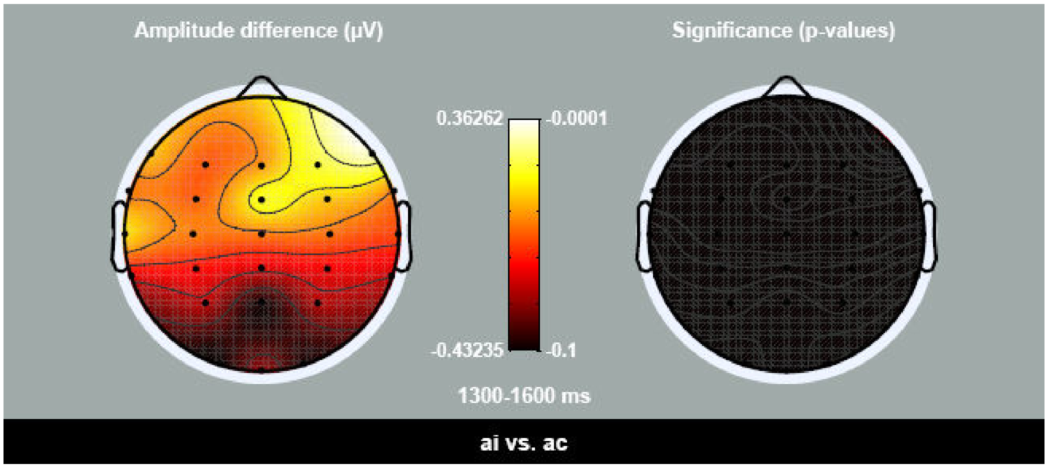


Figure 6. Differences in amplitude and significance levels for attractive-congruent (ac) and attractive-incongruent (ai) trials. There were no significant differences between trials were an attractive face was followed by a positive word and trials were an attractive face was followed by a negative word.

Discussion

As had been predicted the statistical analysis revealed that the N400 component was significantly larger when an unattractive face was followed by a positive word compared to when an unattractive face was followed by a negative word. It appears then that the N400 component is indeed sensitive to stereotype violations. Stimuli contradicting a held stereotype elicit an electrophysiological response similar to the one that is elicited following semantic violations as well as knowledge violations. This supports the view that stereotypes are knowledge structures. Even though stereotypes are generalized pictures of a heterogeneous group of people, the present results suggest that they are still treated by the brain as if they are knowledge, just like knowing the color of trains in your country is knowledge (cf. Hagoort et al., 2004).

However, the same interaction did not take place for the attractive faces. The N400 component was not significantly larger when an attractive face was followed by a negative word, than when an attractive face was followed by a positive word. At first glance it might seem strange that there were significant differences in EEG activity following presentations of unattractive-congruent and unattractive-incongruent pairs, but not after attractive-congruent and attractive-incongruent pairs, after all, one might argue, the stereotype is referred to as the "beautiful is good stereotype". This is indeed true, yet, as mentioned in the introduction it has recently been questioned whether this name is the most accurate. The results in this experiment are consistent with the research showing that we tend to associate negative attributes with unattractive faces more than we associate positive attributes with attractive faces (Griffin & Langlois, 2006). Thus the present experiment lends support to these authors suggestion that it would be more accurate to name the stereotype “ugly is bad”. Simply put the results in this study seem to suggest that our brain finds it quite disturbing to imagine an unattractive person with some positive characteristic such as “social” or “nice”. It is comparably easier to imagine a beautiful person with some negative characteristic such as egocentric or mean.

Because this test is based on an automatic ERP component and therefore does not rely on any overt response from the participants, it has great potential as a measure of stereotypes. Explicit measures, as mentioned, are sensitive to social desirability effects and experimenter expectancy effects. Since the neurophysiology of these biases is not well known it would be premature to say that they do not influence automatic ERP components. However, if social desirability did have an effect on the N400 component one would not expect an “ugly is bad effect” as was found in this experiment. Implicit measures which have been used previously to measure stereotypes likewise have flaws associated with them which an N400 measure may be able to avoid. The implicit association test measure, as mentioned before is reaction time dependent, the N400 is not. It would be interesting to find out if the N400 measure like the IAT responds to salience (see Rothermund & Wentura, 2004), thus bringing in a confound into the equation, or whether perhaps the N400 measure, is immune to such effects. Further experiments would be required to determine this. Just like the IAT, a N400 measure of stereotypes would also prevent biases normally associated with questionnaire measures of stereotypes such as social desirability and experimenter expectancy effects. It would be interesting to see if the results found using this method correlate with the results found using other implicit and explicit measures. For example, since the N400 measure employed in this experiment has many things in common with the IAT one would expect a high correlation between these two measures. Obviously it would also be interesting to know whether this measure can predict openly held stereotypes as well as to what extent this measure correlates with actual behavior. Since it is also an implicit measure it would be odd if results on an IAT measure would be very different from the results obtained using this method. Is there a correlation between peoples´ expressed stereotypes and their stereotypes as measured in the present experiment? Perhaps more importantly, to what extent does this measure correlate with behavior?

There was also a significant main effect of attractiveness. This suggests that when shown a face the brain activation patterns differ depending on whether the face is attractive or not. This is consistent with conclusions from previous studies in which brain activity following attractive or unattractive faces was measured using fMRI (O´Doherty et al., 2003; Winston et al., 2007). In these studies it was shown that faces rated as attractive versus unattractive produce different patterns of blood flow in the brain, which is a measure of the amount of energy used by cells in that region. Studies such as these as well as the present one could perhaps begin to explain how information about attractiveness is processed by the brain. What fMRI studies cannot answer is how fast our brain responds to information about attractiveness. Because an EEG measure was used in this experiment it was possible to show that the brain reacts to information about attractiveness as early as 250 milliseconds following presentation of a face. This information could not have been obtained in an fMRI study.

In conclusion, this study has shown that an N400 component appears when we are presented with an unattractive face followed by a positive characteristic, but not when an attractive face is followed by a negative word. It was also shown that EEG is sensitive to whether a face is attractive or unattractive.

References

Banaji, M. R. (2001). Stereotypes, social psychology of, In *International Encyclopedia of the Social & Behavioral Sciences*, pp. 15100-15104. Elsevier science Ltd

Barrett, S. E., & Rugg, M. D. (1989). Event related potentials and the semantic matching of faces. *Neuropsychologia, 27(7),* 913-922.

Boaz, T. L., Perry, N. W., Raney, R., Fischler, I. S. & Shuman, D. (1991). Detection of guilty knowledge with event-related potentials. *Journal of Applied Psychology, 76(6),* 788-795.

Braun, C., Gruendl, M., Marberger, C. & Scherber, C. (2001). *Beautycheck - Ursachen und Folgen von Attraktivitaet. Report*. Retrieved January 10th, 2007, from http://www.beautycheck.de/ english/bericht/bericht.htm.

Cunningham, M. R., Roberts, A. R., Barbee, A. P. & Druen, P. B. (1995). "Their ideas of beauty are, on the whole, the same as ours": Consistency and variability in the cross-cultural perception of female physical attractiveness. *Journal of Personality and Social Psychology, 68(2),* 261-279.

Dion, K. K., Berscheid, E. & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology, 1972,* 24, 285-290.

Eagly, A. H., Ashmore, R. D., Makhijani, M. G. & Longo, L. C. (1991). What is *beautiful is good*, but…: A meta-analytic review of research on the physical attractiveness stereotype. *Psychological Bulletin,* 110, 109-128.

Faustmann, A., Murdoch, B.E., Finnigan, S.P. & Copland, D.A. (2005). Event-related brain potentials elicited by semantic and syntactic anomalies during auditory sentence processing. *Journal of the American Academy of Audiology, 16(9)*,708-25.

Frieze, I. H., Olson, J. E. & Russel, J. (1991). Attractiveness and income for men and women in management. *Journal of Applied Social Psychology, 21(3),* 1039-1057.

Greenwald, A. G., McGhee, D. E. & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74(6),* 1464-1480.

Griffin, A. M. & Langlois, J. H. (2006). Stereotype directionality and attractiveness stereotyping: Is beauty good or is ugly bad? *Social Cognition, 24(2)*, 187-206.

Gross, A. E. & Crofton, C. (1977). What is good in beautiful. *Sociometry, 40(1),* 85-90.

Hagoort, P., Hald, L., Bastiaansen, M. & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science, 304,* 438-441.

Hamilton, D. L. & Sherman, J. W. (1994). Stereotypes. In *Handbook of Social Cognition.*

Henderson, J. J. & Anglin, J. M. (2003). Facial attractiveness predicts longlevity. *Evolution and Human Behavior, 24*, 351-356.

Jasper, H. H. (1958). The ten-twenty electrode system of the International Federation. *Electroencephalography & Clinical Neurophysiology, 10,* 371-375.

Kanazawa, S. & Kovar, J. L. (2004). Why beautiful people are more intelligent. *Intelligence, 32,* 227-243.

Kumar, N. & Debruille, B. (2004). Semantics and N400: insights for schizophrenia. *Journal of Psychiatry and Neuroscience, 29(2)*, 89-98.

Kutas, M., Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences, 4(12),* 463-469.

Kutas, M. & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 27(4427),* 203-205.

Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M. & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin, 126(3),* 390-423.

Langlois, J. H., Roggman, L. A., Casey, R. J. & Ritter, J. M. (1987). Infant preferences for attractive faces: Rudiments of a stereotype. *Developmental Psychology,* 23, 363-369.

Little, A. C., Burt, D. M. & Perrett, D. I. (2006). What is beautiful is good: Face preferences reflect desired personality. *Personality and Individual Differences, 41,* 1107-1118.

Luck, S. J., (2005). *An Introduction to the Event-Related Potential Technique 1st ed*. Massachusetts institute of technology.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (in press). The Implicit Association Test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Automatic Processes in Social Thinking and Behavior*. Psychology Press.

O´Doherty, J., Winston, J., Critchley, H., Perrett, D., Burt, D. M. & Dolan, R. J. (2003). Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia, 41*, 147-155.

Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review of Psychology, 57,* 199-226.

Riniolo, T. C., Johnson, K. C., Sherman, T. R. & Misso, J. A. (2006). Hot or not: Do professors perceived as physically attractive receive higher student evaluations. *Journal of General Psychology, 133(1),* 19-35.

Rosenberg, S. (1977). New approaches to the analysis of personal constructs in person perception. In A. L. Land & J. K. Cole (Eds.), *Nebraska Symposium on Motivation 1976,* 24, pp. 179-242. Lincoln: University of Nebraska Press.

Rothermund, K. & Wentura, D. (2004). Underlying processes in the implicit association test. Dissociating salience from associations. *Journal of Experimental Psychology: General, 133(2),* 139-165.

Rugg, M.D. & Coles, M.G.H. (1995). *Electrophysiology of Mind: Event-Related Brain Potentials and Cognition.* 1st ed. New York: Oxford University Press.

Shaughnessy, J. J., Zechmeister, E. B. & Zechmeister, J. S. (2005). Survey research. In *Research Methods in Psychology 7th ed.* 143-193. McGraw-Hill companies.

Shackelford, T. K. & Larsen, R. J. (1999). Physical attractiveness and physical health. *Evolution and Human Behaviour, 20,* 71-76.

Sigall, H. & Ostrove, N. (1975). Beautiful but dangerous: Effects of offender attractiveness and nature of the crime on juridic judgment. *Journal of Personality and Social Psychology, 31(3),* 410-414.

Thornhill, R. & Gangestad, S. W. (1999). Facial attractiveness. *Trends in Cognitive Sciences. 3(12)*. 452-460.

Winston, J. S., O´Doherty, J., Kilner, J. M., Perrett, D. I. & Dolan, R. J. (2007). Brain systems for assessing facial attractiveness. *Neuropsychologia, 45*, 195-206.

Appendix A – Words used in the experiment

Positive words

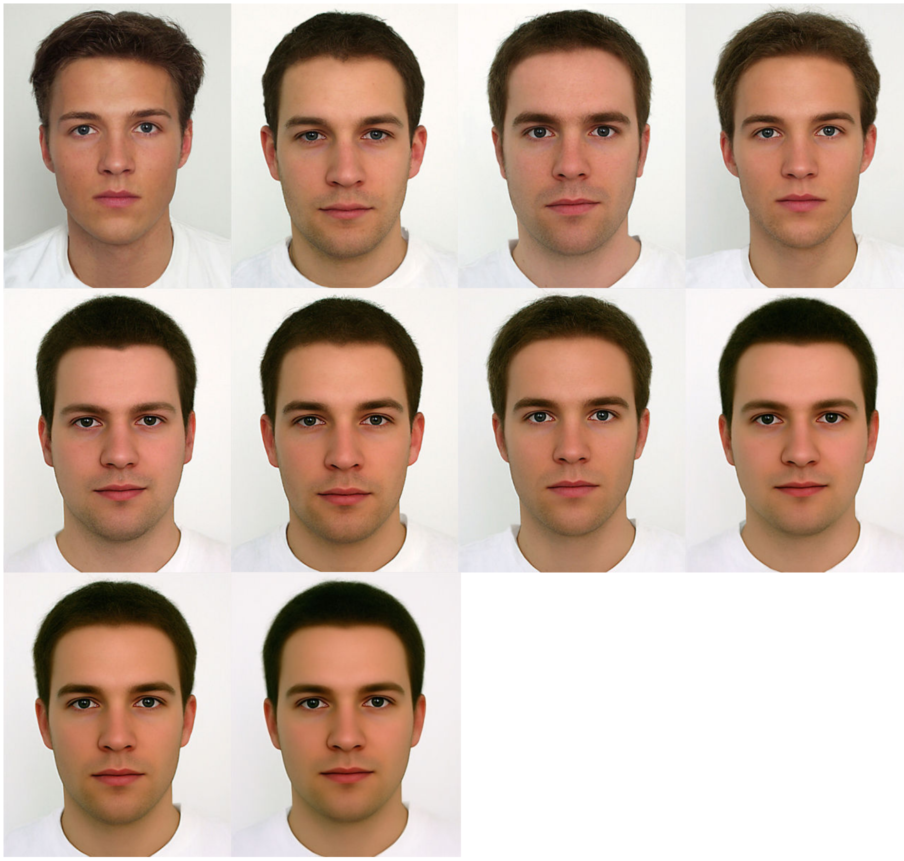
Alert
Begåvad
Behaglig
Charmerande
Driftig
Duktig
Elegant
Engagerad
Entusiastisk
Fantasifull
Påhittig
Fredlig
Fyndig
Företagsam
Förlåtande
Förnuftig
Förståndig
Givmild
Kapabel
Godhjärtad
Hederlig
Hjälpsam
Hjärtlig
Hoppfull
Humoristisk
Hänsynsfull
Initiativrik
Karismatisk
Omtyckt
Produktiv
Klipsk
Klok
Klyftig
Kompetent
Kreativ
Lycklig
Modig
Munter
Mysig
Omsorgsfull
Omtänksam
Pigg
Lyhörd
Resonabel
Rolig
Schysst
Skarpsinnig
Skärpt
Smart
Snabbtänkt
Social
Sympatisk
Tapper
Tillförlitlig
Tillmötesgående
Tilltalande
Trovärdig
Uppfinningsrik
Uppmärksam
Uppskattad
Varmhjärtad
Välvillig
Vänlig
Älskvärd
Ärlig
Ödmjuk
Nyfiken
Ambitiös
Bussig
Dynamisk
Flink
Flitig
Utåtriktad
Framgångsrik
Framstående
Fridsam
Generös
Godsint
Hygglig
Intelligent

Negative words

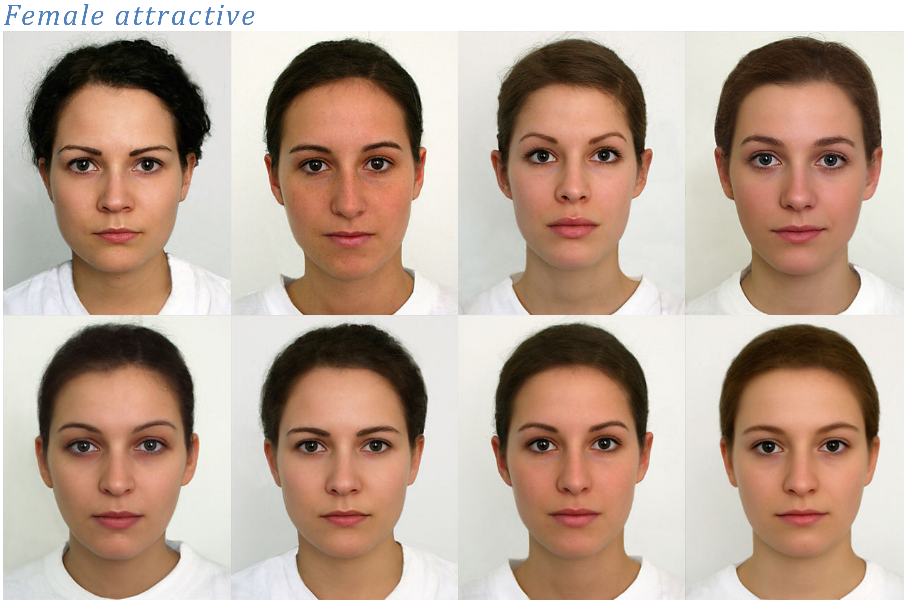
Tungsint
Ansvarslös
Apatisk
Arrogant
Asocial
Avskydd
Avskyvärd
Betydelselös
Dryg
Dum
Dyster
Kall
Egenkär
Egocentrisk
Egotrippad
Elakartad
Falsk
Fanatisk
Fantasilös
Gnällig
Grinig
Hänsynslös
Instabil
Kallsinnig
Klagande
Korkad
Labil
Frånstötande
Orolig
Lömsk
Missanpassad
Motbjudande
Nonchalant
Naiv
Oartig
Obegåvad
Obehaglig
Oberäknelig
Oengagerad
Ogästvänlig
Ohederlig
Ointressant
Rädd
Okänslig
Kylig
Impopulär
Opålitlig
Oresonlig
Oschysst
Oseriös
Deprimerad
Osympatisk
Otrevlig
Otrygg
Ovänlig
Oärlig
Pessimistisk
Självupptagen
Skygg
Slarvig
Slö
Slösaktig
Snobbig
Stel
Stötande
Svekfull
Obetydlig
Tanklös
Trist
Trumpen
Tråkig
Trög
Sluten
Spänd
Vek
Viljelös
Vresig
Vårdslös
Ynklig
Ängslig

Appendix B – Faces used in the experiment

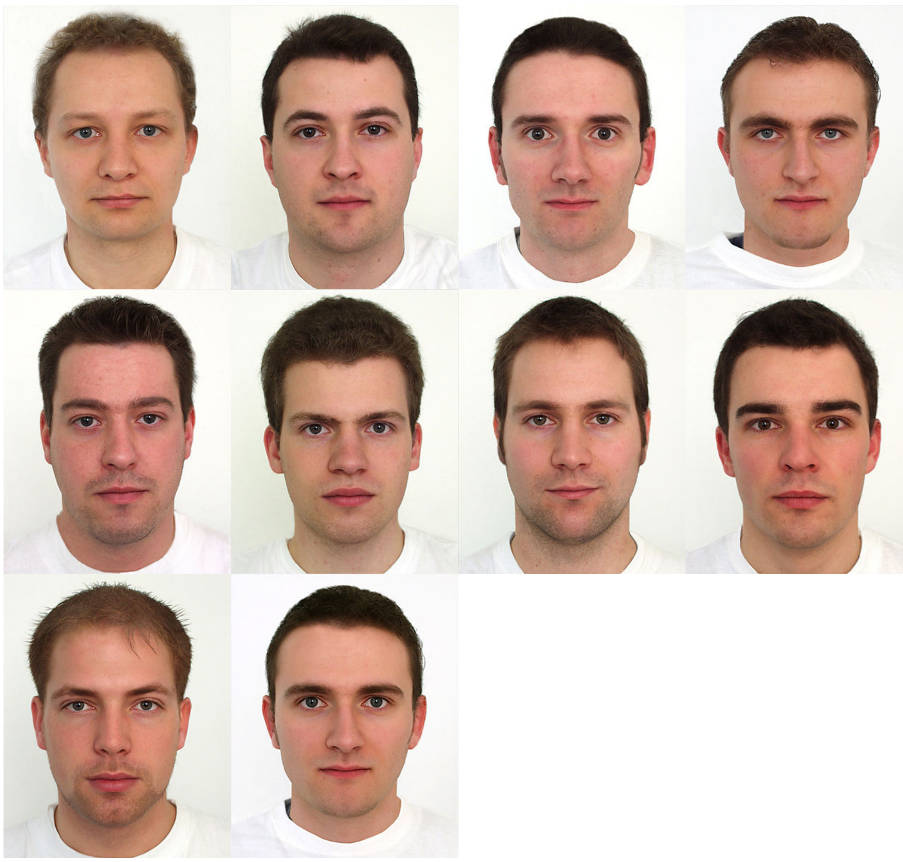
Male attractive



Female attractive



Male unattractive



Female unattractive